

# University of Cincinnati

Date: 3/22/2023

**I. Neelakshi Chatterjee, hereby submit this original work as part of the requirements for the degree of Master of Science in Statistics.**

It is entitled:

**Regression Analysis(Bayesian and Simple linear) of Pulmonary <sup>129</sup>Xe ADC on Voxel MRI Data: A Comparison of CF Patients and Healthy Controls AND Optimizing Under sampled Voxel MRI Data for Retaining T2\* Information: Finding the Point of Cessation.**

Student's name: **Neelakshi Chatterjee**

This work and its defense approved by:

Committee chair: Dan Ralescu, Ph.D.

Committee member: Zackary Cleveland, Ph.D.

Committee member: Md Monir Hossain, Ph.D.

Committee member: Anca Ralescu, Ph.D.



44458

**REGRESSION ANALYSIS(Bayesian and Simple  
linear) OF PULMONARY  $^{129}\text{Xe}$  ADC ON VOXEL  
MRI DATA: A COMPARISON OF CF PATIENTS  
and HEALTHY CONTROLS.**

**AND**

**OPTIMIZING UNDER SAMPLED VOXEL MRI  
DATA FOR RETAINING  $T_2^*$  INFORMATION:  
FINDING the POINT OF CESSATION.**

Author: Neelakshi Chatterjee

Thesis for MS (Statistics)

Department of Mathematical Sciences

College of Arts and Science

University of Cincinnati

Spring 2023

## ABSTRACT:

**Project 1:** ADC(Apparent Diffusion Co-efficient), a biomarker helping in the diagnosis of tumors, is a measure based on the magnitude of diffusion calculated with voxel level DW MRI(Diffusion-weighted Magnetic Resonance Imaging). ADC is expressed as the exponent of the DW signals ( $S_i$ ) with a product component of b values (fixed for 4 levels).  $S_i = A_i * e(-b * ADC)$  - where  $A_i$  are initial signal values. Accounting the problem of limited data due to constraining on the repetition of total body MR measurements, Bayesian setup seemed appropriate. This has been previously done as a part of the project in the lab using Rice distribution(expressed as a mixture of chi sq and Poisson distribution) as likelihood and uniform prior - which I replicated using different parameters. Thereafter I investigated the effect of ADC on age on patients and controls as separate groups and also on an individual basis. This results in quite a contrast - showing high ADC estimates for controls as a group than CF patients but vice versa for some older patients. This was the motivation behind regression analysis in two setups -

- 1) Bayesian Regression and
- 2) Simple Linear Regression.

Both the regression equations show similar trends and give a strong understanding of the contracting results observed which is also in sync with the biological claim which suggests the ADC estimates are higher in older patients due to microstructural damage caused by years of infections and inflammation. Thus, leading to comparison of CF patients and healthy controls.

**Project 2:** Using a similar data structural setup which is voxel MRI data in the next project, the point of cessation was found for under sampling of MR images. Using radial UTE(ultra-echo-time) sequences to overcome some imaging limitations(rapid cardiac or respiratory movement, relatively low lung tissue density, etc.) while acquiring an MRI of the lung leads to long acquisition times. The norm often used in pulmonary imaging is to under sample to reduce the acquisition time. That motivated me to investigate to find the optimized level of under sampling where the information that is specifically  $T_2^*$  mapping remains unaffected. The data that we received

was accumulated in two different setups and the  $T_2^*$  has been computed. The analysis conducted leads to finding the point of cessation to retain the information in  $T_2^*$ . The two setups are described as follows:

1) One setup is the simulated data. The simulated data we have is the Phantom dataset. It has been generated at three different 3D isotropic matrix sizes:  $64^3$ ,  $96^3$  and  $128^3$ .

2) The other dataset used here is MRI on 14 real mice which has been generated at the isotropic matrix size of  $128^3$ .

Now in these two setups, different techniques such as one sample Hotelling  $t^2$  test and paired t-test have been used depending on the data to address our question concerned. Thus, we conclude about attaining the tolerable level of under-sampling has a nonsignificant effect on the parameter of interest that is  $T_2^*$ .



## **Acknowledgement:**

I would like to begin by extending my heartfelt gratitude to Prof. Dan Ralescu for being the chair of the committee on a very short notice and helping and guiding me through the research and process of thesis for masters degree in Statistics.

I would like to acknowledge my mentor at Cincinnati Children's Hospital Medical Center , Prof. Monir Hossain for guiding me throughout the research and helping me in the process of finishing the thesis.

I would like to thank Prof Zackery Cleveland for his continued support in the projects and him experienced bio medical input in the projects leading to their success.

I would like to thank Prof Anca Ralescu for her time and help in writing the thesis down in such a short notice.

I would also like to thank Dr. Abdullah S Bdaiwi, a post doctoral fellow in Prof Cleveland's lab , for his help in understanding the data.

I would also like to thank Ian R. Stecker for agreeing to share his data for analysis.

Again, I would like to thank you all the professors mentioned above for their time and co-operation in the process.

# Contents

1	Exordium :Project 1	1
2	Data Acquisition and Data Explanation:	3
3	Statistical Methods:	5
4	Experiment Setup:	9
5	Results:	11
6	Excursis and Future Work:	21
7	Exordium:Project 2	23
8	Data Acquisition:	24
9	Statistical Methods:	27
10	Experiment Setup:	32
11	Results:	40
12	Excursis and Future Work:	51
13	References:	55

# List of Figures

1.1	Visualization of MRI of a control patient.It shows ADC estimates inside of lung measured in $cm^2/sec$ . . . . .	1
1.2	Visualization of MRI of a CF patient.It shows ADC estimates inside of lung measured in $cm^2/sec$ . . . . .	1
5.1	Histogram with density of ADC estimates(inside the lung area) of CF patients . . . . .	12
5.2	Scatter plot of ADC estimates(inside the lung area) of CF patients vs age . . . . .	13
5.3	Histogram with density of ADC estimates(inside the lung area) of Controls . . . . .	15
5.4	Scatter plot of ADC estimates(inside the lung area) of Controls vs age . . . . .	16
5.5	Trace plot of one patient - Checking for convergence of MCMC	16
5.6	Trace plot of another patient - Checking for convergence of MCMC . . . . .	17
11.1	Visualization of pvalue of Hotelling $t^2$ . . . . .	44
11.2	Visualization of pvalue of Paired t test . . . . .	49

# List of Tables

5.1	Table showing the individual ADC estimates for some older CF patients . . . . .	11
5.2	Table showing the individual ADC estimates for some older controls . . . . .	14
5.3	ANOVA table for Simple Linear Regression . . . . .	18
10.1	A table showing the number of means for each level of difference for Phantom Dataset . . . . .	34
10.2	A table showing the number of means for each level of difference for Vivo real mice data . . . . .	38
11.1	A table p-values of Hotelling $t^2$ , Paired t and ICC for 64 <sup>3</sup> . . .	42
11.2	A table p-values of Hotelling $t^2$ , Paired t and ICC for 96 <sup>3</sup> . . .	42
11.3	A table p-values of Hotelling $t^2$ , Paired t and ICC for 128 <sup>3</sup> . .	43
11.4	A table demonstrating p-values of Paired t for Mice 1 - 7 . . .	47
11.5	A table demonstrating p-values of Paired t for Mice 7 - 14 . .	47
11.6	A table showing ICC for Mice 1 - 7 . . . . .	48
11.7	A table showing ICC for Mice 7-14 . . . . .	48

# Chapter 1

## Exordium :Project 1

ADC(Apparent Diffusion Coefficient) is well validated measure of alveolar airspace size(Microstructural level) which is related to surface area. ADC is not only a validated measure of lung micro-structure but assesses alveolar-size changes due to normal aging and disease progression.

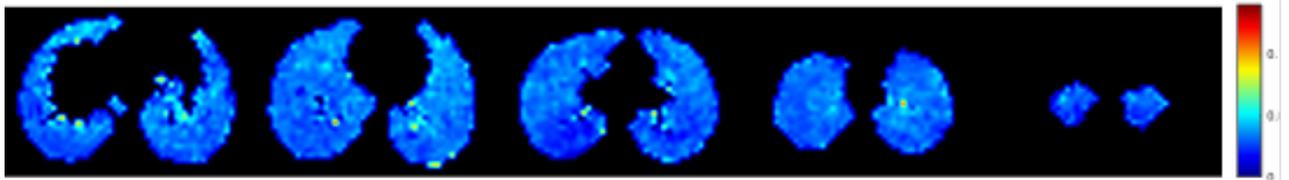


Figure 1.1: Visualization of MRI of a control patient.It shows ADC estimates inside of lung measured in  $cm^2/sec$ .

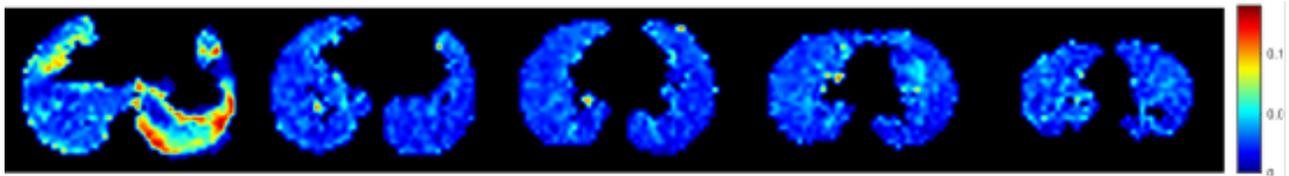


Figure 1.2: Visualization of MRI of a CF patient.It shows ADC estimates inside of lung measured in  $cm^2/sec$ .

We can see how the MRI of a CF patient looks distinctly different from a healthy person.

ADC is estimated from the Diffusion-weighted  $^{129}\text{Xe}$  MRI data assume - the imaging signals were attenuated according to the Stejskal-Tanner model (i.e., exponential decay).

$$S_1 = S_0 * e(-b_i * ADC)$$

Taking log on both sides we get the following form of the equation:

$$ADC = \frac{1}{b_i} * \ln \frac{S_0}{S_1}$$

is the mathematical expression of ADC where  $\frac{S_0}{S_1}$  signal intensity of diffusion weighted image at different  $b_i$  values(MR acquisition parameters) and  $S_0$  is the signal intensity at  $b = 0$  . The b values reflect the strength and timing of gradients used to generate diffusion weighted images. Now as explained in the abstract Bayesian set up was used for ADC estimates. The method for ADC estimates were carried out previously by Dr. Abdullah S Bdaiwi.I replicated it using different parameters and method. We will discuss in detail about the Bayesian set up used but the focal point of this project is to regress the ADC against age , and the dummy binary variable indicating the occurrence of disease and their interactions.

This regression has been carried out in two setups with the goal of comparison and attainment of more precise model.

1. First set up has been carried out in the simplest way possible - Getting the estimates from and using those to model a simple regression model with the variables mentioned above.
2. The next setup has been inspired from the Bayesian set up used for the ADC estimates. In this set up the a two stage Bayesian Model has been used to attain the regression model using the same Bayesian set up used for ADC estimates and then using those estimates and incorporating the same in a Bayesian regression set up with priors imposed on the regression coefficients.

The projects end on a comparison note of the two models with many prospects of future work.

## Chapter 2

# Data Acquisition and Data Explanation:

### Data Acquisition:

Data is collected in a Hyperpolarized Xenon MRI setup.

Previously Helium ( $^3\text{He}$ ) was used but since  $^{129}\text{Xe}$  is more affordable and abundantly available nowadays  $^{129}\text{Xe}$  is used more.

Hyper-polarized(Magnetized) (HP)  $^{129}\text{Xe}$ -MRI provides non-invasive structure(not requiring the introduction of instruments into the body) methods to quantify lung function. The  $^{129}\text{Xe}$  apparent diffusion coefficient (ADC) being a well validated measure of alveolar airspace size. DW-MRI images from 38 healthy-controls (age:  $18\pm 9.7$  year) and 39 CF patients (age:  $14.7\pm 7.9$  year) with 5 b-values [0, 6.25, 12.5, 18.75 and 25 s/cm<sup>2</sup>] were acquired in the Imaging Research Center at Cincinnati Children's Hospital Medical Center. Assuming - the imaging signals were attenuated according to the Stejskal-Tanner model (i.e., exponential decay), it is expected that the image noise follow the Rician distribution, which as expressed by a non-central  $\chi^2$  - Poisson-hierarchical model, was used to describe the data distribution. A uniform distribution was used for specifying the prior probability distribution of ADC.

## Data Explanation:

Let us look at the data type and structure. The files are from matlab software. Now let us see the structure of the data that we received: The data is given in voxel to voxel level along with b values. Thus we have a 4 D array where voxel has x, y and z(which reflects the number of slices in MRI) that is 3 dimension and the b values adding on to the 4th dimension.

$$4D = [\text{number in x axis , number in y axis , number of slices , b values}]$$

The other component of the dataset is the mask which is a binary variable that indicates if the voxel is inside the lung area or not. (This is a filter that we use in image analysis).

The age of each patient has been noted in the dataset.

b values which are fixed values are noted down in the dataset. It is a factor that reflects the strength and timing of the gradients used to generate diffusion-weighted images. The higher the b-value, the stronger the diffusion effects.

A dummy variable named CF has been created for indicating the cases and controls.

The sigma matrix has been calculated for each patient - the dimension of which differs from patient to patient.

nu0 is a parameter that is the signal intensity for b value 0. It has been given for each patient.

# Chapter 3

## Statistical Methods:

### Likelihood:

In the project we use the likelihood of the voxel magnitude which follows Rice distribution.

Let  $M_1, M_2, \dots, M_n$  be random sample from Rice distribution with parameters  $\gamma_0 \exp(-\alpha * b)$  and " $\sigma^2$ " where  $\gamma_0$  signal intensity for  $b = 0$ . Consider a variable transformation:  $R_i = \frac{M_i^2}{\sigma^2}$ , where  $R_i$  follows a non-central  $\chi^2$  distribution with  $2P + 2$  degrees of freedom.  $P$  in the degrees of freedom can be expressed as a Poisson distribution with mean  $\gamma$ .

### Prior:

The prior information that has been used for ADC estimates is : The ADC values range from 0 - 0.14  $cm^2/sec$  in the Xenon MRI set up (Even in other set ups the ADC value are generally positive and small in magnitude.) Thus, the prior of Uniform distribution with a and b parameters as 0 and 0.14 has been used for finding the ADC estimates.

### Posterior Distribution:

Now since both the likelihood and the prior are proper thus, the posterior distribution exists. Due to the computational complexity of posterior distribution rjags has been used to apply MCMC to generate the posterior means

which helps us to estimate ADC. "rjags" package in R either use Metropolis Hastings or Gibbs Sampling.

### Metropolis Hastings:

Metropolis-Hastings is a popular Markov chain Monte Carlo (MCMC) algorithm used to generate sample from a probability distribution which is in most cases are posterior distribution in Bayesian setup when it is difficult or impossible to sample from it directly. The algorithm works by proposing a new sample from a proposal distribution and then accepting or rejecting this proposal based on the ratio or probability of the target distribution to the proposal distribution at the proposed sample which is also called acceptance probability.

The steps of the Metropolis-Hastings algorithm are as follows:

**Result:** Sequence of samples from the target distribution  
Choosing an initial value for the Markov chain,  $x_0$ ; **for**  $i = 1, 2, 3, \dots$   
**do**  
    We have to propose a new state,  $x'$ , from a proposal distribution,  $q(x'|x)$ ; Then main part of the algorithm is to calculate the acceptance ratio,  $A$ , given by  $A = \min(1, \alpha)$ , where  $\alpha = \frac{p(x') \cdot q(x|x')}{p(x) \cdot q(x'|x)}$ , where  $p(x)$  is the target probability density function and  $q(x'|x)$  is the proposal density function for moving from  $x$  to  $x'$ ; Depending on the probability we will either accept the proposal with probability  $A$ , otherwise retain the current state  $x$ ; Setting the current state to  $x'$  or  $x$  depends on whether the proposal was accepted or not;  
**end**

#### **Algorithm 1:** Metropolis-Hastings Algorithm

The algorithm generates a sequence of samples that are drawn from the posterior distribution. The samples are generally correlated with each other since the algorithm depends on the previous generated sample, but the correlation decreases as the number of iterations increases. That is an indication of convergence. It is important to check if the MCMC algorithm has converged or not. The algorithm requires some tuning of the proposal distribution to ensure that it is neither too narrow nor too wide.

## Gibbs Sampling:

Gibbs Sampler is another Markov Chain Monte Carlo(MCMC) method which helps in generating samples from a particular probability distribution. In Bayesian set up it is generally posterior distribution. This is one of the most common method of generating samples in Bayesian set up. In this algorithm at each step, Gibbs sampler updates one variable(in a multivariate setting) and get a sample of new value for the variable from its conditional distribution. This process is repeated for each variable and the resulting sample sequence converges to the said posterior distribution.

1. **Input:** First step is to initialize values for all variables  $x_1, x_2, \dots, x_n$ , number of iterations  $T$ , and conditional distributions  $p(x_i|x_{-i})$  for each variable  $x_i$
2. **Output:** Get samples  $x^{(1)}, x^{(2)}, \dots, x^{(T)}$
3. Then initialize  $x = (x_1, x_2, \dots, x_n)$
4. **For**  $t = 1$  to  $T$  **do**
  - (a) **For**  $i = 1$  to  $n$  **do**
    - i. For sample  $x_i$  from  $p(x_i|x_{-i}^{(t)})$
    - ii. The algorithm updates  $x_{-i}^{(t)}$  with the new value of  $x_i$
  - (b) It saves the current values of  $x$  as  $x^{(t)}$
5. **Return**  $x^{(1)}, x^{(2)}, \dots, x^{(T)}$

## Bayesian Regression:

In Bayesian regression, we start by specifying a prior distribution for the parameters of the regression model. This prior distribution represents our beliefs about the parameters before we observe any data. The choice of prior distribution can have a significant impact on the posterior distribution that we obtain, and it can also reflect any prior knowledge or assumptions that we have about the parameters.

For example, in Bayesian linear regression with a single independent variable, we might specify a prior distribution for the intercept and slope coefficients as follows:

$$\beta_0 \sim \mathcal{N}(0, \sigma_0^2)$$

$$\beta_1 \sim \mathcal{N}(0, \sigma_1^2)$$

Here, we are assuming that the intercept and slope coefficients follow normal distributions with mean 0 and variances  $\sigma_0^2$  and  $\sigma_1^2$ , respectively. The choice of these prior variances can reflect our prior beliefs about the likely range of values for the coefficients. For example, if we believe that the coefficients are likely to be small, we might choose smaller values for the prior variances.

Once we have specified the prior distribution, we can use Bayes' theorem to obtain the posterior distribution for the parameters, given the observed data. Specifically, the posterior distribution is proportional to the likelihood of the data, multiplied by the prior distribution:

$$P(\boldsymbol{\beta}|\mathbf{y}, \mathbf{X}) \propto P(\mathbf{y}|\mathbf{X}, \boldsymbol{\beta})P(\boldsymbol{\beta})$$

Here,  $\boldsymbol{\beta}$  represents the vector of regression coefficients,  $\mathbf{y}$  represents the vector of observed dependent variable values, and  $\mathbf{X}$  represents the matrix of independent variable values. The likelihood of the data given the parameters is given by the product of the individual data point likelihoods:

$$P(\mathbf{y}|\mathbf{X}, \boldsymbol{\beta}) = \prod_{i=1}^n \mathcal{N}(y_i | \beta_0 + \beta_1 x_{i1} + \cdots + \beta_p x_{ip}, \sigma^2)$$

Here,  $\mathcal{N}$  represents the normal probability density function, and  $\sigma^2$  represents the variance of the dependent variable.

To obtain the posterior distribution, we can use MCMC (Markov Chain Monte Carlo) methods to sample from the joint posterior distribution. This allows us to estimate the posterior mean and credible intervals for the parameters, as well as to make predictions for new data points by integrating over the posterior distribution.

In summary, Bayesian regression with priors allows us to incorporate prior knowledge and assumptions into the analysis, and provides a way to quantify uncertainty and make predictions based on the posterior distribution. The choice of prior distribution can reflect our prior beliefs about the parameters, and can have a significant impact on the posterior distribution. MCMC methods can be used to obtain samples from the posterior distribution for inference and prediction.

# Chapter 4

## Experiment Setup:

Model:

The model with the likelihood of Rice distribution and prior of Uniform distribution has been implemented as described above using MCMC to get the ADC estimates. R software has been used and the package rjags has been used for the analysis. "rjags" uses Gibbs sampling or Metropolis Hastings or a combination of both depending on the set up of the experiment. The following parameters were used for "rjags":

Using Metropolis Hastings (Using R-jags).

Number of runs : 10000

Burn-in period : 5000

Number of chain : 1

Thining parameter : 5

Posterior Sample: 1000

Convergence Diagnostic : Trace plot.

Although we are using "rjags" but R codes for Metropolis Hastings or

Gibbs sampling can be written.

## REGRESSION:

### SET UP 1: SIMPLE LINEAR REGRESSION

The ADC estimates are used to set up a simple linear regression using `lm` function of R software.

### SET UP 2: BAYESIAN LINEAR REGRESSION

In this model we have set a two stage Bayesian model: The first stage remains the same as described above

In the second stage for ADC for within the lung area the desired model was written down and the priors used are the ones generally used for Bayesian regression model normal distribution with mean 0 and variance 100000. This is was carried out within the same model set up of the ADC estimates. Now the same set up have been used for the two stage Bayesian model. The following are the parameters used:

Using Metropolis Hastings (Using R-jags).

Number of runs : 10000

Burn-in period : 5000

Number of chain : 1

Thining parameter : 5

Posterior Sample: 1000

Although we are using "rjags" but R codes for Metropolis Hastings or Gibbs sampling can be written.

# Chapter 5

## Results:

The ADC estimates which are found in this setting are estimated as mean of all voxels inside the lung area for each patients. A summary of the ADC estimates(inside the lung area) with CF patients is given as follows:

CF Patients:

1. Mean : 0.03159
2. Median : 0.03091
3. Standard Deviation : 0.00365297
4. CI 95% : (0.02581033 , 0.03863896)

ADC estimates	age
0.0341	23.30
0.0381	26.8
0.0379	31.6
0.0373	37.6

Table 5.1: Table showing the individual ADC estimates for some older CF patients

NOTE:

Just looking at the mean of all patients we can conclude that ADC for CF patients are comparatively lower than that of controls which are listed on Pg. 14.

But if we dive deep into individual estimates we will find some older CF patient having higher ADC estimates than that of age matched controls. For example if we look at the ADC estimate of patient with age 31.6 years we see the ADC estimate is higher than that of a control of age 30.6.

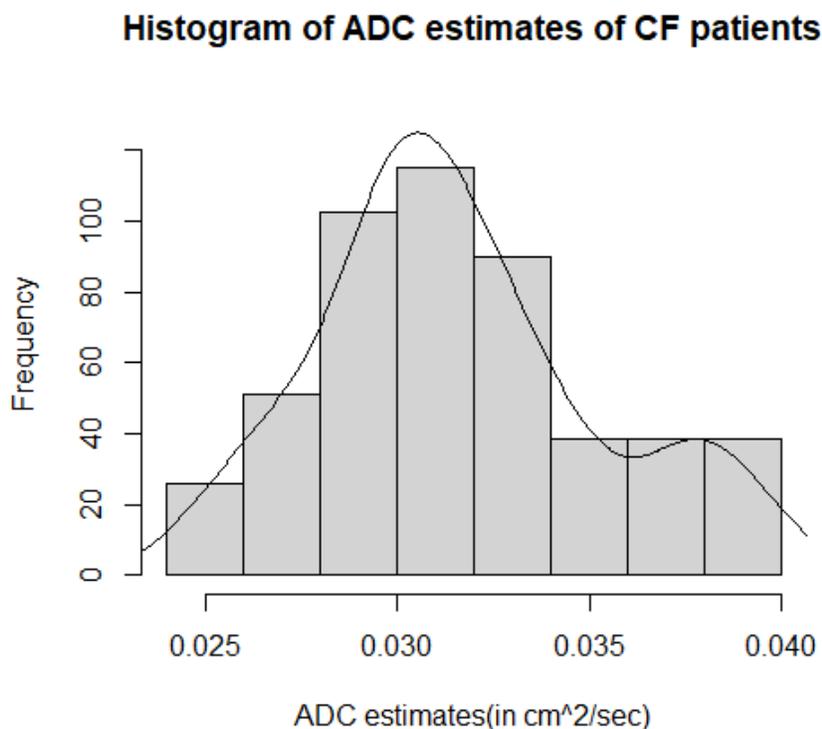


Figure 5.1: Histogram with density of ADC estimates(inside the lung area) of CF patients

**Scatter plot of ADC estimates of CF patients vs age**

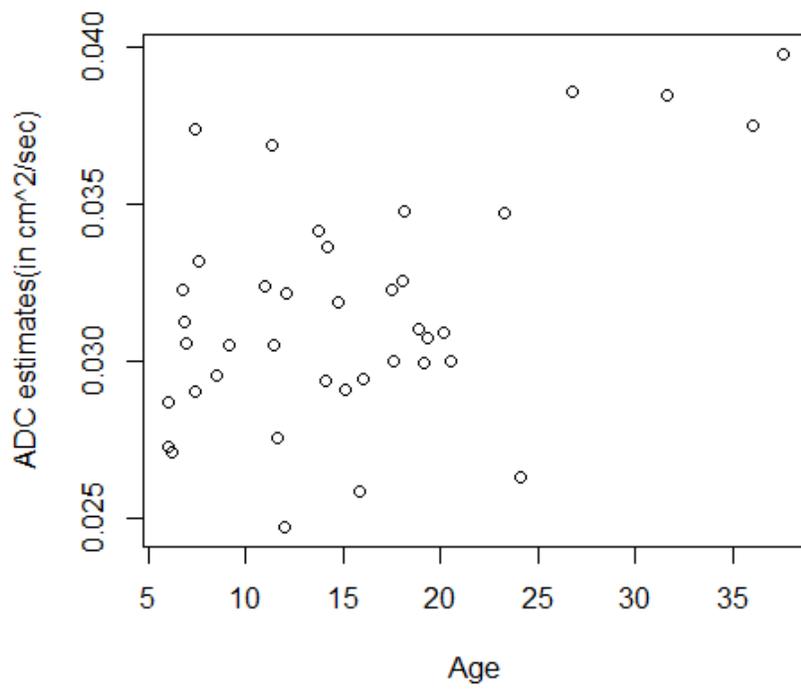


Figure 5.2: Scatter plot of ADC estimates (inside the lung area) of CF patients vs age

Control Patients:

1. Mean : 0.03162
2. Median : 0.03098
3. Standard Deviation : 0.003698021
4. CI 95% : (0.02578125 0.03866951)

ADC estimates	age
0.0389	23.7
0.0310	27.6
0.0360	30.6
0.0326	39

Table 5.2: Table showing the individual ADC estimates for some older controls

NOTE:

Just looking at the mean of controls we can conclude that ADC for controls are comparatively higher than that of CF patients which are listed on Pg. 11. Also apparent from the histogram for controls on Pg. 15 compared to the histogram of CF patients on Pg. 12.

But if we dive deep into individual estimates we will find some older CF patient having higher ADC estimates than that of age matched controls. For example if we look at the ADC estimate of patient with age 31.6 years we see the ADC estimate is higher than that of a control of age 30.6.

**Histogram of ADC estimates of Controls**

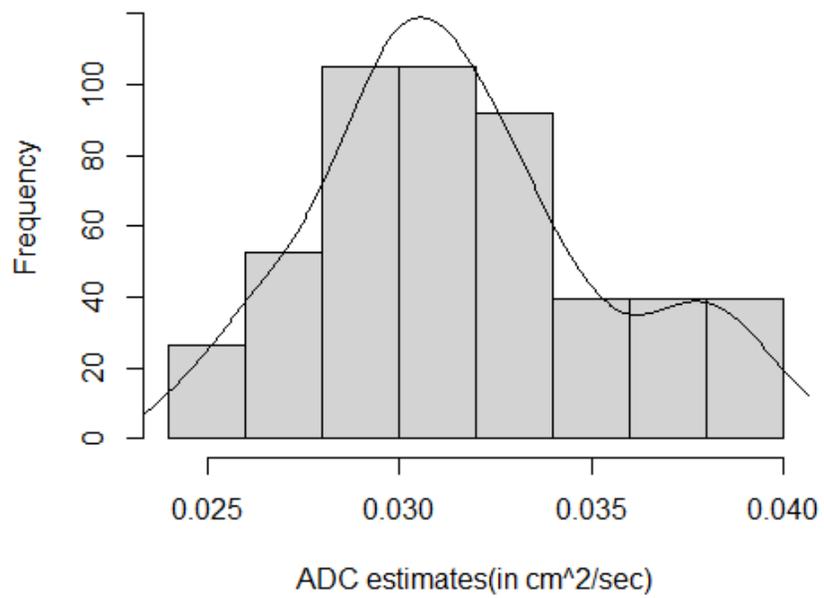


Figure 5.3: Histogram with density of ADC estimates(inside the lung area) of Controls

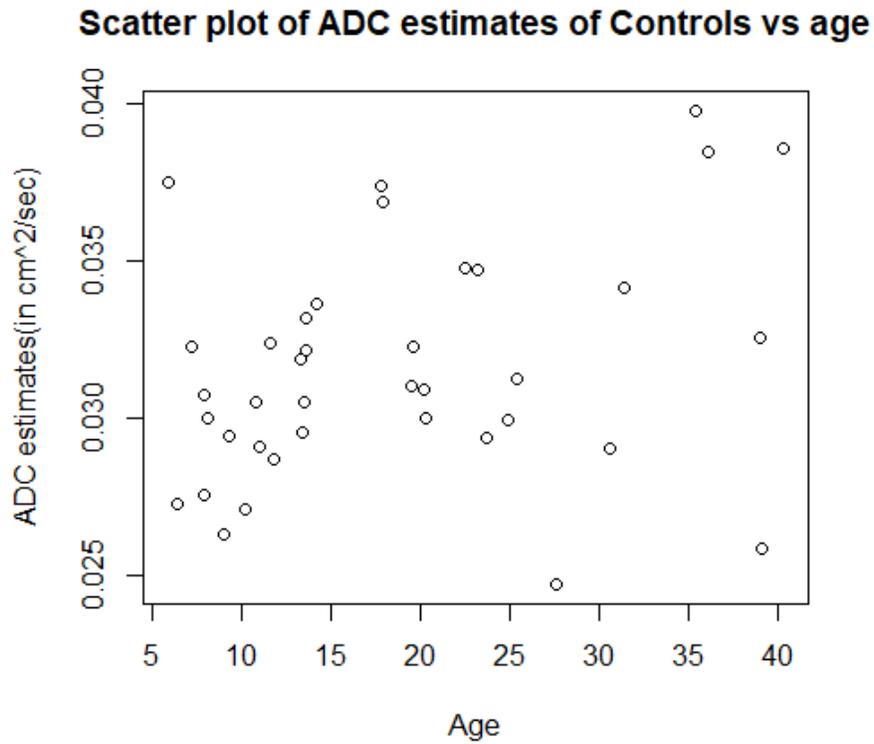


Figure 5.4: Scatter plot of ADC estimates (inside the lung area) of Controls vs age

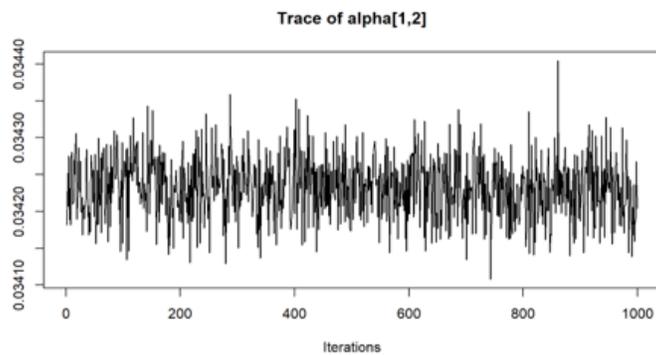


Figure 5.5: Trace plot of one patient - Checking for convergence of MCMC

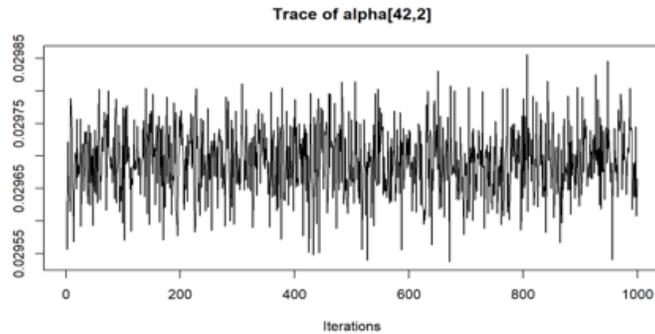


Figure 5.6: Trace plot of another patient - Checking for convergence of MCMC

Note: The trace plots does not show any visible patterns so we can conclude that the MCMC has converged.

We find that generally for CF patients the ADC value is lower than that of healthy patients which supports the biological reasoning that CF lungs are heavier than in healthy subjects, resulting in smaller ADC. But further investigation considering age of individuals we see an interesting observation. *For older CF patients we observe higher ADC as compared to that of healthy patients. This is interesting and it also supports the claim that is due to micro structural damage caused by years of infections and inflammation.*

**This is the reason which inspired us to carry out the regression of ADC estimates on age primarily.**

## Regression Equation:

SET UP 1:

### SIMPLE LINEAR REGRESSION

The dependent variable in the model is the ADC estimates within the lung area.

The independent variables in the model used are age and a dummy variable indicating the case and control patient.

The model also includes the interaction effects of all the independent variables mentioned above.

The equation is as follows:

$$ADC = 0.02871 + 0.0002353 * Age - 0.0007369 * CF - 7.096e-07 * Age * CF$$

Further analysis:

The variable sex was also included in the regression model. The ANOVA table for the same is given below:

Variables	p-values
(Intercept)	$< 2.e^{16}$
Age	0.00209
CF	0.65962
Sex	0.61348
Age:CF	0.99170
Age: Sex	0.84796
CF : Sex	0.99180

Table 5.3: ANOVA table for Simple Linear Regression

From the ANOVA table, we can conclude that the significant variable is age. So in the final model we only used age and the dummy variable CF so that we get different equations for controls and CF patients.

The equation for the CF patients:

Putting  $CF = 1$  :

$$ADC = 0.0275574 + 0.0002431 * Age$$

The equation for the Controls:

Putting  $CF = 0$  :

$$ADC = 0.02834 + 0.0002422 * Age$$

SET UP 2:

### BAYESIAN LINEAR REGRESSION

With a connection from the simple linear regression we carried out the Bayesian Regression with age and the dummy variable CF. The priors used are explained in the experiment setup - Chapter 4.

The equation we get from the Bayesian Linear Regression is as follows:

$$ADC = 0.02863 + 0.0002379 * Age - 0.000425326 * CF + 0.00000077 * Age * CF$$

The equation for the CF patients:

Putting  $CF = 1$  :

$$ADC = 0.02820 + 0.0002387 * Age$$

The equation for the Controls:

Putting  $CF = 0$  :

$$ADC = 0.02863 + 0.0002379 * Age$$

### **Conclusion:**

From the results above we see that the equations we get from both setups are quite similar. Again further research can be carried out for looking into the error of the two setups for further comparison.

Again, if we look at the intercept in the regression line for control and CF separately, we also observe that the coefficient for controls is higher than that of CF.

If we look at the coefficients of the age in the regression line for control and CF separately, we also observe that the coefficient for CF is higher than that of Controls. We can also observe that for older people the age part of the regression may increase giving those elevated results - that's what we saw in individual estimates.

# Chapter 6

## Excursis and Future Work:

### Excursis:

We can report the following results:

1. We can conclude that ADC estimates for the controls are higher than that of the CF. (From histograms and means of all patients and controls reported)
2. We also observe that for some older CF patients have elevated ADC estimates than their age-matched control.
3. If we look at the intercept in the regression line for control and CF separately we also observe that the intercept is higher than that of CF and it is also much higher compared to the coefficients of the independent variable age.
4. Again, if we look at the coefficient of independent variable age for control and CF separately we also observe that the coefficient for CF is higher than that of controls .

Thus the observation of ADC estimates being higher for a CF patient than a control patient of the same age is due to the higher co-efficient for age variable gives higher ADC estimates but only for older patients

since the co-efficient for age variable is much less compared to the intercept so this can only be observed for older patients.

### **Future Work:**

This project opens up the scope of many future research. Some of the options are noted below:

In terms of coding, instead of using jags , any other package like stan or bugs can be used. Moreover, own code for metropolis hastings can also be used.

A difficulty of MH algorithm in this setup is the time consumption. Ways to improve the same can be explored.

The regression project can be carried out extensively by including many other variables which have been collected during the experiment like height, sex , lung volume etc.

While doing so it is also important to keep in mind that some of the variables are correlated with others upto a certain point, for example, our key variable age is correlated with both height and lung volume up to a certain point. Incorporating those is very important for having a sound model.

# Chapter 7

## Exordium:Project 2

Limitations of relatively low lung tissue density, rapid cardiac and respiratory motion and rapid transverse relaxation ( $T_2^*$ ) faced while conducting MRI of the lung are largely overcome by radial ultra-short echo-time (UTE) sequences. But this solution can sometimes lead to long acquisition times relative to the conventional Cartesian sequences due to the fact requirement of using additional radial projections from inefficiently sampling k-space. Therefore the norm often used in pulmonary imaging is to under sample to reduce the acquisition time. But again this solution introduced some biases – reduction of SNR(Signal-to-noise ratio), the introduction of image artifacts and degrade true image resolution. Here it is investigated to see to which level of under sampling the information that is specifically  $T_2^*$  mapping remains unaffected and their difference (taken from the full sampled image) is statistically significant. The investigation has been carried out on the data collected by R. Stecker, in the context of mouse imaging and Shepp-Logan Phantom data is also used to investigate the trends. Detailed discussion is carried out on how much under-sampling of images can be conducted keeping in mind the metric  $T_2^*$  from a statistical point of view.

# Chapter 8

## Data Acquisition:

C57BL/6J mice (Jackson Laboratory, Bar Harbor, ME) (N = 14) were housed at Cincinnati Children's Hospital. Imaging was performed on a 7 T horizontal, 30-cm bore, small animal scanner (Bruker BioSpec, Billerica, MA) with a home-built quadrature transmit/receive birdcage coil (inner diameter = 35 mm; length = 50 mm). Before experiments, the scanner frequency, shim settings, and animal position were determined using Bruker's built-in routines. During experiments, free-breathing mice were anesthetized with 1-3% isoflurane (mixed with air) and imaged head-first and supine at the magnet isocenter. After imaging, mice recovered in an isolated heated cage, set to 36°C, and returned to the barrier room until fully recovered from anesthesia.

### ACQUISITION OF FULLY-SAMPLED IMAGE

Fully-sampled, free induction decay (FID) data were acquired using a 3D radial sequence with five interleaved echo times (TE). Following are some of the parameters used for image acquisition: A flip angle of  $\alpha = 6.3^\circ$ ; 4.3  $\mu\text{s}$  RF block pulse; 51,472 radial projections per image; 128 points on free induction decay (FID); field-of-view (FOV) =  $32 \times 32 \times 64$  (mm<sup>3</sup>); image matrix size =  $128 \times 128 \times 128$  (voxels<sup>3</sup>); receiver bandwidth (BW) = 277 kHz; TEs = [0.08, 0.25, 0.50, 1.25, and 2.50] ms; and TR = 8.24 ms. The total acquisition time for each multi-TE scan was 90 minutes.

## UNDER SAMPLING OF RAW k-SPACE DATA:

Under sampling was done by randomly removing FIDs (Free Induction Decay) from fully sampled k-space data for all echo time images. Each FID was assigned a unique index ranging from 1 to the total number of radial projections, and a specified percentage of these indices was selected for removal by a Mersenne Twister random number generator. The fraction of unique FIDs—and thus k-space trajectories—removed from the fully-sampled data increased in increments of 10% (i.e., 90%, 80%, 70%... to 10% sampled). This was iterated 10 times for each fully sampled image for 14 mice at each echo time and level of sampling. For every iteration and each level of Nyquist sampling the random number generator seed was changed to the date and time of execution for each iteration to maintain unique sets. A total of 1,260 distinct under-sampled and 14 fully-sampled multi-TE sets (5 TE images per set) were generated.

## SIMULATION OF SHEPP-LOGAN PHANTOM WITH GOLDEN ANGLE SAMPLING:

Multi-TE simulations were performed in MATLAB using digital 3D Shepp-Logan phantoms that were modified to include additional ellipsoids to mimic mouse lung vasculature. The phantom signal intensity was chosen to match the one observed in the mice lung dataset. A  $T_2^*$  of 0.4 ms was assigned to the simulated lung parenchyma and a  $T_2^*$  of 3.16 ms for the simulated lung vasculature. Simulated radial k-space data were sampled similarly to in mice data using golden angle sampling, with TEs = [0.08, 0.25, 0.50, 1.25, and 2.50] ms; TR = 7 ms. K-space data were then gridded by rounding to the nearest Cartesian grid point to reduce the time for phantom simulations. The phantom was generated at three different 3D isotropic matrix sizes: 643, 963, and 1283. The total radial projections for each matrix size image were: 12,868 for 643, 28,953 for 963, and 51,472 radial projections for 1283.

## CALCULATION OF $T_2^*$ :

Binary masks of the thoracic cavity from the fully sampled mice images were automatically segmented using a deep-learning algorithm (based on the model of U-Net Convolution Neural Network). This analysis was confined to the parenchymal tissues by drawing a region-of-interest (ROI) of the fully sampled images and the mean of this ROI has been used to weigh the signal values within the lung mask. Voxels with weighted signal falling below the threshold value of 0.8(were treated as lung parenchyma) were used for calculation of SNR and  $T_2^*$ . SNR was calculated from TE = 0.08 ms image by averaging, a non-weighted, lung parenchyma signal and dividing by the standard deviation of the signal within an ROI drawn outside the mouse body in a region free from obvious image artifacts.  $T_2^*$  was estimated voxel-by-voxel via iterative nonlinear least squares fitting of the signal intensity as a function of TE to

$$S_1 = S_0 \exp \frac{-TE}{T_2^*} + c$$

,where c is a constant representing the non-zero noise floor.

# Chapter 9

## Statistical Methods:

### One sample Hotelling $t^2$ test:

The Phantom simulations and the vivo mice datasets have several iterations. To incorporate all the iterations in the dataset the statistical test has been conducted in a multivariate setup where each iteration has been looked at as a variable. The null hypothesis being tested here is the mean of each iteration is equal to 0.

$$H_0 : (\mu_1, \mu_2, \mu_3, \mu_4, \dots, \mu_n) = (0, 0, 0, 0, \dots, 0)$$

$$H_1 : (\mu_1, \mu_2, \mu_3, \mu_4, \dots, \mu_n) \neq (0, 0, 0, 0, \dots, 0)$$

The level of significance is 0.05. This test is known as one sample Hotelling  $t^2$ . Here the  $n$  depends on the number of iterations in each data set-up which is 10 for each isotropic matrix size phantom data and 7 for each real mouse. The Hotelling  $t^2$  statistic is

$$t^2 = n(\bar{X} - \mu_0)'S^{-1}(\bar{X} - \mu_0)$$

where  $\bar{X}$  is the multivariate sample mean and

$S$  is the sample covariance variance matrix and

$\mu_0$  is the vector we want to test for here in our case  $\mu_0$  is 0 vector.

The statistic represents the distance between the sample mean and  $\mu_0$ . Thus, the statistic is expected to have low value if  $\bar{X}$  is near  $\mu_0$  high value if they are different.

Under the null hypothesis, the test statistic follows F distribution with parameters p and n - p.

$$t^2 \sim T_{p,n-1}^2 = \frac{p(n-1)}{n-p} \cdot F_{p,n-p}$$

Again, if  $\bar{X}$  is near  $\mu_0$  which is 0 is our then p-value is expected to be greater than 0.05 and if they are different then the p-value should be less than 0.05.

In this situation, the parameters used are:

1.  $\mu_0$  is the 0 vector.
2.  $\bar{X}$  is the sample mean vector across all the iterations of the calculated  $T^2$ .
3.  $S$  is calculated by taking the variance and covariance matrix across all the iterations of the calculated  $T^2$ .

The p-value of all these tests are noted down and listed in tables in the Results chapter(Page number can be found from List of Tables)

## Why Hotelling $T^2$ :

To incorporate all these small variabilities it is better to have a multivariate setup than a univariate setup like taking the mean of all iterations and comparing it to fully sampled MRI by a paired t test. Since the iterations are repeatedly taken on the same mouse so they cannot be said to be independent to each other. So taking into account the covariance between

the iterations will give us some information. But to cross-check the results of Hotelling  $T_2^*$ , paired t-test was also carried out.

### Paired T Test:

Apart from the Hotelling  $T_2^*$ , paired t-test has also been carried out. The null hypothesis being tested here is as follows:

$$\begin{aligned} H_0 : \mu_U &= \mu_F \\ H_1 : \mu_U &\neq \mu_F \end{aligned}$$

where  $\mu_U$  and  $\mu_F$  stand for the mean of the under-sampled MR images over all iterations and the mean of fully sampled MR images over all iterations. Paired t-test are carried out for each level of under-sampled MR Images. Thus, 9 paired t-test has been carried out for each mouse. In total, 126 paired t-test has been carried out.

### Outliers:

If any calculated  $T_2^*$  is greater than 10ms then those voxels are ignored. The steps used to carry out the identification and ignorance of the voxels are as follows

Step 1 - Summary of all iterations of each level of under-sampling and even fully sampled of all iterations are carried out to identify which level of under-sampled images have voxels of more than 10ms of calculated  $T_2^*$ .

Step 2 - Following step 1, a counter has been setup to see how many voxels have  $T_2^*$  greater than 10ms for those levels of under-sampled images.

Step 3 - Next, the locations of those voxels are taken into account which have the  $T_2^*$  values greater than 10ms for those levels of under-sampled images.

Step 4 - Following this step, those voxels are ignored for all iterations which have values greater than 10ms for those level of under-sampled images.

Step 5- Thereafter the Hotelling  $t^2$  test and paired t-test has been conducted.

### **Intraclass Correlation Co-efficient[ICC]:**

The intraclass correlation coefficient (ICC), is a descriptive statistic that can be used when quantitative measurements are made on units that are organized into groups. It describes how strongly units in the same group resemble each other.

In other words, it is a statistical measure used to assess the reliability or consistency of measurements made by different raters or observers, or by the same rater on different occasions. ICC measures the proportion of the total variability in the data that can be attributed to between-group differences, relative to the total variability. In other words, it quantifies the degree to which observations within a group are similar, compared to the variation between different groups.

ICC values range from 0 to 1, with higher values indicating greater reliability or consistency in the measurements. A value of 0 indicates that there is no agreement between the raters or observers, while a value of 1 indicates perfect agreement.

ICC is commonly used in various fields such as psychology, medicine, and education where it is important to ensure that measurements are consistent and accurate. ICC can help researchers determine whether their data is sufficiently reliable to draw valid conclusions. It can also be used to identify sources of measurement error or variability and to guide improvements in measurement procedures.

The formula for calculating ICC is:

$$ICC = \frac{MS_b - MS_w}{MS_b + (k - 1) \times MS_w} \quad (9.1)$$

where  $MS_b$  is the between-group mean square,  $MS_w$  is the within-group mean square, and  $k$  is the number of raters or measurement occasions.

**In our case for the Phantom dataset setup :** The calculated ICC reflects how organized are the units within the 10 iterations of every level of under-sampling.

**Vivo mice dataset setup :** The calculated ICC reflects how the iterations for each level of under-sampling for each mouse resemble each other.

So in our case, the within-group mean squares are the ones within each iteration and the between-group are the ones between the iterations.

R packages pshyc and irr have been used for calculation and cross-checking.

# Chapter 10

## Experiment Setup:

### Experimental Setup:

#### Phantom Simulations:

As previously mentioned Phantom dataset used here has 10 iterations and the corresponding calculation of the  $T_2^*$  for total radial projection of 12,868 for  $64^3$ , 28,953 for  $96^3$ , and 51,472 radial projections for  $128^3$  for the fully sampled and also the under-sampled data set with decrements of 10 from the fully sampled at each stage that is 90% sampled, 80% sampled, 70% sampled, 60% sampled, 50% sampled, 40% sampled, 30% sampled, 20% sampled and 10% sampled. None of the levels of under sampled images have outliers in the calculated  $T_2^*$ . Now apart from the investigations carried out in the manuscript, some statistical tests were also conducted to scrutinize the statistical significance of such under-sampling on the parameter in context  $T_2^*$ . We are carrying out Hotelling  $t^2$  test. The hypothesis being tested has been explained in detail previously. The mathematical expression is as follows:

$$\begin{aligned} H_0 &: (\mu_1, \mu_2, \mu_3, \mu_4, \dots, \mu_n) = (0, 0, 0, 0, \dots, 0) \\ H_1 &: (\mu_1, \mu_2, \mu_3, \mu_4, \dots, \mu_n) \neq (0, 0, 0, 0, \dots, 0) \end{aligned}$$

The steps below are followed to carry out Hotelling  $t^2$  test:

**Step 1:** The first step carried out to do this was to multiply the binary

mask with the calculated  $T_2^*$  to incorporate the region of interest (ROI) images on a voxel to voxel level for all the three dataset 12,868 for  $64^3$  , 28,953 for  $96^3$  , and 51,472 radial projections for  $128^3$  and also for each of 10 iterations.

**Step 2:** Now since only the region in the lung is the priority in this experiment - only the voxels with nonzero voxels were taken into account after the multiplication with the binary mask for all the three dataset  $64^3$  ,  $96^3$  , and  $128^3$  voxels and also for each of 10 iterations.

**Step 3:** The non-zero  $T_2^*$  measurements were 1066 out of  $64^3$  , 4522 out of  $96^3$  , and 11950 out of  $128^3$  voxels and also for each of 10 iterations.

**Step 4:** Therefore, differences of calculated  $T_2^*$  for each under-sampled image were taken from fully sampled images (like 100% sampled - 90% sampled: 100% sampled – 80% sampled and so on) on a voxel to voxel level basis but only for the non-zero weighted  $T_2^*$  after multiplication with the binary mask - that is  $T_2^*$  difference were taken for 1066 out of  $64^3$  , 4522 out of  $96^3$  , and 11950 out of  $128^3$  voxels and also for each of 10 iterations.

**Step 5:** To investigate the dependence of the under-sampling on  $T_2^*$  the parameter of interest is the mean of  $T_2^*$  differences calculated for each iteration for 1066 out of  $64^3$  , 4522 out of  $96^3$  , and 11950 out  $128^3$  voxels.

**Step 6:** Now after all the calculations left with 10 different means are left corresponding to each iteration for each level of under-sampling. Mathematically speaking we have:

Number of means	Level of $T_2^*$ differences
10 (one for each iteration)	100% sampled – 90% sampled
10 (one for each iteration)	90% sampled – 80% sampled
10 (one for each iteration)	80% sampled – 70% sampled
10 (one for each iteration)	70% sampled – 60% sampled
10 (one for each iteration)	60% sampled – 50% sampled
10 (one for each iteration)	50% sampled – 40% sampled
10 (one for each iteration)	40% sampled – 30% sampled
10 (one for each iteration)	30% sampled – 20% sampled
10 (one for each iteration)	20% sampled – 10% sampled

Table 10.1: A table showing the number of means for each level of difference for Phantom Dataset

So there are 10 means for each of three datasets -1066 out of  $64^3$  , 4522 out of  $96^3$  , and 11950 out  $128^3$

The set-up is therefore in a multivariate format for each dataset. Therefore to look into the effect of  $T_2^*$  due to under-sampling one possible way is to look compare the means of the  $T_2^*$  differences. voxels.

**Step 7:** Comparison of means of  $T_2^*$  differences in multivariate setup can be done by Hotelling  $t^2$  test where we test the means of each variable or column(in a matrix setup) is equal to zero or not.

**Step 8:** Thus, Hotelling  $t^2$  test was carried out for each level of  $T_2^*$  difference based on of 10 iterations taken as each variable.

**Step 9:** So 9 Hotelling  $t^2$  tests were carried out for each dataset 1066 out of  $64^3$  , 4522 out of  $96^3$  and 11950 out of  $128^3$  voxels, so in total 27 tests were carried out.

**Step 10:** Similarly, we also perform paired t-test in the same setup and compare and contrast the results.

## Paired T Test setup:

As previously mentioned Phantom dataset used here has 10 iteration and the corresponding calculation of the  $T_2^*$  for total radial projection of 12,868 for  $64^3$  , 28,953 for  $96^3$  and 51,472 radial projections for  $128^3$  for the fully sampled and also the under-sampled data set with a decrement of 10% from the fully sampled at each stage that is 90% sampled, 80% sampled, 70% sampled, 60% sampled, 50% sampled, 40% sampled, 30% sampled, 20% sampled and 10% sampled. None of the levels of under-sampled images have outliers in the calculated  $T_2^*$ .

Now apart from carrying out Hotelling  $t^2$  , some other statistical tests like paired t-test were also conducted to scrutinize the statistical significance of such under-sampling on the parameter in context  $T_2^*$ . The hypothesis being tested has been mentioned before.

The steps that are followed to calculate the paired t-test are as follows:

Step 1: Reading calculated  $T_2^*$  under-sampled and fully sampled MR Images.

Step 2: Next these images were multiplied with the binary mask to get those voxels that are inside the periphery of the lungs.

Step 3: The following step was the check for any outliers. Any calculated  $T_2^*$  greater than 10ms was ignored.

Step 4: Next step was to calculate the mean of all the iterations at each level of under-sampling (such as 10%, 20%, and so on) on a voxel to voxel basis (considering the voxels inside the periphery of the lung).

Step 5: Then the equality of the means ( $\mu_U$  and  $\mu_F$ ) is checked by conducting a paired t using the R software.

## Experimental Setup:

### Vivo Mice Data:

As previously mentioned data are collected from 14 mice at Cincinnati Children Hospital Medical Centre and 10 iterations are taken for each mice. But it is to be remembered that the data on vivo mice has the parameter of  $128^3$  voxel size only. Thus, the corresponding calculation of the  $T_2^*$  has been carried on  $128^3$  voxels for the fully sampled and also the under-sampled data set with a decrement of 10% from the fully sampled at each stage that is 90% sampled, 80% sampled, 70% sampled, 60% sampled, 50% sampled, 40% sampled, 30% sampled, 20% sampled and 10% sampled.

Now apart from the investigations carried out in the manuscript, some statistical tests were also conducted to scrutinize the statistical significance of such under-sampling on the parameter in context  $T_2^*$ . In terms of data structure - the data structure of each real mouse is the same as that of 3D isotropic matrix size of  $128^3$  of Phantom data. Thus, the calculation is carried out in an exact similar manner as described for Phantom simulations for each of the 14 mice.

Using the steps described in the outlier section we find the following:

Mice 3 have outliers at 90%, 80% and 70% sampled images. Mice 5 have outliers at 90% sampled images. Mice 10 have outliers at 90% sampled images.

These voxels have been ignored while carrying on the following calculation:

**Step 1:** As described in Phantom Simulation data first step carried out to do this was to multiply the binary mask with the calculated  $T_2^*$  to incorporate the region of interest (ROI) images on a voxel to voxel level for all the 14 mice dataset of  $128^3$  separately and also for each of 7 iterations for each level of sampled under-sampling.

**Step 2:** Now since only the region in the lung is the priority in this experiment - only the voxels with nonzero voxels were taken into account after the multiplication with the binary mask for all the 14 mice dataset of  $128^3$ .

**Step 3:** The non-zero  $T_2^*$  measurements varied from 5000-8000 out of  $128^3$  voxels depending on the shape and size of the lungs of 14 mice and also for each of 7 iterations of each level of sampled under-sampling.

**Step 4:** Therefore, differences of calculated  $T_2^*$  for each under-sampled images were taken from fully sampled images on a voxel to voxel level basis but only for the non-zero weighted  $T_2^*$  after multiplication with the binary mask - that is  $T_2^*$  difference was taken for those 5000-8000 voxels out of  $128^3$  voxels and also for each of 7 iterations of each level of sampled under-sampling.

**Step 5:** To investigate the dependence of the under-sampling on  $T_2^*$  the parameter of interest is the mean of  $T_2^*$  differences calculated for each iteration for those 5000-8000 voxels out of  $128^3$  voxels.

**Step 6:** Now after all the calculations 7 different means depending on the number of unique iterations are left corresponding to each iteration for each level of under-sampling. Mathematically speaking we have the following for each of the 14 mice:

Number of means	Level of $T_2^*$ differences
7 (one for each iteration)	100% sampled – 90% sampled
7 (one for each iteration)	90% sampled – 80% sampled
7 (one for each iteration)	80% sampled – 70% sampled
7 (one for each iteration)	70% sampled – 60% sampled
7 (one for each iteration)	60% sampled – 50% sampled
7 (one for each iteration)	50% sampled – 40% sampled
7 (one for each iteration)	40% sampled – 30% sampled
7 (one for each iteration)	30% sampled – 20% sampled
7 (one for each iteration)	20% sampled – 10% sampled

Table 10.2: A table showing the number of means for each level of difference for Vivo real mice data

So there are either 7 means for each of 14 datasets - for those 5000-8000 voxels(those inside lung area) out of  $128^3$  voxels. The set-up is therefore in a multivariate format for each dataset. Therefore to look into the effect of  $T_2^*$  due to under-sampling one possible way is to look compare the means of the  $T_2^*$  differences.

**Step 7:** In the case of vivo real data set we observe that paired t-test are working better than Hotelling  $t^2$ .

**Step 8:** Comparison of means of  $T_2^*$  differences in multivariate setup can be done by Hotelling  $t^2$  test where we test the means of each variable or column(in a matrix setup) is equal to zero or not.

**Step 9:** Thus, Hotelling  $t^2$  test was carried out for each level of  $T_2^*$  the difference based on of either 7 iterations taken as each variable.

**Step 10:** So 9 Hotelling  $t^2$  tests were carried out for each mouse on those 5000-8000 voxels(inside the lung area) out of  $128^3$  voxels and doe the three isotropic size for Phantom data, so in total 153 tests were carried out.

### Paired T-Test setup:

As previously mentioned in vivo real mice dataset used here has 7 iterations for each of the 14 mice and the corresponding calculation of the  $T_2^*$  for the fully sampled and also the under-sampled data set with a decrement of 10% from the fully sampled at each stage that is 90% sampled, 80% sampled, 70% sampled, 60% sampled, 50% sampled, 40% sampled, 30% sampled, 20% sampled and 10% sampled. None of the levels of under-sampled images have outliers in the calculated  $T_2^*$ .

Now apart from carrying out Hotelling  $t^2$ , some other statistical tests like paired t-test were also conducted to scrutinize the statistical significance of such under-sampling on the parameter in context  $T_2^*$ . The hypothesis being tested has been mentioned before.

The steps that are followed to calculate the paired t-test are as follows:

**Step 1:** The first step was to read the calculated  $T_2^*$  under-sampled and fully sampled MR Images.

**Step 2:** Next these images were multiplied with the binary mask to get those voxels that are inside the periphery of the lungs.

**Step 3:** The following step was the check for any outliers. Any calculated  $T_2^*$  greater than 10ms was ignored.

**Step 4:** Next step was to calculate the mean of all the iterations at each level of under-sampling (such as 10%, 20% and so on) on a voxel to voxel basis (considering the voxels inside the periphery of the lung).

**Step 5:** Then the equality of the means ( $\mu_U$  and  $\mu_F$ ) is checked by conducting a paired t using the R software.

# Chapter 11

## Results:

### Hotelling $t^2$ Test Results:

#### Phantom Dataset:

For  $96^3$  and  $128^3$  voxel datasets, we see that in most cases the p-value keeps on increasing steadily through the levels of  $T_2^*$  differences from the less sampled images to the more sampled images.

#### 1. FOR $64^3$ voxels

We see that from the level of  $T_2^*$  difference of 100% sampled and 20% sampled onwards all the levels of  $T_2^*$  difference (that is 100% sampled – 30% sampled, 100% sampled – 40% sampled and so on till 100% sampled -90% sampled) have p-value greater than 0.05. The hypothesis being tested here is the means of the 10 iterations are equal to 0 or not. We can conclude that till 20% sampled images the metric in context  $T_2^*$  is not significantly affected by the under-sampling process.

#### 2. FOR $96^3$ voxels

We see that from the level of  $T_2^*$  difference of 100% sampled and 40% sampled onwards all the levels of  $T_2^*$  difference (that is 100% sampled – 50% sampled, 100% sampled – 60% sampled and so on till 100% sampled -90% sampled) have p-value greater than 0.05. The hypothesis being tested here is the means of the 10 iterations are equal to 0 or not. We can conclude that till 50% sampled images the metric in context  $T_2^*$  is not significantly affected by the under-sampling process.

### 3. FOR $128^3$ voxels

We see that from the level of  $T_2^*$  difference of 100% sampled and 70% sampled onwards all the levels of  $T_2^*$  difference (that is 100% sampled - 80% sampled and 100% sampled -90% sampled) have p-value greater than 0.05. The hypothesis being tested here was is the means of the 10 iterations are equal to 0 or not. Till 80% sampled images the metric in context  $T_2^*$  is not significantly affected by the under-sampling process

## Summarized Results:

### Hotelling $t^2$ Test and Paired t Test

100%-%Sampled	Hotelling $t^2$	Paired t	ICC
100-10	0.0369	0.2606	0.67
100-20	0.6449	0.2088	0.5
100-30	0.2776	0.9086	0.39
100-40	0.7865	0.9756	0.3
100-50	0.2039	0.9269	0.17
100-60	0.4145	0.9403	0.14
100-70	0.2815	0.7372	0.05
100-80	0.2736	0.6447	0.019
100-90	0.4891	0.4267	-0.0001

Table 11.1: A table p-values of Hotelling  $t^2$  , Paired t and ICC for  $64^3$

100%-%Sampled	Hotelling $t^2$	Paired t	ICC
100-10	<0.0001	<0.0001	0.49
100-20	<0.0001	<0.0001	0.32
100-30	<0.0001	<0.0001	0.17
100-40	<0.0001	0.1625	0.07
100-50	0.1220	0.8429	0.029
100-60	0.0981	0.8919	0.011
100-70	0.1117	0.6947	0.02
100-80	0.4905	0.4880	-0.0006
100-90	0.1708	0.6369	0.007

Table 11.2: A table p-values of Hotelling  $t^2$  , Paired t and ICC for  $96^3$

100%-%Sampled	Hotelling $t^2$	Paired t	ICC
100-10	<0.0001	<0.0001	0.43
100-20	<0.0001	<0.0001	0.25
100-30	<0.0001	<0.0001	0.14
100-40	<0.0001	0.0008	0.10
100-50	<0.0001	0.0012	0.057
100-60	0.0003	0.0046	0.029
100-70	0.0180	0.0264	0.0099
100-80	0.1304	0.0195	-2.9e-16
100-90	0.0868	0.0163	-5.6e-07

Table 11.3: A table p-values of Hotelling  $t^2$  , Paired t and ICC for  $128^3$

**NOTE:** It is observed that the ICC (Intraclass Correlation) among 10 iterations are low especially 50% sampled onwards. Negative ICC can be interpreted as really low value of ICC. In the above table we see that the negative ICCs are very close to 0. This suggests that the iterations especially from the difference of 100% sampled and 50% sampled onwards, the iterations are heterogeneous. Thus, we can see that Hotelling  $t^2$  test is working better for the phantom simulations compared to the paired t-test since the Hotelling  $t^2$  can capture all the variability across the iterations and even in between them as it takes into account the variance-covariance matrix of the iterations.

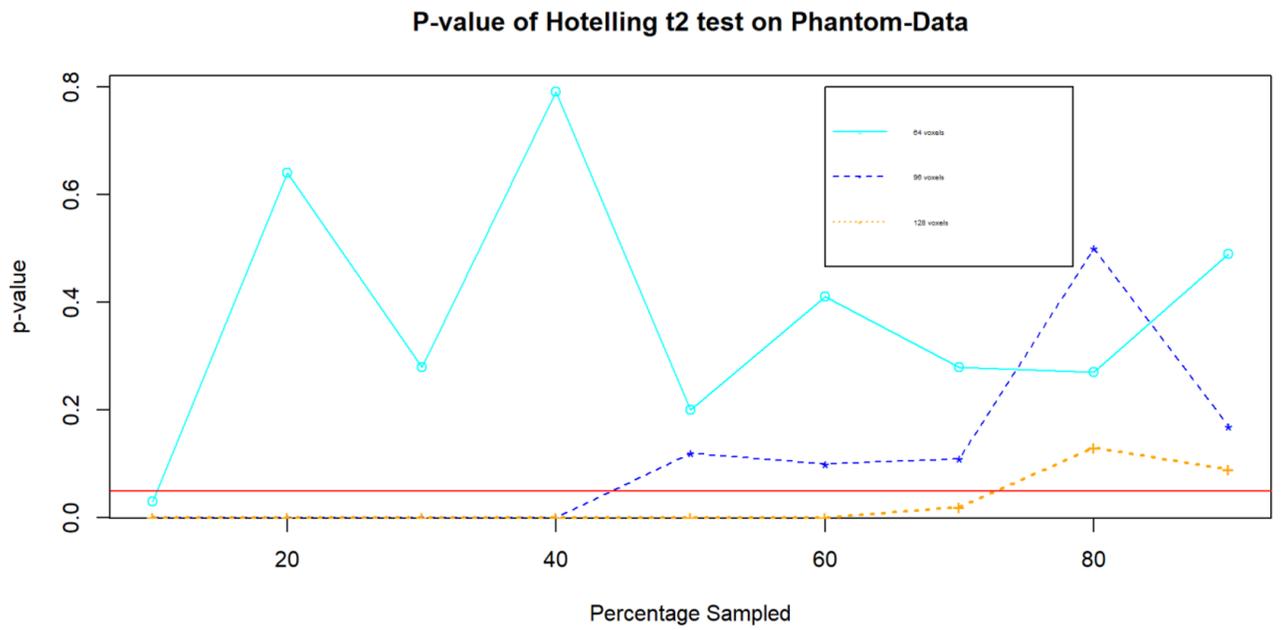


Figure 11.1: Visualization of pvalue of Hotelling  $t^2$

NOTE: The horizontal red line represents the 5% significance level. All the points plotted above it give nonsignificant results.

Thus, for the cyan colored plot (showing the plot of phantom data of 3D isotropic size  $64^3$  20% sampled or 80% under-sampled doesn't affect the calculation of  $T_2^*$ ).

Similarly, for the blue colored plot (showing the plot of phantom data of 3D isotropic size  $96^3$  50% sampled or 50% under-sampled doesn't affect the calculation of  $T_2^*$ ).

Lastly, for the yellow colored plot (showing the plot of phantom data of 3D isotropic size  $128^3$  80% sampled or 20% under-sampled doesn't affect the calculation of  $T_2^*$ ).

## Paired t Test Results:

### Vivo Real Mice Dataset:

The p-value keeps on increasing steadily through the levels of  $T_2^*$  from the less sampled images to the more sampled images. The null hypothesis being tested here is: the mean  $T_2^*$  across all iterations for a particular level of under sampled images is equal to the mean across all iterations of fully sampled images. We see three categories of observations across all 14 mice. The observations are mentioned below in 3 categories along with the mouse number which has the same observations:

#### 1. FOR MICE 1,2,6,7,8,9,12 and 14: 8 mice in total

It has been observed that all p-values are less than 0.05 across all levels of  $T_2^*$  though it increases steadily as the level of sampling of MR images increases from 10% sampled to 90% sampled. The null hypothesis being tested here was if the mean of the iterations of a certain level of under sampling is equal the mean of the iterations of a certain level of fully sampled MR images. Thus, from the paired t-test we can conclude

that all levels  $T_2^*$  differences are significant that is the under-sampling is affecting the calculated  $T_2^*$ .

## 2. FOR Mice 3,5 ,10,11 and 13: 5 mice in total

It has been observed that 90% sampled MR images have a p-value that is greater than 0.05. Thus, it can be concluded that for 90% sampled MR image calculated  $T_2^*$  is not affected by the result of 10% under sampling whereas calculated  $T_2^*$  of all other level of sampling 10%-80% sampled images have been affected by the under sampling process. Thus the cutoff point here is 90%. It has also been observed that in most cases the p-value increases steadily as the level of sampling of MR images increases from 10% sampled to 90% sampled.

NOTE: Mice 3, 5 and 10 had 1 outlier each which have been ignored while carrying out the t-test.

## 3. FOR Mouse 4:

It has been observed that 80% sampled and 90% sampled MR images have a p-value that is greater than 0.05. Thus, it can be concluded that for 80% sampled MR image calculated  $T_2^*$  is not affected by the result of 10% under-sampling whereas calculated  $T_2^*$  of all other level of sampling 10%-80% sampled image have been affected by the under-sampling process. Thus the cutoff point here is 80%. It has also been observed that the p-values do not follow any trend like increasing or decreasing like the other mice.

## Summarized Results:

### Paired t Test ,Hotelling $t^2$ and ICC :

100%-%Sampled	1	2	3	4	5	6	7
100-10	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001
100-20	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001
100-30	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001
100-40	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001
100-50	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001
100-60	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001
100-70	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001
100-80	0.0008	<0.0001	0.0106	0.4046	<0.0001	<0.0001	<0.0001
100-90	0.0047	0.0003	0.1312	0.4596	0.0713	0.0001	0.0002

Table 11.4: A table demonstrating p-values of Paired t for Mice 1 - 7

100%-%Sampled	8	9	10	11	12	13	14
100-10	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001
100-20	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001
100-30	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001
100-40	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001
100-50	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001
100-60	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001
100-70	<0.0001	<0.0005	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001
100-80	0.0077	0.0004	0.0004	0.0012	0.0007	<0.0001	0.0055
100-90	0.0018	0.0358	0.1058	0.2123	<0.0001	0.3541	0.0271

Table 11.5: A table demonstrating p-values of Paired t for Mice 7 - 14

100%-%Sampled	8	9	10	11	12	13	14
100-10	0.66	0.63	0.59	0.64	0.65	0.17	0.63
100-20	0.62	0.59	0.55	0.60	0.58	0.18	0.61
100-30	0.58	0.53	0.48	0.56	0.52	0.18	0.56
100-40	0.49	0.46	0.39	0.49	0.40	0.45	0.51
100-50	0.43	0.41	0.33	0.09	0.32	0.34	0.46
100-60	0.36	0.34	0.23	0.06	0.25	0.27	0.42
100-70	0.28	0.25	0.22	0.04	0.20	0.24	0.38
100-80	0.28	0.21	0.17	0.02	0.18	0.16	0.37
100-90	0.27	0.21	0.16	0.01	0.17	0.15	0.34

Table 11.6: A table showing ICC for Mice 1 - 7

100%-%Sampled	8	9	10	11	12	13	14
100-10	0.62	0.67	0.64	0.68	0.67	0.7	0.67
100-20	0.56	0.59	0.54	0.61	0.62	0.67	0.62
100-30	0.5	0.54	0.47	0.55	0.57	0.6	0.58
100-40	0.43	0.46	0.46	0.48	0.55	0.53	0.53
100-50	0.39	0.40	0.39	0.44	0.48	0.49	0.47
100-60	0.36	0.38	0.38	0.39	0.50	0.43	0.43
100-70	0.34	0.34	0.36	0.34	0.47	0.41	0.44
100-80	0.33	0.34	0.35	0.32	0.52	0.41	0.43
100-90	0.32	0.33	0.33	0.31	0.50	0.43	0.46

Table 11.7: A table showing ICC for Mice 7-14

Hotelling  $t^2$  p-values were all  $<0.05$ .

**NOTE:**

It is observed that the ICCs (Intraclass Correlation) across the iterations are not as low as that compared to the Phantom simulation dataset. In fact in most cases except the difference of 100% sampled and 90% sampled, the ICC is much greater than 0. This suggests that the iterations are much homogenous in nature compared to the ones in Phantom simulations. Thus, we can see that paired t test is working better for the real mice datasets compared to the Hotelling  $t^2$  test since there is not much variability to capture.

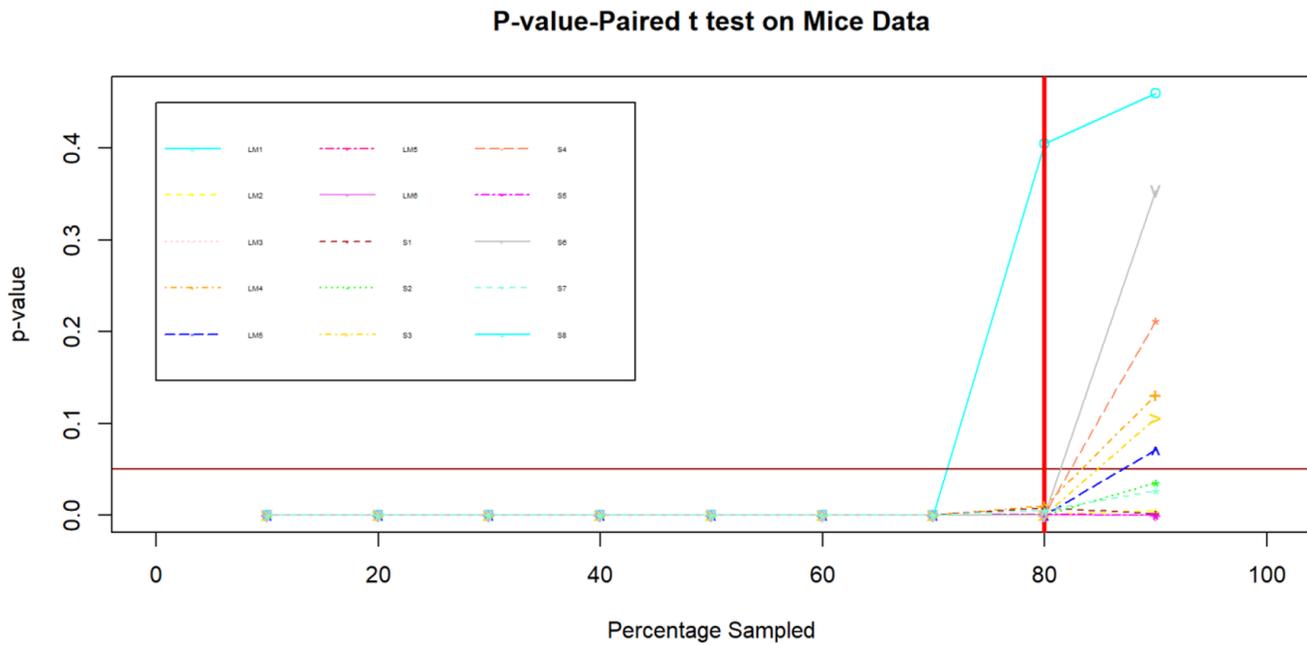


Figure 11.2: Visualization of pvalue of Paired t test

NOTE: The horizontal red line represents the 5% significance level. All the points plotted above it give nonsignificant results. The vertical red line shows the minimum level of under-sampling which is giving non-significant result.

We see for 6 mice the points are above the red line.

Out of those 6, plots for 5 mice (grey, blue, yellow, orange and tangerine-colored plots) show that 90% sampled doesn't affect calculation of  $T_2^*$ .

Lastly, one mouse (the cyan-colored plot) shows 80% sampled or 20% under sampled doesn't affect calculations of  $T_2^*$ .

# Chapter 12

## Excursis and Future Work:

### Excursis:

#### Phantom Dataset:

From the results we see in the results chapter(Chapter 11) we can go ahead and conclude the following for the Phantom dataset:

1. For  $64^3$  voxel MRIs, it can be said that for upto only 20% sampled images that are 80% undersampled MRIs the metric in context,  $T_2^*$  is not significantly affected by the under-sampling process. That is a very interesting result. under-sampling upto 80% can bring down the cost and time significantly.
2. For  $96^3$  voxel MRIs it can be said that for upto 50% sampled images that are 50% undersampled MRIs the metric in context  $T_2^*$  is not significantly affected by the under-sampling process.
3. For  $128^3$  voxel MRIs it can be said that for upto 80% sampled images that are 20% undersampled MRIs the metric in context  $T_2^*$  is not significantly affected by the under-sampling process.

In conclusion, we do observe that the significant level of under-sampling

is inversely proportional to the number of voxels in the MRI that is as the number of voxels increases the level of significant under-sampling decreases.

**NOTE:**

As observed before it is very important to remember that since we see that ICC indicates very less correlations between the iterations of the Phantom data so to capture the vast variability Hotelling  $t^2$  test works better here compared to paired t-test. So all the above conclusions have been drawn by observing the p-values derived from Hotelling  $t^2$  test.

### **Vivo Real Mice Dataset:**

From the results we see in the results chapter we can go ahead and conclude the following for the Vivo real mice dataset:

1. For mice 1,2,6,7,8,9,12 and 14(8 mice) we see from the p-values of the paired t-test that none of the levels of under-sampling gives as significant results that are in conclusion for these mice under-sampling even 10% is affecting the  $T_2^*$  results.
2. For mice 3,5,10,11 and 13(5 mice) we can conclude the cutoff point here is 90% sampled that is only 10% under-sampling do not significantly affect the parameter of interest  $T_2^*$ .
3. For mice 4 it can be concluded that for 80% sampled MR image calculated  $T_2^*$  is not affected by the result of 20% under-sampling. Thus in conclusion taking all the events into account it can be said that in most cases 20% under-sampling can be carried out.

## NOTE:

As observed before it is very important to remember that since we see that ICC indicates considerably higher correlations between the iterations of the vivo real mice data, especially when compared to the Phantom data there is not much variability among the iterations to capture so paired t test works better than Hotelling  $t^2$  test. So all the above conclusions have been drawn by observing the p-values derived from Hotelling  $t^2$  test.

## **Summarized Conclusion: From both Phantom Dataset and Vivo Real Mice Dataset**

Keeping in mind that the vivo real mice dataset MRI has been conducted in a  $128^3$  setup we will be comparing the results of  $128^3$  voxel of the Phantom dataset with the vivo real mice dataset. **We observed that even if the results have been obtained in two different test setups due to the observed ICC but the results are synonymous that is in both cases 80% sampled that is same as 20% under sampling has been obtained as the cutoff point. In other words, 20% under sampling does not disrupt the information obtained by the parameter of interest which is  $T_2^*$**

## **Future Work:**

There are quite a few scopes for future work in this project.

1. To get a much better understanding of the cessation point the same experiment can be carried out for MRI with isotropic sizes  $64^3$  and  $96^3$ .
2. Doing the same will also allow us to venture into the relationship between the voxel sizes and the cessation point of the under-sampling.
3. Other approaches based on distance can be used especially on a voxel by voxel level can be carried out at different levels of under-sampling to find the significant level of under-sampling.
4. The study can be extended to human MRI which will be great in terms of optimization of time and expense.

# Chapter 13

## References:

1. Estimation approaches for ADC in Rice-Distributed MR Images [Stefano Baraldo , Francesca Ieva ,Luca Mairandi and Anna Maria Paganoni]
2. Noise in Magnitude Magnetic Resonance Images - ARTURO CA´RDENAS-BLANCO (Ottawa Health Research Institute, Ottawa, ON, Canada) CRISTIAN TEJOS PABLO IRARRAZAVAL(Department of Electrical Engineering, Pontificia Universidad Cato´lica de Chile, Santiago, Chile) IAN CAMERON( Biomedical Imaging Center, Pontificia Universidad Cato´lica de Chile, Santiago, Chile Department of Physics, Carleton University, Ottawa, ON, Canada)
3. Sampling Density Compensation in MRI: Rationale and an Iterative Numerical Solution - James G. Pipe and Padmanabhan Menon
4. Improving hyperpolarized  $^{129}\text{Xe}$  ADC mapping in pediatric and adult lungs with uncertainty propagation – Abdullah S. Bdaiwa, Peter J. Niedbalski, Md M. Hossain, Mathew M. Willmering, Laura L. Walkup, Hui Wang, Robert P. Thomen, Kai Ruppert, Jason C. Woods, Zackary I. Cleveland.
5. Manuscript – “Impact of under-sampling on Preclinical Lung  $T_2^*$  Mapping with 3D Radial UTE MRI at 7T” - Ian R. Stecker a, b1, Abdullah Bdaiwi a, b, Peter J. Niedbalski b2 , and Zackary I. Cleveland a, b,c\*

a Department of Biomedical Engineering, University of Cincinnati, Cincinnati, OH 45221

b Center for Pulmonary Imaging Research, Cincinnati Children's Hospital Medical Center, Cincinnati, OH 45229

c Department of Pediatrics, University of Cincinnati, Cincinnati, OH 45221

6. Image Quality Statistics and their use in Steganalysis and Compression by

İsmail Avcıbaşı B.S. in E.E., Uludağ University, 1992

M.S. in E.E., Uludağ University, 1994

Submitted to the Institute for Graduate Studies in Science and Engineering in partial fulfillment of the requirements for the degree of Doctor of Philosophy

Boğaziçi University 2001

7. Estimation of optimal b-value sets for obtaining apparent diffusion coefficient free from perfusion in non-small cell lung cancer.

Kishor Karki, Geoffrey D. Hugo, John C. Ford, Kathryn M. Olsen, Sidharth Saraiya, Robert Groves, and Elisabeth Weiss

Phys Med Biol. Author manuscript; available in PMC 2016 Oct 21.