University of California,
Irvine

Expanding the Clinical Relevance of *In-Vivo* Magnetic Resonance Spectroscopy

DISSERTATION

submitted in partial satisfaction of the requirements
for the degree of

DOCTOR OF PHILOSOPHY

In Biological Sciences

By

Aaron Thomas Gudmundson

Dissertation Committee:
Professor Craig Stark, Chair
Professor Leslie Thompson
Professor Marcelo Wood

2022

Table of Contents

List of Abbreviations:

AD	Alzheimer's Disease
ADC	Addiction
ADHD	Attention Deficit Hyperactivity Disorder
ADNI	Alzheimer's Disease Neuroimaging Initiative
AGNOSTIC	A Generalized Neural-network Open-Source Training dataset Insensitive to Acquisition Confines
AL	*ad libitum*
APOE	Apolipoprotein E
APP	Amyloid Precursor Protein
Asc	Ascorbate
Asp	Aspartate
ATN	Amyloid, Tau Tangles, and Neurodegeneration
ATP	Adenosine Triphosphate
Aut	Autism
Bip	Bipolar
BPP	Bloembergen-Purcell-Pound
Canc	Cancer
COHERENC	Comprehensive Open-Source Hosted Electronically Relaxation and Concentration Neurometabolite Values for Clinical Research
COVID-19	Coronavirus Disease 2019
CNN	Convolutional Neural Network
CPMG	Carr-Purcell Meiboom Gill
CR	Caloric Restriction

CRLB	Cramer–Rao Lower Bounds
CSF	Cerebral Spinal Fluid
CSO	Centrum Semiovale
csv	Comma Separated Values
D1	Diabetes Type-1
Dem	Dementia
Depr	Depression
DWI	Diffusion Weighted Imaging
dt	Dwell Time
E4	Apolipoprotein 4 Allele
EOAD	Early-Onset Alzheimer's Disease
ETrm	Essential Tremor
FBMI	Focused Beam Microwave Irradiation
FCN	Fully Connected Network
Fib	Fibromyalgia
FWMH	Full-Width at Half Maximum
GABA	γ-amino-butyric acid
Glc	Glucose
Gln	Glutamine
Glu	Glutamate
GPC	Glycerophosphocholine
GSH	Glutathione
Hz	Hertz

J	Jules
KI	Knock-In
Lac	Lactate
LASER	Localization by Adiabatic Selective Refocusing
LOAD	Late-Onset Alzheimer's Disease
LSTM	Long Short-Term Memory
LTP	Long-Term Potentiation
MCI	Mild Cognitive Impairment
MEGA	Mescher-Garwood Spectral Editing Technique
Mgrn	Migraine
mM	Millimolalilty
MM	Macromolecule
MNIST	Modified National Institute of Standards and Technology dataset
MR	Magnetic Resonance
MRI	Magnetic Resonance Imaging
MRS	Magnetic Resonance Spectroscopy
MRSI	Magnetic Resonance Spectroscopy Imaging
MS	Multiple Sclerosis
MTL	Multi-Task Learning
Myo	Myoinositol
NAA	N-Acetyl Aspartate
NAAG	N-Acetyl Aspartate Glutamate
NAD	Nicotinamide Adenine Dinucleotide

NIA	National Institute of Aging
NMN	Nicotinamide Mononucleotide
NMR	Nuclear Magnetic Resonance
NN	Neural Network
npz	Numerical Python Zipped Archive File
NR	Nicotinamide Riboside
OCD	Obsessive Compulsive Disorder
OSPREY	Open-Source Processing, Reconstruction, and Estimation
OUL	Objects in Updated Locations
Pain	Chronic Pain Diagnosis
PC	Perinatal Conditions
PCC	Posterior Cingulate Cortex
PCho	Phosphocholine
PCr	Phosphocreatine
PD	Parkinson's Disease
Pers	Personality Disorder
PET	Positron Emission Tomography
ppm	parts per million
PC	Principal Components
PRESS	Point Resolved Spectroscopy
PRISMA	Preferred Reporting Items for Systematic Reviews and Meta-Analyses
PSEN	Presenilin (*1 or 2*) Gene
Psy	Psychosis

PTSD	Post-Traumatic Stress Disorder
RELU	Rectified Linear Unit
RF	Radiofrequency
RNN	Recurrent Neural Network
ROS	Reactive Oxygen Species
SCI	Subjective Cognitive Impairment
Schz	Schizophrenia
Scy	Scylloinositol
Seiz	Seizure Disorder
SIR	Silent Information Regulator
SIRT	Silent Information Regulator 2
sLASER	Semi-Localization by Adiabatic Selective Refocusing
SNR	Signal to noise ratio
SPECIAL	Spin-echo Full-Intensity Acquired Localized Spectroscopy
SPM	Statistical Parametric Mapping
STAC	Sirtuin-activating compound
STEAM	Stimulated Echo Acquisition Mode
Str	Stroke
SVD	Single Value Decomposition
Sw	Spectral width
T	Tesla
Tau	Taurine
TBI	Traumatic Brain Injury

TCA	Tricarboxylic Acid Cycle
tCho	Total Choline-containing compounds
tCr	Total Creatine-containing compounds
TE	Echo Time
tNAA	Total N-Acetyl Aspartate-Containing Compounds
TR	Repetition Time
VAPOR	Variable Power radiofrequency pulses with Optimized Relaxation delays

List of Figures:

List of Tables:

Acknowledgements:

Research reported in this dissertation was supported by the National Institute of Aging of the National Institutes of Health under award number T32AG00096. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

VITA:
Aaron Thomas Gudmundson

Field of Study:

Neurobiology and behavior

Education:

B.A.	Department of Psychology School of Social Sciences	University of California, Irvine	2017
M.S.	Department of Neurobiology & Behavior School of Biological Sciences	University of California, Irvine	2020
Ph.D.	Department of Neurobiology & Behavior School of Biological Sciences	University of California, Irvine	2022

Selected Publications:

1. Gong T, Hui SCN, Zöllner HJ, Britton M, Song Y, Chen Y, **Gudmundson AT**, Hupfeld KE, Murali-Manohar S, Porges EC, Oeltzschner G, Chen W, Wang G, Edden RAE (*2022*). Neurometabolic timecourse of healthy aging. bioRxiv 2022.06.08.495050

2. **Gudmundson AT**, Koo A, Virovka A, Amirault A, Soo M, Stark CEL (*In Prep*). COHERENC: A freely available online database for Magnetic Resonance Spectroscopy Metabolite Concentrations and Relaxation Parameters in the Healthy and Diseased Brain. (*In Prep*)

3. **Gudmundson AT**, Koo A, Virovka A, Stark CEL (*In Prep*). Deep Learning for MRS Data and AGNOSTIC benchmark dataset.

Selected Posters:

1. **Gudmundson AT**, Boucquey V, Stark SM, Stark CEL (2015). Analysis of the Neurobiological Mechanisms Supporting Sequence Memory Using High Resolution Functional Magnetic Resonance Imaging within the Hippocampus, Southern California Conference for Undergraduate Research.
2. **Gudmundson AT**, Boucquey V, Stark SM, Stark CEL (2016). Analysis of the Neurobiological Mechanisms Supporting Sequence Memory Using High Resolution Functional Magnetic Resonance Imaging within the Prefrontal Cortex, Honors Transfer Counsel of California Student Research.
3. **Gudmundson AT**, Boucquey V, Stark SM, Stark CEL (2017). Item-item and

Item-Position Strategy Use in a Cross-Species Sequence Memory Task, Graduate Diversity Research Conference at the University of California, Irvine.

4. **Gudmundson AT**, Stark SM, Stark CEL (2017) The effect of aging on fMRI activity during a sequence memory task, Society for Neuroscience

Selected Talks:

1. **Gudmundson AT** (2018) The effect of aging on fMRI activity during a sequence memory task, Neurobiology and Behavior NeuroBlitz at University of California, Irvine.
2. **Gudmundson AT** (2018) Physiological and Artifactual Noise Reduction in Magnetic Resonance Imaging, Campus Center for Neuroimaging at University of California, Irvine.
3. **Gudmundson AT** (2019) approaches to decoding cognitive strategies involved in temporal memory, Neurobiology and Behavior NeuroBlitz at University of California, Irvine.
4. **Gudmundson AT** (2019) The neural biochemistry that supports neurotransmitter production and release, 5th International Symposium on GABA and Advanced MRS.

Courses Taught:

N113L	Neurobiology Lab	University of California, Irvine School of Biological Sciences Dept. of Neurobiology & Behavior	2019
N115	Advanced Neurobiology (Part-A)	University of California, Irvine School of Biological Sciences Dept. of Neurobiology & Behavior	2020
N115	Advanced Neurobiology (Part-B)	University of California, Irvine School of Biological Sciences Dept. of Neurobiology & Behavior	2020

Funding

T32 Aging and Alzheimer's Training Grant	National Institute for Health	2019-2020
T32 Aging and Alzheimer's Training Grant	National Institute for Health	2020-2021

Outreach Activities

Community College Mentor	Irvine Valley College	2017-2022
Symposium Organizer	S. OC Community College District	2018

ABSTRACT OF THE DISSERTATION:

Expanding the clinical relevance of *in-vivo* Magnetic Resonance Spectroscopy

by

Aaron Thomas Gudmundson

Doctor of Philosophy in Biological Sciences

University of California, Irvine, 2022

Professor Craig E.L Stark

Proton (^{1}H) Magnetic Resonance Spectroscopy (*MRS*) is a powerful tool for studying brain metabolites *in-vivo*. As a clinical research tool, consistent and accurate quantification is a necessity, which has driven the MRS community to prioritize producing reliable and standardized analysis software packages. Though, low signal to noise, artifacts, and overlapping signals still provide significant challenges, especially for deep brain regions and multi-voxel acquisition. Data quality can be improved through collecting multiple transients or employing multi-scan (*editing, nulling, etc.)* methods, but time constraints may limit what can be done during a scan session.

This thesis aims to address some of the fundamental problems associated with performing MRS in the brain by developing advanced methods surrounding MRS data processing and analysis. Specifically, I'll describe the results accomplishments from 2 projects. In part-1, I describe developing the open-source COHERENC database and meta-analysis to provide a metabolic profile in both the healthy and clinical brain. Then, in Part-2, I describe the development of a neural network training and benchmark dataset (*AGNOSTIC*) for MRS as well as our work in creating deep-learning-based data analysis techniques.

1. Introduction:

1.1. Neuroimaging:

Neuroimaging modalities are critical components for clinical research and medical diagnostics. While different forms of neuroimaging techniques are available, there is great value in continued development of safe and non-invasive methods. This is especially important in the context of the brain where there is great risk in surgical operations. Magnetic Resonance (MR) is a popular method as it is capable of providing functional, structural, and/or chemical information. Unlike radioactive methods such as Positron Emission Tomography (*PET*), MR uses magnetic induction to measure the internal state of a system by taking advantage of the physical property known as *Spin (discussed further in section 2.1.1.).*

1.2. Magnetic Resonance Spectroscopy:

Magnetic Resonance Spectroscopy (*MRS*) is one of many subtypes of MR-based techniques used commonly by chemists, biologists, and some clinicians. MRS is a quantitative (*represents the actual number of nuclei*) method used to non-invasively investigate the chemical composition and concentrations within a target. In the brain specifically, MRS is generally capable of measuring brain metabolites that have a concentration greater than or equal to 1 millimolar/kilogram (mM). Metabolites can be isolated from one another as their unique chemical structure and stereochemistry creates a different interaction within a magnetic field and thus produces a different unique frequency profile that can be measured for different molecules.

More than a dozen different compounds can be studied through MRS, from antioxidants, neurotransmitters, and those involved in neural energetics. Of these, the most commonly measured metabolites include N-acetyl aspartate (NAA), which primarily exists in neurons and serves as a proxy for neuronal density, Choline-containing (tCho) compounds which reflect

phospholipid homeostasis, and Creatine-containing (tCr) compounds involved in neural energetics (Moffett et al., 1991; Tallan, 1957; Urenjak et al., 1992). Myo-Inositol (Myo) whose exact function remains elusive (but is often thought to reflect neuroinflammation) and the combined signal (Glx) from Glutamate (Glu) and Glutamine (Gln) are also among the most commonly observed metabolites as they appear to be dysregulated in a variety of diseased populations (Öz et al., 2014). Over the last decade, editing techniques, analysis software, and high field (>3T) spectroscopy have made y-amino butyric acid (GABA) and Glu (*independent from Gln*), the primary inhibitory and excitatory neurotransmitters, more common targets (Edden et al., 2014; Mikkelsen et al., 2017, 2019). While less common, the antioxidants Glutathione (GSH) and Ascorbate (Asc) can also be identified (Choi et al., 2021; Oeltzschner et al., 2019; Saleh et al., 2016; Terpstra et al., 2006).

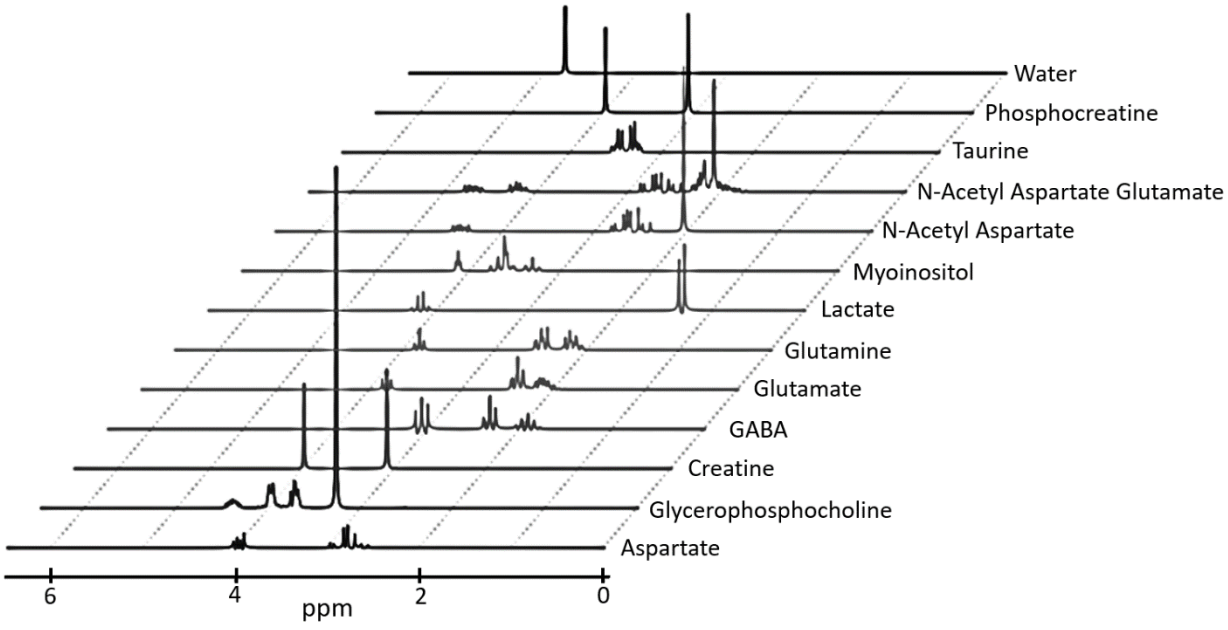


Figure 1.1.: Basis set including some commonly observed brain metabolites.

While anatomical interrogations through Magnetic Resonance Imaging (MRI) and evaluations into structural integrity through Diffusion Weighted Imaging (DWI) have become commonplace, collecting static and/or dynamic chemical information through MRS has yet to be widely adopted. Despite the wealth of biomarkers, MRS still faces many challenges that limits its use to a relatively small number of dedicated research facilities. Understandably, developing MRS protocols and analysis software requires a much deeper understanding of spin physics and quantum mechanics compared to other forms of MR (Hanson, 2008). Though, there are no significant barriers for clinicians to simply employ such methods once developed, challenging the MRS field to take an introspective look into the roadblocks that stand in the way.

As such, in 2011, the MRS community developed a consensus group consisting of experienced neuroscientists and spectroscopists. In 2014, the group outlined the clinical utility for a wide array of brain disease and outlined what research steps would need to be taken to make MRS more clinically relevant (Öz et al., 2014). In recent years, many of the technical hurdles were addressed, especially those involved in the methodological standardization for data acquisition (Choi et al., 2021; Cudalbu et al., 2021; Kreis et al., 2020; Lin et al., 2021; Near et al., 2021; Öz et al., 2021). There have been a handful of follow-ups to the original consensus paper documenting these successes, but also listing new issues that should be addressed (Wilson et al., 2019). Today, what is most needed, is the development of more advanced software for data analysis and quantification that can overcome many of the limitations that lead to poor data quality, reproducibility, and accuracy (Wilson et al., 2019). Furthermore, standardized and open-source software packages are encouraged in order to appropriately compare data collected across research sites.

1.3. Challenges Facing MRS:

1.3.1. Low Signal to Noise:

Many issues in MR revolve around low signal to noise (SNR), requiring long acquisitions to provide appropriate measurements. Unfortunately, MRS is specifically hampered by low SNR to an even larger extent that its MR counterparts. While most forms of MR take advantage of the highly abundant water molecules within the brain, MRS targets low concentrations (~1mM) chemicals. To accomplish this, an MRS sessions can be 30 minutes to 1 hour. During this time, multiple transients (*or averages*) are acquired every ~2-3 seconds each and then averaged to remove uncorrelated noise as shown in Figure 1.1. This method can be time-consuming, especially for deeper brain regions where the strength of the metabolite signal falls off following the inverse-square law. Subject motion across the session can also greatly impede data quality and provide significant barriers for populations that have difficulties remaining still for the prolonged periods required to retrieve high SNR data. Finally, the temporal sensitivity from averaging across several minutes puts many functional measurements of dynamic chemical processes outside of reach.

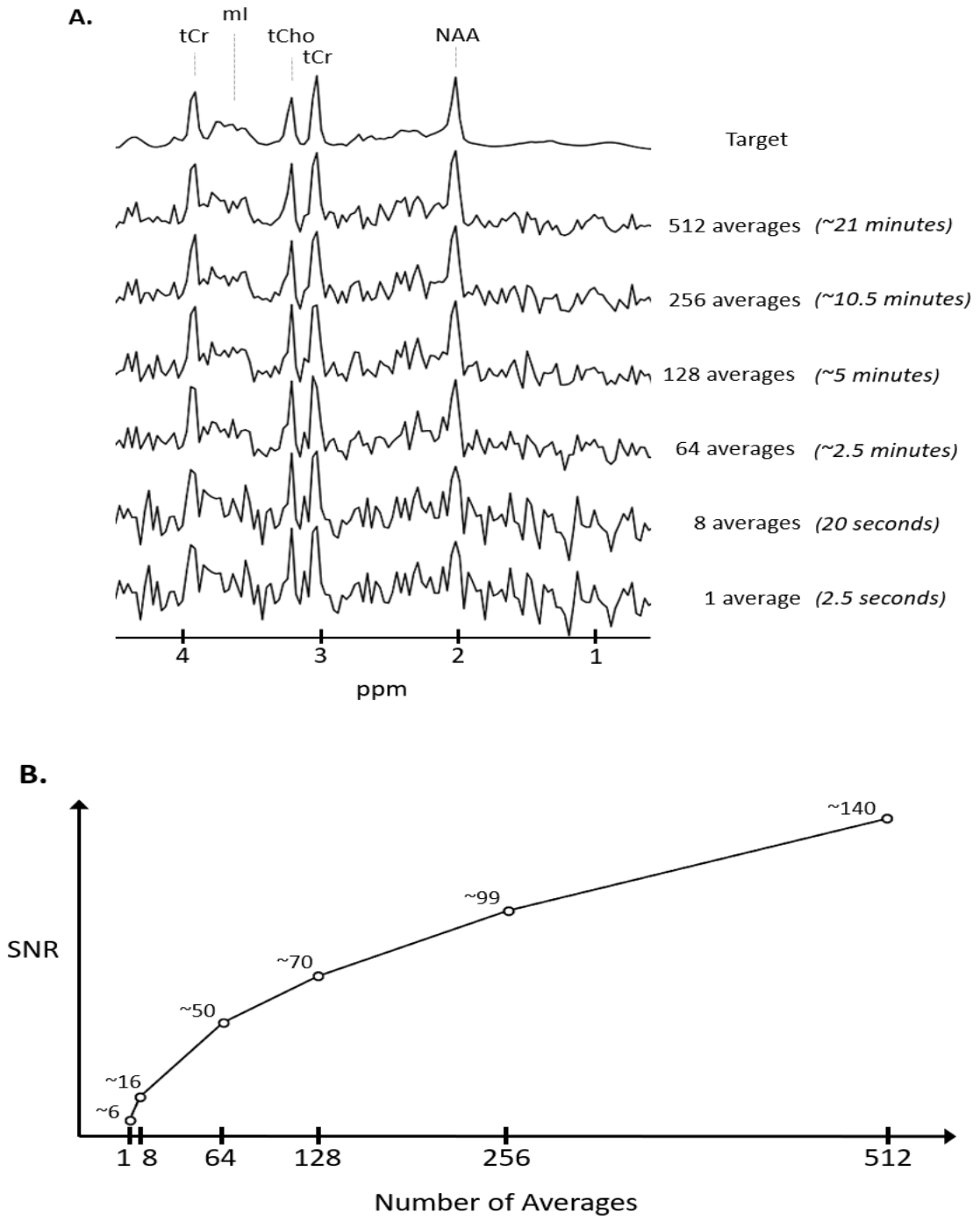


Figure 1.2.: (A) Simple example of the average signal after a single and multiple transients and (B) resulting SNR after collecting multiple averages for an initially low (~6) SNR acquisition. SNR is defined as amplitude of NAA singlet (2.008ppm) divided by the standard deviation of the noise.

1.3.2. Overlapping Signals:

Another issue that spectroscopy faces is disentangling a multitude of overlapping signals. Figure 1a shows a breakdown of many of the individual chemicals observed within the brain. Figure 1b shows an example of the raw trace that could be acquired during an MRS sessions. It can be seen that many chemicals share the same or similar frequencies, or peaks. Take for instance, y-amino-butyric acid (GABA), which resides at 3ppm along with Creatine, Phosphocreatine, and underlying Macromolecule signals as shown in Figure 1.2. To make matters more difficult, GABA is ostensibly completely covered by the total creatine signal where the concentration is ~10x less. While each chemical has a unique spectral signature, these overlapping signals pose difficulties for traditional quantification software.

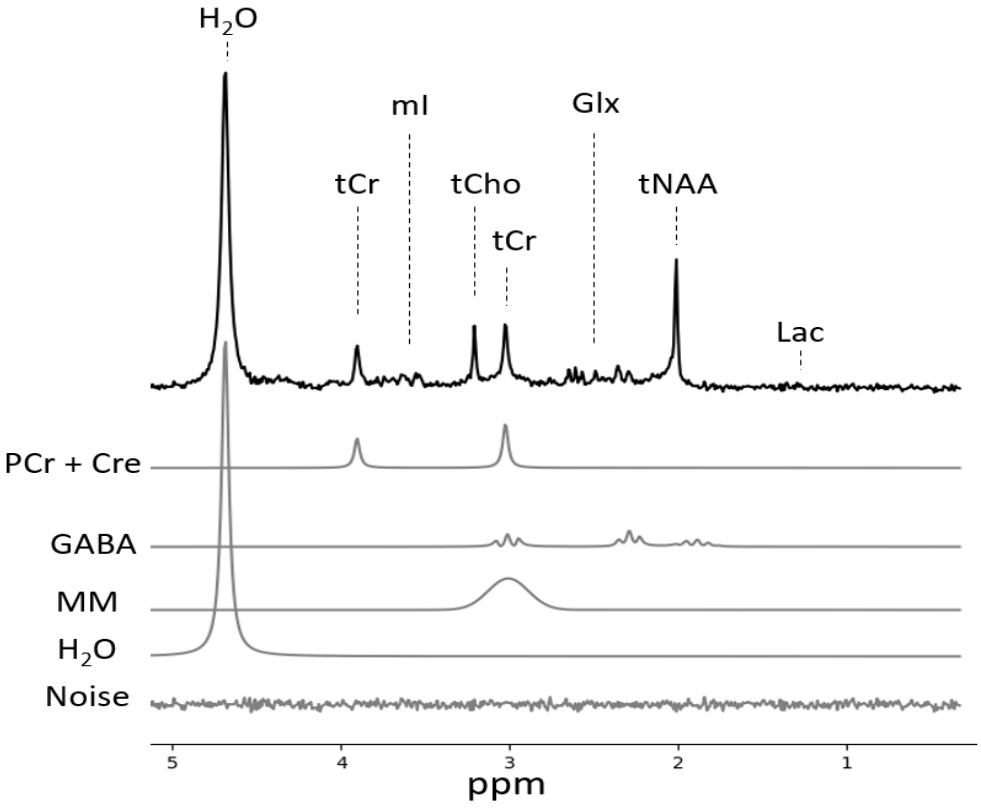


Figure 1.3.: The GABA peak of interest centered at ~3ppm is overlapped by tCr, macromolecules, and noise. The residual water peak (not to scale) is shown to the left of the metabolites with long peak tails that extend underneath and distort metabolite signals.

1.3.3. Residual Water:

The water signal is a single peak (singlet) that shows up at ~4.68ppm has an amplitude approximately 100x of the metabolite amplitudes. During acquisition, a series of water suppression pulses precede the data acquisition in order to suppress the overwhelming signal. Following the water suppression, a residual water signal remains as seen in Figure 1.2. Even after water suppression, the residual water peak can be between 1x-20x greater in size relative to the metabolite signals. Despite this signal being downfield from the primary metabolites of interest (1.00ppm-4.00ppm), the residual water signal can have long tails extending off the base of its peak into the metabolite peaks. The residual water signal tends to contain many components with different phase and amplitudes from incomplete water suppression which makes its removal in post-processing quite challenging. Without full removal, the residual water signal can greatly impact the metabolite amplitudes leading to unreliable concentration estimates.

1.3.4. Spurious Echoes:

Spurious echoes are the result of unwanted signals appearing for a small window of time during the acquisition period. The result is a large fluctuating waveform overlapping and obstructing the targeted metabolite data. During the scan protocol, a number of steps are taken (*crusher gradients, phase cycling, outer volume suppression, etc.*) to eliminate the undesirable signals. However, these spectral artifacts still show up within data, primarily due to technical malfunctions.

1.4. Research Objectives:

In line with these ideals, my primary research goal has been and continues to be overcoming limitations to make MRS more universally available. In this dissertation, I describe the research we completed towards developing methods surrounding MRS data processing and

analysis. Specifically, I'll describe the results accomplishments from 2 projects. In part-1, I describe developing the open-source COHERENC database and meta-analysis of healthy and clinical brain MRS data. Part-2 describes our development of a neural network training dataset for MRS as well as our work in developing deep-learning-based data analysis techniques.

1.4.1. Objective 1: COHERENC: A freely available online database for Magnetic Resonance Spectroscopy Metabolite Concentrations and Relaxation Parameters in the Healthy and Diseased Brain (*Paper Abstract*):

Proton (^{1}H) Magnetic Resonance Spectroscopy (MRS) is a noninvasive tool capable of studying brain metabolites *in-vivo*. The noninvasive nature of this method provides a powerful tool for researchers and clinicians, but also makes determining ground truths and expectations challenging. As such, standardization within the field has become a priority leading to the development of universal pulse sequences, methodological consensus recommendations from leaders in the field, and a prioritization for the development of open-source preprocessing and quantification software packages. As ground-truths are scarcely available for the validation of these various methodologies, data simulations have become an important tool. Especially for the development of deep learning and machine learning algorithms, these simulations must be able to produce accurate spectra which capture all the nuances of *in-vivo* data. Otherwise, we run the risk that such algorithms will misattribute signal variance or otherwise distort our signals. Therefore, to support these endeavors, we sought to determine the physiological ranges and relaxation rates of brain metabolites which can be used both in data simulations and as reference estimates in the human brain. To do this, we've systematically reviewed the literature to build a freely available database that currently includes nearly 500 spectroscopy articles. In Part-1 of this article we describe the systematic approach to developing the database and assess the

heterogeneity across studies. In Part-2, we employ a meta-analytic approach to explore metabolite concentrations and T2 relaxation. First, we determine the physiological ranges of concentrations of brain metabolites in the healthy brain. Next, we model how these metabolite concentrations change across 25 clinical populations. Finally, we use multiple meta-regression model to determine how T2 relaxation changes in human and animal studies under various acquisition protocols.

1.4.2. Objective 2: Deep learning for MRS Data and the AGNOSTIC Benchmark Dataset (Paper Abstract):

Proton (^{1}H) Magnetic Resonance Spectroscopy (*MRS*) is a powerful method capable of studying brain metabolites *in-vivo*. As a clinical research tool, accurate and reliable quantification is a necessity, which has driven the MRS community to prioritize the reliability and standardization quantification software packages. Though, low signal to noise, artifacts, and overlapping signals still provide significant challenges, especially for deep brain regions and multi-voxel acquisition. Data quality can be improved through collecting multiple transients or employing multi-scan (*editing, nulling, etc.)* methods, but time constraints may limit what can be done during a scan session. Machine learning and deep neural networks have the potential to aid in many of these fundamental challenges associated with MRS data. While recent neural network models have already demonstrated some proof of concept, a software package that utilizes deep learning has yet to be shown to either complement traditional methods or perform as well as standard techniques in practice. Furthermore, little is known about how changes to network architecture, hyperparameters, and training sets influence the overall success. Unlike fields like computer vision, a benchmark dataset (*i.e., ImageNet*) does not exist for MRS which makes developing these models or evaluating performance across models impossible. As such, we

sought to develop the first training dataset for MRS. Here, we provide the AGNOSTIC training dataset for MRS. Included in the dataset are 500,000 healthy and clinical examples. Using this dataset, we further demonstrate the performance and limitations of various neural networks models.

1.5. Future Directions:

Here, I include a short discussion on developing new acquisition protocols to evaluate age-related pathology and Late-Onset Alzheimer's Disease. Unfortunately, due to the COVID-19 pandemic, the actual experiments were not able to be carried out. With the pandemic, I directed my focus solely towards advancing analysis methodologies. However, the research goals and hypotheses described within this section will be an area I plan to revisit in the future.

1.5.1. Characterization of Metabolic Dysregulation and Efficacy of Caloric Restriction using in-vivo $^{13}C/^{31}P$ MRS in aging and APOE as model of Late-Onset Alzheimer's Disease:

Numerous research reports are now pointing towards metabolic dysfunction as a key driver of Late-Onset Alzheimer's Disease (*LOAD*). It is therefore crucial to characterize the changes that occur within metabolic pathways and develop metabolic-specific therapeutics. Oxidative stress, lipid peroxidation, and decreased glucose utilization have been identified as early features of LOAD which may specifically point towards impairments in neuronal TCA Cycle. Importantly, aging and the APOE ε4 allele, the top two risk factors for LOAD, are also known to have declines in the rate of neuronal TCA Cycle ($V_{TCA\text{-}Neuron}$) and are vulnerable to damage from peroxidated lipids. In the present longitudinal study, we will use a Knock-In humanized-APOE mouse model to investigate the individual contributions and interaction between aging (*6-months, 12-months, and 18-months*) and the APOE allele (*low-risk ε3 vs. high-risk ε4*) while evaluating the efficacy of 6-months of Caloric Restriction (*CR*) relative to ad

libitum (*AL*) feeding. Using *in-vivo* ^{13}C and ^{31}P NMR spectroscopy, we will directly measure the rate of neuronal TCA Cycle ($V_{TCA\text{-}Neuron}$) and the concentration of NAD^+ (*and NADH*), respectively. Following the final timepoint (18-months) we will collect post-mortem brain tissue to further determine the predictive value of $V_{TCA\text{-}Neuron}$ and NAD^+ concentrations on astrocyte lipid droplets and Sirtuin activity, respectively. Here, we hypothesize that restoring neuronal TCA cycle through CR will prevent LOAD-related cognitive decline through ameliorating age and *ε4 metabolic impairment*. A critical advantage of this approach is that it is directly translatable into non-invasive studies in humans. As such, in this study, we focus on providing a robust profile of LOAD prognosis and potential interventions through CR while developing safe and non-invasive NMR-based biomarkers.

1.6. References:

Choi, I. Y., Andronesi, O. C., Barker, P., Bogner, W., Edden, R. A. E., Kaiser, L. G., Lee, P., Marjańska, M., Terpstra, M., & de Graaf, R. A. (2021). Spectral editing in 1H magnetic resonance spectroscopy: Experts' consensus recommendations. *NMR in Biomedicine*, *34*(5), 1–18. https://doi.org/10.1002/nbm.4411

Cudalbu, C., Behar, K. L., Bhattacharyya, P. K., Bogner, W., Borbath, T., de Graaf, R. A., Gruetter, R., Henning, A., Juchem, C., Kreis, R., Lee, P., Lei, H., Marjańska, M., Mekle, R., Murali-Manohar, S., Považan, M., Rackayová, V., Simicic, D., Slotboom, J., … Mlynárik, V. (2021). Contribution of macromolecules to brain 1H MR spectra: Experts' consensus recommendations. *NMR in Biomedicine*, *34*(5), 1–24. https://doi.org/10.1002/nbm.4393

Edden, R. A. E., Puts, N. A. J., Harris, A. D., Barker, P. B., & Evans, C. J. (2014). Gannet: A batch-processing tool for the quantitative analysis of gamma-aminobutyric acid-edited MR spectroscopy spectra. *Journal of Magnetic Resonance Imaging*, *40*(6), 1445–1452.

https://doi.org/10.1002/jmri.24478

Hanson, L. G. (2008). Is quantum mechanics necessary for understanding magnetic resonance? *Concepts in Magnetic Resonance Part A, 32A*(5), 329–340. https://doi.org/10.1002/cmr.a.20123

Kreis, R., Boer, V., Choi, I. Y., Cudalbu, C., de Graaf, R. A., Gasparovic, C., Heerschap, A., Krššák, M., Lanz, B., Maudsley, A. A., Meyerspeer, M., Near, J., Öz, G., Posse, S., Slotboom, J., Terpstra, M., Tkáč, I., Wilson, M., & Bogner, W. (2020). Terminology and concepts for the characterization of in vivo MR spectroscopy methods and MR spectra: Background and experts' consensus recommendations. In *NMR in Biomedicine* (Issue May 2020). https://doi.org/10.1002/nbm.4347

Lin, A., Andronesi, O., Bogner, W., Choi, I. Y., Coello, E., Cudalbu, C., Juchem, C., Kemp, G. J., Kreis, R., Krššák, M., Lee, P., Maudsley, A. A., Meyerspeer, M., Mlynarik, V., Near, J., Öz, G., Peek, A. L., Puts, N. A., Ratai, E. M., … Mullins, P. G. (2021). Minimum Reporting Standards for in vivo Magnetic Resonance Spectroscopy (MRSinMRS): Experts' consensus recommendations. *NMR in Biomedicine*, *34*(5), e4484. https://doi.org/10.1002/nbm.4484

Mikkelsen, M., Barker, P. B., Bhattacharyya, P. K., Brix, M. K., Buur, P. F., Cecil, K. M., Chan, K. L., Chen, D. Y. T., Craven, A. R., Cuypers, K., Dacko, M., Duncan, N. W., Dydak, U., Edmondson, D. A., Ende, G., Ersland, L., Gao, F., Greenhouse, I., Harris, A. D., … Edden, R. A. E. (2017). Big GABA: Edited MR spectroscopy at 24 research sites. *NeuroImage*, *159*(July), 32–45. https://doi.org/10.1016/j.neuroimage.2017.07.021

Mikkelsen, M., Rimbault, D. L., Barker, P. B., Bhattacharyya, P. K., Brix, M. K., Buur, P. F., Cecil, K. M., Chan, K. L., Chen, D. Y. T., Craven, A. R., Cuypers, K., Dacko, M., Duncan,

N. W., Dydak, U., Edmondson, D. A., Ende, G., Ersland, L., Forbes, M. A., Gao, F., … Edden, R. A. E. (2019). Big GABA II: Water-referenced edited MR spectroscopy at 25 research sites. *NeuroImage*, *191*(March), 537–548. https://doi.org/10.1016/j.neuroimage.2019.02.059

Moffett, J. R., Aryan Namboodiri, M. A., Cangro, C. B., & Neale, J. H. (1991). Immunohistochemical localization of N-acetylaspartate in rat brain. *NeuroReport*, *2*(3), 131–134. https://doi.org/10.1097/00001756-199103000-00005

Near, J., Harris, A. D., Juchem, C., Kreis, R., Marjańska, M., Öz, G., Slotboom, J., Wilson, M., & Gasparovic, C. (2021). Preprocessing, analysis and quantification in single-voxel magnetic resonance spectroscopy: experts' consensus recommendations. *NMR in Biomedicine*, *34*(5), 1–23. https://doi.org/10.1002/nbm.4257

Oeltzschner, G., Saleh, M. G., Rimbault, D., Mikkelsen, M., Chan, K. L., Puts, N. A. J., & Edden, R. A. E. (2019). Advanced Hadamard-encoded editing of seven low-concentration brain metabolites: Principles of HERCULES. *NeuroImage*, *185*(September 2018), 181–190. https://doi.org/10.1016/j.neuroimage.2018.10.002

Öz, G., Alger, J. R., Barker, P. B., Bartha, R., Bizzi, A., Boesch, C., Bolan, P. J., Brindle, K. M., Cudalbu, C., Dinçer, A., Dydak, U., Emir, U. E., Frahm, J., González, R. G., Gruber, S., Gruetter, R., Gupta, R. K., Heerschap, A., Henning, A., … Kauppinen, R. A. (2014). Clinical Proton MR Spectroscopy in Central Nervous System Disorders. *Radiology*, *270*(3), 658–679. https://doi.org/10.1148/radiol.13130531

Öz, G., Deelchand, D. K., Wijnen, J. P., Mlynárik, V., Xin, L., Mekle, R., Noeske, R., Scheenen, T. W. J., Tkáč, I., Andronesi, O., Barker, P. B., Bartha, R., Berrington, A., Boer, V.,

Cudalbu, C., Emir, U. E., Ernst, T., Fillmer, A., Heerschap, A., … Wilson, M. (2021). Advanced single voxel 1H magnetic resonance spectroscopy techniques in humans: Experts' consensus recommendations. *NMR in Biomedicine*, *34*(5), 1–18. https://doi.org/10.1002/nbm.4236

Saleh, M. G., Oeltzschner, G., Chan, K. L., Puts, N. A. J., Mikkelsen, M., Schär, M., Harris, A. D., & Edden, R. A. E. (2016). Simultaneous edited MRS of GABA and glutathione. *NeuroImage*, *142*, 576–582. https://doi.org/10.1016/j.neuroimage.2016.07.056

Tallan, H. H. (1957). STUDIES ON THE DISTRIBUTION OF N-ACETYL-l-ASPARTIC ACID IN BRAIN. *Journal of Biological Chemistry*, *224*(1), 41–45. https://doi.org/10.1016/S0021-9258(18)65008-2

Terpstra, M., Marjanska, M., Henry, P. G., Tkáč, I., & Gruetter, R. (2006). Detection of an antioxidant profile in the human brain in vivo via double editing with MEGA-PRESS. *Magnetic Resonance in Medicine*, *56*(6), 1192–1199. https://doi.org/10.1002/mrm.21086

Urenjak, J., Williams, S. R., Gadian, D. G., & Noble, M. (1992). Specific Expression of N-Acetylaspartate in Neurons, Oligodendrocyte-Type-2 Astrocyte Progenitors, and Immature Oligodendrocytes In Vitro. *Journal of Neurochemistry*, *59*(1), 55–61. https://doi.org/10.1111/j.1471-4159.1992.tb08875.x

Wilson, M., Andronesi, O., Barker, P. B., Bartha, R., Bizzi, A., Bolan, P. J., Brindle, K. M., Choi, I., Cudalbu, C., Dydak, U., Emir, U. E., Gonzalez, R. G., Gruber, S., Gruetter, R., Gupta, R. K., Heerschap, A., Henning, A., Hetherington, H. P., Huppi, P. S., … Howe, F. A. (2019). Methodological consensus on clinical proton MRS of the brain: Review and recommendations. *Magnetic Resonance in Medicine*, *82*(2), 527–550.

https://doi.org/10.1002/mrm.27742

Chapter 2. Magnetic Resonance Physics:

Nuclear magnetic resonance (NMR) was first observed in 1938 by the American physicist, Isidor Rabi. After being proposed by Cornelius Gorter, Rabi and his team demonstrated the ability to measure the energy absorption/emission of lithium and chlorine following nuclear induction by an oscillating magnetic field. Nearly 10 years after its initial discovery, Edwards Purcell and Felix Bloch, working in separate labs independently, both demonstrated the ability to measure NMR in condensed matter (water and paraffin). In the following decades, NMR spectroscopy would become a research tool, primarily for field of chemistry and physics. Later in the 1970s, work by Paul Lauterbur and Peter Mansfield enabled spatial localization, and thus magnetic resonance imaging, by introducing a linear field gradient. These early experimental results laid the foundation for the development of the magnetic resonance tools used today.

2.1 Magnetic Resonance Physics:

2.1.1. Quantum Spin:

MR is the process by which an atomic nucleus absorbs/emits energy by transitioning between quantum spin energy levels. This quantum spin (commonly denoted as *I*) describes the intrinsic angular momentum of a particle and exists in $\frac{1}{2}$ steps (0, $\frac{1}{2}$, 1, $\frac{3}{2}$ …). Just like a common bar magnet, particles that have a non-zero spin state have a magnetic moment, or magnetic dipole, that interacts with electromagnetic fields. In the context of MR, we are specifically concerned with nuclear (*from the atomic nucleus*) spin. There are many non-zero spin nuclei that

can be studied with MR, including, but not limited to: ^{1}H, ^{2}H, ^{13}C, ^{15}N, ^{17}O, and ^{31}P. Here, we will focus exclusively upon spin $\frac{1}{2}$ nuclei with an emphasis on the ^{1}H, or proton.

2.1.2. Zeeman Splitting:

As previously stated, magnetic resonance requires a transition between spin states. The presence of a magnetic field leads to Zeeman splitting (*see figure 2.1*) which produces 2 energy levels for a spin $\frac{1}{2}$, characterized by the quantum number m, where m = $\pm\frac{1}{2}$. The 2 spin energy levels are commonly referred to as spin up (α; m = + $\frac{1}{2}$) and spin down (β; m = - $\frac{1}{2}$) and their energy levels are given in equation 2.1. These two energy levels are perhaps better understood to be eigenvalues of the Hamiltonian ($\widehat{H}$), or energy operator. This means that spin can be represented as a linear combination of these two eigenvalues multiplied by a constant and thus a superposition of both states.

$$E_{\alpha,\beta} = m\hbar\gamma B_0 \tag{2.1.}$$

These energy levels can be seen to be the dependent of the nuclear gyromagnetic ratio (γ) which is unique to each nucleus, the quantum number (*m*), and the field strength (B_0) in Tesla. The energy transitions between the 2 states can therefore be shown to be:

$$\Delta E_{\boldsymbol{\alpha}\leftrightarrows\boldsymbol{\beta}} = \hbar\gamma B_0 \tag{2.2.}$$

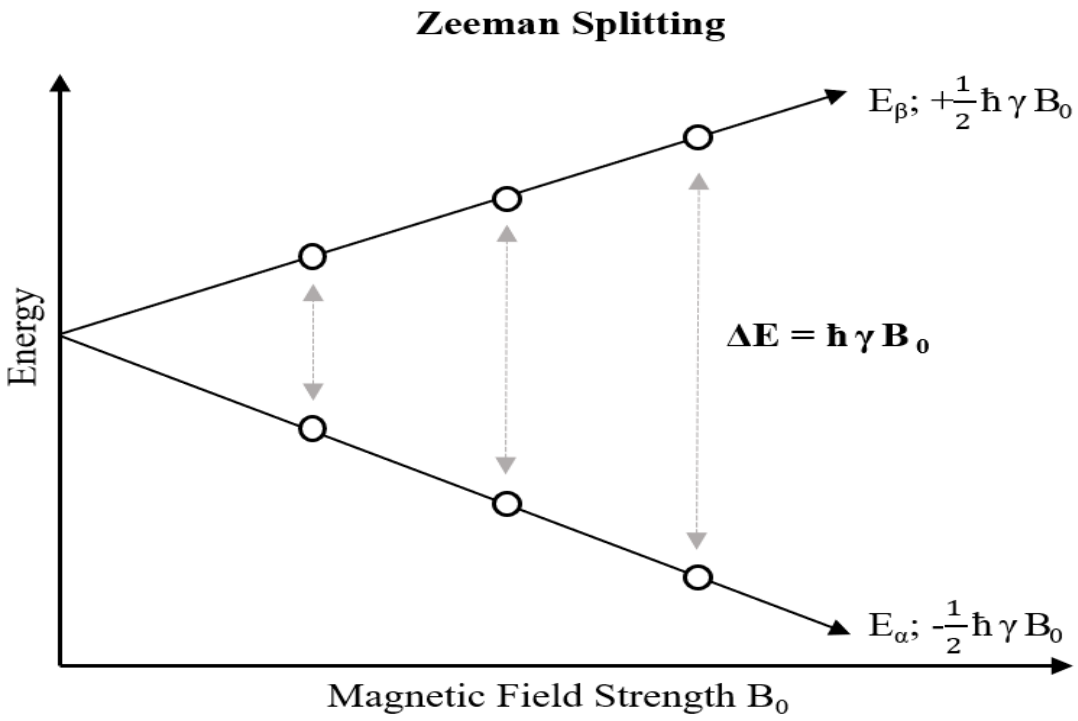


Figure 2.1.: Zeeman Splitting of a spin $\frac{1}{2}$. The energy levels are shown for the two spin states for an increasing field strength (B_0). The energy required for a transition between states can be calculated from the quantum number (m; $\pm\frac{1}{2}$), the reduced plank constant (ħ), nucleus' gyromagnetic ratio (γ), and the field strength (B_0) in Tesla.

In addition to increasing the energy between spin states, the nuclear spins will also tend toward the lowest energy state with increasing field strength. The distribution among spin states can be modeled by the Boltzmann Distribution shown in Equation 2.3 for a spin $\frac{1}{2}$ system.

$$\frac{\alpha}{\beta} = e^{(-\frac{\hbar\gamma B_0}{\kappa T})} \tag{2.3.}$$

2.1.3. Magnetic Moment and Precession:

In the absence of a magnetic field the spin angular momentum results in random orientations of the nuclear magnetic moments. When a magnetic field (B_0) is introduced, a torque is created that acts perpendicular to the field and spin angular momentum described above. This drives nuclei to precess at the Larmor frequency (ω_0) shown in equation 2.4.

$$\omega_0 = \gamma B_0 \tag{2.4.}$$

Additionally, the torque encourages nuclei to orient and point with a magnetic moment along the axis of the magnetic field. Increasing the magnetic field will encourage more nuclei to align along the magnetic field axis, with a slight preference in the direction of the applied field. As described by the Boltzmann distribution the preference towards the lower energy spin state will lead to a net magnetization, or bulk magnetization vector, that points in the direction of the applied field.

2.1.4. MR Signal:

In order to generate a signal, radiofrequency (*RF*) pulses (B_1) are applied orthogonal to the main magnetic field axis or B_0. When this happens, the bulk magnetization vector is rotated into the transverse plane as the spins are now precessing with B_1. Once the RF pulse is removed, a receive coil can be used to measure the electrical current that arises from the bulk magnetization returning to the B_0 axis. While nuclear spin is a quantum effect, the observed signal in magnetic resonance is <u>not</u>. This may seem counterintuitive, but the actual measured signal during an experiment originates from the movement of the bulk magnetization vector that arises from the resulting torque of average spin state.

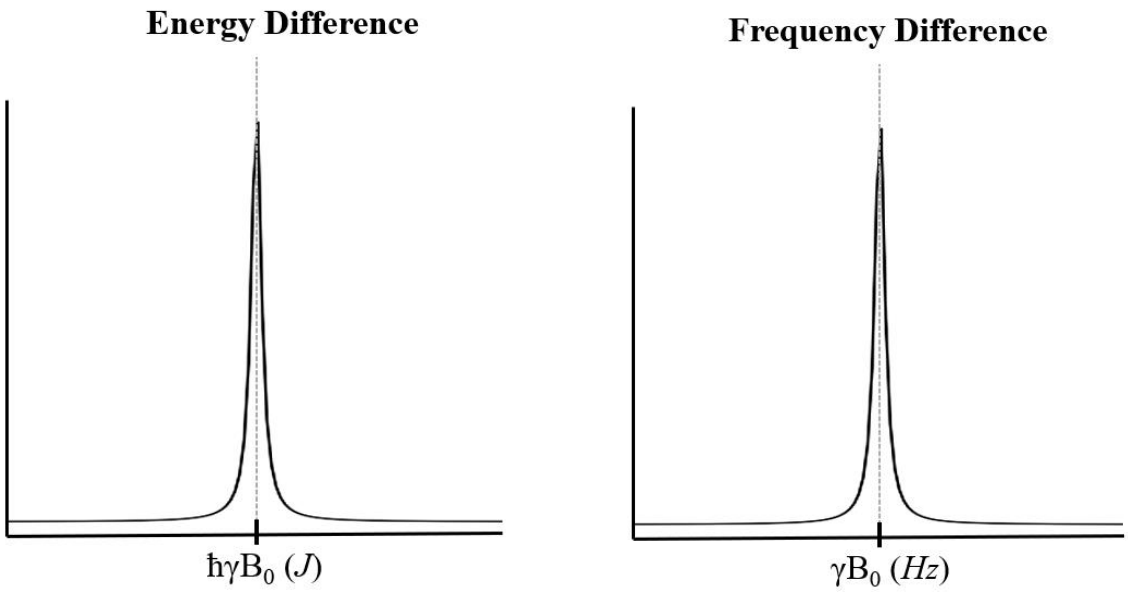


Figure 2.2.: Spectral peak location shown represented as an energy compared to frequency. This conversion can be made as energy is directly related to frequency by Planck's constant as shown in equation 2.5.

2.1.5. Frequency Units:

As angular momentum is commonly given in units of frequency, it's natural and convention for spectroscopists to also maintain calculations in frequency units. Equation 2.5. gives a breakdown showing the conversion from energy units in Jules (J) to frequency units in either Hertz (Hz) or Radians per second by removing Planck's constant. Additionally, Figure 2.2. shows an example peak that results from an energy transition.

$$\begin{aligned} \Delta E_{\alpha\to\beta} &= h\nu_{\alpha\to\beta} \\ \frac{\Delta E_{\alpha\to\beta}}{h} &= \nu_{\alpha\to\beta} \\ \nu_{\alpha\to\beta} &= \frac{\gamma B_0}{2\pi} \quad \text{(radians/second)} \end{aligned} \tag{2.5.}$$

$$\nu_{\alpha \rightarrow \beta} = \gamma B_0 \quad \text{(Hertz)}$$

2.1.6. Chemical Shift:

Thus far, we've primarily described this method with a single resonant frequency, characterized by the Larmor frequency, that corresponds with the hydrogen nucleus. In all actuality, over the course of an MRS experiment we observe hydrogen nuclei from a vast number of hydrogen nuclei from different molecules that have slightly different resonant frequencies depending on their local magnetic and chemical environment. So, rather than targeting a single frequency, we instead use RF pulses that cover a range of frequencies, referred to as the bandwidth or spectral width (sw). As the Larmor frequency is dependent on the static magnetic field strength, the location of these peaks will be different across scanners that typically employ fields between 1.5 Tesla-7.0 Tesla for human subjects. As such a final convention is to normalize the frequency axis to units of parts per million (ppm) by dividing the frequency axis by the Larmor frequency. Despite this sounding like it may be a measure of concentration, it's actually a simple and convenient representation of frequency units that is standardized.

3. COHERENC: A freely available online database for Magnetic Resonance Spectroscopy Metabolite Concentrations and Relaxation Parameters in the Healthy and Diseased Brain (*Paper In Prep*)

3.1. Title:

COHERENC: A freely available online database for Magnetic Resonance Spectroscopy Metabolite Concentrations and Relaxation Parameters in the Healthy and Diseased Brain.

3.2. Abstract:

Proton (^{1}H) Magnetic Resonance Spectroscopy (*MRS*) is a noninvasive tool capable of studying brain metabolites *in-vivo*. The noninvasive nature of this method provides a powerful tool for researchers and clinicians, but also makes determining ground truths and expectations challenging. As such, standardization within the field has become a priority leading to the development of universal pulse sequences, methodological consensus recommendations from leaders in the field, and a prioritization for the development of open-source preprocessing and quantification software packages. As ground-truths are scarcely available for the validation of these various methodologies, data simulations have become an important tool. Especially for the development of deep learning and machine learning algorithms, these simulations must be able to produce accurate spectra which capture all the nuances of *in-vivo* data. Otherwise, we run the risk that such algorithms will misattribute signal variance or otherwise distort our signals. Therefore, to support these endeavors, we sought to determine the physiological ranges and relaxation rates of brain metabolites which can be used both in data simulations and as reference estimates in the human brain. To do this, we've systematically reviewed the literature to build a freely available database that currently includes nearly 500 spectroscopy articles. In Part-1 of this article we describe the systematic approach to developing the database and assess the

heterogeneity across studies. In Part-2, we employ a meta-analytic approach to explore metabolite concentrations and T2 relaxation. First, we determine the physiological ranges of concentrations of brain metabolites in the healthy brain. Next, we model how these metabolite concentrations change across 25 clinical populations. Finally, we use multiple meta-regression model to determine how T2 relaxation changes in human and animal studies under various acquisition protocols.

3.3. Introduction:

3.3.1. Background:

Proton (^{1}H) Magnetic Resonance Spectroscopy (*MRS*) is a noninvasive tool that can be used to study *in-vivo* brain metabolites in the human. At least a dozen different compounds are commonly studied, from antioxidants to those involved in neural energetics. Of these, the most commonly measured metabolites include N-acetyl aspartate (*NAA*) which generally serves as a proxy for neuronal density, Choline-containing (*tCho*) compounds which reflect phospholipid homeostasis, and Creatine-containing (*tCr*) compounds which, in addition to their role in neural energetics, have commonly served as an internal reference compound. Myo-Inositol (*Myo*) whose exact function remains elusive (but is often attributed in some form to neuroinflammation) and the combined signal (*Glx*) from Glutamate (*Glu*) and Glutamine (*Gln*) are also among the most commonly observed metabolites as they appear to be dysregulated in a variety of diseased populations. Over the last decade, editing techniques, analysis software, and high field (*>3T*) spectroscopy have made y-amino butyric acid (*GABA*) and Glutamate, the primary inhibitory and excitatory neurotransmitters, more common targets (Edden et al., 2014; Oeltzschner et al., 2019; Saleh et al., 2016). While less common, the antioxidants Glutathione (*GSH*) and Ascorbate (*Asc*) can also be identified. A full description of the metabolites included

in the COHERENC can be found in review works by Harris, Henning, and Ross & Bluml (*A. D. Harris et al., 2017; Henning, 2018; Ross & Bluml, 2001)*.

^{1}H-MRS has continued gaining popularity, but the method has largely remained a research tool that requires experienced teams. However, given how biologically informative these metabolites are to a variety of clinical populations, the field has sought to improve the reliability and standardize methods to make it more universally accessible. Recent developments of universal acquisition methods and consensus recommendations from experts have provided the ability to effectively acquire data (Choi et al., 2021; Cudalbu et al., 2021; Kreis et al., 2020; Lin et al., 2021; Near et al., 2021; Öz et al., 2021). Additionally, an admirable philosophy of open-source software and analysis techniques have laid down a strong base for the widespread use of MRS. However, validating these methods remains challenging, as the lack of ground-truths poses a limitation.

A number of free software packages are available that can perform full quantum mechanical density matrix simulations that include all the required information for metabolite chemical shifts and coupling constants (Hogben et al., 2011; Simpson et al., 2017; Smith et al., 1994; Starčuk & Starčuková, 2017). To a large extent these simulation tools are all that is needed to develop pulse sequences. However, validating preprocessing and quantification methods requires another layer of complexity by including appropriate metabolite concentrations and relaxation rates as well as realistic signal to noise and acquisition artifacts. These factors have become especially important for machine learning algorithms and deep learning approaches which have difficulty generalizing to features not captured within simulated training data. The Fit Challenge pioneered one of the first large-scale efforts to create realistic signals to test the

performance of the various software packages and successfully demonstrated a need for further advanced in simulations (*Marjańska et al., 2022)*.

3.3.2. Purpose:

In order to further these goals, we aimed to create an online open-source database of metabolite concentrations and relaxation parameters in healthy and diseased subject populations. Nearly 500 papers have been added to the database with resulting concentrations and relaxation rates. For each value added, the publication information, experimental parameters, and data acquisition methods are included. This article is organized into two parts describing the database and simulation parameters, respectively. Part-1 describes the systematic approach to identifying spectroscopy articles for inclusion and further details the experimental information available within the database. Part-2 is a meta-analysis providing the physiological ranges of metabolites concentrations and relaxation rates. Specifically, in Part-2, we performed three separate analyses: 1) investigation into healthy brain metabolite concentrations, 2) a model of how these concentrations change in 25 clinical populations, 3) a model to predict and account for the highly variable T2 results. The breadth of information from articles spanning different subject populations using various acquisition protocols and analysis pipelines, make this work unique. In addition to researchers developing analyses new tools, we hope this work will further contribute towards the greater adoption of MRS within the clinical community.

3.4 Methods:

3.4.1. Database Methods:

3.4.1.1. Search Methods:

In building the database, we utilized the PRISMA guidelines to identify publications and determine eligibility for inclusion (Moher et al., 2009; Page et al., 2021). Searches were

conducted on the PubMed, Web of Science, and Scopus databases. Separate searches were carried out to specifically identify publications that either quantified metabolite concentrations or T2 relaxation, herein referred to as concentration study and relaxation study, respectively. The original search for both was conducted in August of 2021. An additional follow-up search was then conducted at the time of writing, March 2022, to ensure all publications through 2021 were included. Upon searching, no limitation for publication date was set and only articles available in English were included. Figure 1. shows the PRISMA flowchart for both the concentration and relaxation studies and reflects the process of building the database. The information for the total number of articles that went into the meta-analyses and modeling performed is described within the 'study description' section within the Results.

Concentration Study Search:

PubMed:

((Magnetic Resonance Spectroscopy [mh] OR "Magnetic Resonance Spectroscopy" OR MRS OR MRSI OR "1H-MRS" OR "In Vivo NMR" OR "In-Vivo NMR" OR "MR Spectroscopy" OR "MRSpectroscopy") AND (Brain [mh] OR Neurosciences [mh] OR Neurology [mh] OR Neuroscience OR Neurosciences OR Brain OR Midbrain OR Neurology OR Cognition OR Cortical OR Hippocampus OR "Frontal Lobe" OR "Parietal Lobe" OR "Occipital Lobe" OR "Temporal Lobe" OR Amygdala OR Cortex OR Cerebellum OR Cerebrum OR "Brain Stem") AND (Humans [mh] OR Human OR Humans) AND (Proton OR Protons OR 1H OR "Hydrogen Nucleus" OR "1H-MRS" OR "1H MRS") NOT (Review))

Web of Science:

ALL=(("Magnetic Resonance Spectroscopy" OR MRS OR MRSI OR "1H-MRS" OR "In Vivo NMR" OR "In-Vivo NMR" OR "MR Spectroscopy" OR "mrspectroscopy") AND (Neuroscience OR Neurosciences OR Brain OR Midbrain OR Neurology OR Neuroradiology OR Cognition OR Cortical OR Hippocampus OR "Frontal Lobe" OR "Parietal Lobe" OR "Occipital Lobe" OR "Temporal Lobe" OR Amygdala OR Cortex OR Cerebellum OR Cerebrum OR "Brain Stem") AND (Human OR Humans) AND (Proton OR Protons OR 1H OR "Hydrogen Nucleus" OR "1H-MRS" OR "1H MRS") NOT (Review))

Scopus:

((("magnetic resonance spectroscopy" OR "mrs" OR "1h-mrs" OR "mrsi" OR "NMR Spectroscopy") AND ("human") AND ("1h" OR "proton" OR "1h-mrs" OR "hydrogen nucleus") AND ("neuroscience" OR "brain") AND ("concentration") AND NOT ("rat" OR

"mouse" OR "mice" OR "rodent" OR "Animal" OR "Dog" OR "ex vivo" OR "ex-vivo" OR "vitro" OR "13C" OR "31P" OR "23Na" OR "17O" OR "Systematic Review")))

Relaxation Study Search:

PubMed:

("T2 Relaxation" OR "Transverse Relaxation" OR "Spin-spin Relaxation" OR "Carr-Purcell Meiboom-Gill") AND ("Magnetic Resonance Spectroscopy" [mh] OR "NMR Spectroscopy" OR "Magnetic Resonance Spectroscopy" OR "In-Vivo Spectroscopy" OR "Ex-Vivo Spectroscopy") AND (Brain [mh] OR Neurosciences [mh] OR Neurology [mh] OR "Brain" OR "Neuroscience" OR "Neurology" OR "Phantom") NOT ("Fingerprinting" OR "CEST" OR "Chemical Exchange Saturation Transfer" OR "31P" OR "(31)P" OR "13C" OR "(13)C" OR "23Na" OR "(23)Na" OR "17O" OR "15N" OR "14N" OR "19F" OR "(19)F" OR "Systematic Review" OR "Food Storage" [mh])

Web of Science:

ALL=(("NMR Spectroscopy" OR "Magnetic Resonance Spectroscopy" OR "In-Vivo Spectroscopy" OR "Ex-Vivo Spectroscopy") AND (Neuroscience OR Neurosciences OR Brain OR Neurology OR Phantom) AND ("T2 Relaxation" OR "Transverse Relaxation" OR "Spin-spin Relaxation" OR "Carr-Purcell Meiboom-Gill") NOT ("Fingerprinting" OR "CEST" OR "Chemical Exchange Saturation Transfer" OR "31P" OR "(31)P" OR "13C" OR "(13)C" OR "23Na" OR "(23)Na" OR "17O" OR "15N" OR "14N" OR "19F" OR "(19)F" OR "Systematic Review" OR "Food Storage"))

Scopus:

((("magnetic resonance spectroscopy" OR "mrs" OR "1h-mrs" OR "mrsi" OR "NMR Spectroscopy") AND ("human") AND ("1h" OR "proton" OR "1h-mrs" OR "hydrogen nucleus") AND ("neuroscience" OR "brain") AND ("T2 Relaxation" OR "Transverse Relaxation" OR "Spin-spin Relaxation" OR "Carr-Purcell Meiboom-Gill") AND NOT ("13C" OR "31P" OR "23Na" OR "17O" OR "Systematic Review")))

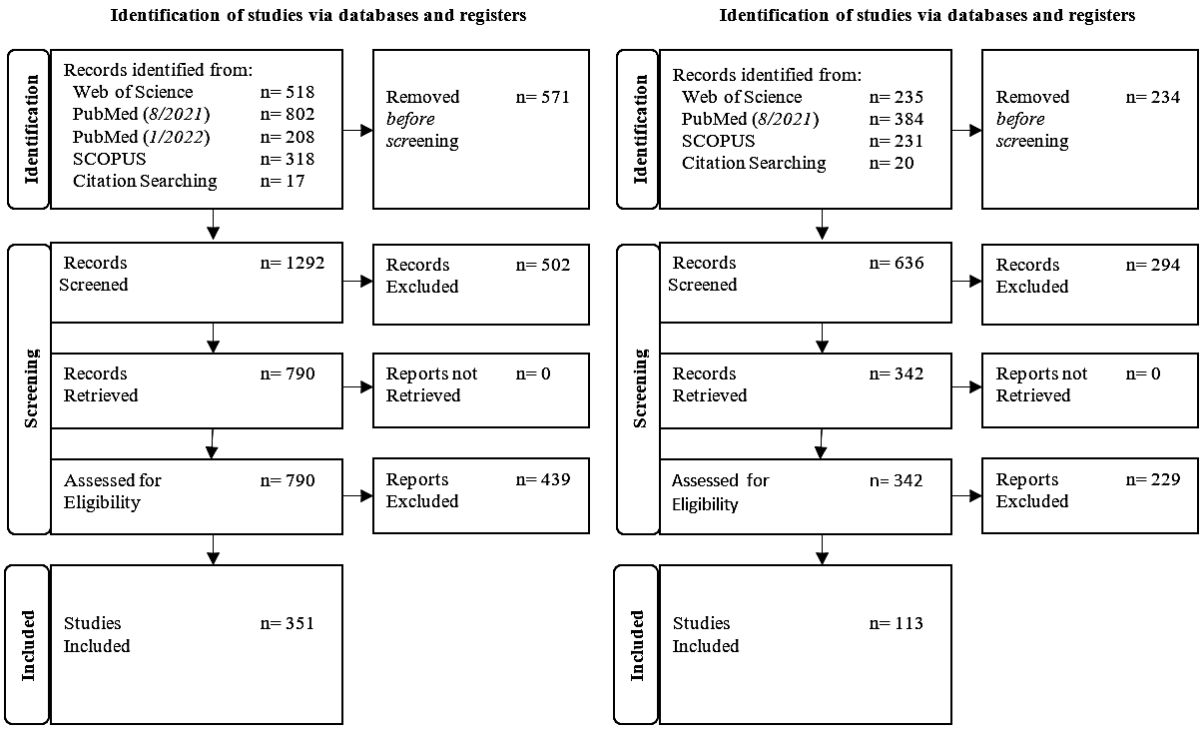


Figure 3.1. PRISMA Flow Charts for the Concentration (left) and T2 Relaxation (right) studies

3.4.1.2. Inclusion/Exclusion Criteria:

For both the Concentration and T2 studies, only *in-vivo* brain ^{1}H-MRS data from primary sources were considered. Reviews, meta-analyses, and book chapters were excluded. Conference posters were typically excluded since they are not peer-reviewed with some exceptions when information was less studied. Only original research was included to avoid repeated inclusions of data. Finally, to be included, a mean and standard deviation needed to be available. For studies that reported statistical results (*t-statistics, p-values, etc.*) without values, we contacted authors by email for inclusion. Median and quartiles were converted to mean and standard deviation using the methods outlined in (Wan et al., 2014) or *(Greco et al., 2015)* to handle normal *(<50% change between (Median - Q1)/(Q3-Median)*) or skewed distributions *(>50% change between (Median - Q1)/(Q3-Median)*). Articles that presented values in the form of bar or scatter plots

were included by manually determining mean and standard deviations with the assistance of an in-house Python software package that maps pixel values to figure axes.

Concentration-Specific:

For the Concentration Study, only human research was considered. Articles were included if they reported at least one metabolite concentration provided as referenced to creatine ($\frac{1}{tCr}$), institutional units (*IU*), or moles/mass, such as millimoles/gram or moles/kilogram (mM). Due to the high volume of articles (*10,506*) returned for the Concentration study, articles were initially limited to 2018-2022. In some cases, articles were retrieved from those identified in earlier years to ensure three or more studies were included in less commonly studied instances, such as certain clinical populations (e.g., Alzheimer's Disease) or difficult to measure metabolites (e.g., Ascorbate). This provided an abbreviated subset of 1,863 articles.

Relaxation-Specific:

Articles were included in the Relaxation study that reported at least one metabolite T2 relaxation time in either a time-based measure (*seconds*) or rate-based measure (*1/seconds*). While this work aims to determine MRS features in the human brain, the relaxation study included all species as a handful of metabolites have not yet been well studied outside of animal models. A total of 870 articles were discovered following the database searches.

***3*.4.1.3. Selection:**

After removing articles that met the exclusion criteria, articles' titles and abstracts were reviewed for relevance. Once confirmed relevant, articles were downloaded to make a final decision on inclusion/exclusion.

3.4.3.1. Concentration Study:

Of the original 1,863 articles, 571 articles were removed prior to screening leaving 1,292 articles. After screening, 790 articles remained and were retrieved and assessed for relevance. A total of 351 articles were determined to be eligible for inclusion in the database and analysis.

3.4.3.2. Relaxation Study:

Of the original 870 articles, 234 were removed prior to screening and 636 articles were further screened. 342 articles were then retrieved and assessed for eligibility. Finally, 113 articles remained and are included in the database and analyses.

3.4.1.4. Curation:

All study information was recorded on a set of two Google spreadsheets where one included all the publication/experiment information with an assigned index number while the other contained all the mean and standard deviation values and corresponding index number. The spreadsheets were later converted using the Python programming language to an SQL database using MySQL and SQLAlchemy (Brown & Wilson, 2012). The database consists of a separate Reference, Experiment, and Values Table for the Concentration and Relaxation studies. All the data is available in either a comma separated values (*.csv*) file or a MySQL folder.

3.4.2. Statistical Analysis:

Data were analyzed using in-house python scripts that utilized numpy, pandas, scipy, statsmodels, matplotlib, and sci-kit learn (C. R. Harris et al., 2020; Hunter, 2007; McKinney, 2010; Pedregosa et al., 2011; Seabold & Perktold, 2010; Virtanen et al., 2020). Rather than an effect size, linear changes were used to directly provide the necessary simulation parameters. Linear changes were calculated using the Ratio Of Means method which has been show to provide an accurate estimate of ratio summary statistics when summary statistics are only provided for groups separately *(*Friedrich et al., 2008, 2011). Heterogeneity was determined

using I^2 (J. Higgins et al., 2002; J. P. T. Higgins & Thompson, 2002). Combined effects were determined using a Random Effects model unless otherwise noted (Borenstein et al., 2009). Meta-regression analyses were performed using a weighted multiple regression model. Both the combined effects and meta-regression used the inverse variance weighting scheme (Hedges & Olkin, 1985). While all data is present in the database, the meta-analyses were carried out when 3 or more studies were available for the particular metabolite, group, or field strength.

3.5. Results:

3.5.1. Database:

The COHERENC database currently contains 461 publications with each entry containing the publication information, experiment details, parameters of the data acquisition, and the mean and standard deviation of resulting values. The database is open-source and available online at github.com/StarkLab/Database. While this study is not meant to be a systematic review, we used the PRISM guidelines to ensure an unbiased and wide-reaching approach was taken to identify and screen publications. A submission form and a correction form are also available for other researchers to submit new data or make corrections.

	Units		Units
Hardware:		**Acquisition:**	
Scanner Manufacturer		Localization Sequence	
Magnetic Field Strength	Tesla	Water Suppression Sequence	
		Acquisition Bandwidth	Hz
Study Populations:		Number of Datapoints	
Index		Number of Averages	
Population		Repetition Time (*TR*)	ms
Control Group		Echo Time (*TE*)	ms
Treatment or Conditions		Inversion Time	ms
Visit or Session Number		T2 Filter	
Total Number of Subjects			
Number of Subjects Analyzed		**Voxel:**	
Sex	m/f	Anatomical Region	
Age (*Mean/Std. Dev.*)	years	Hemisphere Location	left/right/both
		Dimensions (x, y, z)	mm
Analysis:		Volume	cm^3
Preprocessing Software			
Fitting/Quantification Software			
Segmentation Software			
Partial Volume Correction	yes/no		
Relaxation Correction	yes/no		
Tissue Fractions	percent		

Table 3.1. Information available for entries in the database

3.5.2. Healthy Metabolite Concentrations:

Studies that investigated healthy individuals or included control groups were included in the healthy population analysis. Of the 351 studies included, 286 studies investigated a healthy population or included a healthy control group. Subjects were broken down into early life (<2 years of age), adolescent (5-14 years of age), young adult (18-45 years of age) and aged adult (>65 years of age). There were 9, 20, 206, and 45 studies within the four age categories, respectively. The physiological ranges were determined within the each of the four age categories. Values are provided visually in Figure 2 for the young and aged adults for $\frac{1}{tCr}$, IU, and mM.

The heterogeneity (*I2*) within the data ranged between 14%-98%. To combat the high level of inconsistency between studies, we utilized a Random Effects model (Borenstein et al., 2009). Of the 22 metabolites included within the database, 19 metabolites had sufficient data to determine the combined effects. Figure 2 depicts the resulting weighted mean and 95% confidence intervals across mM, IU, and $\frac{1}{tCr}$.

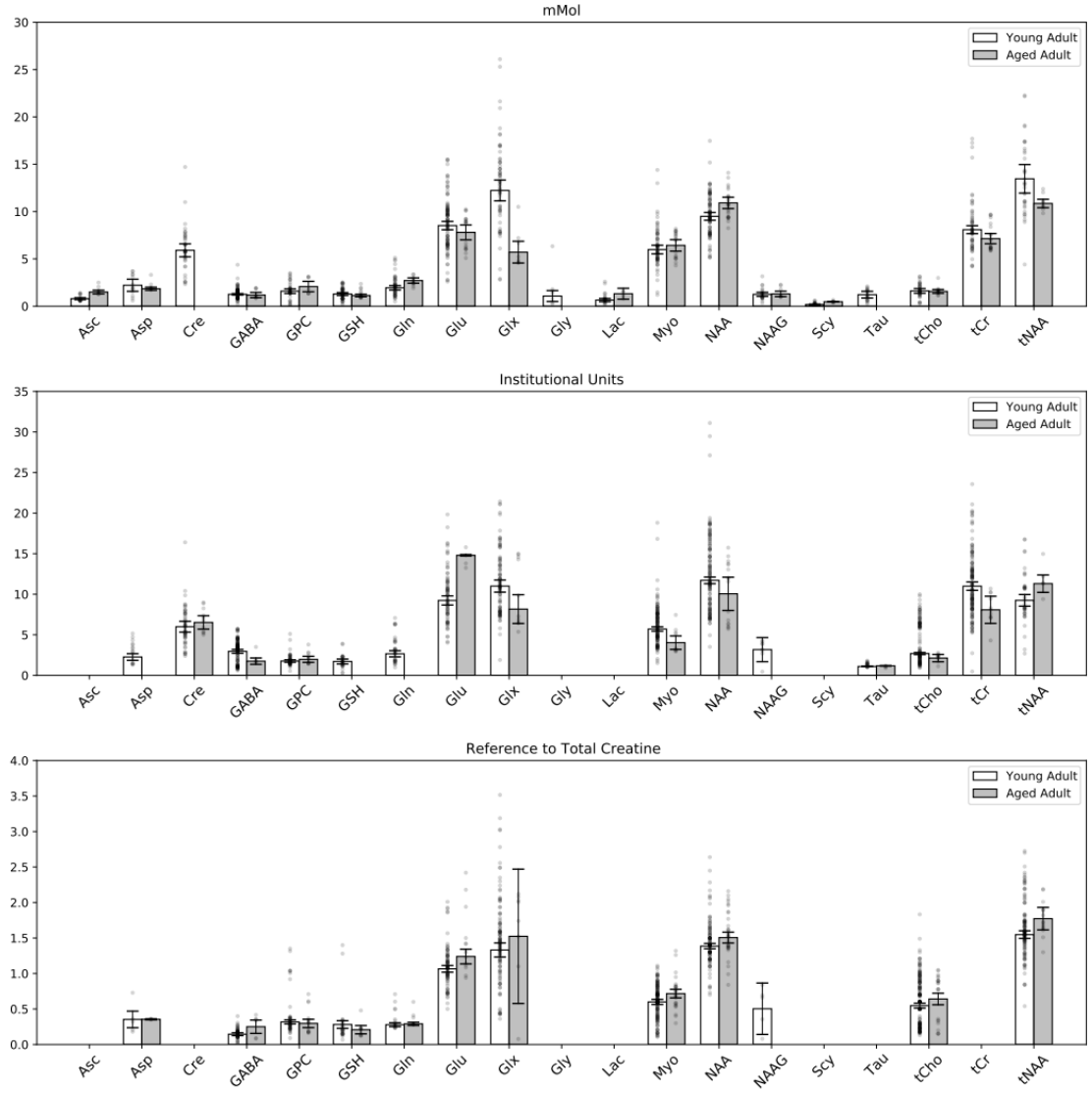


Figure 3.2.: Brain Metabolite Concentrations in Young and Aged Healthy Adults from studies that reported results as A) relative to total Creatine ($\frac{1}{tCr}$), B) Institutional Units (*IU*), C) molas/mass (*mM*).

3.5.3. Clinical Metabolite Concentrations:

Studies that investigated clinical groups and included a control group were included in the clinical population analysis. While we included clinical studies that did not include a control group in the database, we chose to focus on studies that had direct comparisons to avoid confounds involving the variation among studies. There were 195 publications consisting of 25 unique clinical groups. Each clinical population was modeled as a linear change relative to their respective control group. As such, a value of 1.0 would indicate no difference between the clinical and control groups. Figure 3 depicts the resulting weighted mean and 95% confidence intervals across mM, IU, and $\frac{1}{tCr}$.

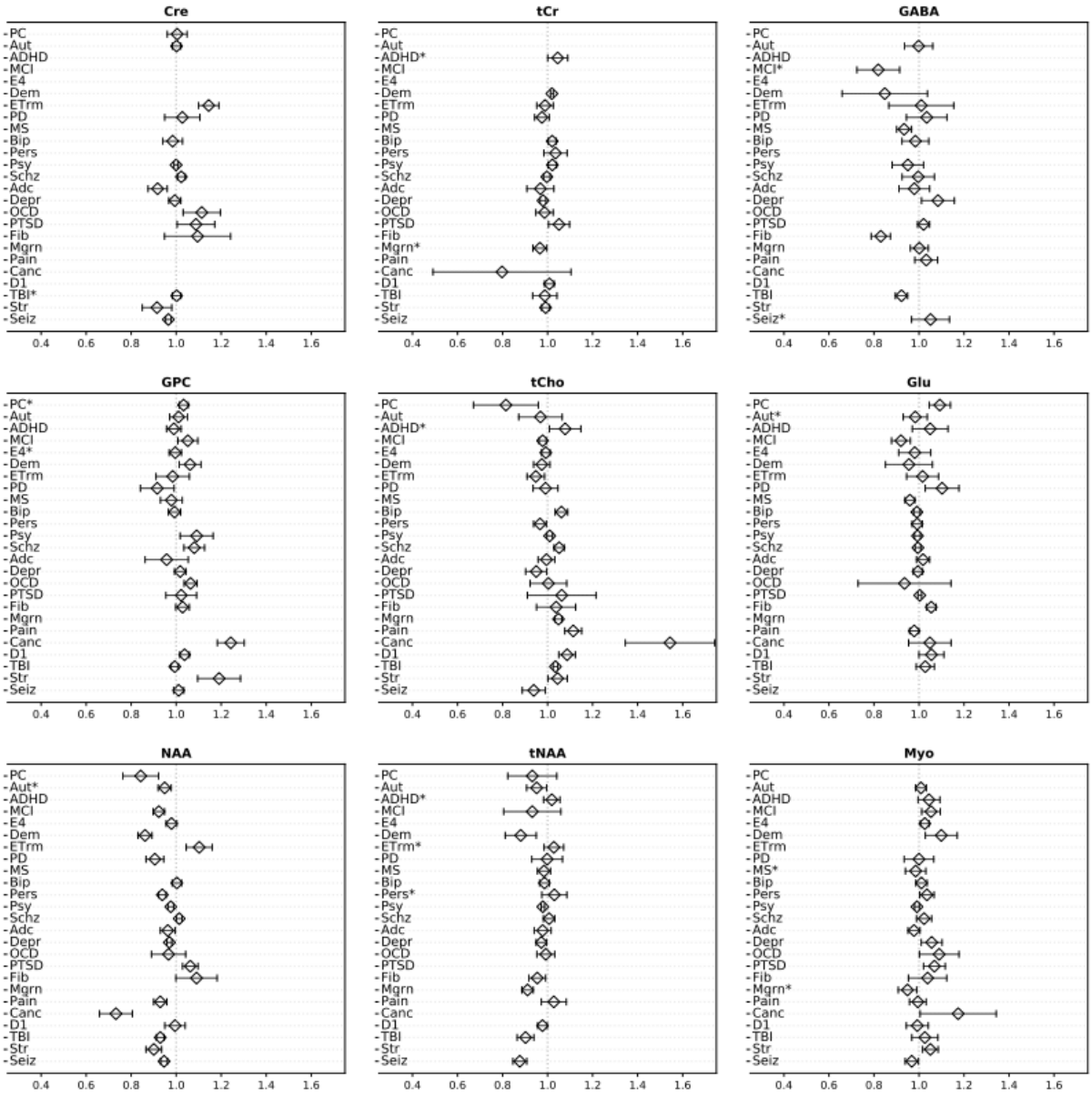


Figure 3.3.: Commonly investigated metabolite concentrations modeled in diseased populations. PC = Perinatal Complications, Aut = Autism, ADHD = Attention Deficit Hyper Activity, MCI = Mild Cognitive Impairment, E4 Apolipoprotein 4 Carriers, Dem = Dementia, ETrm = Essential Tremor, PD = Parkinson's Disease, MS = Multiple Sclerosis, Bip = Bipolar, Pers = Personality Disorder, Psy = Psychosis, Schz = Schizophrenia, Adc = Addiction, Depr = Depression, OCD = Obsessive Compulsive Disorder, PTSD = Post-Traumatic Stress Disorder, Fib = Fibromyalgia, Mgrn = Migraine, Pain = Chronic Pain, Canc = Cancer, D1 = Diabetes Type 1, TBI = Traumatic Brain Injury, Str = Stroke, Seiz = Seizure Disorder. * Denotes the use of Fixed Effect Model.

3.5.4. T2 relaxation

Studies that investigated healthy subjects or included a healthy control group were included in the T2 relaxation analysis. Of the 113 included studies, 76 studies were included in the analysis. Among these studies heterogeneity was high (*over 80%*) for all metabolites. In contrast to metabolite concentrations, the high degree of variability is explained by the known differences in T2 when measured across different acquisition schemes. To effectively model this fact, we employed a multiple meta-regression to capture this variability. Within the model, we included 6 variables: 1) metabolite, 2) field strength, 3) localization pulse sequence, 4) T2 filter, 5) tissue type, and 6) subject species. The model was rerun leaving one datapoint out each time for prediction (i.e., 628 individual leave-one-out regression models were run). The model achieved a median adjusted R^2 of .782 (*Q1 = 78.17; Q3 = .7819*). Predictions for each of these models yielded a median error of 26.76 ms (*Q1 = 12.20ms; Q3 = 54.65ms*). Figure 4 contains histograms of the adjusted R^2 and prediction error across all models.

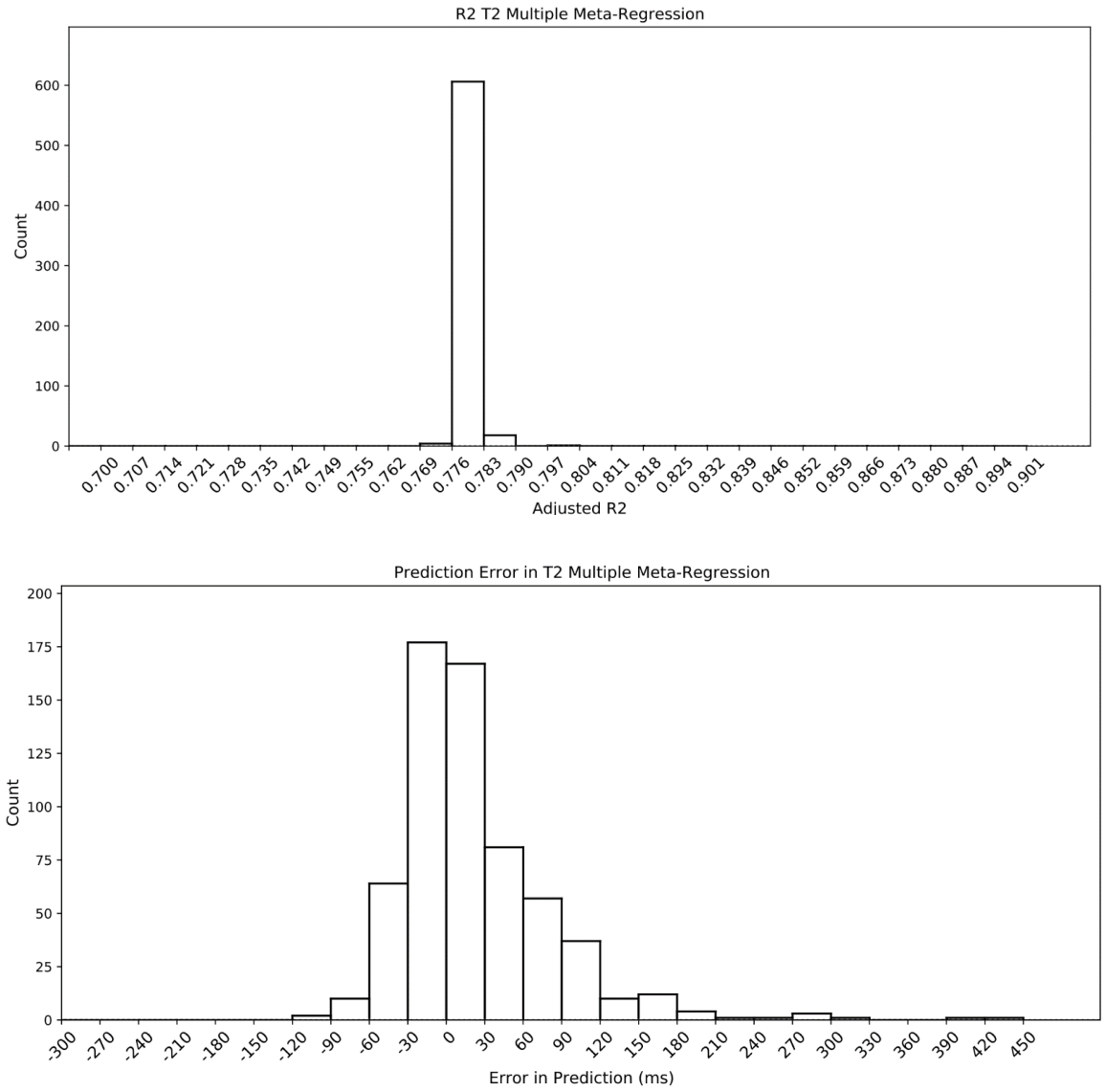


Figure 3.4.: The multiple meta-regression model was run iteratively with each iteration leaving out 1 datapoint. A) Adjusted R^2 for all resulting models. B) Error in the model predictions vs. the actual reported T2 value.

3.6. Discussion:

3.6.1. Physiological Ranges of Brain Metabolites in the Healthy Adults:

One of our core goals was to approximate the true levels of MRS accessible metabolites via a large data mining and unification approach. Ours was certainly not the first effort to provide typical concentration values or ranges. Physiological ranges of brain metabolites have been proposed previously for the brain using experimentally determined ranges (de Graaf, 2019; Govindaraju et al., 2000). Values were also proposed for the Fit Challenge that were said to lie within normal ranges (Marjańska et al., 2022). Here, however, we took a comprehensive approach through the literature that unified measures across hundreds of studies and appropriately weighed studies to determine the physiological ranges of 22 brain metabolites. Additionally, the focus here was on recent publications (<5 years old) where data was quantified using the most current and advanced methodologies. Reassuringly, many values here reflect similar ranges as previously proposed supporting prior estimations (Figure 5).

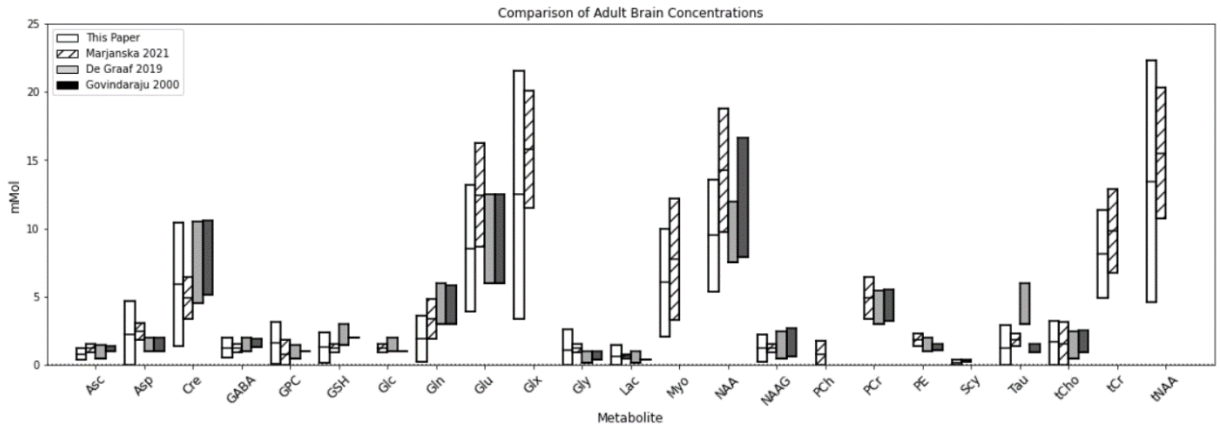


Figure 3.5. Comparison of brain metabolite concentrations in published works.

In terms of quantification, metabolite concentrations are most commonly provided in 3 different units, ratio to total creatine ($\frac{1}{tCr}$), institutional units (*IU*), and/or millimolar/kilogram (*mM*). Other units are reported within the literature, but were not considered due to how rare they are and to avoid inconsistency among comparisons. $\frac{1}{tCr}$ has been one of the most dominant methods for quantifying metabolite levels as it requires little expert intervention and can be performed with even the most basic software packages, including the vendor-provided software installed on scanner consoles. Though, due to its simplicity, $\frac{1}{tCr}$ fails to account for tissue compositions within acquisition voxels, changes in tCr concentrations in different populations, and does not account for metabolite-specific relaxation properties. Despite these limitations, the ease and speed at which results can be determined is powerful. Metabolite levels reported in IU aim to take some of these factors into account to provide a more accurate measurement. These units typically use tissue water as a reference, rather than the total Creatine signal. Tissue water is generally accepted to be more stable than total creatine concentrations and the concentration ranges are maintained within strict physiological ranges. These units may be preferred when some corrections are necessary, but conversion to absolute concentrations presents a technical hurdle. Caution should be used when comparing metabolite levels in IU outside of the study as these values still represent arbitrary units. Finally, mM is has now become the most preferred method for reporting metabolite levels as it reflects an absolute concentration. This is possible as the computation requires several correction factors that are applied to the raw data, including accounting for differences in tissue composition, metabolite-specific relaxation rates, and reference to a known concentration. Reporting values in mM has also become increasingly available in recent years as many commonly used software packages will make all the computations with little, to no, user intervention.

The metabolic profile provided here represents progress towards effectively simulating realistic data. As previously stated, developing methodology surrounding data analysis is limited by the lack of ground-truths. While still objective, current methods to evaluate methodological performance are restricted to outcomes of lower Cramer–Rao Lower Bounds (CRLB) or reductions in standard deviation, rather than error when compared against the true value. Ultimately, having appropriate ground truths to evaluate methods is a major step forward and this work gets us closer to a real approximation of this truth. Additionally, having these ground-truths will continue to motivate work utilizing deep learning and machine learning algorithms as these methods have shown exceptional potential in so many areas. While many recent papers have demonstrated proof of concept around using these approaches with MRS, there has yet to be a fully developed neural network software package available for the field to adopt.

3.6.2. Physiological Ranges of Brain Metabolites in Clinical Populations:

Here, we also presented a linear model demonstrating the relationship between healthy and clinical populations. As far as we know, this is the first study to provide a basis to determine physiological ranges of brain metabolites in such a wide array of clinical populations. Rather than using effect size, we derived a linear model to demonstrate the expected ranges of metabolites within a given population. In combination with the results from the healthy population, one simply has to multiply the metabolite coefficients to produce a metabolic profile for each population.

The most commonly measures metabolites are shown in Figure 3. Many of the results are similar to what has been shown in systematic reviews and more traditional meta-analyses. For instance, investigations into cancer populations have yielded elevated choline signals (Daqqaq, 2019; Wang et al., 2016). For populations with dementias, such as Alzheimer's Disease,

increased concentrations of Myo and decreased concentrations of NAA support previous findings (Kantarci et al., 2000; Zhang et al., 2014). Differing neurometabolic profiles are also present when comparing Parkinson's Disease and Essential Tremor, which otherwise present with similar symptomology (Buard et al., 2022; Handforth & Lang, 2021; Kendi et al., 2005). The results presented here may further provide a priori when planning research analyses.

Ultimately, for MRS to become more relevant in the clinical world, tools must be developed around clinical populations. Though, the diversity and uncertainty of concentrations across populations has previously posed a great challenge. Using the database and values provided here, we encourage using data simulations that encompass diseased populations which will allow for flexible models that can generalize well to healthy and diseased populations.

3.6.3. Multiple Meta-Regression to Explain Heterogeneity of Metabolite T^2 Relaxation Results:

In addition to metabolite concentrations, for simulations and some data analysis approaches depend on accurate estimates of the T^2 relaxation rates for metabolites. After the excitation pulse in a magnetic resonance experiment, relaxation describes the process by which energy is released from a given spin to the surrounding lattice (T^1) and dephasing of quantum phase coherence (T^2) resulting in the observed loss of signal. Here, we focused upon T^2, which arguably plays a greater role on the timescale of MRS experiments.

The loss of phase is further described as T^2* which is the combination of both a spin's intrinsic relaxation components (T^2) and its extrinsic relaxation components (T^2'). Bloembergen-Purcell-Pound (BPP) theory predicts that the spin-spin relaxation would remain relatively constant across experiments as well as different field strengths. However, the measured T^2* is known to change across acquisition schemes and decrease as field strength increases. To a first

order approximation these results are explained from the movement of molecules through different local magnetic fields. For this reason, the ideal method for acquiring T^2 data involves the use of T^2 filters, such as Carr-Purcell Meiboom Gill (*CPMG*), which partially reverses the effect of movement. These T^2 filters can be added into regular T^2 acquisition schemes, which involve modeling the decay rate of signal intensity across a series of echo times.

Unfortunately, collecting T^2 data in human subjects remains a challenging task. T^2 filters do not typically come standard from the manufacturer, and one must consider the time-constraints for collecting multiple echo times during a scanning session. Measuring coupled spins poses further difficulty as the series of echo times must be carefully chosen with respect to the evolution timing and patterns. Finally, overlapping spins may have considerably different rates of decay which necessitates more data acquisition to properly model multiple decay rates or may even require the use of more advanced editing techniques.

Ultimately, the variance that results from the wide array of acquisition schemes has resulted in uncertainty of how to effectively build T^2 relaxation into simulations. Furthermore, a lack of experimental data in humans, especially for coupled spin systems, creates a major limitation for extrapolating many metabolites. Even when combining data across species, the available results are sparse with very few metabolites having data across all field strengths or under specific acquisition schemes. Though, as previously stated, this aspect of *in-vivo* data may prove to be a crucial component for building effective machine learning algorithms and deep learning models that can generalize to true *in-vivo* data.

Here, we present a model that could account for a large degree of the variance in results. The model included 6 variables: 1) metabolite, 2) field strength, 3) localization pulse sequence, 4) T2 filter, 5) tissue type, and 6) subject species. Following a leave-one-out validation approach,

nearly 80% of the variance could be attributed to the 6 factors. As such, the error in prediction was low with approximately 25% of the prediction errors less than 10ms, 50% of prediction errors less than 25ms, and nearly 75% of prediction errors under 50ms. Upon further investigation, the instances where prediction error was high were primarily from a small subset of papers and appear to represent outliers in the data. Future studies should determine whether the values from these studies can be replicated or more effectively modeled. It will also be important to measure T^2 data in clinical populations. Ultimately, this model provides a robust solution towards including T^2 relaxation within simulations and unifying the disparate data currently available.

3.6.4. Further Research:

Future research should aim to further model the effects of different data acquisition and analysis methods. One potential area would be an investigation into how different preprocessing, fitting, segmentation, and quantification software packages impacts concentration results. While we included tissue type, one could determine the influence of specific brain regions using the information available within database. While few articles included details about acquisition parameters, such as bandwidth, number of averages, and vector-size, these factors may provide further explanation.

3.6.5. Limitations:

Unfortunately, due to typos or errors, we expect that there may be discrepancies in reported values and/or the information surrounding the methods. We anticipate the effects of these errors in reporting will likely only contribute a small effect as each individual datapoint only makes a small contribution. Still, we encourage others to submit corrections as they are discovered.

Another limitation is likely present in that data may be correlated within studies. Within the combined effect models, each individual datapoint from studies was included, rather than attempting to collapse across all results for a given study. We opted to take this approach to capture the expected differences within studies that reported results from different acquisition schemes, multiple brain regions, or different tissue (e.g. grey matter vs. white matter) types.

3.7. Conclusions:

Here, we provide a new database organizing brain metabolite results from a little under 500 publications. This database is freely available online where users can view and contribute data. Using the database, we've determined physiological ranges of 22 brain metabolites in the healthy infant, adolescent, adult, and aged brain. We've further provided linear model that describes changes in metabolite concentrations for 25 different clinical populations. Finally, using a multiple meta-regression we've developed a model to predict metabolite relaxation rates across varied acquisition schemes.

3.8. References:

Borenstein, M., Hedges, L. V., Higgins, J. P. T., & Rothstein, H. R. (2009). *Introduction to Meta-Analysis*. John Wiley & Sons, Ltd. https://doi.org/10.1002/9780470743386

Brown, A., & Wilson, G. (Eds.). (2012). SQLAlchemy. In *The Architecture of Open Source Applications Volume II: Structure, Scale, and a Few More Fearless Hacks*. aosabook.org. http://aosabook.org/en/sqlalchemy.html

Buard, I., Lopez-Esquibel, N., Carey, F. J., Brown, M. S., Medina, L. D., Kronberg, E., Martin, C. S., Rogers, S., Holden, S. K., Greher, M. R., & Kluger, B. M. (2022). Does Prefrontal Glutamate Index Cognitive Changes in Parkinson's Disease? *Frontiers in Human*

Neuroscience, *16*(April), 1–9. https://doi.org/10.3389/fnhum.2022.809905

Choi, I. Y., Andronesi, O. C., Barker, P., Bogner, W., Edden, R. A. E., Kaiser, L. G., Lee, P., Marjańska, M., Terpstra, M., & de Graaf, R. A. (2021). Spectral editing in 1H magnetic resonance spectroscopy: Experts' consensus recommendations. *NMR in Biomedicine*, *34*(5), 1–18. https://doi.org/10.1002/nbm.4411

Cudalbu, C., Behar, K. L., Bhattacharyya, P. K., Bogner, W., Borbath, T., de Graaf, R. A., Gruetter, R., Henning, A., Juchem, C., Kreis, R., Lee, P., Lei, H., Marjańska, M., Mekle, R., Murali-Manohar, S., Považan, M., Rackayová, V., Simicic, D., Slotboom, J., … Mlynárik, V. (2021). Contribution of macromolecules to brain 1H MR spectra: Experts' consensus recommendations. *NMR in Biomedicine*, *34*(5), 1–24. https://doi.org/10.1002/nbm.4393

Daqqaq, T. S. (2019). Role of radiological intervention in brain tumor: A meta-analysis. *International Surgery*, *104*(7), 383–389. https://doi.org/10.9738/INTSURG-D-20-00014.1

de Graaf, R. A. (2019). In Vivo NMR Spectroscopy. In *Basic life sciences* (Vol. 55). John Wiley & Sons, Ltd. https://doi.org/10.1002/9781119382461

Edden, R. A. E., Puts, N. A. J., Harris, A. D., Barker, P. B., & Evans, C. J. (2014). Gannet: A batch-processing tool for the quantitative analysis of gamma-aminobutyric acid-edited MR spectroscopy spectra. *Journal of Magnetic Resonance Imaging*, *40*(6), 1445–1452. https://doi.org/10.1002/jmri.24478

Friedrich, J. O., Adhikari, N. K. J., & Beyene, J. (2008). The ratio of means method as an alternative to mean differences for analyzing continuous outcome variables in meta-analysis: A simulation study. *BMC Medical Research Methodology*, *8*. https://doi.org/10.1186/1471-2288-8-32

Friedrich, J. O., Adhikari, N. K. J., & Beyene, J. (2011). Ratio of means for analyzing continuous outcomes in meta-analysis performed as well as mean difference methods. *Journal of Clinical Epidemiology*, *64*(5), 556–564. https://doi.org/10.1016/j.jclinepi.2010.09.016

Govindaraju, V., Young, K., & Maudsley, A. A. (2000). Proton NMR chemical shifts and coupling constants for brain metabolites. *NMR in Biomedicine*, *13*(3), 129–153. https://doi.org/10.1002/1099-1492(200005)13:3<129::AID-NBM619>3.0.CO;2-V

Greco, T., Biondi-Zoccai, G., Gemma, M., Guérin, C., Zangrillo, A., & Landoni, G. (2015). How to impute study-specific standard deviations in meta-analyses of skewed continuous endpoints? *World Journal of Meta-Analysis*, *3*(5), 215. https://doi.org/10.13105/wjma.v3.i5.215

Handforth, A., & Lang, E. J. (2021). Increased Purkinje Cell Complex Spike and Deep Cerebellar Nucleus Synchrony as a Potential Basis for Syndromic Essential Tremor. A Review and Synthesis of the Literature. *Cerebellum*, *20*(2), 266–281. https://doi.org/10.1007/s12311-020-01197-5

Harris, A. D., Saleh, M. G., & Edden, R. A. E. (2017). Edited 1H magnetic resonance spectroscopy in vivo: Methods and metabolites. *Magnetic Resonance in Medicine*, *77*(4), 1377–1389. https://doi.org/10.1002/mrm.26619

Harris, C. R., Millman, K. J., van der Walt, S. J., Gommers, R., Virtanen, P., Cournapeau, D., Wieser, E., Taylor, J., Berg, S., Smith, N. J., Kern, R., Picus, M., Hoyer, S., van Kerkwijk, M. H., Brett, M., Haldane, A., del Río, J. F., Wiebe, M., Peterson, P., … Oliphant, T. E. (2020). Array programming with NumPy. *Nature*, *585*(7825), 357–362. https://doi.org/10.1038/s41586-020-2649-2

Hedges, L. V., & Olkin, I. (1985). *Statistical methods for meta-analysis*. Academic Press.

Henning, A. (2018). Proton and multinuclear magnetic resonance spectroscopy in the human brain at ultra-high field strength: A review. *NeuroImage*, *168*, 181–198. https://doi.org/10.1016/j.neuroimage.2017.07.017

Higgins, J. P. T., & Thompson, S. G. (2002). Quantifying heterogeneity in a meta-analysis. *Statistics in Medicine*, *21*(11), 1539–1558. https://doi.org/10.1002/sim.1186

Higgins, J., Thompson, S., Deeks, J., & Altman, D. (2002). Statistical heterogeneity in systematic reviews of clinical trials: A critical appraisal of guidelines and practice. *Journal of Health Services Research and Policy*, *7*(1), 51–61. https://doi.org/10.1258/1355819021927674

Hogben, H. J., Krzystyniak, M., Charnock, G. T. P., Hore, P. J., & Kuprov, I. (2011). Spinach - A software library for simulation of spin dynamics in large spin systems. *Journal of Magnetic Resonance*, *208*(2), 179–194. https://doi.org/10.1016/j.jmr.2010.11.008

Hunter, J. D. (2007). Matplotlib: A 2D Graphics Environment. *Computing in Science & Engineering*, *9*(3), 90–95. https://doi.org/10.1109/MCSE.2007.55

Kantarci, K., Jack, C. R., Xu, Y. C., Campeau, N. G., O'Brien, P. C., Smith, G. E., Ivnik, R. J., Boeve, B. F., Kokmen, E., Tangalos, E. G., & Petersen, R. C. (2000). Regional metabolic patterns in mild cognitive impairment and Alzheimer's disease - A H-1 MRS study. *NEUROLOGY*, *55*(2), 210–217. https://doi.org/10.1212/WNL.55.2.210

Kendi, A. T. K., Tan, F. U., Kendi, M., Erdal, H. H., & Tellioğlu, S. (2005). Magnetic resonance spectroscopy of the thalamus in essential tremor patients. *Journal of Neuroimaging*, *15*(4),

362–366. https://doi.org/10.1177/1051228405279039

Kreis, R., Boer, V., Choi, I. Y., Cudalbu, C., de Graaf, R. A., Gasparovic, C., Heerschap, A., Krššák, M., Lanz, B., Maudsley, A. A., Meyerspeer, M., Near, J., Öz, G., Posse, S., Slotboom, J., Terpstra, M., Tkáč, I., Wilson, M., & Bogner, W. (2020). Terminology and concepts for the characterization of in vivo MR spectroscopy methods and MR spectra: Background and experts' consensus recommendations. In *NMR in Biomedicine* (Issue May 2020). https://doi.org/10.1002/nbm.4347

Lin, A., Andronesi, O., Bogner, W., Choi, I. Y., Coello, E., Cudalbu, C., Juchem, C., Kemp, G. J., Kreis, R., Krššák, M., Lee, P., Maudsley, A. A., Meyerspeer, M., Mlynarik, V., Near, J., Öz, G., Peek, A. L., Puts, N. A., Ratai, E. M., … Mullins, P. G. (2021). Minimum Reporting Standards for in vivo Magnetic Resonance Spectroscopy (MRSinMRS): Experts' consensus recommendations. *NMR in Biomedicine*, *34*(5), e4484. https://doi.org/10.1002/nbm.4484

Marjańska, M., Deelchand, D. K., Kreis, R., Alger, J. R., Bolan, P. J., Borbath, T., Boumezbeur, F., Fernandes, C. C., Coello, E., Nagraja, B. H., Považan, M., Ratiney, H., Sima, D., Starčuková, J., Soher, B. J., Wilson, M., & van Asten, J. J. A. (2022). Results and interpretation of a fitting challenge for MR spectroscopy set up by the MRS study group of ISMRM. *Magnetic Resonance in Medicine*, *87*(1), 11–32. https://doi.org/10.1002/mrm.28942

McKinney, W. (2010). Data Structures for Statistical Computing in Python. *Proceedings of the 9th Python in Science Conference*, *1*(Scipy), 56–61. https://doi.org/10.25080/majora-92bf1922-00a

Moher, D., Liberati, A., Tetzlaff, J., & Altman, D. G. (2009). Preferred reporting items for

systematic reviews and meta-analyses: The PRISMA statement. *BMJ (Online)*, *339*(7716), 332–336. https://doi.org/10.1136/bmj.b2535

Near, J., Harris, A. D., Juchem, C., Kreis, R., Marjańska, M., Öz, G., Slotboom, J., Wilson, M., & Gasparovic, C. (2021). Preprocessing, analysis and quantification in single-voxel magnetic resonance spectroscopy: experts' consensus recommendations. *NMR in Biomedicine*, *34*(5), 1–23. https://doi.org/10.1002/nbm.4257

Oeltzschner, G., Saleh, M. G., Rimbault, D., Mikkelsen, M., Chan, K. L., Puts, N. A. J., & Edden, R. A. E. (2019). Advanced Hadamard-encoded editing of seven low-concentration brain metabolites: Principles of HERCULES. *NeuroImage*, *185*(September 2018), 181–190. https://doi.org/10.1016/j.neuroimage.2018.10.002

Öz, G., Deelchand, D. K., Wijnen, J. P., Mlynárik, V., Xin, L., Mekle, R., Noeske, R., Scheenen, T. W. J., Tkáč, I., Andronesi, O., Barker, P. B., Bartha, R., Berrington, A., Boer, V., Cudalbu, C., Emir, U. E., Ernst, T., Fillmer, A., Heerschap, A., … Wilson, M. (2021). Advanced single voxel 1H magnetic resonance spectroscopy techniques in humans: Experts' consensus recommendations. *NMR in Biomedicine*, *34*(5), 1–18. https://doi.org/10.1002/nbm.4236

Page, M. J., Moher, D., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., Shamseer, L., Tetzlaff, J. M., Akl, E. A., Brennan, S. E., Chou, R., Glanville, J., Grimshaw, J. M., Hróbjartsson, A., Lalu, M. M., Li, T., Loder, E. W., Mayo-Wilson, E., Mcdonald, S., … Mckenzie, J. E. (2021). PRISMA 2020 explanation and elaboration: Updated guidance and exemplars for reporting systematic reviews. *The BMJ*, *372*. https://doi.org/10.1136/bmj.n160

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., & Duchesnay, E. (2011). Scikit-learn: Machine Learning in {P}ython. *Journal of Machine Learning Research*, *12*(12), 2825–2830.

Ross, B., & Bluml, S. (2001). Magnetic resonance spectroscopy of the human brain. *The Anatomical Record*, *265*(2), 54–84. https://doi.org/10.1002/ar.1058

Saleh, M. G., Oeltzschner, G., Chan, K. L., Puts, N. A. J., Mikkelsen, M., Schär, M., Harris, A. D., & Edden, R. A. E. (2016). Simultaneous edited MRS of GABA and glutathione. *NeuroImage*, *142*, 576–582. https://doi.org/10.1016/j.neuroimage.2016.07.056

Seabold, S., & Perktold, J. (2010). Statsmodels: Econometric and Statistical Modeling with Python. *Proceedings of the 9th Python in Science Conference, Scipy*, 92–96. https://doi.org/10.25080/majora-92bf1922-011

Simpson, R., Devenyi, G. A., Jezzard, P., Hennessy, T. J., & Near, J. (2017). Advanced processing and simulation of <scp>MRS</scp> data using the <scp>FID</scp> appliance (<scp>FID-A</scp>)—An open source, <scp>MATLAB</scp> -based toolkit. *Magnetic Resonance in Medicine*, *77*(1), 23–33. https://doi.org/10.1002/mrm.26091

Smith, S. A., Levante, T. O., Meier, B. H., & Ernst, R. R. (1994). Computer Simulations in Magnetic Resonance. An Object-Oriented Programming Approach. In *Journal of Magnetic Resonance, Series A* (Vol. 106, Issue 1, pp. 75–105). https://doi.org/10.1006/jmra.1994.1008

Starčuk, Z., & Starčuková, J. (2017). Quantum-mechanical simulations for in vivo MR spectroscopy: Principles and possibilities demonstrated with the program NMRScopeB.

Analytical Biochemistry, *529*, 79–97. https://doi.org/10.1016/j.ab.2016.10.007

Virtanen, P., Gommers, R., Oliphant, T. E., Haberland, M., Reddy, T., Cournapeau, D., Burovski, E., Peterson, P., Weckesser, W., Bright, J., van der Walt, S. J., Brett, M., Wilson, J., Millman, K. J., Mayorov, N., Nelson, A. R. J., Jones, E., Kern, R., Larson, E., … Vázquez-Baeza, Y. (2020). SciPy 1.0: fundamental algorithms for scientific computing in Python. *Nature Methods*, *17*(3), 261–272. https://doi.org/10.1038/s41592-019-0686-2

Wan, X., Wang, W., Liu, J., & Tong, T. (2014). Estimating the sample mean and standard deviation from the sample size, median, range and/or interquartile range. *BMC Medical Research Methodology*, *14*(1), 1–13. https://doi.org/10.1186/1471-2288-14-135

Wang, Q., Zhang, H., Zhang, J. S., Wu, C., Zhu, W. J., Li, F. Y., Chen, X. L., & Xu, B. N. (2016). The diagnostic performance of magnetic resonance spectroscopy in differentiating high-from low-grade gliomas: A systematic review and meta-analysis. *European Radiology*, *26*(8), 2670–2684. https://doi.org/10.1007/s00330-015-4046-z

Zhang, N., Song, X., Bartha, R., Beyea, S., D'Arcy, R., Zhang, Y., & Rockwood, K. (2014). Advances in High-Field Magnetic Resonance Spectroscopy in Alzheimer's Disease. *Current Alzheimer Research*, *11*(4), 367–388. https://doi.org/10.2174/1567205011666140302200312

4. Deep learning for MRS Data and the AGNOSTIC Benchmark Dataset (*Paper In Prep*):

4.1. Title:

Deep Learning for MRS Data and the AGNOSTIC Benchmark Dataset.

4.2. Abstract:

Proton (^{1}H) Magnetic Resonance Spectroscopy (*MRS*) is a powerful method capable of studying brain metabolites *in-vivo*. As a clinical research tool, accurate and reliable quantification is a necessity, which has driven the MRS community to prioritize the reliability and standardization quantification software packages. Though, low signal to noise, artifacts, and overlapping signals still provide significant challenges, especially for deep brain regions and multi-voxel acquisition. Data quality can be improved through collecting multiple transients or employing multi-scan (*editing, nulling, etc.)* methods, but time constraints may limit what can be done during a scan session. Neural networks have the potential to aid in many of these fundamental challenges associated with MRS data. While recent neural network models have already demonstrated some proof of concept, a software package that utilizes deep learning has yet to be shown to either complement traditional methods or perform as well as standard techniques in practice. Furthermore, little is known about how changes to network architecture, hyperparameters, and training sets influence the overall success. Unlike fields like computer vision, a benchmark dataset (*i.e., ImageNet*) does not exist for MRS which makes developing these models or evaluating performance across models impossible. As such, we sought to develop the first training dataset for MRS. Here, we provide the AGNOSTIC training dataset for MRS. Included in the dataset are 500,000 (healthy and clinical) examples. Using this dataset, we demonstrate the performance and limitations of various neural networks models that will need to be addressed prior to their widespread use.

4.3. Introduction:

4.3.1. Background:

Proton (^{1}H) Magnetic Resonance Spectroscopy (*MRS*) is a noninvasive tool that can study brain metabolites *in-vivo*. Though, low signal to noise (*SNR*), artifacts, and overlapping signals can greatly impede the quality of data and reliability of results (Cudalbu et al., 2021; Near et al., 2021; Öz et al., 2021). Furthermore, these challenges can be greatly amplified when acquiring data from deep brain structures or during MRS Imaging (MRSI) which are often needed for clinical applications (Edelstein et al., 1986; Maudsley et al., 2010, 2021). Deep neural networks have demonstrated success with similar issues in signal processing and computer vision and thus offer great potential for MRS community.

Deep learning is considered a subcategory of machine learning that uses a series of computational layers to process information. These layers make up an artificial neural network, as they model the inner neuronal system within the brain. These neural networks can contain hundreds of layers where each layer learns to extract specific features of data. Some popular layer types include, “fully connected” where each node of a layer is connected with every node in the adjacent layers, “convolutional” where a kernel function sweeps across an input, and “recurrent” where the layer maintains an internal state that is updated based on past and future inputs. Generally, convolutional and recurrent layers excel in spatial and temporal-based data, respectively, while fully connected layers perform well in less-structured data. These networks are considered ‘shallow’ with only a few layers or ‘deep’ containing many layers, potentially with millions of trainable. As such, these networks generally require hundreds of thousands of training examples. During this training period, each data input is passed through a network that has all its units initialized with random values. The output from the final layer is then compared

against a ground-truth or target which provides an error metric. Finally, a gradient is computed which tunes each of the units in such a way that will decrease the error of the outputs. While training can take significant time to complete, these neural networks are capable of learning highly complex features allowing them to greatly outperform even the most experienced individuals.

Recently, several methods have provided a proof of concept through utilizing fully-connected networks (FCN), Convolutional Neural Networks (CNN), and Recurrent Neural Networks (RNN) (Chandler et al., 2019; Kyathanahally et al., 2018; Lee & Kim, 2019, 2020; Ma et al., 2022; Tapper et al., 2021). Still, despite the potential, these methods have yet to be used in practice or shown to generalize outside of small datasets that use a single fixed acquisition protocol. Before the field moves towards adopting these methods to inform research and clinical decisions, there must be appropriate tools to evaluate performance. Perhaps the first barrier in this goal is the lack of a benchmark dataset akin to what MNIST and ImageNet have provided for the field of Computer Vision (Fei-Fei et al., 2010; Li Deng, 2012). Without having such a dataset, getting started with neural network development can be daunting and further comparing performance across models is not possible. A second barrier is the inherent black box nature of deep learning, which is especially problematic for *in-vivo* MRS data which has no ground truth.

4.3.2. Purpose:

Here, we focus upon two objectives: 1) development of a general unbiased and freely available benchmark dataset and 2) evaluating the performance and capabilities of different neural network architectures. For Part-1, we propose the use of the AGNOSTIC (**A G**eneralized **N**eural-network **O**pen-**S**ource **T**raining dataset **I**nsensitive to Acquisition **C**onfines). In Part-2, we use this dataset to test the performance of a variety of FCN, CNN, and RNN on residual

water removal, macromolecules removal, frequency and phase correction, and denoising. Finally, we demonstrate how these neural network methods have the ability to effectively separate MM from metabolite signals and decrease the variance of results across a large *in-vivo* dataset while using less than 20% of the transients.

4.4. Methods:

4.4.1. AGNOSTIC:

Data in the AGNOSTIC dataset was simulated using the full quantum mechanical density matrix formalism (Blum, 1981; Fano, 1957; Farrar, 1990; Sørensen et al., 1984). Calculations were carried out using an in-house python programming script using numpy (Harris et al., 2020). A total of 270 basis sets were created across 18 field strengths (1.4Tesla- 3.1Tesla; stride of .1T) and 15 echo times (10ms-80ms; stride of 5ms). Ideal pulses were used and followed the Point Resolved Spectroscopy (*PRESS*) pulse sequence protocol. Spectral width in ppm was chosen to be ~62.63ppm, centered on 4.7ppm, for all field strengths which is equivalent to 8000hz (*Dwell Time = 125microseconds)* at 3.0Tesla. Data acquisition was started immediately following the last refocusing pulse, instead of the top of the echo, to allow inclusion of points before the echo. Each metabolite signal was output as a numpy array with a fixed length of 16684 complex time points (*300 points before the top of echo and 16384 points after the top of echo*) down the rows. Individual metabolite spins were then across the columns. Different spectral widths/dwell times were achieved by subsampling the 16684 data points to as many as every 8^{th} datapoint (i.e., at 3.0T every 8^{th} data point results in a dwell time of 1ms, for a spectral width of 1000hz, and still contains a sizeable 2048 data points).

The dataset is structured as a zipped numpy archive file (.npz) and can be opened as a python dictionary object. This dictionary contains complex-valued numpy arrays with the

metabolite, macromolecule, water, and noise time-domain data which can be combined in different ways depending on the target of interest. For instance, a denoising model may want to target the combined metabolite, MM, and water signal without the noise. Within the file, all the acquisition parameters (field strength, echo time, spectral width, etc.), simulation parameters (signal to noise, full-width half-max, concentrations, T^2 relaxation, etc.), and data augmentation options are included.

4.4.1.1. Metabolite Data:

21 Brain Metabolites were included: Alanine (Ala), Ascorbate (Asc), Aspartate (Asp), Creatine (Cre), y-Amino-Butyric Acid (GABA), Glucose (Glc), Glutamine (Gln), Glutamate (Glu), Glycerophosphocholine (GPC), Glutathione (GSH), Glycine (Gly), Water (H2O), Lactate (Lac), Myo-Inositol (Myo), N-Acetyl-Aspartate (NAA), N-Acetyl-Aspartate-Glutamate (NAAG), Phosphocholine (PCho), Phosphocreatine (PCr), Phosphoethanolamine (PhE), Scyllo-Inositol (Scy), and Taurine (Tau). GABA was modeled as either a triplet of doublets or a triplet (Govindaraju et al., 2000; Near et al., 2012). Glucose was modeled as either glucose alpha or glucose beta.

Metabolite concentrations were taken from the COHERENC database (Gudmundson, In Prep). For all spectra, we began with ±2 standard deviations from the mean of healthy values. For clinical spectra, we applied ±2 standard deviations of the linear change relative to healthy control following the same meta-analytic approach from Gudmundson, In Prep.

Metabolite intrinsic T^2 relaxation was modeled with an exponential decay, or Lorentzian line shape in the frequency domain (Juchem & de Graaf, 2017). Values were calculated using the multiple meta-regression model from Gudmundson, In Prep. Specifically, values were computed at 1.5T, which theoretically is least influenced by extrinsic T^2 contributions (Bloembergen et al.,

1948; De Graaf et al., 2006; Michaeli et al., 2002). To create a range, we took the lowest and highest values after computing across various pulse sequences and tissue types. Finally, we used 10% of the 95% confidence interval to create a conservative range of possible relaxation rates, assuming that a large degree of T^2* decay changes can be attributed to extrinsic contributions. Using these ranges, metabolite intrinsic decays were randomly drawn from a uniform distribution.

Metabolite extrinsic T^2 relaxation was modeled with a Gaussian decay. The rate of decay was randomly chosen, using a uniform distribution, to achieve a full-width half-max (*FWHM*) of the NAA CH3 singlet (2.008ppm) between 5hz-18hz. A small amount of jitter was added such that each signal would contain a similar, but not exactly the same, amount of Gaussian decay.

4.4.1.2 Macromolecule Data:

14 Macromolecule (*MM*) signals were included at 0.92ppm, 1.21ppm, 1.39ppm, 1.67ppm, 2.04ppm, 2.26ppm, 2.56ppm, 2.70ppm, 2.99ppm, 3.21ppm, 3.62ppm, 3.75ppm, 3.86ppm, 4.03ppm (Cudalbu et al., 2021; Giapitzakis et al., 2018). A small amount of jitter *(+/- 0.03ppm*) was added to these ppm values to account for the observed variance in reported ppm locations of these.

These signals were simulated as a singlet with uniform random exponential decay between 20ms-60ms (Murali-Manohar et al., 2020). The concentration and FWHM parameters are modeled within healthy ranges (Murali-Manohar, Giapitzakis). Just as the metabolite signals, the FWHM was achieved by increasing the rate of gaussian decay.

4.4.1.3 Noise:

Noise was generated using a random normal distribution. The noise amplitude was such that the signal to noise (SNR) of the NAA singlet (2.008ppm) was between 5-80 (*NAA Height/StdDev of the noise*). In addition to the noise vector, the amplitudes are stored within the archive file to generate additional noise signals with a similar SNR.

4.4.1.4 Residual Water:

The residual water basis signal was simulated as a singlet at 4.7ppm. This water signal was used to then create 5 unique water components with different ppm locations, phase, and amplitudes (Lin et al., 2019). The ppm location, phase, and amplitude ranges are described in table 1. The final water signal was scaled such that it was 1x-20x larger than the maximum value of the metabolite signal. A random number, from 0 to 5, was chosen to include a different number of components across spectra. The final time-domain water signal includes the summation of these components. The water components used along with their corresponding ppm location, phase, and amplitudes are stored within the numpy archive file.

Component	**Location (*ppm*)**		**Phase (*degrees*)**		**Amplitude (*arbitrary*)**	
1	4.679	4.711	-10	10		1.00
2	4.599	4.641	15	45	.35	.55
3	4.801	4.759	-60	-30	.35	.55
4	4.449	4.541	45	-70	.10	.25
5	4.859	4.901	105	135	.10	.25

Table 4.1.: Residual water simulation parameters.

4.4.1.5 Frequency and Phase Shifts:

Frequency and phase shifts are also available for inclusion. Frequency shifts of ±0.313ppm can be applied. 0^{th} order phase shifts from ±1.57 radians are available. 1^{st} order phase shifts are available from ±0.34 radians with a pivot point from 2ppm-5ppm. Users are welcome to use these pre-defined shifts or create their own to apply to the raw signals depending on their goals.

Shifts	**Range**	
0^{th} Order Phase	-1.57 radians	1.57 radians
1^{st} Order Phase	-0.34 radians	0.34 radians
Frequency	-0.313 ppm	0.313 ppm

Table 4.2.: Predefined Phase and Frequency Shifts

4.4.1.6 Data Augmentation:

Data can be further augmented by removing signals from the final spectrum. Metabolite, MM, residual water, and noise can easily be left off by not adding them into the final time-domain signal. We have also preselected the most prominent signals, NAA, Cre, and Cho or one MM signal to be excluded in 20% of examples if desired. Since the metabolite and MM signal are the summation of all their respective signals, we've included a separate vector, "Drop_Sig," which includes one of the isolated metabolite or MM signals to subtract from the final combined signal. Other data augmentation methods can be used and applied to the raw data as users see fit.

4.4.2. In-Vivo data:

4.4.2.1. Participants:

In-vivo data was provided through a collaboration between the University of California, Irvine, Johns Hopkins University, and the Shandong First Medical University. One hundred and two healthy volunteers were recruited from the Shandong Provincial Hospital with local IRB approval. There were approximately 20 participants (10 female and 10 male) per decade of life, from 20s-60s. Participants were excluded if they had a history of neurological or psychiatric illnesses.

4.4.2.2. Magnetic Resonance Protocol:

Data was collected using a 3T Philips Ingenia CX MRI scanner. Each scan session began with a T1-weighted MPRAGE scan (TR/TE 6.9ms/3.2ms; FA 8°) with $1mm^3$ isotropic resolution for voxel positioning and tissue segmentation. Next, water suppressed (VAPOR) metabolite data was acquired using both a PRESS and Semi-Localized Adiabatic Selective Refocusing (sLASER) protocol. Water unsuppressed data was collected for metabolite referencing. Voxels (30mm x 26mm x 26mm) were collected from both the Centrum Semiovale (CSO) and Posterior Cingulate Cortex (PCC). The PRESS acquisition had a TR/TE/Bandwidth of 2000ms/30ms/2000Hz, while the sLASER had a TR/TE/Bandwidth of 2000ms/30ms/2000Hz. A total of 96 transients were collected from each of the 4 voxels. Tissue segmentation was performed in OSPREY which employs SPM12 (Friston et al., 1994).

4.4.3. Neural Network Models:

4.4.3.1. Convolutional Autoencoder:

Figure 3.1. shows the convolutional autoencoder model used during training. This model is fully convolutional with an input and output of 4096 (timepoints) x 2 (real and imaginary

components) x 1 (convolutional filter). Strided-convolutions (stride=2), rather than max-pooling, are used to downsample the input. The central-most layer (Layer 9) is the bottleneck latent space with shape 4 (timepoints) x 128 (convolutional filters). Following the latent space, data funnels into two pathways (labeled as A and B) simultaneously in order to produce fully processed metabolite and macromolecule signals. The full network contains 1,237,786 parameters (*1,237,786 trainable and 5,584 non-trainable*). There are approximately 27 convolutional blocks where each convolutional block consists of 2 convolutional layers of stride-1 or stride 2. Batch Normalization, a Leaky Rectified Linear Unit (RELU) activation function are used on each convolutional layer.

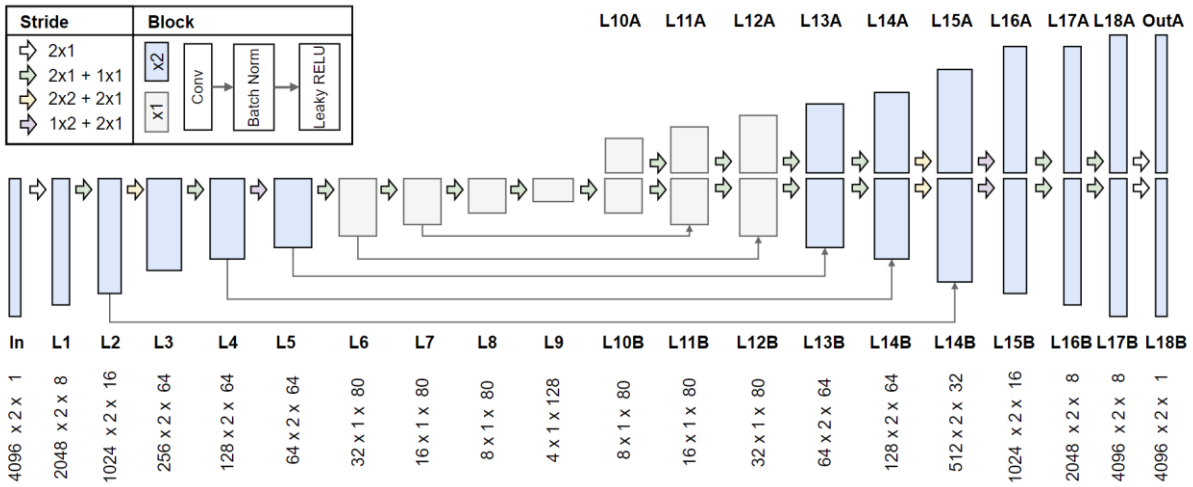


Figure 4.1.: Breakdown of the fully convolutional autoencoder used for multi-task learning.

4.5. Results:

4.5.1. AGNOSTIC Latent Space Representation:

To determine how well AGNOSTIC represented actual in-vivo data, we investigated the latent space of the convolutional auto-encoder. Here, we passed 1,000 never before seen examples (600 AGNOSTIC and 400 in-vivo) through the network and collected the resulting

values from the latent space. The latent space vector, or center-most layer, is of shape 4 (timepoints) x 128 (convolutional filters) and was unraveled to create a vector of 512 datapoints. We then performed a single value decomposition (SVD) on the array of 1000 unseen examples x 512 datapoints. Figure 3.2 shows principal components (PC) 1 and 2. The data is further separated into the 4 different acquisitions (2 protocols and 2 brain regions) performed on each of the participants. The blue datapoints show the AGNOSTIC data, while the non-blue points show the in-vivo data.

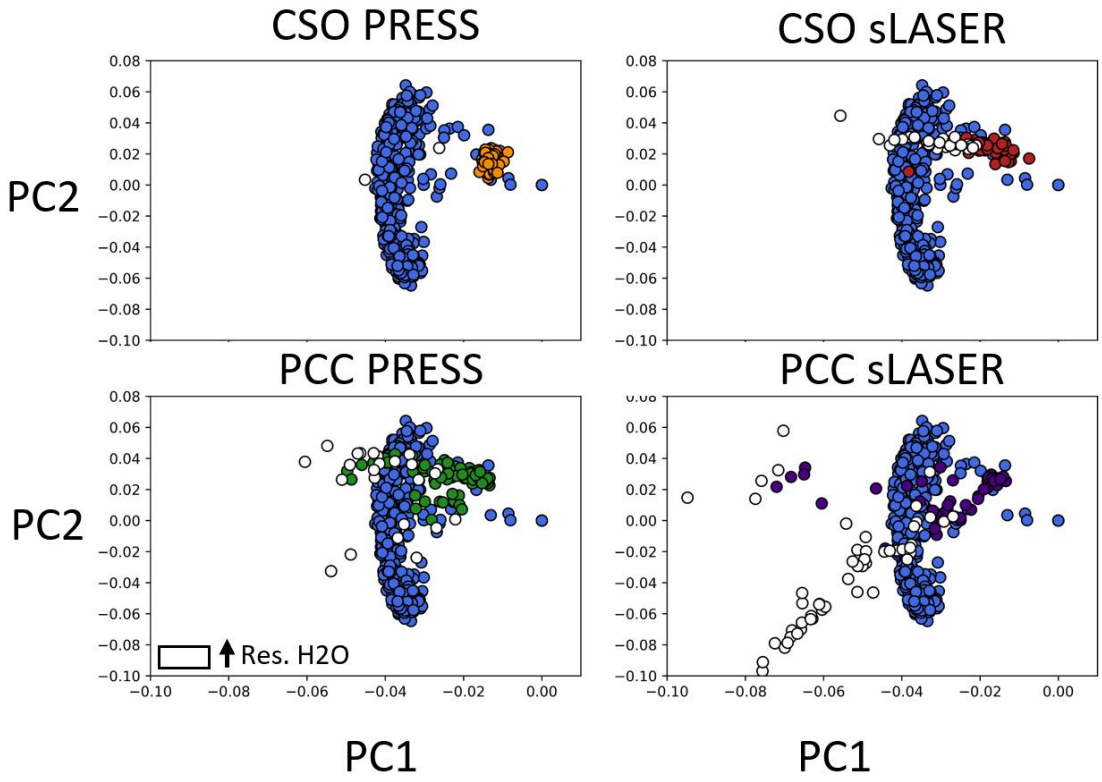


Figure 4.2.: Latent Space Representation of Simulated data (blue) compared to in-vivo (colored) data. White datapoints show data with high residual water signal. Data is from Centrum Semiovale (CSO) and Posterior Cingulate Cortex (PCC) from both the PRESS and sLASER protocols.

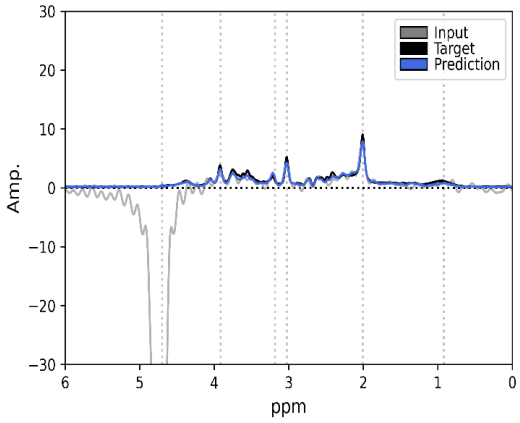
4.5.2. Multi-Target Learning Results without Phase and Frequency Correction:

To determine how well a convolutional neural network performed on MRS data, we trained an initial model on data from AGNOSTIC. Figure 4.3 shows 3 examples where the grey represents the frequency domain input, black represents the ground-truth target, and blue represents the neural network's prediction following training.

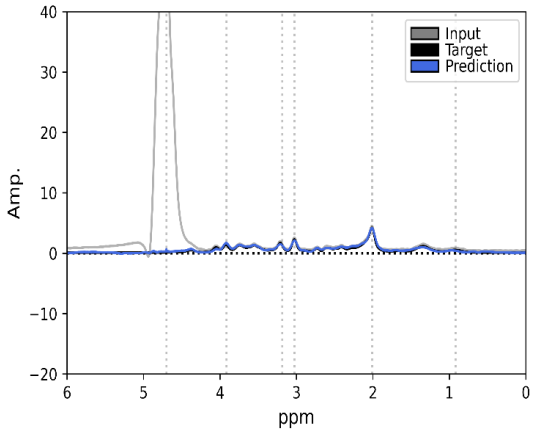
4.5.3. Multi-Target Learning Results with Phase and Frequency Correction:

Using the initial convolutional model, we then passed in data from AGNOSTIC that contained phase and frequency shifts. Performance was significantly impacted and the network failed to capture the necessary corrections. The left-hand column in Figure 4.4 shows the same 3 examples from 4.5.2. where the grey represents the frequency domain input, black represents the ground-truth target, and blue represents the neural network's prediction following training.

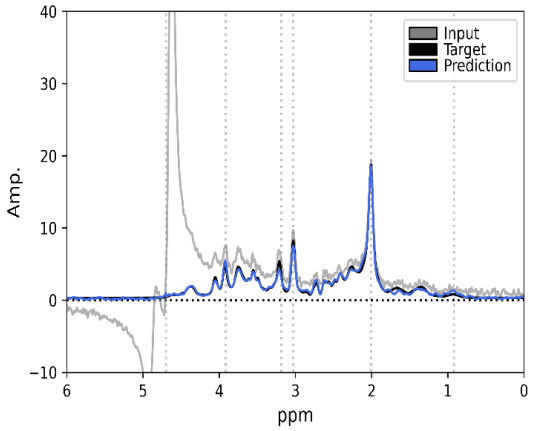
Next, we retrained the convolutional model using data from AGNOSTIC that contained phase and frequency shifts. Performance was restored and the network captured the required phase and frequency shifts. The right-hand column in Figure 4.4 shows the same 3 examples from 4.5.2. where the grey represents the frequency domain input, black represents the ground-truth target, and blue represents the neural network's prediction following training.



Points	256
Freq.	0.0Hz
0^{th} Ph	0.0^{o}
H_2O	Neg
SNR	~25
FWHM	7.5Hz



Points	512
Freq.	0.0Hz
0^{th} Ph	0.0^{o}
H_2O	Pos
SNR	~65
FWHM	8.4Hz



Points	1024
Freq.	0.0Hz
0^{th} Ph	0.0^{o}
H_2O	Neg
SNR	~60
FWHM	6.8Hz

Figure 4.3.: Multi-Task Learning Convolutional Autoencoder performance on synthetic data with no phase or frequency shifts.

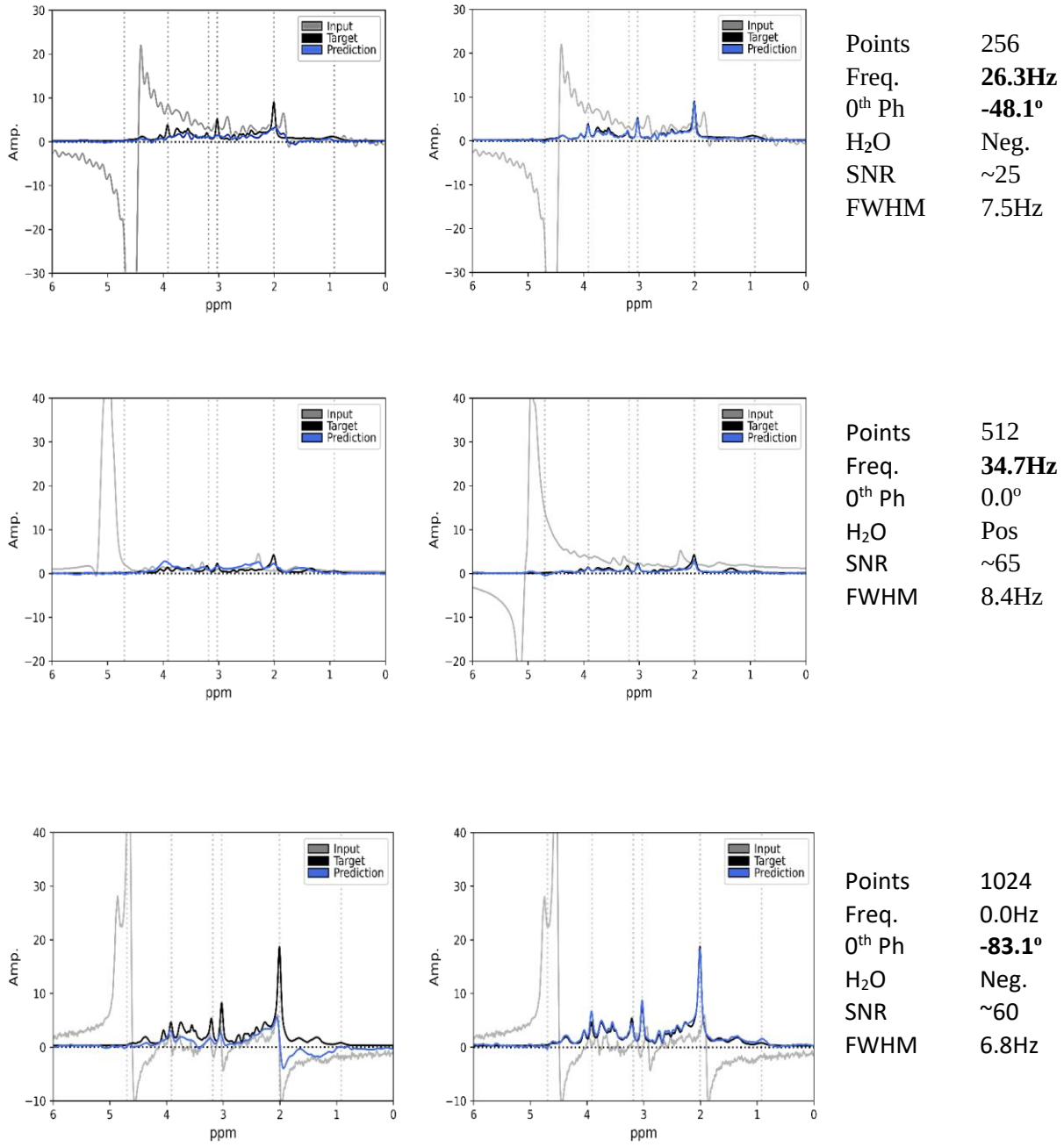


Figure 4.4.: Demonstration of Multi-Task Learning Convolutional Autoencoder. The Left column shows the same spectra where training was performed without phase and frequency shifts. In the middle column, the network was retrained incorporating phase and frequency shifts. The phase and frequency shifts and other spectral parameters are shown on the right-hand side.

4.5.4. Multi-Target Learning Results with Simulated Tumor Data:

Finally, using the previously retrained model, we showed that performance was once again impaired when using a simulated tumor dataset. The left-hand column in Figure 4.5. contains a red-box over the Lactate Methyl signal at 1.31ppm. Here, the network failed to capture the introduction of the Lactate signal as this was not typically present in the healthy data. On the right-hand side, network performance is then restored upon retraining with both healthy and clinical data.

4.5.5. Multi-Target Learning Results with *In-Vivo* Data:

Finally, we used the *in-vivo* data to determine how well the model was performing. Here, we show the metabolite (left-hand column) and Macromolecule (right-hand column) outputs. The network performed with high accuracy with correlations ranging from 75%-90%

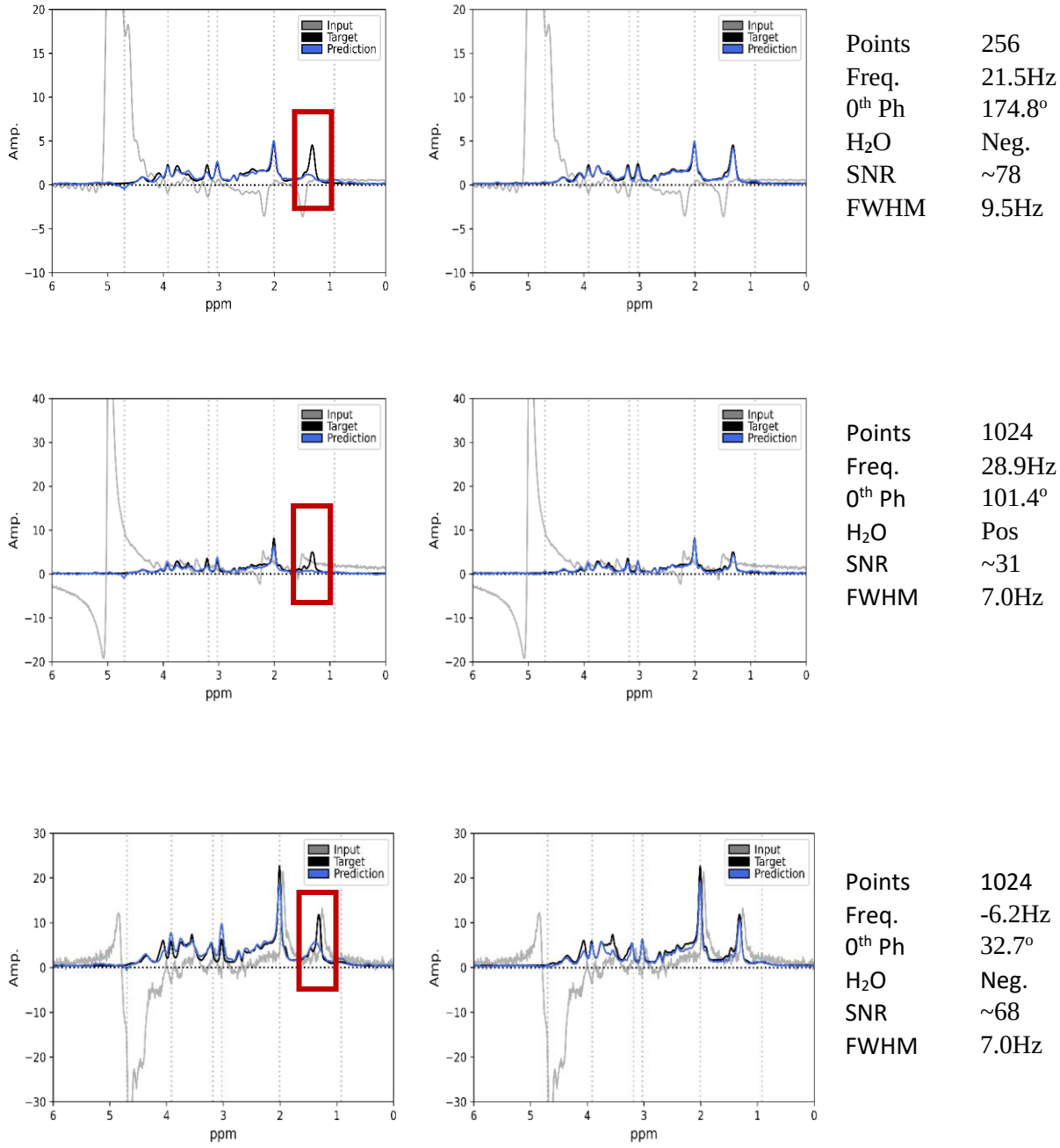


Figure 4.5.: Demonstration of Multi-Task Learning Convolutional Autoencoder. The Left column shows the same spectra where training was performed with data only sampled from healthy populations. In the middle column, the network was retrained incorporating data from clinical populations. The phase and frequency shifts and other spectral parameters are shown on the right-hand side.

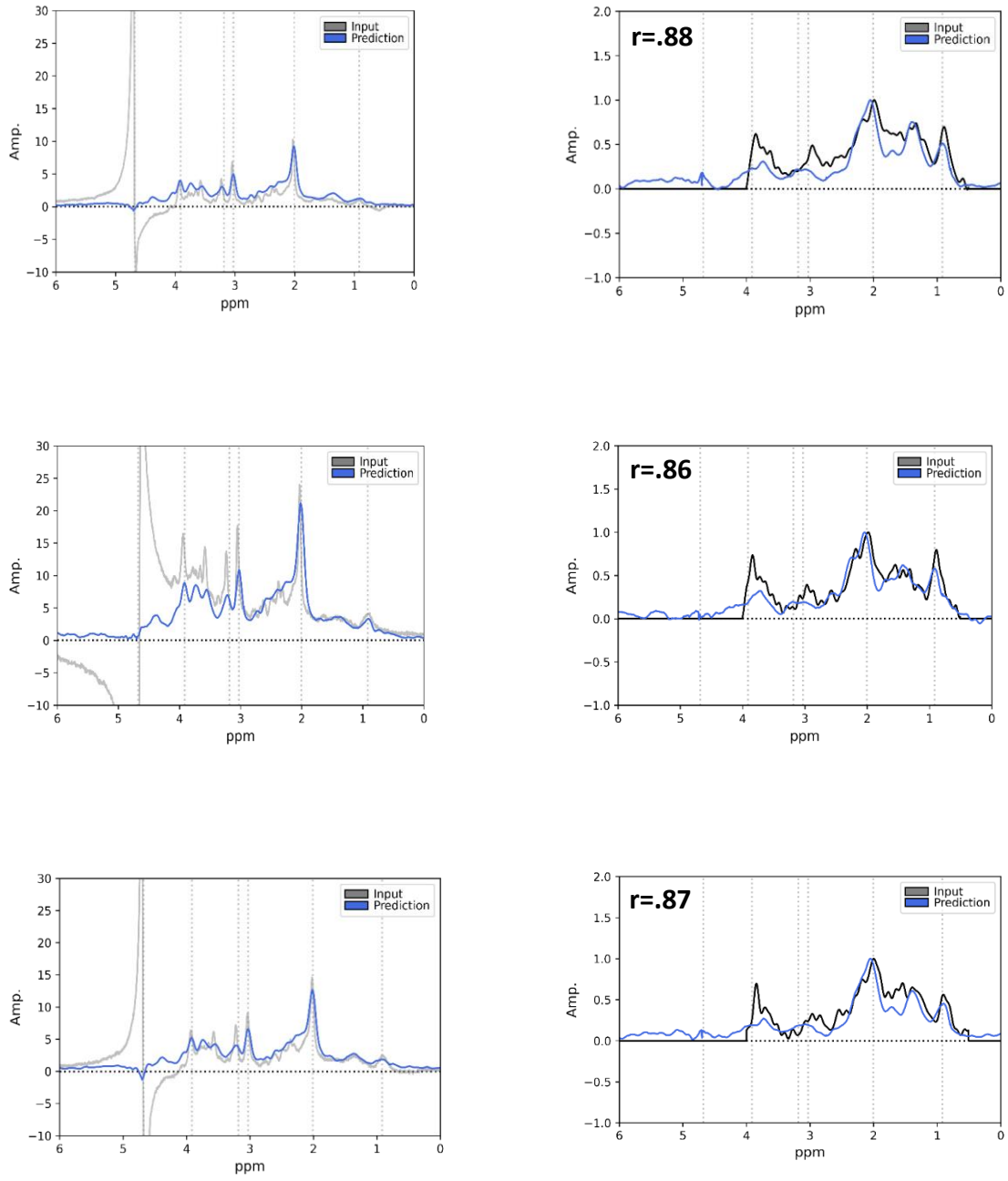


Figure 4.6.: Demonstration of Multi-Task Learning Convolutional Autoencoder for *in-vivo* data. The Left column shows the predictions on then metabolite data while the right columns shows the macromolecule spectra compared against the metabolite-nulled macromolecule data. The Pearson correlation value (r) is shown in the upper left-hand corner of the MM data, showing strong correlations to the actual experimentally acquired data.

4.6. Discussion:

Our work was motivated by many previous studies. Frequency and 0^{th} order phase correction have been performed for frequency-domain data successfully by using either fully-connected or convolutional neural networks (Ma et al., 2022; Tapper et al., 2021). Convolutional neural networks have also been shown to have success in removing spurious echoes, or ghosting, artifacts by using a short-time Fourier transform to produce 2 dimensional (frequency x time) MRS data (Kyathanahally et al., 2018). Fully connected auto-encoders were shown to disentangle metabolite from MM signals in the time-domain (Li et al., 2020). Convolutional auto-encoders have also been used to perform full preprocessing routines and isolate noise-free metabolite data in the frequency domain (Lee & Kim, 2019, 2020). Convolutional neural networks have also be trained to quantify relative metabolite concentrations in the frequency domain (Chandler et al., 2019). Recurrent neural networks have not been used as much, but in one case were shown to perform well for tumor vs. pseudo tumor classification using long short-term memory (LSTM) and Bidirectional LSTM models (Dandıl & Karaca, 2021).

As these many examples demonstrate, neural networks promise to be a great step forward in MRS data analysis. Similar to the early days or MRS, standardization of neural network methods has not yet been addressed. Studies have used a variety of basis sets and parameters to create synthetic data and generally show proof of concept for small single protocol healthy in-vivo datasets.

In this work, we sought to make progress in two areas of advanced MRS data analysis. First, we developed and proposed the use of the AGNOSTIC benchmark dataset for training and evaluating performance across various models. The dataset was intended to be **A G**eneralized representation of MRS data, encompassing the features of various acquisition protocols. Through

its flexibility, a wide array of models that use convolutional, recurrent, or fully-connected **N**eural networks can be trained and test against. It is also made completely **O**pen-**S**ource in order to provide the field at large with standardized resource. AGNOSTIC also includes 500,000 (healthy and clinical) examples which can be partitioned for **T**raining, validation, and testing. The data is fully randomized with uniform distributions of FWHM, noise, correction factors, and more to facilitate models that are **I**nsensitive to **C**onfines of simplistic data. To further demonstrate these ideals, we used the latent space of trained CNN to show that *in-vivo* data with the simulated data found within AGNOSTIC. While we designed the dataset with neural networks in mind, it can further be used as a means to compare traditional preprocessing and quantification software packages.

Next, using AGNOSTIC we further demonstrated the strengths and weaknesses of different model architectures. In particular, models that are unfamiliar with specific features, such as phase and frequency shifts, were incapable of maintaining their accuracy. Additionally, models that have are only trained with healthy representations of MRS data will misattribute or ignore signals which can restrict their usage from clinical settings.

Using a convolutional neural network, we demonstrated that a single network is capable of correcting phase and frequency shifts, removing residual water and MM signals, and isolating metabolites from noise. This network generalized well to the high degree of variability in the training examples and further reproduces the results seen in past autoencoder models. These Multi-Task Learning (*MTL*) networks appear to benefit from a greater degree of regularization and attention as these networks must rely on efficient generalizations of inputs to capture all targets simultaneously (Argyriou et al., 2008, 2006). However, it is less understood whether these networks are capable of capturing small details that a more targeted network would.

As is seen in other fields, single-target neural networks may perform better in instances where multiple tasks interfere with one another thereby causing negative transfer. Another advantage of targeted networks is that they provide users with a greater degree of confidence in seeing step-by-step manipulations as opposed to a complete transformation of data within a single output. Ultimately, finding the balance of using auxiliary tasks and negative transfer should be further explored.

4.7. Limitations:

Our study is not without limitations and marks only the next step towards the use of neural networks for MRS data analysis. While data within AGNOSTIC is varied across field strength and echo time, it was all synthesized with a PRESS basis set. A truly general effort would capture the spin coupling patterns found within stimulated echo acquisition methods (STEAM and SPECIAL) and adiabatic sequences (LASER and semi-LASER).

In creating the synthetic data basis sets, we used ideal pulses that perform perfect rotations and thus fails to capture any pulse imperfections and/or spatial effects that are observed with 3-dimensional voxels. Here, we opted to capture the coupling, increased frequency dispersion, and evolution of spins by spending more time creating basis sets across many field strengths and echo times. Including spatial features with a variety of voxel sizes would greatly lengthen the time required to create the basis sets. However, just as in linear combination modeling, the accuracy in basis set may play a critical role in the success of neural network models.

It is also important to note that spectral editing through j-difference methods such as basing or MEscher and GArwood (MEGA) was not performed here and may introduce a

significant change to metabolite representations (Mescher et al., 1996; Star-Lack et al., 1998). It is unknown to what extent these factors play as neural networks may be capturing very low-level features of time or frequency-based signals. Still, future work and updated versions of AGNOSTIC aim to include an even greater degree of MRS data variability.

With respect to the neural network models designed here, we don't expect these models to stand the test of time. As in other fields, advances in neural network models come at a high rate. Instead, the models here primarily designed to illustrate the strengths and weaknesses of neural networks, compare architectures, as well as provide some initial benchmarks for the field. Future models may aim to utilize deeper architectures, pretrained weights, or simply optimize hyperparameters such as learning rate or loss function. Regardless, we look forward to the continued progress in deep learning.

4.8. Conclusion:

In this work we have developed AGNOSTIC, a dataset for neural network development, and demonstrated the ability of various neural network architectures to process and quantify MRS data. AGNOSTIC was designed to be flexible in order to generalize well with *in-vivo* data. Both healthy and clinical representations are included within the dataset and provide a starting point for using neural network-based diagnostic tools. Investigations into the latent space provide an effective way to determine whether synthetic data is being represented the same as *in-vivo* data within networks. We showed that, as in other domains, the training set must be very carefully chose as neural networks fail to capture features of data that were not included during training. Here, small imperfections, such as phase and frequency shifts, could impact accuracy and their effectiveness in *in-vivo* data. Ultimately, a range of different models can be used and different models may benefit from time-domain vs frequency domain data structures.

4.9. References:

Argyriou, A., Evgeniou, T., & Pontil, M. (2008). Convex multi-task feature learning. *Machine Learning*, *73*(3), 243–272. https://doi.org/10.1007/s10994-007-5040-8

Argyriou, A., Evgeniou, T., & Pontil, M. (2006). *Multi-Task feature learning* (Advances in Neural Information Processing Systems (Ed.); pp. 41–48). Advances in Neural Information Processing Systems. https://papers.nips.cc/paper/2006/file/0afa92fc0f8a9cf051bf2961b06ac56b-Paper.pdf

Bloembergen, N., Purcell, E. M., & Pound, R. V. (1948). Relaxation effects in nuclear magnetic resonance absorption. *Physical Review*, *73*(7), 679–712. https://doi.org/10.1103/PhysRev.73.679

Blum, K. (1981). *Density Matrix Theory and Applications* (1st ed.). Springer US. https://doi.org/10.1007/978-1-4615-6808-7

Chandler, M., Jenkins, C., Shermer, S. M., & Langbein, F. C. (2019). *MRSNet: Metabolite Quantification from Edited Magnetic Resonance Spectra With Convolutional Neural Networks*. 1–12. http://arxiv.org/abs/1909.03836

Cudalbu, C., Behar, K. L., Bhattacharyya, P. K., Bogner, W., Borbath, T., de Graaf, R. A., Gruetter, R., Henning, A., Juchem, C., Kreis, R., Lee, P., Lei, H., Marjańska, M., Mekle, R., Murali-Manohar, S., Považan, M., Rackayová, V., Simicic, D., Slotboom, J., … Mlynárik, V. (2021). Contribution of macromolecules to brain 1H MR spectra: Experts' consensus recommendations. *NMR in Biomedicine*, *34*(5), 1–24. https://doi.org/10.1002/nbm.4393

Dandıl, E., & Karaca, S. (2021). Detection of pseudo brain tumors via stacked LSTM neural networks using MR spectroscopy signals. *Biocybernetics and Biomedical Engineering*,

41(1), 173–195. https://doi.org/10.1016/j.bbe.2020.12.003

De Graaf, R. A., Brown, P. B., McIntyre, S., Nixon, T. W., Behar, K. L., & Rothman, D. L. (2006). High magnetic field water and metabolite proton T1 and T 2 relaxation in rat brain in vivo. *Magnetic Resonance in Medicine*, *56*(2), 386–394. https://doi.org/10.1002/mrm.20946

Edelstein, W. A., Glover, G. H., Hardy, C. J., & Redington, R. W. (1986). The intrinsic signal-to-noise ratio in NMR imaging. *Magnetic Resonance in Medicine*, *3*(4), 604–618. https://doi.org/10.1002/mrm.1910030413

Fano, U. (1957). Description of States in Quantum Mechanics by Density Matrix and Operator Techniques. *Reviews of Modern Physics*, *29*(1), 74–93. https://doi.org/10.1103/RevModPhys.29.74

Farrar, T. (1990). Density matrices in NMR spectroscopy: Part I. *Concepts in Magnetic Resonance*, *2*, 1–12.

Fei-Fei, L., Deng, J., & Li, K. (2010). ImageNet: Constructing a large-scale image database. *Journal of Vision*, *9*(8), 1037–1037. https://doi.org/10.1167/9.8.1037

Friston, K. J., Holmes, A. P., Worsley, K. J., Poline, J. -P, Frith, C. D., & Frackowiak, R. S. J. (1994). Statistical parametric maps in functional imaging: A general linear approach. *Human Brain Mapping*, *2*(4), 189–210. https://doi.org/10.1002/hbm.460020402

Giapitzakis, I. A., Avdievich, N., & Henning, A. (2018). Characterization of macromolecular baseline of human brain using metabolite cycled semi-LASER at 9.4T. *Magnetic Resonance in Medicine*, *80*(2), 462–473. https://doi.org/10.1002/mrm.27070

Govindaraju, V., Young, K., & Maudsley, A. A. (2000). Proton NMR chemical shifts and coupling constants for brain metabolites. *NMR in Biomedicine*, *13*(3), 129–153. https://doi.org/10.1002/1099-1492(200005)13:3<129::AID-NBM619>3.0.CO;2-V

Harris, C. R., Millman, K. J., van der Walt, S. J., Gommers, R., Virtanen, P., Cournapeau, D., Wieser, E., Taylor, J., Berg, S., Smith, N. J., Kern, R., Picus, M., Hoyer, S., van Kerkwijk, M. H., Brett, M., Haldane, A., del Río, J. F., Wiebe, M., Peterson, P., … Oliphant, T. E. (2020). Array programming with NumPy. *Nature*, *585*(7825), 357–362. https://doi.org/10.1038/s41586-020-2649-2

Juchem, C., & de Graaf, R. A. (2017). B0 magnetic field homogeneity and shimming for in vivo magnetic resonance spectroscopy. *Analytical Biochemistry*, *529*, 17–29. https://doi.org/10.1016/j.ab.2016.06.003

Kyathanahally, S. P., Döring, A., & Kreis, R. (2018). Deep learning approaches for detection and removal of ghosting artifacts in MR spectroscopy. *Magnetic Resonance in Medicine*, *80*(3), 851–863. https://doi.org/10.1002/mrm.27096

Lee, H. H., & Kim, H. (2019). Intact metabolite spectrum mining by deep learning in proton magnetic resonance spectroscopy of the brain. *Magnetic Resonance in Medicine*, *82*(1), 33–48. https://doi.org/10.1002/mrm.27727

Lee, H. H., & Kim, H. (2020). Deep learning-based target metabolite isolation and big data-driven measurement uncertainty estimation in proton magnetic resonance spectroscopy of the brain. *Magnetic Resonance in Medicine*, *84*(4), 1689–1706. https://doi.org/10.1002/mrm.28234

Li Deng. (2012). The MNIST Database of Handwritten Digit Images for Machine Learning

Research [Best of the Web]. *IEEE Signal Processing Magazine*, *29*(6), 141–142. https://doi.org/10.1109/MSP.2012.2211477

Li, Y., Wang, Z., & Lam, F. (2020). Separation of Metabolite and Macromolecule Signals for 1 H-Mrsi Using Learned Nonlinear Models. *Proceedings - International Symposium on Biomedical Imaging*, *2020-April*, 1725–1728. https://doi.org/10.1109/ISBI45749.2020.9098365

Lin, L., Považan, M., Berrington, A., Chen, Z., & Barker, P. B. (2019). Water removal in MR spectroscopic imaging with L2 regularization. *Magnetic Resonance in Medicine*, *82*(4), 1278–1287. https://doi.org/10.1002/mrm.27824

Ma, D. J., Le, H. A. M., Ye, Y., Laine, A. F., Lieberman, J. A., Rothman, D. L., Small, S. A., & Guo, J. (2022). MR spectroscopy frequency and phase correction using convolutional neural networks. *Magnetic Resonance in Medicine*, *87*(4), 1700–1710. https://doi.org/10.1002/mrm.29103

Maudsley, A. A., Andronesi, O. C., Barker, P. B., Bizzi, A., Bogner, W., Henning, A., Nelson, S. J., Posse, S., Shungu, D. C., & Soher, B. J. (2021). Advanced magnetic resonance spectroscopic neuroimaging: Experts' consensus recommendations. *NMR in Biomedicine*, *34*(5), 1–22. https://doi.org/10.1002/nbm.4309

Maudsley, A. A., Domenig, C., Ramsay, R. E., & Bowen, B. C. (2010). Application of volumetric MR spectroscopic imaging for localization of neocortical epilepsy. *Epilepsy Research*, *88*(2–3), 127–138. https://doi.org/10.1016/j.eplepsyres.2009.10.009

Mescher, M., Tannus, A., O'Neil Johnson, M., & Garwood, M. (1996). Solvent suppression using selective echo dephasing. *Journal of Magnetic Resonance - Series A*, *123*(2), 226–

229. https://doi.org/10.1006/jmra.1996.0242

Michaeli, S., Garwood, M., Zhu, X.-H., DelaBarre, L., Andersen, P., Adriany, G., Merkle, H., Ugurbil, K., & Chen, W. (2002). ProtonT2 relaxation study of water, N-acetylaspartate, and creatine in human brain using Hahn and Carr-Purcell spin echoes at 4T and 7T. *Magnetic Resonance in Medicine*, *47*(4), 629–633. https://doi.org/10.1002/mrm.10135

Murali-Manohar, S., Borbath, T., Wright, A. M., Soher, B., Mekle, R., & Henning, A. (2020). T2 relaxation times of macromolecules and metabolites in the human brain at 9.4 T. *Magnetic Resonance in Medicine*, *84*(2), 542–558. https://doi.org/10.1002/mrm.28174

Near, J., Harris, A. D., Juchem, C., Kreis, R., Marjańska, M., Öz, G., Slotboom, J., Wilson, M., & Gasparovic, C. (2021). Preprocessing, analysis and quantification in single-voxel magnetic resonance spectroscopy: experts' consensus recommendations. *NMR in Biomedicine*, *34*(5), 1–23. https://doi.org/10.1002/nbm.4257

Near, J., Leung, I., Claridge, T., Cowen, P., & Jezzard, P. (2012). Chemical shifts and coupling constants of the GABA spin system. *Proc. Intl. Soc. Mag. Reson. Med.*, *20*(1993), 4386. http://cds.ismrm.org/protected/12MProceedings/files/4386.pdf

Öz, G., Deelchand, D. K., Wijnen, J. P., Mlynárik, V., Xin, L., Mekle, R., Noeske, R., Scheenen, T. W. J., Tkáč, I., Andronesi, O., Barker, P. B., Bartha, R., Berrington, A., Boer, V., Cudalbu, C., Emir, U. E., Ernst, T., Fillmer, A., Heerschap, A., … Wilson, M. (2021). Advanced single voxel 1H magnetic resonance spectroscopy techniques in humans: Experts' consensus recommendations. *NMR in Biomedicine*, *34*(5), 1–18. https://doi.org/10.1002/nbm.4236

Sørensen, O. W., Eich, G. W., Levitt, M. H., Bodenhausen, G., & Ernst, R. R. (1984). Product

operator formalism for the description of NMR pulse experiments. *Progress in Nuclear Magnetic Resonance Spectroscopy*, *16*, 163–192. https://doi.org/10.1016/0079-6565(84)80005-9

Star-Lack, J., Spielman, D., Adalsteinsson, E., Kurhanewicz, J., Terris, D. J., & Vigneron, D. B. (1998). In Vivo Lactate Editing with Simultaneous Detection of Choline, Creatine, NAA, and Lipid Singlets at 1.5 T Using PRESS Excitation with Applications to the Study of Brain and Head and Neck Tumors. *Journal of Magnetic Resonance*, *133*(2), 243–254. https://doi.org/10.1006/jmre.1998.1458

Tapper, S., Mikkelsen, M., Dewey, B. E., Zöllner, H. J., Hui, S. C. N., Oeltzschner, G., & Edden, R. A. E. (2021). Frequency and phase correction of J-difference edited MR spectra using deep learning. *Magnetic Resonance in Medicine*, *85*(4), 1755–1765. https://doi.org/10.1002/mrm.28525

5. Conclusions:

The goal of this body of work was to help move MRS towards greater clinical relevance. As demonstrated in Chapter 3, MRS has been used for more than 25 clinical diseases associated with the brain. Though, in comparison to standard MRI, MRS is scarcely used. The field has made numerous efforts to

5.1 COHERENC Database and Meta-Analysis:

In chapter 2, we describe the research effort towards building the **C**omprehensive **O**pen-Source **H**osted **E**lectronically **RE**laxation & Concentration **N**eurometabolite Values for **C**linical Research (*COHERENC*) database for brain MRS results. The database hosts results from ~500 publications and contains information on the concentration and relaxation results from both healthy and clinical populations. Importantly, we've made the database freely available online where users can view and contribute their own data. We've also used the database to determine the physiological ranges of 22 brain metabolites across the healthy lifespan. We've further provided linear model that describes changes in metabolite concentrations and giving an entire expected metabolic profile for 25 different clinical populations. Finally, we use a multiple meta-regression model to predict metabolite relaxation rates across varied acquisition schemes, accounting for ~80% of the variance found between studies.

5.2 AGNOSTIC and Neural Network-based MRS data analysis:

In Chapter 3 we describe our work in building **A G**eneralized **N**eural-network **O**pen-source **S**pectroscopy **T**raining dataset **I**nsensitive to acquisition **C**onfines (AGNOSTIC) for flexible neural network training and benchmarking. The dataset contains 500,00 healthy and clinical examples of MRS data. Our goal was to produce a flexible dataset that could be used for

different targets and represent the high degree of variance found within MRS data. A total of 270 basis sets were used to simulate data from 18 field strengths (1.4T-3.1T) and 15 echo times (10ms-80ms). Using latent space representations, we further show that our neural network models are, in fact, representing the synthetic data similar to the *in-vivo* data. We further built a series of networks to demonstrate the strengths and weaknesses of using neural networks for MRS data analysis. Ultimately, a range of different models can be used so long as the synthetic data is well representative of the *in-vivo* data, including, for instance, metabolite concentrations from clinical populations.

5.3 Future Directions:

5.3.1. Clinical Application of ^{13}C and ^{31}P MRS for Aging and Late-Onset Alzheimer's Disease:

Evidence suggests that metabolic dysregulation may be a primary early mechanism that leads to Late-Onset Alzheimer's Disease (*LOAD*). Although the late-stage aberrant build-up of proteins (*amyloid and tau*) have become the hallmark of disease, the earliest features are oxidative stress, lipid peroxidation, and decreased glucose utilization as many as 40 years earlier than disease onset(Farkas et al., 1982; Friedland et al., 1983; Praticò et al., 2002; Praticò & Sung, 2004; Reiman et al., 2004). These factors point towards impairments in the TCA Cycle metabolic pathway, as is seen in Alzheimer's Disease, and provide insight as to why aging and the APOE ε4 allele remain the greatest risk factors for disease onset(Bailey et al., 2015; Demetrius et al., 2015; Dong & Brewer, 2019; L. Liu et al., 2017; Monsell et al., 2015; Williams et al., 2020).

In the healthy young brain, the production of energy (*ATP*) is primarily facilitated by the synchronous metabolic activity of the TCA Cycle in neurons and Aerobic Glycolysis in

astrocytes (*Astrocyte-Neuron Lactate Shuttle*) (Pellerin & Magistretti, 1994). In both aging and carriers of the APOE ε4 allele, this synchrony deteriorates as neurons appear to become independent of astrocytes, glucose utilization declines, and the rate of neuronal TCA cycle ($V_{TCA\text{-}Neuron}$) decreases as much as 25% in the human(Boumezbeur et al., 2010; Jagust & Landau, 2012). Studies in APOE have shown that the ε4 allele is a partial loss-of-function allele limiting neurons' ability to transport excess lipids and thus vulnerable to damage from peroxidated lipids(Bailey et al., 2015; L. Liu et al., 2017). Extrapolating from these suggests that disease prognosis would be exacerbated by such an interaction in LOAD, as decreased TCA cycle leads to increased concentration of neuronal lipids.

Ultimately, both aging and ε4 allele carriers have independently been shown to have a shift in metabolic identity which implicates NAD^+ concentrations and the SIRT1-7 (*Sirtuin*) epigenetic regulators. While NAD^+ is primarily known for its role as an electron acceptor in oxidative phosphorylation, it is highly versatile and also serves as a cofactor in metabolic enzymes that promote balance between glycolysis and TCA Cycle(Guarente, 2000). Sirtuins are NAD^+-dependent, thus nutrient sensing, epigenetic regulators that well-known for being master regulators of metabolism(Haigis & Sinclair, 2010; Houtkooper et al., 2012). Across aging, both NAD^+ and Sirtuin activity are known to decline and APOE further reduces Sirtuin activity(Guarente, 2011; Massudi, Grant, Braidy, et al., 2012; Sinclair & Guarente, 2006; Theendakara et al., 2013). This has prompted the use of Caloric Restriction (*CR*), a metabolic intervention believed to be a conserved survival mechanism and has been shown to be beneficial in aging and AD models by increasing $V_{TCA\text{-}Neuron}$, NAD^+ concentrations, and Sirtuin activity(Anderson & Weindruch, 2010; Lin et al., 2014; Qin et al., 2006; Sinclair, 2005).

In LOAD, how $V_{TCA\text{-}Neuron}$ is affected, what prompts the loss of identity, and the effectiveness of intervention studies are unknown. All of which have been difficult to evaluate without a definitive *in-vivo* outcome measure for metabolism. Non-invasive *in-vivo* NMR spectroscopy (*^{13}C and ^{31}P*) has the potential to serve as an early biomarker, as it is long proven capable of measuring $V_{TCA\text{-}Neuron}$ as well as NAD^+ in the human brain (Boumezbeur et al., 2010; Lu et al., 2016; G. F. Mason et al., 1995; Xin et al., 2018). *Using Knock-In APOE3$^{+/+}$ and APOE4$^{+/+}$ mice, we aim to characterize the interaction between aging and APOE allele on $V_{TCA\text{-}Neuron}$, determine the impact of NAD^+ and its predictive value for Sirtuin activity, and evaluate the efficacy of 6-months of Caloric Restriction (CR) vs Ad Libitum (AL).* Notably, these methods are directly translatable to the human and studies of LOAD detection and have the potential to monitor disease trajectory over time.

5.3.1.1.: Decreased $V_{TCA\text{-}Neuron}$ drives LOAD through an interaction between aging and APOE ε4 allele:

Previous work has demonstrated that aging subjects have ~25% reductions in $V_{TCA\text{-}Neuron}$ and that ε4 carriers are particularly vulnerable due to diminished ability to transport lipids (Bailey et al., 2015; Boumezbeur et al., 2010; L. Liu et al., 2017). Using *in-vivo* ^{1}H-^{13}C NMR spectroscopy, we aim to characterize $V_{TCA\text{-}Neuron}$ to determine allele variant effects at 3 aging timepoints young (*6-months*), middle-aged (*12-months*), and old-age (*18-months*). Upon completion of the final (*18-month*) timepoint, we will evaluate lipid droplets in brain tissue to determine if $V_{TCA\text{-}Neuron}$ is reflective of increased lipid deposition and vulnerability to lipid peroxidation. Additionally, we will also evaluate glutathione concentrations to determine potential for decreased antioxidant protection.

5.3.1.2.: Aim 2: Decreases in NAD^+ impact $V_{TCA\text{-}Neuron}$ and are predictive of Sirtuin activity in LOAD:

NAD^+ levels have been shown to decrease across aging, having detrimental effects on health(Massudi, Grant, Braidy, et al., 2012; Massudi, Grant, Guillemin, et al., 2012). Sirtuins are NAD^+-dependent epigenetic regulators that decline across age and have decreased activity in APOE4 relative to APOE3 carriers (Guarente, 2011; Sinclair & Guarente, 2006; Theendakara et al., 2013) . Using ^{1}H-^{31}P NMR spectroscopy, we will measure NAD^+ concentrations at 3 aging timepoints young (*6-months*), middle-aged (*12-months*), and old-age (*18-months*). Using the final 18-month measurement we will correlate concentrations of NAD^+ with Sirtuin activity in post-mortem hippocampal tissue to test the validity of NMR as an *in-vivo* measure.

5.3.1.3.: Aim 3. Caloric restriction will increase $V_{TCA\text{-}Neuron}$, NAD^+ concentration and Sirtuin activity:

CR has been shown to restore $V_{TCA\text{-}Neuron}$ and increase Sirtuin activity in aging and AD(Anderson & Weindruch, 2010; Lin et al., 2014; Qin et al., 2006). While some studies have used short-term, rather than life-long, caloric restriction, the effect of a mid-life intervention is not well understood yet is more relevant to clinical application in humans(Fontana et al., 2004; Hori et al., 1992; Lin et al., 2014; Walford et al., 2002). Here, we aim to determine the efficacy of a 40% decrease in calories over a 6-month period (*12-months to 18-months*) following the same protocol used in the NIA's Caloric Restricted Colony(Turturro et al., 1999). We predict that a mid-life metabolic intervention will rescue metabolic impairment and that in-vivo measures are sensitive to changes in $V_{TCA\text{-}Neuron}$ and NAD^+.

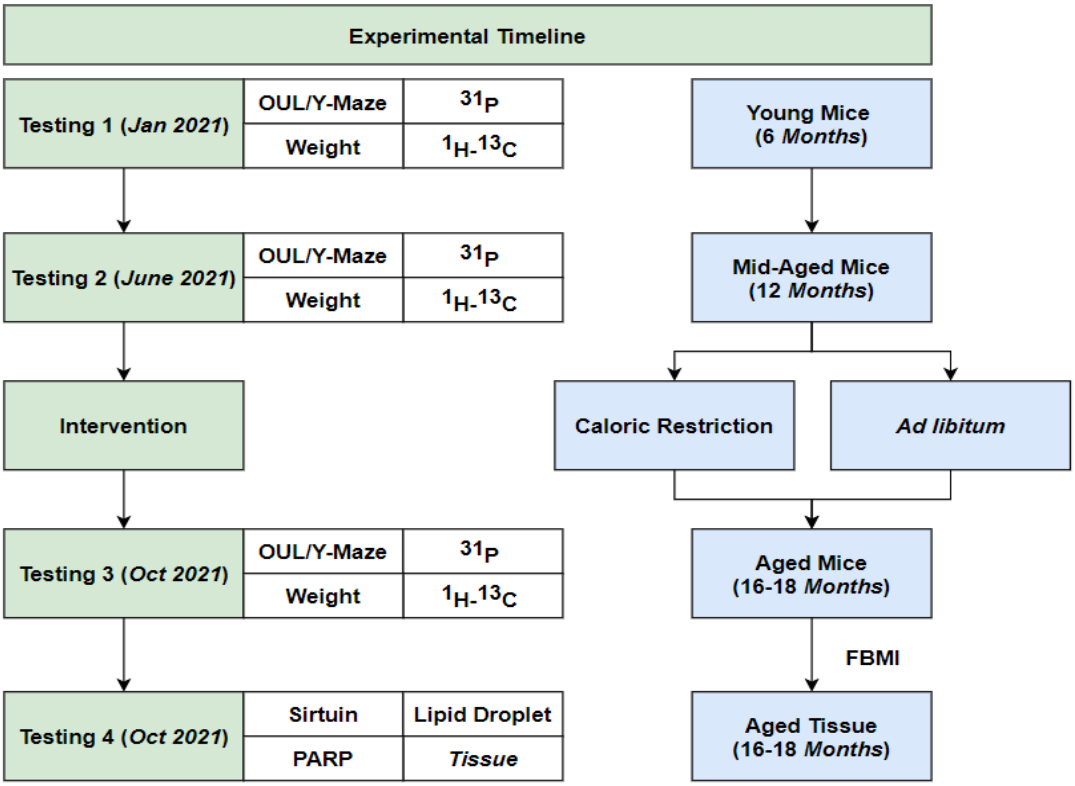


Figure 5.1.: Experimental timeline to longitudinally measure metabolism and neural energetics using *in-vivo* MRS and tissue samples.

5.4. References:

Anderson, R. M., & Weindruch, R. (2010). Metabolic reprogramming, caloric restriction and aging. *Trends in Endocrinology and Metabolism*, *21*(3), 134–141. https://doi.org/10.1016/j.tem.2009.11.005

Bailey, A. P., Koster, G., Guillermier, C., Hirst, E. M. A., MacRae, J. I., Lechene, C. P., Postle, A. D., & Gould, A. P. (2015). Antioxidant Role for Lipid Droplets in a Stem Cell Niche of Drosophila. *Cell*, *163*(2), 340–353. https://doi.org/10.1016/j.cell.2015.09.020

Boumezbeur, F., Mason, G. F., De Graaf, R. A., Behar, K. L., Cline, G. W., Shulman, G. I., Rothman, D. L., & Petersen, K. F. (2010). Altered brain mitochondrial metabolism in healthy aging as assessed by in vivo magnetic resonance spectroscopy. *Journal of Cerebral Blood Flow and Metabolism*, *30*(1), 211–221. https://doi.org/10.1038/jcbfm.2009.197

Demetrius, L. A., Magistretti, P. J., & Pellerin, L. (2015). Alzheimer's disease: The amyloid hypothesis and the Inverse Warburg effect. *Frontiers in Physiology*, *6*(JAN), 1–20. https://doi.org/10.3389/fphys.2014.00522

Dong, Y., & Brewer, G. J. (2019). Global Metabolic Shifts in Age and Alzheimer's Disease Mouse Brains Pivot at NAD+/NADH Redox Sites. *Journal of Alzheimer's Disease*, *71*(1),

119–140. https://doi.org/10.3233/JAD-190408

Farkas, T., Ferris, S. H., & Wolf, A. P. (1982). 18F-2-Deoxy-2-fluoro-D-glucose as a tracer in the positron emission tomographic study of senile dementia. *American Journal of Psychiatry*, *139*(3), 352–353. https://doi.org/10.1176/ajp.139.3.352

Fontana, L., Meyer, T. E., Klein, S., & Holloszy, J. O. (2004). Long-term calorie restriction is highly effective in reducing the risk for atherosclerosis in humans. *Proceedings of the National Academy of Sciences*, *101*(17), 6659–6663. https://doi.org/10.1073/pnas.0308291101

Friedland, R. P., Budinger, T. F., Ganz, E., Yano, Y., Mathis, C. A., Koss, B., Ober, B. A., Huesman, R. H., & Derenzo, S. E. (1983). Regional cerebral metabolic alterations in dementia of the alzheimer type: Positron emission tomography with [18f]fluorodeoxyglucose. *Journal of Computer Assisted Tomography*, *7*(4), 590–598. https://doi.org/10.1097/00004728-198308000-00003

Guarente, L. (2000). Sir2 links chromatin silencing, metabolism, and aging. *Genes and Development*, *14*(9), 1021–1026. https://doi.org/10.1101/gad.14.9.1021

Guarente, L. (2011). Sirtuins, aging, and metabolism. *Cold Spring Harbor Symposia on Quantitative Biology*, *76*, 81–90. https://doi.org/10.1101/sqb.2011.76.010629

Haigis, M. C., & Sinclair, D. A. (2010). Mammalian Sirtuins: Biological Insights and Disease Relevance. *Annual Review of Pathology: Mechanisms of Disease*, *5*(1), 253–295. https://doi.org/10.1146/annurev.pathol.4.110807.092250

Hori, N., Hirotsu, I., Davis, P. J., & Carpenter, D. O. (1992). Long-term potentiation is lost in aged rats but preserved by calorie restriction. *NeuroReport*, *3*(12), 1085–1088. https://doi.org/10.1097/00001756-199212000-00013

Houtkooper, R. H., Pirinen, E., & Auwerx, J. (2012). Sirtuins as regulators of metabolism and healthspan. *Nature Reviews Molecular Cell Biology*, *13*(4), 225–238. https://doi.org/10.1038/nrm3293

Jagust, W. J., & Landau, S. M. (2012). Apolipoprotein E, not fibrillar β-amyloid, reduces cerebral glucose metabolism in normal aging. *Journal of Neuroscience*, *32*(50), 18227–18233. https://doi.org/10.1523/JNEUROSCI.3266-12.2012

Lin, A. L., Coman, D., Jiang, L., Rothman, D. L., & Hyder, F. (2014). Caloric restriction impedes age-related decline of mitochondrial function and neuronal activity. *Journal of Cerebral Blood Flow and Metabolism*, *34*(9), 1440–1443. https://doi.org/10.1038/jcbfm.2014.114

Liu, L., MacKenzie, K. R., Putluri, N., Maletić-Savatić, M., & Bellen, H. J. (2017). The Glia-Neuron Lactate Shuttle and Elevated ROS Promote Lipid Synthesis in Neurons and Lipid Droplet Accumulation in Glia via APOE/D. *Cell Metabolism*, *26*(5), 719-737.e6. https://doi.org/10.1016/j.cmet.2017.08.024

Lu, M., Zhu, X.-H., & Chen, W. (2016). In vivo 31P MRS assessment of intracellular NAD metabolites and NAD+/NADH redox state in human brain at 4 T. *NMR in Biomedicine*, *29*(7), 1010–1017. https://doi.org/10.1002/nbm.3559

Mason, G. F., Gruetter, R., Rothman, D. L., Behar, K. L., Shulman, R. G., & Novotny, E. J. (1995). Simultaneous determination of the rates of the TCA cycle, glucose utilization, α-ketoglutarate/glutamate exchange, and glutamine synthesis in human brain by NMR. *Journal of Cerebral Blood Flow and Metabolism*, *15*(1), 12–25.

Massudi, H., Grant, R., Braidy, N., Guest, J., Farnsworth, B., & Guillemin, G. J. (2012). Age-associated changes in oxidative stress and NAD+ metabolism in human tissue. *PLoS ONE*, *7*(7). https://doi.org/10.1371/journal.pone.0042357

Massudi, H., Grant, R., Guillemin, G. J., & Braidy, N. (2012). NAD + metabolism and oxidative stress: The golden nucleotide on a crown of thorns. *Redox Report*, *17*(1), 28–46. https://doi.org/10.1179/1351000212Y.0000000001

Monsell, S. E., Kukull, W. A., Roher, A. E., Maarouf, C. L., Serrano, G., Beach, T. G., Caselli, R. J., Montine, T. J., & Reiman, E. M. (2015). Characterizing apolipoprotein E ε4 carriers and noncarriers with the clinical diagnosis of mild to moderate Alzheimer dementia and minimal β-amyloid peptide plaques. *JAMA Neurology*, *72*(10), 1124–1131. https://doi.org/10.1001/jamaneurol.2015.1721

Pellerin, L., & Magistretti, P. J. (1994). Glutamate uptake into astrocytes stimulates aerobic glycolysis: A mechanism coupling neuronal activity to glucose utilization. *Proceedings of the National Academy of Sciences of the United States of America*, *91*(22), 10625–10629. https://doi.org/10.1073/pnas.91.22.10625

Praticò, D., Clark, C. M., Liun, F., Lee, V. Y.-M., & Trojanowski, J. Q. (2002). Increase of Brain Oxidative Stress in Mild Cognitive Impairment. *Archives of Neurology*, *59*(6), 972. https://doi.org/10.1001/archneur.59.6.972

Praticò, D., & Sung, S. (2004). Lipid Peroxidation and Oxidative imbalance: Early functional events in Alzheimer's disease. *Journal of Alzheimer's Disease*, *6*(2), 171–175. https://doi.org/10.3233/JAD-2004-6209

Qin, W., Yang, T., Ho, L., Zhao, Z., Wang, J., Chen, L., Zhao, W., Thiyagarajan, M., MacGrogan, D., Rodgers, J. T., Puigserver, P., Sadoshima, J., Deng, H., Pedrini, S., Gandy, S., Sauve, A. A., & Pasinetti, G. M. (2006). Neuronal SIRT1 activation as a novel mechanism underlying the prevention of alzheimer disease amyloid neuropathology by calorie restriction. *Journal of Biological Chemistry*, *281*(31), 21745–21754. https://doi.org/10.1074/jbc.M602909200

Reiman, E. M., Chen, K., Alexander, G. E., Caselli, R. J., Bandy, D., Osborne, D., Saunders, A. M., & Hardy, J. (2004). Functional brain abnormalities in young adults at genetic risk for late-onset Alzheimer's dementia. *Proceedings of the National Academy of Sciences of the United States of America*, *101*(1), 284–289. https://doi.org/10.1073/pnas.2635903100

Sinclair, D. A. (2005). Toward a unified theory of caloric restriction and longevity regulation. *Mechanisms of Ageing and Development*, *126*(9), 987–1002. https://doi.org/10.1016/j.mad.2005.03.019

Sinclair, D. A., & Guarente, L. (2006). Unlocking the Secrets of Longevity Genes. *Scientific American*, *294*(3), 48–57. https://doi.org/10.1038/scientificamerican0306-48

Theendakara, V., Patent, A., Libeu, C. A. P., Philpot, B., Flores, S., Descamps, O., Poksay, K. S., Zhang, Q., Cailing, G., Hart, M., John, V., Rao, R. V., & Bredesen, D. E. (2013). Neuroprotective Sirtuin ratio reversed by ApoE4. *Proceedings of the National Academy of Sciences of the United States of America*, *110*(45), 18303–18308. https://doi.org/10.1073/pnas.1314145110

Turturro, A., Witt, W. W., Lewis, S., Hass, B. S., Lipman, R. D., & Hart, R. W. (1999). Growth curves and survival characteristics of the animals used in the biomarkers of aging program. *Journals of Gerontology - Series A Biological Sciences and Medical Sciences*, *54*(11), 492–501. https://doi.org/10.1093/gerona/54.11.B492

Walford, R. L., Mock, D., Verdery, R., & MacCallum, T. (2002). Calorie Restriction in Biosphere 2: Alterations in Physiologic, Hematologic, Hormonal, and Biochemical Parameters in Humans Restricted for a 2-Year Period. *The Journals of Gerontology Series A: Biological Sciences and Medical Sciences*, *57*(6), B211–B224. https://doi.org/10.1093/gerona/57.6.B211

Williams, H. C., Farmer, B. C., Piron, M. A., Walsh, A. E., Bruntz, R. C., Gentry, M. S., Sun, R. C., & Johnson, L. A. (2020). APOE alters glucose flux through central carbon pathways in astrocytes. *Neurobiology of Disease*, *136*(January), 104742. https://doi.org/10.1016/j.nbd.2020.104742

Xin, L., Ipek, Ö., Beaumont, M., Shevlyakova, M., Christinat, N., Masoodi, M., Greenberg, N., Gruetter, R., & Cuenoud, B. (2018). Nutritional Ketosis Increases NAD+/NADH Ratio in Healthy Human Brain: An in Vivo Study by 31P-MRS. *Frontiers in Nutrition*, *5*(July), 1–8. https://doi.org/10.3389/fnut.2018.00062