Using Machine Learning techniques for identification of Chronic Traumatic Encephalopathy related Spectroscopic Biomarkers

Marcia S Louis*, Michael Alosco†, Benjamin Rowland‡, Huijun Liao§, Joseph Wang*,
Inga Koerte§, Martha Shenton§, Robert Stern†, Ajay Joshi*, Alexander P Lin‡
*Department of Electrical and Computer Engineering, †Department of Neurology,
Boston University, Boston, MA
‡Center for Clinical Spectroscopy,§ Psychiatric Neuroimaging Laboratory,
Brigham and Women's Hospital, Boston, MA

***Abstract*—Contact sports athletes, military personnel, and civilians that suffer from multiple head traumas have the potential to develop Chronic Traumatic Encephalopathy (CTE), a progressive, degenerative brain disease diagnosed only post-mortem by characteristic tau deposition in the brain. There is, therefore, a need for in-vivo diagnosis for CTE to diagnose and manage this disease, while the individual is still alive. However, there is no definitive in-vivo diagnosis because of heterogeneous clinical symptoms that often overlap with other neurodegenerative diseases. Magnetic Resonance Spectroscopy (MRS) can be a suitable candidate for CTE diagnosis as multiple head trauma changes the neurochemicals in the brain that can be detected using MRS. These changes can be subtle, and group differences are not sufficient for clinical diagnosis. This paper proposes a machine learning based approach to capture the neuro-spectroscopic signatures corresponding to CTE-related impairments in NFL players. The classification model uses concentration estimates of metabolites to classify between 'Impaired and 'Non-impaired players. The model using the metabolite concentrations of creatine, choline, N-acetyl-aspartate, glutamate, and macromolecules achieved Area Under the Curve (AUC) of 0.72 and prediction accuracy of 75%. While these metabolites have been shown to be altered in previous concussion studies, other metabolites may improve the diagnostic accuracy. In order to include more metabolites, two-dimensional correlated spectroscopy (L-COSY), which resolves overlapping metabolites, was also acquired. The L-COSY model which included 15 metabolites, increased prediction accuracy to 87% with AUC of 0.83. With the aid of machine learning, these metabolites may serve as potential biomarkers that correspond to the CTE-related impairments that will allow for CTE diagnostics in athletes prior to their death.**

***Index Terms*—Chronic Traumatic Encephalopathy; Machine Learning; Proton Magnetic Resonance Spectroscopy;**

'

I. Introduction

Individuals in contact sports, like American football experience long-term neurological difficulties from head injuries. Repetitive concussive and subconcussive brain trauma can lead to long-term neurological problems such as Chronic traumatic encephalopathy (CTE). CTE is a neurodegenerative disease that can occur in individuals with histories of repetitive brain trauma [1]. Early symptoms include mood disturbance, behavior change, depression, aggression and impaired memory. Dementia as well as motor disturbances can develop in some people [2]. Autopsy studies of the individuals with these symptoms are shown to have a distinct pattern of neuropathological changes known as tau-immunoreactive proteins in the cerebral cortex of the brain [3].

Even though, the neuropathological diagnosis of CTE is well documented there is no definitive in-vivo diagnosis, because of the heterogeneous clinical symptoms that often overlap with other neurodegenerative diseases. An in-vivo diagnosis is needed to provide an in life and personalized approach to manage the condition. Magnetic resonance spectroscopy (MRS) is a specialized noninvasive technique to record neurochemicals in the brain. MRS can be a suitable candidate for CTE diagnosis as studies have shown neurochemical changes in the brain as a result of repetitive brain trauma [4].

One-dimensional MRS allows metabolites on the MR timescale to be monitored noninvasively in a clinical MRI scanner [5]. However, many of the metabolites overlap in the one- dimensional MR spectrum making it difficult to determine the precise level of change in each chemical population. Two-dimensional MRS resolves overlapping metabolites [6], thus providing chemical specificity not available to one-dimensional MRS. However, the changes in metabolite concentration from one-dimensional and two-dimensional MRS can be subtle, and group differences are not sufficient for clinical diagnosis.

In this paper, we model the identification of CTE related biomarkers as a supervised machine learning problem. Classifiers are designed to learn pattern from data and classify them to one of a known set of classes. We use metabolite concentrations from one- dimensional and two-dimensional MRS data as features for the classification model. The classification model achieved an accuracy of 75% for one-dimensional MRS data. With the two- dimensional MRS data the classification accuracy increased to 87%. Therefore, with the aid of machine learning, it is possible to use a multivariate approach to find

978-1-5386-1235-4/17/$31.00 ©2017 IEEE

differences that correspond to the symptomatic changes that have been exhibited in athletes prior to their death.

II. Materials and Methods

A. Subjects

The subjects for this study are from the National Institutes of Health (NIH)-funded "Diagnosing and Evaluating Traumatic Encephalopathy using Clinical Tests" (DETECT) study [7]. Inclusion criteria are ages 40-69 years and professional NFL players with minimum of 2 years of active playing time. The study consist of 65 NFL retired professional male athletes (55.9$\pm$7.8 years). They have an active playing time in NFL of 3 to 14 years (7.9$\pm$2.4 years). Each athlete had experienced repetitive concussive and sub-concussive trauma during his career. They have reported clinical symptoms associated with either prolonged post-concussive syndrome or early CTE, including headaches, confusion, impaired judgment, memory loss, impulse control problems, aggression and depression [2].

B. MRS Data Acquisition

All athletes underwent proton MRS scan on a 3T MRI scanner (Siemens Verio, vB17) using a 32 channel head coil. Routine axial 3D-MPRAGE was acquired and reconstructed in three planes to localize the regions of interest to the posterior cingulate gyrus. This region was selected due to high data quality and as studies have shown it to be sensitive to brain injury [4] and changes in CTE [8]–[10]. For one-dimensional MRS we use short- echo point-resolved (PRESS) single voxel spectroscopy pulse sequence in the posterior cingulate gyrus of the brain (TE = 30 ms, TR = 2 s, 128 averages, 8 cc volume). We acquired two-dimensional MRS using single voxel Localized Correlated Spectroscopy (L-COSY) pulse sequence in the posterior cingulate gyrus of the brain (TE = 30 ms, TR = 2.5 s, 64 increments of 0.8 ms, 8 averages 27 cc volume, acquired vector size of 2048 points and acquisition time of 512 ms thus providing a spectral width of 2000 Hz and 1250 Hz in F2 and F1, respectively) [4], [6], [11]. 58 out of 65 athletes underwent the L-COSY scan.

C. MRS Data Processing

The raw PRESS data of each athlete are frequency corrected and phase aligned using Suspect [12], a Python package for processing Magnetic Resonance spectroscopy data. The residual water peak is suppressed using the reference water signal acquired along with the PRESS data. The data is then processed using LC model [9] software. The LC model software estimates the concentration of metabolites by fitting a baseline spectrum on the Fourier transformed data spectrum and measure the area under the peaks. The model provides Cramer-Rao lower bound values as a metric for the goodness of fit. The total number of metabolite concentrations estimated is 35, out of which only 9 metabolites had Cramer-Rao lower bounds of less than 20%. We used concentration estimates of these 9 metabolites as features for the classification algorithm.

Similarly, raw L-COSY data were processed using commercial available, two-dimensional spectral processing software, Felix (Accelrys, San Diego, CA, USA). The residual water is suppressed to a certain extent by the built-in Gaussian shaped convolution-based method. The Felix software processing parameters are F2 domain (skewed sine-squared window, 2,048 points, magnitude mode), F1 domain (sine-squared window, linear prediction to 96 points, zero-filling to 512 points, magnitude mode) [4]. The area under the cross peaks and diagonal resonances peaks were measured using the software, and the volume of each peak was normalized to the diagonal resonance peak of Creatine at 3.02 ppm for comparable results across all the scans. In total 62 cross peaks and diagonal resonances peaks concentrations were measured.

D. Neuropsychological Assessment

As part of the DETECT study, all athletes underwent a neuropsychological test battery for evaluating attention, visual and verbal memory, psychomotor speed and executive function. The behavioral and mood changes of these athletes are assessed via personal interview, semi-structured interviews and standardized self-report. The raw scores of these tests were standardized to Z-score values as described in previous studies in a larger cohort from which these subjects were drawn [13]. The Z-score values correspond to four-factors: 1) behavior/mood, 2) psychomotor speed/executive function, 3) verbal memory, and 4) visual memory. A Z-scores below -1 for Factors 2,3, and 4 or above +1 for Factor 1 were considered to have impairments. For this study, athletes with signs of one or more impairments based on the four-factors were labeled as "Impaired class". The remaining athletes formed "Non-Impaired class". This criterion yielded 32 "Impaired" athletes and 33 "Non-Impaired" athletes.

III. Classifier Design and Predictive Analysis

In this paper, we design two classification models. The first model is using data from PRESS sequence and the second model using L-COSY sequence. The labels for supervised classification model are deduced from Neuropsychological assessment. The dataset is randomly split into 70% training dataset and 30% testing dataset.

Since the sample size is limited, feature selection is critical for the classification model to be accurate. The distribution of PRESS data is shown in Figure 1, metabolites glycerophosphorylcholine (GPC), glycerophosphorylcholine+phosphocholine (GPC+PCh), and createine+phosphocreatine (Cr+PCr), have different distribution for 'Non-Impaired' and 'Impaired' class. However, GPC, GPC+PCh, macromolecules at 2.0 ppm (MM20), macromolecules and lipid at 2.0 ppm (MM20+Lip20) have similar distributions and might have similar information. The similarity is further confirmed by univariate analysis using Analysis of Variance algorithm. From Figure 2 GPC, GPC+PCh and MM20, MM20+Lip20 have similar variance scores for classification. Also, total creatine have high scores because of their high variance and an important feature for classification. The best number of features were selected using ANOVA and Linear Support Vector Machine Classifier. Figure 3 shows that using 60-80% of the feature has

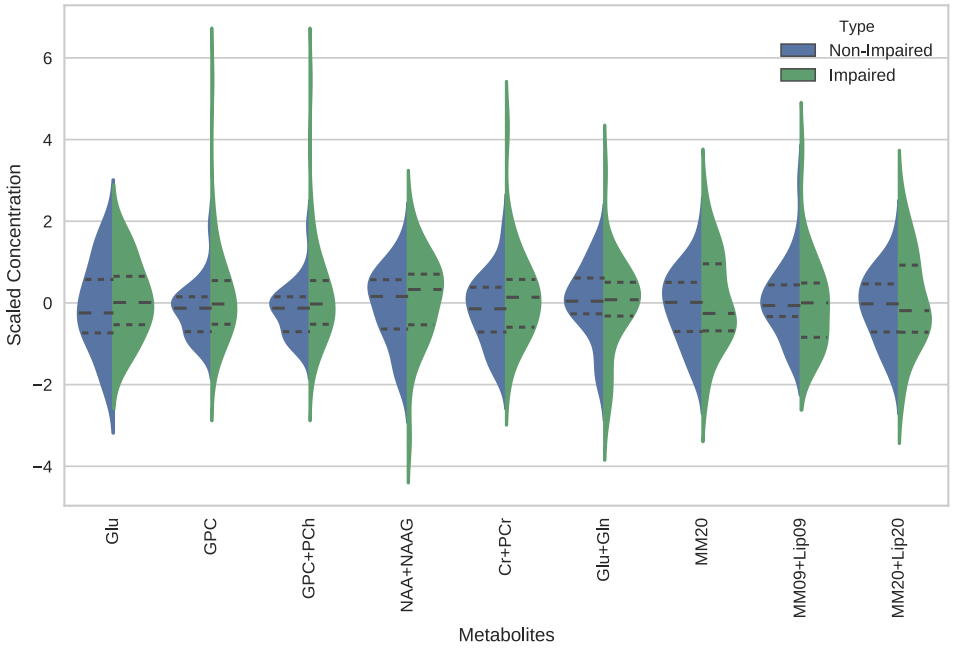

Fig. 1. Distribution of the metabolites concentration values grouped by the class. Dotted line show the quartiles of the distribution

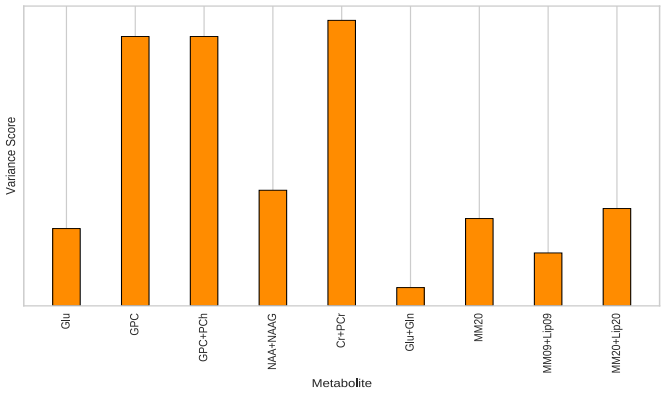

Fig. 2. Univariate feature analysis using ANOVA algorithm

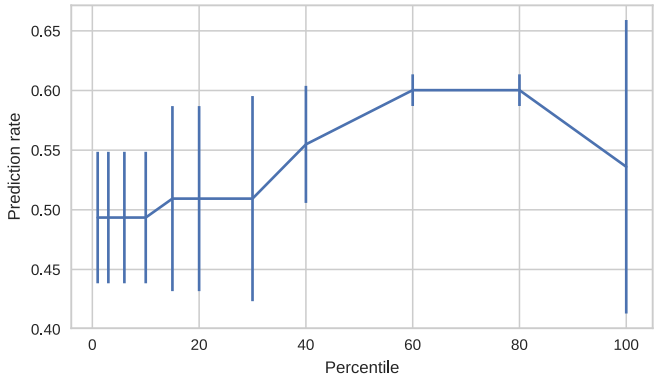

Fig. 3. Performance of the SVM-Anova varying the percentile of features selected for 1D PRESS MRS

high mean classification accuracy and low standard deviation. Following similar analysis for L-COSY data using 20% of metabolites has low standard deviation shown in Figure 4. Therefore, from the analysis 6 metabolites were selected for PRESS data and 15 metabolites were selected for L-COSY data to train the classification model.

IV. Results

A. Model Selection

In our analysis, we considered three type of supervised machine learning algorithms for classification. Support Vector Machine with radial basis function (SVM-RBF), it uses kernel trick for non-linearly separable feature and support vectors for classification. The K-Nearest Neighbors (KNN) applies k neighbors class information for making a classification. Lastly, Random forest (RF) are ensemble classifiers that have multiple decision trees as weak learners. Each decision tree is trained on a randomly selected sub-feature set and uses majority voting for classification. We used three dimensionality reduction techniques to analyze the possibility of reducing the number of features further. Principal Component Analysis (PCA) computes linear component of the data using Singular Value Decomposition (SVD), while Kernel PCA (KPCA) uses kernel trick and SVD to project data to a lower dimensional. Isometric mapping (ISOMAP) is a manifold learning technique that projects data points into a lower dimension transform space by preserving the distance between points in the original space. Additionally, we used Analysis of Variance to understand the

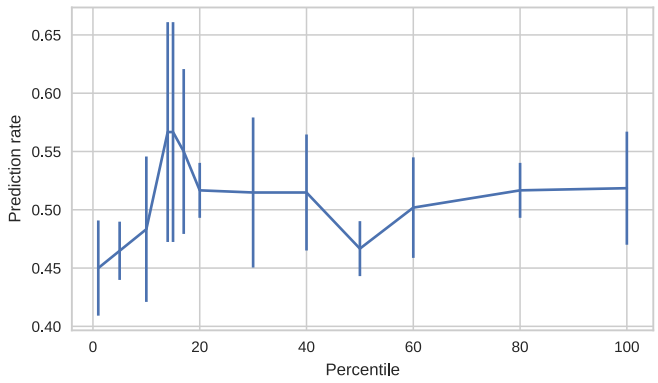

Fig. 4. Performance of the SVM-Anova varying the percentile of features selected for 2D LCOSY MRS

importance of each feature. For all our analysis the dataset was divided randomly into 70% training set and 30% testing set.

TABLE I
MEAN ACCURACY AND STANDARD DEVIATION OF 5-FOLD CROSS-VALIDATION ON TRAINING SET

	PCA	KPCA	ISOMAP	ANOVA
PRESS MRS				
SVM-RBF	0.65±0.075	0.64±0.059	0.67±0.075	0.66±0.075
RF	0.72±0.075	0.71±0.075	0.70±0.075	0.73±0.074
KNN	0.71±0.064	0.72±0.593	0.73±0.060	0.71±0.054
LCOSY MRS				
SVM-RBF	0.75±0.099	0.75±0.088	0.72±0.104	0.75±0.099
RF	0.73±0.093	0.72±0.074	0.75±0.105	0.77± 0.079
KNN	0.77±0.100	0.70±0.079	0.69±0.074	0.77±0.074

TABLE II
EVALUATION RESULTS OF CLASSIFICATION MODELS ON TESTING SET

Algorithm	Accuarcy	F1 Score	AUC
PRESS MRS			
Random Classifier	45%	0.44	0.5
SVM-RBF	60%	0.59	0.65
RF	70%	0.71	0.67
KNN	75%	0.74	0.72
LCOSY MRS			
Random Classifier	61%	0.60	0.5
SVM-RBF	38%	0.34	0.28
RF	61%	0.60	0.68
KNN	87%	0.89	0.83

We performed a hyperparameters grid search for the 12 models using the training set. The 5-fold Cross Validation results of the grid search are given in Table I. Overall KNN classifier has higher classification accuracy for PRESS and L-COSY datasets. All the 12 models are evaluated on the testing set, and the three metrics used for evaluation are accuracy, F1 Score and Area Under the Curve (AUC) of the receiver operating characteristic curve. The results are given in Table II; we the Random Classifier as the baseline classifier. The Random Classifier randomly assigns a label to each sample in the testing set based on a uniform distribution. SVM-RBF classifier performs poorly for L-COSY dataset because of over-fitting in-spite of using 5 fold cross-validation while training the classification model. This is due to the limited number of training samples, which can be addressed by using a larger cohort. The best configuration for PRESS data based on Table II is ISOMAP with 4 components for dimensionality reduction and KNN with 3 neighbors for classification. The best configuration for L-COSY data is with 9 features selected using ANOVA feature scores and KNN with 19 neighbors for classification.

B. Association of MRS Biomarkers with CTE Impairments

Based on the PRESS dataset the features relevant for classification were MM20+Lip20, Cr+PCr, GPC+PCh, macromolecules and lipids at 0.9 ppm (MM09+Lip09), N-acetylasparte+N-acetylaspartylglutamate (NAA+NAAG), and glutamate (Glu) (Table III. NAA+NAAG as a biomarker of neuronal health, Cr+PCr as a biomarker for brain energetics, GPC+PCh as a marker of diffuse axonal injury, and glutamate as a marker of excitotoxicity have been previously identified in other studies of traumatic brain injury [14] however identification of macromolecules is novel and may be a result of underlying large protein changes that are not as visible on the MR timescale.

TABLE III
METABOLITES OF INTEREST FROM PRESS DATA

	Non-Impaired				Impaired			
Metabolite	max	min	mean	std	max	min	mean	std
Cr+PCr	6.518	4.727	5.506	0.419	7.749	4.678	5.708	0.565
GPC+PCr	1.371	0.824	1.007	0.111	1.999	0.756	1.075	0.221
Glu	7.583	5.258	6.427	0.616	7.584	5.527	6.514	0.538
MM09+Lip09	11.859	6.258	8.0906	1.170	12.932	6.066	7.947	1.445
MM20+Lip20	14.439	8.059	11.267	1.522	16.068	7.321	11.597	2.077
NAA+NAAG	8.710	6.373	7.629	0.601	9.108	5.524	7.767	0.722

Metabolites of interest from L-COSY dataset (Table IV) are: glutathione (GSH), imidizole (Imi) (Pearson correlation (r) $\geq$ 0.8 with uridine diphosphate glucose (UPDG-1, UPDG-2), UPDG-1 (r $\geq$ 0.8 glutamate+glutamine, NAA, GPC+PCh, ethanolamine (Eth), GPC), 'UPDG-2' (r $\geq$ 0.8 GPC-2), Glx, myoinositol, isocitrate (Icit), isoleucine (Ile), and phenylalanine (Phe). Glutathione is an antioxidant implicated in neuroinflammation and myoinositol have been shown to be altered in sports-related head injury [15]. Glx and Phe changes have also been implicated in contact sports athletes [4]. However the other metabolites identified on the 2D L-COSY have not been previously observed in other studies. UPDG is involved in glycogen metabolism in the brain [16] which may be disrupted by brain injury and the amino acids such as Imi, Ile, and ICit may also be affected by injury which may provide interesting insight into potential treatments for injury [17].

TABLE IV
METABOLITES OF INTEREST FROM L-COSY DATA (NORMALIZED TO CREATINE PEAK AT 3.02 PPM)

	Non-Impaired				Impaired			
Metabolite	max	min	mean	std	max	min	mean	std
GSH_3	0.162	0.046	0.079	0.031	0.203	0.047	0.100	0.041
Glx_1	0.033	0.006	0.020	0.006	0.186	0.010	0.029	0.032
Icit	0.114	0.018	0.052	0.027	0.183	0.014	0.068	0.039
Ile	0.047	0.007	0.018	0.009	0.067	0.009	0.023	0.013
Imi-1	0.253	0.114	0.158	0.043	0.349	0.128	0.184	0.059
Phe	0.019	0.003	0.010	0.003	0.021	0.005	0.011	0.003
UPDG_1	0.163	0.037	0.068	0.040	0.681	0.040	0.113	0.134
UPDG_2	0.029	0.003	0.010	0.006	0.054	0.005	0.017	0.013
Myoinositol	0.101	0.010	0.035	0.023	0.162	0.015	0.048	0.037

V. CONCLUSION

While NAA, creatine, choline, glutamate, and myoinositol have been described to be changed in other concussion studies [4], it is of great interest that the macromolecules and other metabolites had high relevance for classification. This would imply that changes in this area of the spectrum may have hidden diagnostic value for CTE-related research and is further explored using methods such as 2D correlated spectroscopy. The improvement in classifier accuracy with L-COSY, which provides additional metabolite measures, supports this additional finding. It is also evident from the analysis that the neurochemical changes in brain corresponds with the neuropsychological test and clinical evaluation. This reflects the heterogeneity of changes found in concussion studies and demonstrates the strong value of machine learning methods to evaluate changes in CTE. The current multicenter study (DIAGNOSE-CTE) will provide a much larger cohort and additional insight into the MRS biomarks for CTE.

ACKNOWLEDGMENTS:

This work was supported by grants from the National Institutes of Health (NIH; P30 AG13846; R01NS078337; R56 9500304025; U01NS093334; 1U01NS086659-01) as well as the joint Boston University-Brigham and Women's Hospital Fellowship.

REFERENCES

[1] B. E. Gavett, R. C. Cantu, M. Shenton, A. P. Lin, C. J. Nowinski, A. C. McKee, and R. A. Stern, "Clinical appraisal of chronic traumatic encephalopathy: current perspectives and future directions," *Current opinion in neurology*, vol. 24, no. 6, pp. 525–531, 2011.

[2] R. A. Stern, D. H. Daneshvar, C. M. Baugh, D. R. Seichepine, P. H. Montenigro, D. O. Riley, N. G. Fritts, J. M. Stamm, C. A. Robbins, L. McHale *et al.*, "Clinical presentation of chronic traumatic encephalopathy," *Neurology*, vol. 81, no. 13, pp. 1122–1129, 2013.

[3] A. C. McKee, T. D. Stein, C. J. Nowinski, R. A. Stern, D. H. Daneshvar, V. E. Alvarez, H.-S. Lee, G. Hall, S. M. Wojtowicz, C. M. Baugh *et al.*, "The spectrum of disease in chronic traumatic encephalopathy," *Brain*, vol. 136, no. 1, pp. 43–64, 2013.

[4] A. P. Lin, S. Ramadan, R. A. Stern, H. C. Box, C. J. Nowinski, B. D. Ross, and C. E. Mountford, "Changes in the neurochemistry of athletes with repetitive brain trauma: preliminary results using localized correlated spectroscopy," *Alzheimer's research & therapy*, vol. 7, no. 1, p. 13, 2015.

[5] C. E. Mountford, P. Stanwell, A. Lin, S. Ramadan, and B. Ross, "Neurospectroscopy: the past, present and future," *Chemical reviews*, vol. 110, no. 5, pp. 3060–3086, 2010.

[6] M. A. Thomas, K. Yue, N. Binesh, P. Davanzo, A. Kumar, B. Siegel, M. Frye, J. Curran, R. Lufkin, P. Martin *et al.*, "Localized two-dimensional shift correlated mr spectroscopy of human brain," *Magnetic Resonance in Medicine*, vol. 46, no. 1, pp. 58–67, 2001.

[7] I. K. Koerte, J. Hufschmidt, M. Muehlmann, Y. Tripodis, J. M. Stamm, O. Pasternak, M. Y. Giwerc, M. J. Coleman, C. M. Baugh, N. G. Fritts *et al.*, "Cavum septi pellucidi in symptomatic former professional football players," *Journal of neurotrauma*, vol. 33, no. 4, pp. 346–353, 2016.

[8] S. Ashwal, B. A. Holshouser, S. K. Shu, P. L. Simmons, R. M. Perkin, L. G. Tomasi, D. S. Knierim, C. Sheridan, K. Craig, G. H. Andrews *et al.*, "Predictive value of proton magnetic resonance spectroscopy in pediatric closed head injury," *Pediatric neurology*, vol. 23, no. 2, pp. 114–125, 2000.

[9] A. C. McKee, R. C. Cantu, C. J. Nowinski, E. T. Hedley-Whyte, B. E. Gavett, A. E. Budson, V. E. Santini, H.-S. Lee, C. A. Kubilus, and R. A. Stern, "Chronic traumatic encephalopathy in athletes: progressive tauopathy after repetitive head injury," *Journal of Neuropathology & Experimental Neurology*, vol. 68, no. 7, pp. 709–735, 2009.

[10] B. D. Ross, T. Ernst, R. Kreis, L. J. Haseler, S. Bayer, E. Danielsen, S. Blüml, T. Shonk, J. C. Mandigo, W. Caton *et al.*, "1h mrs in acute traumatic brain injury," *Journal of Magnetic Resonance Imaging*, vol. 8, no. 4, pp. 829–840, 1998.

[11] S. Ramadan, O. C. Andronesi, P. Stanwell, A. P. Lin, A. G. Sorensen, and C. E. Mountford, "Use of in vivo two-dimensional mr spectroscopy to compare the biochemistry of the human brain to that of glioblastoma," *Radiology*, vol. 259, no. 2, pp. 540–549, 2011.

[12] I. J. L. A. Rowland B, Mariano LJ, "An open-source software repository for magnetic resonance spectroscopy data analysis tools," *International Society for Magnetic Resonance in Medicine MR Spectroscopy Workshop*, 2016.

[13] M. L. Alosco, J. Jarnagin, Y. Tripodis, M. Platt, B. Martin, C. E. Chaisson, C. M. Baugh, N. G. Fritts, R. C. Cantu, and R. A. Stern, "Olfactory function and associated clinical correlates in former national football league players," *Journal of neurotrauma*, vol. 34, no. 4, pp. 772–780, 2017.

[14] T. S. Ng, A. P. Lin, I. K. Koerte, O. Pasternak, H. Liao, S. Merugumala, S. Bouix, and M. E. Shenton, "Neuroimaging in repetitive brain trauma," *Alzheimer's research & therapy*, vol. 6, no. 1, p. 10, 2014.

[15] I. K. Koerte, A. P. Lin, M. Muehlmann, S. Merugumala, H. Liao, T. Starr, D. Kaufmann, M. Mayinger, D. Steffinger, B. Fisch *et al.*, "Altered neurochemistry in former professional soccer players without a history of concussion," *Journal of neurotrauma*, vol. 32, no. 17, pp. 1287–1293, 2015.

[16] M. M. Adeva-Andany, M. González-Lucán, C. Donapetry-García, C. Fernández-Fernández, and E. Ameneiros-Rodríguez, "Glycogen metabolism in humans," *BBA clinical*, vol. 5, pp. 85–100, 2016.

[17] J. A. Elkind, M. M. Lim, B. N. Johnson, C. P. Palmer, B. J. Putnam, M. P. Kirschen, and A. S. Cohen, "Efficacy, dosage, and duration of action of branched chain amino acid therapy for traumatic brain injury," *Frontiers in neurology*, vol. 6, 2015.