COMPARISON OF MEGA-PRESS AND SHORT ECHO TIME PRESS ON CLASSIFICATION OF IDH MUTATION USING MACHINE LEARNING AT 3T

by

Ayhan Gürsan

B.Sc., Electronics and Comm. Engineering, Istanbul Technical University, 2016

Submitted to the Institute of Biomedical Engineering
in partial fulfillment of the requirements
for the degree of
Master of Science
in
Biomedical Engineering

Boğaziçi University
2019

COMPARISON OF MEGA-PRESS AND SHORT ECHO TIME PRESS ON CLASSIFICATION OF IDH MUTATION USING MACHINE LEARNING AT 3T

APPROVED BY:

Assoc. Prof. Dr. Esin Öztürk Işık .....
(Thesis Advisor)

Prof. Dr. Alp Dinçer .....

Assoc. Prof. Dr. Bora Garipcan .....

DATE OF APPROVAL: 23 July 2019

ACKNOWLEDGMENTS

First of all, I would like take this opportunity to thank Assoc. Prof. Dr. Esin Öztürk Işık for her support, generous guidance and patience. She played a great role in my career and education. I also would like to thank Prof. Dr. Alp Dinçer for his generous help and teachings not only in this project but also on clinical side of the MRI/S world. In addition, I should also thank Assoc. Prof. Dr. Bora Garipcan and Burak Altun for their assistance on MRS phantom preparations.

I would like to thank my colleagues for the time we share in our laboratory. Gökçe Hale Hatay and Ozan Genç supported me with their knowledge in this project. Also, Dilek Betül Arslan and Başak Bayrambaş showed their kindness with their friendship in tough times.

Finally, I would like to thank my family. They always supported and encouraged me from the beginning of this journey.

This study has been funded by TÜBİTAK project #216S432.

ACADEMIC ETHICS AND INTEGRITY STATEMENT

I, Ayhan Gürsan, hereby certify that I am aware of the Academic Ethics and Integrity Policy issued by the Council of Higher Education (YÖK) and I fully acknowledge all the consequences due to its violation by plagiarism or any other way.

Name :

Signature:

Date:

ABSTRACT

COMPARISON OF MEGA-PRESS AND SHORT ECHO TIME PRESS ON CLASSIFICATION OF IDH MUTATION USING MACHINE LEARNING AT 3T

Malignant glioma is a type of frequent and lethal cancer the brain. Recent World Health Organization (WHO) criteria has included genetic mutations in glioma classification. One of these mutations, isocitrate dehydrogenase (IDH) is common in grades II and III gliomas, and has been related to metabolism of the cancer tissue. IDH mutant gliomas have better prognosis than IDH wild type ones. As a result of this mutation, an onco-metabolite 2-HydroxyGlutarate (2HG) accumulates in tumor tissue. Detection of IDH mutation before surgical procedure could play an important role in treatment planning. Magnetic resonance spectroscopy (MRS) is a noninvasive technique that could be used to provide IDH mutation information. In this study, first, a 3D printed MRS phantom was designed and produced to analyze spatial distribution performances of MRS sequences. Then, 82 glioma patients, whose IDH status have been determined by immunohistochemistry, have been included. Short echo time Point Resolved Spectroscopy (PRESS) and Mescher-Garwood PRESS (MEGA-PRESS) MRS sequences were acquired on a 3T Siemens MRI scanner. Metabolite concentrations have been estimated with LCModel spectal fitting program using corresponding basis sets. Machine learning models have been developed to determine IDH mutation using metabolite concentrations as features. Our results indicated that a decision tree model using features from short TE PRESS profile could detect IDH mutation with 75% accuracy, while maximum accuracy attainable with MEGA-PRESS was 68%. The MRS phantom that was produced as a part of this study could be used as a validation tool for new MRS sequences. Future studies will aim to detect other genetic alterations in gliomas on a larger patient cohort.

Keywords: glioma, IDH, magnetic resonance spectroscopy, machine learning.

ÖZET

MAKİNE ÖĞRENME KULLANARAK IDH MUTASYONUNUN SINIFLANDIRILMASINDA MEGA-PRESS VE KISA EKO ZAMANLI PRESS'İN 3T'DA KARŞILAŞTIRILMASI

Malign gliom, beyinde sık görülen ve ölümcül bir kanser türüdür. Son Dünya Sağlık Örgütü (WHO) kriterleri, gliom sınıflandırmasına genetik mutasyonları dahil etmiştir. Bu mutasyonlardan biri, izositrat dehidrogenaz (IDH), derece II ve derece III gliomlarda yaygın olup, kanser dokusunun metabolizması ile ilgilidir. IDH mutant gliomlar, IDH vahşi tiplerinden daha iyi prognoza sahiptir. Bu mutasyonun bir sonucu olarak onko-metabolit 2-HidroksiGlutarat (2HG), tümör dokusunda birikir. IDH mutasyonunun cerrahi prosedürden önce tespiti tedavi planlamasında önemli bir rol oynayabilir. Manyetik rezonans spektroskopi (MRS), IDH mutasyon bilgisi sağlamak için kullanılabilecek noninvaziv bir tekniktir.Bu çalışmada, ilk olarak, 3 boyutlu basılmış bir MRS fantom, MRS sekanslarının uzaysal dağılım performanslarını analiz etmek için tasarlanmış ve üretilmiştir. Daha sonra çalışmaya, IDH durumu immünohistokimya ile saptanmış 82 gliom hastası dahil edilmiştir. Kısa eko zamanlı Nokta Çözünürlüklü Spektroskopi (PRESS) ve Mescher-Garwood PRESS (MEGA-PRESS) MRS dizileri, 3T Siemens MRG tarayıcısında elde edilmiştir. IDH mutasyonu tespitinde makine öğrenme modellerinde öznitelik olarak metabolit konsantrasyonları kullanılmıştır. Sonuçlarımız, kısa TE PRESS profilindeki öznitelikleri kullanan karar ağacı modelinin, IDH mutasyonunu % 75 doğrulukla tespit edebileceğini, MEGA-PRESS ile elde edilebilecek maksimum doğruluğun ise % 68 olduğunu göstermiştir. Bu çalışmada üretilen MRS fantomu, yeni MRS darbe sekansları için bir doğrulama aracı olarak kullanılabilir. Gelecekteki çalışmalarımız gliomlardaki diğer genetik değişiklikleri daha büyük bir hasta populasyonunda tespit etmeyi hedeflemektedir.

Anahtar Sözcükler: gliom, IDH, manyetik rezonans spektroskopi, makine öğrenme.

TABLE OF CONTENTS

ACKNOWLEDGMENTS . . . iii
ACADEMIC ETHICS AND INTEGRITY STATEMENT . . . iv
ABSTRACT . . . v
ÖZET . . . vi
LIST OF FIGURES . . . ix
LIST OF TABLES . . . xi
LIST OF ABBREVIATIONS . . . xii
1. INTRODUCTION . . . 1
2. BACKGROUND . . . 3
2.1 Gliomas and IDH Mutation . . . 3
2.2 Magnetic Resonance Spectroscopy . . . 3
2.2.1 Point Resolved Spectroscopy . . . 6
2.2.2 Spectral Editing Using Mescher-Garwood PRESS . . . 7
2.2.3 Detection of 2HG . . . 8
2.3 Machine Learning Classification . . . 9
2.3.1 k-Nearest Neighbor Classifier . . . 10
2.3.2 Decision Trees . . . 11
2.3.3 Support Vector Machine . . . 12
2.3.4 Ensemble Methods . . . 13
2.4 Literature Review . . . 13
3. MATERIALS and METHODS . . . 15
3.1 MRS Phantom . . . 15
3.2 Data Acquisition . . . 18
3.2.1 MRS Data Acquisition . . . 18
3.2.2 Histopathological Tests . . . 19
3.3 MRS Data Analysis . . . 19
3.4 Machine Learning Models . . . 22
3.5 Statistical Analysis . . . 22
4. RESULTS . . . 23

4.1 MRS Phantom Results 23
4.2 Comparison of Metabolite Levels Between MEGA-PRESS and Short TE PRESS 27
4.3 Machine Learning Results 28
5. DISCUSSION 32
6. CONCLUSION 35
7. List of publications produced from the thesis 36
APPENDIX A. Phantom solution ingredients 37
REFERENCES 38

LIST OF FIGURES

Figure 2.1	2-HydroxyGlutarate accumulates in cytoplasm or mitochondria in IDH1 and IDH2 mutant gliomas, respectively [1].	4
Figure 2.2	An example of healthy brain MR spectrum (left) and voxel placement on T2 weighted MR images at three orthogonal planes (right).	5
Figure 2.3	RF pulse scheme of PRESS with corresponding gradients [2].	6
Figure 2.4	MEGA-PRESS pulse sequence scheme.	8
Figure 2.5	Simulated spectra of 2-HG and other metabolites that have overlapping peaks at similar frequencies [3].	9
Figure 2.6	A demonstrative decision boundary in kNN algorithm for binary classification.	10
Figure 2.7	Decision boundaries of a univariate tree on feature space (left) and corresponding decision tree scheme (right) [4].	11
Figure 2.8	An example of hyperplane separation in binary classification by support vectors [4].	12
Figure 3.1	Top (a), frontal (b) and axial MRI (c) views of the 3D phantom design.	17
Figure 3.2	LCModel analysis of MEGA-PRESS difference (a) and short TE PRESS (c) spectra and voxel placements (b, d) for an IDH-mut grade III glioma.	21
Figure 4.1	LCModel analysis results of the solution prepared in a vial.	23
Figure 4.2	Grid placement of multi-voxel PRESS acquisition (a) LCModel analysis of IDH-mut (b) and IDH-wt (c) compartments. Metabolite maps for Cho (d), Cr (e) and NAA (f).	25
Figure 4.3	The grid placement of MEGA-LASER acquisition (a), 2HG map generated from difference spectrum (b) and LCModel analysis of difference spectrum taken from IDH-mut (c), IDH-wt (d) a healthy brain (e) compartments.	26
Figure 4.4	Bland-Altman scatter plot of Cho/Cr measurements.	28
Figure 4.5	Bland-Altman scatter plot of NAA/Cr measurements.	29

Figure 4.6 Bland-Altman scatter plot of Glx/Cr measurements. 30

LIST OF TABLES

Table 3.1	Metabolite concentrations in millimolar for each prepared solution.	16
Table 4.1	The differences of actual concentration values of the solutions and LCModel results.	24
Table 4.2	Bland-Altman statistical test results for Cho/Cr, NAA/Cr and Glx/Cr metabolites between short TE PRESS and MEGA-PRESS.	27
Table 4.3	The classification accuracy results obtained from MRS profile of short TE PRESS with zero filling.	28
Table 4.4	The classification accuracy results obtained from MRS profile of MEGA-PRESS with zero filling.	29
Table 4.5	The classification accuracy results obtained from MRS profile of short TE PRESS with NaN filling.	30
Table 4.6	The classification accuracy results obtained from MRS profile of MEGA-PRESS with NaN filling.	31
Table A.1	Names, molecular weight, concentrations in millimolar and CAS numbers for the chemicals used in the phantom solutions.	37

LIST OF ABBREVIATIONS

2HG	2-HydroxyGlutarate
Bagging	Bootstrap Aggregating
CHESS	Chemical Shift Selective Saturation
Cho	Choline
Cr	Creatine
CRLB	Cramer-Rao Lower Bound
CSF	Cerebrospinal Fluid
FID	Free Induction Decay
FWHM	Full Width at Half Max
GABA	γ -Amino Butyric Acid
GAMMA	General Approach to Magnetic Resonance Mathematical Analysis
Gln	Glutamine
Glu	Glutamate
Glx	Glutamate + Glutamine
Glyc	Glycine
GSH	Glutathione
IDH	Isocitrate dehydrogenase
IDH-mut	IDH Mutant
IDH-wt	IDH wildtype
IHC	Immunohistochemistry
Ins	Myo-Inositol
kNN	k-Nearest Neighbor
Lac	Lactate
MEGA	Mescher-Garwood
MEGA-LASER	Mescher Garwood Localization by Adiabatic Refocusing
MRI	Magnetic Resonance Imaging
MRS	Magnetic resonance spectroscopy
MRSI	Magnetic resonance spectroscopic Imaging

NAA	N-acetyl-aspartate
NaN	Not a Number
PRESS	Point Resolved Spectroscopy
RF	Radio Frequency
SNR	Signal-to-Noise Ratio
SVM	Support Vector Machine
TE	Echo Time
TERT	Telomerase Reverse Transcriptase
VESPA	Versatile Simulation, Pulses and Analysis
WHO	World Health Organization

1. INTRODUCTION

Brain tumors could be identified as abnormal cell proliferation or a mass in the brain. There are many types of brain tumors and these types could be classified in different aspects. Malignant tumors tend to grow and spread over healthy tissues. Primary brain tumors originate at the brain tissue, while secondary tumors grow first in another part of the body, and then spread to the brain tissue. Primary brain tumors are named after the cell type that tumors originate. Two of the primary tumors with highest occurrence rate are gliomas and meningiomas, that originate the from glial cells and brain membrane, respectively [5, 6].

Malignant glioma is one of the most frequent and lethal cancer that occurs in the central nervous system. Gliomas account for almost 30% of all primary brain tumors, and 80% of all malignant ones, and are responsible for the majority of deaths from primary brain tumors [6]. Recent World Health Organization (WHO) criteria has included genetic mutations in glioma classification [7]. Isocitrate dehydrogenase (IDH) gene has been indicated as an important marker for brain tumor metabolism. IDH mutations are common in grades II and III gliomas [8]. IDH mutant (IDH-mut) glioma patients have better prognosis than IDH wild type (IDH-wt) [9]. The current gold standard for identification of IDH mutation is biopsy. As a result of this mutation, an onco-metabolite called 2-HydroxyGlutarate (2HG), accumulates in tumor tissue. Magnetic resonance spectroscopy (MRS) is a noninvasive technique that could be used to measure the concentrations of different chemical compounds within the tissue of interest. Different MRS techniques, such as short echo time (TE) Point Resolved Spectroscopy (PRESS) and Mescher-Garwood PRESS (MEGA-PRESS), have been developed and compared in research centers for detecting 2HG in IDH-mut gliomas, but the cohort sizes were relatively small compared to the clinical studies [10–13]. 2HG has a very similar chemical structure with other abundant metabolites in the brain, which makes 2HG detection challenging even with specially developed sequences. Machine learning methods have recently been developed for IDH mutation classification using

radiomics and multimodal MRI features [14–17], but to our knowledge, there has not been any studies that compares the performance of short TE PRESS and MEGA-PRESS for IDH mutation classification using machine learning.

A preoperative noninvasive detection method for IDH mutation could be helpful for personalized treatment planning. The main purpose of this thesis is to investigate the relative performances of a short TE PRESS and MEGA-PRESS MRS sequences for noninvasive preoperative detection of IDH mutation in glioma patients using machine learning methods.

2. BACKGROUND

2.1 Gliomas and IDH Mutation

Gliomas, which originate from the glial cells that are the support cells for neurons, occur in the brain and spinal cord of the central nervous system. Gliomas are classified into 4 grades based on WHO criteria, and 3 types on their histological appearance as astrocytic, oligodendroglial, and ependymal [9, 18]. These classifications strongly impact clinical decisions in treatment planning. In 2016, WHO classification of gliomas included genetic mutations such as IDH and 1p/19q codeletion. The necessity of this addition comes from those cases, for which the genotype comes over histological phenotype [19]. Magnetic resonance imaging (MRI) is mainly used for grading gliomas in the clinics, yet histopathological tests, applied on resected tissue, give detailed information on tumor specifications [19].

Isocitrate dehydrogenase is an enzyme in cytoplasm (IDH1) and mitochondria (IDH2), which catalyzes the conversion of isocitrate to a-ketoglutarate [20]. Mutations on IDH genes have been frequently seen in low grade (II-III) gliomas [21]. These mutations alter enzymatic activities of the tumor cells, and as a result of this alteration, 2HG accumulates to levels that could detected by MRS sequences (Figure 2.1). Importance of detecting IDH mutation comes from its effect on prognosis of the patients. Median overall survival has been reported to be statistically significantly higher in IDH-mut gliomas than IDH-wt ones [21, 22].

2.2 Magnetic Resonance Spectroscopy

MRS is a noninvasive technique that could be used to measure the concentrations of different chemical compounds within the tissue. When patient is inside an MRI scanner, protons with different chemical structure and different neighboring atoms ex-

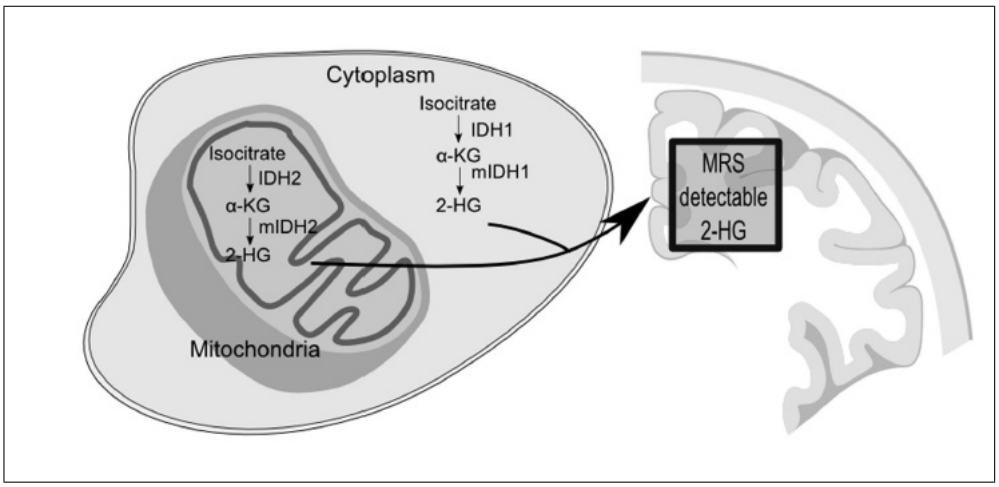

Figure 2.1 2-HydroxyGlutarate accumulates in cytoplasm or mitochondria in IDH1 and IDH2 mutant gliomas, respectively [1].

perience different amounts of magnetic field, which causes shifts in their precessional frequencies. Protons in molecules are excited with a radio frequency (RF) electromagnetic pulse just like in MRI, and after a certain time, which is determined by timings in pulse scheme, signal emitted from protons is acquired. This signal is sinusoidal with an exponential decay and is called a free induction decay (FID). An appropriate acquisition time should be selected for application, since signal loss will occur in time. The beginning of data acquisition time is called the echo time (TE). Fourier Transform is applied to the MR signal to get a spectrum, and each proton group is represented with a peak at a certain frequency, which corresponds to their precessional frequency. MRS could be applied for a single-voxel or multi-voxel imaging (MRSI). Since water concentration is extremely high compared to other molecules, water suppression techniques must be used to get useful information from an MR spectrum. Figure 2.2 shows an example MR spectrum for healthy brain acquired at TE=68 ms (left) and placement of the voxel on the corresponding T2 weighted MR images at 3T.

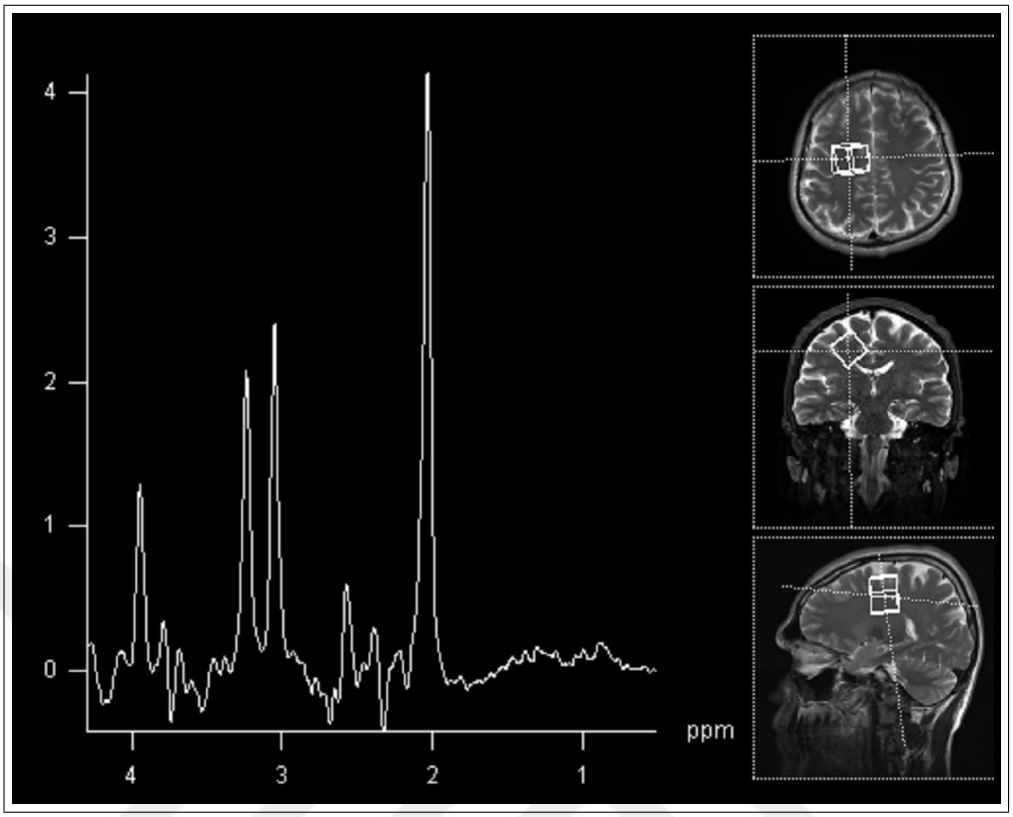

Figure 2.2 An example of healthy brain MR spectrum (left) and voxel placement on T2 weighted MR images at three orthogonal planes (right).

Proton MRS acquired from healthy human brain consists of three major peaks. Intensity values of these peaks may be altered with diseases. N-acetyl-aspartate (NAA) is the most abundant metabolite with a peak at 2 ppm. Concentration of NAA provides information on neuronal density of the voxel [23]. Energy metabolism of the tissues in volume of interest is inspected with creatine (Cr), which resonates at 3.0 and 3.8 ppm. Cr is also commonly used as a reference for relative quantification of other metabolites [24]. Another major metabolite in the spectrum is choline (Cho) with 9 resonating protons represented on a single peak at 3.2 ppm. Especially, brain tumor studies have an interest in Cho concentrations as Cho elevation is associated with an increase in cellular density [25]. There are a few other metabolites in a brain MR spectrum to be mentioned, such as glutamate (Glu) and glutamine (Gln). Because of their very similar structures, these two metabolites often are reported as Glu+Gln (Glx). While Glu is known to be an excitatory neurotransmitter, Gln works as a storage form of Glu in astrocytes [26]. Myo-Inositol (Ins) is another metabolite, which

is believed to play a role in cell growth and storage of glucose [26]. Glycine (Glyc) acts as an inhibitory neurotransmitter. Elevation in Glyc has been reported in brain tumors [27]. Glutathione (GSH) is a metabolite with low concentration that is present in astrocytes, and decreasing levels of GSH has been associated with cell death [28]. Lactate (Lac) peak could also be seen in brain tumors, which accumulates in tumor site as an end product of anaerobic glycolysis [25].

2.2.1 Point Resolved Spectroscopy

PRESS is one of the common pulse sequences used in MRS. In PRESS sequence, protons are excited with 90°-180°-180° slice selective pulses using simultaneous magnetic field gradients to create 3 orthogonal intersecting planes [29]. A stimulated echo is generated as a result of these three slice selective RF pulses at TE, which is determined by time intervals between RF pulses. Chemical shift selective saturation (CHESS) pulses are used before PRESS sequence block to suppress water signal [30]. PRESS is widely used because of its high signal-to-noise ratio (SNR) compared to other MRS techniques. Figure 2.3 shows RF pulse schemes along with gradients for PRESS.

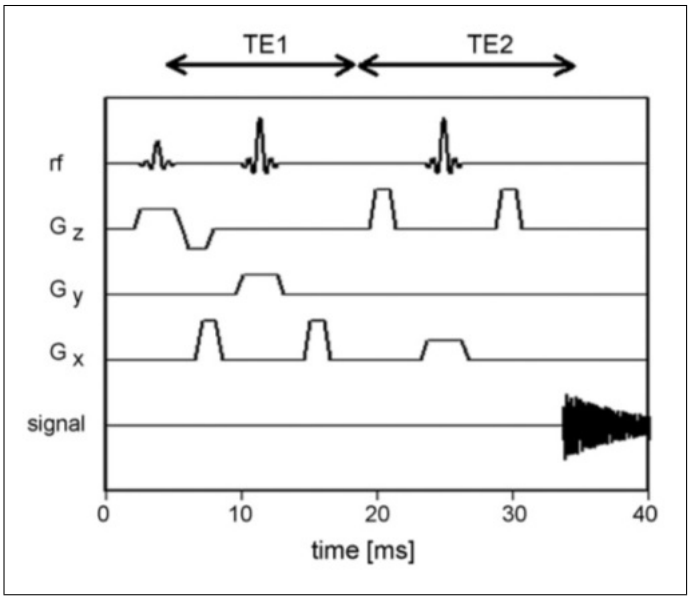

Figure 2.3 RF pulse scheme of PRESS with corresponding gradients [2].

2.2.2 Spectral Editing Using Mescher-Garwood PRESS

Mescher-Garwood (MEGA) is a frequency selective refocusing technique that could be placed into MRS sequences such as PRESS. In addition to slice selective pulses, MEGA employs frequency selective refocusing pulses to dephase transverse magnetization. Since MEGA water suppression relies on two 180° Gaussian shaped pulses, it could be also used for editing J-coupled spin systems. J-coupling results in peak splitting and changes line shape and signal amplitude with TE. If these refocusing pulses are double banded, both water suppression and spectral editing could be done simultaneously [31]. MEGA-PRESS data acquisition is done sequentially for ON and OFF editing pairs. At ON stage, refocusing pulse is applied and at OFF stage refocusing pulse is applied to its frequency symmetric with respect to the water frequency for simplicity of sequence programming. As a result of the refocusing pulses, only J-coupled systems that will have a peak at the chosen frequency will be affected. Spin evolution of these coupled systems will be prevented when the refocusing pulse is ON. After both acquisitions are done, OFF spectra will be subtracted from ON spectra to get difference spectra. Since metabolites that are not affected with refocusing pulse will have the same peaks on both spectra, they will cancel out each other and only metabolites that have peaks at the refocusing frequency will be observed in the difference spectra. MEGA-PRESS has been used for detection of γ -amino butyric acid (GABA), which is an inhibitory neurotransmitter that has overlapping peaks with Glu and Gln [31–33]. A MEGA-PRESS pulse sequence scheme is given in Figure 2.4.

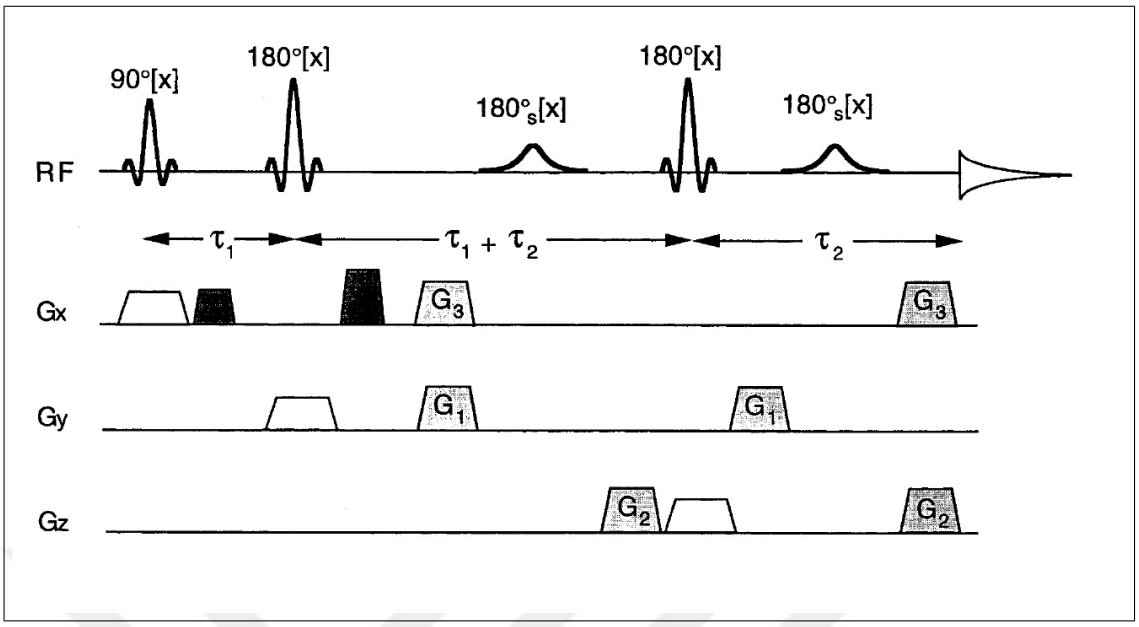

Figure 2.4 MEGA-PRESS pulse sequence scheme.

2.2.3 Detection of 2HG

2HG molecule is a 5-spin system with J-coupling interactions. 2HG protons give rise to peaks located at 4.02, 2.25 and 1.9 ppm (representing 1, 2 and 2 protons respectively) [34]. All peaks of 2HG overlap with other metabolite peaks, such as NAA, Cr, Gln, Glu and GABA [35]. Moreover, concentration levels of these overlapping metabolites, such as Glu, also alter with IDH mutation status of the tumor [36]. Figure 2.5 shows some of the overlapping metabolites, which has been reported to be altering with IDH mutation. These overlaps make detection of 2HG challenging.

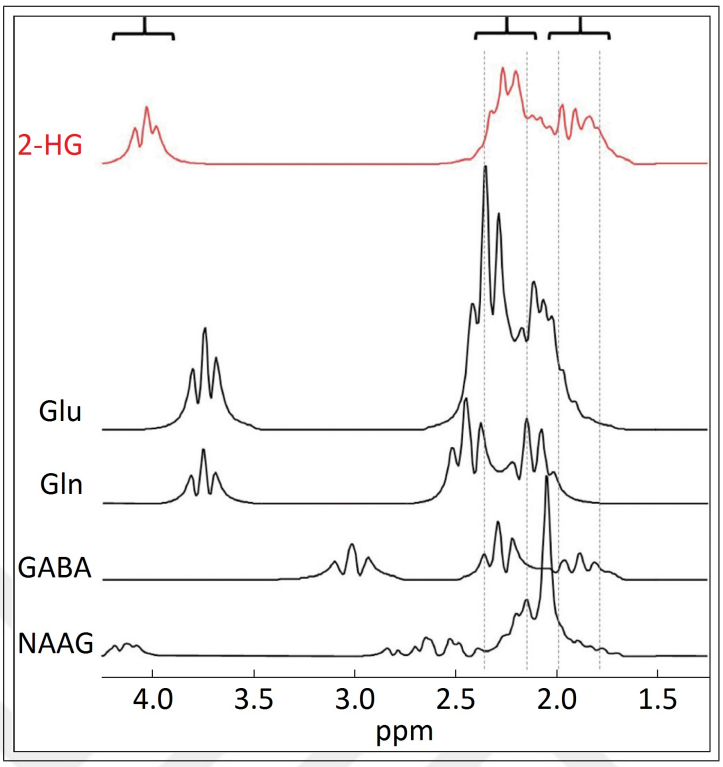

Figure 2.5 Simulated spectra of 2-HG and other metabolites that have overlapping peaks at similar frequencies [3].

.

2.3 Machine Learning Classification

Artificial intelligence methods such as machine learning, have been of interest for the last decade, with increased availability of data for data driven approaches. Machine learning methods need substantial amount of well-structured data to learn [4]. This data actually consists of features extracted from observations. The features help machine learning models by numerically quantifying a specialty of the sample. Machine learning methods could be divided into two parts depending on the task. If the samples have two or more classes with provided class labels, then supervised machine learning methods could be applied. On the other hand, if the samples only consists of features without a group label, then unsupervised machine learning methods could be used to cluster similar samples. Classification tasks is one of the most popular study areas for supervised machine learning methods.

Models learn from features of the training set samples and predict class of the samples in the test set. If the data is not big enough to separate as training and test sets, then other techniques, such as k-fold cross validation, could be used. All data is shuffled and divided to subgroups consisting of k samples. Last sample of these group is used for testing prediction accuracy. Different classification models were tested in this study with k-fold cross validation.

2.3.1 k-Nearest Neighbor Classifier

k-Nearest Neighbor (kNN) is one of the first models that is tried on a new dataset in machine learning, because of its simplicity. kNN is a non-parametric method and assumes that same classes should have similar features, which could be translated as samples from same classes should be close to each other in the feature space [4]. Predictions are made by calculating the distances between test sample and all samples in the feature space. Figure 2.6 shows an example decision boundary in kNN algorithm for binary classification. Euclidian and Minkowski distances could be used for continuous features. Classification is made by nearest samples to the test sample in a voting way. Hyperparameter k stands for how many similar samples in training set would be used in voting. Setting an odd number for k could help avoiding ties in voting. kNN method comes with a drawback as it is computationally expensive in each prediction, and requires large amount of memory to store all training set for calculating distances.

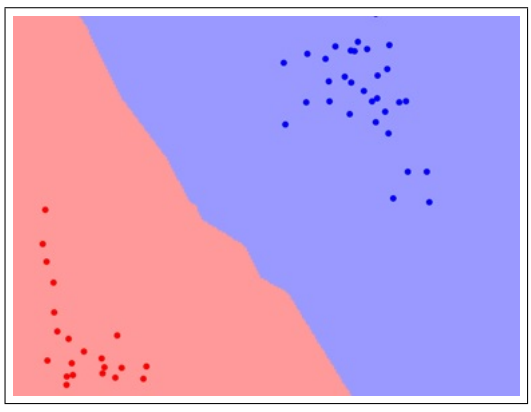

Figure 2.6 A demonstrative decision boundary in kNN algorithm for binary classification.

2.3.2 Decision Trees

Decision tree is a nonparametric supervised machine learning model that is widely used in classification. A hierarchical model, which consists of decision nodes and terminal leaves, is used to divide data into smaller groups with recursive splits [4]. Decision nodes discriminate samples based on features evaluating an impurity measure and lead to branches until a leaf is reached. These leaves represent one of the class labels and give the output of the classifier. Univariate trees could be given as a simple example of tree structures in machine learning. These trees use only one feature in each decision node, while proceeding from roots to leaves by branches. In binary split, decision nodes compare the input with a threshold value and split based on outcome of this comparison. Figure 2.7 gives an example of univariate tree decision boundaries on feature space (left) with the corresponding hierarchical model of the tree (right).

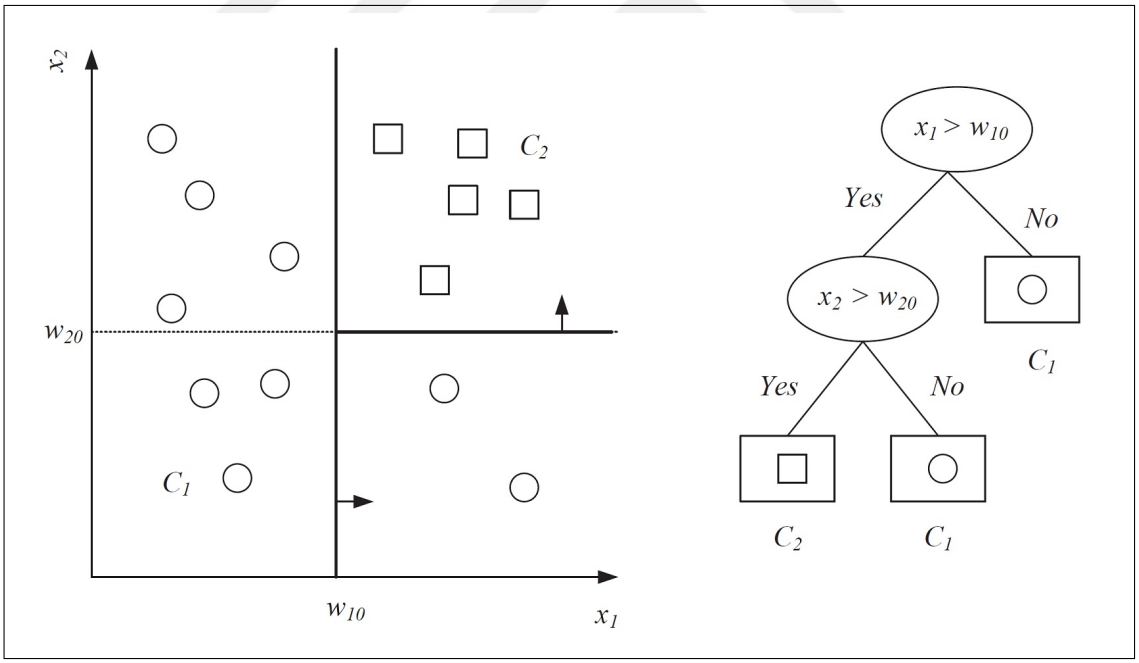

Figure 2.7 Decision boundaries of a univariate tree on feature space (left) and corresponding decision tree scheme (right) [4].

2.3.3 Support Vector Machine

Support vector machine (SVM) is a discriminant based kernel method that constructs hyperplanes, which utilizes margins on boundaries [4]. As a kernel-based algorithm, SVMs have single optimum solution that could be found in a convex optimization. This has a great advantage, since learning procedure does not need hyperparameters such as learning rate, initialization, and convergence as in numerical methods. First, a linear model is constructed and class boundaries are found with a margin. Support vectors are extracted from feature space of the training set, and hyperplanes are constructed to divide each class. Those samples that sit close to the boundaries between classes are uncertain cases and give an estimate of generalization error. Figure 2.8 gives an example of SVM as encircled support vectors separate hyperplanes for binary classification.

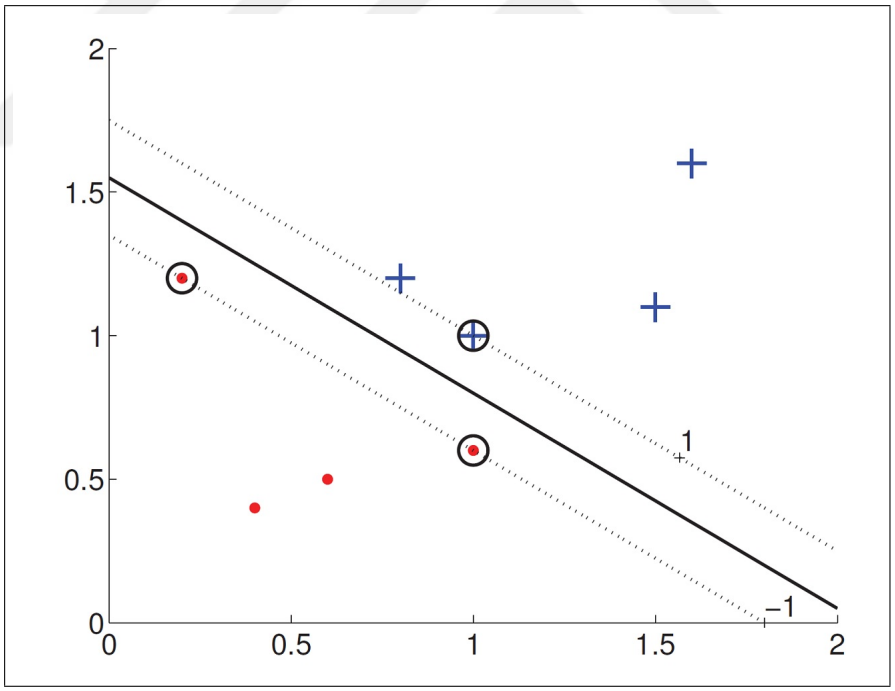

Figure 2.8 An example of hyperplane separation in binary classification by support vectors [4].

2.3.4 Ensemble Methods

Some of the machine learning methods that are widely used have been introduced previously. These methods are generally successful on most studies. Nevertheless, combination of these methods in the same classifier, like parametrics with non-parametrics or in itself, may also be used to boost accuracy, sensitivity and specificity of the model. This approach is named as ensemble methods in machine learning [4]. The simplest way of ensemble methods is voting, which is combination of models in their sum of votes. In addition to that, bootstrap aggregating (bagging) is also a voting method, but models are trained with a subset of training set with possibility of overlapping samples in trained models. Another ensemble method to achieve higher accuracy is boosting. In boosting, learning models are cascaded, and each model learns from previous model's mistakes.

2.4 Literature Review

There have been several studies in the literature concentrated on detection of 2HG as it is one of most popular topics in MR spectroscopy field. In one of the first studies, Pope et al. used short echo time (30 ms) PRESS sequence and stated that specificity of 2HG detection might be an issue, since 26% false positive detection rate was seen along with 100% sensitivity [10]. Furthermore, Choi et al. optimized echo time and pulse sequence parameters in PRESS with numerical simulations and phantom studies and applied long (97 ms) echo time PRESS and detected 2HG signal at 2.25 ppm in all IDH mutant patients in their study cohort [13]. Another study by the same group, compared short (35 ms) and long (97 ms) echo time PRESS and reported that at TE=97 ms a better 2HG estimation could be achieved, but TE=35 ms allowed for a quantification of full signal from other metabolites [12]. In a recent study, Branzoli et al. employed both 97 ms PRESS and MEGA-PRESS at 68 ms and found out that 97 ms PRESS acquisition may result in lower sensitivity and specificity values, especially when strict Cramer-Rao Lower Bound (CRLB) threshold, such as 15% or 30%, was applied [37]. In addition to that, same study stated that spectral editing methods could

be a better option for monitoring 2HG levels in IDH-mut patients to follow therapy response. Also, other sequences were developed to face problems that are frequently seen in MRS. In one of these studies, Andronesi et al. used spectral editing technique Mescher-Garwood Localization by Adiabatic Refocusing (MEGA-LASER) in addition to 2D LASER COSY and 1D LASER. Relative quantification of 2HG was compared to Gln+Glu and it was said that 2HG/Gln+Glu >1 could be a specific threshold for IDH mutation detection [11]. Ultra-high field scanners were also used on detection of 2HG in gliomas. For instance, Emir et al. reported that sensitivity and spectral resolution could be increased by using 7T scanners for detection of 2HG [1]. Yet, 7T scanners are still far from being common in clinics and still have inhomogeneity issues. All of these studies concentrated on detection of 2HG at a specific range of precessional frequencies using target oriented sequences rather than utilizing all metabolites that are expected in the MR spectrum.

Machine learning and pattern recognition techniques have been used earlier for classification of brain tumors. One of the earliest studies was combination of MRS and MRI features and investigation of brain tumor classification utilizing feature reduction techniques [38]. Furthermore, the study employed estimated metabolite concentrations as features in classification algorithms. Similarly, the whole spectra that was acquired from the tumor region was used in another study and 90% accuracy was achieved in the test set, predicting glioblastoma, low-grade glioma, and meningioma spectra [39].

3. MATERIALS and METHODS

3.1 MRS Phantom

MRS sequences are first validated with phantoms before in vivo studies. Yet, commercially available MRS phantoms have uniform structure, which limits assessment of sequence performances in terms of spatial localization accuracy of multi-voxel acquisitions and resultant quantification of varying metabolite concentrations. There have been studies that produced multi-compartment MRS phantoms using vials, but those phantoms lack geometric similarity with a human head to simulate coil loading and B0 inhomogeneity [40]. As 3D printers become widespread, rapid prototyping has been more available at research centers. Previous studies employed 3D printers to produce MR phantoms for other modalities or generic use [41, 42]. As a part of this thesis, a 3D printed human head sized multi-compartment MRS phantom was designed and printed for validating sequences on detection of brain metabolites including 2HG for IDH-mut and IDH-wt gliomas.

MRS solutions were prepared in two steps: first a single vial to optimize solution preparation procedure, then the solutions that would be used in the phantom. Table 3.1 shows metabolite concentrations of each solution that were prepared in this process. The single vial solution was prepared in 100 ml polyethylene jar. Other solutions were prepared in volumes corresponding to compartment sizes in phantom. Gln was not added to the phantom solutions, because of the instability of this metabolite. The vial was scanned on the same day of preparation to avoid Gln degradation. Metabolite ratios were chosen to mimic normal brain white matter, IDH-mut, and IDH-wt brain tumors. The molecular weights, concentration values and CAS numbers of all the chemicals used in phantom preparation are given in Appendix A. For each metabolite in the solutions, its amount was calculated using the corresponding concentration values and molecular weights listed in Appendix A. First, ionized water, in half of the solution volume, was poured to a beaker with a magnetic stirrer fish inside and pH

was fixed to 7.2 by titration with sodium hydroxide (NaOH) and potassium monophosphate (KH_2PO_4) buffer solutions. When the pH is fixed, metabolites were added to the solutions one by one starting from the highest weight to the lowest one. After that, Gadolinium ($C_{14}H_{20}GdN_3O_{10}$) in 0.047% and sodium azide (NaN_3) in 0.1% of the solution weight were added to the solutions for shortening T1 and preventing bacterial growth, respectively. After all ingredients were added, pH of the solution would normally be acidic because of the nature of the chemicals used, so pH was fixed to 7.2 again with buffer solutions in the amounts of 1-2 ml. In the final step, ionized water was added to the solution to match the volume of solution to the desired compartment.

Table 3.1
Metabolite concentrations in millimolar for each prepared solution.

	Vial	IDH-mut	IDH-wt	Normal Brain
NAA	6.32	7.50	6.32	12.50
Myo-Inositol	3.80	4.50	3.80	7.50
Creatine	6.32	6.32	6.32	10.00
Choline	5.00	4.50	5.00	3.58
Lactate	6.00	5.25	6.00	0.00
Glutamate	5.75	5.75	5.75	12.50
GABA	3.00	3.00	3.00	3.00
2HG	4.73	5.20	0.00	0.00
Glutamine	5.75	0.00	0.00	0.00

The phantom model was designed using Solid Edge (Siemens PLM Software, Plano, Texas). For the design of the phantom, a half sphere with a diameter of 17 cm was used as the outer shell. The thickness of the outer shell was set to 1 cm to simulate the skull. Three sphere shaped compartments with volumes of 35 ml each were placed inside the phantom with a thickness of 3 mm at their perimeter. The compartments were placed at least 1 cm away from each other (a) and also from the outer shell (b) to avoid susceptibility issues (Figure 3.1). The compartment perimeters did not generate an MR signal (c), which shows the suitability of the material for MRS phantom design.

Acrylonitrile butadiene styrene (ABS) was used as the 3D printing material, as

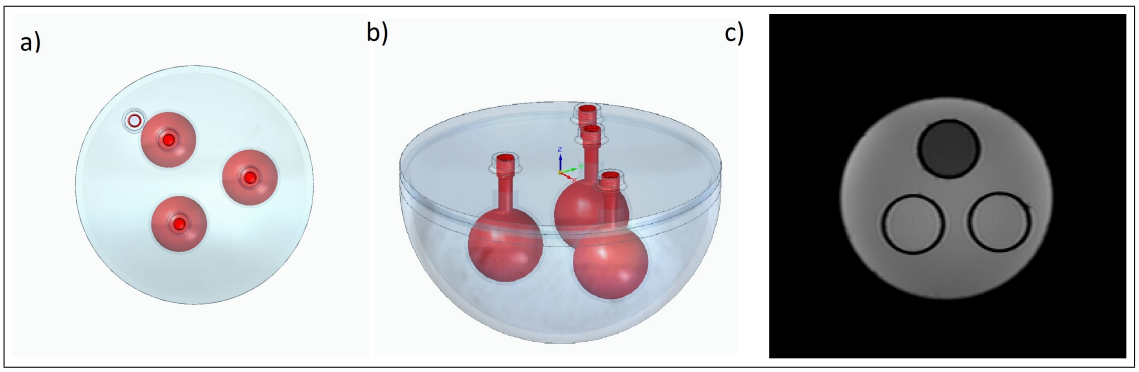

Figure 3.1 Top (a), frontal (b) and axial MRI (c) views of the 3D phantom design.

it is waterproof and does not generate any detectable MR signal. All parts were printed in highest density on M200 (Zortrax, Olsztyn, Poland) 3D printer and carefully cleaned to avoid contamination. The lid was sealed with hot glue before filling the phantom. Two of the compartments were filled with solutions that mimic the spectra of IDH-mut and IDH-wt brain tumors. Metabolite concentration ratios of these compartments were selected from previous studies [36]. Last compartment was filled with water and additive salts to mimic cerebrospinal fluid (CSF). Rest of the volume was filled with a solution mimicking normal brain white matter.

The vial and the phantom were scanned on a 3T Siemens Prisma scanner using a 32 channel head coil at room temperature. Single voxel PRESS data was acquired from vial (TR=2000ms, TE=30ms, NA=32, voxel size=8ml, BW=2000Hz, 1024 spectral points). A water unsuppressed spectrum was acquired from the same voxel for absolute metabolite quantification (NA=4). A multi-voxel PRESS data was acquired covering all compartments of the phantom (TR=2000ms, TE=30ms, voxel size=0.5ml, spatial resolution=16x16, BW=2000Hz, 1024 spectral points). Also, MEGA-LASER sequence was used for detecting 2HG (TR=2000ms, TE=68ms, Resolution=16x16, voxel size=0.5ml, BW=1100Hz, 512 spectral points) [43]. LCModel spectral fitting program was used for quantification of metabolites using respective basis sets [44]. For multi-voxel acquisition, metabolite maps were generated in MATLAB R2018a (MathWorks Inc., Natick, MA). Metabolites with a CRLB of higher than 25% were excluded from the analysis.

3.2 Data Acquisition

For MEGA-PRESS and short TE PRESS comparison, a total of 82 glioma patients, whose MRI/MRS scans were available with pathology reports, were included. Acıbadem Mehmet Ali Aydınlar University Ethics committee approved the study protocol and written consent was obtained from the patients before the scan. A clinical MR imaging protocol for glioma patients was followed before MRS data acquisition. An experienced radiologist conducted all steps of the MRS acquisitions and evaluated the results visually before saving the data to the database.

3.2.1 MRS Data Acquisition

The MRI/S data were obtained on a 3T whole body MR scanner (Prisma, Siemens Medical Systems, Erlangen, Germany) with a 32 channel head coil at Acıbadem Altunizade Hospital. The data acquisition sequences of PRESS and MEGA-PRESS were supplied by the vendor. T2 weighted and T2 FLAIR MR images were used as reference by the radiologist for voxel placement. MRS data were acquired from the solid tumor region excluding gross hemorrhage, edema and necrosis. An automatic shimming scheme was followed by manual first and second order shimming for best spectral results. The number of signal averages was set to a value between 128 and 256, depending on the situation of the patient in the scanner and expected spectrum quality from the placed voxel.

For PRESS sequence, a short echo time (TE=30ms) was chosen to avoid T2 signal loss and TR value was set to 2000ms. At this echo time macromolecule and lipid contamination was still expected to occur to an extent, especially in heterogeneous tumor sites [45]. Nevertheless, short echo time acquisitions are advantageous with higher signal intensity, and were used in previous 2HG detection studies in gliomas [10].

For MEGA-PRESS sequence, echo time was chosen as 68ms and TR value was set to 1600ms. At this echo time, middle part of the 2HG triplet at 4.02 ppm is

inverted since J-coupling constants between neighbor protons resonating at 1.90 ppm are 7Hz (leading to an inversion at TE=142ms) and 4.1Hz (inversion at TE=243ms). At difference spectra, the outer peaks of the triplet at 4.02 ppm are magnified, while the one in the middle is almost suppressed. Also, this echo time is widely used in studies for quantification of J-coupled metabolites such as Glu, Gln and GABA [32, 46, 47] .

3.2.2 Histopathological Tests

Immunohistochemistry (IHC) and sequencing methods are widely used in the clinics for detection of IDH mutation in surgical specimen excised from gliomas [48]. IDH status of each sample was assessed with DIA-H09 antibody. Also, sequencing was used for IDH negative samples for possible false positive results of IHC.

3.3 MRS Data Analysis

LCModel spectral fitting program was used for quantification of the metabolites [44]. LCModel software uses basis sets to estimate metabolite concentrations of a given spectrum. These basis sets include each metabolite's peak pattern at a given TE, and LCModel linearly combines these patterns for quantification. The basis set for MEGA-PRESS sequence was simulated for TE=68 ms by using General Approach to Magnetic Resonance Mathematical Analysis (GAMMA/PyGAMMA) simulation library of Versatile Simulation, Pulses and Analysis (VESPA) [49] using known metabolite chemical shifts and coupling constants [34, 35]. First, simulation package of the MEGA-PRESS pulse sequence was downloaded from VESPA contributions website. This simulation package includes all parameters used in Siemens MEGA-PRESS sequence. Most of the metabolites, that were expected in brain MRS, comes as default in VESPA software [50]. For those metabolites that should be added into the basis set, chemical shift values and j-coupling constants were defined in "Add metabolite" section under "Management". When the setup was ready, "New simulation" was clicked under "Experiment" and all the parameters were set and metabolites of interest were

included in the simulated signal. For a reasonable computation time, some of the metabolites were either divided to their moieties, such as GSH, or peaks that do not have J-coupling, were downscaled to less protons, such as trimethyl in choline compounds. These metabolites could be mixed in the export of results or upscaled by a coefficient. After all the metabolites were simulated with the same parameters that were used in the MRS scan, makebasis function of LCModel was used to combine these metabolites into a basis set. A short echo time basis set containing all metabolites including 2HG existed in the LCModel program. CRLB was used as a quality measure of the quantification process, as concentration values of expected metabolites could only be estimated. LCModel quantified metabolite concentration estimations with their standard deviations, and ratio to total Cr estimations were used as the output.

As this retrospective study used only already available clinical data, data loss caused by quality assurance was inevitable. MR spectra with high full width at half max (FWHM) values or low quality caused by lipid contamination in either or both of the MRS acquisitions were excluded from the analysis. Seventeen patients of the cohort had to be excluded from the study because of poor SNR or high FWHM values either in short TE PRESS or MEGA-PRESS spectrum. The rest of the cohort included 26 IDH-mut and 39 IDH-wt gliomas. Figure 3.2 shows LCModel analysis results for short TE PRESS and difference spectrum of MEGA-PRESS sequence for an example IDH-mut grade III glioma.

Using strict CRLB thresholds, such as 20%, could result in losing most of the metabolite estimations, but would result in a more confident analysis. On the other hand, this approach would harm the accuracy of the machine learning models, because as stated before, machine learning algorithms require large datasets for learning. In this thesis, metabolites with higher than 30% CRLB values were either filled with zero or not a number (NaN) as a quality control in the quantification process. By this way, the effect of uncertain metabolite estimations was investigated, for either setting that metabolite as nonexistent in the voxel (as filled with zero) or not estimated properly in the quantification (as filled with NaN).

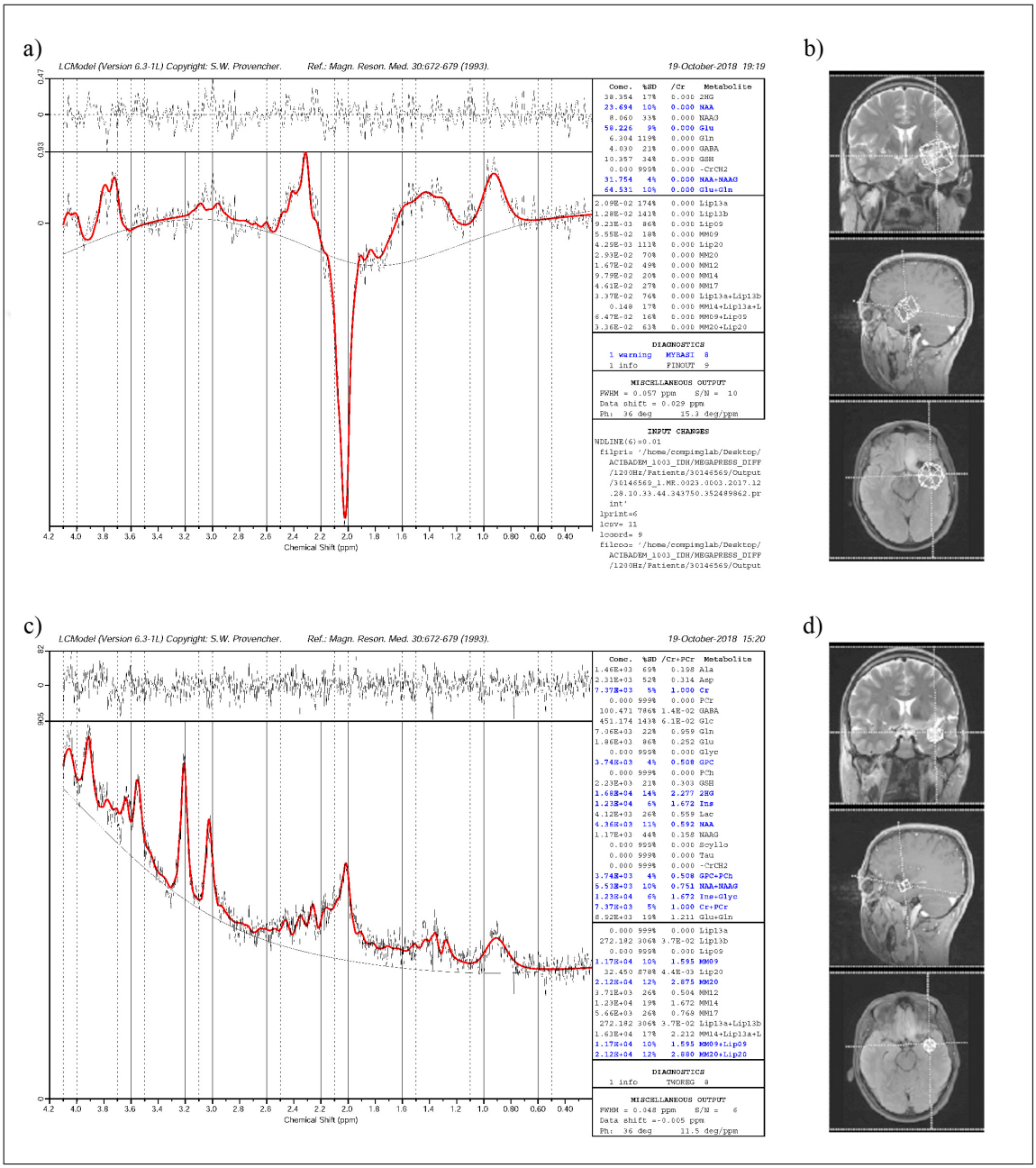

Figure 3.2 LCModel analysis of MEGA-PRESS difference (a) and short TE PRESS (c) spectra and voxel placements (b, d) for an IDH-mut grade III glioma.

The correlations between quantification of metabolites was also investigated. For those metabolites that had correlation coefficient below -0.5, their sum was used as a feature. 2HG, total of Ins and Glyc (Ins+Glyc), GSH, Lac, Cho, NAA and Glx (Glu+Gln) to Cr ratios were used in machine learning for both PRESS and MEGA-PRESS sequences. 2HG, GSH, NAA and Glx were estimated from the difference spectrum for the MEGA-PRESS sequence.

3.4 Machine Learning Models

Classification Learner app in MATLAB R2018a was used to construct machine learning models, such as decision trees, SVM, kNN, and ensemble of bagged trees to classify IDH-mut and IDH-wt gliomas. 10-fold cross validation was used to evaluate the classifier performances. Default MATLAB settings were used while training models. Models giving the highest accuracy was executed hundred times with default parameters and average accuracy, sensitivity and specificity values were reported.

3.5 Statistical Analysis

LCModel outputs were also used for the statistical analysis of the study. Sensitivity and specificity of both methods was calculated using pathology reports. Bland-Altman method was used for analysis agreement of the two MRS data acquisition methods [51]. Mean value and standard deviation of differences was calculated to estimate mean bias and to measure variations on NAA/Cr, Cho/Cr and Glx/Cr estimations of short TE PRESS and MEGA-PRESS.

4. RESULTS

4.1 MRS Phantom Results

Figure 4.1 shows LCModel analysis results for the first solution prepared in a vial. Table 4.1 gives the differences of actual and estimated concentration values of metabolites in the vial. While Gln estimation was 55% less, 2HG estimation was 79.5% higher than the actual concentration value. Other metabolites were quantified well with up to 20% error margin.

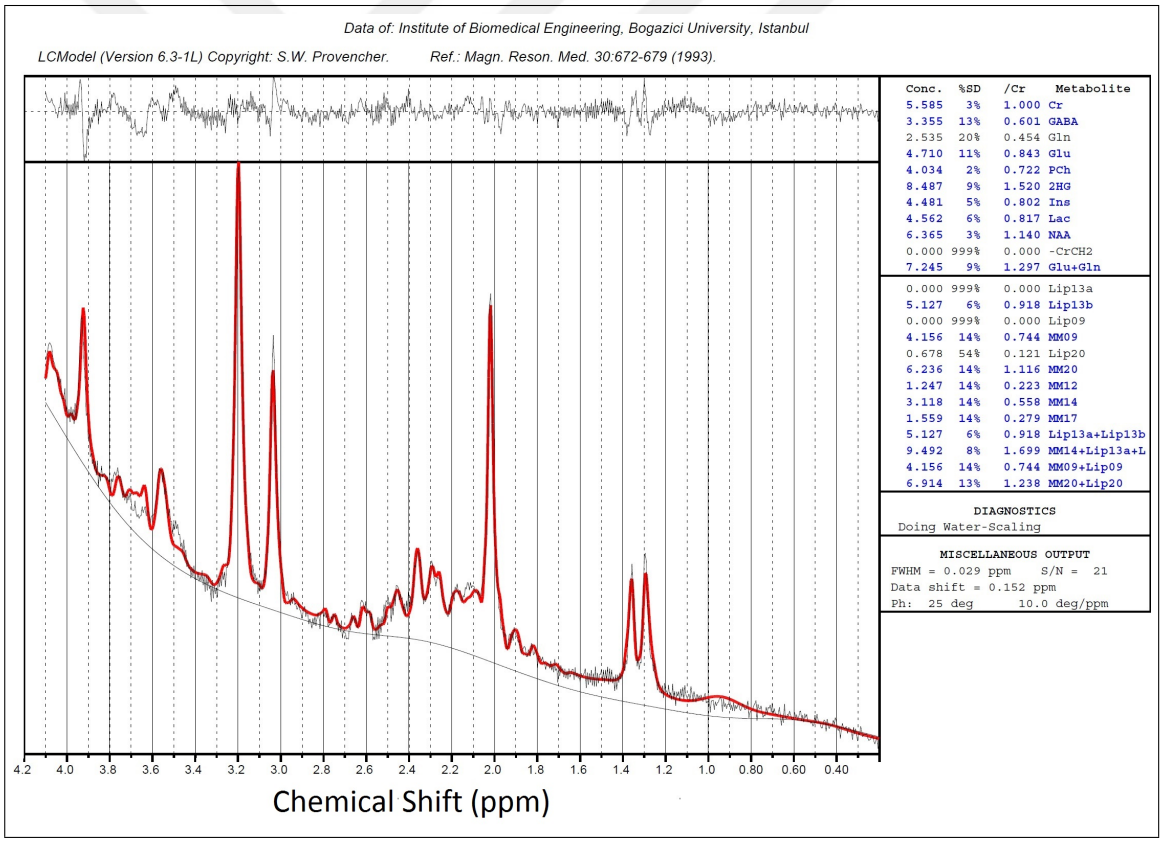

Figure 4.1 LCModel analysis results of the solution prepared in a vial.

Table 4.1
The differences of actual concentration values of the solutions and LCModel results.

	Vial(mM)	LCModel output	Difference
NAA	6.32	6.37	+0.8%
Myo-Inositol	3.80	4.48	+17.9%
Creatine	6.32	5.59	-11.5%
Choline	5.00	4.03	-19.4%
Lactate	6.00	4.56	-21.9%
Glutamate	5.75	4.71	-18.0%
Glutamine	5.75	2.54	-55.8%
2HG	4.73	8.49	+79.5%
GABA	3.00	3.36	+12.0%

In Figure 4.2, grid placement of short echo time multi-voxel PRESS on phantom image (a), LCModel analysis results of voxels taken from IDH-mut (b), and IDH-wt (c) compartments, and metabolite maps generated from LCModel results of Cho (d), Cr (e) and NAA (f) values in multiple voxels are given. Metabolite concentration estimates with a CRLB of higher than 25% were excluded from the analysis. The grid placement of MEGA-LASER on phantom image(a) and 2HG map generated from difference spectrum (b) are given in Figure 4.3. It could be seen that, 2HG was detected only in voxels placed on the IDH-mut compartment. LCModel analysis of difference spectrum taken from IDH-mut tumor (c), IDH-wt tumor (d), and healthy brain (e) mimicking compartments are given in the bottom row. Edited signal of 2HG giving rise to peaks at 4.02 ppm could only be seen in the IDH-mut compartment of the phantom.

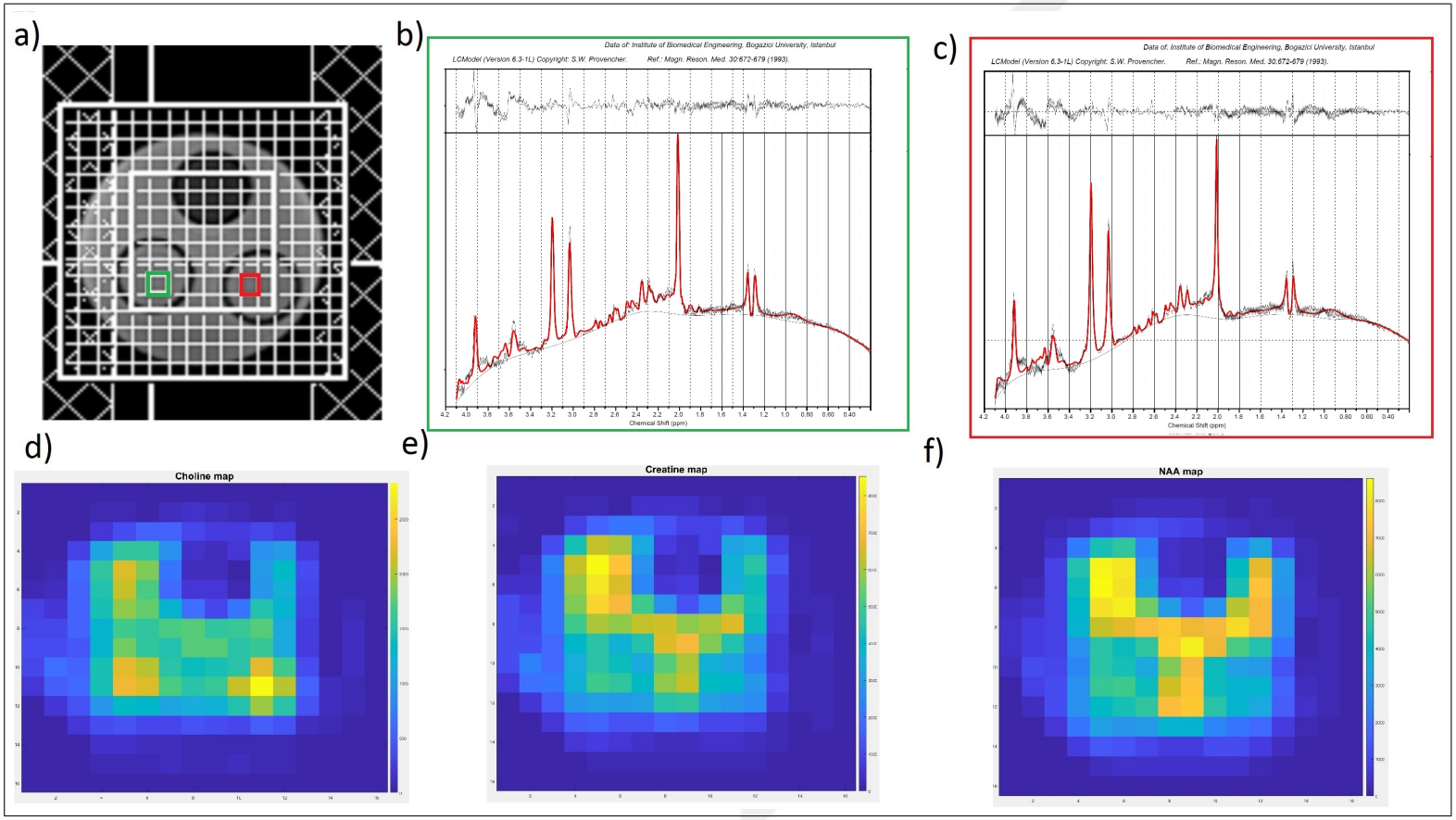

Figure 4.2 Grid placement of multi-voxel PRESS acquisition (a) LCModel analysis of IDH-mut (b) and IDH-wt (c) compartments. Metabolite maps for Cho (d), Cr (e) and NAA (f).

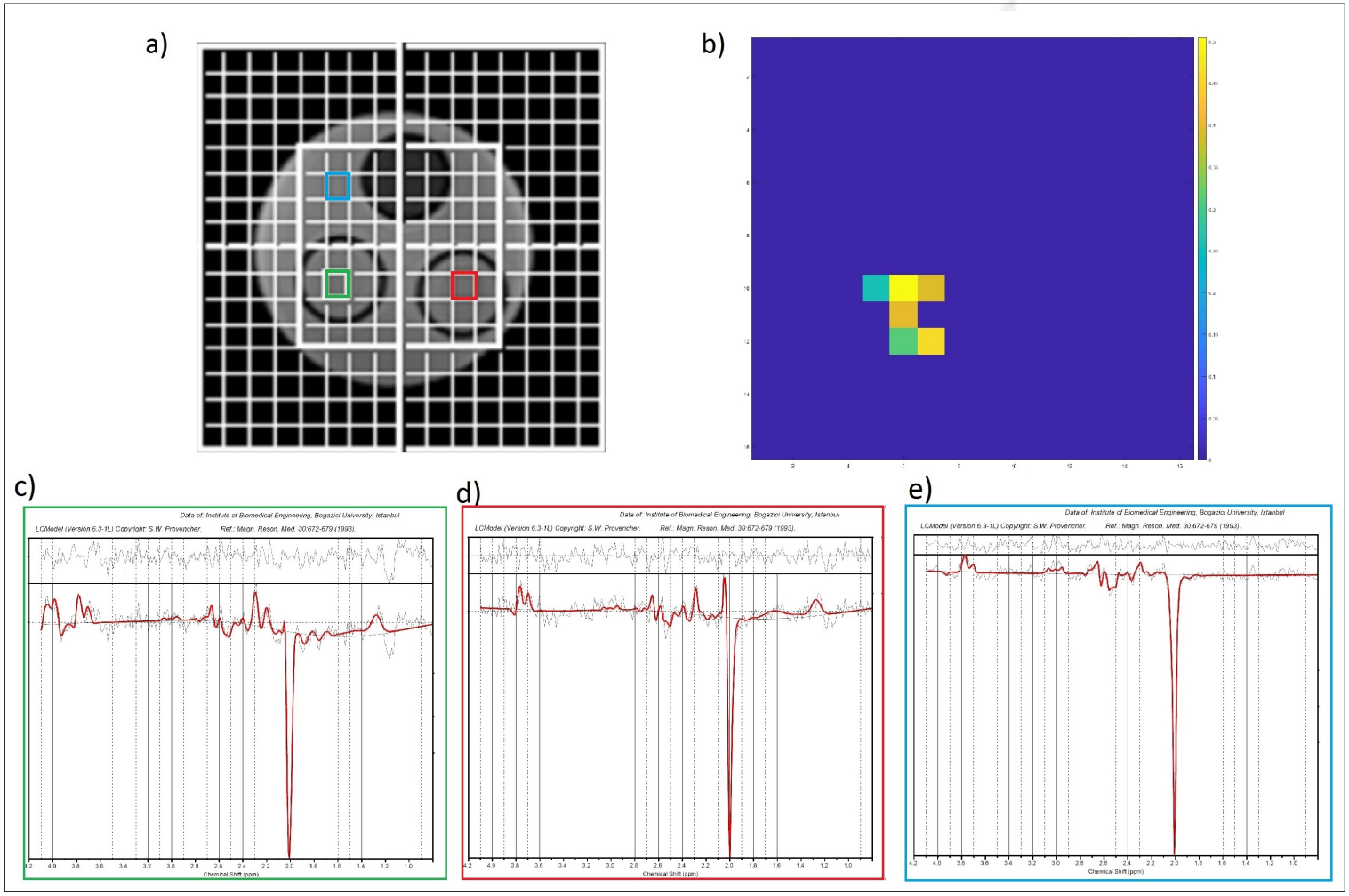

Figure 4.3 The grid placement of MEGA-LASER acquisition (a), 2HG map generated from difference spectrum (b) and LCModel analysis of difference spectrum taken from IDH-mut (c), IDH-wt (d) a healthy brain (e) compartments.

4.2 Comparison of Metabolite Levels Between MEGA-PRESS and Short TE PRESS

Bland-Altman results are given in Table 4.2. There were only 2 outliers for NAA/Cr and Glx/Cr, while the number of outliers was 4 for Cho/Cr. Mean difference value was close to zero for NAA/Cr (-0.2795) and Cho/Cr (0.1696) estimations. Also, standard deviations were low for NAA/Cr (0.2769) and Cho/Cr (0.1603) between acquisition methods. On the other hand, for Glx/Cr both mean difference (1.3647) and standard deviation (0.9646) were relatively high compared to the other two metabolites. The scatter plots of Bland-Altman test for Cho/Cr, NAA/Cr and Glx/Cr are given in figures 4.4, 4.5, and 4.6, respectively.

Table 4.2
Bland-Altman statistical test results for Cho/Cr, NAA/Cr and Glx/Cr metabolites between short TE PRESS and MEGA-PRESS.

	Cho/Cr	NAA/Cr	Glx/Cr
Number of outliers	4	2	2
Mean difference	0.1696	-0.2795	1.13647
Standard deviation	0.1603	0.2869	0.9646

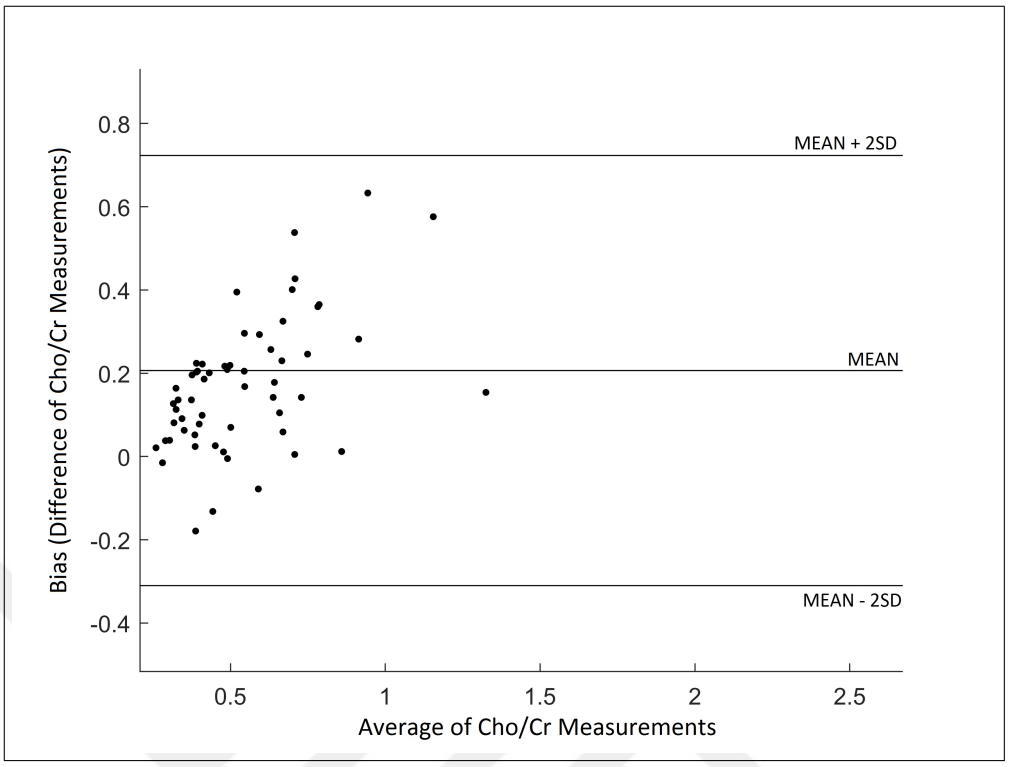

Figure 4.4 Bland-Altman scatter plot of Cho/Cr measurements.

4.3 Machine Learning Results

The highest average classification accuracy of 75% was obtained with decision trees for short TE PRESS when zero filling was used for for missing values (Table 4.3). While decision trees model also had the highest average sensitivity value of 75%, ensemble of bagged trees model resulted in the highest average specificity value of 83%.

Table 4.3
The classification accuracy results obtained from MRS profile of short TE PRESS with zero filling.

	Accuracy	Sensitivity	Specificity
Decision trees	0.75 ± 0.04	0.75 ± 0.06	0.75 ± 0.04
Ensemble of bagged trees	0.73 ± 0.03	0.57 ± 0.06	0.83 ± 0.03
SVM	0.69 ± 0.03	0.63 ± 0.05	0.74 ± 0.02
kNN	0.66 ± 0.03	0.63 ± 0.06	0.69 ± 0.04

Between the models that used MEGA-PRESS features with zero filling, the highest average accuracy of 68% and the highest average specificity of 81% were ob-

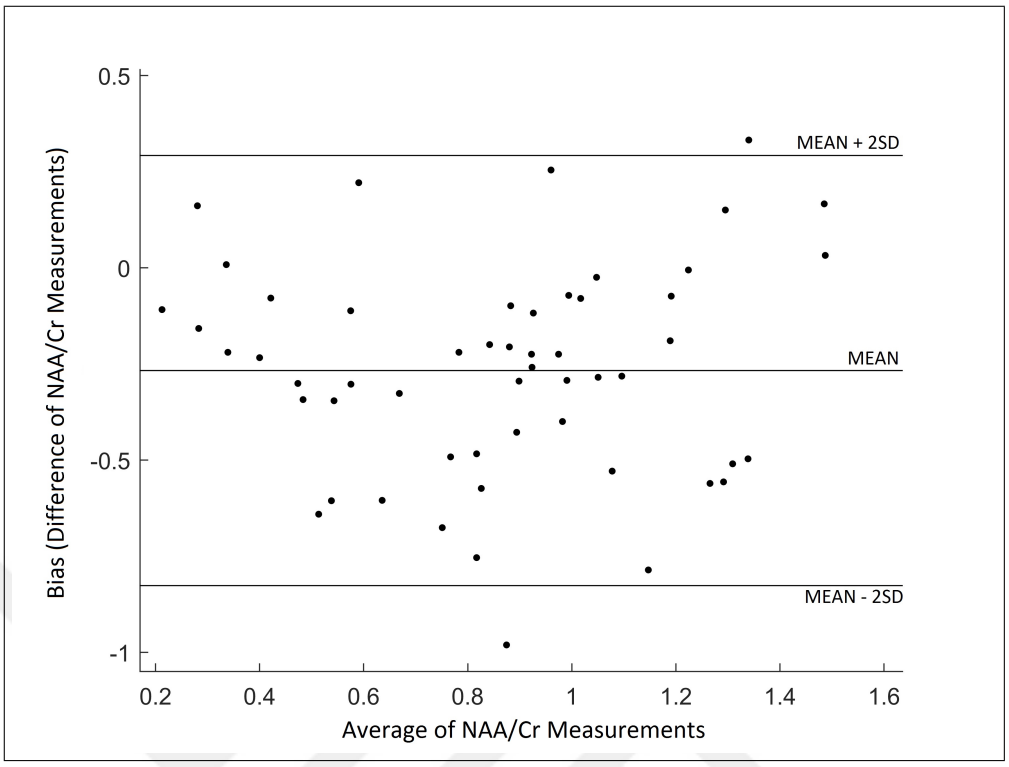

Figure 4.5 Bland-Altman scatter plot of NAA/Cr measurements.

tained with ensemble of bagged trees (Table 4.4). However, a kNN model resulted in the highest sensitivity of 59%, along with the second best accuracy performance among all trained models.

Table 4.4
The classification accuracy results obtained from MRS profile of MEGA-PRESS with zero filling.

	Accuracy	Sensitivity	Specificity
Decision trees	0.58 ± 0.04	0.36 ± 0.08	0.73 ± 0.06
Ensemble of bagged trees	0.68 ± 0.04	0.48 ± 0.06	0.81 ± 0.05
SVM	0.53 ± 0.03	0.44 ± 0.04	0.59 ± 0.04
kNN	0.64 ± 0.03	0.59 ± 0.05	0.67 ± 0.08

The highest average classification accuracy and the highest average specificity obtained were 72% and 91% with a kNN model for short TE PRESS with NaN filling (Table 4.5). Also, decision trees model had the highest average sensitivity of 69% with the second best average accuracy of 70%. In addition, an SVM model did not work in this setting as the number of missing values were high and SVM could not be trained

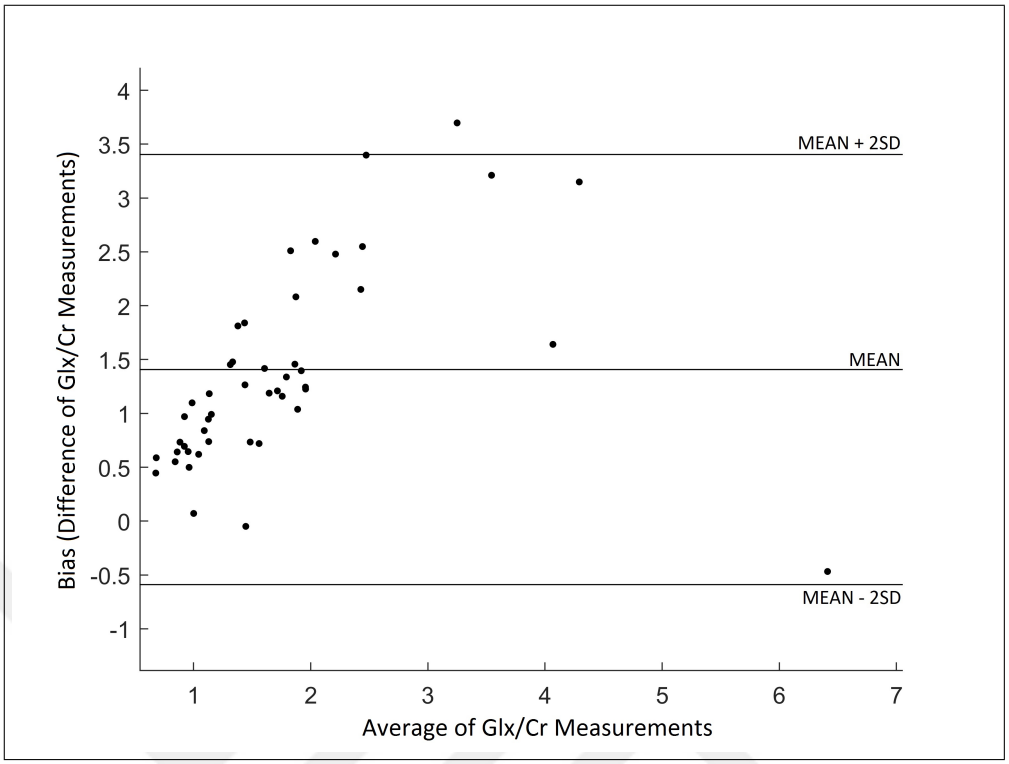

Figure 4.6 Bland-Altman scatter plot of Glx/Cr measurements.

with NaN values.

Table 4.5
The classification accuracy results obtained from MRS profile of short TE PRESS with NaN filling.

	Accuracy	Sensitivity	Specificity
Decision trees	0.70 ± 0.04	0.69 ± 0.07	0.70 ± 0.04
Ensemble of bagged trees	0.70 ± 0.03	0.63 ± 0.06	0.75 ± 0.03
kNN	0.72 ± 0.02	0.42 ± 0.01	0.91 ± 0.01

The highest average classification accuracy of 66% was obtained with ensemble of bagged trees when MEGA-PRESS was used with NaN filling (Table 4.6). Also, a kNN model had the highest average specificity value of 90%, but average sensitivity was very poor for the same model. Similarly, SVM models did not work with NaN values.

Table 4.6
The classification accuracy results obtained from MRS profile of MEGA-PRESS with NaN filling.

	Accuracy	Sensitivity	Specificity
Decision trees	0.64 ± 0.04	0.53 ± 0.06	0.71 ± 0.04
Ensemble of bagged trees	0.66 ± 0.03	0.48 ± 0.06	0.78 ± 0.04
kNN	0.61 ± 0.01	0.17 ± 0.03	0.90 ± 0.02

5. DISCUSSION

In this study, a multi-compartment MRS phantom was designed and printed using 3D printer technology. Chemical solutions were prepared to mimic IDH-mut and IDH-wt tumor tissues for MRS sequences. A prototype of the MRS phantom was produced and used as a validation tool. Furthermore, the classification performances of machine learning models were compared for IDH mutation detection based on metabolites of short TE PRESS and MEGA-PRESS sequences acquired from same patients. Even though, the dataset was not optimal for machine learning applications, reasonable accuracies were achieved in some models.

In short term, ABS was a suitable material for the design of the MRS phantom in terms of signal generation and magnetic susceptibility. However, we observed an unexpected peak at 2.2 ppm in the spectrum when the phantom was scanned for the second time six months later. We suspect that ABS might have interacted with the solution inside the phantom and dissolve in the process. In addition to that, 3D printing is still a costly procedure in terms of time and money. In the printing process, we observed that the printer may use printing patterns that could cause leakages in the future steps. Also, coating materials that could prevent these leakages may cause susceptibility artifacts and unexpected peaks in the spectrum.

LCModel is widely used in MRS studies as it needs minimal user interference in data analysis. Nevertheless, LCModel quantification is optimized for in-vivo data, and even with the phantom settings, the quantification of phantom solutions was less accurate for metabolites having overlapping peaks, such as Gln, Glu, and 2HG. Yet, the main metabolites of the brain spectrum were quantified well, and the intensity values in generated metabolite maps were in agreement with the metabolite ratios in the corresponding compartments of the phantom. Therefore, a 3D printed phantom could be used as a validation tool for optimizing spatial localization and testing quantification performances of new MRS sequences before in-vivo experiments.

The pipeline constructed for this study utilized multiple quality control steps. As stated in methods, all MRS data were collected in routine clinical procedures. For this reason, data of 17 patients were excluded because of poor SNR and FWHM values. In addition to that, small size of IDH-mut group caused imbalance in cohort. This imbalance affected the sensitivity of the machine learning algorithms. Also, some of the metabolites that were picked as features had overlapping peaks with other metabolites in addition to low concentrations in tissue. As a result of applying a strict CRLB threshold, some portion of the features were assigned either zero or NaN. Using zeros for missing values resulted in a higher accuracy and sensitivity. When the missing values were filled with NaN, only decision trees had a good balance between sensitivity and specificity for both MRS techniques. On the other hand, models based on MEGA-PRESS profiles failed to achieve sensitivity values of higher than 50% except kNN. Although average accuracy of kNN was second best among other models, sensitivity and specificity values were more reasonable compared to others (Table 4.4).

Previous MRS studies, which used PRESS and MEGA-PRESS for 2HG detection had smaller patient cohorts than our study [10] [12] [13] [37]. Additionally, these studies also reported issues with sensitivity or specificity of the technique as more strict CRLB thresholds were applied [12] [37]. In one of these studies, MEGA-PRESS was utilized with same parameters but longer acquisition time and Branzoli et al. stated that MEGA-PRESS results were better compared to optimized PRESS [37].Similarly, our results indicated that MEGA-PRESS had less false positive results as there were no other peaks expected at 4.02 ppm on difference spectrum. Yet, we observed that MEGA-PRESS resulted with false negative outcomes with patients that may have low 2HG concentration.

According to Bland-Altman test results, mean difference values and standard deviations were low for Cho/Cr and NAA/Cr ratios, which indicated that short TE PRESS and MEGA-PRESS sequences were similar in terms of their metabolite estimations. The main reason behind the variance of Glx/Cr estimations between the two sequences could be the overlapping peaks of Glx. As Glx was estimated from the difference spectrum in MEGA-PRESS, the editing scheme might have nulled out other

overlapping peaks resulting in a more reliable quantification of Glx. Mean difference values could be lower with T2 correction of metabolite concentrations as echo time differences between sequences could affect quantification results. Yet, for T2 correction, T2 relaxation times of each metabolite should be known, but these constants are only calculated for limited number of metabolites in healthy brain regions so far [52].

This study had some limitations caused by imbalance in groups and cohort size. Hyperparameter tuning was not applied on machine learning models as the groups were small that might have caused overfitting. However, a larger and more balanced cohort would allow advance machine learning models and optimization applications.

Future studies will explore the design of an anthropomorphic head MRS phantom with details of the brain structure, and T2 relaxation times of the healthy brain and tumor tissues will be matched using agarose. As a further application, machine learning algorithms will be utilized on other genetic alterations that are known to affect prognosis of the gliomas, such as telomerase reverse transcriptase (TERT) promoter and 1p/19q codeletion [22]. Also, higher CRLB threshold values could be applied to see how the machine learning algorithms would handle less reliable metabolite concentration estimations. Additionally, after a feature elimination process, features that survive for both sequences could be used together to achieve even higher sensitivity and specificity.

6. CONCLUSION

In this project, we have successfully produced a 3D printed multi-compartment head MRS phantom, which could help with MRS sequence development studies for detecting the spatial distribution of metabolites for brain tumor studies, including IDH mutation. The accuracy, sensitivity, and specificity performances of different machine learning models have been compared on two MRS data acquisition methods for detecting IDH mutation in gliomas. Since all the data was acquired as a part of the clinical routine on a relatively large cohort, it could be said that our results reflect real world performances of machine learning algorithms for these MRS sequences. The results of the study indicated that, IDH mutation in gliomas could be predicted with machine learning algorithms using short TE PRESS or MEGA-PRESS.

7. List of publications produced from the thesis

1. Comparison of MEGA-PRESS and Short Echo Time PRESS on Classification of IDH Mutation Using Machine Learning at 3T, A. Gursan, G. H. Hatay, C. Yakicier, M.N. Pamir, K. Ozduman, A. Dincer, E. Ozturk-Isik, " *International Society of Magnetic Resonance in Medicine, ISMRM, 2019 ISMRM 27th Annual Meeting in Montreal, Canada*, Vol. 27, pp. 2533, May. 11-16, 2019.

APPENDIX A. Phantom solution ingredients

Table A.1
Names, molecular weight, concentrations in millimolar and CAS numbers for the chemicals used in the phantom solutions.

Ingredient	Molecular weight	Concentration(mM)	CAS number
N-Acetyl-L-aspartic acid (NAA)	175.14	6.32	997-55-7
Myo-Inositol	180.16	3.80	87-89-9
Creatine Hydrate	149.15	6.32	6020-87-7
Choline Chloride	139.62	5.00	67-78-1
DL-Lactic Acid	96.01	6.00	16891-53-5
L-Glutamatic Acid	187.13	5.75	03-04-06
Gamma-Amino-N-Butyric Acid	103.12	3.00	56-12-2
L-alpha-Hydroxyglutaric Acid (2HG)	192.08	4.73	63512-50-5
L-Glutamine	146.14	5.75	56-85-9
Sodium Azide	65.01	15.38	26628-22-8
Magnavist	547.57	0.85	80529-93-7

REFERENCES

1. Emir, U. E., S. J. Larkin, N. de Pennington, N. Voets, P. Plaha, R. Stacey, K. Al-Qahtani, J. McCullagh, C. J. Schofield, S. Clare, P. Jezzard, T. Cadoux-Hudson, and O. Ansorge, "Noninvasive quantification of 2-hydroxyglutarate in human gliomas with idh1 and idh2 mutations," *Cancer Res*, Vol. 76, no. 1, pp. 43–9, 2016.

2. Klose, U., "Measurement sequences for single voxel proton mr spectroscopy," *Eur J Radiol*, Vol. 67, no. 2, pp. 194–201, 2008.

3. Kim, H., S. Kim, H. H. Lee, and H. Heo, "In-vivo proton magnetic resonance spectroscopy of 2-hydroxyglutarate in isocitrate dehydrogenase-mutated gliomas: A technical review for neuroradiologists," *Korean J Radiol*, Vol. 17, no. 5, pp. 620–32, 2016.

4. Alpaydin, E., *Introduction to Machine Learning*, The MIT Press, 2014.

5. Ferlay, J., D. M. Parkin, and E. Steliarova-Foucher, "Estimates of cancer incidence and mortality in europe in 2008," *Eur J Cancer*, Vol. 46, no. 4, pp. 765–81, 2010.

6. Ostrom, Q. T., H. Gittleman, G. Truitt, A. Boscia, C. Kruchko, and J. S. Barnholtz-Sloan, "Cbtrus statistical report: Primary brain and other central nervous system tumors diagnosed in the united states in 2011-2015," *Neuro Oncol*, Vol. 20, no. suppl4, pp. iv1–iv86, 2018.

7. Louis, D. N., A. Perry, G. Reifenberger, A. von Deimling, D. Figarella-Branger, W. K. Cavenee, H. Ohgaki, O. D. Wiestler, P. Kleihues, and D. W. Ellison, "The 2016 world health organization classification of tumors of the central nervous system: a summary," *Acta Neuropathol*, Vol. 131, no. 6, pp. 803–20, 2016.

8. Yan, H., D. W. Parsons, G. Jin, R. McLendon, B. A. Rasheed, W. Yuan, I. Kos, I. Batinic-Haberle, S. Jones, G. J. Riggins, H. Friedman, A. Friedman, D. Reardon, J. Herndon, K. W. Kinzler, V. E. Velculescu, B. Vogelstein, and D. D. Bigner, "Idh1 and idh2 mutations in gliomas," *N Engl J Med*, Vol. 360, no. 8, pp. 765–73, 2009.

9. Weller, M., W. Wick, K. Aldape, M. Brada, M. Berger, S. M. Pfister, R. Nishikawa, M. Rosenthal, P. Y. Wen, R. Stupp, and G. Reifenberger, "Glioma," *Nat Rev Dis Primers*, Vol. 1, p. 15017, 2015.

10. Pope, W. B., R. M. Prins, M. Albert Thomas, R. Nagarajan, K. E. Yen, M. A. Bittinger, N. Salamon, A. P. Chou, W. H. Yong, H. Soto, N. Wilson, E. Driggers, H. G. Jang, S. M. Su, D. P. Schenkein, A. Lai, T. F. Cloughesy, H. I. Kornblum, H. Wu, V. R. Fantin, and L. M. Liau, "Non-invasive detection of 2-hydroxyglutarate and other metabolites in idh1 mutant glioma patients using magnetic resonance spectroscopy," *J Neurooncol*, Vol. 107, no. 1, pp. 197–205, 2012.

11. Andronesi, O. C., G. S. Kim, E. Gerstner, T. Batchelor, A. A. Tzika, V. R. Fantin, M. G. Vander Heiden, and A. G. Sorensen, "Detection of 2-hydroxyglutarate in idh-mutated glioma patients by in vivo spectral-editing and 2d correlation magnetic resonance spectroscopy," *Sci Transl Med*, Vol. 4, no. 116, p. 116ra4, 2012.

12. Choi, C., S. K. Ganji, R. J. DeBerardinis, K. J. Hatanpaa, D. Rakheja, Z. Kovacs, X. L. Yang, T. Mashimo, J. M. Raisanen, I. Marin-Valencia, J. M. Pascual, C. J. Madden, B. E. Mickey, C. R. Malloy, R. M. Bachoo, and E. A. Maher, "2-hydroxyglutarate detection by magnetic resonance spectroscopy in idh-mutated patients with gliomas," *Nat Med*, Vol. 18, no. 4, pp. 624–9, 2012.

13. Choi, C., S. Ganji, K. Hulsey, A. Madan, Z. Kovacs, I. Dimitrov, S. Zhang, K. Pichumani, D. Mendelsohn, B. Mickey, C. Malloy, R. Bachoo, R. Deberardinis, and E. Maher, "A comparative study of short- and long-te (1)h mrs at 3 t for in vivo detection of 2-hydroxyglutarate in brain tumors," *NMR Biomed*, Vol. 26, no. 10, pp. 1242–50, 2013.

14. De Looze, C., A. Beausang, J. Cryan, T. Loftus, P. G. Buckley, M. Farrell, S. Looby, R. Reilly, F. Brett, and H. Kearney, "Machine learning: a useful radiological adjunct in determination of a newly diagnosed glioma's grade and idh status," *J Neurooncol*, Vol. 139, no. 2, pp. 491–499, 2018.

15. Eichinger, P., E. Alberts, C. Delbridge, S. Trebeschi, A. Valentinitsch, S. Bette, T. Huber, J. Gempt, B. Meyer, J. Schlegel, C. Zimmer, J. S. Kirschke, B. H. Menze, and B. Wiestler, "Diffusion tensor image features predict idh genotype in newly diagnosed who grade ii/iii gliomas," *Sci Rep*, Vol. 7, no. 1, p. 13396, 2017.

16. Lu, C. F., F. T. Hsu, K. L. Hsieh, Y. J. Kao, S. J. Cheng, J. B. Hsu, P. H. Tsai, R. J. Chen, C. C. Huang, Y. Yen, and C. Y. Chen, "Machine learning-based radiomics for molecular subtyping of gliomas," *Clin Cancer Res*, Vol. 24, no. 18, pp. 4429–4436, 2018.

17. Zhou, H., K. Chang, H. X. Bai, B. Xiao, C. Su, W. L. Bi, P. J. Zhang, J. T. Senders, M. Vallieres, V. K. Kavouridis, A. Boaro, O. Arnaout, L. Yang, and R. Y. Huang, "Machine learning reveals multimodal mri patterns predictive of isocitrate dehydrogenase and 1p/19q status in diffuse low- and high-grade gliomas," *J Neurooncol*, Vol. 142, no. 2, pp. 299–307, 2019.

18. Hoshide, R., and R. Jandial, "2016 world health organization classification of central nervous system tumors: An era of molecular biology," *World Neurosurg*, Vol. 94, pp. 561–562, 2016.

19. Perry, A., and P. Wesseling, "Histologic classification of gliomas," *Handb Clin Neurol*, Vol. 134, pp. 71–95, 2016.

20. Yen, K. E., M. A. Bittinger, S. M. Su, and V. R. Fantin, "Cancer-associated idh mutations: biomarker and therapeutic opportunities," *Oncogene*, Vol. 29, p. 6409, 2010.

21. Parsons, D. W., S. Jones, X. Zhang, J. C. Lin, R. J. Leary, P. Angenendt, P. Mankoo, H. Carter, I. M. Siu, G. L. Gallia, A. Olivi, R. McLendon, B. A. Rasheed, S. Keir, T. Nikolskaya, Y. Nikolsky, D. A. Busam, H. Tekleab, J. Diaz, L. A., J. Hartigan, D. R. Smith, R. L. Strausberg, S. K. Marie, S. M. Shinjo, H. Yan, G. J. Riggins, D. D. Bigner, R. Karchin, N. Papadopoulos, G. Parmigiani, B. Vogelstein, V. E. Velculescu, and K. W. Kinzler, "An integrated genomic analysis of human glioblastoma multiforme," *Science*, Vol. 321, no. 5897, pp. 1807–12, 2008.

22. Eckel-Passow, J. E., D. H. Lachance, A. M. Molinaro, and et al., "Glioma groups based on 1p/19q, idh, and tert promoter mutations in tumors," *N Engl J Med*, Vol. 372, no. 26, pp. 2499–508, 2015.

23. Birken, D. L., and W. H. Oldendorf, "N-acetyl-l-aspartic acid: a literature review of a compound prominent in 1h-nmr spectroscopic studies of brain," *Neurosci Biobehav Rev*, Vol. 13, no. 1, pp. 23–31, 1989.

24. Rackayova, V., C. Cudalbu, P. J. W. Pouwels, and O. Braissant, "Creatine in the central nervous system: From magnetic resonance spectroscopy to creatine deficiencies," *Analytical Biochemistry*, Vol. 529, pp. 144–157, 2017.

25. Rudkin, T. M., and D. L. Arnold, "Proton magnetic resonance spectroscopy for the diagnosis and management of cerebral disorders," *JAMA Neurology*, Vol. 56, no. 8, pp. 919–926, 1999.

26. Ross, B. D., "Biochemical considerations in 1h spectroscopy. glutamate and glutamine; myo-inositol and related metabolites," *NMR Biomed*, Vol. 4, no. 2, pp. 59–63, 1991.

27. Kinoshita, Y., H. Kajiwara, A. Yokota, and Y. Koga, "Proton magnetic resonance spectroscopy of brain tumors: an in vitro study," *Neurosurgery*, Vol. 35, no. 4, pp. 606–13; discussion 613–4, 1994.

28. Cooper, A. J., and B. S. Kristal, "Multiple roles of glutathione in the central nervous system," *Biol Chem*, Vol. 378, no. 8, pp. 793–802, 1997.

29. Bottomley, P. A., "Spatial localization in nmr spectroscopy in vivo," *Ann N Y Acad Sci*, Vol. 508, pp. 333–48, 1987.

30. Haase, A., J. Frahm, W. Hanicke, and D. Matthaei, "1h nmr chemical shift selective (chess) imaging," *Phys Med Biol*, Vol. 30, no. 4, pp. 341–4, 1985.

31. Mescher, M., H. Merkle, J. Kirsch, M. Garwood, and R. Gruetter, "Simultaneous in vivo spectral editing and water suppression," *NMR Biomed*, Vol. 11, no. 6, pp. 266–72, 1998.

32. Mullins, P. G., D. J. McGonigle, R. L. O'Gorman, N. A. Puts, R. Vidyasagar, C. J. Evans, M. R. S. o. G. Cardiff Symposium on, and R. A. Edden, "Current practice in the use of mega-press spectroscopy for the detection of gaba," *Neuroimage*, Vol. 86, pp. 43–52, 2014.

33. Oeltzschner, G., M. Butz, T. J. Baumgarten, N. Hoogenboom, H. J. Wittsack, and A. Schnitzler, "Low visual cortex gaba levels in hepatic encephalopathy: links to blood ammonia, critical flicker frequency, and brain osmolytes," *Metab Brain Dis*, Vol. 30, no. 6, pp. 1429–38, 2015.

34. Bal, D., and A. Gryff-Keller, "1h and13c nmr study of 2-hydroxyglutaric acid and its lactone," *Magnetic Resonance in Chemistry*, Vol. 40, no. 8, pp. 533–536, 2002.

35. Govindaraju, V. ; Young, K. . M. A., "Proton nmr chemical shifts and coupling constants for brain metabolites," *NMR Biomed*, Vol. 13, pp. 129–153, 2001.

36. Ozturk-Isik, E., S. Cengiz, A. Ozcan, C. Yakicier, M. Pamir, K. Ozduman, and A. Dincer, "Magnetic resonance spectroscopic differences of diffuse glioma groups classified by idh and tert promoter mutations at 3t," in *ISMRM-ESMRMB* (ISMRM, ed.), Vol. 26, p. 955, ISMRM, 2018.

37. Branzoli, F., A. L. Di Stefano, L. Capelle, C. Ottolenghi, R. Valabregue, D. K. Deelchand, F. Bielle, C. Villa, B. Baussart, S. Lehericy, M. Sanson, and M. Marjanska, "Highly specific determination of idh status using edited in vivo magnetic resonance spectroscopy," *Neuro Oncol*, Vol. 20, no. 7, pp. 907–916, 2018.

38. Simonetti, A. W., W. J. Melssen, F. Szabo de Edelenyi, J. J. van Asten, A. Heerschap, and L. M. Buydens, "Combination of feature-reduced mr spectroscopic and mr imaging data for improved brain tumor classification," *NMR Biomed*, Vol. 18, no. 1, pp. 34–43, 2005.

39. Raschke, F., E. Fuster-Garcia, K. S. Opstad, and F. A. Howe, "Classification of single-voxel 1h spectra of brain tumours using lcmodel," *NMR Biomed*, Vol. 25, no. 2, pp. 322–31, 2012.

40. Woo, D. C., B. S. Kim, S. L. Jung, H. J. Park, H. S. Rhim, G. H. Jahng, and B. Y. Choe, "Development of a cone-shape phantom for multi-voxel mr spectroscopy," *J Neurosci Methods*, Vol. 162, no. 1-2, pp. 101–7, 2007.

41. Gopalan, K., J. Tamir, A. Arias, and M. Lustig, "Toward 3d printed, anatomy-mimicking, quantitative mri phantoms," in *ISMRM*, Vol. 25, p. 3733, ISMRM, 2017.

42. Kasten, J. A., T. Vetterli, F. Lazeyras, and D. Van De Ville, "3d-printed shepp-logan phantom as a real-world benchmark for mri," *Magn Reson Med*, Vol. 75, no. 1, pp. 287–94, 2016.

43. Hnilicova, P., M. Povazan, B. Strasser, O. C. Andronesi, M. Gajdosik, U. Dydak, J. Ukropec, D. Dobrota, S. Trattnig, and W. Bogner, "Spatial variability and reproducibility of gaba-edited mega-laser 3d-mrsi in the brain at 3 t," *NMR Biomed*, Vol. 29, no. 11, pp. 1656–1665, 2016.

44. Provencher, S. W., "Estimation of metabolite concentrations from localized in vivo proton nmr spectra," *Magnetic Resonance in Medicine*, Vol. 30, no. 6, pp. 672–679, 1993.

45. Majos, C., M. Julia-Sape, J. Alonso, M. Serrallonga, C. Aguilera, J. J. Acebes, C. Arus, and J. Gili, "Brain tumor classification by proton mr spectroscopy: Comparison of diagnostic accuracy at short and long te," *American Journal of Neuroradiology*, Vol. 25, no. 10, pp. 1696–1704, 2004.

46. O'Gorman, R. L., L. Michels, R. A. Edden, J. B. Murdoch, and E. Martin, "In vivo detection of gaba and glutamate with mega-press: reproducibility and gender effects," *J Magn Reson Imaging*, Vol. 33, no. 5, pp. 1262–7, 2011.

47. Riese, F., A. Gietl, N. Zolch, A. Henning, R. O'Gorman, A. M. Kalin, S. E. Leh, A. Buck, G. Warnock, R. A. Edden, R. Luechinger, C. Hock, S. Kollias, and L. Michels, "Posterior cingulate gamma-aminobutyric acid and glutamate/glutamine are reduced in amnestic mild cognitive impairment and are unrelated to amyloid deposition and apolipoprotein e genotype," *Neurobiol Aging*, Vol. 36, no. 1, pp. 53–9, 2015.

48. Zou, Y., H. X. Bai, Z. Wang, and L. Yang, "Comparison of immunohistochemistry and dna sequencing for the detection of idh1 mutations in gliomas," *Neuro Oncol*, Vol. 17, no. 3, pp. 477–8, 2015.

49. Smith, S. A., T. . Levante, B. H. Meier, and R. R. Ernst, "Computer simulations in magnetic resonance. an object oriented programming approach," *J Magn Reson*, Vol. 106A, no. 1, pp. 75–105, 1994.

50. *Versatile Simulation Pulses and Analysis*, https://scion.duhs.duke.edu/vespa/.

51. Bland, J. M., and D. G. Altman, "Measuring agreement in method comparison studies," *Stat Methods Med Res*, Vol. 8, no. 2, pp. 135–60, 1999.

52. Wyss, P. O., C. Bianchini, M. Scheidegger, I. A. Giapitzakis, A. Hock, A. Fuchs, and A. Henning, "In vivo estimation of transverse relaxation time constant (t2) of 17 human brain metabolites at 3t," *Magn Reson Med*, Vol. 80, no. 2, pp. 452–461, 2018.