Radiotherapy Treatment Assessment using DCE-MRI

by

Chunhao Wang

Medical Physics Graduate Program
Duke University

Date:_______________________

Approved:

___________________________
Fang-Fang Yin, Co-Supervisor

___________________________
Zheng Chang, Co-Supervisor

___________________________
Oana Craciunescu, Chair

___________________________
Janet K. Horton

___________________________
James MacFall

Dissertation submitted in partial fulfillment of the requirements
for the degree of Doctor of Philosophy
in the Medical Physics Graduate Program
in the Graduate School of Duke University

2016
ABSTRACT

Radiotherapy Treatment Assessment using DCE-MRI

by

Chunhao Wang

Medical Physics Graduate Program
Duke University

Date:_______________________
Approved:

___________________________
Fang-Fang Yin, Co-Supervisor

___________________________
Zheng Chang, Co-Supervisor

___________________________
Oana Craciunescu, Chair

___________________________
Janet K. Horton

___________________________
James MacFall

An abstract of a dissertation submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy in the Medical Physics Graduate Program in the Graduate School of Duke University

2016
Abstract

The goal of modern radiotherapy is to precisely deliver a prescribed radiation dose to delineated target volumes that contain a significant amount of tumor cells while sparing the surrounding healthy tissues/organs. Precise delineation of treatment and avoidance volumes is the key for the precision radiation therapy. In recent years, considerable clinical and research efforts have been devoted to integrate MRI into radiotherapy workflow motivated by the superior soft tissue contrast and functional imaging possibility. Dynamic contrast-enhanced MRI (DCE-MRI) is a noninvasive technique that measures properties of tissue microvasculature. Its sensitivity to radiation-induced vascular pharmacokinetic (PK) changes has been preliminary demonstrated. In spite of its great potential, two major challenges have limited DCE-MRI's clinical application in radiotherapy assessment: the technical limitations of accurate DCE-MRI imaging implementation and the need of novel DCE-MRI data analysis methods for richer functional heterogeneity information.

This study aims at improving current DCE-MRI techniques and developing new DCE-MRI analysis methods for particular radiotherapy assessment. Thus, the study is naturally divided into two parts. The first part focuses on DCE-MRI temporal resolution as one of the key DCE-MRI technical factors, and some improvements regarding DCE-MRI temporal resolution are proposed; the second part explores the potential value of
image heterogeneity analysis and multiple PK model combination for therapeutic response assessment, and several novel DCE-MRI data analysis methods are developed.

I. Improvement of DCE-MRI temporal resolution. First, the feasibility of improving DCE-MRI temporal resolution via image undersampling was studied. Specifically, a novel MR image iterative reconstruction algorithm was studied for DCE-MRI reconstruction. This algorithm was built on the recently developed compress sensing (CS) theory. By utilizing a limited k-space acquisition with shorter imaging time, images can be reconstructed in an iterative fashion under the regularization of a newly proposed total generalized variation (TGV) penalty term. In the retrospective study of brain radiosurgery patient DCE-MRI scans under IRB-approval, the clinically obtained image data was selected as reference data, and the simulated accelerated k-space acquisition was generated via undersampling the reference image full k-space with designed sampling grids. Two undersampling strategies were proposed: 1) a radial multi-ray grid with a special angular distribution was adopted to sample each slice of the full k-space; 2) a Cartesian random sampling grid series with spatiotemporal constraints from adjacent frames was adopted to sample the dynamic k-space series at a slice location. Two sets of PK parameters’ maps were generated from the undersampled data and from the fully-sampled data, respectively. Multiple quantitative measurements and statistical studies were performed to evaluate the accuracy of PK maps generated from the undersampled data in reference to the PK maps generated from the fully-
sampled data. Results showed that at a simulated acceleration factor of four, PK maps could be faithfully calculated from the DCE images that were reconstructed using undersampled data, and no statistically significant differences were found between the regional PK mean values from undersampled and fully-sampled data sets. DCE-MRI acceleration using the investigated image reconstruction method has been suggested as feasible and promising.

Second, for high temporal resolution DCE-MRI, a new PK model fitting method was developed to solve PK parameters for better calculation accuracy and efficiency. This method is based on a derivative-based deformation of the commonly used Tofts PK model, which is presented as an integrative expression. This method also includes an advanced Kolmogorov-Zurbenko (KZ) filter to remove the potential noise effect in data and solve the PK parameter as a linear problem in matrix format. In the computer simulation study, PK parameters representing typical intracranial values were selected as references to simulated DCE-MRI data for different temporal resolution and different data noise level. Results showed that at both high temporal resolutions (<1s) and clinically feasible temporal resolution (~5s), this new method was able to calculate PK parameters more accurate than the current calculation methods at clinically relevant noise levels; at high temporal resolutions, the calculation efficiency of this new method was superior to current methods in an order of 10^2. In a retrospective of clinical brain DCE-MRI scans, the PK maps derived from the proposed method were comparable with
the results from current methods. Based on these results, it can be concluded that this new method can be used for accurate and efficient PK model fitting for high temporal resolution DCE-MRI.

II. Development of DCE-MRI analysis methods for therapeutic response assessment. This part aims at methodology developments in two approaches. The first one is to develop model-free analysis method for DCE-MRI functional heterogeneity evaluation. This approach is inspired by the rationale that radiotherapy-induced functional change could be heterogeneous across the treatment area. The first effort was spent on a translational investigation of classic fractal dimension theory for DCE-MRI therapeutic response assessment. In a small-animal anti-angiogenesis drug therapy experiment, the randomly assigned treatment/control groups received multiple fraction treatments with one pre-treatment and multiple post-treatment high spatiotemporal DCE-MRI scans. In the post-treatment scan two weeks after the start, the investigated Rényi dimensions of the classic PK rate constant map demonstrated significant differences between the treatment and the control groups; when Rényi dimensions were adopted for treatment/control group classification, the achieved accuracy was higher than the accuracy from using conventional PK parameter statistics. Following this pilot work, two novel texture analysis methods were proposed. First, a new technique called Gray Level Local Power Matrix (GLLPM) was developed. It intends to solve the lack of temporal information and poor calculation efficiency of the commonly used Gray Level
Co-Occurrence Matrix (GLCOM) techniques. In the same small animal experiment, the dynamic curves of Haralick texture features derived from the GLLPM had an overall better performance than the corresponding curves derived from current GLCOM techniques in treatment/control separation and classification. The second developed method is dynamic Fractal Signature Dissimilarity (FSD) analysis. Inspired by the classic fractal dimension theory, this method measures the dynamics of tumor heterogeneity during the contrast agent uptake in a quantitative fashion on DCE images. In the small animal experiment mentioned before, the selected parameters from dynamic FSD analysis showed significant differences between treatment/control groups as early as after 1 treatment fraction; in contrast, metrics from conventional PK analysis showed significant differences only after 3 treatment fractions. When using dynamic FSD parameters, the treatment/control group classification after 1\textsuperscript{st} treatment fraction was improved than using conventional PK statistics. These results suggest the promising application of this novel method for capturing early therapeutic response.

The second approach of developing novel DCE-MRI methods is to combine PK information from multiple PK models. Currently, the classic Tofts model or its alternative version has been widely adopted for DCE-MRI analysis as a gold-standard approach for therapeutic response assessment. Previously, a shutter-speed (SS) model was proposed to incorporate transcytolemmal water exchange effect into contrast agent concentration quantification. In spite of richer biological assumption, its application in
therapeutic response assessment is limited. It might be intriguing to combine the information from the SS model and from the classic Tofts model to explore potential new biological information for treatment assessment. The feasibility of this idea was investigated in the same small animal experiment. The SS model was compared against the Tofts model for therapeutic response assessment using PK parameter regional mean value comparison. Based on the modeled transcytosomeal water exchange rate, a biological subvolume was proposed and was automatically identified using histogram analysis. Within the biological subvolume, the PK rate constant derived from the SS model were proved to be superior to the one from Tofts model in treatment/control separation and classification. Furthermore, novel biomarkers were designed to integrate PK rate constants from these two models. When being evaluated in the biological subvolume, this biomarker was able to reflect significant treatment/control difference in both post-treatment evaluation. These results confirm the potential value of SS model as well as its combination with Tofts model for therapeutic response assessment.

In summary, this study addressed two problems of DCE-MRI application in radiotherapy assessment. In the first part, a method of accelerating DCE-MRI acquisition for better temporal resolution was investigated, and a novel PK model fitting algorithm was proposed for high temporal resolution DCE-MRI. In the second part, two model-free texture analysis methods and a multiple-model analysis method were
developed for DCE-MRI therapeutic response assessment. The presented works could benefit the future DCE-MRI routine clinical application in radiotherapy assessment.
Contents

Abstract .......................................................................................................................... iv
List of Tables ................................................................................................................... xv
List of Figures .................................................................................................................. xvi
List of Key Abbreviations .............................................................................................. xxi
Acknowledgements ......................................................................................................... xxv

1. Background .................................................................................................................. 1
  1.1 Clinical significance ................................................................................................. 1
  1.2 Previous developments of DCE-MRI theory ............................................................ 3
    1.2.1 Contrast mechanism ......................................................................................... 3
    1.2.2 Qualitative signal curve analysis ..................................................................... 4
    1.2.3 Quantitative pharmacokinetic analysis ............................................................. 5
  1.3 Challenges facing DCE-MRI and its application in treatment assessment ............. 13
    1.3.1 Temporal resolution ....................................................................................... 13
    1.3.2 Spatial resolution ............................................................................................. 14
    1.3.3 Accurate T₁ measurement ............................................................................... 15
    1.3.4 Arterial input function ..................................................................................... 16
    1.3.5 Data analysis for treatment assessment ........................................................... 17

2. Study focus and the overall structure ......................................................................... 20

3. Improving DCE-MRI temporal resolution with a novel total generalized variation
   (TGV) image reconstruction technique ........................................................................ 22
   3.1 Introduction ............................................................................................................ 22
3.2 Theory of TGV .............................................................................................................. 26

3.3 Materials and methods ............................................................................................... 28
  3.3.1 Reference data ...................................................................................................... 28
  3.3.2 Radial-based sampling scheme ............................................................................ 29
  3.3.3 Cartesian-based sampling scheme ....................................................................... 30
  3.3.4 Image reconstruction ......................................................................................... 33
  3.3.5 Pharmacokinetic analysis .................................................................................. 34
  3.3.6 Quantitative Accuracy Analysis ......................................................................... 36

3.4 Results ....................................................................................................................... 37
  3.4.1 Radial-based sampling scheme results ............................................................... 37
  3.4.2 Cartesian-based sampling scheme results ........................................................... 43

3.5 Discussion .................................................................................................................. 50

3.6 Conclusion .................................................................................................................. 52

4. Development of an efficient calculation method for high temporal resolution DCE-
  MRI analysis ................................................................................................................... 54
  4.1 Introduction .............................................................................................................. 54
  4.2 Materials and methods ........................................................................................... 55
    4.2.1 Theory of the DLLSQ method ......................................................................... 55
    4.2.2 Single voxel simulation .................................................................................... 59
    4.2.3 2D Simulation .................................................................................................... 61
    4.2.4 In vivo study ...................................................................................................... 61

4.3 Results ....................................................................................................................... 62
7.1 Introduction.................................................................................................................. 99
7.2 Theory of dynamic FSD ............................................................................................... 100
7.3 Retrospective small animal study .............................................................................. 102
7.4 Key Results .................................................................................................................. 103
7.5 Discussion .................................................................................................................... 110
7.6 Conclusion ................................................................................................................... 115

8. Integrating transcytosomeal water exchange analysis in multi-model PK analysis for DCE-MRI treatment response assessment .......................................................... 117

8.1 Introduction .................................................................................................................. 117
8.2 Theory of SS model .................................................................................................... 119
8.3 Retrospective small animal study .............................................................................. 122
  8.3.1 Evaluation of SS model ............................................................................................. 122
  8.3.2 Design of novel biomarkers .................................................................................... 124
8.4 Key Results .................................................................................................................. 125
8.5 Discussion .................................................................................................................... 136
8.6 Conclusion ................................................................................................................... 141

9. Summary ....................................................................................................................... 142

Appendix A ....................................................................................................................... 146
Appendix B ....................................................................................................................... 148
Appendix C ....................................................................................................................... 154
References ......................................................................................................................... 156
Biography .......................................................................................................................... 168
List of Tables

Table 1: The results of Error in Volume Mean (EIVM), Total Relative Error (TRE) and Cross-Correlation (CC) within the 3D ROI using radial-based undersampling scheme. 43

Table 2: The results of Error in Volume Mean (EIVM), Total Relative Error (TRE) and Cross-Correlation (CC) within the 3D ROI using Cartesian-based undersampling scheme. .................................................................46

Table 3: The comparison of PK parameter accuracies of a selected patient with different sampling strategy.................................................................52

Table 4: Total calculation times for 5000 runs of all methods at different temporal resolutions with CNR = 10. Calculation times are in unit of seconds.................................66

Table 5: Total Relative Error (TRE) and Cross-Correlation (CC) values of the 2D simulation results.................................................................70

Table 6: The median values of the investigated parameters in the in vivo study. ..........71

Table 7: The results of treatment/control groups’ classification using SVM ..................88

Table 8: Statistical analysis results of the derived features (Defined by Eqs (B1) – (B22) in Appendix B).........................................................................96

Table 9: The SVM results of treatment/control groups’ classification using different techniques for dynamic FSD analysis study .........................................................109

Table 10: The SVM results of treatment/control groups’ classification using dynamic FSD parameters from different temporal resolution.........................................114

Table 11: The results of treatment/control groups’ classification using SVM with different input options ........................................................................136
# List of Figures

Figure 1: An example of radiotherapy workflow using modern LINAC (a) and the proposed workflow with treatment assessment component (b) ......................................................... 2

Figure 2: Schematic representation of the contrast agent exchange process taking place on the microvessel scale ........................................................................................................ 3

Figure 3: An example of breast MR signal intensity curve categorization based on 3TP method. The Pattern 3 curve (blue) was categorized as malignant .................................................. 5

Figure 4: A simple scheme of the two-compartment Tofts model ........................................ 8

Figure 5: Concept of sliding window (SW) reconstruction using 2D radial sampling profile. Each full sets at \( t_1 \) and \( t_3 \) consists of four partial acquisitions which is undersampled azimuthally. ................................................................................................................... 24

Figure 6: An example of staircasing artifact of total variation (TV) processing. The denoise results (red curve) based on the noisy signal (blue curve) suffers staircasing artifact as opposed to ideal diagonal line (not shown) ..................................................... 25

Figure 7: Generation of the 32-ray radial sampling pattern at a golden ratio angular increment ..................................................................................................................................... 30

Figure 8: Illustration of a 2D sampling profile with variable sampling density. The central red region is fully sampled and the rest in gray level is sampled with varying sampling density ............................................................................................................. 31

Figure 9: Schematic of SW with variable density Cartesian sampling profiles. Each black line represents a sampled readout line ..................................................................................... 33

Figure 10: A comparison of the original post-injection image (a) and the reconstructed image using the radially undersampled data (b). The intensity difference between (a) and (b) is shown in (c). The red contour indicates the location of GTV as region-of-interest (ROI) ............................................................................................................. 38

Figure 11: The comparison of the measured \( C_p(t) \) from the CA concentration maps calculated from original clinical images (blue curve) and the images reconstructed using reduced sampling data with radial-based scheme (red curve) ........................................................... 38
Figure 12: A comparison of $K_{\text{trans}}$ map calculated using the original full sampling data (a) and using the reduced sampling data with radial-based scheme (b); The histogram about relative error of (b) versus (a) of all 3D ROI voxels (c); The percent difference map of the ROI region (d).................................................................40

Figure 13: A comparison of $F_B$ map calculated using the original full sampling data (a) and using the reduced sampling data with radial-based scheme (b); The histogram about relative error of (b) versus (a) of all 3D ROI voxels (c); The percent difference map of the ROI region (d).................................................................................................................41

Figure 14: A comparison of $V_B$ map calculated using the original full sampling data (a) and using the reduced sampling data with radial-based scheme (b); The histogram about relative error of (b) versus (a) of all 3D ROI voxels (c); The percent difference map of the ROI region (d).................................................................................................................42

Figure 15: A comparison of the original post-injection image (a) and the reconstructed image using the Cartesian-based undersampled data (b). The intensity difference between (a) and (b) is shown in (c). The red contour indicates the location of GTV as region-of-interest (ROI)..................................................................................................................44

Figure 16: The comparison of the measured $C_p(t)$ from the CA concentration maps calculated from original clinical images (blue curve) and the images reconstructed using reduced sampling data with Cartesian-based scheme (red curve).........................................................44

Figure 17: A comparison of $K_{\text{trans}}$ map calculated using the original full sampling data (a) and using the reduced sampling data with Cartesian-based scheme (b); The histogram about relative error of (b) versus (a) of all 3D ROI voxels (c); The percent difference map of the ROI region (d).................................................................................................................47

Figure 18: A comparison of $F_B$ map calculated using the original full sampling data (a) and using the reduced sampling data with Cartesian-based scheme (b); The histogram about relative error of (b) versus (a) of all 3D ROI voxels (c); The percent difference map of the ROI region (d).................................................................................................................48

Figure 19: A comparison of $V_B$ map calculated using the original full sampling data (a) and using the reduced sampling data with Cartesian-based scheme (b); The histogram about relative error of (b) versus (a) of all 3D ROI voxels (c); The percent difference map of the ROI region (d).................................................................................................................49
Figure 20: An example CA concentration curve processing of the DLLSQ method in single voxel simulation. (a) The comparison of $C_p(t)$; (b) The comparison of $C_p(t)$ derivative; (c) The comparison of $C_t(t)$; (d) The comparison of $C_t(t)$ derivative. The derivative of noisy $C_p(t)$ and $C_t(t)$ are not shown for the best illustration.

Figure 21: Calculated values of (a) $K_{\text{trans}}$, (b) $k_{ep}$, and (c) $v_p$ as a function of CNR with a temporal resolution of 1s. The horizontal dashed lines represent the known true values: $K_{\text{trans}} = 0.0890$, $k_{ep} = 0.3201$, and $v_p = 0.0230$. Each error bar represents the SD for 5000 simulations.

Figure 22: Calculated values of (a) $K_{\text{trans}}$, (b) $k_{ep}$, and (c) $v_p$ as a function of temporal resolution at CNR = 10. The horizontal dashed lines represent the known true values: $K_{\text{trans}} = 0.0890$, $k_{ep} = 0.3201$, and $v_p = 0.0230$. Each error bar represents the SD for 5000 simulations.

Figure 23: Calculated results of $K_{\text{trans}}$ (left column), $k_{ep}$ (middle column), and $v_p$ (right column) with the improved ILLSQ and NLSQ methods at varying CNR (1st row) and temporal resolutions (2nd row). The horizontal dashed lines represent the known true values: $K_{\text{trans}} = 0.0890$, $k_{ep} = 0.3201$, and $v_p = 0.0230$. Each error bar represents the SD for 5000 simulations.

Figure 24: 2D simulation results of $K_{\text{trans}}$. (a) True values; (b) DLLSQ results; (c) ILLSQ results; (d) NLSQ results; (e) difference map of the DLLSQ results; (f) difference map of the ILLSQ results; (g) difference map of the NLSQ results.

Figure 25: 2D simulation results of $k_{ep}$. (a) True values; (b) DLLSQ results; (c) ILLSQ results; (d) NLSQ results; (e) difference map of the DLLSQ results; (f) difference map of the ILLSQ results; (g) difference map of the NLSQ results.

Figure 26: 2D simulation results of $v_p$. (a) True values; (b) DLLSQ results; (c) ILLSQ results; (d) NLSQ results; (e) difference map of the DLLSQ results; (f) difference map of the ILLSQ results; (g) difference map of the NLSQ results.

Figure 27: In vivo study results of a selected patient. 1st row: $K_{\text{trans}}$ results of the DLLSQ (a), the ILLSQ (b), and the NLSQ method (c); 2nd row: $k_{ep}$ results of the DLLSQ (d), the ILLSQ (e), and the NLSQ method (f); 3rd row: $v_p$ results of the DLLSQ (g), the ILLSQ (h), and the NLSQ method (i).

Figure 28: An example of the GLCOM generation on the horizontal direction. Red and blue color shows the derivation of $P(1,2)$ and $P(3,4)$ respectively.
Figure 29: An example of the GLRLM generation on the horizontal direction. Red and blue color shows the derivation of $P(3,1)$ and $P(2,3)$ respectively. ........................................ 79

Figure 30: An example of box counting method for Apollonian Gasket analysis. (a) The binary object is successively divided using boxes of decreasing grid size $s$. (b) The relationship between $\log s$ and $\log n$ where $n$ is the number of boxes occupied by the object. For a perfect fractal object, the curve should be a straight line. Box counting dimension is computed as the slope of the line (approximately 1.4 in this example) ....... 80

Figure 31: The diagram of the small animal experiment. MR = DCE-MRI scan; Tx = Treatment delivery............................................................................................................ 83

Figure 32: The comparisons of tumor volume $V$ and its relative value between treatment/control groups during the treatment course. * indicates that statistically significant difference was found between treatment/control groups ........................................ 86

Figure 33: The longitudinal change of mean $K_{\text{trans}}$ (a), relative $K_{\text{trans}}$ (b), $K_{\text{trans}}$ CV (c), $K_{\text{trans}}$ kurtosis (d), $K_{\text{trans}}$ skewness (e), $K_{\text{trans}} d_1$ (f), and $K_{\text{trans}} d_2$ (g). * indicates that statistically significant difference was found between treatment/control groups ........................................ 87

Figure 34: The comparison of CA concentration curves of treatment/control groups at Day0 (left) and at Day 9 (right) ........................................................................................................ 93

Figure 35: The comparison of auto correlation (AC) curves of GLLPM (left column), 4D GLCOM (middle column) and 3D GLCOM (right column) at Day0 (top row) and Day9 (bottom row). Red: Treatment group; Blue: Control group ........................................ 94

Figure 36: The comparison of inverse variance (InvVar) curves of GLLPM (left column), 4D GLCOM (middle column) and 3D GLCOM (right column) at Day0 (top row) and Day9 (bottom row). Red: Treatment group; Blue: Control group ........................................ 95

Figure 37: A diagram of a 3D surface (a) and its upper blanket (above the surface) and lower blanket (below the surface) (b). ........................................................................................................ 100

Figure 38: An example of pre-treatment DCE-MRI scan. (a) Post-injection MR image of a selected slice; (b) The 3D tumor’s average CA concentration curve (blue) and dynamic FSD curve; (c) Pre-injection $T_{10}$ map; (d) $AUC_{\text{MR}}$ map generated in the same time window as FSD curve; (e) CA distribution 2 minutes after injection; (f) $K_{\text{trans}}$ distribution ............ 104

Figure 39: The longitudinal change of mean $AUC_{\text{MR}}$ (a), relative $AUC_{\text{MR}}$ (b), $AUC_{\text{MR}}$ CV (c), $AUC_{\text{MR}}$ kurtosis (d), $AUC_{\text{MR}}$ skewness (e), $AUC_{\text{MR}} d_1$ (f), and $AUC_{\text{MR}} d_2$ (g) .................. 106
Figure 40: The demonstration of dynamic FSD analysis. (a) and (b): dynamic FSD curves on all scan days of a treatment group animal and a control group animal (The Day0 curve in (b) has been shown in Figure 38(b)); (c) and (d): the longitudinal change of\( AUC_{FSD} \) and its relative value; (e) and (f): the longitudinal change of\( ME_{FSD} \) and its relative value.*indicates that statistically significant difference was found between treatment/control groups .......................................................................................................................... 108

Figure 41: The comparison of dynamic FSD curves generated using the original 9.9s temporal resolution (red), simulated 19.8s (green) and 29.7s (blue) temporal resolutions. The red curve was shown in Figure 6(b). ................................................................................................................... 113

Figure 42: An example of a selected animal’s scans at three scan days. Left column: Day0; middle column: Day2; right column: Day9. First row: post-injection T1w DCE image; 2nd row: \( K_{trans} \) maps from the Tofts model; 3rd row: \( K_{trans} \) maps from the SS model; 4th row: joint histogram of \( K_{trans} \) from the two models. The tumor mean values of this animal are: Tofts: Day0: 0.429 min\(^{-1}\), Day2: 0.361 min\(^{-1}\), Day9: 0.153 min\(^{-1}\); SS: Day0: 0.661 min\(^{-1}\), Day2: 0.559 min\(^{-1}\), Day9: 0.229min\(^{-1}\) ................................................. 127

Figure 43: Transcytolemmal water exchange analysis at three scan days. Left column: Day0; middle column: Day2; right column: Day9. First row: the \( \tau_i \) maps across the tumor; 2nd row: the identified BV (red area) within the tumor; 3rd row: the \( \tau_i \) histograms. The tumor mean values \( \tau_i \) of this animal are: Day0: 0.159s; Day2: 0.170s; Day9: 0.193s ......... 128

Figure 44: The comparison of model fitting quality. (a): \( \chi^2 \) ratio map of the Day0 scan of the same animal used before; (b) \( \chi^2 \) ratio histogram; (c) the comparison of BIC values of all analyzed scans.................................................................................................................. 129

Figure 45: The comparisons of tumor mean \( K_{trans} \) statistics using Tofts model (a) and using SS model (b). * indicates statistical significance................................................................. 130

Figure 46: The comparisons of tumor mean \( \tau_i \). * indicates statistical significance .......... 131

Figure 47: The comparisons of mean \( K_{trans} \) from SS model in BV(a), mean \( \tau_i \) in BV (b), \( d_1 \) (c) and \( d_i \) (d) of \( K_{trans} \) from SS model in BV. * indicates statistical significance......................... 132

Figure 48: The comparisons of tumor mean \( \Delta K_{trans} \) (a), tumor mean normalized \( \Delta K_{trans} \) (b), BV mean \( \Delta K_{trans} \) (c) and BV mean normalized \( \Delta K_{trans} \) (d). * indicates statistical significance ................................................................................................................. 134
### List of Key Abbreviations

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>CA</td>
<td>Contrast agent</td>
</tr>
<tr>
<td>EES</td>
<td>Extravascular extracellular space</td>
</tr>
<tr>
<td>ECU</td>
<td>Early contrast uptake</td>
</tr>
<tr>
<td>TTP</td>
<td>Time to (CA curve) peak</td>
</tr>
<tr>
<td>AUC</td>
<td>Area under curve</td>
</tr>
<tr>
<td>PK</td>
<td>Pharmacokinetic</td>
</tr>
<tr>
<td>FXL</td>
<td>Fast-Exchange Limit</td>
</tr>
<tr>
<td>$R_1$</td>
<td>Longitudinal relaxation rate ($=1/T_1$)</td>
</tr>
<tr>
<td>$R_{10}$</td>
<td>Native longitudinal relaxation rate</td>
</tr>
<tr>
<td>$r_1$</td>
<td>Relativity rate constant of CA</td>
</tr>
<tr>
<td>$v_p$</td>
<td>Volume fraction of blood plasma</td>
</tr>
<tr>
<td>$v_e$</td>
<td>Volume fraction of EES</td>
</tr>
<tr>
<td>$K^{trans}$</td>
<td>CA extravasation rate constant from blood plasma to EES</td>
</tr>
<tr>
<td>$k_{ep}$</td>
<td>CA transfer rate constant from EES back to blood plasma</td>
</tr>
<tr>
<td>$C_p(t)$</td>
<td>Arterial input function (AIF)</td>
</tr>
<tr>
<td>SS</td>
<td>Shutter-Speed</td>
</tr>
<tr>
<td>$\Delta t$</td>
<td>Temporal resolution</td>
</tr>
<tr>
<td>GTV</td>
<td>Gross tumor volume</td>
</tr>
<tr>
<td>FOV</td>
<td>Field-of-view</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Description</td>
</tr>
<tr>
<td>--------------</td>
<td>-------------</td>
</tr>
<tr>
<td>ROI</td>
<td>Region-of-interest</td>
</tr>
<tr>
<td>SW</td>
<td>Sliding Window</td>
</tr>
<tr>
<td>CS</td>
<td>Compress Sensing</td>
</tr>
<tr>
<td>TV</td>
<td>Total variation</td>
</tr>
<tr>
<td>TGV</td>
<td>Total generalized variation</td>
</tr>
<tr>
<td>$F_B$</td>
<td>(Brain) Blood flow</td>
</tr>
<tr>
<td>$V_B$</td>
<td>(Brain) Blood volume</td>
</tr>
<tr>
<td>EIVM</td>
<td>Error in volume mean</td>
</tr>
<tr>
<td>CC</td>
<td>Cross correlation</td>
</tr>
<tr>
<td>TRE</td>
<td>Total relative error</td>
</tr>
<tr>
<td>NLSQ</td>
<td>Nonlinear least-squares (method)</td>
</tr>
<tr>
<td>ILLSQ</td>
<td>Linear least-squares (method) based on (PK model’s) integral form</td>
</tr>
<tr>
<td>DLLSQ</td>
<td>Linear least-squares (method) based on (PK model’s) derivative form</td>
</tr>
<tr>
<td>KZ filter</td>
<td>Kolmogorov-Zurbenko filter</td>
</tr>
<tr>
<td>CNR</td>
<td>Contrast-to-noise ratio</td>
</tr>
<tr>
<td>GLCOM</td>
<td>Gray level co-occurrence matrix</td>
</tr>
<tr>
<td>GLRLM</td>
<td>Gray level run length matrix</td>
</tr>
<tr>
<td>GLLPM</td>
<td>Gray level local power matrix</td>
</tr>
<tr>
<td>Symbol</td>
<td>Definition</td>
</tr>
<tr>
<td>--------</td>
<td>------------</td>
</tr>
<tr>
<td>$d_i$</td>
<td>Information dimension (of Rényi dimensions)</td>
</tr>
<tr>
<td>$d_i$</td>
<td>Correlation dimension (of Rényi dimensions)</td>
</tr>
<tr>
<td>CV</td>
<td>Coefficient of variation</td>
</tr>
<tr>
<td>SVM</td>
<td>Supporting vector machine</td>
</tr>
<tr>
<td>FSD</td>
<td>Fractal signature dissimilarity</td>
</tr>
<tr>
<td>$AUC_{FSD}$</td>
<td>Area under FSD time curve</td>
</tr>
<tr>
<td>$ME_{FSD}$</td>
<td>Maximum enhancement of FSD time curve</td>
</tr>
<tr>
<td>$AUC_{MR}$</td>
<td>Area under MR signal enhancement time curve</td>
</tr>
<tr>
<td>$\tau_i$</td>
<td>Mean residence time of water molecules in cell (its inverse is transcytolemmal water exchange rate)</td>
</tr>
<tr>
<td>$p_e$</td>
<td>Fraction of EES water molecules</td>
</tr>
<tr>
<td>$f_w$</td>
<td>Fraction of water molecules that accessible to CA particles in EES</td>
</tr>
<tr>
<td>BV</td>
<td>Biological subvolume</td>
</tr>
<tr>
<td>BIC</td>
<td>Bayesian Information Criterion</td>
</tr>
<tr>
<td>$K_{F,trans}$</td>
<td>$K_{trans}$ calculated from Tofts model</td>
</tr>
<tr>
<td>$K_{S,trans}$</td>
<td>$K_{trans}$ calculated from SS model</td>
</tr>
<tr>
<td>$K_{S,BV,trans}$</td>
<td>$K_{trans}$ calculated from SS model in BV</td>
</tr>
<tr>
<td>$\tau_i, BV$</td>
<td>$\tau_i$ calculated in BV</td>
</tr>
</tbody>
</table>
Acknowledgements

I would like to firstly thank my committee member for their guidance and help during my Ph.D. research. In particular, I would like to thank my advisors, Dr. Fang-Fang Yin and Dr. Zheng Chang, who have provided me with tremendous help and support throughout my research. Thank you for taking your time from the busy clinical schedule to discuss with me not only scientific topics but also in aspects of life as a young man in graduate school. I was very fortunate to have worked under the supervision of these two brilliant scientists. Dr. Oana Craciunescu has been very helpful for providing valuable comments as an experienced clinical physicist and a prestigious researcher. I am proud of having such a big name in DCE-MRI research as my committee chair. I would also like to thank Dr. Janet Horton for giving me the precious chance working with her on clinical projects. Her insightful suggestions from a radiation oncologist’s perspective helped this work to better focus on clinical challenges and practicality. I am also very grateful for Dr. James MacFall’s involvement of my research. He has provided me with solid training of medical imaging courses for my minor specialty and has been very generous to devoting his time for this work. I also want to express my sincere appreciation to Dr. Ergys Subashi. I have been working closely with him during his clinical residency. Though he was very busy with extensive clinical training, he was generous about sharing his precious data and spending time with me on discussions of our collaborations, which led to several conference abstracts and
journal articles. My thanks also go to Dr. Jackie Wu for her kind support of my work when I started my research; to Dr. Lei Ren and Dr. Jing Cai for their selfless support of my research with their resources; to Yang Sheng and Xiao Liang for their help in scientific discussions on my first two projects; to Wendy Harris, Kyle Lafata and Lei Zhang for their great suggestions on my recent projects; to Dr. Anna Rodrigues, Dr. Austin Faught and Dr. Lulin Yuan for giving me clinical trainings after working hours; to Dr. You Zhang, Dr. Taoran Li and Dr. Taoran Cui for their suggestions and advices for choosing residency programs and for my career plans; to all my dear friends at Duke for making my life full of fun that made me happy everyday.

Lastly, I would like to thank my honored family for supporting me pursing my passion in medical physics. I was lucky to grow up in a family atmosphere of knowledge and wisdom. My mom and dad in China have provided me with endless love. I also want to thank my girlfriend Hangjie, who has been extremely understanding and supportive for my work. Thank you for being here with me to cheer me up during the tough study life.

The research was supported by a master research grant from Varian Medical System. Dr. James Dobbins III, Dr. Ehsan Samei, Dr. Tim Turkington, Dr. Anuj Kapadia, Dr. Olga Baranova and Ms. Katherine Hand from the Duke Medical Physics Graduate Programs are supportive throughout my Duke time.
1. Background

1.1 Clinical significance

Generally, the workflow of radiotherapy involves three important components: pre-treatment imaging, treatment planning, and treatment delivery (Khan, 2010). As an example, Figure 1(a) demonstrates a typical workflow of radiotherapy using modern LINAC. The first step is usually the treatment simulation to obtain CT images. In the following treatment planning stage, the treatment plan is developed primarily based on CT data for treatment dose calculation, while other imaging modalities including MRI and PET may provide substantial information for treatment target delineation (Torresin et al., 2015). The treatment plan can be delivered on a modern LINAC with on-board image guidance for accurate treatment dose delivery. The current workflow may, however, miss one potentially important step: treatment assessment (Figure 1(b)). Effective treatment assessment would provide potentially valuable information to determine the treatment effectiveness. In the proposed concept of individualized radiotherapy, treatment assessment provides essential information that permits treatment interventions for the optimal outcome (Yaromina et al., 2012). Treatment assessment would also provide valuable insights on the future development of treatment techniques (Wang et al., 2014).
As a functional MR imaging technique, DCE-MRI tracks the pharmacokinetics of injected contrast agent (CA) as they pass through the tumor vasculature. Treatment assessment using DCE-MRI is the process of using such technique before, during and after the treatment course to evaluate the changes of microvascular functional information (Chang and Wang, 2015). It is sensitive to the radiotherapy induced alterations in vascular permeability, extracellular extravascular and vascular volume, and blood flow (O’Connor et al., 2007). In the proposed concept of individualized radiotherapy, the captured microvascular functional change can be utilized to optimize the radiotherapy plan along the treatment course in aspect of fractionation altering, treatment target refinement and dose escalation for improved therapeutic outcome.
1.2 Previous developments of DCE-MRI theory

1.2.1 Contrast mechanism

In DCE-MRI, a serial T₁ weighted MR images were acquired before, during, and after an intravenous administration of low-molecular weight contrast agent (CA). Figure 2 provides a schematic representation of CA dynamics at the microscopic scale. The injected CA will rapidly distribute in the blood plasma as in the blood circulation. Owing to the possible microvascular leakiness, CA will enter the extravascular extracellular space (EES) together with water protons. In EES, the CA cannot pass through the cell membrane but may re-enter the blood plasma through the leaking capillary wall and exit the local tissue in venous flow. Most clinically used CA contains Gadolinium (Gd) ions, which are paramagnetic and interact with nearby proton to shorten the T₁ time in local tissue. This sequence of events results in a wash-in phase and a potential wash-out phase on the MR signal intensity time curve.

Figure 2: Schematic representation of the contrast agent exchange process taking place on the microvessel scale
### 1.2.2 Qualitative signal curve analysis

As an early qualitative attempt, the three time-points (3TP) method has been proposed for MR signal time curve characterization, as being demonstrated in Figure 3 (Degani et al., 1997). In this method, one pre-injection \( t_0 \) and two post-injection \( t_1, t_2 \) time points are selected to represent the wash-in and wash-out features of a MR signal curve. Based on the signal intensity of the selected time points, each MR signal curve can be categorized according to one of three patterns coded by color hues: (1) slow wash out \( I(t_1) < I(t_2) \), coded red; (2) moderate wash-out with \( I(t_1) \approx I(t_2) \); (3) fast wash-out \( I(t_1) > I(t_2) \).

The wash-in rate \( [I(t_1) - I(t_0)] / (t_1 - t_0) \) can be coded by color intensity in each pattern. Previous studies have shown that the MR signal curve pattern can be used to emphasize the benign (Pattern 1), malignant (Pattern 3) and uncertain (Pattern 2) nature of breast tissue (Hauth et al., 2006). The MR signal curve of each voxel can be color-coded with its wash-out pattern on the anatomical MR images as a pseudo-colored map for better visualization. Since only three time points are required for analysis with potential high spatial resolution capability, the 3TP method has been widely reported in DCE-MRI clinical literatures (Sansone et al., 2015). Other quantitative metrics such as early contrast uptake (ECU, the percent increase of MR signal intensity at the 1st time point after CA injection) and time to peak (TTP, the time between CA injection and maximum MR signal intensity time point) have been proposed to characterize MR signal intensity time curve (Martincich et al., 2004).
Figure 3: An example of breast MR signal intensity curve categorization based on 3TP method. The Pattern 3 curve (blue) was categorized as malignant

1.2.3 Quantitative pharmacokinetic analysis

As a phenomenological approach with simple assumptions, the aforementioned analyses of MR signal time curve has been frequently adopted in preclinical and clinical studies. However, its qualitative nature limits the comprehensive description of CA dynamics in physiological contexts. In quantitative DCE-MRI analysis, the CA dynamics is described by certain pharmacokinetic (PK) models, in which the outputs can be interpreted as parameters of known physiological process. Because of this merit, the PK model analysis has become a standard procedure in quantitative DCE-MRI studies.

The quantitative PK analysis works on the CA concentration time curves. Unlike dynamic CT or PET, the relationship between MR signal intensity and CA concentration is not linear. It is usually assumed that water exchange between intracellular space and EES (transcytolemmal water exchange) is very fast (also known as fast-exchange limit FXL). Since in most tissues most water is intracellular, the FXL leads to an ideal scenario
that all CA particles are combined with the mobile water molecules in EES as a homogeneous distribution without gradients (Yankeelov et al., 2003). As a result, the CA concentration in a voxel \( C(t) \) can be estimated as a linear dependence of longitudinal relaxation rate \( R_1 (=1/T_1) \) change:

\[
C(t) = \frac{(R_1(t) - R_{10})}{r_1}
\]

where \( R_{10} \) is the native relaxation rate before the CA injection. \( r_1 \) is the longitudinal relaxivity of the CA which is a constant at a given \( B_0 \) strength. To measure \( R_{10} \), calibration scans are acquired before the CA injection. Currently, the fast T1-weighted spoiled gradient echo (SPGR) sequences are widely for DCE-MRI scan as they allow good CA sensitivity, high signal-to-noise ratio (SNR), adequate volume coverage and fast data acquisition (Parker and Buckley, 2005). Using the SPGR sequence, the magnitude of the MR signal at a voxel is given by:

\[
M(\vec{r}) = M_0(\vec{r}) \frac{\sin(\alpha) \cdot (1 - e^{-TR/T_1(\vec{r})})}{1 - \cos(\alpha) e^{-TR/T_2^*(\vec{r})}} \cdot e^{-TE/T_2^*(\vec{r})}
\]

where \( M_0(\vec{r}) \) is proportional to the equilibrium longitudinal magnetization and the gain function of the imaging system. TR and TE denote repetition time and echo time, respectively. \( \alpha \) represents the flip angle. The \( T_2^* \) term is usually assumed to be unity with short TE. A common \( R_{10} \) mapping approach is multiple flip angle method in which calibration scans are acquired at a series of flip angles \( \alpha = \{\alpha_1, \alpha_2, \ldots, \alpha_N\} \). Then \( T_1 \) can be solved from a linear least-squares fitting problem:
\[ \hat{y} = m \cdot \hat{x} + b \]

\[ \hat{y}_i = \frac{S(\alpha_i)}{\sin(\alpha_i)}, \hat{x}_i = \frac{S(\alpha_i)}{\tan(\alpha_i)} \]

\[ m = \exp(-TR/T_1) \]

where \( S(\alpha_i) \) is the measured MR signal intensity in Eq. (2). As a special case of Eq. (3), dual flip angle method was proposed as a simpler approach (Schabel and Morrell, 2009). Specifically, \( \rho \) is defined as the MR signal intensity ratio using different flip angles \( \phi \) and \( \psi \). Then \( T_1 \) value can be calculated as:

\[ f(\rho) = \frac{\rho \sin(\psi \cos(\psi) - \cos(\phi \sin(\psi))}{\rho \sin(\phi \sin(\psi) - \sin(\psi))} \]

\[ T_1 = TR / \ln[f(\rho)] \]

Eqs. (3) and (4) can be used for both \( R_{10} \) and \( R_1 \) calculation. Combining with Eq. (3) or (4) with Eq. (2) and Eq. (1), each voxel’s CA concentration time curve after the CA injection can be derived from MR signal intensity curves.

The interpretation of functional information in DCE-MRI depends on the selection of PK model. Currently, most PK models in DCE-MRI consist of different compartments. A compartment is defined as an amount or a volume of material and does not necessarily describe a singular physical location. So far, the most widely used PK model is the one proposed by Tofts and Kermode in 1991 (Tofts and Kermode, 1991). In this model, the physiological kinetics in Figure 1 is described by the CA bidirectional transendothelium movement between two compartments, blood plasma and EES (Figure 4):
Figure 4: A simple scheme of the two-compartment Tofts model

where $v_p$ and $v_e$ represent the volume fractions of blood plasma and EES, respectively. $K_{\text{trans}}$ is the extravasation rate constant of CA from blood plasma to EES, and $k_{cp}$ is the rate constant describing the return of CA from EES to blood plasma. $K_{\text{trans}}$ and $k_{cp}$ are related as $k_{cp} = K_{\text{trans}}/v_e$. The measured CA concentration $C_t(t)$ in a voxel consists of two components:

$$C_t(t) = C_{\text{EES}}(t) + v_p C_p(t)$$  \hspace{1cm} (5)

where $C_{\text{EES}}(t)$ is the CA concentration in EES and $C_p(t)$ is the CA concentration in blood plasma. The $C_{\text{EES}}(t)$ term can also be expressed by the Kety Rate Law as the convolution of $C_p(t)$ with an exponential term (Kety, 1960):

$$C_{\text{EES}}(t) = K_{\text{trans}} \int_0^t C_p(t') \exp\left[-\frac{K_{\text{trans}}}{v_e} (t - t')\right] dt'$$  \hspace{1cm} (6)

In some tissues such as human breast, the blood plasma volume fraction $v_p$ is argued to be very small, so the contribution from $C_p(t)$ in Eq. (5) is neglected. Then Eq. (6) can be rewritten as Eq. (7), which is referred as the standard Tofts model:

$$C_t(t) = K_{\text{trans}} \int_0^t C_p(t') \exp\left[-\frac{K_{\text{trans}}}{v_e} (t - t')\right] dt'$$  \hspace{1cm} (7)
Another version of Eq. (7) incorporates the effects of possible significant vascular signal and thus adds an additional vascular signal term:

$$C_t(t) = v_p C_p(t) + K_{\text{trans}} \int_0^t C_p(t') \exp\left[-\frac{K_{\text{trans}}}{v_e} (t - t')\right] dt'$$  

(8)

The above equation is frequently called as the extended Tofts model. It was argued that extended Tofts model might be more reliable in the region with higher vascular signal (Roberts et al., 2006).

The rate constant $K_{\text{trans}}$ is the key parameter in Tofts models. It contains the combined information of capillary wall surface, capillary permeability, and blood flow from blood vessel to tissue and has served as the primary quantitative imaging biomarker in many clinical studies involving DCE-MRI analysis (Yankeelov and Gore, 2009; Zahra et al., 2007). If the delivery of the contrast medium to a tissue is insufficient (flow-limited situation), then blood perfusion will be the dominant factor of CA enhancement in tissue. In this situation, $K_{\text{trans}}$ approximates to tissue blood flow due to high microvascular permeability (Tofts et al., 1999). On the other hand, if the blood perfusion is sufficient and the CA extravasation out of the vasculature does not deplete the intravascular CA concentration (non-flow-limited situation), then the CA transport from blood plasma to EES is the major factor of tissue CA enhancement, and thus $K_{\text{trans}}$ approximates to the product of vascular permeability rate and microvessel surface area (d’Arcy et al., 2006). For Gd-based low molecular weight CA, the mixed situation occurs most commonly, and the influence of permeability-surface area is believed to outweigh
the perfusion effect when the CA injection dose is sufficient. In addition to $K_{\text{trans}}$, the semi-quantitative initial Area Under the CA concentration time Curve (iAUC) is another frequently reported metric in quantitative DCE-MRI analysis. Studies revealed that there were no statistically significant differences in reproducibility between iAUC and $K_{\text{trans}}$ (Roberts et al., 2006). In addition to $K_{\text{trans}}$, iAUC has also been selected as the primary parameter-of-interest for clinical DCE-MRI investigations by Quantitative Imaging Biomarker Alliance (QIBA) of Radiological Society of North America (RSNA) (DCE MRI Technical Committee, 2012).

Besides the Tofts models, some other PK models have been developed to improve the physiological description of compartmental models. For example, both Tofts models assume that CA will transfer back from EES to blood plasma immediately after its arrival in EES. As an alternative approach, an adiabatic approximation to the tissue homogeneity (AATH) model was proposed with the assumption that CA concentration in EES changes slowly relative to that in blood plasma (St Lawrence and Lee, 1998). The solution to this model is similar to those of Tofts models with increased mathematical complexities. Some other studies revealed that FXL assumption may not always be valid, especially when the CA concentration is very high and not all CA particles have access to tissue water molecules (Brix et al., 2010). In this circumstance, the simple linear relationship between the longitudinal relaxation rate change $\Delta R_1$ and CA concentration in Eq. (1) is not guaranteed. To account for this issue, the shutter-speed
(SS) model (will be discussed later in Chapter 8) was developed to add the effect of a finite transcytolemmal water exchange rate to the compartment models (Li et al., 2005). The SS model parameterizes the transcytolemmal water exchanges rates as an additional PK parameter and corrects the $K_{trans}$ calculation with a nonlinear expression of CA concentration as a function of $R_i$. So far, the Tofts models are used as the primary PK models for clinical DCE-MRI studies because of their simplicity and straightforward physiological description. In contrast, though models like shutter-speed model are argued to have more realistic physiological details, the complex mathematical theory and unclear interpretation of additional parameters limit the application of these models in DCE-MRI investigations, especially in DCE-MRI therapeutic response assessment studies.

In the current DCE-MRI analysis workflow, the PK parameters ($K_{trans}$, $v_e$ and $v_p$) are commonly solved from the least-squares fitting of Eq. (7) or Eq. (8). The knowledge of $C_p(t)$, which is also known as arterial input function (AIF), must be determined as known information prior to the model fitting. This knowledge can be achieved by imaging a major arterial blood pool during DCE-MRI scan (Port et al., 2001). However, this measurement is usually unavailable in absence of qualified arterial structure in the dynamic imaging field-of-view. A commonly used approach is to use simple mathematical models as approximations of AIF wash-in and wash-out. The bi-
exponential decay is a simple and reliable approach (d’Arcy et al., 2006; Loveless et al., 2012):

\[ C_p(t) = D[a_1 \exp(-m_1 t) + a_2 \exp(-m_2 t)] \] (9)

where \( D \) is the CA administration dose as per unit of bodyweight. The two terms correspond to the fast dynamic equilibrium of CA between blood plasma and EES (represented by \( a_1 \) and \( m_1 \)) and the slow renal removal of CA (represented by \( a_2 \) and \( m_2 \)).

Human AIF can also be modeled as a mixture of two Gaussians and an exponentially modulated sigmoid function based on the population-averaged measurement results (Parker et al., 2006):

\[ C_p(t) = \sum_{n=1}^{2} \frac{A_n}{\sigma_n \sqrt{2\pi}} \exp \left( -\frac{(t - T_n)^2}{2\sigma_n^2} \right) + \alpha \exp(-\beta t) / (1 + \exp(-s(t - \tau))) \] (10)

where \( A_n, T_n \) and \( \sigma_n \) are the constants that describe the position and shape of the \( n^{th} \) Gaussian; \( \alpha \) and \( \beta \) are the amplitude and decay constant of the exponential term; and \( s \) and \( \tau \) are the width and center of the sigmoid. As another approach, some studies proposed a compartmental model fitting method without the knowledge of AIF (Yankeelov et al., 2005a). Instead of using explicit AIF knowledge, this method measures CA concentration time curve of a selected reference region to approximate the CA dynamics in the blood plasma. Nevertheless, this method requires the accurate knowledge of the reference region’s PK parameter values, while such knowledge are still difficult to be justified.
1.3 Challenges facing DCE-MRI and its application in treatment assessment

As an early effort of promoting DCE-MRI clinical application, a QIBA DCE-MRI Technical Committee was formed by academia scientists, imaging device manufacturers, biopharmaceutical industry and professional societies. This committee aims at defining basic technical standards and quality control for consistency DCE-MRI measurement. In the committee’s published profile (DCE MRI Technical Committee, 2012), many technical standards, such as parameter terminology, data archiving and data distribution, have been established. In addition, guidelines regarding hardware configuration, imaging post-processing and parametric image analysis have been proposed for better data reproducibility across imaging platforms. However, these guidelines were primarily designed for clinical trial usage for novel anti-angiogenic agents development. Despite the fact that many phase I trials and investigator-led studies of cancer treatment regimes and vascular-descrupting drugs have incorporated the functional data from DCE-MRI (O’Connor et al., 2012; Padhani, 2002), some challenges concerning the imaging technical difficulties and guidelines of imaging analysis for DCE-MRI therapeutic response assessment remain unsolved.

1.3.1 Temporal resolution

In DCE-MRI, the temporal resolution ($\Delta t$) represents the imaging time of a single 2D/3D volume. The reported clinically feasible temporal resolution ranges from several seconds (brain tissue) to $10^2$ seconds (breast tissue). It is well documented that the
accuracy and precision of the PK analysis is affected by the temporal resolution. To ensure the accurate calculation of $K_{\text{trans}}$ and other PK parameters, higher temporal resolution is desirable. Simulation and imaging studies have revealed that the error of the $K_{\text{trans}}$ calculation increases as the temporal resolution degrades (Heisen et al., 2010; Di Giovanni et al., 2010). Poor temporal resolutions may also introduce intensity uncertainties in the CA concentration time curve measurement. Additionally, suboptimal temporal resolution may affect the individual AIF measurement. Specifically, a fast acquisition with 1 second or less temporal resolution has been claimed for the accurate recording of the rapid wash-in and wash-out processes of the AIF (Henderson et al., 1998). As a result, improving temporal resolution via increasing the image sampling rate is desired in DCE-MRI.

1.3.2 Spatial resolution

There is always a trade-off between spatial resolution and temporal resolution with currently available MR equipment and acquisition techniques (Kuhl et al., 2005). The vast majority of DCE-MRI sampling strategies acquire relatively thick slices with possibly sacrificed in-slice resolution for improving temporal resolution. However, the volume averaging effect makes it difficult to characterize the vascular heterogeneity inside the tumor (Subashi et al., 2013). In treatment assessment with longitudinal imaging studies, the insufficient spatial resolution will reduce the reproducibility of imaging location. Isotropic or quasi-isotropic spatial resolution with <1 mm in-slice
resolution is desired (Turnbull, 2009). Such stringent requirement may be satisfied by reducing the image volume coverage. However, for early stage disease imaging, the selection of limited imaging volume may not capture the region-of-interest with functional abnormality.

### 1.3.3 Accurate $T_1$ measurement

To determine CA concentration, pre-injection $T_{10}$ value and post-injection $T_1$ values must be accurately derived regardless of PK model selection. For dynamic image acquisition, rapid high-resolution $T_1$ mapping techniques must be adopted. Of particular, multiple flip angle method expressed in Eq. (3) has been widely accepted as standard practice with better precision and speed (Cheng and Wright, 2006). However, this method is sensitive to the transmit radiofrequency field ($B_1$) inhomogeneity which may lead to systematic errors in $T_1$ estimation (Deoni et al., 2005). The $B_1$ field inhomogeneity problem could be more prominent at high field strengths. One approach to correct this problem is to derive an actual flip-angle map with additional procedures (Yarnykh, 2007). At the same time, the multiple flip-angle method has been optimized in the consideration of imperfect $B_1$ field with simple $T_1$ correction (Parker et al., 2001; Cheng and Wright, 2006; Sung et al., 2013). In spite of their scientific merits, the application of the improved methods is somewhat limited in the clinical implementation due to additional requirements of sequence parameter optimization.
1.3.4 Arterial input function

The accuracy of AIF information affects the accuracy of functional parameter calculation in the PK analysis (Mustapha et al., 2015). The individual AIF measurement is desired for its direct reflection of patient-specific physiology. This measurement needs to be done in a major arterial structure in which the signal enhancement is sufficiently high. Unfortunately, this is usually not feasible. Due to the stringent requirement of temporal and spatial temporal resolution, it is very difficult to cover a such qualified arterial structure within the dynamic imaging field-of-view. For a limited number of studies with individual AIF measurement capability, the measurement procedure is far from being standardized (Farjam et al., 2014b). Some studies acquired the AIF information outside the DCE-MRI field-of-view through an additional dynamic scan. However, this alternative strategy may not be acceptable because the sampled arterial structure may differ from the vessel supplying the tumor, and thus the accuracy of the acquired AIF is questionable (Li et al., 2011; Parker and Buckley, 2005). In addition, the special procedure was difficult for clinical protocols (Port et al., 2001). Because of these potential issues, the use of population averaged AIF results with simplid mathematical models has become a standard approach. In spite of its popularity, this approach ignores the individual physiological variability and may introduce additional errors in the longitudinal studies for treatment assessment.
1.3.5 Data analysis for treatment assessment

For treatment assessment, PK parameters derived from both pre- and post-treatment DCE-MRI scans need to be acquired. The accurate derivation of PK parametric maps is in demand for accurate therapeutic response assessment. Recently, many advanced MR techniques regarding pulse sequence design, sampling strategy optimization and image reconstruction have been proposed to improve DCE-MRI quality. Nevertheless, the implementation of such techniques in the routine clinical DCE-MRI protocol is still preliminary and challenging. With the potentially compromised data acquisition, the accuracy of PK model fitting could be questionable.

In one of our early works evaluating the therapeutic effect of a novel single-dose preoperative breast radiotherapy, regional mean PK parameters from the Tofts model were compared before and after the treatment delivery (Wang et al., 2015). The Tofts model fitting of gross tumor volume (GTV) was failed. One possible reason is that the image noise effect was strong in the relatively small GTV volume (<1 cc) and the quantitative derivation of CA concentration curve was compromised. In addition, the adopted temporal resolution was about one minute, which is suboptimal in terms of technical capability and yet acceptable in clinical DCE-MRI protocol. Furthermore, the AIF information in this work was approximated by the published bi-exponential decay model, and the ignorance of individual physiological variability in this approach may lead to PK analysis failure and/or erroneous results. Considering these potential
problems, the observed of treatment effects could be unreliable to some extent. On the other hand, model-free DCE-MRI analysis methods are robust and have less requirements on technical quality. Treatment assessment based on verified model-free methods could be accurate and efficient approaches for clinical protocols.

When using PK parameter statistics for therapeutic response assessment, the PK model fitting process is performed on a defined region-of-interest that encompasses the all tumor or with some surrounding normal tissues (Zahra et al., 2007). A single averaged CA concentration time curve can be generated for PK analysis if the region-of-interest is relatively small (Thomas et al., 2005; Liu et al., 2005). Alternatively, the PK parameters can be extracted in a voxel-by-voxel pattern within the ROI. The mean value of a PK parameter across the region-of-interest is commonly evaluated as the primary metric for classic DCE-MRI treatment assessment (Evelhoch et al., 2004). However, a single mean value has limited capability to reveal the anatomical and functional heterogeneity information, which is critical for treatment assessment as the therapeutic induced function changes may vary across different subregions in the region-of-interest (Aerts et al., 2014). Some other first order statistics of PK parametric maps (or DCE images), such as median and variance, as well as histogram descriptors including kurtosis, skewness and interquartile range, may provide the intensity distribution information but has limited spatial heterogeneity information. Some studies have evaluated the treatment response in the separated subregions, while the subregion
identification was based on chosen threshold value with limited justification (Zahra et al., 2009; Marzola et al., 2003). Nevertheless, the treatment assessment based on the current metrics from DCE-MRI PK analysis is still preliminary and challenging.
2. Study focus and the overall structure

This study will investigate new methodology to address limitations of using DCE-MRI in the area of radiotherapy treatment assessment. The structure of this study is divided into two parts. In the first part, the temporal resolution is the topic of interest in DCE-MRI improvement, and two projects are conducted. First, in Chapter 3, a novel iterative MR reconstruction method is investigated to improve DCE-MRI temporal resolution using undersampled k-space data in image acquisition. Specifically, two undersampling strategies are designed to enhance image reconstruction accuracy, and the accuracies of PK parametric maps derived from simulated accelerated DCE-MRI are quantitatively evaluated. In Chapter 4, a new PK model fitting algorithm is proposed for PK parameter calculation of high temporal resolution DCE-MRI. The calculation accuracy and efficiency of this algorithm are compared against the current PK model fitting methods in both computer simulations and in vivo case demonstrations.

In the second part of this study, new DCE-MRI analysis methods are developed for therapeutic response assessment application. Based on the rationale that functional response induced by therapeutic effect could be heterogeneous across the treatment area, PK model-free analysis methods for treatment response assessment using DCE-MRI heterogeneity information are developed using a small animal anti-angiogenesis drug experiment. As an early effort, the potential use of classic fractal dimension analysis as functional heterogeneity evaluation tool is validated in Chapter 5. Then, two
novel image texture analysis methods, gray level local power matrix method (Chapter 6) and dynamic fractal signature dissimilarity method (Chapter 7), are developed and are demonstrated in the same experiment. In the last Chapter 8, the feasibility of combining biological information from multiple PK models for therapeutic response assessment is investigated. In the same aforementioned small animal experiment, the PK parameters from a novel shutter-speed (SS) model are derived and are analyzed for therapeutic response assessment. Subsequently, novel PK biomarkers are designed using PK parameters from SS model and from the commonly used Tofts model.
3. Improving DCE-MRI temporal resolution with a novel total generalized variation (TGV) image reconstruction technique

3.1 Introduction

As discussed in section 1.3, temporal resolution (Δt) is an important factor that affects the DCE-MRI quantitative PK analysis accuracy. High temporal resolution during the DCE-MRI acquisition is always in demand. A straightforward way for improving the temporal resolution is to reduce the imaging field-of-view (FOV) to increase the imaging sampling rate. However, the limited imaging volume may not be optimal for early stage disease imaging where the primary region-of-interest (ROI) is difficult to be determined. In the area of treatment assessment, the reduced FOV imaging may not able to provide sufficient information regarding the normal tissue response when the treatment area is relatively large. Similarly, the reduction of spatial resolution is another simple way by which both the temporal resolution and CNR may be improved. However, precaution needs to be taken for the acceptable level of imaging heterogeneous PK function within the ROI.

So far, various fast imaging concepts and methods have been proposed to for fast MR imaging acquisition. Though many of them were not primarily developed for DCE-MRI, their potential applications in DCE-MRI are always of interest. As a benefit of advanced MR imaging equipment, the concept of parallel imaging was proposed to accelerate the imaging readout with multiple elements of coils. When covering the same
imaging volume, each coil element can be used to acquire the image with reduced acquisition time, and the final image can be solved from the data acquired from each element. In general, the final image can be reconstructed from the aliased images in the image domain (Pruessmann et al., 1999) or from the filled k-space data which is calculated using the acquired partial k-space components in the frequency domain (Cheng, 2012; Griswold et al., 2002). The main drawback of parallel imaging for DCE-MRI is a decrease in signal-to-noise ratio (SNR) which is majorly affected by the number of coil elements.

In DCE-MRI imaging, the keyhole imaging strategy was proposed as an early solution to temporal resolution improvement. Prior to the dynamic acquisition with CA, a full reference set of k-space data is acquired first. Subsequently, after the CA injection, only the central region of k-space is acquired, and each set of low-frequency data is then inserted into the reference data set to reconstruct the high-resolution image. It has been claimed that keyhole imaging is well suited to the large and homogeneous ROI imaging, and the small regions of contrast change cannot be well resolved (d’Arcy et al., 2002). Sliding window (SW) is another technique for improving the temporal resolution of dynamic imaging without loss of spatial resolution (Rasche et al., 1995). This technique acquires the data in a number of incompletely sampled subsets which can be combines to produce a fully sampled k-space data set for reconstruction (Figure 5). Radial sampling and spiral sampling profiles lend themselves well to the SW technique because
of their high sampling ratio near the k-space center. SW can also incorporates the keyhole imaging technique where the center and peripheral regions of k-space are sampled with different frame rates (Subashi et al., 2013).

\[ \text{Figure 5: Concept of sliding window (SW) reconstruction using 2D radial sampling profile. Each full sets at } t_1, t_2, \text{ and } t_3 \text{ consists of four partial acquisitions which is undersampled azimuthally.} \]

In recent years, the theory of compress sensing (CS) proved that at the violation of classic Nyquist theorem, a small sampling fraction in certain patterns is sufficient to reconstruct compressible signals using a linear reconstruction subject to a specified penalty (Lustig et al., 2008; Lustig et al., 2007; Donoho, 2006). To date, CS application in dynamic MR has been successfully demonstrated in many studies. As for DCE-MRI, several studies have been conducted in the past to investigate CA rate constants using the reconstructed images with undersampled patient and animal data (Chen et al., 2010; Smith et al., 2011; Smith et al., 2012; Wang et al., 2010; Han et al., 2012). In these studies, the concept of total variation (TV) based on the first order derivative calculation was commonly adopted in the constrained image reconstruction as to minimize the gradient of the reconstructed image (Block et al., 2007). In spite of its popularity, the concept of
TV relies on the assumption that MR image consists of piecewise constant regions, yet this assumption may not always be valid in human MR imaging with the potential inhomogeneity of exciting field (Knoll et al., 2011). With the existence of staircasing artifacts (Figure 6), the accuracy of TV-based reconstructed image may be compromised. The staircasing artifacts may further affect the accuracy of PK parameter map calculation in DCE-MRI. Another problem in these studies is that the AIF was approximated by published models or previously reported measurements, and the potential errors of individual AIF measurement under the undersampling effect has not been well demonstrated. Furthermore, few studies have been conducted to explore the feasibility of using undersampled k-space data for quantitative blood perfusion analysis which evaluates the blood flow and blood volume in brain tissue.

![Figure 6: An example of staircasing artifact of total variation (TV) processing. The denoise results (red curve) based on the noisy signal (blue curve) suffers staircasing artifact as opposed to ideal diagonal line (not shown).](image)

As a novel mathematic concept derived from TV, total generalized variation (TGV) was recently proposed as with second order derivative in its definition (Bredies et al., 2011).
al., 2010). It has been demonstrated in 2D brain parallel imaging, but the feasibility of its application in DCE-MRI has not been reported. We aimed to explore the feasibility of using undersampled k-space data and TGV image reconstruction method for quantitative DCE-MRI PK analysis. Specifically, eight clinical brain DCE-MRI scans were analyzed retrospectively, and two optimized k-space data sampling patterns were examined. In the PK analysis, two commonly used PK models were analyzed, and the generated PK parametric maps were compared against the ones that are derived from fully sampled image data. The individualized AIF information was adopted in the analysis as for a more comprehensive investigation of quantitative PK analysis with undersampled DCE-MRI data.

3.2 Theory of TGV

Generally, iterative MR image reconstruction is solved as an inverse problem:

$$\hat{u} = \arg\min_u ||F(u) - k||_2$$  \hspace{1cm} (11)

where $F$ is the forward operator, i.e., Fourier transform for MR image reconstruction, $\hat{u}$ is the MR image to be reconstructed and $k$ is the measured k-space data. It is known that subsampling in k-space leads to artifacts in the reconstructed images with specific structures. Thus, penalty terms which are sensitive to aliasing structures need to be included in Eq. (11):

$$\hat{u} = \arg\min_u ||F(u) - k||_2 + \sum_i \beta_i R_i(u)$$  \hspace{1cm} (12)
where $R_i$ represents each penalty term as a function of the image to be reconstructed, and $\beta_i$ is the weighting factor of each penalty term. The definition of the TGV penalty term includes both first order derivative and second order derivative of the image $u$. It is organized as (Bredies et al., 2010):

$$TGV^2(u) = \min_v \left\{ \int_\Omega |\nabla u - v| dx + \alpha \int_\Omega \frac{1}{2} |\nabla v + \nabla v^T| dx \right\}$$

where the minimum is taken over all vector fields on $\Omega$, and $v$ is an introduced test function in the same order as the first-order derivative of the reconstructed image ($\nabla u$). At the neighborhood of edges on the image, the test function $v$ approximates the first order derivative of the image, and the first integration term in Eq. (13) preserves the edge information as the optimal outcome. On the other hand, the symmetric derivative of $v$ in the second integration term of Eq. (13) improves the smoothness of the image based on the observation that second order derivative ($\nabla^2 u$) is locally small within the smooth region of the image. Thus, the TGV penalty term regularizes the features on both smooth and sharp region of the image. The parameter $\alpha$ is used to keep a balance between the first and second order derivatives of the image, and it was argued that $\alpha = 2$ yields the best reconstructed images and does not have to be tuned (Knoll et al., 2011).

By incorporating Eq. (13) into Eq. (12), the TGV-based iterative reconstruction method can be summarized as:

$$\hat{u} = \arg\min_u ||F(u) - k||^2 + \beta TGV^2(u)$$

(14)
which can be approximated using the iterative-regularized Gaussian-Newton (IRGN) method (Uecker et al., 2008; Hohage, 1997; Knoll et al., 2012). At each step $n$ for the given $u^n$, update the $\delta u$ of the minimization problem:

$$\min_{\delta u} ||F'(u^n)\delta u + F(u^n) - k||^2 + 2\beta_k TGV^2(u^n + \delta u)$$

which can be estimated efficiently using the first-order primal dual algorithm (Chambolle and Pock, 2011; Esser et al., 2010; Chan et al., 1999). $\beta_k$ is the TGV term weighting factor and it varies as the iteration number increases.

### 3.3 Materials and methods

#### 3.3.1 Reference data

Eight sets of brain DCE-MRI scans from six patients with recurrent malignant glioma were selected from an IRB-approved clinical study as the reference images. The scans were acquired in the axial plane using a bird-cage quadrature head coil and a 3-dimensional spoiled-gradient recalled-echo (SPGR) sequence on a clinical 1.5T scanner (General Electric, Fairfield, CT). Each set of scan consists of pre-injection calibration volumes at various flip angles and 60 post-injection volumes, and the temporal resolution was about 5.25s of each volume (image matrix size: 256x256; acquisition matrix size: 256x128). The k-space data of the clinical scans was saved as the full sampling data in this study, and the undersampled k-space data was generated via sampling the full data with specific undersampling masks.
3.3.2 Radial-based sampling scheme

Because of a good balance of central/peripheral sampling weights with good low-frequency coverage and potentially improved point spread function (PSF) (Lauzon and Rutt, 1998; Nagel et al., 2009), an optimized radial-based undersampling profile with multiple spokes was firstly investigated. Specifically, in this work, the undersampled k-space data was generated by sampling each slice of the full sampling data with a 32 non-Cartesian-ray radial sampling profile defined on a Cartesian grid (Smith et al., 2012). Related to the Golden Ratio, the sampling profile was generated by a constant azimuthal increment of $\Delta \beta = 111.25^\circ$ (Figure 7) (Winkelmann et al., 2007). This sampling pattern acquired approximately 11.5% of full k-space and simulated an accelerated image acquisition by a factor of $128/32 = 4$. The simulated acquisition time using the presented radial sampling pattern was about $5.25/4 = 1.31s$. Compared to the even-distributed sampling profile, the sampling efficiency of this golden ratio based profile has no difference (Tsai and Nishimura, 2000). Radial rays were evenly spaced with time, and only three different azimuthal gaps occurred in comparison with the evenly distributed rays (Kohler, 2004). Thus, this sampling profile allows a posteriori adjustment of the temporal resolution (Kazantsev et al., 2005).
3.3.3 Cartesian-based sampling scheme

Although radial sampling profile has its advantages for limited k-space acquisition, it may suffer from gradient delay, off-resonance effects and strong streaks, which require additional physics works for implementation. Besides, image reconstruction from radially sampled data requires certain regridding approaches, which are often far from trivial. (Rasche et al., 1995). In contrast, sampling profiles defined in Cartesian coordinate system are straightforward and are suitable for easy implementation. However, because of evenly distributed sampling density, more readout lines are required to achieve an adequate k-space coverage for backprojection. Early efforts were made to integrate SW technique into Cartesian undersampling profile sequences (d’Arcy et al., 2002). Although the feasibility was demonstrated, the achieved acceleration factor was limited without proper indication of possible application in DCE-MRI.

Based on above issues, a novel Cartesian-based sampling profile sequence with integrated SW technique was designed as the 2\textsuperscript{nd} sampling scheme of this work. First, a
variable Poisson Ellipsoid sampling density function was built as demonstrated in Figure 8.

![Illustration of a 2D sampling profile with variable sampling density. The central red region is fully sampled and the rest in gray level is sampled with varying sampling density.](image)

For each 2D profile, the central 15% of k-space along the phase encoding direction was fully sampled (indicated by red color). Then, the rest k-space region was treated as high frequency (HF) sampling region, and each side of the region was evenly divided into four subregions along the phase encoding direction. Each subregion was assigned a uniform sampling density \( \rho_i \) which decreases exponentially from k-space center (indicated by gray levels with colorbar):

\[
\rho_i = \rho_0 e^{-r_i/\tau} \tag{16}
\]

where \( \rho_0 \) is the scaling factor to ensure the probability density normalization. The parameter \( r_i \) indicates the relative length (in phase encoding direction) if the \( i \)th HF subregion in reference to the total HF region length, and the dimensionless decay constant \( \tau \) was chosen as 1. In this work, HF data in 10% of the original k-space size was
acquired in the designated HF region. Thus, an acceleration of 4 was achieved by sampling each 2D reference data with 25% data size (15% in central region, 10% in HF region).

For a series of 2D k-space data at a certain location in the dynamic scan, the SW technique was applied by inserting HF data acquired in two adjacent frames into the HF region of the k-space to be constructed. Figure 9 shows the schematic of SW in HF region. In the first row, at each time point, k-space data was acquired using the variable density profile in Figure 8. For the k-space to be reconstructed at time point \( t_1 \) in the 2nd row, the HF region was filled by HF data from \( t_0, t_1 \) and \( t_2 \) points, while central k-space region was filled with data only from \( t_1 \) central k-space data. To achieve a higher k-space coverage ratio, temporal constraints were added to Eq. (16) to ensure that sampled HF regions have no overlap within each k-space triplet. Thus, the filled k-space data at each time point was in about 45% of the original k-space data size.
3.3.4 Image reconstruction

Using each data sampling scheme, the undersampled data was then used for the image reconstruction on Cartesian grid as Eqs. (12) - (14). At the beginning of the reconstruction, no reference image was needed. The initial value of the TGV term weighting factor $\beta_k$ in Eq. (15) was set as 1 and its attenuation factor along iterations was empirically chosen as 5. The maximum iteration step number was set to 8 in a balanced consideration of acceptable image reconstruction accuracy as well as tolerable computation cost.
3.3.5 Pharmacokinetic analysis

PK analysis with same procedures was conducted using the reference clinical image series as well as the image series reconstructed from undersampled k-space data with two difference sampling schemes, respectively. The pre-injection $T_{10}$ map and post-injection $T_1(t)$ were calculated voxel-by-voxel using the standard dual flip angle method as in Eq. (3), and flip angle 5° and 15° were used. Then, CA concentration maps at each post-injection time point was calculated by Eq. (1), where the longitudinal relaxivity of the CA at $B_0 = 1.5T$ was 4.3 mM-1s-1. To improve the PK model fitting quality, an improved texture preserving variational denoising method was applied to the CA concentration maps to reduce the spatial noise effect (Gilboa et al., 2003). As a typical approach of voxel-by-voxel analysis, the spatial averaging strategy has been reported in various clinical studies for PK analysis (Cabrera et al., 2013; Sourbron et al., 2009). In this texture preserving denoising method, the noise reduction was optimized with a spatially varying local power constraint $P_z$ on image $I$ (Bauerle et al., 2010; Gilboa et al., 2003):

$$P_z(x,y) = \int_{\Omega} \left( I(\tilde{x},\tilde{y}) - E(I(\tilde{x},\tilde{y})) \right)^2 w_{x,y}(\tilde{x},\tilde{y}) d\tilde{x} d\tilde{y}$$

and $\int_{\Omega} P_z(x,y) dx dy = var(I)$

where $w_{x,y}(\tilde{x},\tilde{y})$ is a radially symmetric smoothing window centered at point-of-interest $(x,y)$, and $E(\cdot)$ denotes the expected value. The detailed mathematical deduction is not discussed in this thesis due to the specific scope; as the key point, this method has been reported to improve the image quality in terms of SNR at no change of spatial
patterns in the comparative study against the classic variational denoising technique and moving average kernel (Buades et al., 2005; Gilboa et al., 2003).

Two commonly used PK models were investigated in this study. First, the extended Tofts model (Eq. (8)) was analyzed to quantify $K_{\text{trans}}$, which has served as the primary quantitative imaging biomarker in many oncologic DCE-MRI studies. Then, the two-compartment exchange model was used to quantify the blood flow and blood volume (Brix et al., 2004; Sourbron et al., 2009). Similar to the schematic shown in Figure 4, this two-compartment exchange model describe the conservation of CA mass as:

$$C_t(t) = F_B \int_0^t C_p(u) \cdot [e^{-K_+(t-u)} + E(e^{-K_-(t-u)} - e^{-K_+(t-u)})] \, du \tag{18}$$

where $F_B$ stands for blood flow, $K_{\pm}$ are related to the rate constants of CA exchange between blood plasma and tissue-of-interest, and $E$ is a parameter associated with $K_{\pm}$ and circulation transit time $T_T$. The AIF information $C_p(t)$ in this work was measured in the way as described below: first, an arterial structure within the intracranial imaging volume was carefully contoured in 3D fashion by a clinical professional. Then a histogram of all arterial structure voxels’ time-to-peak ($T_{\text{max}}$) values was generated (Farjam et al., 2014a). Then, Next, the $T_{\text{max}}$ of the $C_p(t)$ was determined by the peak position of the generated histogram, and the $C_p(t)$ was calculated by averaging the CA concentration of the voxels that constituted the histogram’s peak. Fitting Eq. (18) with $C_p(t)$ and $C_t(t)$ by Levenberg-Marquardt algorithm with zero initial conditions, $T_T$ can be calculated as:
\[ T_T^{-1} = K_+ - E(K_+ - K_-) \]  \hspace{1cm} (19)

and blood volume \( V_B \) can be estimated as:

\[ V_B = F_B \cdot T_T \]  \hspace{1cm} (20)

### 3.3.6 Quantitative Accuracy Analysis

The investigated PK parameters \((K_{\text{trans}}, F_B \) and \( V_B)\) were calculated in a voxel-by-voxel pattern, and two sets of parameter maps were generated from the original clinical scan image (full sampling set) and the reconstructed images using undersampled data (reduced sampling set), respectively. For the purpose of quantitative accuracy evaluation, a 3D region-of-interest (ROI) was defined using Gross Tumor Volume (GTV) from the radiotherapy treatment planning system. As the most commonly reported metrics in clinical studies, the means of \( K_{\text{trans}}, F_B \) and \( V_B \) across the ROI were compared for relative differences. We used the ‘error in volume mean’ (EIVM) for the PK parameters estimation accuracy (Smith et al., 2011):

\[
EIVM(P) = \frac{\overline{P_r(x,y,z)} - \overline{P_f(x,y,z)}}{\overline{P_f(x,y,z)}} \times 100\% \]  \hspace{1cm} (21)

where the \( P_r(x,y,z) \) and \( P_f(x,y,z) \) denote a certain PK parameter’s distributions within the ROI from the reduced sampling set and the full sampling set, respectively, and the overhead bar represents the average across ROI. The accuracy was further evaluated with the ROI difference map of each PK parameter, cross-correlation (CC), and total relative error (TRE) defined as (Chang and Xiang, 2006):
\[ TRE = \sqrt{\frac{\sum_{x,y,z}[P_r(x,y,z) - P_f(x,y,z)]^2}{\sum_{x,y,z}[P_f(x,y,z)]^2}} \]  

(22)

where the \( P_r(x, y, z) \) and \( P_f(x, y, z) \) follow the same definition as in Eq. (21). To determine the significance levels of the differences between the PK parameter results using full sampling data and reduced sampling data, Wilcoxon signed-rank test was adopted. Statistical significance was defined as \( \alpha < 0.05 \) with multiple comparison correction (Chen et al., 2007).

3.4 Results

3.4.1 Radial-based sampling scheme results

This section presents the results using the radial-based undersampling scheme which has been explained in section 3.3.2. Figure 10 shows a 2D comparison of the original clinical DCE image (a), the reconstructed image using the undersampled data (b) and the difference map (c) from a selected patient. The images were from a post-injection volume, and the red contour line indicates the ROI for quantitative analysis. As can be seen, (b) preserved most structural details in (a), though some details were lost at high intensity region. The average TRE value of all patients’ reconstructed images was about 0.113. Figure 11 presents the comparison of measured AIF information, and \( C_{p,f}(t) \) and \( C_{p,r}(t) \) denote the AIF measured on the CA concentration maps derived from original clinical image (blue) and the reconstructed image (red), respectively. As can be seen, the peak positions of the two curves were identical, and the peak enhancements of
$C_{p,f}(t)$ and $C_{p,r}(t)$ were 0.350 mmol/L and 0.366 mmol/L, respectively. The shape features of $C_{p,f}(t)$ were generally preserved by $C_{p,r}(t)$, though $C_{p,r}(t)$ had an overestimated intensity both before and after the first pass peak.

Figure 10: A comparison of the original post-injection image (a) and the reconstructed image using the radially undersampled data (b). The intensity difference between (a) and (b) is shown in (c). The red contour indicates the location of GTV as region-of-interest (ROI).

Figure 11: The comparison of the measured $C_p(t)$ from the CA concentration maps calculated from original clinical images (blue curve) and the images reconstructed using reduced sampling data with radial-based scheme (red curve)
Figure 12 demonstrates the comparison of $K_{\text{trans}}$ maps that correspond to the same imaging position shown in Figure 10. The $K_{\text{trans}}$ map calculated using the full sampling data and using the reduced sampling data with radial-based sampling scheme are presented in (a) and (b), respectively. The histogram (c) presents the distribution of the percent difference of the 3D ROI $K_{\text{trans}}$ in (b) versus (a). The percent difference map within the ROI is shown in (d). As can be observed, the two $K_{\text{trans}}$ maps were morphologically similar, and a $K_{\text{trans}}$ intensity-elevated region within the ROI (indicated by white arrows) was clearly identified on both maps. The heterogeneity of $K_{\text{trans}}$ intensity within the identified region in (a) was successfully reproduced in (b). The histogram in (c) suggested an overall acceptable 3D $K_{\text{trans}}$ estimation accuracy using the reduced sampling data. The median value of absolute percent difference in (c) was 6.70%. The difference map in (d) also suggested a good voxel-based quantitative calculation accuracy of $K_{\text{trans}}$ at the center of ROI, though minor discrepancies of $K_{\text{trans}}$ underestimation in (b) were found at ROI peripheral region.
Figure 12: A comparison of $K_{\text{trans}}$ map calculated using the original full sampling data (a) and using the reduced sampling data with radial-based scheme (b); The histogram about relative error of (b) versus (a) of all 3D ROI voxels (c); The percent difference map of the ROI region (d).

Figure 13 demonstrates the blood flow $F_B$ results in the same organization as Figure 12. On both Figure 13 (a) and (b), a circular region with elevated blood flow was identified approximately at the $K_{\text{trans}}$ hot region shown in Figure 12 (a) and (b).

Compared to the $F_B$-elevated region in (a), the identified region in (b) had a minor difference at its right posterior boundary. The difference map in (d) shows a good accuracy of $F_B$ estimation in (b), and the left anterior region indicated a minor $F_B$
underestimation in (a). The results of blood volume $V_B$ are illustrated in Figure 14. As is
the pattern indicated by $K^{\text{trans}}$ and $F_B$ results, a region with elevated $V_B$ values was
indicated in both (a) and (b). The $V_B$ difference map (d) also exhibited the good $V_B$
calculation accuracy using the reduced sampling data set.

![Figure 13](image_url)

Figure 13: A comparison of $F_B$ map calculated using the original full sampling data (a)
and using the reduced sampling data with radial-based scheme (b); The histogram
about relative error of (b) versus (a) of all 3D ROI voxels (c); The percent difference
map of the ROI region (d)
Figure 14: A comparison of $V_B$ map calculated using the original full sampling data (a) and using the reduced sampling data with radial-based scheme (b); The histogram about relative error of (b) versus (a) of all 3D ROI voxels (c); The percent difference map of the ROI region (d).

Table 1 summarizes each PK parameter’s EIVM, TRE and CC within the defined ROI of all 8 scans. The EIVM values demonstrated in Table 1 were all around 5% or less except one $V_B$ EIVM value that indicated an 8.7% error. These results suggest that just as the commonly reported metrics in clinical studies, the mean values of the investigated PK parameters can be accurately calculated using the undersampled data. The TRE values varied from 0.08 to 0.15 with an average value of 0.11. As a description of the
morphological similarity of two images, most CC values were higher than 0.95, and the observed smallest CC was 0.83. In the Wilcoxon signed-rank tests, the statistical significance level of differences of $K_{\text{trans}}$, $F_B$ and $V_B$ were 0.528, 0.889 and 0.123, respectively. These results suggested that there is no significant differences between the PK maps calculated from the full sampling set and the one calculated from the reduced sampling set.

Table 1: The results of Error in Volume Mean (EIVM), Total Relative Error (TRE) and Cross-Correlation (CC) within the 3D ROI using radial-based undersampling scheme.

<table>
<thead>
<tr>
<th>Patient No.</th>
<th>$K_{\text{trans}}$</th>
<th>$F_B$</th>
<th>$V_B$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>EIVM</td>
<td>TRE</td>
<td>CC</td>
</tr>
<tr>
<td>1</td>
<td>-0.65%</td>
<td>0.108</td>
<td>0.951</td>
</tr>
<tr>
<td>2</td>
<td>5.17%</td>
<td>0.110</td>
<td>0.977</td>
</tr>
<tr>
<td>3*</td>
<td>-1.53%</td>
<td>0.104</td>
<td>0.990</td>
</tr>
<tr>
<td>3#</td>
<td>0.85%</td>
<td>0.129</td>
<td>0.979</td>
</tr>
<tr>
<td>4</td>
<td>2.66%</td>
<td>0.102</td>
<td>0.960</td>
</tr>
<tr>
<td>4#</td>
<td>-1.76%</td>
<td>0.129</td>
<td>0.926</td>
</tr>
<tr>
<td>5</td>
<td>5.90%</td>
<td>0.103</td>
<td>0.941</td>
</tr>
<tr>
<td>6</td>
<td>&lt;0.01%</td>
<td>0.133</td>
<td>0.944</td>
</tr>
</tbody>
</table>

3.4.2 Cartesian-based sampling scheme results

This section presents the results using the Cartesian-based undersampling scheme which has been explained in section 3.3.3. Figure 15 shows the comparison of reference DCE image (a) and the reconstructed image using undersampled data (b) in the same organization as Figure 10. The DCE image in (b) preserves most structural details in (a), and the difference map in (c) indicates the discrepancies at boundaries. The
average TRE values of all reconstructed images using the Cartesian undersampling scheme was about 0.109, and this number was very close to the corresponding number (0.113) when using the aforementioned radial-based sampling scheme.

Figure 15: A comparison of the original post-injection image (a) and the reconstructed image using the Cartesian-based undersampled data (b). The intensity difference between (a) and (b) is shown in (c). The red contour indicates the location of GTV as region-of-interest (ROI).

Figure 16: The comparison of the measured $C_p(t)$ from the CA concentration maps calculated from original clinical images (blue curve) and the images reconstructed using reduced sampling data with Cartesian-based scheme (red curve).

Figure 16 presents the comparison of measured AIF information from original clinical image (blue) and the reconstructed image (red), respectively. Similar to the
results in Figure 11, the red curve has same peak position and fluctuation patterns as of the blue curve. The peak enhancements the red curve was 0.382 mmol/L, which is more deviant from the 0.350 mmol/L value of the blue curve than the red curve in Figure 11 (0.366 mmol/L) with radial-based sampling scheme.

Figure 17, Figure 18 and Figure 19 summarizes the $K_{\text{trans}}$, $F_b$ and $V_b$ results in the same fashion as the corresponding figures in section 3.4.1. For a short conclusion, all three PK parametric maps that were reconstructed using the Cartesian-based undersampled data were very similar to the reference maps in terms of morphological features. For quantitative evaluations, each PK parameter’s EIVM, TRE and CC of all 8 scans are summarized as Table 2. The EIVM results in Table 2 were all around 5% or less, but in some cases the results were worse than the corresponding values in Table 1. The largest EVIM value in Table 2 is 16.67%. All observed CC values in Table 2 were higher than 0.9, and this result indicates the good morphological match of PK parametric maps using undersampled data. The average TRE values of all calculated PK parametric maps was about 0.12, and this number is higher than the corresponding result (0.11) using the radial sampling scheme. No statistically significant results were found between the mean $K_{\text{trans}}$ ($p = 0.950$), $F_b$ ($p = 0.401$) and $V_b$ ($p = 0.888$) using reference data and using the undersampled data with Cartesian-based scheme.
Table 2: The results of Error in Volume Mean (EIVM), Total Relative Error (TRE) and Cross-Correlation (CC) within the 3D ROI using Cartesian-based undersampling scheme.

* Images are selected from this scan; #Second scan of this patient

<table>
<thead>
<tr>
<th>Patient No.</th>
<th>$K^{trans}$</th>
<th>$F_B$</th>
<th>$V_B$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>EIVM</td>
<td>TRE</td>
<td>CC</td>
</tr>
<tr>
<td>1</td>
<td>6.92%</td>
<td>0.135</td>
<td>0.982</td>
</tr>
<tr>
<td>2</td>
<td>1.40%</td>
<td>0.135</td>
<td>0.967</td>
</tr>
<tr>
<td>3*</td>
<td>2.01%</td>
<td>0.082</td>
<td>0.994</td>
</tr>
<tr>
<td>3#</td>
<td>4.55%</td>
<td>0.147</td>
<td>0.935</td>
</tr>
<tr>
<td>4</td>
<td>-1.45%</td>
<td>0.137</td>
<td>0.980</td>
</tr>
<tr>
<td>4#</td>
<td>-4.41%</td>
<td>0.139</td>
<td>0.938</td>
</tr>
<tr>
<td>5</td>
<td>-7.89%</td>
<td>0.108</td>
<td>0.990</td>
</tr>
<tr>
<td>6</td>
<td>2.81%</td>
<td>0.111</td>
<td>0.988</td>
</tr>
</tbody>
</table>
Figure 17: A comparison of $K_{trans}$ map calculated using the original full sampling data (a) and using the reduced sampling data with Cartesian-based scheme (b); The histogram about relative error of (b) versus (a) of all 3D ROI voxels (c); The percent difference map of the ROI region (d)
Figure 18: A comparison of $F_b$ map calculated using the original full sampling data (a) and using the reduced sampling data with Cartesian-based scheme (b); The histogram about relative error of (b) versus (a) of all 3D ROI voxels (c); The percent difference map of the ROI region (d)
Figure 19: A comparison of $V_B$ map calculated using the original full sampling data (a) and using the reduced sampling data with Cartesian-based scheme (b); The histogram about relative error of (b) versus (a) of all 3D ROI voxels (c); The percent difference map of the ROI region (d)
3.5 Discussion

In this pilot work, the feasibility of using a novel TGV iterative MR reconstruction technique with undersampled k-space data for quantitative PK analysis in DCE-MRI study was investigated. At the simulated acceleration factor of 4, the PK parameters can be accurately calculated from undersampled data in terms of EIVM and TRE evaluations. Through the comparison of figures shown in sections 3.4.1 and 3.4.2, the PK parametric maps from the reduced sampling data could be believed to be reliable for clinical radiotherapy use including boost region delineation and treatment response monitoring. One problem of this work is that only 8 scans from 6 patients were analyzed. This is because to ensure individualized AIF measurement, the lesions have to be close to a major intracranial arterial structure within a limited imaging volume with acceptable spatiotemporal resolution. Because of this rigid requirement, only a small number of scans were qualified for this pilot study. Since the current work is an early demonstration of the proposed method, 8 cases is a good start for pilot study. Certainly, future works with a much larger number of data sets on various tumors are essential towards the clinical application of the proposed DCE-MRI acceleration method.

An acceleration factor of 4 was achieved using both a radial-based multi-ray sampling profile and a Cartesian-based variable density sampling profile with SW technique. To ensure the accurate DCE image reconstruction, central k-space region needs to be sampled properly. Because of the superior central k-space sampling scheme,
the radial sampling profile could use fewer rays than the Cartesian sampling profile to
reach an acceptable central/peripheral k-space coverage balance. Thus, to achieve a same
acceleration factor, SW technique becomes important to guarantee the adequate k-space
coverage. Even though, the comparison of the numerical TRE and EIVM results in Table
1 and Table 2 suggest that the accuracy of PK parameters using Cartesian-based
undersampled data was generally worse than using the radial-based undersampled
data. To further explain this point, another comparison study using the illustrated scan
in sections 3.4.1 and 3.4.2 was conducted with difference undersampling strategy. In
Table 3, results in the first row and in the last row have been demonstrated in Table 1
and Table 2, respectively. When using the Cartesian-based variable sampling profile
without SW technique, the accuracies of all three PK parameters degrade as the total k-
space sampling ratio decreased from 50% to 25%. With the adopted SW strategy in this
work, at 25% sampling ratio, the EIVM and TRE values of all three parameters were
much improved than the results without SW technique. In this work, the effective k-
space coverage at each time point after SW adopted is about 45%; the PK parameter
accuracy in EIVM and TRE of the investigated strategy (last row in Table 3) is between
the results with 50% k-space coverage and 33% k-space coverage. As discussed in
section 3.3.3, the implement of radial sampling scheme could be challenging for clinical
application, but its application of DCE-MRI acceleration has been demonstrated as very
promising; on the other hand, as suggested by this work, the Cartesian-based
undersampling scheme for DCE-MRI acceleration is feasible and yet needs to be further explored and improved for better performance.

Table 3: The comparison of PK parameter accuracies of a selected patient with different sampling strategy.

<table>
<thead>
<tr>
<th>Sampling Strategy</th>
<th>$K_{\text{trans}}$</th>
<th></th>
<th></th>
<th>$F_B$</th>
<th></th>
<th></th>
<th>$V_B$</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>EIVM</td>
<td>TRE</td>
<td>CC</td>
<td>EIVM</td>
<td>TRE</td>
<td>CC</td>
<td>EIVM</td>
<td>TRE</td>
</tr>
<tr>
<td>Radial, 32-ray</td>
<td>-1.53%</td>
<td>0.104</td>
<td>0.990</td>
<td>-3.00%</td>
<td>0.109</td>
<td>0.950</td>
<td>-0.12%</td>
<td>0.096</td>
</tr>
<tr>
<td>Cartesian 50%</td>
<td>-2.08%</td>
<td>0.057</td>
<td>0.999</td>
<td>0.70%</td>
<td>0.042</td>
<td>0.969</td>
<td>-2.56%</td>
<td>0.038</td>
</tr>
<tr>
<td>Cartesian 33%</td>
<td>7.89%</td>
<td>0.119</td>
<td>0.993</td>
<td>2.22%</td>
<td>0.143</td>
<td>0.955</td>
<td>8.57%</td>
<td>0.149</td>
</tr>
<tr>
<td>Cartesian 25%</td>
<td>8.46%</td>
<td>0.154</td>
<td>0.956</td>
<td>5.97%</td>
<td>0.179</td>
<td>0.902</td>
<td>9.50%</td>
<td>0.172</td>
</tr>
<tr>
<td>Cartesian 25%, SW</td>
<td>2.01%</td>
<td>0.082</td>
<td>0.994</td>
<td>5.02%</td>
<td>0.110</td>
<td>0.977</td>
<td>0.80%</td>
<td>0.084</td>
</tr>
</tbody>
</table>

Currently, the simulated undersampled k-space data was generated by 2D-based sampling profiles, which is a good start as a feasibility study demonstrating the TGV applicability in DCE-MRI acceleration. As a natural extension, the 3D radial/Cartesian sampling strategy may further reduce the scan time, though the second order derivative calculation required by TGV penalty term may become very intensive in 3D image reconstruction. Additionally, the temporal constraints based on TGV or other theories can be further explored to the iterative reconstruction of multiple 3D volumes at the same time. These research directions could be potential future works for further DCE-MRI temporal resolution improvement.

3.6 Conclusion

The feasibility of accelerated brain DCE-MRI has been successfully demonstrated in the quantitative PK parameters. As shown in this work, the quantitative parameters
from two commonly adopted PK models can be accurately calculated from the images reconstructed with two different k-space undersampling schemes at a simulated acceleration factor of 4. The investigated acceleration methods can be used for reliable clinical image acquisition. Future works with a large population are in demand towards the exploration before routine clinical implementation.
4. Development of an efficient calculation method for high temporal resolution DCE-MRI analysis

4.1 Introduction

The quantification of the CA PK parameters and the physiological interpretation of the derived parameters rely on accurate PK model analysis. As was discussed in the section 1.3.1, the temporal resolution of dynamic image acquisition ($\Delta t$) is one of the key factors that affect the accuracy and precision of PK parameters. From the mathematics perspective, higher temporal resolution as with short image acquisition time is desirable. The optimal temporal resolution would be less than 1 second to ensure the potential measurement of individual AIF (Henderson et al., 1998). Currently, the reported temporal resolution ranges from several seconds to several minutes, depending on the imaging site and volume.

Based on this current temporal resolution range, the PK parameters are commonly calculated with curve fitting based on the nonlinear least-squares (NLSQ) methods, and these methods have been incorporated into commercial software packages. In spite of their popularity, NLSQ based curve fitting methods require intensive computation and may lead to erroneous results at the local optima due to uncertainties in the iterative process (Ahearn et al., 2005). As an alternative approach, a linear least-squares method based on the integral form (ILLSQ) of the extended Tofts model was proposed (Murase, 2004). Simulation results suggested that when the signal-to-noise ratio (SNR) of CA concentration curve is less than 10, the ILLSQ method is
superior to the NLSQ methods in aspects of calculation accuracy and efficiency. However, the computation time required by this ILLSQ method is sensitive to the data size. For a fixed dynamic imaging scan time, the number of data point in a CA concentration time curve goes up with improved temporal resolutions. In the ILLSQ method, the computation time required for the integral calculation rapidly increases in the quadratic fashion (Weiss, 2002). As a result, the efficiency of the ILLSQ method may be an issue when the temporal resolution of image acquisition reaches the desired 1 second or less.

To further improve the PK model fitting computational efficiency, a new method for calculating PK parameters for brain DCE-MRI analysis was proposed in this study. In this method, curve fitting based on linear least-squares method was applied to the derivative expression of the extended Tofts model (abbreviated as the DLLSQ method). The accuracy and the efficiency of the proposed DLLSQ method was evaluated against the current ILLSQ and NLSQ methods in the designed simulation studies. The potential clinical application of the DLLSQ method was examined in in vivo brain DCE-MRI cases.

4.2 Materials and methods

4.2.1 Theory of the DLLSQ method

The commonly used extended Tofts model in DCE-MRI was expressed as Eq. (8) above:
\[ C_t(t) = v_p C_p(t) + K^{\text{trans}} \int_0^t C_p(t') \exp[-\frac{K^{\text{trans}}}{v_e}(t - t')] dt' \] (8)

As was discussed above, in the ILLSQ method and NLSQ methods, the integration calculation in the above equation could be time-consuming when the temporal resolution is high. To avoid this integration calculation, the DLLSQ method rewrite Eq. (8) into a derivative form with zero initial conditions:

\[ \frac{dC_t(t)}{dt} = (K^{\text{trans}} + v_p \cdot k_{ep}) \cdot C_p(t) - k_{ep} \cdot C_e(t) + v_p \cdot \frac{dC_p(t)}{dt} \] (23)

where \( k_{ep} = \frac{K^{\text{trans}}}{v_e} \). Compared to the integration calculation, the derivative calculation may be more sensitive than the integration calculation to noise in CA concentration time curves. Studies have indicated that when the temporal resolution was high, the noise in the CA concentration measurements were mainly associated with relatively higher frequencies compared to the features of CA concentration dynamics (Kershaw and Buckley, 2006; Lavini et al., 2007; Schabel and Parker, 2008). To minimize the noise effects, a low-pass filter is in demand to reduce the high frequency intensity fluctuations in the CA concentration curves. The basic requirements for the filtering process should 1) be simple; 2) use minimal implementation time; and 3) preserve the quick wash-in and wash-out features of the CA concentration curves at high temporal resolutions. As an equivalent iterative moving average (MA) filter, the Kolmogorov-Zurbenko (KZ) filter was selected, fulfilling the aforementioned requirements (Yang and Zurbenko, 2010). This filter processes both \( C_p(t) \) and \( C_e(t) \) in the time domain with two parameters, \( m \) and \( k \):
\[
\tilde{C}_X(t) = \sum_{s=-k(m-1)/2}^{k(m-1)/2} \frac{a_s^{m,k}}{m^k} C_X(t + s)
\] (24)

where the coefficients of the convolution kernel \(a_s^{m,k}\) are given by the following polynomial equation:

\[
\sum_{s=-k(m-1)/2}^{k(m-1)/2} t^{s+k(m-1)/2} \cdot a_s^{m,k} = (1 + t + \cdots + t^{m-1})^k
\] (25)

The parameter \(k\) defines the equivalent MA iteration times, and the odd parameter \(m\) defines the width of the MA rectangular window size. Thus, the coefficients \(a_s^{m,k}\) constitute a tapering window which has a finite support of \((m-1)k + 1\).

Generally, the KZ filter is heavily weighted on a length of \(m\sqrt{k}\) around the time point \(t\) and vanishes to zero outside that range, and the effective low-pass cutoff frequency of the KZ filter is approximately \(2\sqrt{6\frac{1-\alpha^{1/2k}}{m^2 - \alpha^{1/2k}}}\), where \(\alpha\) is a pre-selected value between 0 and 1 (Zurbenko, 1986). The computation cost of the KZ filtering is low because of its convolution implementation in time space with no additional transform.

After KZ filtering, Eq. (11), with the processed \(\tilde{C}_p(t)\) and \(\tilde{C}_t(t)\), can be reorganized into a matrix format:

\[
B_D = (A_D^T A_D)^{-1} \cdot A_D^T \cdot \tilde{C}_D
\] (26)

where
\[
\begin{bmatrix}
\bar{C}_p(t_1) - \bar{C}_t(t_1) & \frac{d\bar{C}_p}{dt}(t_1) \\
\bar{C}_p(t_2) - \bar{C}_t(t_2) & \frac{d\bar{C}_p}{dt}(t_2) \\
\vdots & \vdots \\
\bar{C}_p(t_{n-1}) - \bar{C}_t(t_{n-1}) & \frac{d\bar{C}_p}{dt}(t_{n-1})
\end{bmatrix}
\]

\[\hat{A}_D = \begin{bmatrix}
\end{bmatrix}\]

\[
\bar{B}_D = \begin{bmatrix}
K^{\text{trans}} + v_p \cdot k_{ep} \\
k_{ep} \\
v_p
\end{bmatrix}
\]

and

\[
\hat{C}_D = \begin{bmatrix}
\frac{d\bar{C}_t}{dt}(t_1) \\
\frac{d\bar{C}_t}{dt}(t_2) \\
\vdots \\
\frac{d\bar{C}_t}{dt}(t_{n-1})
\end{bmatrix}
\]

The PK parameters \(K^{\text{trans}}, k_{ep}, \) and \(v_p\) were solved using Eq. (23). For numeric implementation, the derivative terms in the matrices \(\hat{A}_D\) and \(\hat{C}_D\) were estimated as

\[
d\bar{C}_X(t_i)/dt = (\bar{C}_X(t_{i+1}) - \bar{C}_X(t_i))/\Delta t, \quad i = 1, 2, \ldots, n - 1.
\]

For a pair of \((\bar{C}_p, \bar{C}_t)\) with \(n\) sampled time points, the computation time required by the derivative calculation in Eq. (23) increases linearly with \(n\), while the computation time required by the integral calculation in Eq. (8) is proportional to \(n^2\) (Weiss, 2002). Thus, the derivative calculation in Eq. (23) is more efficient than integration calculation in Eq. (8) when \(n\) is a large number (i.e., when the temporal resolution is high).
4.2.2 Single voxel simulation

To demonstrate the efficiency and accuracy using the DLLSQ method for calculation of the PK parameters, simulations based on a single voxel were conducted with a group of representative brain PK parameters. First, the AIF information was modeled as \( \hat{C}_p(t) = AIF(t) \cdot (1 - Hct) \), where \( AIF(t) \) was modeled by the published population-average result in Eq. (10), and \( Hct \) is the hematocrit with an assumed value of 0.45 (Jesberger et al., 2006). Next, the noiseless CA concentration curve in the simulated tissue voxel \( \hat{C}_t(t) \) was generated using Eq. (8) with the following PK parameters: \( K_{trans} = 0.0890 \text{ min}^{-1} \), \( k_p = 0.3201 \text{ min}^{-1} \), and \( v_p = 0.0230 \), which were chosen from the population average of reported PK values in previous intracranial DCE-MRI studies (Cabrera et al., 2013; Armitage et al., 2007). Both \( \hat{C}_p(t) \) and \( \hat{C}_t(t) \) were firstly generated using a 5-minute total post-injection sampling time with a temporal resolution of \( \Delta t=10\text{ms} \). Then a Gaussian-distributed random noise was added to the concentration curves \( \hat{C}_p(t) \) and \( \hat{C}_t(t) \) to obtain different levels of CNR, where the CNR was defined as the ratio of the peak CA enhancement (the maximum intensity in a concentration curve) to the standard deviation of the Gaussian noise intensity (Luypaert et al., 2010). The curves \( C_p(t) \) and \( C_t(t) \) used in the simulation were generated by sampling \( \hat{C}_p(t) \) and \( \hat{C}_t(t) \) with the designed temporal resolutions.

During the simulation, the DLLSQ method was investigated against the current ILLSQ and NLSQ methods for a certain range of \( \Delta t \) and CNR. For the NLSQ method, the
Levenberg-Marquardt algorithm was selected for comparison (Moré, 1978). The integration calculation in the ILLSQ and NLSQ methods was implemented using the trapezoidal rule. In the DLLSQ method, the parameter $k$ was chosen as 5, and an odd number for $m$ was empirically determined by a time window of 20 seconds divided by the temporal resolution. A simulation study was conducted for the following two conditions: 1) $\Delta t=1s$ with CNRs ranging from 5 to 100; and 2) a range of $\Delta t$ from 0.1 to 20s with a clinically relevant noise level of CNR = 10 (Cao et al., 2010). At each condition ($\Delta t$, CNR), the simulation was performed with 5000 runs. For each parameter ($K^{\text{trans}}$, $k_0$, and $v_p$), the calculated average values and the standard deviations over 5000 runs from all three methods were recorded. The accuracy was evaluated by the percent difference defined as $\bar{X}/X_0 \times 100\%$, where $X_0$ represents the true value and $\bar{X}$ represents the difference between the calculated average value and $X_0$. The percent differences of each calculated PK parameter using all three methods were recorded for accuracy comparison. The calculation time for each method in 5000 runs were recorded for evaluation of the calculation efficiency.

To further study the impact of noise reduction using the KZ filter, we applied the KZ filter to the standard ILLSQ and NLSQ methods and compared these improved methods, denoted as ILLSQ$_i$ and NLSQ$_i$, to the DLLSQ method. The simulation was performed for the same conditions ($\Delta t$, CNR) as mentioned above. The percent difference of each parameter was calculated for accuracy comparison.
4.2.3 2D Simulation

The DLLSQ method was further evaluated by comparing these results against those obtained using the standard ILLSQ and NLSQ methods with varying PK parameter. The voxels with different PK parameter values were selected from previous clinical brain scans at our institution with the following conditions: $K^{trans}$ ranged from 0 to 0.25 min$^{-1}$, $k_{ep}$ ranged from 0 to 3.00 min$^{-1}$, and $v_p$ ranged from 0 to 0.20 ml/ml. The simulation was performed with 1000 runs for a $\Delta t$=1s and a CNR = 10. In each run, $C_p(t)$ and $C_r(t)$ were generated in the way as described in the single voxel simulation section. For illustration purposes, both the true and calculated parameter values were presented in 2D maps. The accuracy of each method was evaluated using percent difference map, cross-correlation (CC) and the TRE given by Eq. (22).

4.2.4 In vivo study

To demonstrate the feasibility of using the DLLSQ method for potential clinical applications, 5 sets of brain DCE-MRI scans from 4 patients in an IRB-approved clinical study were analyzed. Each set of scan included 60 post-injection 3D volumes, and the temporal resolution during the image acquisition was about 5.25 seconds. A particular region-of-interest (ROI) containing a lesion was contoured prior to the calculation. The individual’s AIF information acquired in the clinical study was used, and the parameters $K^{trans}$, $k_{ep}$, and $v_p$ were calculated in a voxel-by-voxel pattern for a chosen 2D slice. The resulting parameter maps were compared to the calculation results using standard
ILLSQ and NLSQ methods for the same slice. For visual comparison, the parameter distributions within the ROI were illustrated in colored maps superimposed on the pre-injection T1-weighted MR images. For quantitative comparison, each parameter’s median values of all ROI voxels calculated by each method were recorded. The total calculation time for all voxels in the whole slice analysis using each method was recorded for the evaluation of calculation efficiency.

All calculations were carried out in the MATLAB environment (R2012b, MathWorks Inc., Natick, MA) on a workstation with 16GB of RAM and a 3.4 GHz clock rate. The computation time was measured with the functions ‘tic’ and ‘toc’ in MATLAB.

4.3 Results

4.3.1 Single voxel simulation results

Figure 20 shows an example of data processing in the DLLSQ method. As can be seen, the KZ-filtered blue CA curves were much smoother than the red noisy curves. Compared with the black noiseless curves, the positions of the first-pass peak and the curve amplitude in the steady washout phase were preserved. Though the peak enhancement was somewhat compromised after KZ filtering, the accuracy of the least-squares calculation using the derivative curves with two distinct opposite peaks may still be acceptable.
Figure 20: An example CA concentration curve processing of the DLLSQ method in single voxel simulation. (a) The comparison of $C_p(t)$; (b) The comparison of $C_p(t)$ derivative; (c) The comparison of $C_t(t)$; (d) The comparison of $C_t(t)$ derivative. The derivative of noisy $C_p(t)$ and $C_t(t)$ are not shown for the best illustration.

Figure 21 shows the calculation results of $K_{trans}$, $k_{ep}$, and $v_p$ as a function of CNR with $\Delta t=1s$ for the single voxel simulation. In Figure 21, the average values of the calculation results over 5000 runs using all three methods are presented as scatter plots with the error bars representing the standard deviations. The black horizontal lines represent the true values used in $\hat{C}_t(t)$ generation ($K_{trans}=0.0890 \text{ min}^{-1}$, $k_{ep}=0.3201 \text{ min}^{-1}$, and $v_p=0.0230$). The vertical axes were set to the range for the best illustration (i.e. the points out of the figure range were not shown). As shown in Figure 21(a), the DLLSQ method had the best performance for the $K_{trans}$ accuracy, and the percent differences of
the calculated results were all less than 5.0% except for the case with a very low CNR (=5). The accuracy of $K_{\text{trans}}$ using all three methods improved and converged to the same endpoint as the CNR increased. Figure 21(b) suggests that the DLLSQ method provided relatively better estimation of $k_p$ at lower CNR values ($\leq 15$), while the NLSQ and ILLSQ methods were more accurate in estimating $k_p$ at higher CNR values ($\geq 50$). Regarding the $v_p$ estimation in (c), the DLLSQ method results were more accurate than the ILLSQ and NLSQ methods’ results in the investigated CNR range.

![Figure 21: Calculated values of (a) $K_{\text{trans}}$, (b) $k_p$, and (c) $v_p$ as a function of CNR with a temporal resolution of 1s. The horizontal dashed lines represent the known true values: $K_{\text{trans}} = 0.0890$, $k_p = 0.3201$, and $v_p = 0.0230$. Each error bar represents the SD for 5000 simulations](image)

Figure 22 summarizes the simulation results for a range of temporal resolutions $\Delta t$ and a CNR = 10. As shown in Figure 22(a), when $\Delta t < 5$s, the $K_{\text{trans}}$ results using the DLLSQ method were more accurate than those generated by the other two methods with a percent difference <5.0%. As $\Delta t$ was reduced from 5s to 20s, the $K_{\text{trans}}$ calculation accuracy of the DLLSQ method was better or comparable to the other two methods. Results in (b) show that the DLLSQ method estimated the $k_p$ value more accurately than
the other two methods when $\Delta t < 5s$, though the results were mixed when $\Delta t > 5s$. Data in (c) demonstrate that the DLLSQ method was superior to the other two methods in calculating $v_p$ when $\Delta t < 10s$; however, when $\Delta t \geq 10s$, the DLLSQ $v_p$ results had larger errors.

Figure 22: Calculated values of (a) $K^{\text{trans}}$, (b) $k_{ep}$, and (c) $v_p$ as a function of temporal resolution at CNR = 10. The horizontal dashed lines represent the known true values: $K^{\text{trans}} = 0.0890$, $k_{ep} = 0.3201$, and $v_p = 0.0230$. Each error bar represents the SD for 5000 simulations.

Table 4 lists the elapsed time in seconds for 5000 runs in the single voxel simulation at CNR = 10 with a range of temporal resolutions. Generally, the table underlines the prominent high efficiency of the DLLSQ method at higher temporal resolutions. At the desired temporal resolution $\Delta t = 1s$, the DLLSQ method executed faster than the ILLSQ and NLSQ method by a factor of 7.3 and 458, respectively. In addition, for the clinically feasible temporal resolution $\Delta t = 5s$, the efficiency improvement factors for the DLLSQ method were 1.7 and 38 when comparing to the ILLSQ method and the NLSQ method, respectively. The efficiency improvement of the
DLLSQ method decreased when the temporal resolution was degraded, as less data points were available for the calculation given a fixed total sampling time.

**Table 4: Total calculation times for 5000 runs of all methods at different temporal resolutions with CNR = 10. Calculation times are in unit of seconds.**

<table>
<thead>
<tr>
<th>Δt(s)</th>
<th>0.1</th>
<th>0.5</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>10</th>
<th>15</th>
<th>20</th>
</tr>
</thead>
<tbody>
<tr>
<td>DLLSQ</td>
<td>15.46</td>
<td>2.21</td>
<td>1.28</td>
<td>0.98</td>
<td>0.92</td>
<td>0.87</td>
<td>0.84</td>
<td>0.77</td>
<td>0.76</td>
<td>0.76</td>
</tr>
<tr>
<td>ILSQ</td>
<td>7.6x10^2</td>
<td>32.90</td>
<td>9.31</td>
<td>3.21</td>
<td>2.06</td>
<td>1.63</td>
<td>1.41</td>
<td>1.09</td>
<td>1.02</td>
<td>1.00</td>
</tr>
<tr>
<td>NLSQ</td>
<td>2.86 x10^4</td>
<td>1.89 x10^3</td>
<td>5.86x10^3</td>
<td>1.52x10^3</td>
<td>82.53</td>
<td>52.17</td>
<td>31.96</td>
<td>13.04</td>
<td>13.86</td>
<td>12.40</td>
</tr>
</tbody>
</table>

Figure 23 shows the simulation results when the KZ filter was applied to the standard ILSQ and NLSQ methods (ILLSQi and NLSQi methods). With low CNR levels (<10) and small Δt values (≤1s), the ILSQi and NLSQi methods yielded more accurate results than the standard methods with improved precision. Of the three investigated parameters, the calculation accuracy for $v_p$ improved the most using the ILSQi and NLSQi methods. However, the results generated by the DLLSQ method were still more accurate than those generated by the ILSQi and NLSQi methods with the clinically relevant CNR = 10 and the desired Δt = 1s. In contrast, the comparison of DLLSQ method against the ILSQi and NLSQi methods had mixed results at low temporal resolutions (Δt >5s) and at low noise level (CNR > 10).
Figure 23: Calculated results of $K^{\text{trans}}$ (left column), $k_{ep}$ (middle column), and $v_p$ (right column) with the improved ILLSQi and NLSQi methods at varying CNR (1st row) and temporal resolutions (2nd row). The horizontal dashed lines represent the known true values: $K^{\text{trans}} = 0.0890$, $k_{ep} = 0.3201$, and $v_p = 0.0230$. Each error bar represents the SD for 5000 simulations.

### 4.3.2 2D simulation results

The 2D simulation results of DLLSQ, ILLSQ, and NLSQ methods are demonstrated in Figure 24 through Figure 26. The difference maps indicate that the DLLSQ method calculated all three parameters more accurately than the current standard methods. The NLSQ method tended to overestimate the rate constants $K^{\text{trans}}$ and $k_{ep}$ in the high intensity region. In contrast, both ILLSQ and NLSQ methods overestimated low intensity $v_p$ as shown in Figure 26. The 2D simulation results of the ILLSQi and NLSQi methods can be found in Appendix A as Figures A1 to A3. Generally,
the derived parametric maps from the ILLSQ and NLSQ methods were very similar to the ILLSQ and NLSQ methods, respectively.

The quantitative comparison of DLLSQ, ILLSQ, and NLSQ methods is summarized in Table 5. In this table, the TRE values using the DLLSQ method were smaller than the TRE values obtained using the ILLSQ and NLSQ methods. This suggests that the DLLSQ method was more accurate than the other two standard methods in calculating all three parameters. The CC values were all close to 1, indicating that the parameter maps generated by all three methods were all morphologically similar to the true value maps. In the $K_{\text{trans}}$ simulation, the total elapsed time for 1000 simulation runs was 26.82s for the DLLSQ method, 200.58s for the ILLSQ method, and $1.28\times10^4$s for the NLSQ method, respectively. These results suggested that the DLLSQ method can more accurately and more efficiently calculate the brain PK parameters in comparison with the current methods.
Figure 24: 2D simulation results of $K_{\text{trans}}$. (a) True values; (b) DLLSQ results; (c) ILLSQ results; (d) NLSQ results; (e) difference map of the DLLSQ results; (f) difference map of the ILLSQ results; (g) difference map of the NLSQ results.

Figure 25: 2D simulation results of $k_{cp}$. (a) True values; (b) DLLSQ results; (c) ILLSQ results; (d) NLSQ results; (e) difference map of the DLLSQ results; (f) difference map of the ILLSQ results; (g) difference map of the NLSQ results.
Figure 26: 2D simulation results of $v_p$. (a) True values; (b) DLLSQ results; (c) ILLSQ results; (d) NLSQ results; (e) difference map of the DLLSQ results; (f) difference map of the ILLSQ results; (g) difference map of the NLSQ results.

Table 5: Total Relative Error (TRE) and Cross-Correlation (CC) values of the 2D simulation results.

<table>
<thead>
<tr>
<th></th>
<th>$K_{\text{trans}}$</th>
<th>$k_p$</th>
<th>$v_p$</th>
<th>$K_{\text{trans}}$</th>
<th>$k_p$</th>
<th>$v_p$</th>
</tr>
</thead>
<tbody>
<tr>
<td>DLLSQ</td>
<td>0.031</td>
<td>0.990</td>
<td>0.082</td>
<td>0.994</td>
<td>0.058</td>
<td>0.999</td>
</tr>
<tr>
<td>ILLSQ</td>
<td>0.059</td>
<td>0.990</td>
<td>0.091</td>
<td>0.993</td>
<td>0.117</td>
<td>0.988</td>
</tr>
<tr>
<td>NLSQ</td>
<td>0.091</td>
<td>0.953</td>
<td>0.123</td>
<td>0.966</td>
<td>0.125</td>
<td>0.975</td>
</tr>
</tbody>
</table>

4.3.3 *In vivo* study results

Figure 27 presents the calculation results for a selected patient’s data using the DLLSQ (left column), ILLSQ (middle column), and NLSQ (right column) methods, respectively. A large region with elevated parameter intensities was identified by all three methods near the center of the ROI, though differences were noted for the
peripheral region in the ROI as the intensities generated by the different methods varied visually. Table 6 records each parameter’s median value within the ROI. The recorded values from the investigated methods were comparable, though some noticeable differences could be observed. The average calculation time required to analyze the whole slice was 1.59s for the DLLSQ method, 2.31s for the ILLSQ method, and 55.07s for the NLSQ method, respectively. These results suggest that the DLLSQ method could potentially be used for the current clinical DCE-MRI analysis.

**Table 6: The median values of the investigated parameters in the in vivo study.**

* Images are shown as Figure 27; † 2nd scan of this patient after radiotherapy treatment

<table>
<thead>
<tr>
<th>Patient No.</th>
<th>Type of Disease</th>
<th>1*</th>
<th>2</th>
<th>3</th>
<th>3†</th>
<th>4</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Lt Parietal Glioblastoma</td>
<td>Lt Frontal/Parietal Glioblastoma</td>
<td>Rt Parietal Glioblastoma</td>
<td>Rt Parietal Glioblastoma</td>
<td>Lt Frontal Oligodendroglioma</td>
</tr>
<tr>
<td>Median (K_{\text{trans}}) (min⁻¹)</td>
<td>DLLSQ</td>
<td>0.083</td>
<td>0.114</td>
<td>0.027</td>
<td>0.018</td>
<td>0.066</td>
</tr>
<tr>
<td></td>
<td>ILLSQ</td>
<td>0.087</td>
<td>0.128</td>
<td>0.025</td>
<td>0.017</td>
<td>0.072</td>
</tr>
<tr>
<td></td>
<td>NLSQ</td>
<td>0.092</td>
<td>0.161</td>
<td>0.024</td>
<td>0.017</td>
<td>0.076</td>
</tr>
<tr>
<td>Median (k_{\text{vle}}) (min⁻¹)</td>
<td>DLLSQ</td>
<td>0.436</td>
<td>0.491</td>
<td>0.519</td>
<td>0.500</td>
<td>0.321</td>
</tr>
<tr>
<td></td>
<td>ILLSQ</td>
<td>0.401</td>
<td>0.553</td>
<td>0.506</td>
<td>0.553</td>
<td>0.339</td>
</tr>
<tr>
<td></td>
<td>NLSQ</td>
<td>0.669</td>
<td>0.954</td>
<td>0.553</td>
<td>0.449</td>
<td>0.382</td>
</tr>
<tr>
<td>Median (v_p) (10⁻² ml/ml)</td>
<td>DLLSQ</td>
<td>4.021</td>
<td>0.926</td>
<td>0.914</td>
<td>0.796</td>
<td>0.557</td>
</tr>
<tr>
<td></td>
<td>ILLSQ</td>
<td>2.871</td>
<td>0.961</td>
<td>1.043</td>
<td>0.988</td>
<td>0.523</td>
</tr>
<tr>
<td></td>
<td>NLSQ</td>
<td>2.593</td>
<td>1.143</td>
<td>0.988</td>
<td>0.867</td>
<td>0.499</td>
</tr>
</tbody>
</table>
Figure 27: *In vivo* study results of a selected patient. 1st row: $K_{\text{trans}}$ results of the DLLSQ (a), the ILLSQ (b), and the NLSQ method (c); 2nd row: $k_{ep}$ results of the DLLSQ (d), the ILLSQ (e), and the NLSQ method (f); 3rd row: $v_p$ results of the DLLSQ (g), the ILLSQ (h), and the NLSQ method (i).

### 4.4 Discussion

In this work, we developed a new efficient DLLSQ method for calculating the PK permeability parameters based on a linear least-squares fitting of the derivative expression of the PK model. The KZ filter adopted in the DLLSQ method is important in
effective and efficient noise reduction. As pointed out by the detailed mathematical
deduction, the effective low-pass cutoff frequency of the KZ filter is approximately

\[ 2\sqrt{6} \frac{1-\alpha^{1/2k}}{m^2 - \alpha^{1/2k}}, \]

where \( \alpha \) is a pre-selected value between 0 and 1 (Yang and Zurbenko, 2010). In addition, at the edge of the impulse response function, the zero derivatives
make the KZ filter a sharply declining function and provide high frequency resolutions
(Rao et al., 1997). Meanwhile, the implementation of KZ filtering in the time domain is
fast because convolution operations are used. Because of these features, the KZ filtering
process facilitated the accurate calculation at very low computational cost.

When CNR was higher than 20 as shown in Figure 21(b), the calculation results
for \( k_{ep} \) was over-estimated by 6.2% using the DLLSQ method and was less than 1% using
the ILLSQ and NLSQ methods. This observation indicated that the performance of KZ
filter may depend on the choice of \( k \) and \( m \). In this study, the kernel width related to
parameter \( m \) was determined empirically, and the parameter \( k \) was chosen as 5 based on
the information from a previous qualitative DCE-MRI study using polynomial filtering
(Jesberger et al., 2006). Results in Single Voxel Simulation indicated that the selection of
\((k,m)\) in this study improved the calculation accuracy at low CNR levels. However, at
high CNR levels, the selected parameters \((k,m)\) for KZ filtering process may not be
optimal leading to a systematic error in \( k_{ep} \) calculation. To verify this assumption, we
calculated the \( k_{ep} \) using DLLSQ method at CNR = 100 and \( \Delta t =1s \) with different \((k,m)\)
selections. As can be seen in Sup.Table S4, the calculation accuracy was improved when
a shorter time window width and a smaller $k$ value were used. Thus, to ensure the optimal calculation accuracy, relatively smaller $(k,m)$ values should be selected in the KZ filtering for high CNR levels. Future investigations might be required in searching for optimal parameters $(k,m)$ for the best estimation of all PK parameters at different CNR levels and temporal resolutions.

The dependency of DLLSQ method on temporal resolution was studied at a clinically relevant level of CNR = 10. As the temporal resolution decreased from 1s or less to 5s, the $K^{\text{trans}}$ calculation bias using the DLLSQ method increased from about 1.1% to nearly 5.7%. This pattern fits the common expectation that the derivative calculation becomes less accurate as $\Delta t$ increases. When temporal resolution ranged from 5s to 20s, the accuracy of the analysis was also affected by the first-pass peak sampling. As indicated in the source of AIF model used in the single voxel simulation (Parker et al., 2006), the first-pass peak lasted about 20 seconds after the CA injection, and the FWHM of the peak was about 10 seconds. If the temporal resolution is comparable or larger than the peak width, the quick wash-in process in $C_p(t)$ and $C_r(t)$ may be missed if the two sampling points happen to be at two distal ends of the first-pass peak; in this case, the PK parameters may not accurately be calculated (Kershaw and Cheng, 2010). Since high temporal resolutions are of our particular interests for the potential application of the DLLSQ method, the extended study of its performance with low sampling rate is not considered.
In the 2D simulation work, we demonstrated the accuracy of the DLLSQ method in ranges of clinically observed parameter values. The signed difference maps in Figure 24 to Figure 26 suggest that the accuracy of DLLSQ method varies at different parameter intensity levels. Some previous studies have revealed that rate constants $K_{\text{trans}}$ and $k_p$ at different intensity levels may require different temporal resolutions to ensure accurate estimation (Lopata et al., 2007; Luypaert et al., 2010). In this work, the true PK permeability parameters were generated in brain DCE-MRI analysis. Specifically, the used $K_{\text{trans}}$ values were small and $v_p$ values are not ignorable. When the interested parameter intensity goes beyond the investigated range, such as the relatively high $K_{\text{trans}}$ values in human breast, the accuracy of the DLLSQ method needs to be determined. For further investigation, the scrutiny of the DLLSQ method with a large number of PK parameters combinations in a wider intensity-varying range might be important for the potential clinical application at more other clinical sites.

4.5 Conclusion

In this work, an efficient DLLSQ method was developed to calculate the PK parameters for high temporal resolution brain DCE-MRI. Based on the derivative expression of the PK model, the presented DLLSQ method was proved to be more efficient with comparable or improved accuracy in comparison with currently existing PK analysis methods. With the desired high temporal resolution, the DLLSQ method could be used for clinical analysis in brain DCE-MRI studies.
5. Application of classic fractal dimension analysis in DCE-MRI treatment response assessment

5.1 Introduction

Currently, the treatment response of a tumor using DCE-MRI is often evaluated by comparing the morphological descriptors of the tumor volume and first order statistics (mean/median/variance) of the PK parameters from the classic PK analysis before and after treatment. Such comparisons are conducted in the manually selected regions of interest (ROIs) or over the entire tumor volume (O'Connor et al., 2011; Jaffe, 2006). As discussion in section 1.3.5, the current techniques have two limitations. Firstly, the PK analysis accuracy could be questionable under clinical protocols, and the observed therapeutic effects may not be reliable. Secondly, the currently techniques cannot reveal functional heterogeneity within the ROI. Since certain regions could be more sensitive to the treatment effects, heterogeneity evaluation could provide potentially useful biomarkers (Alić et al., 2006). As a result, development of model-free DCE-MRI analysis methods with the heterogeneity evaluation capability is favored.

Recently, digital image texture analysis has gained its popularity in medical image analysis for spatial heterogeneity analysis. As a general concept, texture in a biomedical image can be described as the appearance and the structure of pixel arrangements (Castellano et al., 2004). Texture analysis focuses on the detection and quantification of repeating patterns and nonrandom distributions of pixel intensity values throughout an ROI on a digital image (Gatenby et al., 2013). The spatial
heterogeneity of an ROI is usually quantified by many numerical texture features through a certain feature space, which is created by a specific texture analysis technique. In the recently developed concept of radiomics which refers to the automated extraction of image features with great throughput potentials, heterogeneity assessment using texture analysis has been incorporated as a key component for comprehensive data mining and characterization (Kumar et al., 2012; Lambin et al., 2012; Aerts et al., 2014).

Following the growing interest in radiomics, texture analysis has been applied in preclinical and clinical MR studies for computer-aided diagnosis (CADx) and treatment outcome prediction. The most frequently used texture analysis technique is the gray level texture matrices. As for describing local relationships between voxels’ gray level intensities, texture matrices contain the statistics of gray level intensity variation for a 2D/3D object. Many texture features based on the intensity distribution of texture matrices have been proposed for spatial heterogeneity quantification (Chicklore et al., 2013). Two texture matrices have been widely adopted: gray level co-occurrence matrix (GLCOM) and gray level run-length matrix (GLRLM) (Haralick et al., 1973; Galloway, 1975). The GLCOM describes the information about the gray level distribution of pairs of pixels with specified separation and direction. As a simple example, Figure 28 demonstrates the generation of GLCOM on horizontal direction of a 4x4 image. Each element \( P(i, j) \) of the horizontal GLCOM is the incidence of pairs of pixels on image \( I \) that happen to have \( I(r,c) = i \) and \( I(r,c±1) = j \) where \( r \) and \( c \) indicate the pixel position on \( I \).
The size of the GLCOM is $N \times N$ where $N$ is the number of discrete gray levels in $I$. For a 2D object, GLCOM can be generated at 4 directions ($0^\circ$ (horizontal), $45^\circ$, $90^\circ$ and $135^\circ$) with 1-pixel separation; for a 3D object, GLCOM can be generated at 13 direction. Texture features can be extracted from the GLCOM at each direction or the averaged GLCOM of 4 or 13 directions. Texture features defined originally based on GLCOM are also called as Haralick features (Haralick et al., 1973). Appendix B summarizes the 22 Haralick features with detailed mathematical definition.

![Image of GLCOM generation](image)

**Figure 28:** An example of the GLCOM generation on the horizontal direction. Red and blue color shows the derivation of $P(1,2)$ and $P(3,4)$ respectively.

Compared to GLCOM, GLRLM describes the runs of pixels with the same gray level at a specific direction. Figure 29 shows a simple example of horizontal GLRLM generation as following the pattern of Figure 28. Each element $P(i, j)$ of the horizontal GLRLM is the incidence of runs of pixels on image $I$ that happen to have a length of $j$ pixels at gray level $i$. The size of the GLRLM is $N \times M$ where $N$ is the number of discrete gray levels in $I$ and $M$ is the maximum possible run length in $I$. Similarly, GLRLM can be generated at 4 (2D) or 13 (3D) directions like GLCOM. 11 reported texture features from GLRLM can be found in Appendix B.
Figure 29: An example of the GLRLM generation on the horizontal direction. Red and blue color shows the derivation of $P(3,1)$ and $P(2,3)$ respectively.

Currently, the texture analysis using GLCOM/GLRLM in DCE-MRI can be applied to post-injection images and/or PK parametric maps. Because of their straightforward definition and rather simple implementation, GLCOM/GLRLM based texture analysis has been incorporated into some clinical studies. So far, texture features from these two matrices have been hypothesized to be superior in the diagnosis of breast cancer (Karahaliou et al., 2010; Holli et al., 2010). For the application of treatment assessment, GLCOM/GLRLM has been adopted in some studies as preliminary works of texture analysis application for treatment response assessment (Alic et al., 2011; Xie et al., 2015).

The derivation aforementioned gray level matrices relies on the gray level intensity distribution. Since the MR signal intensity output is affected by the hardware specification, the intra-scanner reproducibility of these gray level matrices could be a potential issue. As an alternative strategy, fractal dimension analysis evaluates the spatial heterogeneity by estimating the complexity of ROI at multiple scales with no direct dependence on the gray level intensity values. The simple box counting method is
one of the approaches in fractal dimension analysis to estimate the shape complexity of a binary object. Figure 30 shows the example of box counting method. The original binary was divided using boxes of decreasing scale \( s \). The number of boxes occupied by the object at scale \( s \) is recorded as \( n \). The box counting method records the fitted slope of \( \log(n) - \log(s) \) curve, which is also called as box counting dimension. For a perfect fractal object like the Apollonian Gasket in this example, the curve should be a perfect straight line.

![Figure 30: An example of box counting method for Apollonian Gasket analysis. (a) The binary object is successively divided using boxes of decreasing grid size \( s \). (b) The relationship between \( \log s \) and \( \log n \) where \( n \) is the number of boxes occupied by the object. For a perfect fractal object, the curve should be a straight line. Box counting dimension is computed as the slope of the line (approximately 1.4 in this example).](image)

In spite of its intuitive nature, the box counting method ignores the intensity variation within the object and requires a hard threshold for binary operation. As a more general method, Rényi dimensions are a generalization of fractal dimensions as a family.
of information measurement (Peitgen et al., 2006). It is expressed by summing the Rényi entropy at different scales:

\[
d_q = \lim_{s \to 0} \frac{\log \sum_i p_i^q}{(1 - q) \log \frac{1}{s}}
\]

\[
d_1 = \lim_{q \to 1} d_q
\]

(30)

where \( s \) is the scale resolution normalized to unity (For an undivided original object, \( s = 1 \)) and \( p_i \) is the normalized intensity value of \( i \)-th voxel of the image such that \( \sum_i p_i = 1 \ \forall i \). When substituting \( q = 0 \) in Eq. (30), the zero order Rényi dimension \( d_0 \) is equivalent to the simple box counting dimension. Two parameters are often used: \( d_1 \) information dimension and \( d_2 \) correlation dimension, and \( d_1 \geq d_2 \).

Previous studies reported that higher \( d_1 \) and \( d_2 \) values of DCE-MRI parametric maps were correlated with more heterogeneity gliomas, and the simulated treatment resulted in an increase of \( d_1 \) and \( d_2 \) values (Rose et al., 2009). Nevertheless, the thorough \textit{in vivo} investigation of fractal dimension analysis in treatment response assessment is limited. In this work, the applicability of PK parametric maps fractal dimensions \( d_1 \) and \( d_2 \) \textit{in vivo} treatment response assessment was examined in an experiment of small animal anti-angiogenesis drug response assessment with treatment/control group knowledge.
5.2 Materials and methods

5.2.1 Small animal experimental protocol

The small animal study was conducted at Duke Center In Vivo Microscopy (CIVM) with the approval by the Institutional Animal Care and Use Committee. The diagram of the experiment is briefly summarized in Figure 31. Sixteen Female \textit{nu/nu} mice with LS-174T (Charles River Laboratories, Wilmington, MA) implanted in the mammary fat pad were followed for four weeks. When tumor volume was approximately 100\(\mu\)L, the mice were randomly assigned into the treatment (\(n = 8\)) group or the control (\(n = 8\)) group. A baseline DCE-MRI scan was acquired at Day0. At Day1, the treatment/control group received bevacizumab (Avastin\textsuperscript{®}, Genentech, South San Francisco, CA) or normal saline via an intraperitoneal injection at a dose of 5 mg/kg or 5 mL/kg, respectively. Thereafter, bevacizumab/saline was administered at the same respective dose twice weekly, and DCE-MRI imaging scans were performed weekly. The first two post-treatment scans at Day2 and Day9 as well as the baseline scan were evaluated since the therapeutic effects of the treatment group were prominent after the first two weeks’ treatments.
5.2.2 Imaging protocol

All DCE-MRI scans were acquired in a 7T small animal MRI scanner (Bruker BioSpin MRI GmbH, Ettlingen, Germany) equipped with self-shielded gradient coils with a maximum strength of 450 mT/m and a rise time of 110 μs. An actively detuned volume RF coil (linear transmit, ID = 72 mm) was used in conjunction with a four-element coil (2x2 linear array, 10x10 mm loops) for surface receive. An interleaved ultra-short-echo radial sampling sequence was adopted for 4D reconstruction using a sliding-window keyhole approach (Subashi et al., 2013). The acquisition parameters were: FOV = 20 mm³, matrix = 128³, TR/TE = 5/0.02 ms, NEX = 1, flip angle α = 10°, temporal resolution = 9.9s. The varying flip angle method with α = {2°, 10°} was used to measure the native relaxation rate before the CA injection with Eq. (3). During the scan, animals were positioned in a custom-made MR-craddle and were maintained under anesthesia by isoflurane delivery via a nose cone. The body temperature was controlled between 36°C and 37°C by circulating warm water. Breathing was monitored through a pneumatic pillow and was maintained at a rate of 50-60 breaths/min via adjusting isoflurane delivery. An automatic syringe pump (KD Scientific Inc., Holliston, MA) was used to
administer Gd-DTPA (Magnevist, Schering AG, Berlin, Germany) as a bolus via a 27-gauge vein at a dose of 0.5 mmol/kg and a flow rate of 2.4mL/min. Dynamic imaging was initiated two minutes prior to the CA injection and lasted for approximately 20 minutes after the CA injection.

For each scan, the tumor volume $V$ was recorded as the primary indicator of therapeutic effect (Jaffe, 2006). In the PK analysis, the CA concentration maps were calculated based on Eq. (1), and the extended Tofts model (Eq. (8)) was fitted. The AIF information was approximated by a reported small animal population measurement result (Loveless et al., 2012). The tumor’s mean $K^{\text{trans}}$ was reported as the primary PK biomarker in this study. The $K^{\text{trans}}$ coefficient of variation (CV), which is defined as the ratio of standard deviation to the mean, was also reported as a measure of $K^{\text{trans}}$ probability distribution dispersion. Rényi dimensions $d_1$ and $d_2$ were recorded as $K^{\text{trans}}$ heterogeneity measurement. In addition, $K^{\text{trans}}$ kurtosis (measurement of ‘peakedness’ of the probability distribution) and skewness (measurement of asymmetry of probability distribution) were recorded to describe $K^{\text{trans}}$ distribution shape features for comparison study.

5.2.3 Statistical analysis

To evaluate the therapeutic effect, for each post-treatment scan, the Mann-Whitney U-test was used to assess the difference of the recorded quantitative values between treatment and control groups. Significance was determined based on a $p$-level
less than 0.05 with multi-comparison correction if applicable (Chen et al., 2007).

Experiments using support vector machine (SVM) in a leave-one-out approach were performed to validate the potential use of the recorded metrics in treatment/control group classification. 2 out of total 48 scans were excluded for the PK analysis due to the uncertain CA injection dose during the DCE-MRI scan.

5.3 Key results

Figure 32 demonstrates the tumor volume $V$ changes during the study. At Day0, the initial tumor volumes of the treatment and control group were $115 \pm 41 \, \mu L$ and $91 \pm 37 \, \mu L$, respectively. As the treatment course continue, the tumor in the treatment group grew slower than the tumors in the control group. The effect of the bevacizumab was more obvious on the comparison of the relative value of $V$, where the relative value was defined as the ratio of post-treatment value to pre-treatment value. After three administrations, the relative $V$ in the control group was significantly higher than that in the treatment group ($p = 0.002$) at Day9.
Figure 32: The comparisons of tumor volume V and its relative value between treatment/control groups during the treatment course. * indicates that statistically significant difference was found between treatment/control groups.

The $K^{\text{trans}}$ analysis results are summarized in Figure 33. The tumor mean $K^{\text{trans}}$ evolution during the treatment course is presented in (a). At Day0, there was no significant difference between the treatment/control groups $K^{\text{trans}}$ value. After three treatment deliveries at Day9, the treatment group had a significantly smaller $K^{\text{trans}}$ value than the control group ($p = 0.021$). The relative $K^{\text{trans}}$ of the two cohorts at both Day2 and Day9 had no significant difference in (b). Similarly, the $K^{\text{trans}}$ CV, kurtosis and skewness showed no significant difference between the two cohorts at both post-treatment days in (c)-(e). In contrast, the fractal dimensions $d_1$ ($p = 0.013$) and $d_2$ ($p = 0.028$) of $K^{\text{trans}}$ map reflected significant difference between the treatment/control group at Day 9 in (f) and (g). These results demonstrated the great potential of using PK parametric maps fractal dimensions for capturing therapeutic response effect.
Figure 33: The longitudinal change of mean $K_{\text{trans}}$ (a), relative $K_{\text{trans}}$ (b), $K_{\text{trans}}$ CV (c), $K_{\text{trans}}$ kurtosis (d), $K_{\text{trans}}$ skewness (e), $K_{\text{trans}}$ $d_1$ (f), and $K_{\text{trans}}$ $d_2$ (g). * indicates that statistically significant difference was found between treatment/control groups.

The results of treatment/control groups’ classification experiments using SVM are summarized in Table 7. When mean $K_{\text{trans}}$ was used as the sole input elements, the classification accuracy at Day9 was 68.8%, and if $K_{\text{trans}}$ CV was added to the input, the accuracy was improved to 75.0% at Day9. In contrast, if classic fractal dimensions $d_1$ and $d_2$ were selected as the input elements, the achieved accuracies were promising at both
Day2 (87.5%) and Day9 (100%). These results show the superiority of fractal dimensions over the current $K_{\text{trans}}$ statistics in therapeutic response assessment.

### Table 7: The results of treatment/control groups’ classification using SVM

<table>
<thead>
<tr>
<th>SVM Input</th>
<th>Classification Accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Day2</td>
</tr>
<tr>
<td>Mean $K_{\text{trans}}$</td>
<td>37.5%</td>
</tr>
<tr>
<td>Rev. Mean $K_{\text{trans}}$</td>
<td>42.9%</td>
</tr>
<tr>
<td>(Mean $K_{\text{trans}}$, CV)</td>
<td>43.8%</td>
</tr>
<tr>
<td>(Kurtosis, Skewness)</td>
<td>50.0%</td>
</tr>
<tr>
<td>$(d_1, d_2)$</td>
<td>87.5%</td>
</tr>
</tbody>
</table>

### 5.4 Conclusion

In this work, the great potential of classic fractal dimensions of DCE-MRI PK parametric maps for in vivo therapeutic response assessment was successfully demonstrated. In a longitudinal small animal experiment using high spatiotemporal DCE-MRI protocol, the investigated fractal dimensions could reflect significant difference between the treatment and the control groups. Compared to the current PK parameter statistics, the classic fractal dimensions were more reliable for treatment/control classification. The application of classic fractal dimensions in clinical DCE-MRI treatment assessment can be seen as promising after more human in vivo verification studies.
6. Development of a novel Gray Level Local Power Matrix (GLLPM) for DCE-MRI texture analysis

6.1 Introduction

As mentioned in section 5.1, GLCOM/GLRLM texture matrices and the relevant texture features are commonly used in medical imaging analysis for straightforward definitions and easy implementations. The current texture matrices can be improved in two aspects. Firstly, the calculation efficiency can be enhanced. Most studies involving GLCOM and/or GLRLM extracted texture features from the averaged matrices at 4 (or 13) directions. Since averaged texture matrices discard the directionality information, the derivation of matrices at each separated direction becomes inefficient in terms of calculation time. It is important to point out that texture features can be derived from the matrices at each direction separately, and yet the rationality of feature extraction at each direction separately has not been justified. Specifically, different features may demonstrate significant changes at different directions in response to treatment, and the determination of a ‘major’ direction for analysis is far from straightforward (Zheng et al., 2009).

Secondly, most reported studies using texture analysis for treatment assessment focused primarily on the evaluation of a chosen contrast-enhanced image volume or PK parametric maps, while the dynamic information of tumor heterogeneity evolution during the CA uptake was not well studied. The introduction of temporal information could further utilize the 4D nature of DCE-MRI and may yield additional factors for
treatment assessment. Although 4D GLCOM has been tentatively proposed for breast lesion segmentation using DCE-MRI, the high computation cost of the reported technique limited its application in therapeutic response assessment (Woods et al., 2007a; Woods et al., 2007b).

To account for these issues, a novel texture analysis matrix, called gray level local power matrix (GLLPM), was designed. Following the spirit of GLCOM technique, the GLLPM method depicts the relationship between voxels’ intensities and voxel neighborhood intensity variation, and no direction specification is required. Additionally, the information on the temporal resolution is incorporated. Based on GLLPM, the dynamic curves of Haralick texture features are generated following the similar fashion in GLCOM techniques, and the dynamics characteristics of the texture feature curves are investigated for potential therapeutic response assessment application.

6.2 Theory of GLLPM

For each voxel \((x_0, y_0, z_0, t_0)\), in a 4D image series \(I\), its neighborhood is defined in a 4D fashion as all connective voxels within a predefined cubic area with a side length of \(k\) voxels:

\[
I_c = I(x_0 \pm k, y_0 \pm k, z_0 \pm k, \{t_0, t_0 + k \cdot \Delta t\})
\]  

(31)

where \(\Delta t\) is the temporal resolution. As shown above, at each time point, the 3D volume acquired within in next \(k\) time points are included for possible temporal
connection. The default selection of $k$ is 1 as in most GLCOM studies. Then, the local power (LP) of each voxel $(x_0, y_0, z_0, t_0)$ is defined as the intensity variance of its corresponding neighborhood:

$$LP(x_0, y_0, z_0, t_0) = \text{var}(I_c), I_c = I(x_0 \pm 1, y_0 \pm 1, z_0 \pm 1, \{t_0, t_0 + \Delta t\})$$

(32)

The generated LP is a 4D volume in the same size is image $I$. Then the LP volume is normalized to its 4D maximum value with $N$ gray levels, where $N$ is the number of discrete gray level of the 4D volume $I$. Each element of the GLLPM, $P(k,l,t)$, is defined as the number of voxels in $I$ that happens to have an intensity value as $k$ and normalized LP value as $l$:

$$P(k,l,t) = \#(x,y,z) \in I(t) | I(x,y,z,t) = k, LP(x,y,z,t) = l,$$

$$k = 1, ..., N, \quad l = 1, ..., N$$

(33)

The generated GLLPM, $P$, is a 3D matrix whose 3rd dimension is time evolution. At each time point $t$, Haralick features can be generated from the 2D GLLPM slice in the same way as GLCOM. By repeating the process over time, the dynamic curve of each Haralick feature can be generated. Since the LP calculation includes all surrounding voxels, no directionality need to be addressed and thus calculation time can be reduced in comparison with GLCOM calculation.

6.3 Retrospective small animal study
The feasibility of using GLLPM based Haralick texture features for therapeutic response assessment was retrospectively examined in the small animal experiment in Section 5.2.1. The Day0 scans and Day9 scans were analyzed. For each scan, the CA concentration maps in the first 10-minute time window were adopted. Prior to the GLLPM generation, the CA concentration maps of each scan were normalized to 64 gray levels. 22 Haralick texture feature curves $F_n, \text{GLLPM}(t)$, $n=1,2,...,22$ were generated. For comparison study, 4D GLCOM and 3D GLCOM series were included. Specifically, the 4D GLCOM matrix was generated using the published workflow (Woods et al., 2007b); on the other hand, each 3D CA concentration volume at a time point was used to generated a 2D GLCOM matrix, and all generated 2D matrices were combined to a 3D matrix whose 3rd dimension also represents time convolution. Both 4D GLCOM and 3D GLCOM were generated in the same time window, and they were used for the generation of texture feature curves set $F_n, \text{4D GLCOM}(t)$, and $F_n, \text{3D GLCOM}(t)$, respectively. The calculation time of GLLPM and 3D/4D GLCOM were recorded using ‘tic’ and ‘toc’ function in MATLAB for efficiency evaluation.

For both pre- and post-treatment scans, the Area Under Curve (AUC) values of all dynamic texture feature curves using GLLPM and 3D/4D GLCOM were recorded. The Mann-Whitney U-test was used to assess the differences of the all AUC values between treatment/control groups. Statistical significance was determined as $\alpha < 0.05$ with multiple comparison correction. The post-treatment dynamic texture feature curves
were fitted with cubic polynomial to describe each feature’s dynamics. For those curves with good fitting ($R^2 > 0.8$), the fitted polynomial coefficients were used for treatment/control group classification using SVM in a leave-one-out approach, and the accuracy of each classification test were recorded.

6.4 Key results

Figure 34 shows the tumor mean CA concentration curve of both treatment and control groups at Day0 can Day9. In the investigated 10-minute post-injection time window, the curves had a steady wash-in at both scan days. Although the curves at Day9 had lower amplitudes, the difference between treatment/control groups were not statistically significant.

![Figure 34: The comparison of CA concentration curves of treatment/control groups at Day0 (left) and at Day 9 (right)](image)

Figure 35 shows the comparison of auto correlation (AC, defined by Eq. (B1) in Appendix B) curves from GLLPM (left column), 4D GLCOM (middle column) and 3D GLCOM (right column) at Day0 (top row) and Day9 (bottom row). As can be seen, at
Day9 after the treatment, the separation of treatment group and control group curves was obvious on all three figures at the bottom row, and the curve AUC differences between treatment/control groups were statistically significant. It is noticeable that AC curves from GLLPM and 4D GLCOM were generally smoother than the curves AC from 3D GLCOM. This observation could be explained by the benefit of additional temporal information of GLLPM and 4D GLCOM.

![Graphs showing AC curves](image)

**Figure 35:** The comparison of auto correlation (AC) curves of GLLPM (left column), 4D GLCOM (middle column) and 3D GLCOM (right column) at Day0 (top row) and Day9 (bottom row). Red: Treatment group; Blue: Control group

Figure 36 shows the comparison of inverse variance (InvVar, defined by Eq. (B17) in Appendix B) curves in the same organization as Figure 35. At Day9, the separation of treatment/control groups' curves from GLLPM was noticeable, while the
curves from 4D/3D GLCOM failed to show differences between treatment/control groups. The AUC values between treatment/control groups had a significant difference at Day9 when using GLLPM; in contrast, no significant difference between treatment/control groups were found when comparing AUC values of 4D/3D GLCOM curves.

![Figure 36: The comparison of inverse variance (InvVar) curves of GLLPM (left column), 4D GLCOM (middle column) and 3D GLCOM (right column) at Day0 (top row) and Day9 (bottom row). Red: Treatment group; Blue: Control group](image)

Table 8 lists the statistical analysis results including the pre- and post-treatment dynamic texture feature AUC differences and the SVM classification experiment results. Of the 22 Haralick texture feature curves, 21 texture feature curves from GLLPM showed significant differences between treatment/control groups at Day9 post-treatment
scan. The corresponding number for 4D and 3D GLCOM were 18 and 18, respectively. 19 post-treatment dynamic texture feature curves from GLLPM can be fitted by cubic polynomial, and the corresponding number for 4D and 3D GLCOM were 19 and 14, respectively. When using the fitted coefficients for treatment/control group classification, the averaged accuracy of 19 tests was \((84.5 \pm 12.1)\%\), which was higher than the averaged accuracy from 4D GLCOM \((73.3 \pm 12.8)\%\) and from 3D GLCOM \((65.6 \pm 10.5)\%\). In terms of calculation efficiency, the mean GLLPM calculation time of all scans was about 127 seconds, while the mean calculation time for 4D and 3D GLCOM was about 41.7 minutes and 388s, respectively. These results suggest that the dynamics characteristics of GLLPM based texture features are more accurate and more efficient than the characteristics of 4D/3D GLCOM texture features in monitoring the therapeutic effect difference between treatment/control groups.

### Table 8: Statistical analysis results of the derived features (Defined by Eqs (B1) – (B22) in Appendix B).

*: statistically significant N.A.: not available because of bad cubic polynomial fitting results \( (R^2<0.8) \)

<table>
<thead>
<tr>
<th>Feature Name</th>
<th>GLLPM Pre-Tx p</th>
<th>GLLPM Post-Tx p</th>
<th>GLLPM SVM Acc.</th>
<th>4D GLCOM Pre-Tx p</th>
<th>4D GLCOM Post-Tx p</th>
<th>4D GLCOM SVM Acc.</th>
<th>3D GLCOM Pre-Tx p</th>
<th>3D GLCOM Post-Tx p</th>
<th>3D GLCOM SVM Acc.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Auto Correlation</td>
<td>0.065 &lt;0.001*</td>
<td>0.065 &lt;0.001*</td>
<td>93.8%</td>
<td>0.065 &lt;0.001*</td>
<td>0.065 &lt;0.001*</td>
<td>81.3%</td>
<td>0.050 0.001*</td>
<td>0.050 0.001*</td>
<td>56.3%</td>
</tr>
<tr>
<td>Cluster Prominence</td>
<td>0.130 &lt;0.001*</td>
<td>0.130 &lt;0.001*</td>
<td>68.8%</td>
<td>0.130 0.001*</td>
<td>0.130 0.001*</td>
<td>87.5%</td>
<td>0.161 &lt;0.001*</td>
<td>0.161 &lt;0.001*</td>
<td>75.0%</td>
</tr>
<tr>
<td>Cluster Shade</td>
<td>0.083 &lt;0.001*</td>
<td>0.195 &lt;0.001*</td>
<td>75.0%</td>
<td>0.195 0.001*</td>
<td>0.195 0.001*</td>
<td>81.3%</td>
<td>0.161 &lt;0.001*</td>
<td>0.161 &lt;0.001*</td>
<td>68.8%</td>
</tr>
<tr>
<td>Cluster Tendency</td>
<td>0.065 &lt;0.001*</td>
<td>0.161 &lt;0.001*</td>
<td>93.8%</td>
<td>0.161 &lt;0.001*</td>
<td>0.161 &lt;0.001*</td>
<td>81.3%</td>
<td>0.050 &lt;0.001*</td>
<td>0.050 &lt;0.001*</td>
<td>56.3%</td>
</tr>
<tr>
<td>Contrast</td>
<td>0.076 &lt;0.001*</td>
<td>0.076 &lt;0.001*</td>
<td>93.8%</td>
<td>0.076 &lt;0.001*</td>
<td>0.076 &lt;0.001*</td>
<td>75.0%</td>
<td>0.279 0.002*</td>
<td>0.279 0.002*</td>
<td>N.A.</td>
</tr>
<tr>
<td>Correlation</td>
<td>0.161 0.002*</td>
<td>0.161 &lt;0.001*</td>
<td>68.8%</td>
<td>0.161 &lt;0.001*</td>
<td>0.161 &lt;0.001*</td>
<td>81.3%</td>
<td>0.161 &lt;0.001*</td>
<td>0.161 &lt;0.001*</td>
<td>N.A.</td>
</tr>
</tbody>
</table>
### Table 1: Texture Matrix Dynamics

<table>
<thead>
<tr>
<th>Matrix Type</th>
<th>Differential Entropy</th>
<th>Energy</th>
<th>Entropy</th>
<th>Homogeneity 1</th>
<th>Homogeneity 2</th>
<th>IMC 1</th>
<th>IMC 2</th>
<th>IDMN</th>
<th>IDN</th>
<th>InvVar</th>
<th>Max Probability</th>
<th>Sum Ave</th>
<th>Sum Entropy</th>
<th>Sum Variance</th>
<th>Variance</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0.105</td>
<td>0.234</td>
<td>0.328</td>
<td>0.065</td>
<td>0.065</td>
<td>0.083</td>
<td>0.721</td>
<td>0.645</td>
<td>0.959</td>
<td>0.065</td>
<td>0.195</td>
<td>0.065</td>
<td>0.382</td>
<td>0.382</td>
<td>0.065</td>
</tr>
<tr>
<td></td>
<td>&lt;0.001*</td>
<td>&lt;0.001*</td>
<td>&lt;0.001*</td>
<td>&lt;0.001*</td>
<td>&lt;0.001*</td>
<td>&lt;0.001*</td>
<td>0.105</td>
<td>0.002*</td>
<td>0.007*</td>
<td></td>
<td>&lt;0.001*</td>
<td>&lt;0.001*</td>
<td>0.065</td>
<td>0.065</td>
<td>&lt;0.001*</td>
</tr>
<tr>
<td></td>
<td>93.8%</td>
<td>87.5%</td>
<td>87.5%</td>
<td>N.A.</td>
<td>N.A.</td>
<td>87.5%</td>
<td>68.8%</td>
<td>87.5%</td>
<td>50%</td>
<td>81.3%</td>
<td>93.8%</td>
<td>100%</td>
<td>87.5%</td>
<td>87.5%</td>
<td>93.8%</td>
</tr>
<tr>
<td></td>
<td>0.234</td>
<td>0.195</td>
<td>0.328</td>
<td>0.161</td>
<td>0.161</td>
<td>0.130</td>
<td>0.328</td>
<td>0.279</td>
<td>0.382</td>
<td>0.083</td>
<td>0.003*</td>
<td>0.083</td>
<td>0.382</td>
<td>0.382</td>
<td>0.083</td>
</tr>
<tr>
<td></td>
<td>&lt;0.001*</td>
<td>&lt;0.001*</td>
<td>&lt;0.001*</td>
<td>0.442</td>
<td>0.087</td>
<td>0.003*</td>
<td>0.065</td>
<td>0.442</td>
<td>62.5%</td>
<td>62.5%</td>
<td>&lt;0.001*</td>
<td>50.0%</td>
<td>&lt;0.001*</td>
<td>&lt;0.001*</td>
<td>50.0%</td>
</tr>
<tr>
<td></td>
<td>81.3%</td>
<td>75.0%</td>
<td>87.5%</td>
<td>43.8%</td>
<td>62.5%</td>
<td>N.A.</td>
<td>N.A.</td>
<td>N.A.</td>
<td>N.A.</td>
<td>N.A.</td>
<td>50.0%</td>
<td>N.A.</td>
<td>N.A.</td>
<td>N.A.</td>
<td>N.A.</td>
</tr>
<tr>
<td></td>
<td>0.130</td>
<td>0.382</td>
<td>0.130</td>
<td>0.065</td>
<td>0.065</td>
<td>0.083</td>
<td>0.721</td>
<td>0.130</td>
<td>0.003*</td>
<td>0.001*</td>
<td>0.003*</td>
<td>0.065</td>
<td>0.382</td>
<td>0.161</td>
<td>0.083</td>
</tr>
<tr>
<td></td>
<td>&lt;0.001*</td>
<td>&lt;0.001*</td>
<td>&lt;0.001*</td>
<td>0.001*</td>
<td>&lt;0.001*</td>
<td>&lt;0.001*</td>
<td>0.065</td>
<td>&lt;0.001*</td>
<td>75.0%</td>
<td>68.8%</td>
<td>&lt;0.001*</td>
<td>&lt;0.001*</td>
<td>62.5%</td>
<td>62.5%</td>
<td>62.5%</td>
</tr>
<tr>
<td></td>
<td>N.A.</td>
<td>N.A.</td>
<td>N.A.</td>
<td>N.A.</td>
<td>N.A.</td>
<td>N.A.</td>
<td>N.A.</td>
<td>N.A.</td>
<td>N.A.</td>
<td>N.A.</td>
<td>N.A.</td>
<td>N.A.</td>
<td>N.A.</td>
<td>N.A.</td>
<td>N.A.</td>
</tr>
</tbody>
</table>

### 6.5 Conclusion

In this work, a gray level local power matrix (GLLPM) was proposed as novel texture matrix for dynamic texture analysis of CA concentration uptake in DCE-MRI. This matrix describes the local relationship between voxel intensity and neighborhood intensity fluctuation in a 4D fashion with temporal consideration. In the retrospective study using a small animal experiment, the feasibility of using GLLPM based Haralick texture feature dynamics for treatment assessment was demonstrated. In the comparison with current 4D and 3D GLCOM techniques, the derived dynamics characteristics of...
GLLPM based Haralick texture features were better than the current GLCOM based techniques in terms of treatment/control therapeutic effect capture accuracy and calculation efficiency. Perspective study of the proposed GLLPM technique in a larger population is in demand for its potential application in clinical DCE-MRI.
7. Development of a dynamic Fractal Signature Dissimilarity (FSD) method for DCE-MRI treatment response assessment

7.1 Introduction

In Chapter 6, the feasibility of using classic fractal dimensions analysis assessing functional heterogeneity for DCE-MRI treatment response assessment has been successfully demonstrated. Compared to the current PK parametric map statistics, the investigated Rényi dimensions $d_1$ and $d_2$ could more accurately monitor the therapeutic effect difference between treatment/control groups. Although the current results are promising, the current approach suffers to limitations. First, the fractal dimension analysis was conducted on $K^{trans}$ maps after PK analysis, and thus the accuracy of the whole workflow was affected by the PK analysis results, which are sensitive to various technical factors in the imaging protocol. Direct evaluations on DCE images would be the ideal model-free approach which may avoid potential problems in PK analysis. Another limitation of the current method is the lack of temporal information as an inherent and advantageous point of DCE-MRI. The dynamics of possible fractal dimensions with temporal evolution could be valuable for capturing therapeutic responses.

Under the consideration of these points, in this work, a novel dynamic fractal signature dissimilarity (FSD) method is proposed as a novel DCE-MRI texture analysis method. Inspired by the current theories of classic fractal dimensions, this methods
evaluates the DCE image heterogeneity evolution during the CA uptake process. The parameters describing the heterogeneity dynamics are hypothesized to be of potential value for better therapeutic response assessment metrics.

### 7.2 Theory of dynamic FSD

In the general theory of fractal dimension, the complexity of an intensity object defined on the 2D lattice can be described by its extruded 3D area (Alic et al., 2011; Lopes et al., 2011). Such surface area can be calculated through the 3D Blanket method (Peleg et al., 1984). In this method, the surface area of a 3D surface $I$ is estimated by measuring the volume between two defined ‘blankets’, which are two extruded surfaces that are not further than $\varepsilon$ voxels above or below the surface to be measured (Figure 37):

These two surfaces are named as upper blanket $u_\varepsilon(i,j)$ and lower blanket $b_\varepsilon(i,j)$ which are defined as follows:

$$
\begin{align*}
    u_\varepsilon(i,j) &= \max\{u_{\varepsilon-1}(k,l)\}, \text{ with } \| (i,j) - (k,l) \| \leq 1 \\
    b_\varepsilon(i,j) &= \min\{b_{\varepsilon-1}(k,l)\}, \text{ with } \| (i,j) - (k,l) \| \leq 1 \\
\end{align*}
$$

$$u_0 = b_0 = I$$

**Figure 37:** A diagram of a 3D surface (a) and its upper blanket (above the surface) and lower blanket (below the surface) (b).
where \((i,j), (k,l) \in K^+\) depicts the spatial coordinates. The selection of \(\varepsilon\) is also called the resolution for blankets calculation and is represented by the voxel numbers.

The surface area of \(I\) at a certain resolution of \(\varepsilon\), \(A(\varepsilon)\), is then defined as:

\[
V(\varepsilon) = \sum \{u_{\varepsilon}(i,j) - b_{\varepsilon}(i,j)\}
\]

\[
A(\varepsilon) = \frac{[V(\varepsilon) - V(\varepsilon - 1)]}{2}
\]

where \(V(\varepsilon)\) is the measured volume between the upper blanket \(u_{\varepsilon}(i,j)\) and lower blanket \(b_{\varepsilon}(i,j)\) and \(V(0) = 0\). Based on the series \((\varepsilon, A(\varepsilon)), S(\varepsilon)\), which we called as fractal signature, is defined as the slope of the best linear fitting of these three points: \((\log(\varepsilon-1), \log(A(\varepsilon-1))), (\log(\varepsilon), \log(A(\varepsilon))), (\log(\varepsilon+1), \log(A(\varepsilon+1)))\). The magnitude of \(S(\varepsilon)\) can be interpreted as the amount of detailed information on the surface \(A(\varepsilon)\) that is lost when the measuring resolution of the Blanket method is worse (i.e., higher value) than \(\varepsilon\).

When \(I\) comes from a perfect fractal object, \(S(\varepsilon)\) should be a constant for all possible \(\varepsilon\) values (Mandelbrot, 1977). In order to describe the tumor heterogeneity change during the CA uptake, we proposed the concept of dynamic FSD by comparing the \(S(\varepsilon)\) from two DCE tumor volumes that were acquired adjacently:

\[
FSD(t) = \frac{\sum_{n=1}^{N} w_n \sum_{\varepsilon=2}^{\varepsilon_{\text{max}}-1} \{[S_{n,t}(\varepsilon) - S_{n,t-1}(\varepsilon)]^2 \cdot \log(\frac{2\varepsilon + 1/2\varepsilon - 1}{2\varepsilon - 1})\}}{[\sum_{\varepsilon} \log(\frac{2\varepsilon + 1/2\varepsilon - 1}{2\varepsilon - 1})][\sum w_n]}, t > 0
\]

\[
w_n = \frac{v_n}{v} \cdot \frac{1}{d_n}
\]

where \(t\) represents the time point after CA injection and \(t = 0\) represents the injection time point, and \(S_{n,t}(\varepsilon)\) is calculated by using \(n\)th slice of the tumor volume at...
time point \( t \). The weighting factor \( w_n \) accounts for the contribution of the tumor’s \( n \)th slice towards the \( FSD(t) \) calculation, in which \( v_n \) is the sub-volume of the tumor’s \( n \)th slice and \( d_n \) is the distance of the \( n \)th slice to the tumor’s central slice. The location of the central slice was determined by the averaged pre-injection volume. For a generated dynamic FSD curve, two heuristic shape features, Area Under Curve (\( AUC_{\text{FSD}} \)) and Maximum Enhancement (\( ME_{\text{FSD}} \)), were selected to characterize the curve feature and were investigated for therapeutic response assessment application.

### 7.3 Retrospective small animal study

The potential application of dynamic FSD method for treatment response assessment was retrospectively examined in the small animal anti-angiogenesis drug treatment experiment as discussed in section 5.2. For this work, one pre-treatment (Day0) and two post-treatment (Day2/9) scans were studied.

With a focus of tumor heterogeneity change during CA uptake process, the DCE volumes acquired in the first 2-minute post-injection time window were adopted for analysis. The minimum \( \varepsilon \) (In Eq. (35)) was selected same as the voxel size (156μm) and \( \varepsilon_{\text{max}} \) was empirically chosen as 10 times of the voxel size. \( AUC_{\text{FSD}} \) and \( ME_{\text{FSD}} \) were recorded as primary biomarkers of dynamic FSD analysis.

For comparison study, the PK parameter \( K_{\text{trans}} \) results, including tumor mean value, CV, kurtosis, skewness and its classic fractal dimension \( d_1 \) and \( d_2 \) in Chapter 5 were included. Additionally, since the study focuses on the CA uptake stage, the AUC
map of DCE enhancement image $AUC_{MR}$ was also studied. For each scan, $AUC_{MR}$ map was generated in the same first two-minute post-injection time window. Similar to the PK parameter $K_{trans}$, the $AUC_{MR}$ tumor mean value, CV, kurtosis, skewness, $d_1$ and $d_2$ were analyzed for comparison.

For each post-treatment scan, the Mann-Whitney U-test was used to assess the difference of the recorded quantitative values between treatment and control groups. Significance was determined based on a $p$-level less than 0.05 with multi-comparison correction if applicable. Experiments using SVM in a leave-one-out approach were performed to validate the potential use of the recorded metrics in treatment/control group classification.

### 7.4 Key Results

Figure 38 demonstrates an example of data analysis using a Day0 scan. The tumor heterogeneity can be easily appreciated on post-injection DCE axial image slice (a), $AUC_{MR}$ map (d) and CA concentration map (e). The $K_{trans}$ map shows in (f) shares some similarities with (d) and (e) in terms of spatial distribution pattern. (b) shows the CA concentration time curve (solid) and dynamic FSD curve (dashed) in the first 2-minute post-injection time window. During the steady CA uptake stage, the dynamic FSD curve had a prominent peak following the CA injection, while the FSD values after the peak were relatively small and stable. This observation means that the tumor
heterogeneity changes drastically at the beginning of the CA uptake; in contrast, the tumor heterogeneity had little change in the following CA steady uptake.

Figure 38: An example of pre-treatment DCE-MRI scan. (a) Post-injection MR image of a selected slice; (b) The 3D tumor’s average CA concentration curve (blue) and dynamic FSD curve; (c) Pre-injection T₁₀ map; (d) AUCₘₐₜ map generated in the same time window as FSD curve; (e) CA distribution 2 minutes after injection; (f) $K^{\text{trans}}$ distribution
The $K_{\text{trans}}$ results has been reported in Figure 33 in Chapter 5. For a short summary here, at Day9, the treatment group had a significant smaller tumor mean $K_{\text{trans}}$ ($p = 0.021$) and significant larger $d_1$ ($p = 0.013$) and $d_2$ ($p = 0.028$) values. None of the investigated $K_{\text{trans}}$ metrics showed significant differences between treatment/control groups at Day2. The $AUC_{\text{MR}}$ analysis results are presented in Figure 39 with the same layout as Figure 33. Of the investigated metrics, none of them showed significant differences between the treatment/control groups at Day2/Day9.
Figure 39: The longitudinal change of mean $AUC_{MR}$ (a), relative $AUC_{MR}$ (b), $AUC_{MR}$ CV (c), $AUC_{MR}$ kurtosis (d), $AUC_{MR}$ skewness (e), $AUC_{MR} \, d_1$ (f), and $AUC_{MR} \, d_2$ (g)

Figure 40 presents the dynamic FSD analysis results. (a) and (b) shows the dynamic FSD curves at all scan days of two selected animals from the treatment group (red) and control group (blue), respectively. As can be seen, the dynamic FSD curves shared common features with an enhancement peak following the CA injection and a relatively stable post-peak tail. The longitudinal change of $AUC_{FSD}$ and its relative value
are presented in (c) and (d). Both cohorts had reduced $AUC_{FSD}$ value after the treatments, and the relative $AUC_{FSD}$ values (post-treatment value/pre-treatment value) showed significant differences between the two cohorts at both Day2 ($p = 0.029$) and Day9 ($p = 0.005$), suggesting the bevacizumab treatment effect. (e) and (f) illustrate the comparison of $ME_{FSD}$ and its relative values between treatment/control groups. At Day0, the baseline $ME_{FSD}$ values of the control group were higher than those in the treatment group with wider distribution, though the difference was not statistically significant. The relative $ME_{FSD}$ values in the treatment group were significantly higher than the control group values at both Day2 ($p = 0.005$) and Day9 ($p = 0.008$).
Figure 40: The demonstration of dynamic FSD analysis. (a) and (b): dynamic FSD curves on all scan days of a treatment group animal and a control group animal (The Day0 curve in (b) has been shown in Figure 38(b)); (c) and (d): the longitudinal change of $AUC_{FSD}$ and its relative value; (e) and (f): the longitudinal change of $ME_{FSD}$ and its relative value.*indicates that statistically significant difference was found between treatment/control groups

The results of treatment/control groups’ classification experiments using SVM are summarized in Table 9. The $K^{\text{trans}}$ related results has been presented as Table 7 and are repeated here for better comparison. As can be seen, when tumor mean $K^{\text{trans}}$ was
used as the only input parameter, the classification accuracy at Day9 was 68.8%. If classic fractal dimensions \( d_1 \) and \( d_2 \) were selected as the input elements, the classification accuracies were promising at both Day2 (87.5%) and Day9 (100%). Compared to the results using \( K_{\text{trans}} \) metrics, the classification accuracies using \( AUC_{\text{MR}} \) metrics were generally lower than the corresponding statistics of \( K_{\text{trans}} \) results, and \( AUC_{\text{MR}} \) metrics demonstrated better classification accuracies at Day2 rather than Day9. In comparison, the classification accuracies were as high as 93.8% at both Day2 and Day9 using the metrics from the dynamic FSD analysis. In short, the treatment/control group classification accuracy at Day2 using the selected dynamic FSD parameters was higher than the accuracy using the classic texture analysis metrics based on \( K_{\text{trans}} \), and the corresponding accuracies at Day9 were comparable. These results suggest that dynamic FSD analysis may be promising in the early detection of therapeutic effects during a fractionated treatment course.

**Table 9: The SVM results of treatment/control groups’ classification using different techniques for dynamic FSD analysis study**

<table>
<thead>
<tr>
<th>SVM Input</th>
<th>Classification Accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Day2</td>
</tr>
<tr>
<td>K\text{trans}</td>
<td></td>
</tr>
<tr>
<td>Mean ( K_{\text{trans}} )</td>
<td>37.5%</td>
</tr>
<tr>
<td>Rev. Mean ( K_{\text{trans}} )</td>
<td>42.9%</td>
</tr>
<tr>
<td>(Mean ( K_{\text{trans}} ), CV)</td>
<td>43.8%</td>
</tr>
<tr>
<td>(Kurtosis, Skewness)</td>
<td>50.0%</td>
</tr>
<tr>
<td>( (d_1, d_2) )</td>
<td>87.5%</td>
</tr>
<tr>
<td>AUC_{\text{MR}}</td>
<td></td>
</tr>
<tr>
<td>Mean ( AUC_{\text{MR}} )</td>
<td>43.8%</td>
</tr>
<tr>
<td>Rev. Mean ( AUC_{\text{MR}} )</td>
<td>64.3%</td>
</tr>
<tr>
<td>(Mean ( AUC_{\text{MR}} ), CV)</td>
<td>31.3%</td>
</tr>
</tbody>
</table>
7.5 Discussion

In this work, the proposed dynamic FSD analysis is a novel image texture analysis method that evaluates the spatial heterogeneity evolution of DCE tumor volume during the CA uptake. This method is very different from the classic PK analysis based on PK model fitting. The study outcomes from the treatment/control groups demonstrated significant differences for certain dynamic FSD features right after the first treatment delivery (Day2). In contrast, the therapeutic effect of the treatment group for the PK parameter $K_{\text{trans}}$ and its heterogeneity measurement were not obvious until three treatments were delivered (Day9). The treatment/control group classification accuracy at Day2 using the selected dynamic FSD parameters was higher than the accuracy using the classic texture analysis metrics based on $K_{\text{trans}}$, and the corresponding accuracies at Day9 were comparable. These results from dynamic FSD analysis can be attributed to the description of CA uptake heterogeneity dynamics which cannot be captured by conventional PK model analysis. This dynamic FSD analysis also demonstrated several technical advantages for potential clinical application. Since it utilizes the DCE images rather than the model fitting of CA concentration distributions, a measurement of the arterial input function is not needed. Also, the calibration scans...
for the native relaxation calculation are not required. In some clinical DCE-MRI applications with low temporal resolution, such as breast DCE-MRI for early stage diseases, the saved scan time is considerable (Wang et al., 2014). In addition, the dynamic FSD analysis only requires a short scan time immediately after the CA injection, while the PK analysis needs longer post-injection scan time for accurate model fitting. This shorter scan time could potentially reduce the effects of intra-scan motion due to the patient discomfort after a certain time.

The dynamic FSD analysis uses the post-injection DCE images and evaluates the heterogeneity evolution during a very short CA uptake time window. The derived FSD value is a quantitative metric that describes the heterogeneity differences of two volumes. In a dynamic FSD curve, a higher FSD value means the tumor CA distribution pattern at this point has a higher degree of difference from the CA distribution pattern at the prior time point. In another word, a higher FSD value means bigger morphological change of DCE image resulted by CA uptake, which is equivalent to higher CA uptake heterogeneity. The FSD curve is originated from the fractal signature $S(\varepsilon)$ series which can be interpreted as the evaluation of frequency component of the intensity distribution. Specifically, high magnitude of $S(\varepsilon)$ at small $\varepsilon$ relate to prominent ‘high-frequency’ gray level variations, while high magnitude of $S(\varepsilon)$ at large $\varepsilon$ relates to substantial ‘low-frequency’ patterns of the gray level distribution. Compared to Fourier analysis, the curve fitting approach in $S(\varepsilon)$ generation could be less sensitive to ‘bad’
voxels with very high/low intensities and thus be potentially robust for image texture analysis and pattern recognition with high noise level (Peleg et al., 1984)

As indicated by the blue curve in Figure 38 (b), the tumor heterogeneity changed dramatically at the initial CA uptake and remained relatively stable during the continuing steady CA uptake. During the experiment, the steady CA uptake could be observed in all analyzed DCE scans during a short time window after injection, while the CA washout dynamic varies after the initial uptake with many early washout cases. As a result, the first 2 minutes post-injection time were selected for the analysis to keep the focus on the study of CA uptake heterogeneity.

Based on the observation of sharp peaks right after the CA injection, it is speculated that the performance of the proposed dynamic FSD method may be affected by the temporal resolution. To verify this hypothesis, an additional comparison study was performed by generating the FSD curve at simulated 19.8s (x2) and 29.7s (x3) temporal resolutions in the same 2-minute time window. Specifically, each FSD value was calculated using two DCE volumes that were acquired with 19.8s or 29.7s intervals. Figure 41 shows the comparison of the dynamic FSD curves using different temporal resolutions from the same scan shown in Figure 38(b). As can be seen, three curves generated with different temporal resolutions shared similar shape patterns with a prominent peak after the immediate CA injection and a relatively stable tail in the following CA uptake. At the simulated 19.8s temporal resolution, the maximum
enhancement of the FSD curve (green) was found at the same position as the original FSD curve (red). Similarly, the curve generated with 29.7s temporal resolution (blue) demonstrated its maximum as its first datum but with a 9.9s shift due to the limited temporal resolution.

Figure 41: The comparison of dynamic FSD curves generated using the original 9.9s temporal resolution (red), simulated 19.8s (green) and 29.7s (blue) temporal resolutions. The red curve was shown in Figure 6(b).

The comparison of $ME_{FSD}$ and $AUC_{FSD}$ statistics from the simulated 19.8s and 29.7s temporal resolutions can be found as Figures C1 and C2 in Appendix C, respectively. At the simulated 19.8s temporal resolution, significant differences of relative $ME_{FSD}$ between treatment/control groups were found at both Day2 and Day9, but significant differences of relative $AUC_{FSD}$ between two groups were only found at Day9; In contrast, when the temporal resolution was reduced to 29.7s, significant differences of relative $ME_{FSD}$ and relative $AUC_{FSD}$ between two groups were observed only at Day9.
The SVM results of treatment/control groups’ classification using simulated dynamic FSD parameters are summarized as Table 10. For a short summary, as the temporal resolution degrades, the classification at both post-treatment scan days became less accurate. These results suggest that high temporal resolution is favored for the dynamic FSD method. However, the performance of the dynamic FSD method at low temporal resolution could be improved because the spatial resolution can be enhanced during the low temporal resolution scan. In this retrospective simulation study, the benefit of enhanced spatial resolution cannot be addressed. Future perspective experiments would be helpful for the comprehensive demonstration of the dynamic FSD method at low temporal resolutions.

Table 10: The SVM results of treatment/control groups’ classification using dynamic FSD parameters from different temporal resolution

<table>
<thead>
<tr>
<th>Temporal Resolution</th>
<th>SVM Input</th>
<th>Classification Accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(AUC_{FSD}, ME_{FSD})</td>
<td>Day2: 93.8%</td>
</tr>
<tr>
<td></td>
<td>(Rev. AUC_{FSD}, Rev.ME_{FSD})</td>
<td>Day9: 93.8%</td>
</tr>
<tr>
<td>9.9s</td>
<td>(AUC_{FSD}, ME_{FSD})</td>
<td>Day2: 92.9%</td>
</tr>
<tr>
<td></td>
<td>(Rev. AUC_{FSD}, Rev.ME_{FSD})</td>
<td>Day9: 92.9%</td>
</tr>
<tr>
<td>19.8s</td>
<td>(AUC_{FSD}, ME_{FSD})</td>
<td>Day2: 68.8%</td>
</tr>
<tr>
<td></td>
<td>(Rev. AUC_{FSD}, Rev.ME_{FSD})</td>
<td>Day9: 75.0%</td>
</tr>
<tr>
<td>29.7s</td>
<td>(AUC_{FSD}, ME_{FSD})</td>
<td>Day2: 50.0%</td>
</tr>
<tr>
<td></td>
<td>(Rev. AUC_{FSD}, Rev.ME_{FSD})</td>
<td>Day9: 68.8%</td>
</tr>
<tr>
<td></td>
<td>(AUC_{FSD}, ME_{FSD})</td>
<td>Day2: 78.8%</td>
</tr>
<tr>
<td></td>
<td>(Rev. AUC_{FSD}, Rev.ME_{FSD})</td>
<td>Day9: 78.8%</td>
</tr>
</tbody>
</table>

The 3D Blanket method which is used in FSD calculation can be further improved. In 3D Blanket method, a 2D intensity object is extruded as 3D surface and the calculation is conducted in a 3D fashion. In this work we calculated the FSD value of a 3D intensity object (i.e., the tumor) as a weighted average of 2D slices’ FSD values, and
this approach was reasonable and yet empirical (Agner et al., 2011; Alic et al., 2011). It might be appealing to extrude the 3D intensity object as a 4D surface for a possible ‘4D Blanket Method’. Nevertheless, this proposed idea requires extensive theoretical works in algebra and topology, and future studies in mathematics will explore the possibilities of such progress.

7.6 Conclusion

In this work, a dynamic FSD method was developed for DCE-MRI therapeutic response assessment. This method is a novel image texture analysis technique which quantitatively evaluates the CA uptake heterogeneity dynamics based on DCE images. In the retrospective study of the small animal anti-angiogenesis drug treatment experiment, the selected parameters from the dynamic FSD method demonstrated significant differences between treatment/control groups as early as after first treatment delivery; in contrast, the classic PK parameter $K_{\text{trans}}$ and its heterogeneity measurement from classic texture analysis reflected the significant differences between treatment/control groups only after three treatment deliveries. After first treatment delivery, the treatment/control group classification using the selected parameters from the dynamic FSD method achieved highest accuracy in comparison with the classifications using existing metrics. These results suggest that the proposed dynamic FSD analysis are promising in monitoring anti-angiogenesis treatment response. Future
works about the mathematical framework and large scale experiments might be essential towards the clinical application of the proposed method.
8. Integrating transcytosomeal water exchange analysis in multi-model PK analysis for DCE-MRI treatment response assessment

8.1 Introduction

Up to date, treatment assessment using DCE-MRI is usually implemented through the comparison of PK parametric map first order statistics. Although the translational work in Chapter 5 demonstrates the value of classic fractal dimensions $K_{trans}$ map as heterogeneity evaluators for treatment assessment, PK analysis is still mandatory for $K_{trans}$ map derivation, which needs to be accurate and precise to ensure the prompt capture of treatment induced functional changes (Chang and Wang, 2015). So far, the most widely used PK model is the one proposed by Tofts and Kermode in 1991 (Tofts and Kermode, 1991) in which the CA kinetics in the microvessel environment is described by the CA bidirectional transendothelium movement between two compartments, blood plasma and extravascular-extracellular space (EES). The conversion of MR signal to CA concentration is frequently reported as a linear relationship between the CA concentration and the change of longitudinal relaxation $R1 = (1/T1)$ after the CA injection (Eq. (1)). However, as discussed in section 1.2.3, such simple linear relationship relies on the assumption that the extravascular space is a single well-mixed medium and thus the interstitium can be treated as a homogeneous solution. Hence, it is required that water exchange from the intercellular space to EES (transcytosomeal water exchange) must be sufficiently fast (FXL condition). However,
biological tissue could be highly compartmentalized on a histological scale (Landis et al., 1999) and the assumption of FXL may not always be satisfied. If the transcytolemmal water exchange is not fast enough to equilibrate the effects of CA in EES, the linear relationship in Eq. (1) would be violated (Labadie et al., 1994). In practice, the Bloch equations should incorporate the effects of the limited transcytolemmal water exchange rate with a bi-exponential decay term of longitudinal relaxation (Landis et al., 2000). The main result of this modification is given as the shutter-speed (SS) model (Yankeelov et al., 2003). In this model, the transcytolemmal water exchange rate is modelled as the inverse of the mean residence time of water molecules in intracellular space, and the relationship between the CA concentration and longitudinal relaxation rate change is expressed as in a nonlinear equation with the presence of limited transcytolemmal water exchange rate.

The SS model has been applied for the PK characterization of head and neck cancer (Kim et al., 2010; Kim et al., 2007), breast cancer (Huang et al., 2008) and prostate cancer (Li et al., 2013). For treatment response assessment, this model has been investigated in a limited number of clinical studies (Huang et al., 2014; Yankeelov et al., 2007b). However, to our best knowledge, a comprehensive comparison of the SS model versus the classic Tofts model in capturing therapeutic response effect has not been fully explored, especially with the comparison of randomized treatment/control groups in a longitudinal experiment setup with multiple post-treatment evaluation. In addition, the
SS model has been investigated as the sole model in PK analysis, while its integration with the Tofts model for multiple model analysis has not been demonstrated. Potential biomarkers that incorporate information from both Tofts model and SS model may contain richer physiological information for better treatment assessment utility. The present work was conducted to address the aforementioned problems using the same small animal experiment in previous chapters. The potential use of first order statistics, histogram descriptors and spatial heterogeneity indices of the PK parametric maps from the SS models for therapeutic response assessment were firstly demonstrated. Additionally, a novel biomarker that reflects the difference of PK rate constants from the SS model and from the Tofts model was designed. The feasibility of this biomarker for better therapeutic response assessment was investigated.

8.2 Theory of SS model

The major breakthrough of SS model from the classic Tofts model is the separation of intracellular water molecules and EES water molecules. In classic Tofts model, these two types of water molecule were considered as same under FXL condition; in SS model at the potential violation of FXL, the two type of water molecules are differentiated by their mean residence time inside the cell \( t_i \) or inside the EES \( t_e \), respectively. The fraction of water molecules in cell and in EES were parameterized as \( p_i \) and \( p_e \), and \( p_i + p_e = 1 \). By conservation of mass:

\[
t_i \cdot p_e = t_e \cdot p_i \quad (37)
\]
The longitudinal relaxation of the system’s signal was modified using a two-pool exchange formalism (Hazlewood et al., 1974; Woessner, 1961). The solution has a bi-exponential form with the total longitudinal relaxation rate $R_1$ described by two components, $R_{1L}$ and $R_{1S}$, and their fractional apparent population constants, $a_L$ and $a_S$, with $a_L + a_S = 1$:

$$M_z = M_0\{1 - \cos \alpha \cdot [a_L \cdot \exp(-t_iR_{1L}) + a_S \cdot \exp(-t_iR_{1S})]\}$$

(38)

where $\alpha$ is the flip angle, $M_z$ is the instantaneous magnetization, $M_0$ is the Boltzmann equilibrium value, and $t_i$ is the running time for recovery by relaxation. The subscripts $L$ represents the apparent component with a larger $T_1$, and $S$ represents the apparent component with a smaller $T_1$. The three unknown parameter, $R_{1L}, R_{1S}$ and $a_S$, were given as (Buckley et al., 2008):

$$R_{1L} = \frac{1}{2}\left[R_{1i} + R_{1e} + r_1 < CA > + \frac{1}{t_i} + \frac{1}{t_e}\right]$$

$$- \frac{1}{2}\left[R_{1i} - R_{1e} - r_1 < CA > + \frac{1}{t_i} \cdot \frac{1}{t_e}\right]^2 + \frac{4}{t_i t_e}$$

(39)

$$R_{1S} = \frac{1}{2}\left[R_{1i} + R_{1e} + r_1 < CA > + \frac{1}{t_i} + \frac{1}{t_e}\right]$$

$$+ \frac{1}{2}\left[R_{1i} - R_{1e} - r_1 < CA > + \frac{1}{t_i} \cdot \frac{1}{t_e}\right]^2 + \frac{4}{t_i t_e}$$

(40)
\[ a_s = \frac{1}{2} - \frac{1}{2} \left( \frac{(R_{1i} - R_{1e} - r_1 < CA >) \cdot (p_e - p_i) + \frac{1}{\tau_i} + \frac{1}{p_e}}{\left[ \left( R_{1i} - R_{1e} - r_1 < CA > + \frac{1}{\tau_i} - \frac{1}{p_e} \right)^2 + \frac{4(1 - p_e)}{\tau_i^2 p_e} \right]^{1/2}} \right) \]  

(41)

where \(<CA>\) is the CA concentration to be solved, and \(r_1\) the longitudinal relaxation rate change constant. As can be seen, when \(\tau_i\) approximates to zero as in FXL and \(p_e \approx 1\), Eqs. (39) - (41) can be reduced to a single longitudinal relaxation rate \(R_1\):

\[ R_1 \approx p_e (r_1 < CA > + R_{1e}) \]  

(42)

To approximate \(<CA>\) in Eqs. (39) - (41), another approximation was proposed to describe \(R_1\) as \(R_1 \approx R_{1L}\) (Landis et al., 2000). Thus, Eq. (39) can be rewritten as the following one with Eq. (37) employed:

\[ R_1 = \frac{1}{2} \left[ R_{1i} + R_{1e} + r_1 < CA > + \frac{1}{\tau_i} + \frac{1 - p_e}{\tau_i p_e} \right] \]

\[ - \frac{1}{2} \left[ \left( R_{1i} - R_{1e} - r_1 < CA > + \frac{1}{\tau_i} + \frac{1 - p_e}{\tau_i p_e} \right)^2 + \frac{4(1 - p_e)}{\tau_i^2 p_e} \right]^{1/2} \]  

(43)

where \(\tau_i = t_i\) is the mean water residence time in cell, and \(<CA>\) can still be described by the Tofts model equation:

\[ < CA > (t) = K^{trans} \int_0^t C_p(t') \exp\left[-\frac{K^{trans}}{v_e}(t - t')\right] dt' \]  

(7)

where \(C_p(t)\) is the AIF information with prior knowledge. The EES water fraction \(p_e\) is connected with \(v_e\) as (Yankeelov et al., 2007b):

\[ p_e = \frac{v_e}{f_w} \]  

(44)
where $f_w$ is the fraction of water that is accessible to CA particles. The combination of Eqs. (42) - (44) is the shutter-speed (SS) model. To employ this model, $f_w$ needs to be selected as a constant, and it has been modeled as 0.8 for human breast tissue, though very few has been provided for other types of tissue (Yankeelov et al., 2007b). Also, it has been proved to be valid that $R_{1i} \equiv R_{1e} = R_{10}$ as the native longitudinal relaxation rate prior to the CA injection (Li et al., 2008). Thus, three unknown PK parameters need to be solved in SS model: $K_{\text{trans}}$, $v_e$ and $\tau_i$.

### 8.3 Retrospective small animal study

#### 8.3.1 Evaluation of SS model

The SS model outputs for treatment response assessment was retrospectively examined in the small animal experiment as discussed in section 5.2. The pre-treatment (Day0) scan and two post-treatment (Day2/9) scans were studied. For each scan, SS model defined by Eqs. (42) - (44) was firstly adopted for PK analysis. In this work, the constant $f_w$ was selected as 0.8, and both $R_{1i}$ and $R_{1e}$ were approximated by $R_{10}$ using dual flip angle method (Eq. (4)). The SS model was solved by the nonlinear least-squares fitting Levenberg-Marquardt method on a voxel-by-voxel base. To avoid the potential fitting error due to the local minimum the Levenberg-Marquardt iteration was repeated with 25 groups of randomly-selected initial points, and the result with the best fitting quality (least $\chi^2$ value, defined by Eq. (45)) was reported.
For comparative study purpose, the classic Tofts model was also analyzed for each scan in the voxel-by-voxel pattern. The CA extravasation rate constant from the Tofts model ($K_{T}^{\text{trans}}$) and the SS model ($K_{S}^{\text{trans}}$) were reported as the primary PK parameter. For SS model, $\tau_{i}$ was also selected as an additional PK parameter. For each recorded parametric map, the tumor mean value was calculated. For each recorded parametric map, the tumor mean value, CV, kurtosis, skewness, and fractal dimensions $d_1$ and $d_2$ were recorded in the same fashion as in Chapter 5.

As a further utilization of transcytolympmal water exchange analysis, for each scan, a biological subvolume (BV) within the tumor was identified based on $\tau_{i}$ intensity distribution. Based on the histogram of $\tau_{i}$, an intensity threshold was determined in an automatic fashion and was used to identify BV as $\tau_{i}$ an intensity-elevated region. Within this region, the CA extravasation rate constant ($K_{S,BV}^{\text{trans}}$) and intracellular water molecule residence time ($\tau_{i,BV}$) were analyzed. The aforementioned PK parameter metrics (mean, CV, kurtosis, skewness, $d_1$ and $d_2$) of $K_{S,BV}^{\text{trans}}$ and $\tau_{i,BV}$ were recorded.

To measure the performance of the two models in aspect of data fitting quality, the map of average weighted residual sum of squares $\chi^2$ was generated (Kim et al., 2007):

$$\chi^2 = \frac{1}{N} \sum_{i=1}^{N} \frac{(S_i - P_i)^2}{P_i}$$

where $N$ is the length of measurement data series, $S_i$ is the measured data ($C(t)$ for Tofts model and $R_i(t)$ for SS model), and $P_i$ is the fitted data by one of the models. As
a commonly used statistical criteria for model selection, the Bayesian information
criterion (BIC) was reported based on the $\chi^2$ maps from two models (Schwarz, 1978):

$$BIC = \log(\chi^2) + (k \log N)/N$$

(46)

where $N$ is still the length of measurement data series, $\chi^2$ is the median value of
the $\chi^2$ map from one model and $k$ is the model freedom degree (2 for Tofts model and 3
for SS model). A smaller BIC value can be interpreted as better data fitting quality.

Each recorded metric was compared longitudinally, and at each post-treatment
scan day, the Mann-Whitney U-test was used to assess the difference of the recorded
metric between treatment and control groups. Significance was determined based on a
$p$-level less than 0.05 with multi-comparison correction if applicable. To validate the
potential use of the recorded metrics from PK parameter analysis for treatment/control
group classification, experiments using support vector machine (SVM) in a leave-one-
out approach were performed at each post-treatment scan day with single/multiple
metric(s) as input. To determine the model fitting quality, the BIC values of the Tofts
model and the SS model of all examined scans were compared using Wilcoxon signed-
rank test with significance level $p < 0.05$.

8.3.2 Design of novel biomarkers

To utilize the physiological information from both $K_T^{trans}$ and $K_S^{trans}$, two metrics
were designed as below to quantify the numerical value difference of the $K^{trans}$ statistics
from two models:
\[ \Delta K_{\text{trans}} = K_S^{\text{trans}} - K_T^{\text{trans}} \]  \hspace{1cm} (47)

and the normalized (Norm.) difference to the value of SS model:

\[ \text{Norm.} \Delta K_{\text{trans}} = \frac{K_S^{\text{trans}} - K_T^{\text{trans}}}{K_T^{\text{trans}}} \]  \hspace{1cm} (48)

\( \Delta K_{\text{trans}} \) depicts the increase of \( K_{\text{trans}} \) value from Tofts model to SS model, which could reflect the dissimilarity of two models. The similar statistics \( \Delta K_{BV}^{\text{trans}} \) and its normalized value were generated within BV. For both \( \Delta K_{\text{trans}} \) and \( \Delta K_{BV}^{\text{trans}} \), the aforementioned statistics (tumor mean value, CV, kurtosis, skewness, and fractal dimensions) in treatment assessment were investigated following the same workflow as in section 8.3.1.

### 8.4 Key Results

Figure 42 shows the analysis of a representative animal from the treatment group at three DCE-MRI scan days (left column: Day0; middle column: Day2; right column: Day9). The first row presents the comparison of DCE volumes about 60 seconds after the CA injection. Because of the high spatial resolution (156\( \mu \)m), the spatial heterogeneity of intensity distribution within the defined tumor can be readily appreciated. The CA extravasation rate constant \( K_T^{\text{trans}} \) maps from the Tofts model and \( K_S^{\text{trans}} \) maps from the SS model are presented in the 2\(^{nd} \) row and the 3\(^{rd} \) row, respectively. As can be seen, the \( K_T^{\text{trans}} \) map and \( K_S^{\text{trans}} \) map at each scan day were morphologically similar with comparable shape patterns. Compared to \( K_T^{\text{trans}} \) maps, the \( K_S^{\text{trans}} \) maps had higher intensity values across the tumor. The \( K_{\text{trans}} \) hotspots were identified at the same
locations on both set of maps, and the hotspots on $K_S^{\text{trans}}$ maps had relatively larger sizes and higher intensities. The joint histogram of $K_T^{\text{trans}}$ and $K_S^{\text{trans}}$ at each scan day are illustrated as the 4th row in Figure 42. The $K^{\text{trans}}$ histograms from two models had peak positions towards the low intensity region (i.e., positive skewness), and the $K_S^{\text{trans}}$ were likely to have more voxels towards the high intensity region. This observation is consistent with the high $K^{\text{trans}}$ values of $K_S^{\text{trans}}$ maps. The bevacizumab treatment effect was obvious on (f) and (i) after three doses, as the $K^{\text{trans}}$ intensities across the tumor decreased in reference to the pre-treatment maps.

Figure 43 shows the $\tau_i$ results of the same animal in Figure 42 from the transcytosomeal water exchange analysis (left column: Day0; middle column: Day2; right column: Day9). The $\tau_i$ maps across the tumor are presented in the 1st row. As can be observed, at each scan day, $\tau_i$ had an intensity-elevated region with very clear and sharp boundary. Accordingly, the identified BV regions are presented in the 2nd row as red areas. Within these BVs, the transcytosomeal water exchange rate was limited, and the FXL condition didn’t hold. The 3rd row of Figure 43 shows the $\tau_i$ histograms across the tumor, and the axes ranges were cropped for the best illustration. At each scan day, the $\tau_i$ histogram had a very tall and distinct peak near the zero value, and this shape feature is in accordance with the sharp boundary of BVs. The increase of probability sum of all non-zero bins suggested an increase of mean $\tau_i$ value along the experiment. Such increase is also supported by the observable intensity elevation from Figure 42 (d) to (f).
Figure 42: An example of a selected animal’s scans at three scan days. Left column: Day0; middle column: Day2; right column: Day9. First row: post-injection T1w DCE image; 2nd row: $K_{trans}$ maps from the Tofts model; 3rd row: $K_{trans}$ maps from the SS model; 4th row: joint histogram of $K_{trans}$ from the two models. The tumor mean values of this animal are: Tofts: Day0: 0.429 min$^{-1}$, Day2: 0.361 min$^{-1}$, Day9: 0.153 min$^{-1}$; SS: Day0: 0.661 min$^{-1}$, Day2: 0.559 min$^{-1}$, Day9: 0.229 min$^{-1}$
Figure 43: Transcytolemmal water exchange analysis at three scan days. Left column: Day0; middle column: Day2; right column: Day9. First row: the $\tau_i$ maps across the tumor; 2nd row: the identified BV (red area) within the tumor; 3rd row: the $\tau_i$ histograms. The tumor mean values $\tau_i$ of this animal are: Day0: 0.159s; Day2: 0.170s; Day9: 0.193s

Figure 44 reports the comparison of data fitting quality using the Tofts model and the SS model. As an example, (a) and (b) provide the data fitting quality of the Day0 scan reported in Figure 42 and Figure 43. The map of $\chi^2$ ratio, which is defined as
\(\chi^2 (SS \ model)/\chi^2 (Tofts \ model)\), is presented as colormap in (a). Within the tumor, the \(\chi^2\) ratio intensities were generally less than one, and this suggests that \(\chi^2\) values of SS model were less than the values of Tofts model. The histogram of \(\chi^2\) ratio within the 3D tumor was presented in (b). The median \(\chi^2\) ratio was 0.228. The BIC values of all analyzed DCE scans from two models were compared in (c). The BIC values of the SS model were significantly lower than the corresponding values of the Tofts model \((p<.0001)\). These results indicate that the SS model has generally improved data fitting qualities than the Tofts model.

![Figure 44](image)

**Figure 44: The comparison of model fitting quality.** (a): \(\chi^2\) ratio map of the Day0 scan of the same animal used before; (b) \(\chi^2\) ratio histogram; (c) the comparison of BIC values of all analyzed scans

As the primary therapeutic response descriptor, the relative tumor volume (tumor growth rate) results of the treatment and the control groups has been summarized as Figure 32 in Chapter 5. As a quick fact, the relative \(V\) of the treatment group was significantly lower than the control group \((p = 0.002)\). Figure 45 summarizes the tumor mean \(K_T^{trans}\) (a) and \(K_S^{trans}\) (b) on all DCE scan days. At Day0, there was no significant difference between the treatment/control groups \(K_T^{trans}\) and \(K_S^{trans}\) value. At
end of follow-up, both treatment and control groups showed a decreasing trend of $K^{trans}$ using either PK model. After three treatment deliveries at Day9, the treatment group had significantly lower $K_T^{trans}$ ($p = 0.021$) and $K_S^{trans}$ ($p = 0.021$) values than the control group. In contrast, other statistics from both models (CV, kurtosis, skewness, $d_1$ and $d_2$) didn’t show significant difference between the treatment/control groups at both post-treatment scan days.

**Figure 45:** The comparisons of tumor mean $K^{trans}$ statistics using Tofts model (a) and using SS model (b). * indicates statistical significance

Figure 46 shows the comparison of tumor mean $\tau_i$ during the experiment. While the treatment group $\tau_i$ increased during the experiment, the control group value was relatively stable. At Day9, the treatment group had significantly higher mean $\tau_i$ values ($p = 0.045$). In contrast, the CV, $d_1$ and $d_2$ statistics didn’t reflect difference after the experiment. It has to be pointed out that kurtosis and skewness were not evaluated since all $\tau_i$ histograms had a very high peak at the first bin (see Figure 43) and thus the histogram descriptors calculation became trivial.
Figure 46: The comparisons of tumor mean $\tau_i$. * indicates statistical significance

Figure 47 shows the results using $K_{S,BV}^{trans}$ and $\tau_{i,BV}$ within the identified BV. As shown in (a), the mean $K_{S,BV}^{trans}$ across the BV had higher intensities in comparison with Figure 45(a), and the decreasing trend of both treatment/control groups was also observed. It is important to point out that the treatment group had significantly lower mean $K_{S,BV}^{trans}$ values at both Day2 ($p = 0.038$) and Day9 ($p = 0.007$), while the treatment group had significantly lower mean $K_{S,BV}^{trans}$ values only at Day9 ($p = 0.021$) in Figure 45(a). This suggests the great value of $K_{S,BV}^{trans}$ for the early capture of anti-angiogenesis drug treatment effect. Similarly to $\tau_i$ results in Figure 7(b), the treatment group had significantly higher $\tau_{i,BV}$ values at Day9 ($p = 0.045$). In (c) and (d), both spatial heterogeneity indices of $K_{S,BV}^{trans}$ in (c) and (d) showed a decreasing trend for both treatment/control groups at Day2; in contrast, from Day2 to Day9, the treatment group demonstrated an increase of $d_1$ and $d_2$ while the control group’s values kept the decreasing trend. At Day9, the treatment group had significantly higher $K_{S,BV}^{trans}$ $d_1$ ($p =$
0.010) and $d_1 (p = 0.021)$ values after the experiment. The other comparisons of $k_{S,BV}^{\text{trans}}$ and $\tau_i,BV$ metrics didn’t show significant difference between treatment/control groups and thus were not reported as figures due to limited space. The results in Figure 8 suggest that $K_{S,BV}^{\text{trans}}$ and $\tau_i,BV$ from the identified BV could provide additional physiological information for early treatment response assessment.

![Figure 47: The comparisons of mean $K^{\text{trans}}$ from SS model in BV(a), mean $\tau_i$ in BV (b), $d_1$ (c) and $d_2$ (d) of $K^{\text{trans}}$ from SS model in BV. * indicates statistical significance](image)

Figure 48 summarizes the comparison of tumor mean $\Delta K^{\text{trans}}$ and its normalized value (1st row) as well as the mean $\Delta K_{BV}^{\text{trans}}$ and its normalized value in BV. When the evaluation was conducted within the whole tumor, the mean $\Delta K^{\text{trans}}$ indicated significant difference ($p = 0.015$) between the treatment and the control groups at Day9.
This result was similar with the $K^{\text{trans}}$ results from both models in Figure 45. The normalized $\Delta K^{\text{trans}}$ in the whole tumor didn’t reflect significant difference at Day9. When the evaluation was conducted in the identified BV, the treatment group had significantly smaller mean $\Delta K^{\text{trans}}_{\text{BV}}$ values on both Day2 ($p = 0.021$) and Day9 ($p = 0.007$). At the same time, the normalized $\Delta K^{\text{trans}}_{\text{BV}}$ value of the treatment group was also significantly lower at Day9 ($p = 0.015$). Although the control group demonstrated an increase of normalized $\Delta K^{\text{trans}}_{\text{BV}}$ as an deviation of decreasing trend observed in most substudies at Day2, the statistical test results indicated no significance ($p = 0.337$). The other statistics about $\Delta K^{\text{trans}}$ and $\Delta K^{\text{trans}}_{\text{BV}}$ didn’t show any significant difference on post-treatment scan days.
Figure 48: The comparisons of tumor mean $\Delta K_{trans}$ (a), tumor mean normalized $\Delta K_{trans}$ (b), BV mean $\Delta K_{trans}$ (c) and BV mean normalized $\Delta K_{trans}$ (d). * indicates statistical significance

Table 11 summarizes the results of treatment/control groups’ classification experiments using SVM. When the tumor mean value of $K^T_{trans}$ or $K^S_{trans}$ was selected as the sole input elements, the classification accuracy at Day9 was 68.8%. In contrast, the accuracy at Day9 using mean $K^T_{trans}$ as the input was as high as 87.5%. If CV was added to the input, the classification accuracy at Day9 using tumor mean $K^T_{trans}$ and its CV from the Tofts model and using tumor mean $K^S_{trans}$ and its CV from the SS model were improved to 87.5% and 87.5%, respectively. In comparison with the test using mean $K^S_{trans}$ and its CV, the Day2 classification accuracy using mean $K^{S, BV}_{trans}$ and its CV were improved from
62.5% to 68.8%, while the Day9 accuracy was also 87.5% without further improvement. The classifications tests using \( \tau_i/\tau_{i,BV} \) metrics were suboptimal than the tests using \( K_{trans} \) metrics from both models, as the highest achievable accuracy at Day9 were 62.5% using mean \( \tau_{i,BV} \) and its CV. For the designed biomarkers with information from two models, the achieved highest classification accuracy at Day9 was also 87.5% using BV mean \( \Delta K_{trans}^{BV} \) and its CV; on the other hand, at Day2, the achieved accuracy was generally smaller than the results using SS \( K_{trans} \) statistics in BV. In a nutshell, the treatment/control group classification using \( K_{trans} \) across the whole tumor from the Tofts model and the SS model were comparable at Day9; when using the identified BV, the classification accuracies at Day2 using BV SS model PK metrics were improved. The classification accuracies using the designed biomarkers at Day9 were comparable with the results using \( K_{trans}^{BV} \) statistics in the lack of desired further improvement.
<table>
<thead>
<tr>
<th>SVM Input</th>
<th>Classification Accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Tofts Model</strong></td>
<td></td>
</tr>
<tr>
<td>Whole tumor</td>
<td></td>
</tr>
<tr>
<td>Mean $K_{trans}^T$</td>
<td>43.8%</td>
</tr>
<tr>
<td>(Mean $K_{trans}^T$, CV)</td>
<td>50.0%</td>
</tr>
<tr>
<td>(Mean $K_{trans}^T$, Kurtosis, Skewness)</td>
<td>31.3%</td>
</tr>
<tr>
<td>(Mean $K_{trans}^T$, $d_1$, $d_2$)</td>
<td>43.8%</td>
</tr>
<tr>
<td><strong>SS Model</strong></td>
<td></td>
</tr>
<tr>
<td>Whole tumor</td>
<td></td>
</tr>
<tr>
<td>Mean $K_{trans}^S$</td>
<td>50.0%</td>
</tr>
<tr>
<td>(Mean $K_{trans}^S$, CV)</td>
<td>62.5%</td>
</tr>
<tr>
<td>(Mean $K_{trans}^S$, Kurtosis, Skewness)</td>
<td>43.8%</td>
</tr>
<tr>
<td>(Mean $K_{trans}^S$, $d_1$, $d_2$)</td>
<td>50.0%</td>
</tr>
<tr>
<td>Mean $\tau_i$</td>
<td>56.3%</td>
</tr>
<tr>
<td>(Mean $\tau_i$, CV)</td>
<td>37.5%</td>
</tr>
<tr>
<td>(Mean $\tau_i$, $d_1$, $d_2$)</td>
<td>50.0%</td>
</tr>
<tr>
<td><strong>BV</strong></td>
<td></td>
</tr>
<tr>
<td>Mean $K_{trans}^{S,BV}$</td>
<td>50.0%</td>
</tr>
<tr>
<td>(Mean $K_{trans}^{S,BV}$, CV)</td>
<td>68.8%</td>
</tr>
<tr>
<td>(Mean $K_{trans}^{S,BV}$, Kurtosis, Skewness)</td>
<td>68.8%</td>
</tr>
<tr>
<td>(Mean $K_{trans}^{S,BV}$, $d_1$, $d_2$)</td>
<td>68.8%</td>
</tr>
<tr>
<td>Mean $\tau_{i,BV}$</td>
<td>43.8%</td>
</tr>
<tr>
<td>(Mean $\tau_{i,BV}$, CV)</td>
<td>43.8%</td>
</tr>
<tr>
<td>(Mean $\tau_{i,BV}$, $d_1$, $d_2$)</td>
<td>56.3%</td>
</tr>
<tr>
<td><strong>Multi-Model</strong></td>
<td></td>
</tr>
<tr>
<td>Whole tumor</td>
<td></td>
</tr>
<tr>
<td>Mean $\Delta K_{trans}^T$</td>
<td>25.0%</td>
</tr>
<tr>
<td>(Mean $\Delta K_{trans}^T$, CV)</td>
<td>31.3%</td>
</tr>
<tr>
<td><strong>BV</strong></td>
<td></td>
</tr>
<tr>
<td>Mean $\Delta K_{trans}^{BV}$</td>
<td>56.3%</td>
</tr>
<tr>
<td>(Mean $\Delta K_{trans}^{BV}$, CV)</td>
<td>62.5%</td>
</tr>
<tr>
<td>Mean $\text{Norm.} \Delta K_{trans}^{BV}$</td>
<td>43.8%</td>
</tr>
<tr>
<td>(Mean $\text{Norm.} \Delta K_{trans}^{BV}$, CV)</td>
<td>56.3%</td>
</tr>
</tbody>
</table>

### 8.5 Discussion

One significance of this work is the first comparative *in vivo* evaluation of SS model parameters versus Tofts model parameters for DCE-MRI therapeutic response assessment with treatment/control group knowledge. As shown in Figure 42 and Figure 45, the tumor mean values of $K_{S}^{trans}$ from the SS model were higher than the tumor
mean values $K_{trans}$ from the classic Tofts model. This result was consistent with the previously reported results about larger $K_S^{trans}$ (Huang et al., 2014; Kim et al., 2007). The inclusion of limited transcytolemmal water exchange rate has been argued to consider the CA particles that cannot be immediately distributed in EES. Thus, the SS model yielded higher CA concentrations and thus higher CA extravasation rate values in comparison with the classic Tofts model (Yankeelov et al., 2005b). The $K^{trans}$ maps identified by the two models were comparable in Figure 42, and the SS model was found to be a better fit for all scans in terms of BIC comparison in Figure 44. It is important to point out that in Figure 44(a), the $\chi^2$ ratio distribution within the tumor was relatively homogeneous. In the area outside the identified BV (shown as Figure 43(d)) where the FXL assumption was acceptable, the SS model fitting quality was also improved as the $\chi^2$ ratio was less than 1 without noticeable hotspots. These results may serve as the evidence that SS model with the effect of transcytolemmal water exchange could describe the microvessel environment more accurately than the classic Tofts model.

The value of the additional PK parameter $\tau_i$ from the transcytolemmal water exchange analysis for therapeutic response assessment was demonstrated. Previously, $\tau_i$ has been shown to potentially valuable for diagnosis assistance (Li et al., 2005). In terms of treatment assessment, however, some studies concluded that $\tau_i$ may not be able to offer extra information for monitoring treatment response (Yankeelov et al., 2007a). In this work, the tumor mean $\tau_i$ of the treatment group increased in the experiment and
was significantly higher than the corresponding values of the control group. As is presented in the theory (Landis et al., 1999), \( \tau_i = V/(P \cdot A) \) where \( P \) is the diffusional permeability of the cell membrane, \( A \) is the cell surface area and \( V \) is the volume of the cell. As a result, the increased \( \tau_i \) of the treatment group could be the results of the increased cell size and/or decreased diffusional cell permeability. The diffusional cell membrane permeability may be reflected by the apparent diffusion coefficient (ADC) from diffusion weighted imaging (DW-MRI). Additional works with DW-MRI protocol would be beneficial of understanding \( \tau_i \) in therapeutic response assessment.

When including the small animal experiment study results in Chapter 5 in which the extended Tofts model was adopted, it would be interesting to see that the tumor mean \( K_{\text{trans}} \) values from all three models (Tofts, extended Tofts, SS) showed significant differences between treatment/control groups at Day9 only. The visual hints in Figure 33 and Figure 42 illustrate the good similarity in morphology of \( K_{\text{trans}} \) map from three models. This observation serves as a convincible evidence of models’ assumption convergence.

The additional BV identified by \( \tau_i \) is the major breakthrough of the SS model. Previous work has predicted that FXL holds when \( \tau_i^{-1}/p_0 \gg |R_{100} - R_{111}| \), where \( p_0 \) is the fractions of water molecules in EES, \( R_{100} \) is the native longitudinal relaxation rate of EES water molecules, and \( R_{11} \) is the longitudinal relaxation rate of intracellular water molecules. (Landis et al., 1999). For the Day0 scan presented in Figure 2, the \( \tau_i \) in BV
ranged from about 10ms to 1s. If $R_{100}$ was approximated by $R_{10}$ and $R_{1i}$ could be roughly estimated as $R_i$, then the observed mean $|R_1 - R_{10}|$ at the maximum enhancement time point was about $0.233 s^{-1}$, and its maximum value was around $0.952 s^{-1}$. Given the range of $p_0$ as $[0.13, 0.95]$ (Donahue et al., 1995), it is possible that FXL could be invalid in some BV voxels with higher $\tau_i$ values. Without further knowledge of $p_0$, it is infeasible to do further quantitative evaluation of FXL condition distribution in BV.

Within the BV, the mean $K_{trans}^{S,BV}$ could capture the therapeutic effect of the treatment group with statistical significance at both Day2 and Day9. When using first order statistics of $K_{trans}$ (Mean value and CV) for treatment/control classification, the reported results using the extended Tofts model was 43.8%/75% at Day2/Day9 in Table 7. For comparison, in Table 11, the corresponding results using Tofts model were 50.0% and 87.5% at Day2 and Day9; when using the SS model, the results were 62.5%/87.5% at Day2/Day9. As can be seen, the classification accuracy at Day9 from all three models were comparable, but the SS model shows an outstanding accuracy at Day2. Furthermore, in the identified BV, the classification results using the same input option at Day2 can be improved to 68.8% as the highest results of all substudies. Through these comparison, the SS model can be argued as the favored model selection for DCE-MRI early treatment response assessment. However, it must be remembered that the identification of BV requires high spatial temporal resolution to assure the potential segmentation in small tumor volumes. In this experiment, the high spatial resolution

139
was achieved with the dedicated small animal MR unit. For clinical scanners, however, the high spatial resolution requirement is challenging. Future works with clinical DCE-MRI scans would be helpful to validate the great value of BV identification in this work.

Two novel biomarkers in Eqs. (47) and (48) were proposed in this work to integrate the biological information from both Tofts model and SS model. These two metrics reflect the numerical differences of $K^{\text{trans}}$ between two models. In SS model theory indicated by Eq. (42), when $\tau_i$ approximates to zero as in FXL, the SS model degrades to the same linear fashion as Eq. (1). The discrepancy of $K^{\text{trans}}$ value between SS model and Tofts model is very likely to be results of FXL deviation degrees, which is an implicit physiological descriptor. The degree of such deviation could reflect tissue properties and could be sensitive to treatment effect. In Figure 48, the $\Delta K^{\text{trans}}$ could differentiate treatment/control groups at Day9, and $\Delta K^{\text{trans}}_{BV}$ could differentiate treatment/control groups at both Day2 and Day9. These results prove the feasibility of the proposed biomarkers in monitoring treatment effect. In terms of treatment/control classification shown in Table 11, the achieved accuracy using mean $\Delta K^{\text{trans}}_{BV}$ and its CV at Day9 was 87.5%, which is same as the results from SS model and from Tofts model. In addition, at Day2, the achieved accuracy 62.5% was lower than the result from SS model in BV 68.8%. Thus, there was no improvement of using the proposed biomarkers for treatment/control group classification. It has to be acknowledged that the proposed biomarkers are good starts and yet to be simple to incorporate richer informatics. In a
potential future projects with larger small animal population, the design of more sophisticated biomarkers using multiple models could further improve the treatment assessment performance.

8.6 Conclusion

This work first compares the use of SS model with transcytolemmal water exchange analysis versus the classic Tofts model for in vivo therapeutic response assessment with treatment/control group assignment. Results show that when using the $K_{\text{trans}}$ information across the tumor, the performances of the two models in treatment/control differentiation were comparable. When the biological subvolume based on the SS model was adopted, the PK parameters’ metrics were capable of capturing the therapeutic effects as early as after the first treatment delivery, while the Tofts model analysis can only demonstrate the therapeutic effects after three treatment deliveries. Another effort is this work was the development of two novel biomarkers that integrates the rate constant information from two models. Within the BV, the certain biomarker can also reflect significant treatment effect at both post-treatment scans. These results suggest the great value of incorporating SS model in DCE-MRI analysis for early therapeutic response assessment.
9. Summary

This study focused on addressing the current challenges of applying DCE-MRI in clinical radiotherapy assessment. The first part of the efforts in this study was spent on improving DCE-MRI temporal resolution. High temporal resolution is in demand for accurate DCE-MRI parametric mapping, while the current clinical equipment cannot meet the technically desired temporal resolution level. To account for this problem, a novel iterative MR reconstruction method using undersampled k-space data was investigated to accelerate DCE-MRI acquisition. A radial-based golden ratio incorporated undersampling strategy and a Cartesian-based spatiotemporally constrained random sampling strategy were designed to improve the parametric map accuracy. The results showed that at a simulated conservative acceleration factor of four, human brain PK maps that were calculated from the DCE images reconstructed using undersampled data were quantitatively accurate in reference to the PK maps generated from fully sampled data. The PK maps generated from simulated accelerated acquisition were clinically acceptable for radiotherapy assessment purpose. The investigated image reconstruction algorithm and the designed image undersampling strategy can be translated to other fast MR imaging application. Another work was developed to address the potential problems of PK model fitting for high temporal resolution DCE-MRI. A new method that based on a derivative deformation of current extended Tofts model with an advanced Kolmogorov-Zurbenko method was developed. In the
computer simulation study, results showed that Results showed that at both high
temporal resolutions (<1s) and clinically feasible temporal resolution (~5s), the PK
parameter results from this new method were more accurate than the results from two
current calculation methods at clinically relevant noise levels. In addition, the
calculation efficiency of this new method was superior to current methods at high
temporal resolution situation. The clinical application of this method is promising.

The second part of this study aims at the development of novel DCE-MRI
analysis method. Two different approaches were considered. The first one is to develop
model-free texture analysis method that can evaluate the spatial heterogeneity evolution
during DCE-MRI. Inspired by the current trend of radiomics research, this approach
utilized the temporal domain dynamics which is an inherent property of DCE-MRI.
Specifically, two techniques, Gray Level Local Power Matrix (GLLPM) and dynamic
fractal signature dissimilarity (FSD), were developed and were evaluated in a
longitudinal small animal anti-angiogenesis drug therapy experiment, which included a
random assignment of treatment/control group assignments and multiple high
spatiotemporal resolution DCE-MRI scan. Through the comparison with current
GLCOM based techniques, the GLLPM derived texture feature dynamics based on CA
concentration maps were demonstrated as superior in differentiating treatment/control
groups after the experiment. On the other hand, the dynamic FSD method evaluated the
tumor heterogeneity dynamics during the CA uptake directly on DCE images. Results
showed that certain metrics from dynamic FSD analysis could capture the significant therapeutic effect at 1st post-treatment scan day earlier than current treatment assessment statistics. The second approach for developing novel DCE-MRI analysis method for treatment assessment is to integrate multiple models in PK analysis. To achieve this goal, a shutter-speed (SS) model with transcytolemmal water exchange analysis was evaluated in the same small animal experiment. Using this model, a biological subvolume was firstly proposed based on transcytolemmal water exchange rate for better therapeutic assessment. The results showed that rate constants in the identified biological subvolume from the SS model was better than the one from classic Tofts model in early capture of significant treatment effect. The introduced transcytolemmal water exchange rate was also demonstrated as valuable for treatment assessment. Additionally, two novel biomarkers were designed to incorporate the rate constant information from both Tofts model and SS model for richer physiological information. The feasibility of these biomarkers in treatment effect monitoring was demonstrated as effective and promising.

The two parts of this study together provide substantial contribution to addressing the limitations of current DCE-MRI techniques for radiotherapy assessment. The concepts and techniques developed in this study, as well as the understandings and inspirations from the key results, provide valuable knowledge to the development of
DCE-MRI technology towards its clinical application of radiotherapy treatment assessment.
Appendix A

Figure A1: 2D simulation results of $K^{\text{trans}}$. (a) True values; (b) DLLSQ results; (c) ILLSQ results; (d) NLSQ results; (e) difference map of the DLLSQ results; (f) difference map of the ILLSQ results; (g) difference map of the NLSQ results.

Figure A2: 2D simulation results of $k_{ep}$. (a) True values; (b) DLLSQ results; (c) ILLSQ results; (d) NLSQ results; (e) difference map of the DLLSQ results; (f) difference map of the ILLSQ results; (g) difference map of the NLSQ results.
Figure A3: 2D simulation results of $v_p$. (a) True values; (b) DLLSQ results; (c) ILLSQ results; (d) NLSQ results; (e) difference map of the DLLSQ results; (f) difference map of the ILLSQ results; (g) difference map of the NLSQ results.
Appendix B

Let \( P(i,j) \) be the GLCOM derived at a specific direction or the average results of all possible directions where \((i,j)\) describes coordinates in \( P \). \( N_g \) is the number of discrete gray levels in the image \( I \).

Define:

\[
p_x(i) = \sum_{j=1}^{N_g} P(i,j) \quad \text{as the marginal probabilities at row direction};
\]

\[
p_y(i) = \sum_{i=1}^{N_g} P(i,j) \quad \text{as the marginal probabilities at column direction};
\]

\[
p_{x-y}(k) = \sum_{i=1}^{N_g} \sum_{j=1}^{N_g} P(i,j), |i-j| = k, k = 0,1, ..., N_g - 1 \quad \text{as the 1st type differential entropy}
\]

\[
p_{x+y}(k) = \sum_{i=1}^{N_g} \sum_{j=1}^{N_g} P(i,j), |i+j| = k, k = 2,3, ..., 2N_g \quad \text{as the 2nd type differential entropy}
\]

\[
\mu_x \text{ as the mean value of } p_x;
\]

\[
\mu_y \text{ as the mean value of } p_y;
\]

\[
\mu \text{ as the mean intensity of the image;
\]

\[
\sigma_x \text{ as the standard deviation of } p_x;
\]

\[
\sigma_y \text{ as the standard deviation of } p_y;
\]

\[
HX = -\sum_{i=1}^{N_g} p_x(i) \log_2(p_x(i)) \quad \text{as the entropy of } p_x;
\]

\[
HY = -\sum_{i=1}^{N_g} p_y(i) \log_2(p_y(i)) \quad \text{as the entropy of } p_y;
\]

\[
HXY1 = -\sum_{i=1}^{N_g} \sum_{j=1}^{N_g} P(i,j) \log_2(p_x(i)p_y(i)) \quad \text{as the 1st type joint entropy};
\]

\[
HXY2 = -\sum_{i=1}^{N_g} \sum_{j=1}^{N_g} p_x(i)p_y(i) \log_2(p_x(i)p_y(i)) \quad \text{as the 2nd type joint entropy};
\]
The Haralick texture features are defined as:

**Autocorrelation**

\[
\text{Autocorrelation} = \sum_{i=1}^{N_g} \sum_{j=1}^{N_g} ijP(i, j) \tag{B1}
\]

**Cluster Prominence**

\[
\text{Cluster Prominence} = \sum_{i=1}^{N_g} \sum_{j=1}^{N_g} [i + j - \mu_x(i) - \mu_y(i)]^4 P(i, j) \tag{B2}
\]

**Cluster Shade**

\[
\text{Cluster Shade} = \sum_{i=1}^{N_g} \sum_{j=1}^{N_g} [i + j - \mu_x(i) - \mu_y(i)]^3 P(i, j) \tag{B3}
\]

**Cluster Tendency**

\[
\text{Cluster Tendency} = \sum_{i=1}^{N_g} \sum_{j=1}^{N_g} [i + j - \mu_x(i) - \mu_y(i)]^2 P(i, j) \tag{B4}
\]

**Contrast**

\[
\text{Contrast} = \sum_{i=1}^{N_g} \sum_{j=1}^{N_g} |i - j|^2 P(i, j) \tag{B5}
\]

**Correlation**

\[
\text{Correlation} = \frac{\sum_{i=1}^{N_g} \sum_{j=1}^{N_g} ijP(i, j) - \mu_x(i)\mu_j(j)}{\sigma_x(i)\sigma_y(j)} \tag{B6}
\]

**Difference Entropy**
Differnece Entropy \[ \Delta \text{Entropy} = \sum_{i=0}^{N_g-1} p_{x-y}(i) \log_2[p_{x-y}(i)] \] (B7)

Dissimilarity

\[ \text{Dissimilarity} = \sum_{i=1}^{N_g} \sum_{j=1}^{N_g} |i - j| p(i,j) \] (B8)

Energy

\[ \text{Energy} = \sum_{i=1}^{N_g} \sum_{j=1}^{N_g} [p(i,j)]^2 \] (B9)

Entropy

\[ \text{Entropy} = -\sum_{i=1}^{N_g} \sum_{j=1}^{N_g} p(i,j) \log_2[p(i,j)] \] (B10)

Homogeneity 1

\[ \text{Homogeneity 1} = \sum_{i=1}^{N_g} \sum_{j=1}^{N_g} \frac{p(i,j)}{1 + |i - j|} \] (B11)

Homogeneity 2

\[ \text{Homogeneity 2} = \sum_{i=1}^{N_g} \sum_{j=1}^{N_g} \frac{p(i,j)}{1 + |i - j|^2} \] (B12)

Informational Measure of Correlation (IMC) 1

\[ \text{IMC1} = \frac{H - HXY1}{\max\{Hx, Hy\}}, H \text{ is entropy} \] (B13)

Informational Measure of Correlation (IMC) 2
\[ IMC2 = \sqrt{1 - e^{-2(HXY^2-H)}} \]  

Inverse Difference Moment Normalized (IDMN)

\[ IDMN = \sum_{i=1}^{N_g} \sum_{j=1}^{N_g} \frac{P(i,j)}{1 + (|i - j|^2)/N^2} \]  

Inverse Difference Normalized (IDN)

\[ IDN = \sum_{i=1}^{N_g} \sum_{j=1}^{N_g} \frac{P(i,j)}{1 + (|i - j|)/N} \]  

Inverse Variance

\[ Inverse\ Variance = \sum_{i=1}^{N_g} \sum_{j=1}^{N_g} \frac{P(i,j)}{|i - j|^2, i \neq j} \]  

Maximum Probability (MP)

\[ MP = \max\{P(i,j)\} \]  

Sum Average

\[ Sum\ Average = \sum_{i=2}^{2N_g} i \cdot p_{x+y}(i) \]  

Sum Entropy

\[ Sum\ Entropy = - \sum_{i=2}^{2N_g} p_{x+y}(i) \log_2(p_{x+y}(i)) \]  

Sum Variance

\[ Sum\ Variance = - \sum_{i=2}^{2N_g} (i - Sum\ Entropy)^2 \log_2(p_{x+y}(i)) \]  

Variance
Sum Variance = \sum_{i=1}^{N_g} \sum_{j=1}^{N_g} (i - \mu)^2 P(i, j) \quad (B22)

Let \( P(i, j) \) be the GLRLM derived at a specific direction or the average results of all possible directions where \((i, j)\) describes coordinates in \( P \) (incidence of gray level run length with \( j \) at intensity \( i \)). \( N_g \) is the number of discrete gray levels in the image \( I \). \( N_r \) is the number of difference run lengths in \( P \). \( N_p \) is the number of voxels in image \( I \). The GLRLM texture features are defined as:

Short Run Emphasis (SRE)

\[
SRE = \frac{\sum_{i=1}^{N_g} \sum_{j=1}^{N_r} P(i, j) / j^2}{\sum_{i=1}^{N_g} \sum_{j=1}^{N_r} P(i, j)} \quad (B23)
\]

Long Run Emphasis (LRE)

\[
LRE = \frac{\sum_{i=1}^{N_g} \sum_{j=1}^{N_r} P(i, j) \cdot j^2}{\sum_{i=1}^{N_g} \sum_{j=1}^{N_r} P(i, j)} \quad (B24)
\]

Gray Level Non-Uniformity (GLN)

\[
GLN = \frac{\sum_{i=1}^{N_g} [\sum_{j=1}^{N_r} P(i, j)]^2}{\sum_{i=1}^{N_g} \sum_{j=1}^{N_r} P(i, j)} \quad (B25)
\]

Run Length Non-Uniformity (RLN)

\[
RLN = \frac{\sum_{j=1}^{N_r} [\sum_{i=1}^{N_g} P(i, j)]^2}{\sum_{i=1}^{N_g} \sum_{j=1}^{N_r} P(i, j)} \quad (B26)
\]

Run Percentage (RP)
\[ RP = \sum_{i=1}^{N_g} \sum_{j=1}^{N_r} \frac{P(i,j)}{N} \]  

Low Gray Level Run Emphasis (LGLRE)

\[ LGLRE = \frac{\sum_{i=1}^{N_g} \sum_{j=1}^{N_r} P(i,j) \cdot i^2}{\sum_{i=1}^{N_g} \sum_{j=1}^{N_r} P(i,j)} \]  

High Gray Level Run Emphasis (HGLRE)

\[ HGLRE = \frac{\sum_{i=1}^{N_g} \sum_{j=1}^{N_r} P(i,j) \cdot i^2}{\sum_{i=1}^{N_g} \sum_{j=1}^{N_r} P(i,j)} \]  

Short Run Low Gray Level Emphasis (SRLGLE)

\[ SRLGLE = \frac{\sum_{i=1}^{N_g} \sum_{j=1}^{N_r} P(i,j)}{\sum_{i=1}^{N_g} \sum_{j=1}^{N_r} P(i,j)} \]  

Short Run High Gray Level Emphasis (SRHGLE)

\[ SRHGLE = \frac{\sum_{i=1}^{N_g} \sum_{j=1}^{N_r} P(i,j) \cdot j^2}{\sum_{i=1}^{N_g} \sum_{j=1}^{N_r} P(i,j)} \]  

Long Run Low Gray Level Emphasis (LRLGLE)

\[ LRLGLE = \frac{\sum_{i=1}^{N_g} \sum_{j=1}^{N_r} P(i,j) \cdot j^2}{\sum_{i=1}^{N_g} \sum_{j=1}^{N_r} P(i,j)} \]  

Long Run High Gray Level Emphasis (LRHGLE)

\[ LRHGLE = \frac{\sum_{i=1}^{N_g} \sum_{j=1}^{N_r} P(i,j) \cdot j^2}{\sum_{i=1}^{N_g} \sum_{j=1}^{N_r} P(i,j)} \]
Appendix C

Figure C1: The demonstration of dynamic FSD parameter analysis at simulated 19.8s temporal resolution. (c) and (d): the longitudinal change of $AUC_{FSD}$ and its relative value; (e) and (f): the longitudinal change of $ME_{FSD}$ and its relative value.
Figure C2: The demonstration of dynamic FSD parameter analysis at simulated 29.7s temporal resolution. (c) and (d): the longitudinal change of $AUC_{FSD}$ and its relative value; (e) and (f): the longitudinal change of $ME_{FSD}$ and its relative value.
References


Ahearn T S, Staff R T, Redpath T W and Semple S I K 2005 The use of the Levenberg-Marquardt curve-fitting algorithm in pharmacokinetic modelling of DCE-MRI data Physics in Medicine and Biology 50 N85-N92


Armitage P A, Schwindack C, Bastin M E and Whittle I R 2007 Quantitative assessment of intracranial tumor response to dexamethasone using diffusion, perfusion and permeability magnetic resonance imaging Magnetic resonance imaging 25 303-10


Block K T, Uecker M and Frahm J 2007 Undersampled radial MRI with multiple coils. Iterative image reconstruction using a total variation constraint Magnetic Resonance in Medicine 57 1086-98

Bredies K, Kunisch K and Pock T 2010 Total generalized variation SIAM Journal on Imaging Sciences 3 492-526

Brix G, Griebel J, Kiessling F and Wenz F 2010 Tracer kinetic modelling of tumour angiogenesis based on dynamic contrast-enhanced CT and MRI measurements European journal of nuclear medicine and molecular imaging 37 30-51


Buades A, Coll B and Morel J M 2005 A review of image denoising algorithms, with a new one Multiscale Model Sim 4 490-530


Cao Y, Li D, Shen Z and Normolle D 2010 Sensitivity of Quantitative Metrics Derived from DCE MRI and a Pharmacokinetic Model to Image Quality and Acquisition Parameters *Acad Radiol* **17** 468-78


Chambolle A and Pock T 2011 A First-Order Primal-Dual Algorithm for Convex Problems with Applications to Imaging *J Math Imaging Vis* **40** 120-45


Chang Z and Xiang Q S 2006 Highly accelerated MRI by skipped phase encoding and edge deghosting with array coil enhancement (SPEED-ACE) *Medical Physics* **33** 3758-66

Chen L Y, Schabel M C and DiBella E V R 2010 Reconstruction of dynamic contrast enhanced magnetic resonance imaging of the breast with temporal constraints *Magnetic resonance imaging* **28** 637-45


Cheng H L M and Wright G A 2006 Rapid high-resolution T1 mapping by variable flip angles: Accurate and precise measurements in the presence of radiofrequency field inhomogeneity *Magnetic Resonance in Medicine* **55** 566-74


d’Arcy J, Collins D, Rowland I, Padhani A and Leach M 2002 Applications of sliding window reconstruction with cartesian sampling for dynamic contrast enhanced MRI *NMR in biomedicine* **15** 174-83

DCE MRI Technical Committee Q I B A 2012 DCE MRI Quantification Profile, Publicly Reviewed Version. Radiological Society of North America)


Deoni S C, Peters T M and Rut B K 2005 High - resolution T1 and T2 mapping of the brain in a clinically acceptable time with DESPOT1 and DESPOT2 *Magnetic resonance in medicine* 53 237-41


Donoho D L 2006 Compressed sensing *Information Theory, IEEE Transactions on* 52 1289-306


Galloway M M 1975 Texture analysis using gray level run lengths *Computer graphics and image processing* 4 172-9

Gatenby R A, Grove O and Gillies R J 2013 Quantitative imaging in cancer evolution and ecology *Radiology* 269 8-15

Hazlewood C F, Chang D C, Nichols B L and Woessner D E 1974 Nuclear magnetic resonance transverse relaxation times of water protons in skeletal muscle Biophysical journal 14 583-606
Henderson E, Rutt B K and Lee T Y 1998 Temporal sampling requirements for the tracer kinetics modeling of breast disease Magnetic resonance imaging 16 1057-73


Kershaw L E and Cheng H L M 2010 Temporal Resolution and SNR Requirements for Accurate DCE-MRI Data Analysis Using the AATH Model Magnet Reson Med 64 1772-80

Kety S 1960 Blood–tissue exchange methods. Theory of blood-tissue exchange and its application to measurement of blood flow Methods in medical research 8 223-7

Khan F M 2010 The physics of radiation therapy (Baltimore, MD: Lippincott Williams & Wilkins)


Knoll F, Bredies K, Pock T and Stollberger R 2011 Second Order Total Generalized Variation (TGV) for MRI Magnetic Resonance in Medicine 65 480-91


Li X, Rooney W D and Springer C S 2005 A unified magnetic resonance imaging pharmacokinetic theory: intravascular and extracellular contrast reagents *Magnetic Resonance in Medicine* **54** 1351-9


Lopes R, Dubois P, Bhouri I, Bedoui M, Maouche S and Betrouni N 2011 Local fractal and multifractal features for volumic texture characterization Pattern Recognition 44 1690-7


Lustig M, Donoho D and Pauly J M 2007 Sparse MRI: The application of compressed sensing for rapid MR imaging Magnetic Resonance in Medicine 58 1182-95


Luypaert R, Sourbron S, Makkat S and de Mey J 2010 Error estimation for perfusion parameters obtained using the two-compartment exchange model in dynamic contrast-enhanced MRI: a simulation study Physics in Medicine and Biology 55 6431-43

Mandelbrot B B 1977 Fractals: form, chance, and dimension: WH Freeman San Francisco)


Moré J J 1978 Numerical analysis: Springer) pp 105-16

Murase K 2004 Efficient method for calculating kinetic parameters using T-1-weighted dynamic contrast-enhanced magnetic resonance imaging Magnetic Resonance in Medicine 51 858-62


Parker G J, Barker G J and Tofts P S 2001 Accurate multislice gradient echo T1 measurement in the presence of non-ideal RF pulse shape and RF field nonuniformity Magnetic resonance in medicine 45 838-45


Peleg S, Naor J, Hartley R and Avnir D 1984 Multiple resolution texture analysis and classification Pattern Analysis and Machine Intelligence, IEEE Transactions on 5 18-23

Port R E, Knopp M V and Brix G 2001 Dynamic contrast-enhanced MRI using Gd-DTPA: interindividual variability of the arterial input function and consequences for the assessment of kinetics in tumors Magnetic resonance in medicine 45 1030-8


Schabel M C and Morrell G R 2009 Uncertainty in T1 mapping using the variable flip angle method with two flip angles Physics in medicine and biology 54 N1

Schabel M C and Parker D L 2008 Uncertainty and bias in contrast concentration measurements using spoiled gradient echo pulse sequences Physics in Medicine and Biology 53 2345-73

Schwarz G 1978 Estimating the dimension of a model The annals of statistics 6 461-4


contrast-enhanced T1-weighted MRI of a diffusable tracer: standardized quantities and symbols Journal of Magnetic Resonance Imaging 10 223-32
Tsai C M and Nishimura D G 2000 Reduced aliasing artifacts using variable-density k-space sampling trajectories Magnetic Resonance in Medicine 43 452-8
Turnbull L W 2009 Dynamic contrast-enhanced MRI in the diagnosis and management of breast cancer NMR in biomedicine 22 28-39
Wang C, Horton J K, Yin F-F and Chang Z 2015 Assessment of Treatment Response With Diffusion-Weighted MRI and Dynamic Contrast-Enhanced MRI in Patients With Early-Stage Breast Cancer Treated With Single-Dose Preoperative Radiotherapy Initial Results Technology in cancer research & treatment 1533034615593191
Weiss M A 2002 Data structures and algorithm analysis in C Pearson Education Asia
Winkelmann S, Schaeffter T, Koehler T, Eggers H and Doessel O 2007 An optimal radial profile order based on the golden ratio for time-resolved MRI ieee T Med Imaging 26 68-76
Woessner D 1961 Nuclear transfer effects in nuclear magnetic resonance pulse experiments The Journal of Chemical Physics 35 41-8
Diffusion Weighted MRI and Dynamic Contrast Enhanced MRI

Medical physics 42 3641-


Yankeelov T E and Gore J C 2009 Dynamic contrast enhanced magnetic resonance imaging in oncology: theory, data acquisition, analysis, and examples Current medical imaging reviews 3 91


Yankeelov T E, Rooney W D, Li X and Springer C S 2003 Variation of the relaxographic “shutter-speed” for transcytolic water exchange affects the CR bolus-tracking curve shape Magnetic resonance in medicine 50 1151-69

Yarnykh V L 2007 Actual flip-angle imaging in the pulsed steady state: a method for rapid three-dimensional mapping of the transmitted radiofrequency field Magnetic resonance in Medicine 57 192-200

Yaromina A, Krause M and Baumann M 2012 Individualization of cancer treatment from radiotherapy perspective Molecular oncology 6 211-21


Zurbenko I 1986 *The spectral analysis of time series*: Elsevier North-Holland, Inc.)
Biography

Chunhao Wang was born on May 4th, 1990 at Beijing City, the capital of the People’s Republic of China. He attend Zhejiang University at Hangzhou, a beautiful city at Southern China near Shanghai, where he earned his Bachelor of Science degree in physics in June 2012 with several national scholarships from Ministry of Education of China. In August 2012, he came to Duke University to pursue his Ph.D. degree in medical physics. He has published eight peer-reviewed articles including first-author scientific papers on Medical Physics (MedPhys), Magnetic Resonance in Medicine (MRM), Physics in Medicine and Biology (PMB) and Technology in Cancer Research and Treatment (TCRT). He also has 10 conference abstracts including four oral presentations at annual meetings of American Association of Physicist in Medicine (AAPM) and American Society of Radiation Oncology (ASTRO). Chunhao has been titled as The Chancellor’s Scholar by Duke University Graduate School and received Douglas Wagenaar Research Fellowship from Duke University Medical Physics Graduate Program. He also received Scientific Abstract Award of ASTRO annual meeting, Educational Stipend Award from International Society of Magnetic Resonance in Medicine (ISMRM), and Outstanding Academic Award from Duke University Medical Physics Graduate Program.