Toward Developing Pharmacokinetic Response Criteria to Chemoradiation in Head and Neck Cancer Patients Using Dynamic Contrast-Enhanced MRI

by

Jennifer Dixon Onxley

Graduate Program in Medical Physics
Duke University

Date:_____

Approved:

_____
Oana I. Craciunescu, Supervisor

_____
James R. MacFall

_____
Shiva K. Das

Thesis submitted in partial fulfillment of the
requirements for the degree of Master of Science in the
Graduate Program in Medical Physics
in the Graduate School of
Duke University

2012

ABSTRACT

Toward Developing Pharmacokinetic Response Criteria to Chemoradiation

in Head and Neck Cancer Patients Using Dynamic Contrast-Enhanced MRI

by

Jennifer Dixon Onxley

Graduate Program in Medical Physics
Duke University

Date:_____

Approved:

_____
Oana I. Craciunescu, Supervisor

_____
James R. MacFall

_____
Shiva K. Das

An abstract of a thesis submitted in partial fulfillment of
the requirements for the degree of Master of Science in the
Graduate Program in Medical Physics
in the Graduate School
of Duke University

2012

Abstract

Purpose: The purpose of this study was to assess the feasibility of using dynamic contrast-enhanced MRI to monitor early treatment-induced changes in pharmacokinetic (PK) parameters in head and neck cancer patients. The intrinsic variability of three parameters (K^{trans}, v_e, and iAUC60) without treatment intervention was measured and compared to the treatment-induced variability.

Materials and Methods: Twenty patients were imaged while undergoing chemoradiation therapy (CRT) for head and neck malignancies. The imaging protocol included two baseline scans one week apart (B1, B2), and a third scan after 1-2 weeks of chemoradiation (ETX – early treatment). The images were acquired on a 1.5T scanner in the coronal plane with a temporal resolution of 10 sec. A population-averaged arterial input function was calculated from plasma concentration curves in both the left and the right carotids of each patient at each time point (B1, B2, ETX). The statistical significance of using a left/right AIF or a time-point-specific AIF versus the population-averaged AIF was evaluated using Bland-Altman plots. To further ensure the correct calculation of PK parameters, the accuracy of the flip angles produced by the MR scanner was measured in phantoms and a volunteer. PK analysis was performed in iCAD (Nashua, NH) based on the modified Tofts model. This study focuses on two PK parameters used in the

Tofts model (K_{trans}, v_e), and one semi-quantitative parameter that was also calculated in iCAD using an uptake integral approach (iAUC60). K^{trans}, v_e, and iAUC60 were averaged over regions of interest (ROIs), some of which covered primary tumors, and others of which covered known nodal metastases. Bland-Altman plots were used to describe the intrinsic variability in each parameter between the two baseline scans. The coefficient of repeatability (CR) between the baseline values was determined from the Bland-Altman plots and compared to the magnitude of the observed treatment-induced changes.

Results: The plasma parameters for the population-averaged AIF were $a_1 = 27.1135$ kg/liter, $a_2 = 17.6486$ kg/liter, $m_1 = 11.7525$ min^{-1}, and $m_2 = 0.2054$ min^{-1}. The use of a left- or right-sided AIF was determined to be unnecessary, as it did not give statistically different PK parameters than the population-averaged AIF. The use of a time-point-specific AIF was not necessary in most cases, though it may give more accurate results when K^{trans} values are > 1 min^{-1}. The flip angle tests revealed high inaccuracies at a flip angle of 5°, so flip angles ≤ 5° were not used in PK analysis. The intrinsic variability of K^{trans}, v_e, and iAUC60 was very high. For nodes, the CRs from the B1–B2 Bland-Altman plots were 0.725 min^{-1} for K^{trans}, 0.315 for v_e, and 18.15 mM-sec for iAUC60. The fractions of node ROIs which showed treatment-induced changes greater than the CR were 3 out of 14 for K^{trans}, 3 out of 17 for v_e, and 7 out of 17 for iAUC60. For primaries, the CRs were

1.385 min^{-1} for K^{trans}, 0.305 for v_e, and 62.85 mM-sec for iAUC60. The fractions of primary ROIs which showed treatment-induced changes greater than the CR were 0 out of 9 for K^{trans}, 1 out of 11 for v_e, and 2 out of 12 for iAUC60.

Conclusion: A population-averaged AIF for head and neck was generated that accounts for differences in right vs. left carotids, day-to-day AIF fluctuations, and treatment-induced AIF changes. In most cases, it is not necessary to use a side-specific or a time-point-specific AIF. When K^{trans} is greater than 1 min^{-1}, PK parameter accuracy may be improved with the use of a time-point-specific AIF. Using the average AIF, large intrinsic fluctuations were observed in ROI-averaged values of K^{trans}, v_e, and iAUC60, making these parameters poor evaluators of early treatment response in head and neck cancer. Nodes were slightly more likely than primaries to show significant treatment-induced changes. Overall, the use of averaged MR-based PK parameters to assess early treatment response is limited and challenging. An analysis of voxel-based variability might be better suited to this task.

Contents

List of Tables

List of Figures

Acknowledgements

First and foremost, thanks to my advisor, Dr. Oana Craciunescu, for her patient guidance and understanding throughout this project. Oana, it has truly been a pleasure working with you. Thanks to Dr. James MacFall and Dr. Shiva Das for their helpful input and for their willingness to serve on my committee. I am also grateful for Dr. David Brizel's and Dr. David Yoo's work in envisioning, establishing, and guiding this entire study from its beginning. To Dr. Naira Muradyan at iCAD, Inc.: your input and your help with the iCAD software have been crucial to the completion of this project. I very much appreciate your contributions. Finally, to my husband Matt, my family, and my friend Bethany: I am deeply grateful for and humbled by your constant, wise encouragement throughout these past two years.

1. Introduction

1.1 Head and Neck Cancer

Every year, approximately 50,000 people in the United States are newly diagnosed with head and neck cancer [1]. The TNM (Tumor, Nodes, Metastasis) staging system is currently the primary means of assessing disease progression and determining optimum treatment. However, even cancers that have been identically staged will sometimes respond differently to the same treatment. In recent years, researchers have identified several functional imaging methods that offer supplemental ways to evaluate disease progression and more accurately predict and assess a patient's response to treatment [2-6]. Dynamic contrast-enhanced magnetic resonance imaging is one such method.

1.2 DCE-MRI

Dynamic contrast-enhanced magnetic resonance imaging (DCE-MRI) is a noninvasive functional imaging technique that monitors the passage of a contrast agent bolus through a patient's vasculature. As it travels, the contrast agent diffuses into interstitial space and out again in a matter of seconds. This diffusion happens to a small extent in all tissues, but it happens to a greater extent in tumors, which are characterized by leaky microvasculature [7]. Signal enhancement is thus greater in tumors than in normal tissues. Analysis of the magnitude and rate of this signal enhancement yields valuable information about tumor perfusion, vascular density, and vascular

permeability (collectively termed "pharmacokinetic parameters") that can be used to evaluate disease progression, define tumor boundaries, and assess response to therapy [8].

1.3 Project Overview

The ultimate goal of this study is to work within the context of head and neck cancer to quantify the changes in pharmacokinetic (PK) parameters that occur specifically because of treatment intervention, versus those changes that occur intrinsically without treatment intervention. To this end, each patient who enrolled in the study underwent two pre-treatment (baseline) MR scans and one scan after one week of chemoradiation therapy. Analysis of the two baseline scans reveals the extent of PK parameter fluctuations without treatment intervention, while subsequent analysis of the early treatment scan reveals the extent of treatment-induced PK parameter changes.

The validity of this study rests entirely on the ability to obtain accurate PK parameters. The model used for contrast agent kinetics (section 2.3) requires (1) accurate maps of the pre-contrast longitudinal relaxation time (T_{10}), and (2) accurate arterial input functions (AIFs) obtained from fast dynamic MR imaging sequences. Furthermore, T_{10} maps rely on the MR scanner's ability to generate accurate flip angles. Figure 1 shows these sequential dependencies.

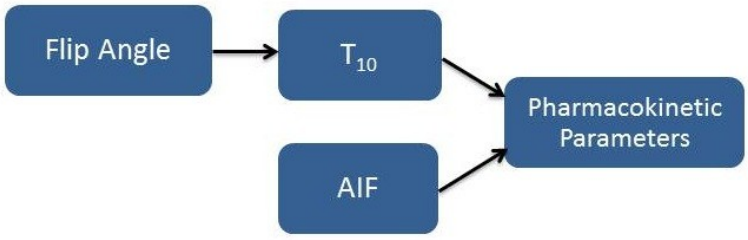


Figure 1: The process of generating correct PK parameters requires accuracy at several steps: flip angle generation, T_{10} calculation, and AIF calculation.

Therefore, to ensure the accuracy of the final PK parameters, we investigate (1) the accuracy of the flip angles produced by our MR scanner, and (2) the feasibility of using a head-and-neck-specific, population-averaged AIF. The final PK parameters are then analyzed at each time point (B1, B2, and early treatment (ETX)) to determine the magnitude of intrinsic vs. treatment-induced changes in their values.

2. Materials and Methods

2.1 Clinical Protocol

Twenty patients were enrolled in this IRB-approved clinical trial in the Department of Radiation Oncology at Duke University Medical Center. All patients had locally advanced, histologically documented squamous cell carcinoma of the head and neck and were undergoing curative intent chemoradiation therapy (CRT). Each patient received three MR scans at three different time points: two pre-treatment baseline scans approximately one week apart (B1 and B2), and one scan that occurred one week into a 7-week course of CRT (ETX – early treatment).

In addition to the DCE-MRI scans, each patient also underwent diffusion-weighted MR scans and ^{18}F-fluorodeoxyglucose (FDG)-PET/CT scans at the same three time points. Diffusion-weighted MRI provides information on the cellular density of a tissue, and FDG-PET/CT provides information on the extent of glucose utilization in a tissue. A further goal of this study, which is beyond the scope of this work, is to quantify the temporal variability of parameters derived from each of these scans, to reveal possible correlations between different functional imaging parameters, and to evaluate their ability to predict treatment outcome.

2.2 MR Imaging Protocol

MRI images were acquired on a 1.5 T scanner (Signa Excite, GE Medical Systems, Waukesha, WI, software versions 14x and 15x) in the Radiation Oncology Department.

Following a single bolus injection of 0.1 mmol/kg Gd-DTPA (Magnevist; Berlex Laboratories, Wayne, NJ), the entire head and neck region was imaged using a dynamic coronal 3D fast gradient echo (fGRE) sequence: TR = 5.8–6.6 ms, TE = 1.8–2.0 ms, field of view = 24x24 cm^2, matrix size = 256x256, temporal resolution = 10 s, number of scans = 19–32, slice thickness = 10 mm, and flip angle = 60° [7]. Imaging was done with a 4-channel flexible phase array coil (GE Healthcare, Waukesha, WI), with a head and neck immobilization device designed during the CT-simulation stage of the radiation treatment.

Additionally, coronal fGRE sequences were acquired at multiple flip angles before contrast injection (α = 2, 5, 10, 15, 20, 30, 45, 60°). These scans were used to calculate pre-contrast longitudinal relaxation times (T_{10}). Finally, a T_1-weighted coronal fast spin echo (fSE) sequence was acquired after contrast injection, which provided better anatomic spatial resolution and aided in ROI delineation. The total scan time was approximately 45 minutes.

2.3 DCE-MRI Theory

2.3.1 Tofts Compartmental Model

Pharmacokinetic analysis in this study is based on the modified Tofts model for contrast agent kinetics [9]. This is a two-compartment model that characterizes the exchange of contrast agent between the vasculature and the extravascular extracellular

space (EES) (Figure 2). The rate constant for this process is K^{trans}, which is related both to the amount of blood flow through the tissue and to the permeability of the vasculature.

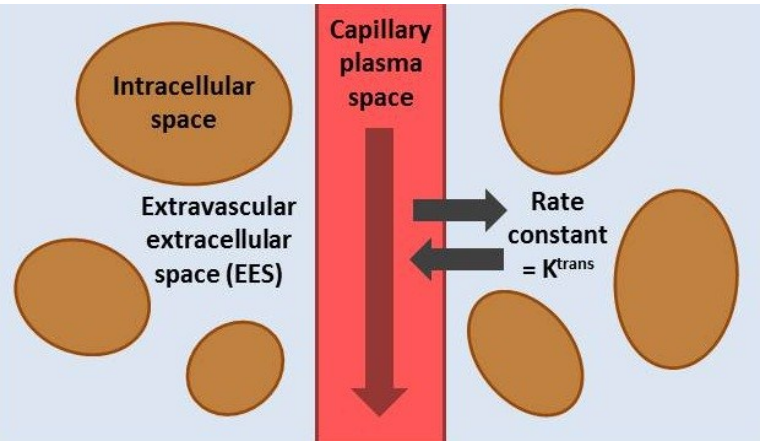


Figure 2: Compartments involved in contrast agent distribution in tissue.

The contrast agent concentrations (mmol/liter) in the plasma, EES, and tissue as a whole are denoted as C_p, C_e, and C_t, respectively. The fractions of tissue volume occupied by plasma and EES are denoted as v_p and v_e, respectively ($0 < v_p, v_e < 1$). The plasma volume v_p is further defined as (1-Hct)*v_b, where v_b is the volume fraction occupied by blood, and Hct is the blood hematocrit, assumed to be 0.425 (an average of 0.4 for women and 0.45 for men) [10]. Similarly, C_p can be further defined as C_b/(1-Hct), where C_b is the contrast agent concentration in blood. The movement of contrast agent across the capillary wall is driven by the concentration difference $C_p - C_e$. Initially assuming that v_p is negligible, and using $C_e = C_t/v_e$, the differential equation that describes this movement is:

$$\frac{dC_t}{dt} = K^{trans}\left(C_p - \frac{C_t}{v_e}\right) \qquad [Eq. 1]$$

$C_P(t)$ is called the arterial input function (AIF) – the time course of contrast agent concentration in the arterial plasma near the tissue of interest. It is modeled by a bi-exponential decay function – a fast component corresponding to plasma-EES equilibration, and a slow component corresponding to the draining of contrast agent from the plasma/EES by the kidneys [11]:

$$C_p(t) = D(a_1 e^{-m_1 t} + a_2 e^{-m_2 t}) \qquad [Eq. 2]$$

where D is the contrast agent dose (mmol/kg), and $a_{1,2}$ and $m_{1,2}$ are plasma parameters. Plugging Eq. 2 into Eq. 1 and solving for tissue concentration C_t yields:

$$C_t(t) = DK^{trans} \sum_{i=1}^{2} a_i \frac{e^{-(K^{trans}/v_e)t} - e^{-m_i t}}{m_i - K^{trans}/v_e} \qquad [Eq. 3]$$

The modified (3-parameter) Tofts model does not assume that v_P (the plasma volume fraction) is negligible. This model must be used when the lesions of interest are highly vascular. The addition of another term to Eq. 3 accounts for the fact that the full expression for C_t is $v_e C_e + V_p C_p$. Eq. 4 is the final equation used for PK analysis:

$$C_t(t) = DK^{trans} \sum_{i=1}^{2} a_i \frac{e^{-(K^{trans}/v_e)t} - e^{-m_i t}}{m_i - K^{trans}/v_e} + v_p D \sum_{i=1}^{2} a_i e^{-m_i t} \qquad [Eq. 4]$$

Non-model-based, semi-quantitative methods are often used in addition to or in place of model-based methods in order to avoid some challenges associated with PK modeling (e.g. computational time, input parameter inaccuracies, fit failures). One of the simplest model-free quantification techniques is an integration of the tissue contrast agent concentration for a specific time interval after injection (Eq. 5). This integration

gives a semi-quantitative parameter known as the initial area under the concentration-time curve (iAUC) [12]:

$$iAUC_t = \int_0^t C_t(t')\ dt' \qquad [Eq.5]$$

The interpretation of iAUC is dependent on the time period over which the integration is performed. Evelhoch [13] suggested that when integration is performed over the first 60-90 seconds post injection, iAUC can be equivalent to Tofts' K^{trans}. This study analyzes iAUC60, for which the upper limit of integration in Eq. 5 is 60 seconds.

2.3.2 MR Data Analysis

For a spoiled gradient echo sequence, assuming TE << T_2^*, the MR signal intensity (S) is given by [14]:

$$S = S_0 \sin\alpha \frac{1 - e^{-TR/T_1}}{1 - \cos\alpha\, e^{-TR/T_1}} \qquad [Eq.6]$$

where α is the flip angle, TR is the repetition time, T_1 is the longitudinal relaxation time, and S_0 is a constant proportional to proton density. For a dynamic scan with contrast injection, T_1 changes as the concentration of the contrast agent changes. If S(0) is the pre-contrast signal intensity (with pre-contrast longitudinal relaxation time T_{10}), and S(t) is the post-contrast signal intensity (with modified longitudinal relaxation time T_1), then signal enhancement E(t) is defined as [7]:

$$E(t) = \frac{S(t) - S(0)}{S(0)} = \frac{S(t)}{S(0)} - 1 = \frac{\left(1 - \cos\alpha\, e^{-TR/T_{10}}\right)\left(1 - e^{-TR/T_1}\right)}{(1 - e^{-TR/T_{10}})(1 - \cos\alpha\, e^{-TR/T_1})} - 1 \qquad [Eq.7]$$

In the presence of Gd-DTPA, T_1 is less than T_{10}. The relaxation rate ($1/T_1$) is assumed to increase linearly with Gd-DTPA concentration [9]:

$$\frac{1}{T_1} = \frac{1}{T_{10}} + R_1 C_t \qquad [Eq. 8]$$

R_1 is the contrast agent relaxivity. A value of $R_1 = 4.33$ sec^{-1}(mmol/liter)$^{-1}$ is used for Gd-DTPA at 1.5 T [15]. Plugging Eq. 8 into Eq. 7 yields:

$$E(t) = \frac{\left(1 - \cos\alpha\, e^{-TR/T_{10}}\right)\left(1 - e^{-TR\left(T_{10}^{-1} + R_1 C_t(t)\right)}\right)}{\left(1 - e^{-TR/T_{10}}\right)\left(1 - \cos\alpha\, e^{-TR\left(T_{10}^{-1} + R_1 C_t(t)\right)}\right)} - 1 \qquad [Eq. 9]$$

T_{10} values are known from variable flip angle scans that are performed before contrast injection. For this MR scanner, the optimal flip angle combination for T_{10} calculation was previously determined to be 10°-45° [7]. The only unknown in Eq. 9 is $C_t(t)$. To determine final PK parameters, $C_t(t)$ is calculated from E(t) using Eq. 9, and K^{trans}, v_e, and v_p are then calculated from $C_t(t)$ using Eq. 4.

2.4 Flip Angle Analysis

As Eq. 9 demonstrates, the accuracy of the final PK parameter values depends on the accuracy of the calculated T_{10} values. There are two common methods used for T_{10} calculation: the variable TR (VTR) method and the variable flip angle (VFA) method.

In the VTR method, multiple scans are taken at different TRs, holding all other scan parameters constant. A flip angle of 90° is used, so that Eq. 6 becomes:

$$S = S_0\left(1 - e^{-TR/T_1}\right) \qquad [Eq. 10]$$

If signal intensity S is known for multiple TRs, the graph of S vs. TR can be fitted to an exponential in the form of Eq. 10 to determine T_1.

A significant weakness of the VTR method is the long acquisition time required. For this reason, T_1 measurement in this study is performed using the VFA method. In this method, two or more MR scans are taken with different flip angles, holding all other scan parameters constant. In the simpler case where just two flip angles are used (α_1, α_2), the ratio of the two signals follows from Eq. 6:

$$\frac{S_{\alpha_1}}{S_{\alpha_2}} = \frac{\sin\alpha_1\left(1-\cos\alpha_2\, e^{-TR/T_{10}}\right)}{\sin\alpha_2\left(1-\cos\alpha_1\, e^{-TR/T_{10}}\right)} \qquad [Eq.\,11]$$

T_{10} values can be analytically calculated from this ratio.

However, as Eq. 11 shows, if the flip angles produced by the MR scanner are not accurate, the accuracy of the calculated T_{10} values will be compromised as well [16, 17]. A method proposed by Stollberger and Wach [17] is used to investigate the accuracy of the flip angles produced by the Radiation Oncology MR scanner. For a gradient echo sequence with $TR > 5T_1$, the exponential terms in Eq. 6 are approximately zero, and the ratio of signal intensities for two scans with flip angles α_1 and α_2 is:

$$\frac{S_{\alpha_2}}{S_{\alpha_1}} = \frac{\sin\alpha_2}{\sin\alpha_1} \qquad [Eq.\,12]$$

If $\alpha_2 = 2\alpha_1$, Eq. 12 becomes:

$$\frac{S_{\alpha_2}}{S_{\alpha_1}} = \frac{\sin 2\alpha_1}{\sin\alpha_1} = \frac{2\sin\alpha_1\cos\alpha_1}{\sin\alpha_1} = 2\cos\alpha_1 \qquad [Eq.\,13]$$

Thus the actual value of flip angle α_1 at each voxel can be estimated by:

$$\alpha_{1,meas} = \cos^{-1}\left(\frac{1}{2}\frac{S_{\alpha_2}}{S_{\alpha_1}}\right) \qquad [Eq. 14]$$

Three phantoms and one volunteer were imaged on the Radiation Oncology MR scanner. The goal of the phantom scans was to determine the actual flip angles the scanner was producing at a wide range of T_1 values. The goal of the volunteer scan was to investigate how the inhomogeneity of the head and neck region affected flip angle accuracy.

Several phantoms were used to cover a wide range of T_1 values. The first phantom was a spherical phantom filled with a low-T_1 (≈80 ms) material. The second phantom was the Magphan® EMR051 phantom, often referred to as the ADNI phantom for its use in the Alzheimer's Disease Neuroimaging Initiative study. The slice of interest for this study contained five spheres with T_1 values ranging from approximately 350 ms to 760 ms. The third phantom was a cylindrical gel phantom chosen because its high T_1 value (≈1400 ms) was close to that of blood. Images of each phantom and the volunteer are shown in Figure 3.

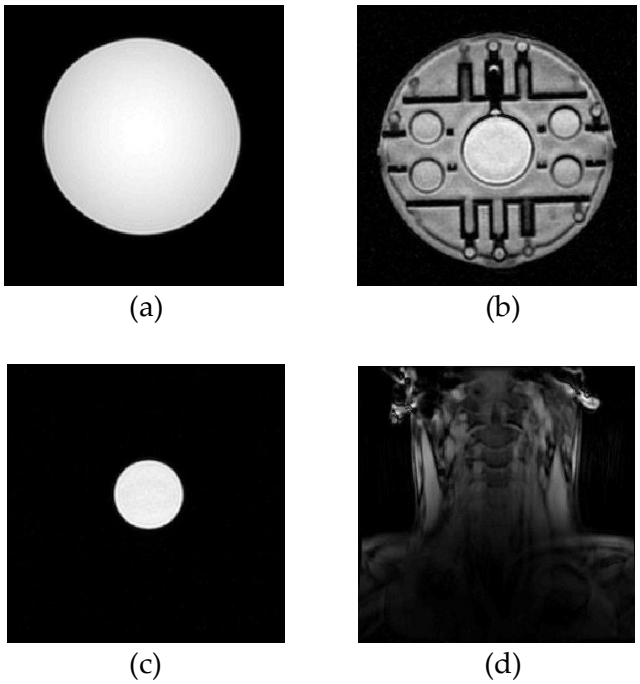


Figure 3: Phantoms and volunteer imaged for FA accuracy study. (a) Low-T_1 phantom; (b) ADNI phantom; (c) High-T_1 phantom; (d) Volunteer.

Each of the phantoms was imaged with three sequences: one sequence that allowed measurement of flip angle accuracy, and two sequences that allowed T_1 calculation by two different methods (VTR and VFA). The volunteer was imaged with the flip angle accuracy sequence and the VFA sequence. Detailed parameters for all scans are listed in Appendix A.

Using the low-T_1 phantom, the five spheres in the ADNI phantom, and the high-T_1 gel phantom, a total of 7 different materials were used to analyze flip angle accuracy. Three ROIs (10 x 10 pixels) were drawn in the low-T_1 phantom: 1 in the center of the phantom and 2 at different distances from the center to the outer edge. The results obtained from these three ROIs were averaged. Five rectangular ROIs were drawn in

the ADNI phantom: 1 to cover the majority of each sphere. One ROI was drawn to cover the majority of the high-T_1 phantom.

Stollberger and Wach's method was applied to each ROI to measure flip angle accuracy. In order to reduce the effects of image noise, the ROI pixel values were averaged, and the method was applied to the average signal intensities. The T_1 value of each of the 7 materials was computed using the VTR method (Eq. 10), in the same ROIs that were used to calculate flip angle accuracy. All ROI analysis was performed in MATLAB, version 7.10.0.499 (R2010a) (MathWorks, Natick, MA).

2.5 T_{10} Variability

A previous study on this MR scanner [7] established that when T_{10} maps cannot be calculated, and when qualitative analysis is sufficient, a T_{10} value of 0.800 s may be used for head and neck primaries and nodes. However, patient-specific, exam-specific T_{10} maps were found to be indispensable to quantitative PK analysis. As part of this study, we sought to confirm these results on another set of patients by analyzing the T_{10} values of primaries and nodes for all time points.

2.6 AIF Analysis

To generate a population-averaged AIF, regions of interest (ROIs) were drawn on coronal subtraction images in both the left and right carotids of each patient at each time point (B1, B2, and ETX). The slice was selected to be centrally located within the coronal 3D slab to allow adequate saturation of the flowing blood signal, and far from

the outer slices of the slab to avoid significant RF field inhomogeneities [18]. Each ROI was 8 x 3 pixels in size [19] and was located in the most enhancing carotid region on the slice to minimize partial volume effects (Figure 4).

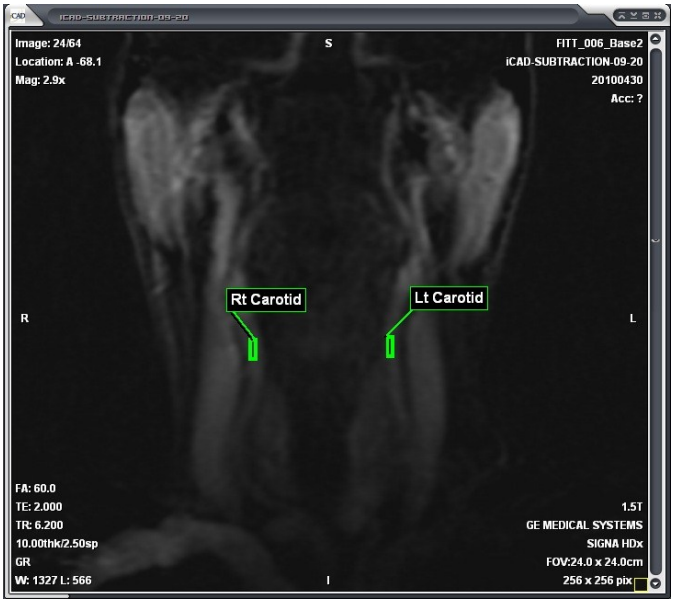


Figure 4: Example of carotid ROIs used to create the AvgAll AIF.

Each ROI was positioned such that its mean T_{10} value was within or close to the range of T_{10} values for arterial blood reported in literature (1.2–1.8 s) [20, 21]. If the carotids were hard to identify, or if a region of appropriate T_{10} values could not be found, or if a carotid did not show the expected degree of enhancement, then that ROI was not included in subsequent analysis. The wash-out phases of the plasma concentration curves obtained from each ROI were averaged to create one final concentration curve, which was then fitted to a bi-exponential decay function [22] as explained in section 2.3:

$$C_p(t) = D(a_1 e^{-m_1 t} + a_2 e^{-m_2 t}),$$

where D = 0.1 mmol/kg. The plasma parameters a_1, a_2, m_1, and m_2 define the population-averaged AIF (AvgAll AIF).

The ROIs were also divided into left carotid ROIs vs. right carotid ROIs. Again, the wash-out phases of each group's plasma concentration curves were averaged and fitted to the bi-exponential decay function to obtain an AvgLeft AIF and an AvgRight AIF. The purpose of generating these two AIFs was to determine whether or not contrast agent kinetics differ significantly between the left and right carotid arteries, since the heart is located on the left side of the body. For five patients with left-sided disease, PK analysis was performed using the AvgAll AIF and the AvgLeft AIF. The resulting two sets of K^{trans} and v_e values were compared using Bland-Altman plots. Similar analysis was performed for five patients with right-sided disease using the AvgAll AIF and the AvgRight AIF.

Finally, to investigate the impact of AIF fluctuations without treatment intervention and treatment-induced AIF changes on PK parameters, the carotid ROIs were divided by time point to obtain AvgB1, AvgB2, and AvgETX AIFs. PK analysis was performed for each MR scan using the corresponding time-point-specific AIF. The resulting K^{trans} and v_e values were compared to those obtained with the AvgAll AIF using Bland-Altman plots. The semi-quantitative parameter iAUC60 is independent of the AIF used and was not analyzed for this part of the study.

2.7 PK Parameter Variability Analysis

Voxel-based PK analysis was performed using post-processing software from iCAD, Inc. (Nashua, NH). The tissue concentration time curve was numerically fitted to the modified Tofts model using the Levenberg-Marquardt algorithm [23] to estimate K^{trans}, v_e and v_p. The fit was deemed acceptable if $R^2 > 0.65$, if values returned were positive, and if v_e and $v_p < 1$. All of the iCAD settings used for these calculations are listed in Appendix B.

Primary tumors and nodes were contoured by a radiation oncologist on either an early subtraction map or on the post-contrast fSE T_1-weighted series and then redirected to the K^{trans} parametric maps. Their entire volumes were used for PK analysis [24]. Only median parameter values are reported in this study. Voxels for which a T_{10} value could not be calculated, voxels that did not enhance, and voxels that could not be fitted to the modified Tofts model were not included in the calculation of the ROI medians. Furthermore, ROIs with median K^{trans} values > 5 min^{-1} were considered non-physiological and were not used in the temporal variability analysis [25].

Temporal variability analysis was performed separately for nodes and primaries. Time points are denoted as B1, B2, AvgB12 (average of B1 and B2), and ETX (early treatment). Each of the following methods of analysis was performed for K^{trans}, v_e, and iAUC60. First, Wilcoxon signed-rank tests were performed between the two baseline measurements and between the baseline and early treatment measurements. Second,

box-and-whisker plots were generated to show each parameter's distribution at each time point. Third, Bland-Altman plots and regression plots were generated to further compare the two baseline measurements. The CRs obtained from these plots were compared to the quantities ETX–B2 and ETX–AvgB12 for each parameter.

2.8 Statistical Methods

Three statistical comparison methods were used in AIF analysis and temporal variability analysis: Wilcoxon signed-rank tests, Bland-Altman plots, and regression plots. Box-and-whisker plots were also used to show data distribution. All plots and tests were done using MedCalc, version 12.2.1.0 (MedCalc Software, Mariakerke, Belgium). The Wilcoxon signed-rank test [26] is used to compare repeated measurements of the same quantity. Unlike the paired t-test, it does not assume that the data is normally distributed. The difference between two measurements is considered statistically significant if the returned p-value is < 0.05.

Bland-Altman plots [27] assess agreement between two different methods of measuring the same quantity. They are scatter plots of the difference between the two methods versus the mean of the two methods. Horizontal lines are drawn at the mean of the differences and at ± 2 times the standard deviation of the differences. The values (mean + 2*SD) and (mean – 2*SD) are the upper and lower limits of agreement, respectively. If differences between the upper and lower limits of agreement are not

clinically significant, then the two methods of measurement may be used interchangeably. The value 2*SD is known as the coefficient of repeatability (CR).

For the temporal variability analysis, Bland-Altman plots are used to show the distribution of intrinsic variability (B1–B2). The CR in this case represents two times the standard deviation of all B1–B2 values for a given parameter. In order for treatment-induced PK parameter changes to be considered significant, the magnitude of the quantity ETX–B2 or ETX–AvgB12 for a given parameter must be greater than the CR for that parameter. (ETX–B2 is the quantity more likely to be measured in clinical settings, where just one baseline is taken right before the start of treatment. However, ETX–AvgB12 is the quantity that more accurately reflects intrinsic temporal variability.)

Regression plots simply show the relationship between two variables. In this study, they are presented next to Bland-Altman plots as an additional way to visualize the differences between two measurements. A line of equality (depicting the ideal case when the two measurements are always equal) is drawn for reference.

Box-and-whisker plots [28] show the distribution of a particular variable. The lower and upper limits of the box represent the 25^{th} and 75^{th} percentile, respectively, and the line between these limits represents the median. The difference between the 75^{th} percentile and the 25^{th} percentile is the interquartile range. The "whiskers" extend from the minimum observation to the maximum observation, excluding outliers. An outlier is

defined as any value less than the 25th percentile minus 1.5 times the interquartile range, or any value greater than the 75th percentile plus 1.5 times the interquartile range.

3. Results

All twenty patients completed all three MR scans, except for one patient who was unable to complete the ETX scan.

3.1 Flip Angle Analysis

The use of Stollberger and Wach's method for flip angle calculation presented some difficulties at low flip angles (5° and 10°). For some ROIs, the method failed because the signal intensities were such that the ratio $S_{\alpha 2}/(2*S_{\alpha 1})$ (Eq. 14) was greater than 1, and taking the inverse cosine of this ratio gave an imaginary number. In these cases, the method was applied pixel-by-pixel, and only the pixels for which the method returned a real number were used in subsequent analysis.

The low-T_1 phantom was found to have a T_1 of 81 ms. The five ADNI spheres had T_1s of 355, 394, 569, 716, and 763 ms. The high-T_1 phantom had a T_1 of 1410 ms. (The true T_1 values of these phantoms are unknown.) Table 1 shows the measured flip angles in each material.

Table 1: Measured FA in 7 different materials.

Nominal FA | T_1 (ms)	81	355	394	569	716	763	1410
5	5.63	11.75	9.80	10.91	10.70	8.43	19.96
10	4.41	12.19	9.65	10.14	9.38	11.72	11.3
20	18.99	21.55	21.05	21.04	20.67	21.60	19.29
30	23.73	32.00	30.89	30.84	30.77	31.44	28.41
40	36.64	42.88	41.50	41.35	41.00	42.19	38.51
60	54.96	64.32	62.15	61.48	61.36	62.69	57.92

Figure 5 shows the above data in graphical form (measured flip angle vs. nominal flip angle), with a line of equality drawn for reference.

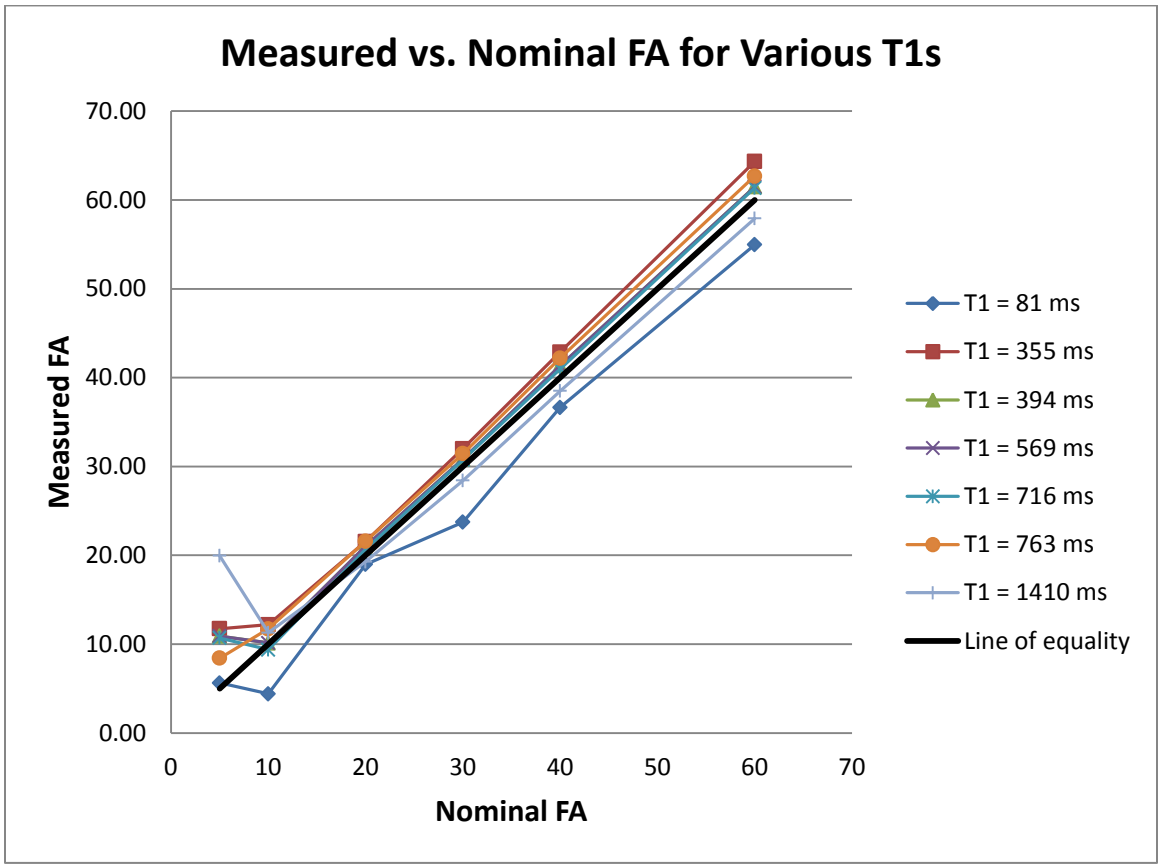


Figure 5: Measured vs. nominal flip angles in 7 different materials.

An alternate way to display the data in Table 1 is shown in Figure 6, which plots the average percent error in the flip angle at various T_1 values for five different nominal flip angles. The data for a flip angle of 5° is not shown because of the high percent errors (maximum error = 300% at T_1 = 1410 ms).

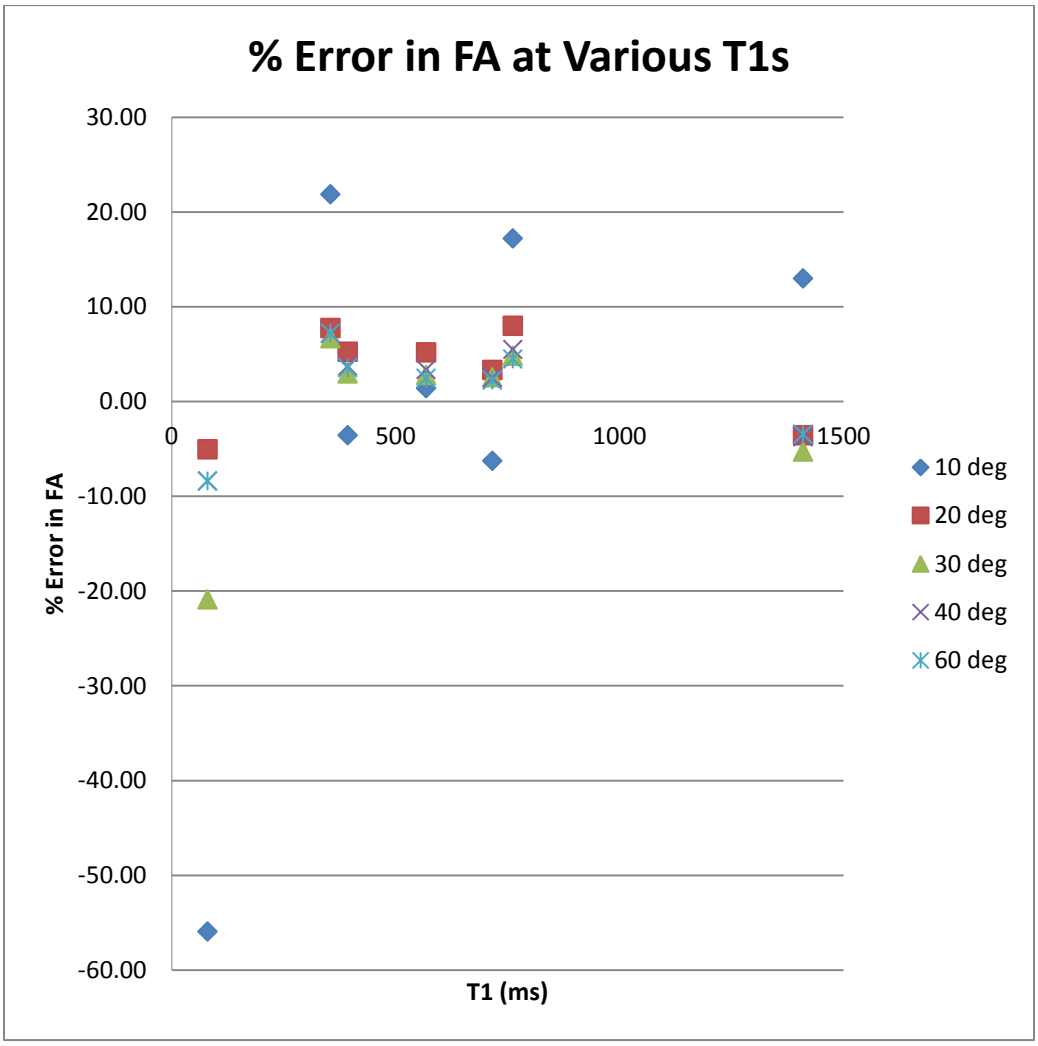


Figure 6: Percent error in flip angle vs. T_1 for five different nominal flip angles.

Figures 5 and 6 show that in phantoms, the highest flip angle inaccuracies occur at low flip angles and low T_1s.

To investigate the effects of head and neck inhomogeneities on flip angle accuracy, flip angle maps of the volunteer were generated at nominal flip angles of 5°, 10°, and 20° by applying Stollberger and Wach's method pixel-by-pixel. These maps are shown in Figure 7.

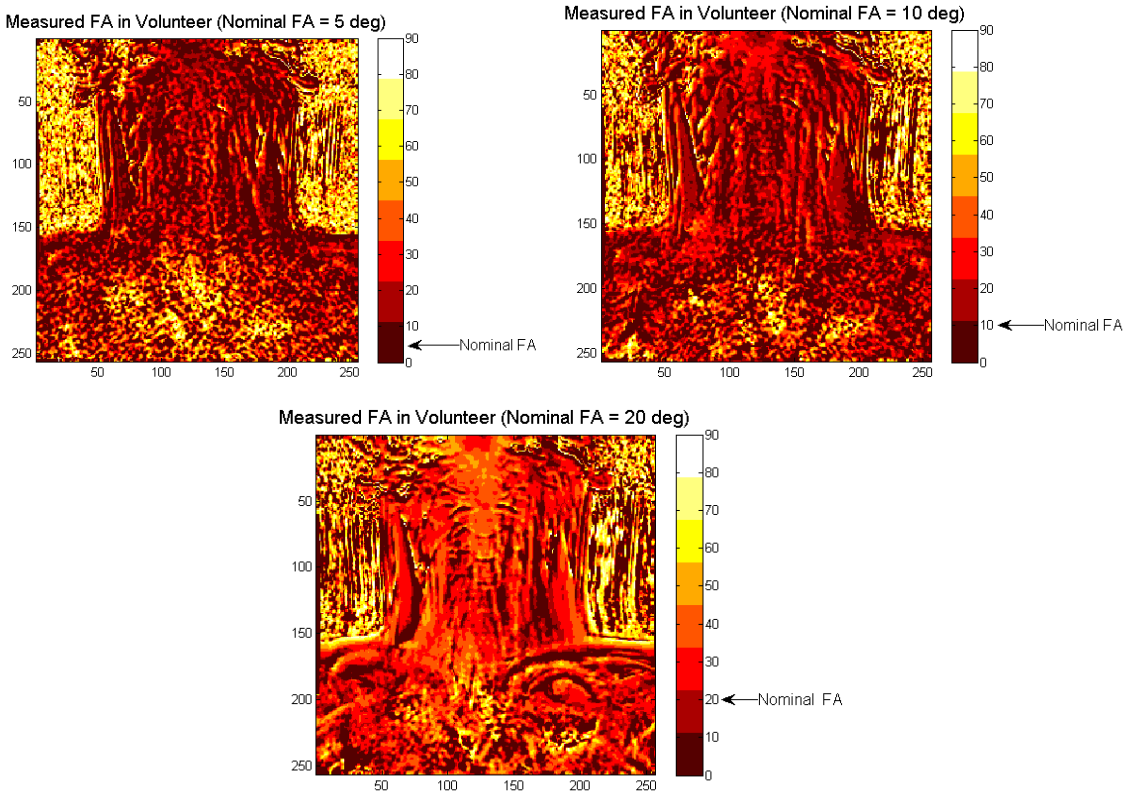


Figure 7: Flip angle maps of the head and neck region of a volunteer.

3.2 T_{10} Variability

We briefly assessed the median T_{10} values of the primaries and nodes at all time points. When primaries, nodes, and time points are all averaged together, the resulting T_{10} value is 0.71 ± 0.29 s. For primaries alone, this value becomes 0.75 ± 0.09 s, and for nodes alone, 0.66 ± 0.11 s. There was significant variability in T_{10} between time points. On average, T_{10} values changed by 39% between B1 and B2 and by 45% between B2 and ETX. This high variation provides evidence that T_{10} calculations should be performed at each time point. The VFA scans listed in the MR protocol (section 2.2) were performed to accomplish exactly this purpose.

3.3 AIF Analysis

3.3.1 Population-Averaged AIF

A total of 93 carotid ROIs were used to generate the population-averaged AIF, with mean ± SD T_{10} values of 1.37 ± 0.25 s. The plasma parameters for the final population-averaged AIF (AvgAll) were a_1 = 27.1135 kg/liter, a_2 = 17.6486 kg/liter, m_1 = 11.7525 min^{-1}, and m_2 = 0.2054 min^{-1}.

A total of 13 primaries and 18 nodes were contoured and used for subsequent AIF and variability analysis. The 2-parameter Tofts model was used initially, but many of the primaries and nodes were so vascular that the majority of their voxels could not be fitted to the 2-parameter model. The analysis was therefore repeated using the modified (3-parameter) Tofts model, and only these PK parameter values are reported.

3.3.2 Left vs. Right Carotid AIF

The plasma concentration curves of 46 left carotid ROIs and 47 right carotid ROIs were averaged to create the AvgLeft and AvgRight AIFs. Their plasma parameters are shown in Table 2, along with the coefficients for the AvgAll AIF for comparison. The three AIFs are graphed in Figure 8.

Table 2: AvgLeft and AvgRight AIF coefficients.

AIF	a_1 (kg/liter)	a_2 (kg/liter)	m_1 (min^{-1})	m_2 (min^{-1})
AvgAll	27.1135	17.6486	11.7525	0.2054
AvgLeft	24.6248	17.3245	11.1080	0.2019
AvgRight	29.6236	17.9358	12.0975	0.2078

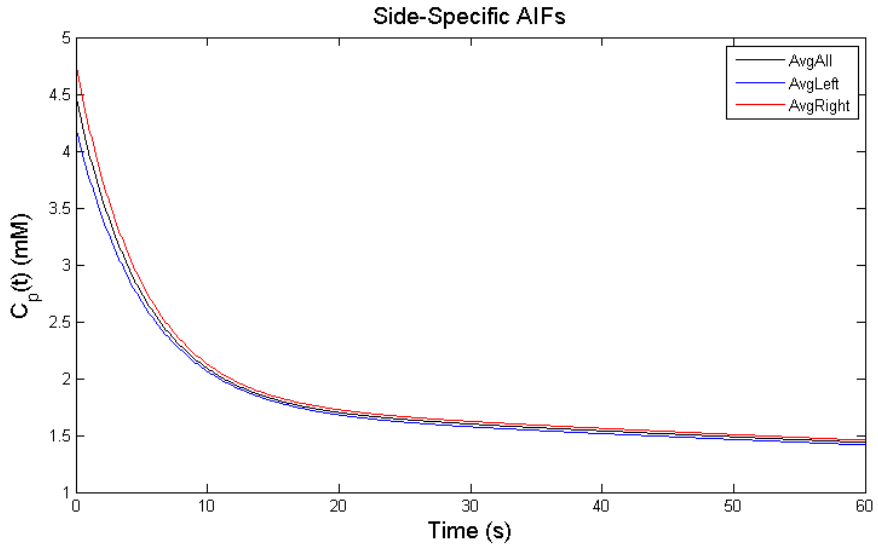


Figure 8: Side-specific AIFs.

To investigate the impact of using a one-sided AIF instead of the AvgAll AIF, PK analysis was performed with the AvgAll, AvgLeft, and AvgRight AIFs, and the resulting K^{trans} and v_e values were compared using Bland-Altman plots and regression plots. Figures 9 and 10 compare AvgAll with AvgLeft, and Figures 11 and 12 compare AvgAll with AvgRight.

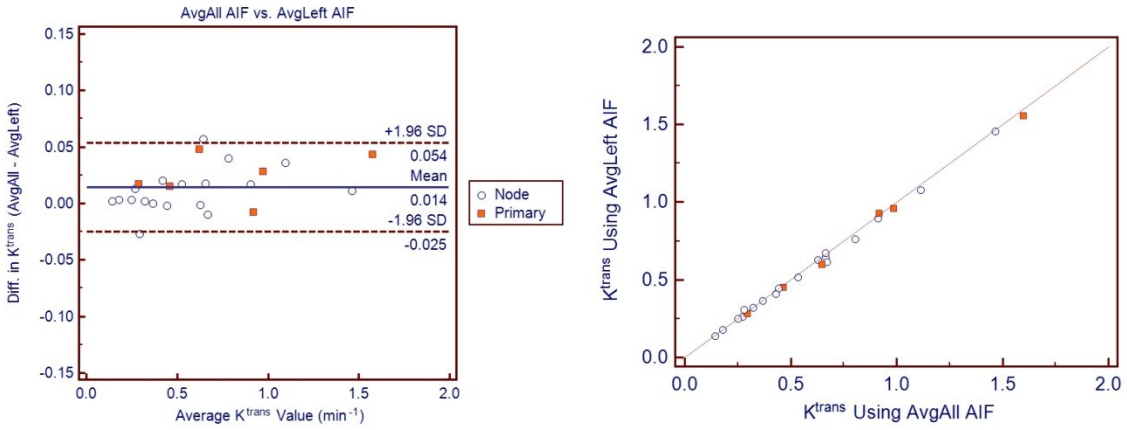


Figure 9: Comparison of AvgAll AIF and AvgLeft AIF for K^{trans}: Bland-Altman plot (left) and regression plot with line of equality (right).

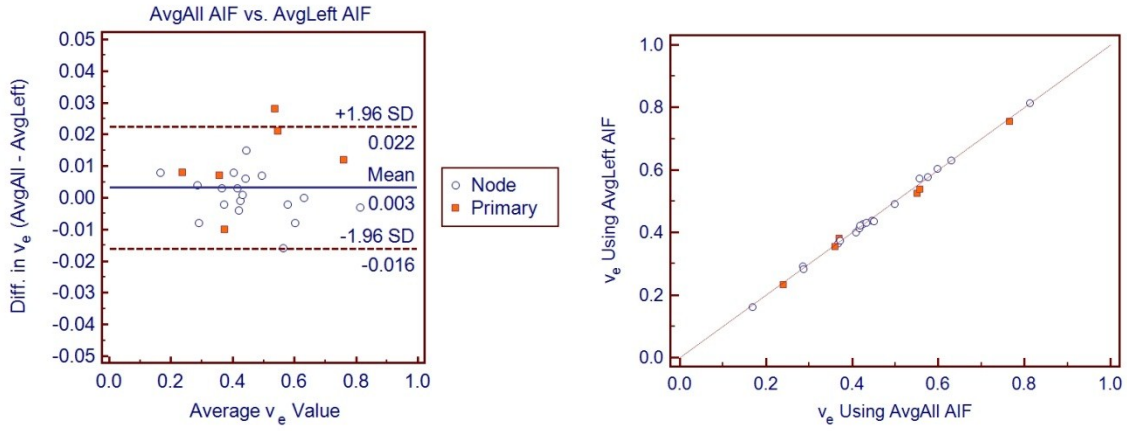


Figure 10: Comparison of AvgAll AIF and AvgLeft AIF for v_e: Bland-Altman plot (left) and regression plot with line of equality (right).

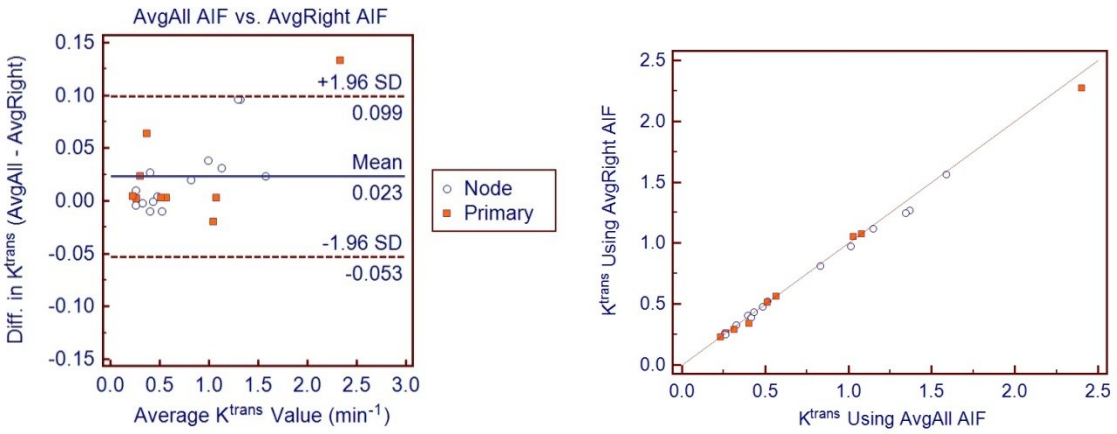


Figure 11: Comparison of AvgAll AIF and AvgRight AIF for K^{trans}: Bland-Altman plot (left) and regression plot with line of equality (right).

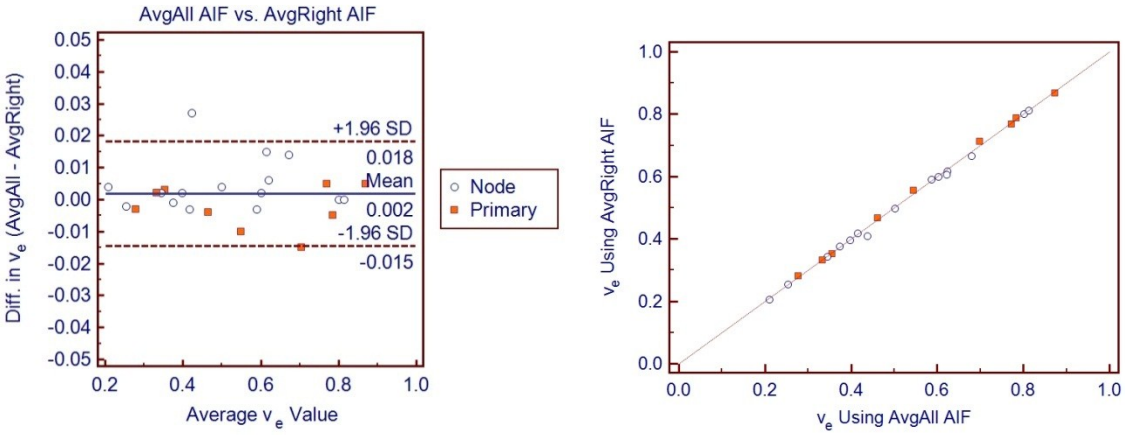


Figure 12: Comparison of AvgAll AIF and AvgRight AIF for v_e: Bland-Altman plot (left) and regression plot with line of equality (right).

The limits of agreement in the Bland-Altman plots in Figures 9-12 are summarized in Table 3.

Table 3: Limits of agreement for K^{trans} and v_e when comparing the AvgAll AIF with the one-sided AIFs.

	K^{trans} limits of agreement	v_e limits of agreement
AvgAll – AvgLeft	-0.025 and 0.054	-0.016 and 0.022
AvgAll – AvgRight	-0.053 and 0.099	-0.015 and 0.018

All of the limits of agreement in Table 3 were very small compared to the range of measured K^{trans} and v_e values. Moreover, all of the points on all four regression plots were close to the line of equality.

3.3.3 Baseline vs. ETX AIF

The plasma concentration curves of 32 B1 carotid ROIs, 33 B2 carotid ROIs, and 28 ETX carotid ROIs were averaged to create the AvgB1, AvgB2, and AvgETX AIFs. Their plasma parameters are shown in Table 4, along with the coefficients for the AvgAll AIF for comparison. The four AIFs are graphed in Figure 13.

Table 4: AvgB1, AvgB2, and AvgETX AIF coefficients.

AIF	a_1 (kg/liter)	a_2 (kg/liter)	m_1 (min^{-1})	m_2 (min^{-1})
AvgAll	27.1135	17.6486	11.7525	0.2054
AvgB1	28.6142	17.5421	14.7221	0.2314
AvgB2	30.8994	17.8612	8.7204	0.1876
AvgETX	22.8739	18.0620	16.6394	0.2085

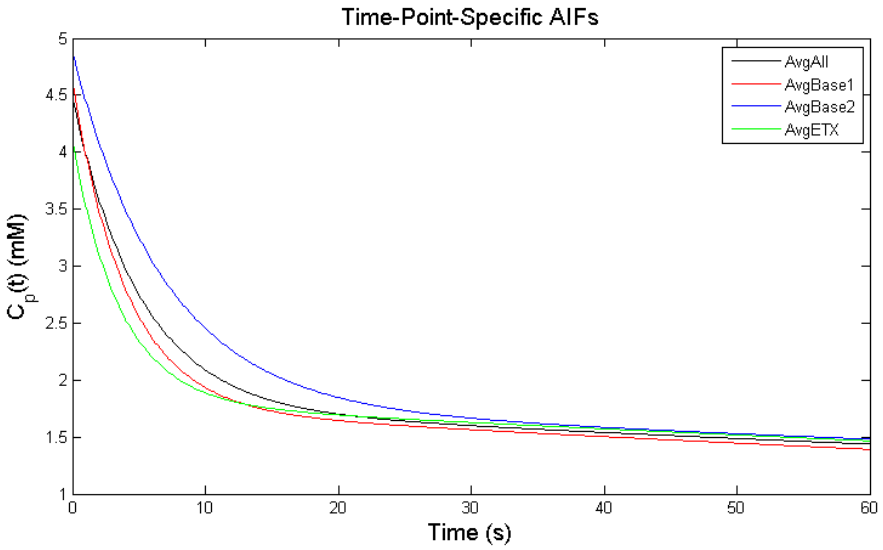


Figure 13: Time-point-specific AIFs.

The AvgAll AIF was compared with each of the time-point-specific AIFs using Bland-Altman plots and regression plots for K^{trans} and v_e (Figure 14-19).

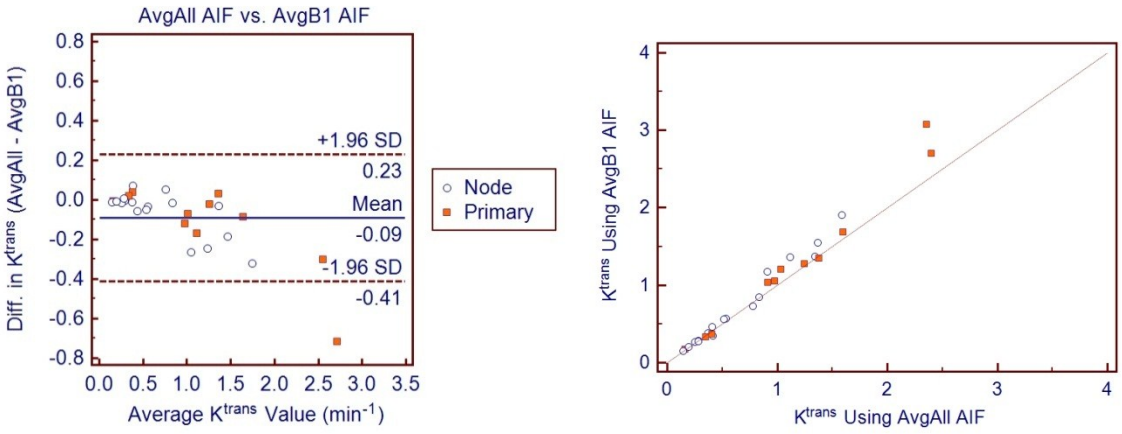


Figure 14: Comparison of AvgAll AIF and AvgB1 AIF for K^{trans}: Bland-Altman plot (left) and regression plot with line of equality (right).

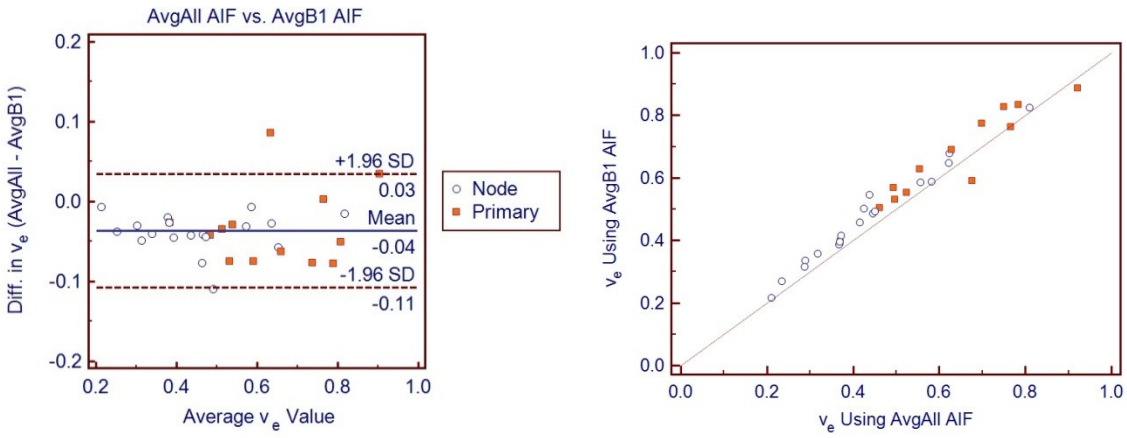


Figure 15: Comparison of AvgAll AIF and AvgB1 AIF for v_e: Bland-Altman plot (left) and regression plot with line of equality (right).

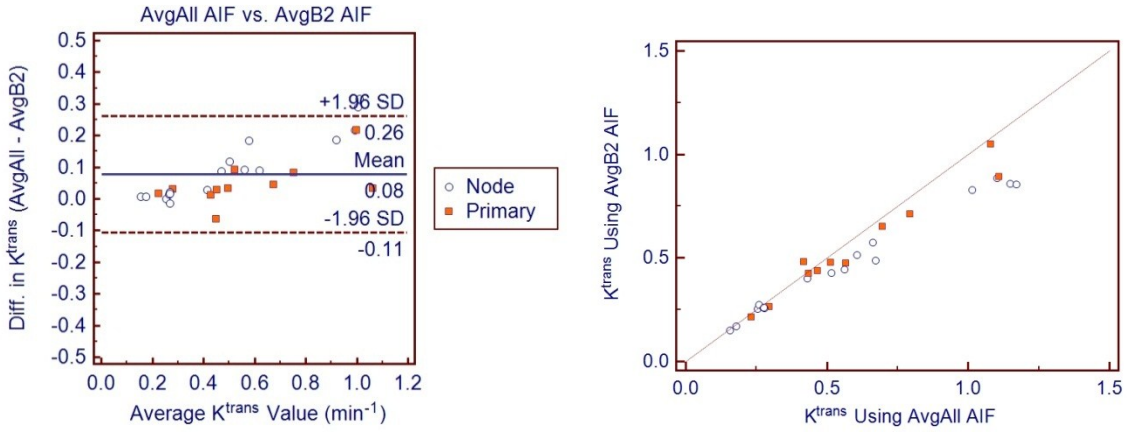


Figure 16: Comparison of AvgAll AIF and AvgB2 AIF for K^{trans}: Bland-Altman plot (left) and regression plot with line of equality (right).

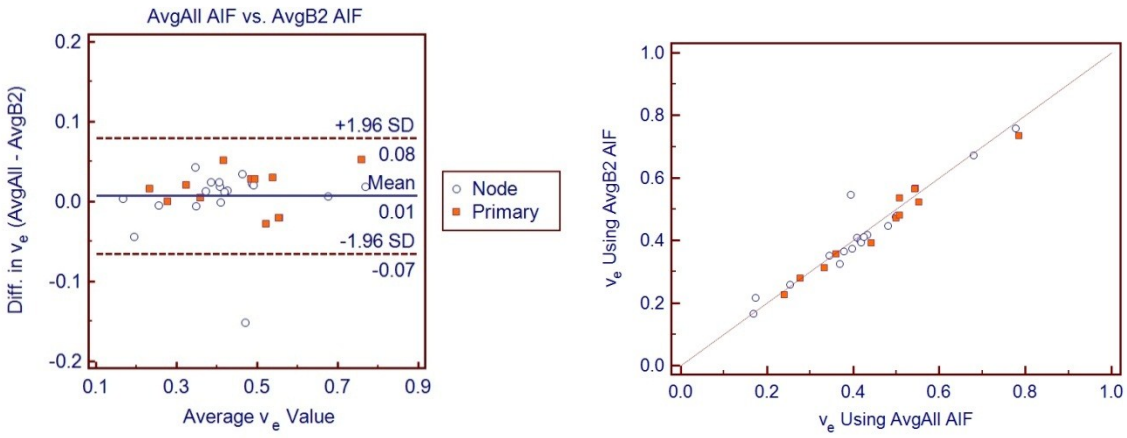


Figure 17: Comparison of AvgAll AIF and AvgB2 AIF for v_e: Bland-Altman plot (left) and regression plot with line of equality (right).

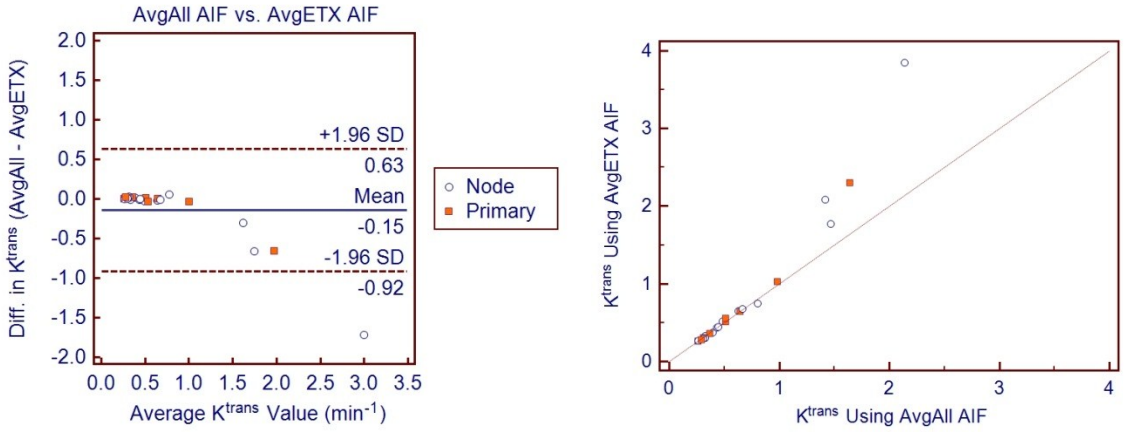


Figure 18: Comparison of AvgAll AIF and AvgETX AIF for K^{trans}: Bland-Altman plot (left) and regression plot with line of equality (right).

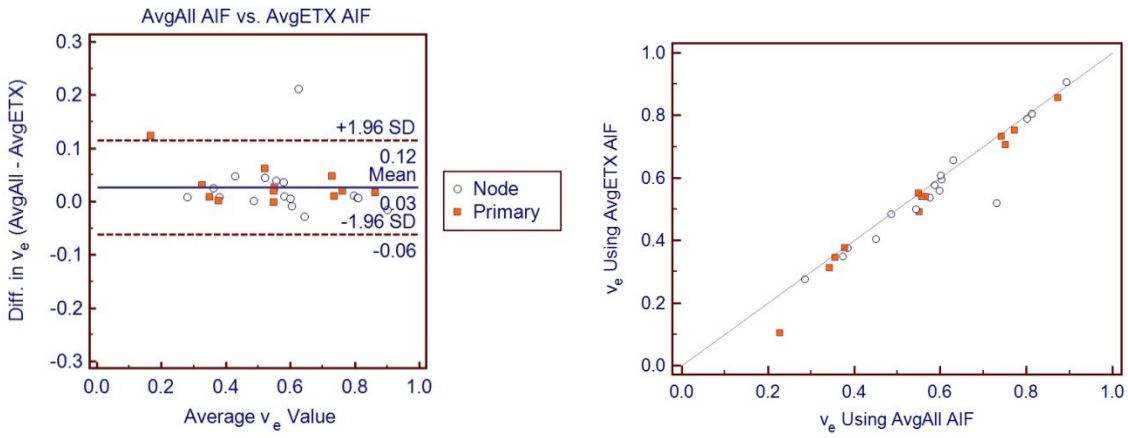


Figure 19: Comparison of AvgAll AIF and AvgETX AIF for v_e: Bland-Altman plot (left) and regression plot with line of equality (right).

The limits of agreement in the Bland-Altman plots in Figures 14-19 are summarized in Table 5.

Table 5: Limits of agreement for K^{trans} and v_e when comparing the AvgAll AIF with the time-point-specific AIFs.

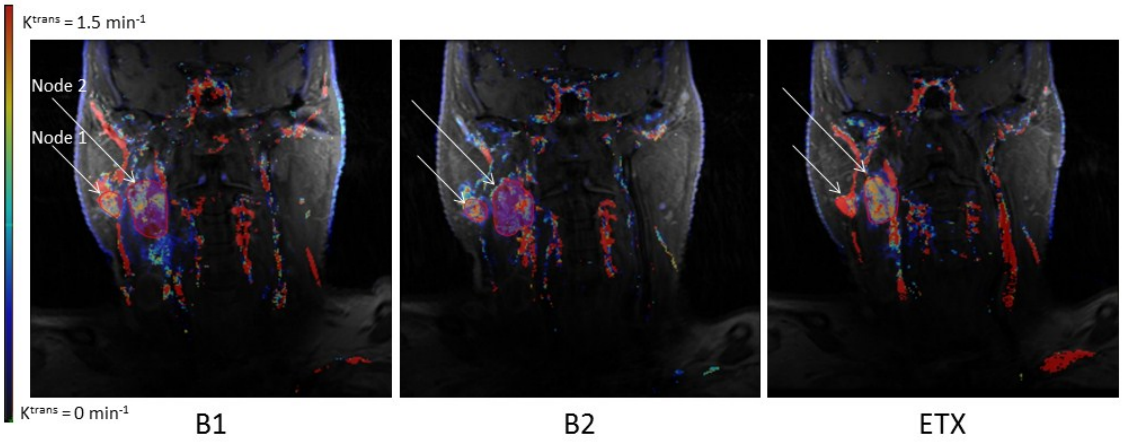
	K^{trans} limits of agreement	v_e limits of agreement
AvgAll – AvgB1	-0.41 and 0.23	-0.11 and 0.03
AvgAll – AvgB2	-0.11 and 0.26	-0.07 and 0.08
AvgAll – AvgETX	-0.92 and 0.63	-0.06 and 0.12

The K^{trans} limits of agreement in Table 5 are particularly large relative to the typical range of measured K^{trans} values. However, in each case (B1, B2, ETX), these large limits for K^{trans} are influenced by only a small number of points. For example, picking an arbitrary value of 0.1 min^{-1} as a clinically insignificant difference in measured K^{trans} values, we calculated the number of points in each K^{trans} Bland-Altman plot whose difference exceeded that number. For B1, only 8 out of the 28 points corresponded to

K^{trans} differences greater than 0.1. For B2, this number was 7 out of 28, and for ETX, 4 out of 23. Furthermore, of these 19 "outliers," 13 corresponded to average K^{trans} values > 1 min^{-1}. There is thus better agreement between the AvgAll AIF and the time-point-specific AIFs at lower values of K^{trans}.

3.4 PK Parameter Variability Analysis

Figure 20 shows one slice of the K^{trans}, v_e, and iAUC60 maps for Patient 1at each time point. The goal of this final part of the study was to quantify the variability in each parameter between baselines and to compare it with the magnitude of treatment-induced changes.



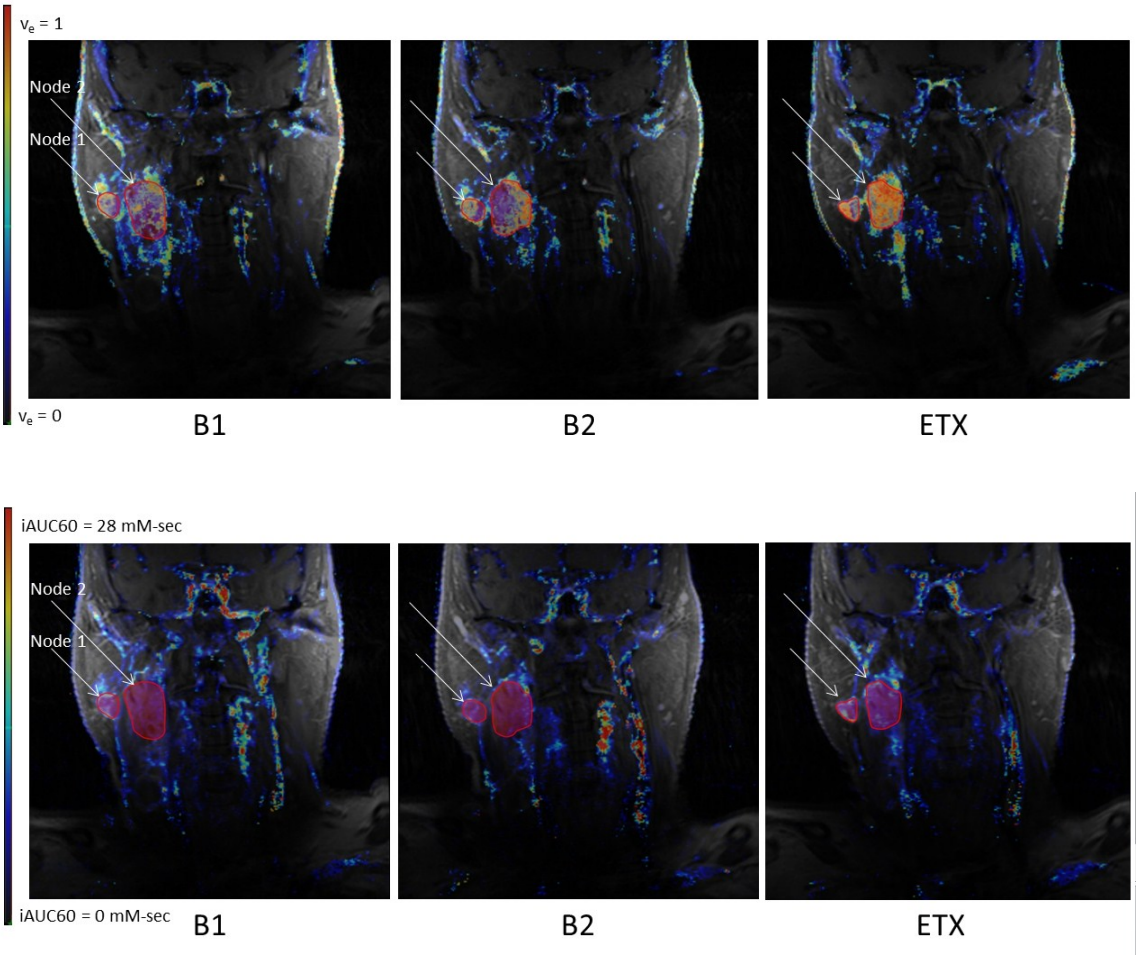


Figure 20: K^{trans}, v_e, and iAUC60 maps at each time point for a representative patient.

Table 6 shows the p-values obtained from the Wilcoxon signed-rank tests comparing B1 to B2, ETX to B2 and ETX to AvgB12.

Table 6: p-values from Wilcoxon tests assessing significance of intrinsic variability (B1/B2) and treatment-induced variability (ETX/B2 and ETX/AvgB12).

		B1/B2 p-value	ETX/B2 p-value	ETX/AvgB12 p-value
Nodes	**K^{trans}**	0.2842	**0.0166**	0.1726
	v_e	0.6397	**0.0004**	**0.0002**
	iAUC60	0.0737	**0.0013**	**0.0032**
Primaries	**K^{trans}**	**0.0186**	1.0000	0.2500
	v_e	**0.0068**	0.1475	0.6377
	iAUC60	**0.0061**	0.1099	0.9097

Based on this table alone, it would seem that nodes consistently showed significant treatment-induced changes, and primaries consistently showed significant intrinsic variability. However, the two rightmost columns in this table cannot be used to make judgments about statistical significance, because they do not take into account the observed variability between the two baselines. It is therefore necessary to find better methods to assess statistical significance of treatment-induced changes.

3.4.1 Nodes

Figure 21 shows the distributions of K^{trans}, v_e, and iAUC60 for nodes at different time points.

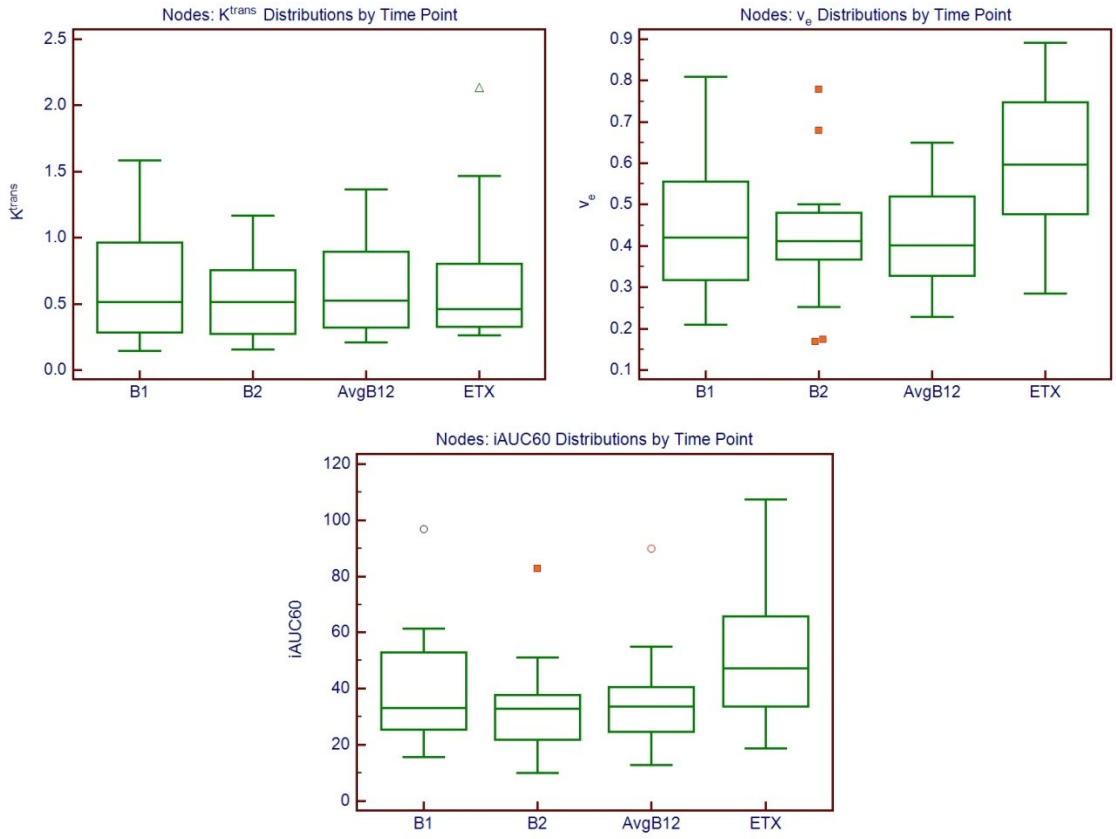


Figure 21: Box-and-whisker plots showing K^{trans}, v_e, and iAUC60 distributions for nodes. (The discrete points represent outliers.)

To more quantitatively contrast intrinsic variability (B1–B2) with treatment-induced variability (ETX–B2 or ETX–AvgB12), Bland-Altman plots were generated to show the magnitude of the intrinsic variability of each parameter (Figures 22-24).

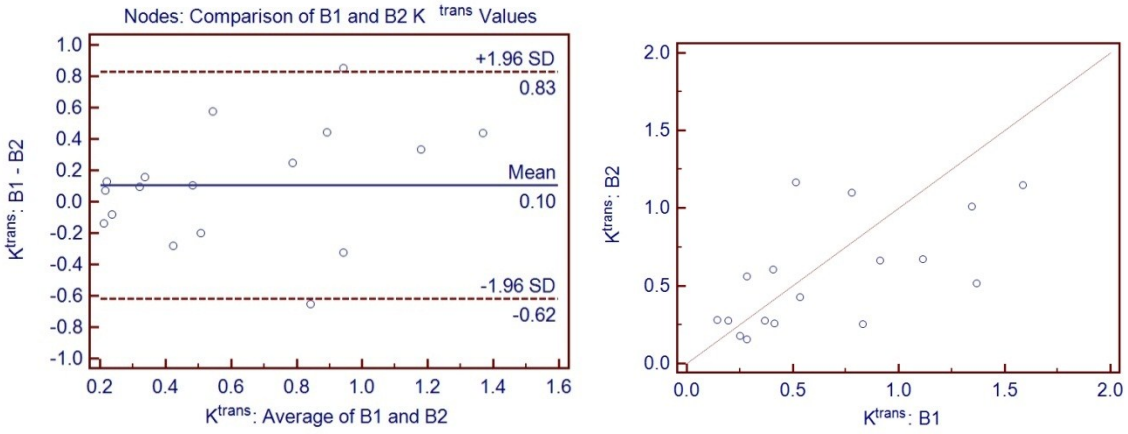


Figure 22: Comparison of B1 and B2 K^{trans} values for nodes: Bland-Altman plot (left) and regression plot with line of equality (right).

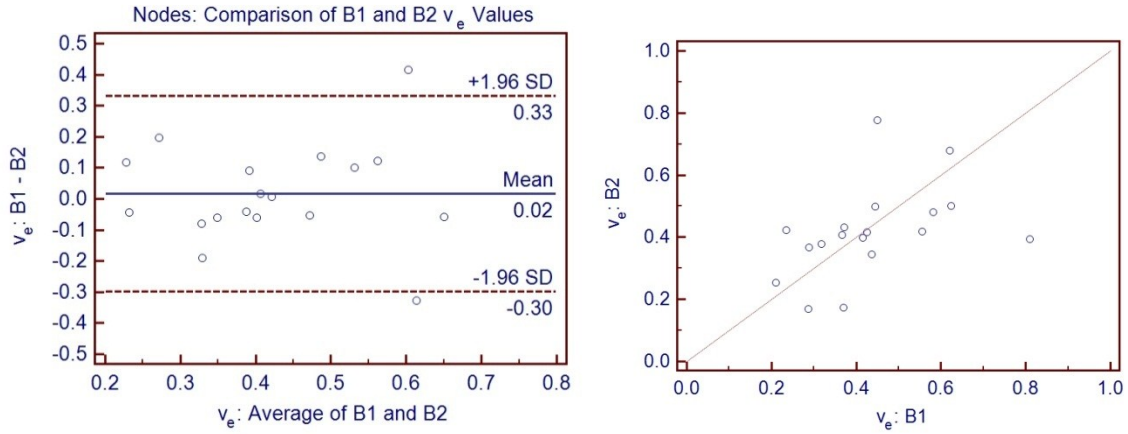


Figure 23: Comparison of B1 and B2 v_e values for nodes: Bland-Altman plot (left) and regression plot with line of equality (right).

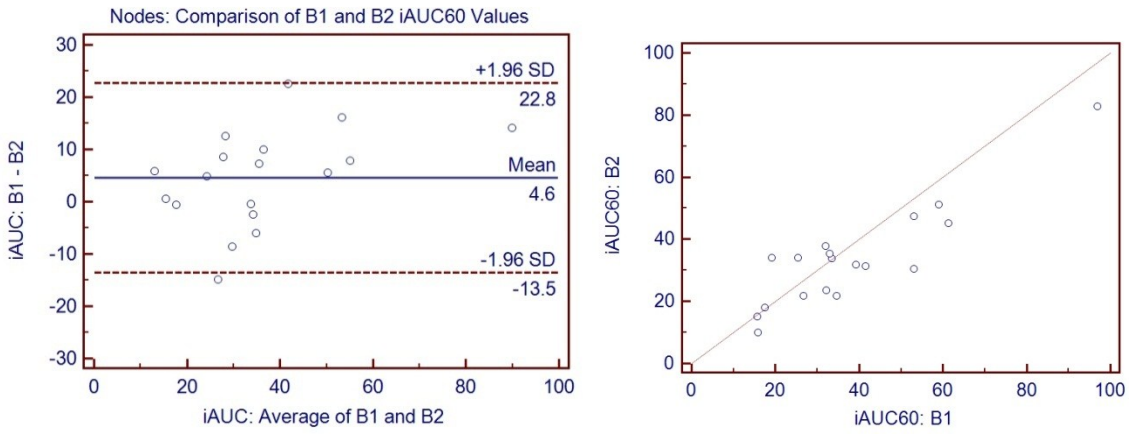


Figure 24: Comparison of B1 and B2 iAUC60 values for nodes: Bland-Altman plot (left) and regression plot with line of equality (right).

The high intrinsic variability of all three parameters is evident in Figures 22-24. In order for treatment-induced changes in these parameters to be considered statistically significant, an ROI's ETX–B2 or ETX–AvgB12 value for a given parameter should be greater than the coefficient of repeatability for that parameter. Table 7 lists the number of times this occurs.

Table 7: Coefficients of repeatability for K^{trans}, v_e, and iAUC60 for nodes, along with the number of ROIs for which the treatment-induced variability exceeded the CR.

	K^{trans} (min^{-1})	v_e	iAUC60 (mM-sec)
CR (from Bland-Altman plot)	0.725	0.315	18.15
# of ROIs whose |ETX–B2| value exceeds CR	3 of 14	3 of 17	7 of 17
# of ROIs whose |ETX–AvgB12| value exceeds CR	3 of 14	4 of 17	9 of 17

There were 9 ROIs that showed significant ETX–B2 treatment-induced changes. Of these, 5 ROIs showed a significant change in just 1 parameter, while 4 showed

significant changes in 2 parameters. Similarly, there were 10 ROIs that showed significant ETX–AvgB12 treatment-induced changes. Of these, 4 showed a change in just 1 parameter, while 6 showed changes in 2 parameters.

3.4.2 Primaries

Figure 25 shows the distributions of K^{trans}, v_e, and iAUC60 for primaries at different time points.

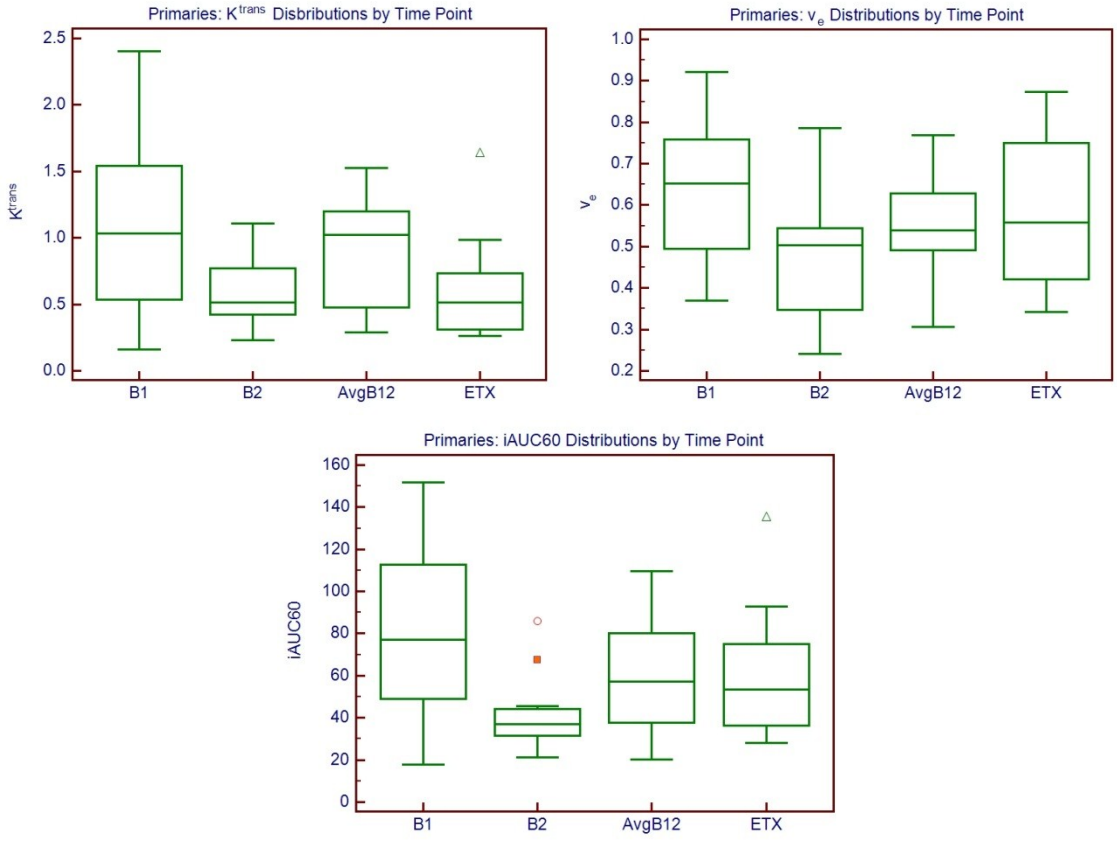


Figure 25: Box-and-whisker plots showing K^{trans}, v_e, and iAUC60 distributions for primaries. (The discrete points represent outliers.)

To more quantitatively compare the intrinsic variability of primary PK parameters with their treatment-induced variability, Bland-Altman plots were generated to show the magnitude of the intrinsic variability of each parameter (Figures 26-28). Regression lines (and their 95% confidence intervals) are drawn on the Bland-Altman plots where there is a clear increase in variability with increasing parameter values.

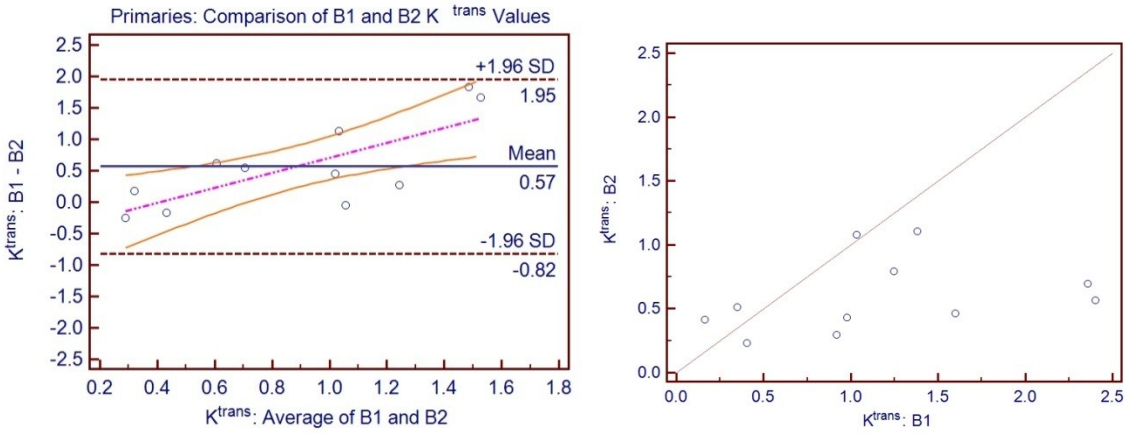


Figure 26: Comparison of B1 and B2 K^{trans} values for primaries: Bland-Altman plot (left) and regression plot with line of equality (right).

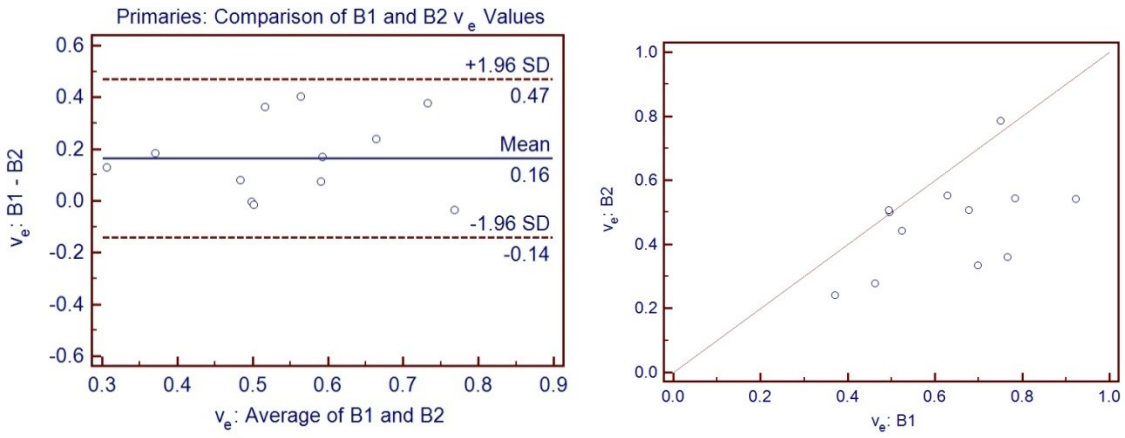


Figure 27: Comparison of B1 and B2 v_e values for primaries: Bland-Altman plot (left) and regression plot with line of equality (right).

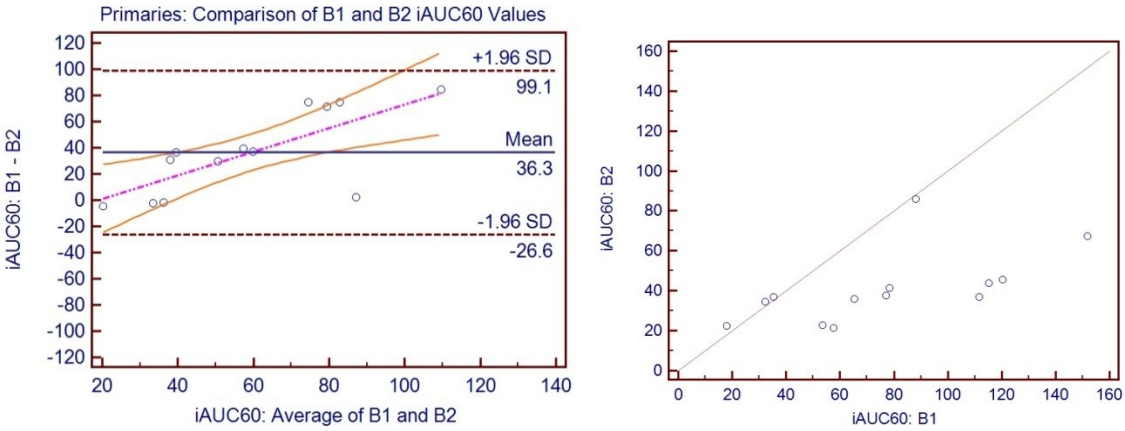


Figure 28: Comparison of B1 and B2 iAUC60 values for primaries: Bland-Altman plot (left) and regression plot with line of equality (right).

Table 8 compares the intrinsic variability observed in Figures 26-28 with treatment-induced changes in each parameter for primaries.

Table 8: Coefficients of repeatability for K^{trans}, v_e, and iAUC60 for primaries, along with the number of ROIs for which the treatment-induced variability exceeded the CR.

	K^{trans} (min^{-1})	v_e	iAUC60 (mM-sec)
CR (from Bland-Altman plot)	1.385	0.305	62.85
# of ROIs whose |ETX–B2| value exceeds CR	0 of 9	1 of 11	2 of 12
# of ROIs whose |ETX–AvgB12| value exceeds CR	0 of 9	2 of 11	1 of 12

There were 2 ROIs that showed significant ETX–B2 treatment-induced changes. Of these, 1 ROI showed a significant change in just 1 parameter, and 1 showed significant changes in 2 parameters. There were also 2 ROIs that showed significant ETX–AvgB12 treatment-induced changes. Of these, 1 showed a change in just 1 parameter, and 1 showed changes in 2 parameters.

4. Discussion

DCE-MRI is playing an increasingly important role in the diagnosis and staging of cancers, the assessment of treatment efficacy, and even the prediction of treatment response [29-32]. It has also been shown to be a valuable tool for each of these tasks for head and neck cancers in particular [3, 33, 34]. Our previous studies have shown the necessity of voxel-based T_{10} calculations for quantitative PK analysis [7] and identified the optimal methods for delineating primary and node ROIs [24]. The end goal of this particular study was to compare the intrinsic variability of three MR-extracted parameters (K^{trans}, v_e, and iAUC60) with their treatment-induced variability to determine the feasibility of using these parameters to monitor early treatment response. However, before studying variability, considerable effort went into studying the accuracy of T_{10} maps and AIFs, since the final PK parameters depend heavily on each [22, 35-38]. For example, Guo et al. [37], using a temporal resolution of 13 s in patients with osteosarcoma, found that the magnitudes of K^{trans}, v_e, and iAUC60 were highly dependent on T_{10} measurement, and that flip angle inaccuracies and imperfect slice profiles cause large uncertainties in T_{10} calculation via the VFA method. Di Giovanni et al. [36], using a temporal resolution of 10 s in the breast, found that errors in flip angle generation and T_{10} calculation caused K^{trans} and v_e errors of up to 88% and 73%, respectively.

While checking the T_{10} values produced by a previously determined flip angle combination (10°-45°) [7], we made several unexpectedly high measurements of the T_{10} value of blood. This prompted a study of the accuracy of the flip angles being produced by our MR scanner.

In section 3.1, we observed that the highest flip angle inaccuracies occurred at low flip angles and low T_1s. There is thus good justification for not using flip angles of 2° and 5° when calculating T_{10} with the VFA method on this particular MR scanner. Furthermore, it should be noted that the T_{10} values of blood, muscle, fat, and head and neck tumors/nodes are approximately 1.2–1.8 s [20, 21], 0.71–1.16 s [39, 40], 0.26–0.46 s [40, 41], and 0.800 s [7], respectively – much larger than 81 ms, where the highest T_1 errors occur.

The flip angle maps of the head and neck region of the volunteer showed the drastic effects of inhomogeneities on flip angle accuracy. For example, the 20° flip angle was consistently observed to be accurate in the phantom studies, but in the volunteer map for that angle, large regions of 30°–40° flip angles are easily distinguishable. Another factor that likely contributed to flip angle inaccuracy in this study was the short TRs used (5.8–6.6 ms). Very short TRs may hinder the ability of the spoiler gradient to completely destroy transverse magnetization between successive RF excitation pulses [42, 43]. Several papers have suggested that flip angle correction is indispensable when using the VFA method [17, 42, 44, 45]. A full investigation of the causes of flip angle

inaccuracies and the needed corrections in head and neck subjects was not possible within the time frame of this project. Therefore, to minimize T_{10} errors as much as possible, we referred back to a previous study performed on this scanner [7] which determined that the optimal flip angle combination when using the VFA method was 10°-45°. It was only at this flip angle combination that the measured T_{10} values for muscle, fat, and blood agreed with those found in literature. For several patients in this study, we verified that the muscle, fat, and blood T_{10} values obtained using 10°-45° were still consistent with literature values. We also tried using different flip angle combinations, as well as combinations of three or more flip angles, but the 10°-45° pair consistently gave the best results. The original high measurements of the T_{10} of blood that prompted the flip angle study remain unexplained, but isolated.

After our study of flip angle accuracy, we moved on to the calculation of a population-averaged AIF. It is assumed throughout this study that the use of a population-averaged AIF does not give statistically different results than the use of individual AIFs. This assumption is supported by previous studies in the abdomen [18], breast [25], prostate [46], osteosarcomas [47], and neck nodal metastases [8]. Particularly, Shukla-Dave et al. [8] determined a population-averaged AIF for head and neck based on data from 11 patients. Their study had better temporal resolution (3.75-7.5 s), but only nodes were imaged, only one carotid was analyzed per patient, and only one time point was considered. Furthermore, their study was only based on 2D

acquisitions, and they averaged plasma parameters rather than concentration curves to obtain their final AIF coefficients. The AvgAll AIF proposed by this study sacrifices temporal resolution for an increased field of view. It also considers both carotids and multiple time points. The increased field of view allowed for PK analysis of both primaries and nodes. It will also allow for future analysis of treatment-induced changes in parotid function. Given the limitations of the Shukla-Dave study and the fundamental differences between their study and ours, we did not compare the effects of interchanging these two AIFs.

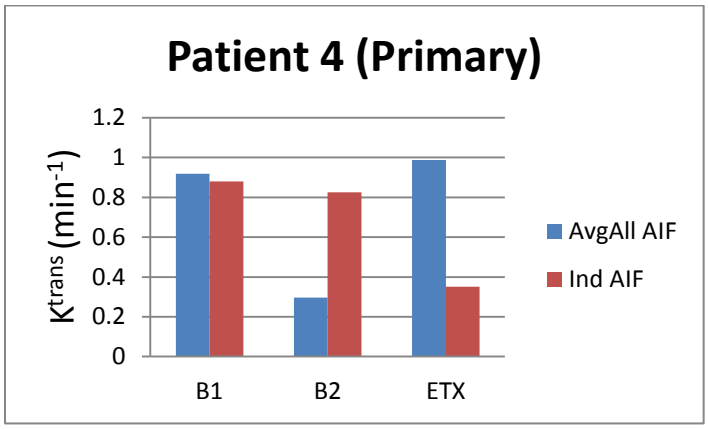
It is important to note that the coefficients of the population-averaged AIF calculated in this study should not be used with other MR scanners. Pabst et al. [48] showed that even under identical measurement conditions, different scanners will measure different signal enhancements. A calibration measurement is needed if DCE-MRI data are to be compared across scanners.

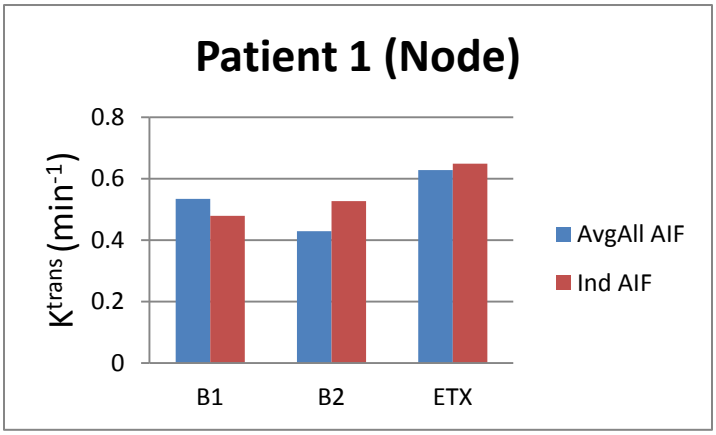
Side-specific AIFs were determined to be unnecessary based on the low limits of agreement obtained from the Bland-Altman plots of AvgAll vs. AvgLeft and AvgAll vs. AvgRight. The necessity of a time-point-specific AIF was not as clear. In most cases, it gave PK parameters very similar to those obtained with the AvgAll AIF. However, in some cases, particularly when K_{trans} is greater than 1 min^{-1}, the use of a time-point-specific AIF may be necessary.

The variability analysis, which used the AvgAll AIF, revealed high intrinsic variability in K^{trans}, v_e, and iAUC60 in both nodes and primaries, with the largest variability observed in primaries. It was rare for treatment-induced changes in PK parameters to exceed the magnitude of this variability. This may preclude the use of these MR-extracted, ROI-averaged parameters for assessment of early treatment response in head and neck cancer. It is important to note that it is not sufficient to perform the Wilcoxon signed-rank test on the baseline data vs. the early treatment data. This does not take into account the baseline variability, and it will give inaccurate, oversimplified results.

As a final addendum to this study, following the suggestion of Port et al. [38] that individual AIFs be used when accurate tissue parameters are needed, we briefly investigated how the use of individual AIFs would affect the magnitude of intrinsic variability in K^{trans}. For a subset of 8 patients, PK analysis was performed using side-specific and time-point-specific AIFs as possible. Some individual plasma concentration curves did not show the expected degree of enhancement, while others, when fitted to a bi-exponential, gave non-physiological plasma parameters (e.g. $a_{1,2}$ on the order of 10^3 – 10^{10} kg/liter). When a side-specific AIF was unavailable, the individual contralateral AIF was used instead. When a time-point-specific AIF was unavailable, the individual AIF corresponding to the next closest time point was used. The effects of using individual AIFs varied: in some cases, the intrinsic variability decreased, while in other cases, it did

not change significantly, and in other cases, it increased. Figure 29 shows examples of each of these cases.





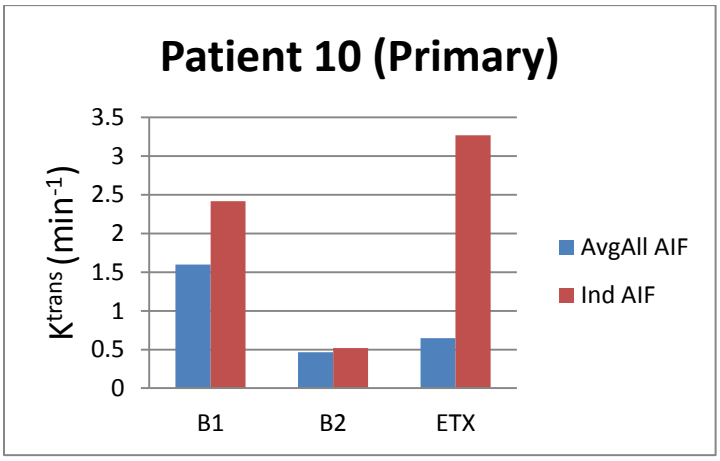


Figure 29: Effects of using individual AIFs: decrease in intrinsic variability (top), insignificant change in variability (middle), or increase in variability (bottom).

The results from all 8 patients are combined in the box-and-whisker plot shown in Figure 30. The blue boxes represent the K^{trans} distributions at each time point using individual AIFs. It is clear that based on this subset of the data, the use of individual AIFs does not reduce the magnitude of intrinsic variability.

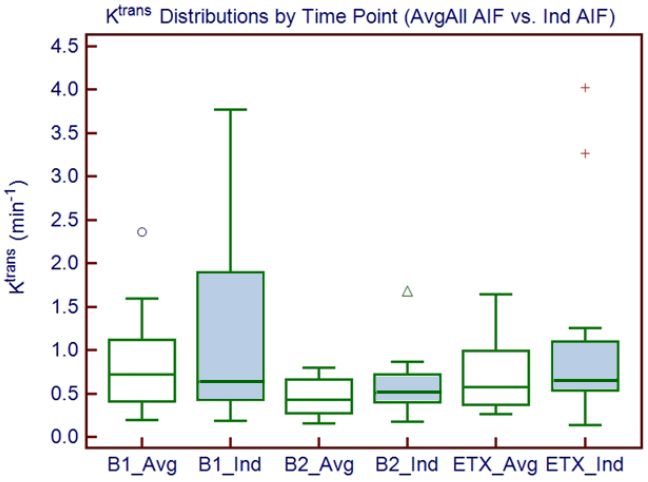


Figure 30: K^{trans} distributions by time point using both AvgAll AIF and individual AIFs for a subset of 8 patients.

There were several limitations to this study. First, we had no means of correcting for flip angle inaccuracies, and we only calculated T_{10} with two flip angles (although T_{10} accuracy did not improve with the use of more flip angles). Second, ROI delineation – both in the carotids and in the primaries/nodes – was likely complicated by partial volume effects because of the 10 mm slice thickness used. Third, the temporal resolution of our dynamic study (10 s) limited our ability to measure the true peak of the arterial plasma concentration curve. Fourth, our software limited us to use of the Tofts model for PK analysis Other models exist which may provide benefits over the Tofts model. For example, the Brix model [49] does not require T_{10} maps. Finally, by deciding to work only with median PK parameter values, we discarded potentially useful information about the distribution of those parameter values within a primary or node. Alternatively, it may be of interest to average the MR signal over a lesion first, and then perform PK analysis on the average signal intensities, rather than averaging the final PK parameters themselves.

Future work will include a more detailed analysis of the variability observed in each voxel, rather than in ROI-averaged values. Furthermore, this DCE-MRI variability data will be combined with diffusion-weighted MRI and FDG-PET/CT variability data. Which of these imaging-based parameters correlate best with clinical outcome? Can any of them be used to predict treatment response? Are MRI studies and FDG-PET/CT studies redundant or complementary?

5. Conclusion

A population-averaged AIF for head and neck was generated from 93 carotid ROIs with a temporal resolution of 10 s. It accounts for differences in right vs. left carotids, intrinsic AIF fluctuations, and treatment-induced AIF changes. This average AIF may be used with both left- and right-sided malignancies both before and after treatment. A time-point-specific AIF may improve PK parameter accuracy in cases when K^{trans} is greater than 1 min^{-1}. Using this population-averaged AIF, large intrinsic fluctuations were observed in ROI-averaged values of K^{trans}, v_e, and iAUC60. The likelihood of treatment-induced changes exceeding the observed magnitude of intrinsic variability was very low, and lower in primaries than in nodes. Because of this high intrinsic variability, these PK parameters may be poor evaluators of early treatment response in head and neck cancer.

Appendix A

The MR imaging parameters for the phantom and volunteer scans are as follows:

Low-T_1 Phantom	FA Accuracy Sequence	Gradient Echo α = 5, 10, 20, 30, 40, 60, 80, 120° TE = 5 ms TR = 1 sec Axial 5 slices, 10 mm thick 256x256 matrix 24x24 cm^2 FOV Head coil (CP)
	VTR Sequence	Gradient Echo α = 90° TE = 5 ms TR = 50, 100, 200, 600, 1000, 3000 ms Axial 5 slices, 10 mm thick 256x256 matrix 24x24 cm^2 FOV Head coil (CP)
	VFA Sequence	Gradient Echo α = 5, 10, 20, 30, 40, 60, 80, 120° TE = 5 ms TR = 17 ms Axial 5 slices, 10 mm thick 256x256 matrix 24x24 cm^2 FOV Head coil (CP)
ADNI Phantom	FA Accuracy Sequence	Gradient Echo α = 5, 10, 20, 30, 40, 60, 80, 120° TE = 3.8 ms TR = 3000 ms Plane through center of 4 colored spheres 1 slice, 3 mm thick 256x256 matrix 24x24 cm^2 FOV Head coil (CP)

	VTR Sequence	Spin Echo α = 90° TE = 20 ms TR = 100, 200, 300, 500, 1000, 2000, 3000 ms Plane through center of 4 colored spheres 1 slice, 3 mm thick 256x256 matrix 24x24 cm^2 FOV Head coil (CP)
	VFA Sequence	Gradient Echo α = 5, 10, 20, 30, 40, 60, 80, 120° TE = 3.8 ms TR = 17 ms Plane through center of 4 colored spheres 1 slice, 3 mm thick 256x256 matrix 24x24 cm^2 FOV Head coil (CP)
High-T_1 Phantom	FA Accuracy Sequence	Gradient Echo α = 5, 10, 20, 30, 40, 60, 80, 120° TE = 3.5 ms TR = 4500 ms Axial 1 slice, 3 mm thick 256x256 matrix 16x16 cm^2 FOV Head coil (CP)
	VTR Sequence	Spin Echo α = 90° TE = 20 ms TR = 102, 202, 302, 502, 1002, 2002, 3002 ms Axial 1 slice, 3 mm thick 256x256 matrix 16x16 cm^2 FOV Head coil (CP)

	VFA Sequence	Gradient Echo α = 5, 10, 20, 30, 40, 60, 80, 120° TE = 3.5 ms TR = 17 ms Axial 1 slice, 3 mm thick 256x256 matrix 16x16 cm^2 FOV Head coil (CP)
Volunteer	FA Accuracy Sequence	Gradient Echo α = 5, 10, 20, 40° TE = 2.9 ms TR = 1000 ms Coronal Slices covering LT and RT carotid arteries, 10 mm thick, 12.5 mm spacing 256x256 matrix 24x24 cm^2 FOV 6" Flex PA coil
	VFA Sequence	Gradient Echo α = 2, 5, 10, 15, 20, 30, 45° TE = 2 ms TR = 6.6 ms Coronal Slices covering LT and RT carotid arteries, 10 mm thick, 12.5 mm spacing 256x256 matrix 24x24 cm^2 FOV 6" Flex PA coil

Appendix B

iCAD settings are as follows:

Protocol	Single Scan Duration	10.00 s
	First Delay	0.00 s
	Injection Start	25.00 s
	Injection Duration	5.00 s
	Second Delay	0.00 s
	Number of Scans	32*
	CA-Arrival Delay	Auto**
	k-Space Center	0.50
	Repetition Time (in msec)	6.40 ms
	Flip Angle (in degrees)	60.00°
	Magnet Strength (Tesla)	1.50 T
	Dose (mmol / kg body weight)	0.10 mmol/kg
	Threshold 1(%)	10
	Fit Coefficient	0.65
	T10 (in seconds)	0.80 s
	Peak Time (in seconds)	50.00 s
	Permeability Cap (min^{-1})	5 min^{-1}
	Fit Options	3 parameter fit (K-trans, k-ep, v-p)
	Contrast Agent	Default
	Do neighborhood similarity matching	Yes
Processing Filters	Noise Threshold	20
	Smoothing Filters	No
	Scaling	No
	Philips Rescaling	No
	Interpolation	No
T10 Mapping	Do T10 Mapping	Yes
	Skip fTP processing for pixels which don't pass T10 mapping	Yes
	Sequence Type	Gradient Echo
	Noise Threshold	0
	Fit coefficient (only used with more than two series)	0.00
Additional Series	Generate Quick iAUC series	Yes
	Initial time interval (in seconds)	60 s

*Usually 32, but varied between 19 and 32

**Had to be set manually to 45 s for some scans (P7 ETX, P11 B2, P12 B1, P17 B1, P19 B2)

References

1. Jemal, A., et al., *Cancer statistics, 2007.* CA Cancer J Clin, 2007. **57**(1): p. 43-66.

2. Ha, P.K., et al., *The role of positron emission tomography and computed tomography fusion in the management of early-stage and advanced-stage primary head and neck squamous cell carcinoma.* Arch Otolaryngol Head Neck Surg, 2006. **132**(1): p. 12-6.

3. Hoskin, P.J., et al., *Dynamic contrast enhanced magnetic resonance scanning as a predictor of response to accelerated radiotherapy for advanced head and neck cancer.* Br J Radiol, 1999. **72**(863): p. 1093-8.

4. Murakami, R., et al., *Impact of FDG-PET/CT imaging on nodal staging for head-and-neck squamous cell carcinoma.* Int J Radiat Oncol Biol Phys, 2007. **68**(2): p. 377-82.

5. Tomura, N., et al., *Dynamic contrast-enhanced magnetic resonance imaging in radiotherapeutic efficacy in the head and neck tumors.* Am J Otolaryngol, 2005. **26**(3): p. 163-7.

6. Wang, J., et al., *Head and neck lesions: characterization with diffusion-weighted echo-planar MR imaging.* Radiology, 2001. **220**(3): p. 621-30.

7. Craciunescu, O., et al., *Dynamic contrast enhanced-MRI in head and neck cancer patients: variability of the precontrast longitudinal relaxation time (T10).* Med Phys, 2010. **37**(6): p. 2683-92.

8. Shukla-Dave, A., et al., *Average arterial input function for quantitative dynamic contrast enhanced magnetic resonance imaging of neck nodal metastases.* BMC Med Phys, 2009. **9**: p. 4.

9. Tofts, P.S., *Modeling tracer kinetics in dynamic Gd-DTPA MR imaging.* J Magn Reson Imaging, 1997. **7**(1): p. 91-101.

10. Purves, W.K., et al., *Life: The Science of Biology*. 7 ed 2003, Sunderland, Mass.: Sinauer Associates and W.H. Freeman.

11. Tofts, P.S. and A.G. Kermode, *Measurement of the blood-brain barrier permeability and leakage space using dynamic MR imaging. 1. Fundamental concepts.* Magn Reson Med, 1991. **17**(2): p. 357-67.

12. Parker, G. and D. Buckley, *Tracer Kinetic Modelling for T1-Weighted DCE-MRI Dynamic Contrast-Enhanced Magnetic Resonance Imaging in Oncology*, A. Jackson, D. Buckley, and G. Parker, Editors. 2005, Springer Berlin Heidelberg. p. 81-92.

13. Evelhoch, J.L., *Key factors in the acquisition of contrast kinetic data for oncology.* J Magn Reson Imaging, 1999. **10**(3): p. 254-9.

14. Haase, A., *Snapshot FLASH MRI. Applications to T1, T2, and chemical-shift imaging.* Magn Reson Med, 1990. **13**(1): p. 77-89.

15. Rohrer, M., et al., *Comparison of magnetic properties of MRI contrast media solutions at different magnetic field strengths.* Invest Radiol, 2005. **40**(11): p. 715-24.

16. Clare, S., M. Alecci, and P. Jezzard, *Compensating for B(1) inhomogeneity using active transmit power modulation.* Magn Reson Imaging, 2001. **19**(10): p. 1349-52.

17. Stollberger, R. and P. Wach, *Imaging of the active B1 field in vivo.* Magn Reson Med, 1996. **35**(2): p. 246-51.

18. Parker, G.J., et al., *Experimentally-derived functional form for a population-averaged high-temporal-resolution arterial input function for dynamic contrast-enhanced MRI.* Magn Reson Med, 2006. **56**(5): p. 993-1000.

19. Cutajar, M., et al., *The importance of AIF ROI selection in DCE-MRI renography: reproducibility and variability of renal perfusion and filtration.* Eur J Radiol, 2010. **74**(3): p. e154-60.

20. Anzai, Y., et al., *MR angiography with an ultrasmall superparamagnetic iron oxide blood pool agent.* J Magn Reson Imaging, 1997. **7**(1): p. 209-14.

21. Sharma, P., et al., *Effect of Gd-DTPA-BMA on blood and myocardial T1 at 1.5T and 3T in humans.* J Magn Reson Imaging, 2006. **23**(3): p. 323-30.

22. McGrath, D.M., et al., *Comparison of model-based arterial input functions for dynamic contrast-enhanced MRI in tumor bearing rats.* Magn Reson Med, 2009. **61**(5): p. 1173-84.

23. Schmitt, P., et al., *Quantitative tissue perfusion measurements in head and neck carcinoma patients before and during radiation therapy with a non-invasive MR imaging spin-labeling technique.* Radiother Oncol, 2003. **67**(1): p. 27-34.

24. Craciunescu, O.I., et al., *Dynamic contrast-enhanced MRI in head-and-neck cancer: the impact of region of interest selection on the intra- and interpatient variability of pharmacokinetic parameters.* Int J Radiat Oncol Biol Phys, 2012. **82**(3): p. e345-50.

25. Li, X., et al., *A novel AIF tracking method and comparison of DCE-MRI parameters using individual and population-based AIFs in human breast cancer.* Phys Med Biol, 2011. **56**(17): p. 5753-69.

26. *Wilcoxon test (paired samples)*. 2012 03/21/12]; Available from: http://www.medcalc.org/manual/wilcoxon.php.

27. Bland, J.M. and D.G. Altman, *Statistical methods for assessing agreement between two methods of clinical measurement.* Lancet, 1986. **1**(8476): p. 307-10.

28. *Box-and-whisker plot*. 2012 03/21/12]; Available from: http://www.medcalc.org/manual/box_and_whisker_plot.php.

29. Collins, D.J. and A.R. Padhani, *Dynamic magnetic resonance imaging of tumor perfusion. Approaches and biomedical challenges.* IEEE Eng Med Biol Mag, 2004. **23**(5): p. 65-83.

30. George, M.L., et al., *Non-invasive methods of assessing angiogenesis and their value in predicting response to treatment in colorectal cancer.* Br J Surg, 2001. **88**(12): p. 1628-36.

31. Padhani, A.R., et al., *Prediction of clinicopathologic response of breast cancer to primary chemotherapy at contrast-enhanced MR imaging: initial clinical results.* Radiology, 2006. **239**(2): p. 361-74.

32. Taylor, J.S., et al., *MR imaging of tumor microcirculation: promise for the new millennium.* J Magn Reson Imaging, 1999. **10**(6): p. 903-7.

33. Konouchi, H., et al., *Evaluation of tumor proliferation using dynamic contrast enhanced-MRI of oral cavity and oropharyngeal squamous cell carcinoma.* Oral Oncol, 2003. **39**(3): p. 290-5.

34. Shah, G.V., et al., *Dynamic contrast-enhanced MR imaging.* Top Magn Reson Imaging, 2004. **15**(2): p. 71-7.

35. Cheng, H.L., *Investigation and optimization of parameter accuracy in dynamic contrast-enhanced MRI.* J Magn Reson Imaging, 2008. **28**(3): p. 736-43.

36. Di Giovanni, P., et al., *The accuracy of pharmacokinetic parameter measurement in DCE-MRI of the breast at 3 T.* Phys Med Biol, 2010. **55**(1): p. 121-32.

37. Guo, J.Y., et al., *Dynamic contrast-enhanced magnetic resonance imaging parameters independent of baseline T10 values.* Magn Reson Imaging, 2009. **27**(9): p. 1208-15.

38. Port, R.E., M.V. Knopp, and G. Brix, *Dynamic contrast-enhanced MRI using Gd-DTPA: interindividual variability of the arterial input function and consequences for the assessment of kinetics in tumors.* Magn Reson Med, 2001. **45**(6): p. 1030-8.

39. Houmard, J.A., R. Smith, and G.L. Jendrasiak, *Relationship between MRI relaxation time and muscle fiber composition.* J Appl Physiol, 1995. **78**(3): p. 807-9.

40. Poon, C.S., et al., *Quantitative magnetic resonance imaging parameters and their relationship to mammographic pattern.* J Natl Cancer Inst, 1992. **84**(10): p. 777-81.

41. Bushberg, J., et al., *The Essential Physics of Medical Imaging*. 2 ed 2002, Philadelphia: Lippincott Williams & Wilkins.

42. Preibisch, C. and R. Deichmann, *Influence of RF spoiling on the stability and accuracy of T1 mapping based on spoiled FLASH with varying flip angles.* Magn Reson Med, 2009. **61**(1): p. 125-35.

43. Yarnykh, V.L., *Optimal radiofrequency and gradient spoiling for improved accuracy of T1 and B1 measurements using fast steady-state techniques.* Magn Reson Med, 2010. **63**(6): p. 1610-26.

44. Cheng, H.L., *T1 measurement of flowing blood and arterial input function determination for quantitative 3D T1-weighted DCE-MRI.* J Magn Reson Imaging, 2007. **25**(5): p. 1073-8.

45. Wang, J., et al., *T1 measurements incorporating flip angle calibration and correction in vivo.* J Magn Reson, 2006. **182**(2): p. 283-92.

46. Meng, R., et al., *Comparison between population average and experimentally measured arterial input function in predicting biopsy results in prostate cancer.* Acad Radiol, 2010. **17**(4): p. 520-5.

47. Wang, Y., et al., *Feasibility of using limited-population-based arterial input function for pharmacokinetic modeling of osteosarcoma dynamic contrast-enhanced MRI data.* Magn Reson Med, 2008. **59**(5): p. 1183-9.

48. Pabst, T., et al., *Understanding why contrast enhancement in dynamic MRI is not reproducible: illustration with a simple phantom.* Breast J, 2001. **7**(3): p. 166-70.

49. Brix, G., et al., *Pharmacokinetic parameters in CNS Gd-DTPA enhanced MR imaging.* J Comput Assist Tomogr, 1991. **15**(4): p. 621-8.