6-2021

Machine Learning Classification of Traumatic Brain Injury Patients Versus Healthy Controls Using Arterial Spin Labeled Perfusion MRI

Vanessa I. Grass
The Graduate Center, City University of New York

MACHINE LEARNING CLASSIFICATION OF TRAUMATIC BRAIN INJURY PATIENTS VERSUS HEALTHY CONTROLS USING ARTERIAL SPIN LABELED PERFUSION MRI

by

VANESSA GRASS

A master's thesis submitted to the Graduate Faculty in Cognitive Neuroscience in partial fulfillment of the requirements for the degree of Master of Science, The City University of New York

2021

Machine Learning Classification of Traumatic Brain Injury Patients versus Healthy Controls using Arterial Spin Labeled Perfusion MRI

by

Vanessa Grass

This manuscript has been read and accepted for the Graduate Faculty in Cognitive Neuroscience in satisfaction of the thesis requirement for the degree of Master of Science.

Date	Junghoon Kim Thesis Advisor
Date	Valerie Shafer Associate Director/Second Reader

THE CITY UNIVERSITY OF NEW YORK

ABSTRACT

Machine Learning Classification of Traumatic Brain Injury Patients versus Healthy Controls using Arterial Spin Labeled Perfusion MRI

by

Vanessa Grass

Advisor: Junghoon Kim, Ph.D.

Traumatic brain injury (TBI) is one of the most common causes of death and disability worldwide, yet accurate *in vivo* detection of TBI neuropathology remains challenging due to complexities in the structural and functional changes observed post-injury as well as limitations in conventional neuroimaging modalities. Although advanced neuroimaging techniques such as arterial spin labeling (ASL) can noninvasively assess cerebral blood flow (CBF) changes observed post-injury, this technique is underutilized in TBI research partly due to the low signal-to-noise-ratio (SNR) inherent in ASL imaging. The aim of the current study is to examine the use of machine learning, specifically a Support Vector Machine (SVM) classifier, in discriminating between healthy controls (n=35) and TBI patients (n=42) using ASL-generated CBF data 3 months post-injury. Identification of the regions of interest (ROIs) most predictive of TBI is also explored as part of this aim. Furthermore, several ASL outlier cleaning methods, such as the Structural Correlation-Based Outlier REjection (SCORE) and prior-guided, slice-wise adaptive outlier cleaning (PAOCSL) algorithms, are examined in relation to improving the SNR and SVM performance. While the classification models tested did not reach statistically significant performance levels, the results were in the direction suggesting that more sophisticated outlier cleaning methods can improve classification accuracy. Potential explanations of the observed low classification accuracy and the implications of our findings on future research are discussed.

ACKNOWLEDGMENTS

I would like to express my gratitude and appreciation for the guidance and support of my mentor and advisor, Dr. Kim. The current study builds upon Dr. Kim's previous impressive work in the field of TBI and I am honored to have had the opportunity to contribute to this area of impactful research. I also want to thank Daniel Brennan, a Ph.D. student in Dr. Kim's lab, who offered indispensable knowledge and help in executing the machine learning and analysis component of my research. Additionally, I am grateful for the expertise and feedback provided by Naomi Gaggi, another Ph.D. student in the lab, who shared essential information as it relates to important experimentation in the field of TBI. Working on this research has been an immense learning experience and I am very thankful for all your help.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

1. Introduction

1.1. Traumatic brain injury: background and pathology

Critical injuries to the brain, which are broadly referred to as traumatic brain injuries (TBIs) in the field of neuroscience, present very serious medical, clinical, and research implications given they are a major cause of death and disability across the world (Humphreys et al., 2013; Hyder et al., 2007; Langlois et al., 2006; Maas et al., 2017). Most TBIs are the result of falls or road traffic accidents, encompassing injuries in which the skull and dura are breached, causing direct damage to the brain or injuries in which the skull and dura remain intact but there is internal damage to the brain (Blennow et al., 2016). TBIs can affect different areas of the brain and the severity can range anywhere from mild, moderate, and/or severe based on several criteria such as injury type, clinical presentation, and neuroimaging results (Blennow et al., 2016). Past research has shown that TBI severity is directly correlated with adverse effects such as cognitive impairment, as well as increased risk of dementia (Blennow et al., 2016). Furthermore, the anatomical and functional changes resulting from TBI could be contributing factors to premature aging of the brain (Cole et al., 2015). However, memory and cognitive deterioration caused by TBI may not always be overt or directly obvious (Langlois et al., 2006), furthering the need for the investigation of reliable biomarkers of TBI identification and mechanistic understanding (Ware et al., 2017), as well as strategies to predict, and ultimately improve, patient outcomes.

Fundamentally, the major complication of TBI is altered brain functioning. Within the first milliseconds of the trauma, during the primary injury phase, tissue damage occurs caused by stretching, straining, and shearing due to acceleration or deceleration forces external to the head (de Rooij & Kuhl, 2018). Minutes to days post-trauma, in the secondary phase of injury, compounding biochemical mechanisms continue to induce further tissue damage (de Rooij & Kuhl, 2018).

Months to years post-trauma, the primary and secondary injuries may lead to structural and functional changes in the brain and a further cascade of neurodegeneration (de Rooij & Kuhl, 2018), made evident by gray matter and white matter atrophy (Harris et al., 2019). In particular, some prominent pathological features of TBI include white matter structural injury, insufficient blood supply to the brain (ischemia), and/or brain inflammation, each with their own unique and complex neuropathology (Blennow et al., 2016).

1.2. Post-traumatic structural and functional changes

1.2.1. Diffuse axonal injury

Although white matter structural injury is not the primary focus of this research, it's important to note that structural changes, for instance, diffuse axonal injury (DAI), are usually intermingled with other neurophysiological changes, such as changes in cerebral blood flow (CBF) (Wang & Li, 2016), which is the focus of this current study. Nevertheless, DAI is understood to be present in most TBI patients that have lost consciousness (Meythaler et al., 2001): loss of consciousness generally demonstrates a TBI of at least moderate severity. DAI targets the white matter tracts of the brain, disrupting the neurofilament subunits within the axonal cytoskeleton, resulting in axonal disintegration and ultimately axonal degeneration (Meythaler et al., 2001).

The complicated nature of TBI-related brain damage makes exact measurement of trauma challenging (Kim et al., 2010), even though neuroimaging structural changes post-TBI is promising. A relevant observation is that computed tomography (CT) or magnetic resonance imaging (MRI) often only shows minimal changes in DAI in TBI patients (Meythaler et al., 2001). More recent advances in neuroimaging to better detect white matter injury include diffusion tensor imaging (DTI), however, even with improved imaging techniques, there is a significant amount of variance across TBI patients as it relates to DAI (Ware et al., 2017). In addition, the DTI technique itself is inherently limited in its scope of white matter imaging capabilities, missing the nuances of

damage to axonal membranes, myelin sheath, and other axonal microstructures (Choi et al., 2019). A new imaging method, myelin water imaging (MWI), examines myelin loss in TBI patients, which may affect neuronal signaling, cognitive function, and suggests TBI causes demyelination of white matter (Choi et al., 2019).

1.2.2. Cerebral blood flow

While DAI has been the focus of a myriad of past TBI studies (Graves & Kreipke, 2015), CBF is an emerging and very relevant biomarker of cognitive functioning and can often predict cognitive decline and brain atrophy (Hua et al., 2008; Humphreys et al., 2013; Hyder et al., 2007), which is regularly seen in TBI patients. Prior understanding of CBF as it relates to TBI mistakenly operated on the assumption that CBF would actually increase after injury to the brain, leading to elevated intracranial pressure (Graves & Kreipke, 2015). More recent studies propose that elevated intracranial pressure is the outcome of edema and not CBF changes (Betrus & Kreipke, 2013). In fact, TBI is frequently associated with significantly reduced global CBF, with a decrease in CBF observed in the acute post-TBI settings (within the first 24 hours) being correlated with unfavorable patient outcomes and poor recovery (Barclay et al., 1985; Marion et al., 1991; Martin et al., 1997). Not surprisingly, blood supply in the brain is critically important regardless of TBI presence, as blood is constantly pumped throughout the brain tissues via a network of cerebral arteries and veins, supplying oxygen, glucose, and other vital nutrients, as well as removing carbon dioxide and other metabolic compounds (Joris et al., 2018).

The arterial blood volume transported to a unit of brain tissue per unit of time is quantified as the clinical measure of CBF (Joris et al., 2018). CBF is typically measured as milliliters (mL) of blood per 100 grams (g) of brain tissue per minute (min) (Joris et al., 2018). As an approximate estimate, the brain receives 1% of its total tissue volume of fresh blood each second, given an average brain tissue density of 1 g/mL (Liu & Brown, 2007). However, gray and white matter have

significantly different CBF measures as the metabolic demands of the respective tissues vary according to their function (Bentourkia et al., 2000). For example, an average measure of CBF in gray matter is about 50 mL/100 g/min (Petrella & Provenzale, 2000), which is consistent with about 1 mL of blood delivered to 100 g of brain tissue per second (Joris et al., 2018). Average CBF measures in white matter are roughly half that of gray matter at 22 mL/100 g/min (Carroll et al., 2008), receiving roughly 0.5 mL of blood to 100 g of brain tissue per second.

A reduction in CBF induced by brain tissue disruption post-TBI is generally the primary cause of post-traumatic cerebral ischemia and overall brain damage (Botteri, et al., 2008). However, the process by which CBF reduction occurs post-TBI remains unclear (Wang & Li, 2016). Neuronal loss in the brain post-TBI, therefore requiring a lesser demand for oxygen and nutrients, may trigger a reduction in CBF (Kim et al., 2010). Furthermore, microvascular injury, in particular, is associated with damage to the neurovascular unit (NVU); this NVU made up of the endothelial lining of the blood vessels, smooth muscle cells, and pericytes (Sandsmark et al., 2019). Disturbed NVU functioning affects cerebrovascular perfusion and reactivity, as well as the blood-brain barrier's effectiveness (Fridley et al., 2015). Together, these may disrupt normal repair systems on the neuronal level, resulting in further injury and loss of neurons in the brain and contributes to overall cognitive dysfunction commonly seen after TBIs (Golding et al., 1999; Golding, 2002; Hlatky et al., 2004; Honda et al., 2016). Treatment of such vascular damage at the NVU level appears to be a promising approach for TBI-related neurodegeneration drug therapies, thus further highlighting CBF measures as an important metric of TBI (Sandsmark et al., 2019).

The interplay between CBF and post-TBI recovery also remains unclear due to research on CBF disruption following TBIs being inconsistent (Kallakuri et al., 2015). For example, research into CBF changes post-TBI is limited in that many studies overlook changes in longer periods post-injury, and instead focus on CBF changes occurring only within a few hours or days following TBI

(Kallakuri et al., 2015). Another limitation of past studies is the focus on measuring CBF changes in specific, and therefore limited, anatomical regions instead of globally across brain regions (Kallakuri et al., 2015). Overall, the severities of TBI are also not well controlled or characterized in many studies (Kenney et al., 2016). Furthermore, past TBI neuroimaging research suffers additional drawbacks related to overlooking subjects with significant focal lesions, unreliable neuroimaging registration methods, using relative versus absolute measures of CBF, and limited repeated neuroimaging assessment due to minimizing exposure to ionizing radiation (Kim et al., 2010). Hence, more research is required to elucidate the true nature of CBF reductions post-TBI.

1.3. Neuroimaging of CBF

1.3.1. History of perfusion imaging

In the previous several decades, there have been several different imaging techniques used to assess CBF including positron emission tomography (PET), single-photon emission computed tomography (SPECT), perfusion CT, diffuse optical spectroscopy, dynamic susceptibility contrast (DSC), MRI imaging, and sonography (Proisy et al., 2016). Each technique has specific technical requirements, tracers, and mathematical modeling approaches that make each uniquely suited for examining certain imaging questions (Proisy et al., 2016). For example, PET with an ^{15}O exogenous tracer typically measures CBF; however, exposure to ionizing radiation is a major disadvantage of this technique since it potentially puts the subject's health at risk, especially for repeated measurements. Functional MRI (fMRI), in particular blood-oxygenation-level-dependent (BOLD) fMRI, is a non-ionizing and completely non-invasive imaging technique. While this MRI technique is widely used, it confounds neuronal and vascular behavior because it measures blood flow, blood volume, and hemoglobin oxygenation all at once (Kwong et al., 1992; Mandeville et al., 1999; Ogawa et al., 1993). This makes the resulting data difficult to interpret when looking at cerebrovascular damage (Kim et al., 2006). Additional limitations of BOLD fMRI are the presence

of low-frequency noise (Friston et al., 2000; Zarahn et al., 1997) and the lack of absolute CBF measures during multiple scanning sessions due to changes in subject activation (Kim et al., 2006). Hence, the aforementioned neuroimaging techniques are not always optimal to measure CBF.

1.3.2. Arterial spin labeling

Arterial spin labeling (ASL) is a relatively recent fMRI-based approach to measure CBF (Detre et al., 2009). Of particular interest to this current study is that ASL can estimate the damage to microvasculature as well as metabolic dysfunction often seen in TBI patients (Ware et al., 2020). Findings by Miller et al. (2001) proposed that ASL-generated CBF maps demonstrate a more linear relationship with neural activity compared to data from BOLD fMRI. Moreover, ASL provides high spatial resolution, absolute CBF quantification without the use of ionizing contrast agents and is obtained in a multi-modal MRI examination (Ware et al., 2020), making it a practical way to measure CBF. Unlike BOLD fMRI, ASL data is free of slow signal drifts, making it more appropriate for studying low frequency brain events (Wang et al., 2013). Additionally, ASL data may have reduced intersubject variability (Aguirre et al., 2002; Kemeny et al., 2005), be less prone to artifacts (Tjandra et al., 2005; Wang et al., 2004), and provide more precise functional localization than BOLD fMRI (Duong et al., 2001; Luh et al., 2000). This potentially reflects a more direct link between CBF and neural activity (Aguirre et al., 2002; Tjandra et al., 2005; Wang et al., 2003a).

During the ASL procedure, CBF is quantified by measuring magnetically-labeled arterial blood water labeled via radio-frequency (RF) pulses (Detre et al., 1992; Williams et al., 1992). Conceived by Williams et al. in 1992, with considerable refinements over the last several decades, the aim of ASL is to generate a “labeled image” and a “control image” (Petcharunpaisan et al., 2010). The labeled and control images differ only in the presence of inflowing magnetized arterial blood water and are otherwise identical in terms of static tissue signal. The RF pulse magnetically

labels the arterial blood water by inverting the hydrogen protons in the imaged brain region (Petersen et al., 2006; Williams et al., 1992; Wolf & Detre, 2007). A delay between acquiring the labeled and control images allows labeled blood to flow to the capillaries, generating a perfusion signal (Peterson et al., 2006; Pollock et al., 2009). The decay of the arterial blood water tracer follows a T1 longitudinal relaxation rate (Detre et al., 1994; Detre et al., 2009), which is a measure of the inverted hydrogen protons returning to equilibrium (Grover et al., 2015). Only small amounts of labeled protons concentrate in the imaged brain region, as the corresponding relaxation rate of the labeled arterial blood water is about 1 to 2 seconds (Detre et al., 1994; Wolf & Detre, 2007). The labeled image is ultimately subtracted from the control image, removing any static effects from the magnetization and other artifacts unrelated to CBF in the brain (Brown et al., 2007; Petcharunpaisan et al., 2010). Absolute CBF is then calculated by referencing established blood flow models which convert the perfusion signal from the subtracted images into a quantitative value (Alsop & Detre, 1996; Buxton et al., 1998).

Some limitations of CBF maps generated via ASL are susceptibility to artifacts and noise (Dolui et al., 2017). Because ASL scans last for several minutes and involve a subtraction technique, the resulting CBF maps are sensitive to subject movement (Petcharunpaisan et al., 2010). These artifacts can be somewhat reduced in the raw data by implementing motion correction algorithms (Dolui et al., 2017). Nevertheless, the signal-to-noise-ratio (SNR) in ASL-generated CBF maps is still very low even after motion correction and only about 1% of the background signal is removed during the subtraction process (Dolui et al., 2017). There are several sources of noise that contribute to this low SNR, such as the interplay of the T1 relaxation rate of arterial blood water and transit delay to the imaged brain area (Liu & Brown, 2007). Other sources of noise relate to additional nuanced complications of tagged arterial blood water perfusing throughout brain tissues and vasculature, such as the variable velocity of arterial blood water and lingering

residual tagged arterial blood water that has not yet reached its target brain region (Liu & Brown, 2007).

To boost the signal in the low SNR, mean CBF maps are generated from a series of ASL labeled and control image slices, known as volumes. However, this provides only a minor increase in the SNR and the resulting mean CBFs maps are still sensitive to outlier image volumes so even a small number of outlier voxels can consequently generate very biased mean CBF maps (Li et al., 2018). Several outlier cleaning strategies have been proposed to mitigate this phenomenon, some more simplistic while others include more sophisticated logic and adaptive outlier rejection criteria. A simple approach, for example, rejects an entire volume based on the mean and standard deviation falling outside some predefined range within an individual CBF map (Tan et al., 2009; Wang et al., 2008). More sophisticated, adaptive approaches iteratively reject outliers based on deviation from a representative mean CBF map (Wang et al., 2013). For example, the Structural Correlation-Based Outlier REjection (SCORE) algorithm by Dolui, et al. (2017), compares outlier volumes with the mean CBF map and rejects them if highly correlated, reasoning the outliers in those volumes influence the mean to a greater degree than nonoutlier volumes. An even more sophisticated, adaptive approach uses a prior-guided, slice-wise adaptive outlier cleaning algorithm named PAOCSL (Li et al., 2018). This algorithm leverages a prior mean CBF map generated from high quality data, using this as outlier rejection criteria at each slice while also considering neighboring slices. Slices are rejected as outliers if they are the least correlated with the mean CBF reference map (Li et al., 2018). PAOCSL has outperformed other outlier cleaning algorithms, benefitting from the fact that this algorithm does not simply discard entire volumes, identifying outliers in a slice-wise manner, retaining more data while detecting outliers at a more refined level.

1.4. Neuroimaging and machine learning

Neuroimaging data has conventionally been analyzed with standard statistical inference

techniques, examining whole-brain, voxel-wise correlations, deriving statistical significance measures such as *p*-values for each voxel in the data (Gaonkar & Davatzikos, 2013). This approach effectively considers each voxel separate from the whole image volume, ignoring correlation with neighboring voxels (Wang et al., 2006). While such an approach allows for ease of interpretation, it fails to adequately and efficiently capture the relationships among different brain regions (Gaonkar & Davatzikos, 2013). Given that CBF maps generated via ASL inherently represent data across different brain regions, it may require a multivariate approach to boost sensitivity (Wang et al., 2006).

Machine learning methods such as support vector machines (SVMs) have become increasingly more popular as researchers have realized that such multivariate pattern analysis (MVPA) provides richer information than univariate analysis, which can only consider one variable at a time (Gaonkar & Davatzikos, 2013). SVMs operate under an algorithm that considers the space of predictive features to find the best plane to separate binary classes. SVMs have been shown to be quite effective in predicting neurological outcomes related to brain activity, structure, and clinical features (Gaonkar & Davatzikos, 2013). In particular, the use of SVMs in the field of neuroimaging has garnered the most attention for its use in BOLD fMRI analysis as it has proven successful in brain state classification using whole-brain (LaConte et al., 2005; Mourão-Miranda et al., 2005) and pre-feature selected data (Cox & Savoy, 2003; Davatzikos et al., 2005; Mitchell et al., 2004; Wang et al., 2003b). In 2006, Wang et al. claimed they were the first to use an SVM to explore the multivariate information structure of ASL perfusion data obtained from a small study of healthy subjects. Since then, several other studies have leveraged SVMs trained on ASL data, most commonly to research age-related cognitive decline seen in Alzheimer's disease or dementia (Bron et al., 2014; Collij et al., 2016; Xekardaki et al., 2015).

1.5. The current study

The primary aim of this research is to use machine learning, specifically an SVM classification model, to classify TBI patients from healthy control subjects. The goal is to identify the most predictive regions of interest (ROIs) in the brain. A secondary aim of this research is to investigate the effectiveness of the PAOCSL outlier cleaning algorithm as compared to no outlier cleaning and the SCORE outlier cleaning algorithm with regards to boosting the SNR in ASL-generated CBF maps. Specifically, this research will examine the effectiveness of these outlier cleaning methods via the performance of an SVM machine learning model, which will classify and detect patient CBF maps from control CBF maps. We hypothesize that the PAOCSL algorithm will generate CBF maps with a greater SNR as evaluated by higher SVM classification performance compared to the CBF maps generated via no outlier cleaning or the SCORE algorithm when classifying patient versus control CBF maps.

As cited by Li et al. (2018), ASL is an up-and-coming technique used to measure CBF. However, the SNR in this neuroimaging technique is less than desirable. Moreover, past research in the field of neuroimaging generally uses purely statistical, univariate, and voxel-wise approaches to analyze data (Gaonkar & Davatzikos, 2013). As Gaonkar and Davatzikos (2013) point out, while univariate analyses may be more interpretable, multivariate approaches, such as machine learning, can provide a more robust depiction of the nuanced relationships within neuroimaging data (such as ASL data) and can also provide better predictive power. SVMs, in particular, are quite effective machine learning models when dealing with high dimensional spaces, such as neuroimaging data, and work well when dealing with small sample sizes (Overton et al., 2020).

While the current study builds upon findings from several previous studies, including Ware et al. (2020) and Li et al. (2018), a literature search on the topics of ASL, CBF, and machine learning indicates that previous research has not used any innovative ASL outlier cleaning

algorithms or SVM classification concerning the study of CBF changes seen post-TBI; prior studies have only used the standard techniques of post-processing subtraction, rudimentary outlier detection, and traditional statistical analyses.

This research is in response to past work as it relates to ASL, CBF, and TBI but with the novel approach of leveraging new and improved outlier detection and cleaning techniques, as well as implementing machine learning to better classify and detect CBF alterations between TBI patients and control subjects. From a clinical perspective, better detection leads to improved prediction of prognosis, ultimately leading to identification of treatment targets such as regional hypoperfusion. These targets may be treated with various neuromodulation methods including transcranial photo-biomodulation, which has been shown to increase CBF in TBI patients (Torre et al., 2019).

2. Methods

2.1. Participants

The institutional review board of the home institution approved this study and informed written consent was granted by the participants or their legally authorized representatives. The subject pool consisted of participants between the ages of 18 and 64 years old with a history of at least a moderate severity, non-penetrating TBI. The subjects presenting with TBI met at least one of the following criteria: a Glasgow Coma Scale (GCS) not related to sedation, paralysis, or intoxication of less than 13 while in the emergency department; 12 or more hours of loss of consciousness, or post-traumatic amnesia (PTA) lasting longer than 24 or more hours. Excluded participants met one or more of the following criteria: a history of previous TBI, central nervous system disease, seizure disorder, schizophrenia, or bipolar disorder; a history of drug or alcohol abuse; pregnancy; inability to complete MRI scanning due to implants, claustrophobia, or

restlessness; non-fluency in English; or incomplete testing and scanning 3 months post-TBI due to severity of disability. Another exclusion criterion included focal intraparenchymal lesions greater than 5 cm^3 and 50 cm^3 for subcortical and cortical lesions, respectively. This resulted in a subject pool of 42 TBI participants which were matched by 35 healthy control volunteers based on age, sex, and years of education. The control participants were subject to the same exclusion criteria mentioned previously with an additional exclusion criterion of any history of TBI resulting in loss or alteration of consciousness. Neuroimaging, specifically ASL imaging measures, were obtained from all moderate-to-severe TBI participants 3, 6, and 12 months post-injury. Neuropsychological measures were also obtained post-injury for the same time periods. For the purposes of this study, only ASL neuroimaging data 3 months post-injury was analyzed.

2.2. Clinical measures

Clinical measures were ascertained during hospitalization or retrieved from medical records. TBI severity was indexed by PTA duration, measured during patient rehabilitative care, or conservatively estimated if patients still exhibited PTA when discharged. All control subjects and TBI patients were subjected to neuropsychological tests in domains often affected by TBI, such as processing speed, executive function, and verbal learning to assess cognitive performance and impairment. While the neuropsychological impact of TBI is not under direct investigation in this current study, previous research the current study is based on did examine the relationship between imaging measures and neuropsychological test performance and demonstrated a moderate correlation (Ware et al., 2020).

2.3. ASL imaging acquisition

ASL imaging data was collected via a pseudo-continuous ASL sequence in which the ASL plane is 9 cm below the center of the imaging volume. The labeling duration and post-labeling delay were 1.5 s. A two-dimensional echo planar sequence was used to acquire the images. The

parameters of a 4 s repetition time (TR), an 18 ms echo time (TE), a field of view of 220 mm, a matrix 64 x 64, and voxel size 3.4 x 3.4 x 7.2 mm^3 were used. A total of 18 sequential slices with a thickness of 6 mm were acquired in an inferior to superior orientation. Signal averaging was executed using 45 label-control image pairs.

2.4. Data preprocessing

ASL preprocessing was applied to the ASL imaging data prior to any machine learning using a standard ASL toolbox pipeline and SPM12 software (Wang et al., 2008). First, the origin of each volume was reset to the center, reorienting the original native space structural T1 data. Following this, the raw ASL data were motion corrected by realigning to each subject's mean functional image. Motion timecourses for the ASL control and label images were estimated by a rigid body transform and the spin and control labeling time paradigm was removed from the motion timecourses via regression. Acquired label and control images were motion corrected using the residuals. Next, the ASL data were coregistered to the reorientated structural T1 image data. To reduce noise, SPM's Gaussian smoothing kernel of 6 mm full width at half maximum was used to smooth the realigned and coregistered ASL image data. Following this, extra-cranial voxels were excluded using perfusion masks based on the mean functional images. Pairwise subtraction of the control and label images quantified the CBF from the ASL images (Wang et al., 2008) and the resulting CBF maps were segmented and normalized to the Montreal Neurological Institute (MNI) template space. Finally, three sets of mean CBF maps were generated using no outlier cleaning (for a baseline measure), as well as the SCORE and PAOCSL outlier cleaning algorithms.

2.5. Dimensionality reduction and feature selection

Since the CBF map data set in this study was quite small in terms of sample size, at a total of 77 participants, but yet quite high in number of variables, with the dimensionality of each 3D ASL image at 79 x 95 x 79 voxels, amounting to 592,895 total voxels, selecting meaningful input

variables (features) for the model was of critical importance. While each voxel could be considered an individual feature, given the large number of voxels in the whole brain ASL image data, it was paramount to reduce the number of features for use in the SVM model to avoid overfitting and the curse of dimensionality (Noble, 2006).

In machine learning, overfitting occurs when the analysis of too many features in a given data set causes the model to essentially "memorize" the data rather than learn a true signal from the features (Hawkins, 2004). This results in a model that may have poor performance and generalizability on yet unseen data (Hawkins, 2004). On the other hand, the curse of dimensionality arises when analyzing a high-dimensional feature space, such that the number of features greatly outnumbers the data samples used for training (Hughes, 1968). In fact, while the number of features increases at a fixed training sample size, the model's performance will indeed increase up to a certain number of additional features; however, past this point the model's performance will steadily decline (Hughes, 1968). This decline in performance is due to the additional features obscuring any significant difference between training samples in the data set, essentially introducing noise, causing the model to not discriminate any relevant pattern in the data (Houle et al., 2010).

For each of the three CBF map data sets handled with the different ASL outlier cleaning methods under study here, dimensionality reduction and feature selection were performed using nilearn (Abraham et al., 2014), a Python based statistical neuroimaging library, using the approach of computing an ROI mask. The ROIs were calculated using the Harvard-Oxford probabilistic atlas, covering 48 cortical regions, sampled at 2 mm, with a threshold probability of 0.25, and combining values from both the right and left hemispheres. In particular, the nilearn class *NiftiLabelsMasker* was implemented to extract and calculate the mean CBF values from a subset of voxels corresponding to the 48 cortical regions rather than the whole brain images. This effectively

reduced the data from a potentially very high dimensional feature space of 592,895 voxels to a more tractable feature space of 48 ROIs. Furthermore, using ROIs as a dimensionality reduction technique has the additional advantage of better interpretability given that a priori neurobiologically derived brain regions should reflect a natural link between units of analysis and physiological function (Eickhoff et al., 2018). See Appendix Figure 1 for a visualization of the Harvard-Oxford atlas.

The ASL image data used in this study were stored in NIfTI file format. Unfortunately, NIfTI files can contain null values, generally produced by prior spatial registration conducted as part of the SPM data preprocessing pipeline. This is problematic because the null values signify that certain brain regions are missing data. To ameliorate the complications caused by any missing data in this analysis, mean values for ROIs containing less than 10% null values were calculated for any subjects meeting this criterion by simply ignoring the null values when taking the mean. For example, one patient had 3.7% null values in the superior frontal gyrus but given that more than 90% of the voxels in that region contained valid data, the mean was calculated from the remaining 96.3% of data in this region. On the other hand, any ROIs with greater than 10% null values were discarded from further analyses. This included two regions, the temporal fusiform cortex (anterior division) and inferior temporal gyrus (anterior division). Both these regions consistently showed up as regions with greater than 10% null values in several subjects across both patients and controls. The data from all three outlier cleaning methods exhibited identical null value percentages across the same ROIs. As the temporal fusiform cortex (anterior division) and inferior temporal gyrus (anterior division) were discarded from the feature set due to lack of reliable data, the final feature set contained 46 ROIs.

2.6. Exploratory data analysis

Subjects were divided by patients and controls and exploratory data analyses were

performed separately for each outlier cleaning method using pandas, an open-source Python data analysis and manipulation tool (McKinney, 2010). The group mean, standard deviation, median, interquartile range, minimum, and maximum mean CBF values were calculated for each of the 46 ROIs. Additionally, the mean differences in control versus patient ROI values were determined, again for each of the three different outlier cleaning methods.

2.7. Machine learning classification and validation

The machine learning classification and validation portion of this study was performed using scikit-learn (Pedregosa et al., 2011), an open source, state of the art machine learning Python library. For each of the three CBF map data sets generated with the different ASL outlier cleaning methods, a linear kernel (a mathematically derived space that transforms the data) SVM model was trained using the 46 ROIs as features. The regularization parameter C was left at the default value of 1.0, penalizing misclassifications with an L2 penalty, which has been shown to make SVMs less prone to overfitting (Schölkopf et al., 2001). A small class imbalance existed between control and patient groups, slightly biased towards patients with patients making up 55.69% of the entire sample. Given how minor the imbalance was, and the fact that it's been shown that SVMs still perform reasonably well on moderately imbalanced data (Akbani et al., 2004), the only measure employed to address the class imbalance was to stratify the data such that the original distribution of classes (patients and controls) remains intact during model training and validation. Because SVMs assume all feature values are centered around zero and have a variance in the same order of magnitude (scikit-learn developers, 2020b), the 46 ROI features were scaled through standardization to have a mean of zero and a standard deviation of one.

Three SVM models were trained, one with the CBF maps preprocessed with no outlier cleaning, another with the CBF maps preprocessed with the SCORE outlier cleaning algorithm, and a third using the CBF maps preprocessed with the PAOCSL outlier cleaning algorithm. The models

were initially evaluated using a 10-fold stratified cross validation technique and the area under the Receiver Operating Characteristic (ROC) curve was calculated for each model trained on the 46 ROI features. K-fold cross validation involves splitting the data into k smaller sets, or "folds", training the model on $k - 1$ of folds for training, validating the model on the remaining data, and measuring the final performance by taking the average across the training folds (scikit-learn developers, 2020a). While the primary use case for cross validation is to avoid overfitting while tuning hyperparameters (which was not done in this study), cross validation is also advantageous when evaluating models trained on small data sets (scikit-learn developers, 2020a). However, single models trained on small data sets containing less than a few hundred samples with a relatively high feature count can have better than expected results on cross validation due to chance alone (scikit-learn developers, 2020a). Despite the dimensionality reduction that was implemented on the CBF map data, the resulting data set under study here is still quite feature rich relative to the small sample size. To address this concern, other groups have investigated the use of permutation-based p-values for assessing model performance (Golland & Fischl, 2003; Golland et al., 2005; Hsing et al., 2003; Molinaro et al., 2005). A permutation test evaluates how likely the observed model performance would be obtained by chance (Ojala & Garriga, 2010). As such, a permutation test was conducted to evaluate the model performances. First, the mean accuracy, precision, and recall across 1,000 iterations of a five-fold stratified cross validation SVM model were calculated using the unpermuted label data (i.e., control or patient label) for data from each outlier cleaning methods. This greatly reduces any results due to variation in the data since each observation in the data is randomly subsampled as training and testing data many times (Ojala & Garriga, 2010), providing more robust estimates of model performance. These values were then compared to a null distribution to ascertain p-values. The null distribution was calculated by taking the average performance of the five-fold cross validation SVM trained on 1,000 different permutations of the

labels being randomly shuffled. The *p*-value was then measured as the fraction of permutations for which the average cross validation score using the model trained on the randomly shuffled label data outperformed the score from the model using the original (unshuffled) data. This effectively determines if there is a dependency between the features and the labels in the data (Ojala & Garriga, 2010).

To further examine model performances, analysis of misclassified samples predicted from a single SVM model trained on 70% of the data and evaluated on the remaining 30% of the data as a hold-out test set was also performed for each set of data generated from the three different outlier cleaning methods. This analysis was more exploratory in nature, since as previously stated, results from a single SVM model trained on a small data set may not be robust enough to generalize well to unseen data. This exploration aimed to get a rough sense of the kind of samples the SVMs misclassified for each of the three data sets.

2.8. ROI feature importance

The most important ROI features for SVM classification were determined using a permutation importance method, again using a scikit-learn implementation. The basic idea behind the permutation feature importance method is as a single feature value is randomly shuffled, breaking the relationship between the feature and label. Any accompanying drop in model performance indicates how important the feature is to the model (Breiman, 2001). Put another way, features with no predictive value do not affect model performance when their values are randomized. Due to limitations in the scikit-learn implementation, an SVM trained on cross validation was not employed for this permutation feature importance method. However, similar to the permutation-based *p*-value test strategy, results due to only chance subsampling were controlled for by training and testing 1,000 iterations of an SVM, using different random stratified subsets of 75% training and 25% testing data for each run in the 1,000 iterations. During each iteration on the

random train-test split of the data, the values of each ROI feature were randomly shuffled for another five iterations and the mean model performance change (compared to the baseline model performance on the unshuffled ROI feature data) was calculated as ROI feature importance. After the 1,000 iterations were complete, the mean ROI feature importances over the 1,000 iterations of the SVM were computed as the final ROI feature importances.

The magnitude of the SVM coefficients (weights assigned to the features) were also evaluated in terms of ROI feature importance. For an SVM with a linear kernel, the coefficients have been shown to be directionally proportional to the most informative features (Guyon et al., 2002). That is, the features with the largest magnitude coefficients impact the classification decision the most. The positive or negative sign of the SVM coefficients can be roughly understood as contributing to the prediction of a class (Rodríguez-Pérez et al., 2017). More specifically in this study, a positive sign contributes to the prediction of TBI patients, whereas a negative sign contributes to the prediction of control subjects. A five-fold stratified, cross-validation SVM model was implement using 1,000 iterations to ensure robust results. The mean coefficient value of these 1,000 iterations resulted in a total of 5,000 coefficients for each ROI. These mean coefficient values were then computed as another estimate of ROI feature importance.

3. Results

3.1. ROI exploratory data analysis

Dimensionality reduction and feature selection resulted in mean CBF values for a total of 46 ROIs using data from each of the three ASL outlier cleaning methods: no outlier cleaning, SCORE, and PAOCSL outlier cleaning. Two regions were discarded due to inadequate brain coverage: the temporal fusiform cortex (anterior division) and the inferior temporal gyrus (anterior division). Additional ROIs with null values that were equally distributed across both control and

patient groups included the temporal pole, temporal fusiform cortex (posterior division), occipital pole, inferior temporal gyrus (posterior division), precentral gyrus, and superior frontal gyrus which were retained in the feature set as these ROIs had only 5% or less null values. The top five largest ROIs in terms of total voxel size included the frontal pole, lateral occipital cortex (superior division), precentral gyrus, postcentral gyrus, and precuneus cortex with the smallest ROI being the supracalcarine cortex. See Appendix Table 1 for total voxel sizes across all 48 ROIs, including the two discarded ROIs.

The control and patient group means and standard deviations were calculated for each of the three outlier cleaning methods across the 46 ROIs. ROIs consistently within the top ten highest group mean CBF values across both control and patient groups for all three outlier cleaning methods included Heschl's gyrus (both H1 and H2), the cingulate gyrus (both anterior and posterior), the parietal operculum cortex, the intracalcarine cortex, and the insular cortex. ROIs consistently within the top ten lowest group mean CBF values included the inferior temporal gyrus (posterior division), the parahippocampal gyrus (anterior division), temporal fusiform cortex (posterior division), subcallosal cortex, and the temporal pole. The majority of mean CBF values were higher for controls versus patients. As a general trend, the standard deviation for each ROI across all three outlier cleaning methods was similar for the control groups, with the frontal medial cortex having the largest standard deviation. The only major point of difference was the standard deviation for the frontal pole was considerably larger in the patient data for no outlier cleaning compared to the data cleaned with the SCORE or PAOCSL algorithms in which, again, the frontal medial cortex was the ROI with the largest standard deviation. The mean CBF differences between control and patient groups were also calculated for each of the three outlier cleaning methods across the 46 ROIs. The insular cortex, inferior frontal gyrus (pars triangularis and pars opercularis), and frontal and temporal poles were common ROIs with the greatest mean CBF

differences across control and patient groups for all three outlier cleaning methods. The top ten ROIs with the largest control minus patient mean CBF differences are plotted in Figure 1 for each outlier cleaning method (also see Appendix Table 1 for mean CBF difference values for all 46 ROIs).

In addition to the control and patient group mean and standard deviation, the control and patient median, interquartile range, minimum, and maximum mean CBF values for each ROI are shown as box plots in Figure 2 to visualize the distribution of data across control and patient groups for each of the three outlier cleaning methods.

3.2. SVM classification performance

The performance of the three SVM models trained on the CBF data preprocessed with no outlier cleaning, with SCORE outlier cleaning, and with PAOCSL outlier cleaning are summarized as the area under the ROC curves (AUC) after a 10-fold cross validation (Figure 3). In the case of the current study, the AUC represents the model's ability to classify patients from control subjects. While it appears as though the SVM model trained on the PAOCSL outlier cleaned data performed best, with an AUC of 0.74, closer examination revealed a large degree of variance in model performances across the 10-fold cross validation for all three SVM models. However, as mentioned previously in the Methods section, machine learning model performance can be sensitive to random chance when cross validation is performed using small data sets with a high feature count. For this reason, a permutation-based *p*-value test evaluating the mean accuracy, mean precision, and mean recall across 1,000 iterations of a 5-fold cross validation SVM model was performed for each of the three outlier cleaning methods. The results of the permutation test are shown in Figure 4. While the mean accuracy score (0.61) was near significance (*p*-value = 0.055) for the SVM trained on the data using the PAOCSL outlier cleaning algorithm and the mean precision score (0.65) was near significance (*p*-value = 0.057) for the SVM trained on the data from the SCORE

outlier cleaning algorithm, the *p*-values were not less than 0.05, therefore none of the SVM models evaluated in this study were found to perform significantly better than chance.

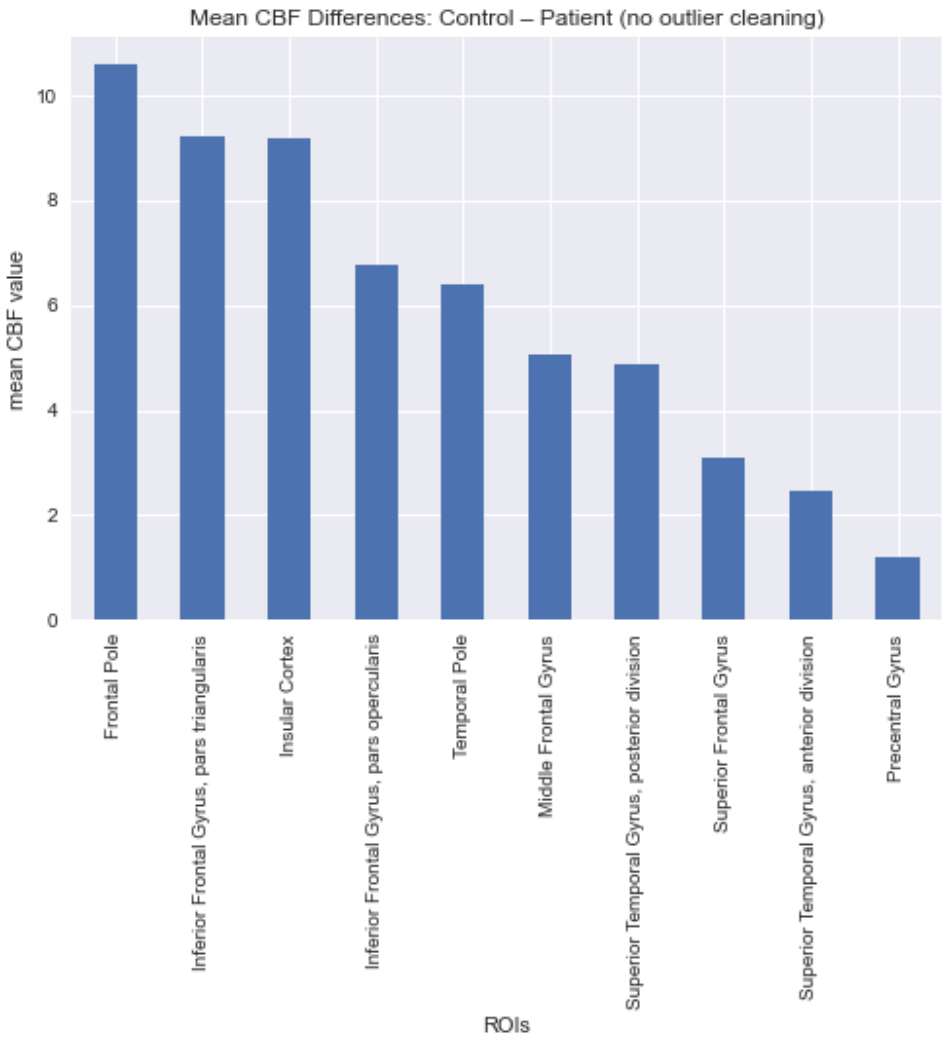
To get an approximate sense of the misclassification errors of each of the three SVM models, samples incorrectly classified as either patient or control were examined. As a general trend, the SVM model trained on the no outlier cleaned data misclassified patients more often than controls as compared to models trained on the SCORE or PAOCSL outlier cleaned data. Albeit the test data set was quite small (n = 24 subjects), there was a degree of overlap between misclassified samples across the three different SVM models. See Appendix Figure 2 for an example of a patient all three models misclassified as a control subject. While all three models misclassified this subject, qualitatively one can see visual differences in the CBF map images, likely the result of the different outlier cleaning methods used.

3.3. Identification of the most predictive ROIs

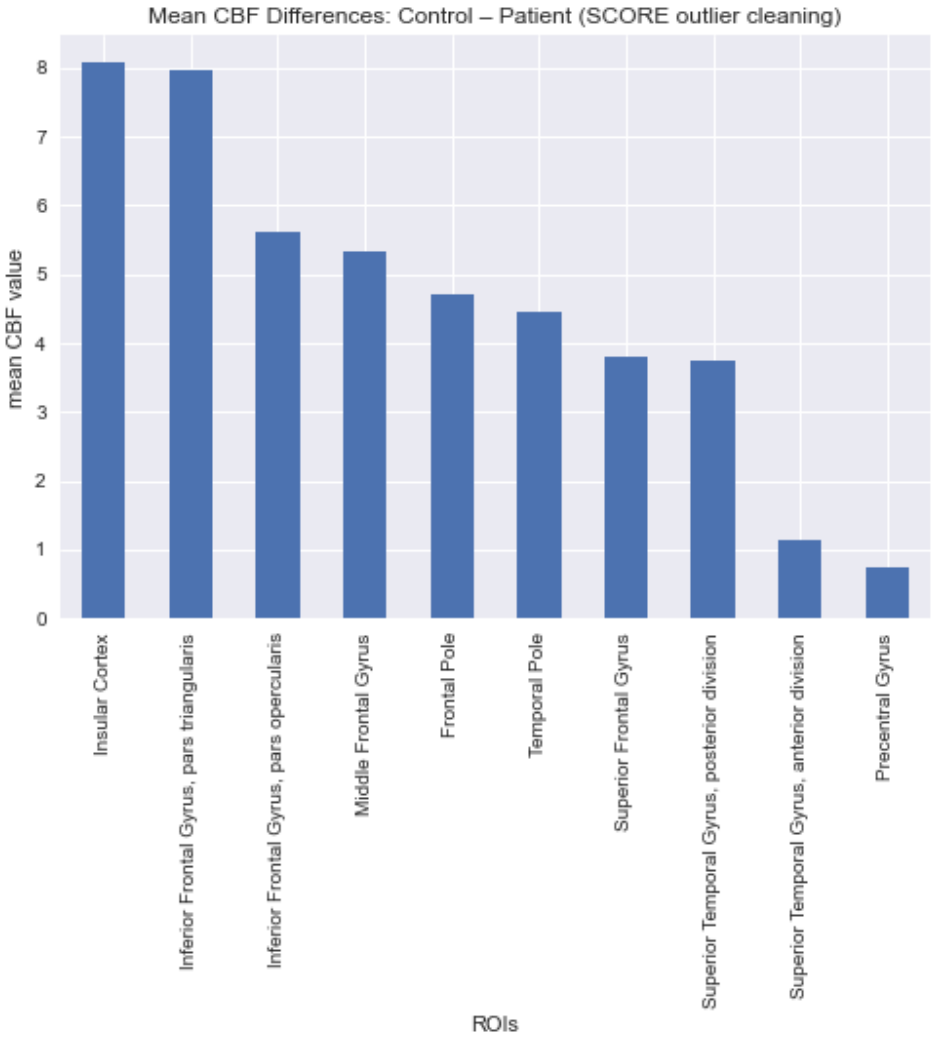
Although the SVM model classification portion of this study did not yield statistically significant results in terms of model performance, feature importances, that is the most predictive ROIs, were still investigated for the SVM model trained on the PAOCSL outlier cleaned data as this model's accuracy was highest and most closely approached significance (mean accuracy = 0.61, *p*-value = 0.055). The top ten most important features from the permutation importance method are plotted in Figure 5a. The most predictive ROIs were also estimated in relation to the coefficients from the SVM trained on the PAOCSL outlier cleaned data and are plotted in Figure 5b. Several of the same ROIs were identified as the most predictive from the permutation importance method as well as from the SVM coefficient analysis. These ROIs included the temporal occipital fusiform cortex, the lingual gyrus, the temporal pole, the inferior frontal gyrus (pars triangularis), the occipital pole, and the parahippocampal gyrus (posterior division).

Figure 1. Mean CBF differences between control and patient groups.
Top ten ROIs with the greatest difference in mean CBF values between control and patient groups using (a) no outlier cleaning, (b) SCORE outlier cleaning, and (c) PAOCSL outlier cleaning.

(a)



(b)


Mean CBF Differences: Control – Patient (SCORE outlier cleaning)
mean CBF value
0
1
2
3
4
5
6
7
8
Insular Cortex
Inferior Frontal Gyrus, pars triangularis
Inferior Frontal Gyrus, pars opercularis
Middle Frontal Gyrus
Frontal Pole
Temporal Pole
Superior Frontal Gyrus
Superior Temporal Gyrus, posterior division
Superior Temporal Gyrus, anterior division
Precentral Gyrus
ROIs

(c)

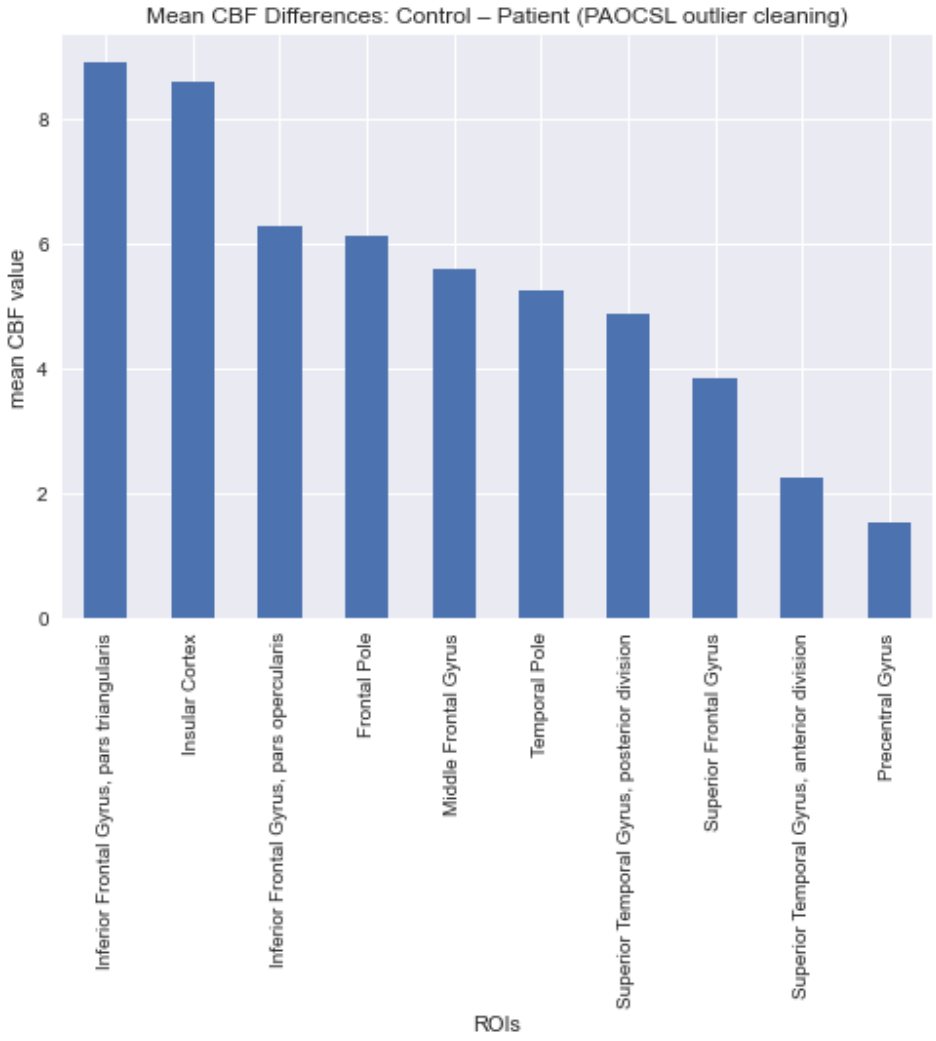
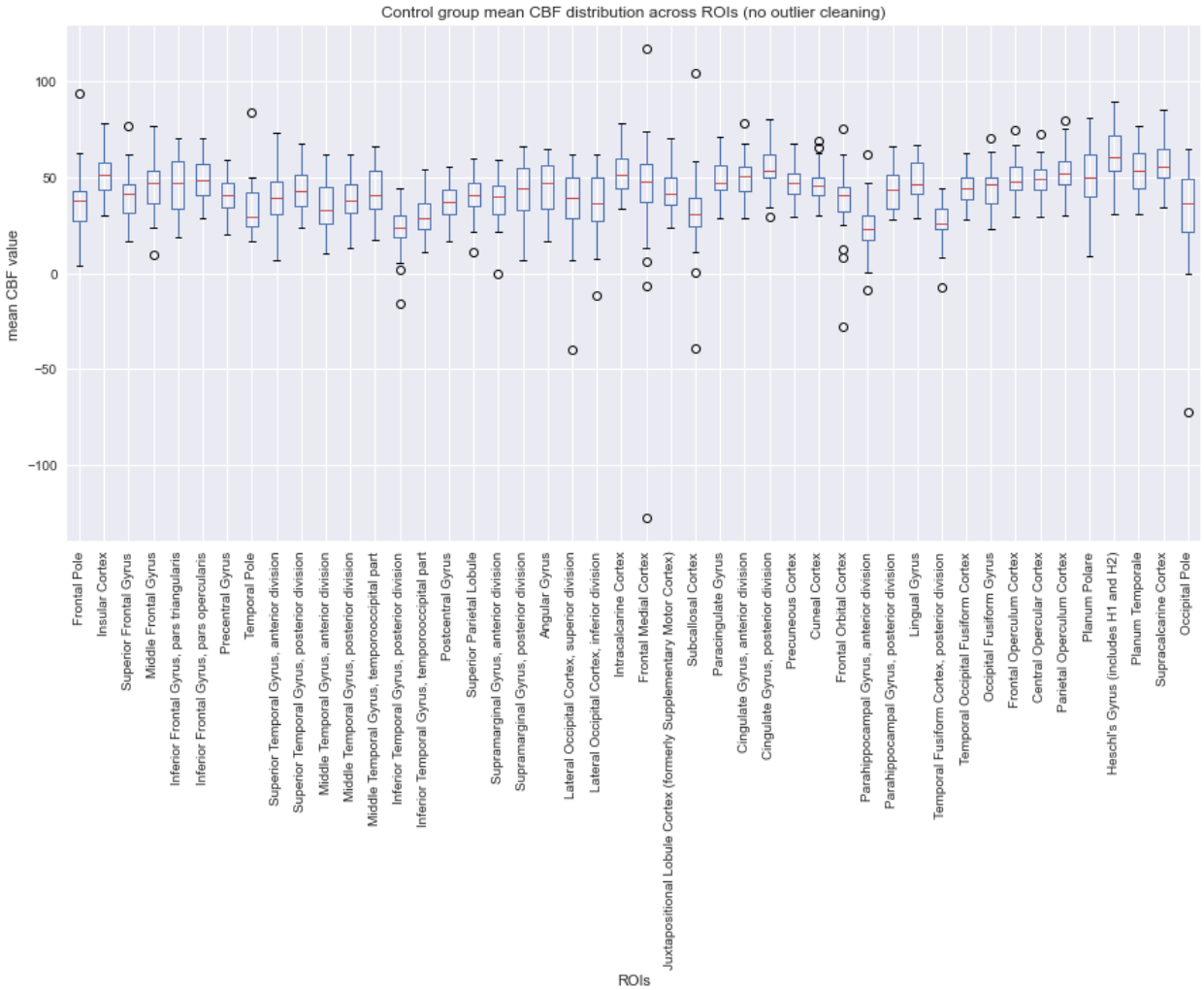
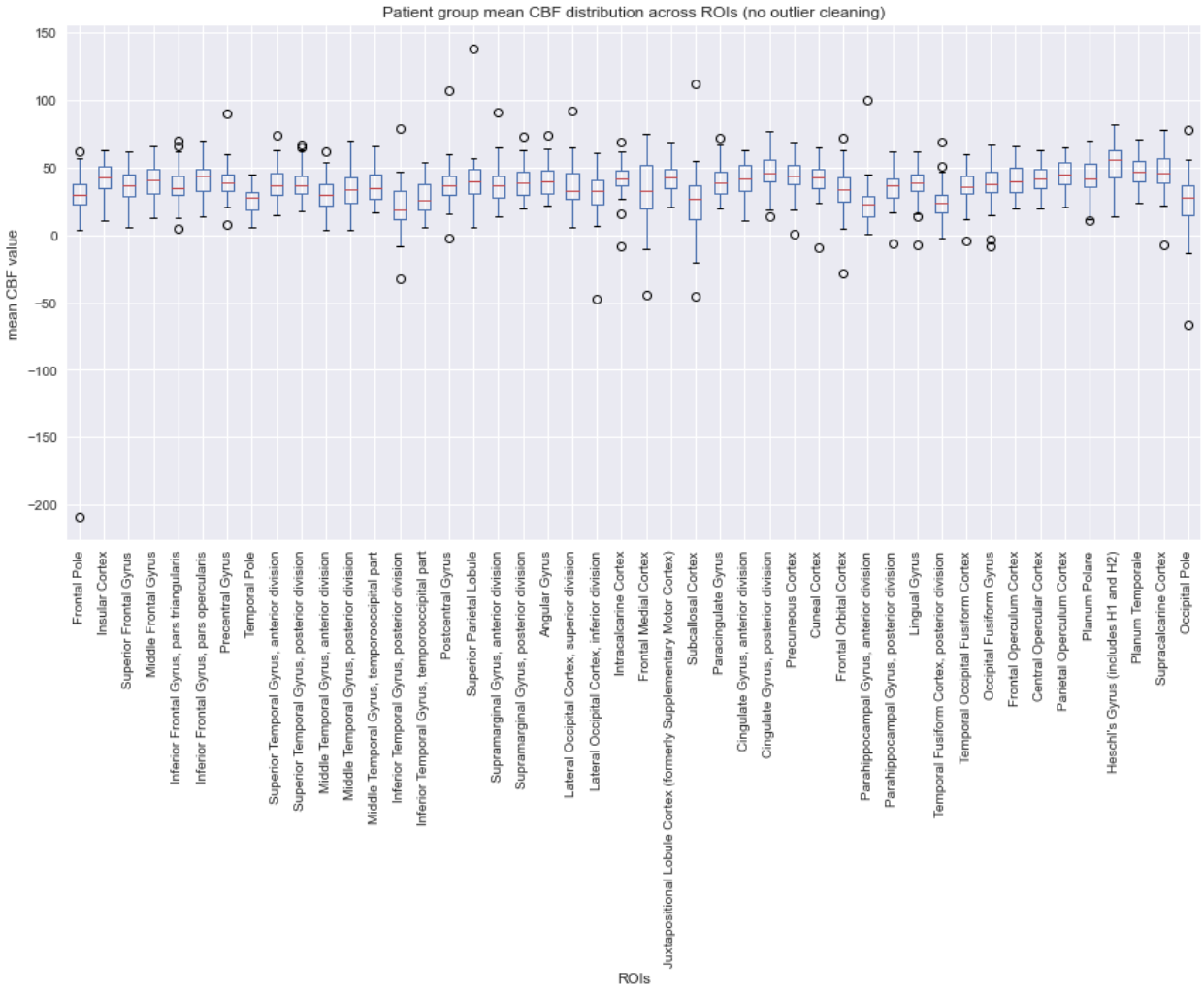

Mean CBF Differences: Control – Patient (PAOCSL outlier cleaning)
mean CBF value
0
2
4
6
8
Inferior Frontal Gyrus, pars triangularis
Insular Cortex
Inferior Frontal Gyrus, pars opercularis
Frontal Pole
Middle Frontal Gyrus
Temporal Pole
Superior Temporal Gyrus, posterior division
Superior Frontal Gyrus
Superior Temporal Gyrus, anterior division
Precentral Gyrus
ROIs

Figure 2. Group mean CBF distributions across 46 ROIs.
Box plots showing (a) control group and (b) patient group with no outlier cleaning, (c) control group and (d) patient group with SCORE outlier cleaning, (e) control group and (f) patient group with PAOCSL outlier cleaning. The red line represents the median, the box represents the interquartile range, the whiskers show the range of the data (1.5 x interquartile range), and outliers are plotted as open circles.

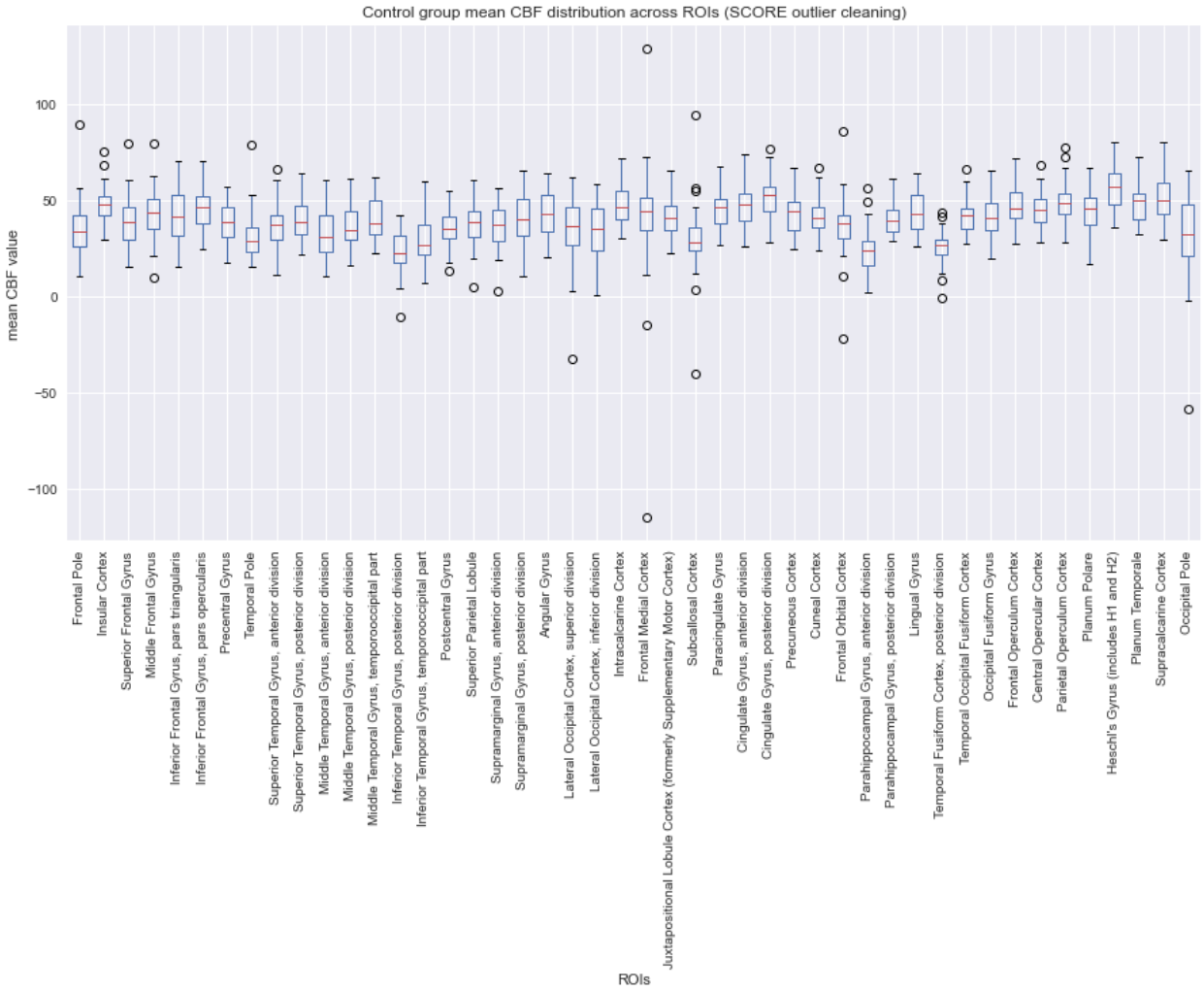
(a)



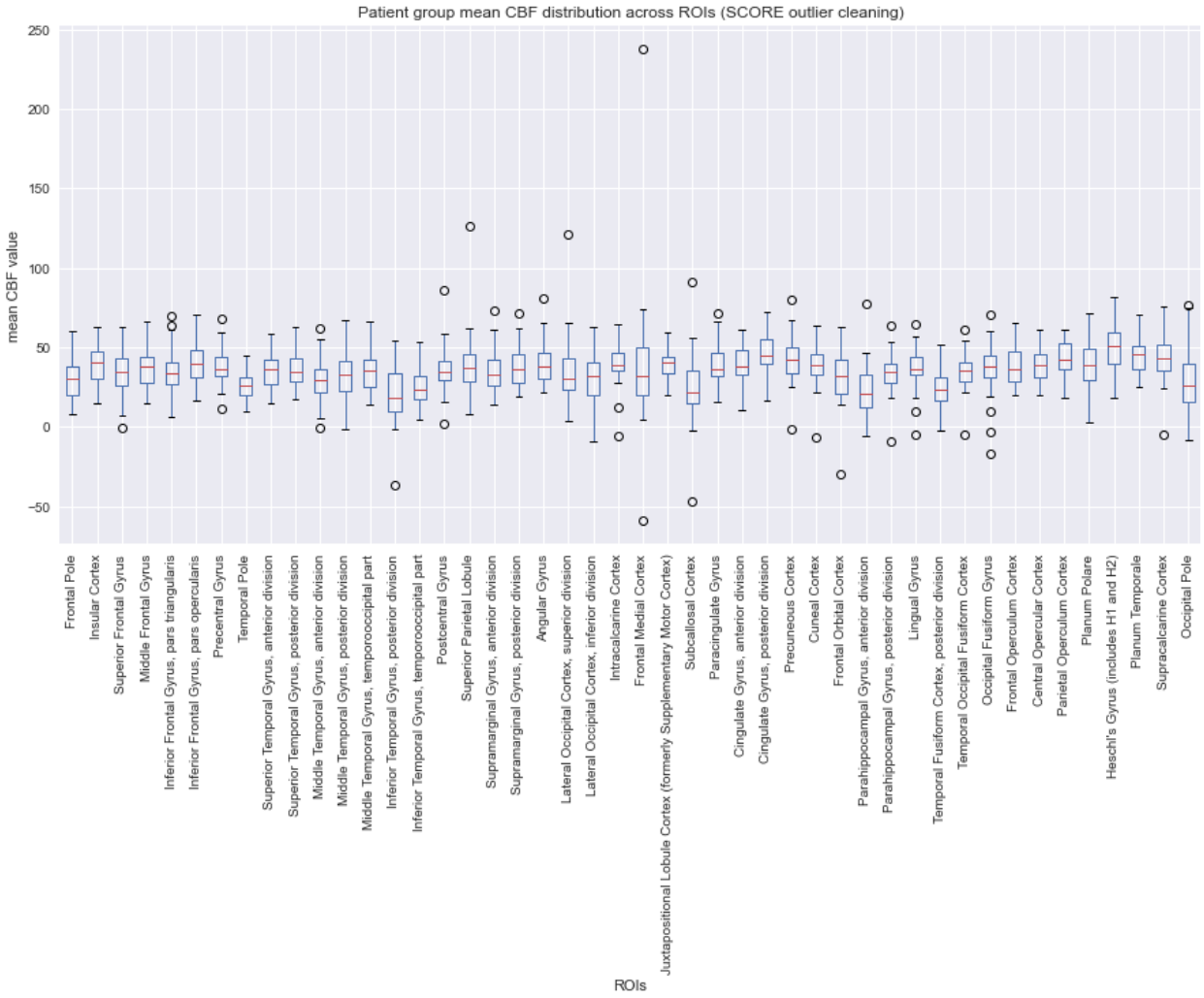
(b)


Patient group mean CBF distribution across ROIs (no outlier cleaning)
mean CBF value
150
100
50
0
−50
−100
−150
−200
Frontal Pole
Insular Cortex
Superior Frontal Gyrus
Middle Frontal Gyrus
Inferior Frontal Gyrus, pars triangularis
Inferior Frontal Gyrus, pars opercularis
Precentral Gyrus
Temporal Pole
Superior Temporal Gyrus, anterior division
Superior Temporal Gyrus, posterior division
Middle Temporal Gyrus, anterior division
Middle Temporal Gyrus, posterior division
Middle Temporal Gyrus, temporooccipital part
Inferior Temporal Gyrus, posterior division
Inferior Temporal Gyrus, temporooccipital part
Postcentral Gyrus
Superior Parietal Lobule
Supramarginal Gyrus, anterior division
Supramarginal Gyrus, posterior division
Angular Gyrus
Lateral Occipital Cortex, superior division
Lateral Occipital Cortex, inferior division
Intracalcarine Cortex
Frontal Medial Cortex
Juxtapositional Lobule Cortex (formerly Supplementary Motor Cortex)
Subcallosal Cortex
Paracingulate Gyrus
Cingulate Gyrus, anterior division
Cingulate Gyrus, posterior division
Precuneous Cortex
Cuneal Cortex
Frontal Orbital Cortex
Parahippocampal Gyrus, anterior division
Parahippocampal Gyrus, posterior division
Lingual Gyrus
Temporal Fusiform Cortex, posterior division
Temporal Occipital Fusiform Cortex
Occipital Fusiform Gyrus
Frontal Opercular Cortex
Central Opercular Cortex
Parietal Opercular Cortex
Planum Polare
Heschl's Gyrus (includes H1 and H2)
Planum Temporale
Supracalcarine Cortex
Occipital Pole
ROIs

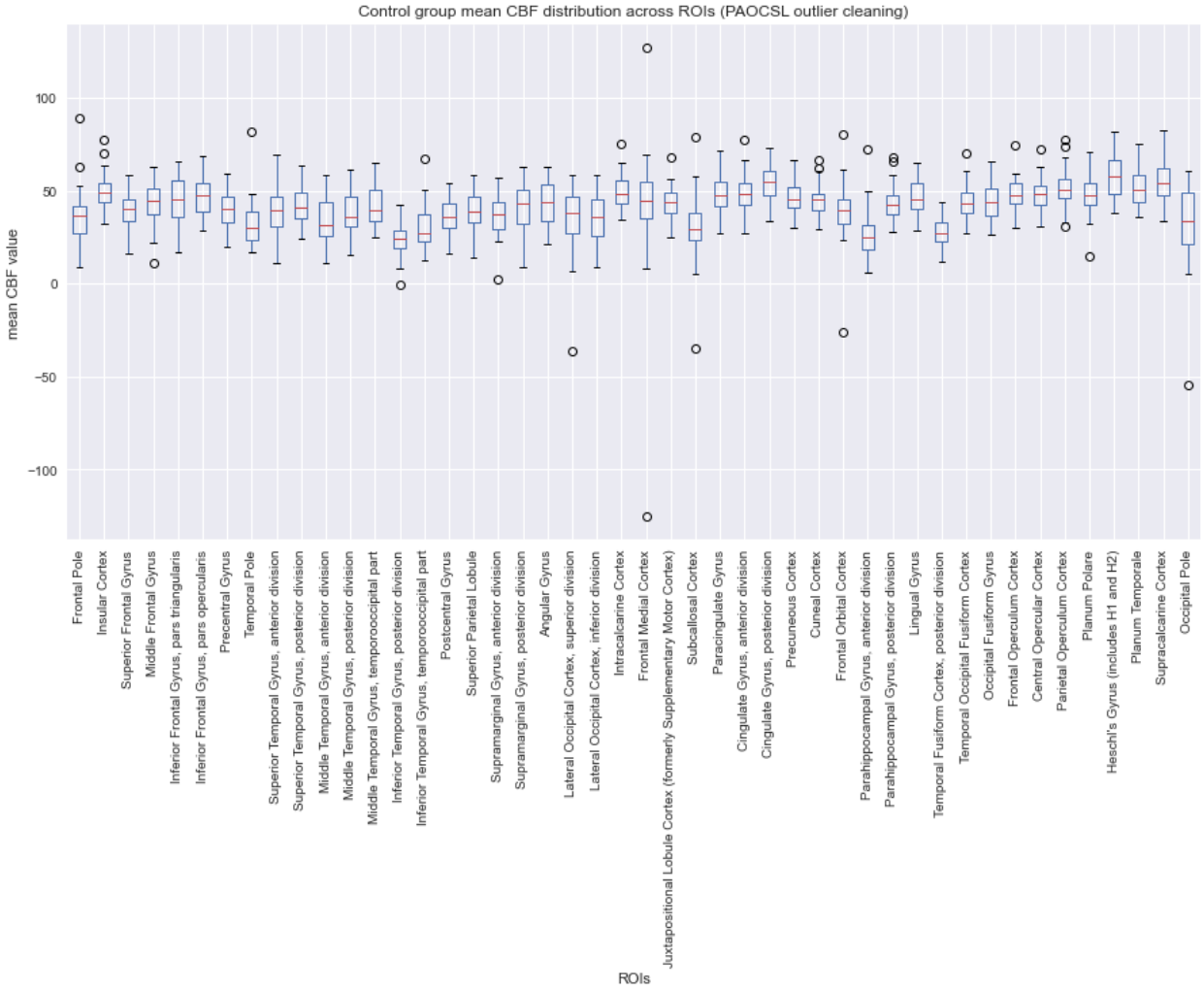
(c)


Control group mean CBF distribution across ROIs (SCORE outlier cleaning)
mean CBF value
100
50
0
−50
−100
Frontal Pole
Insular Cortex
Superior Frontal Gyrus
Middle Frontal Gyrus
Inferior Frontal Gyrus, pars triangularis
Inferior Frontal Gyrus, pars opercularis
Precentral Gyrus
Temporal Pole
Superior Temporal Gyrus, anterior division
Superior Temporal Gyrus, posterior division
Middle Temporal Gyrus, anterior division
Middle Temporal Gyrus, posterior division
Middle Temporal Gyrus, temporooccipital part
Inferior Temporal Gyrus, posterior division
Inferior Temporal Gyrus, temporooccipital part
Postcentral Gyrus
Superior Parietal Lobule
Supramarginal Gyrus, anterior division
Supramarginal Gyrus, posterior division
Angular Gyrus
Lateral Occipital Cortex, superior division
Lateral Occipital Cortex, inferior division
Intracalcarine Cortex
Frontal Medial Cortex
Juxtapositional Lobule Cortex (formerly Supplementary Motor Cortex)
Subcallosal Cortex
Paracingulate Gyrus
Cingulate Gyrus, anterior division
Cingulate Gyrus, posterior division
Precuneous Cortex
Cuneal Cortex
Frontal Orbital Cortex
Parahippocampal Gyrus, anterior division
Parahippocampal Gyrus, posterior division
Lingual Gyrus
Temporal Fusiform Cortex, posterior division
Temporal Occipital Fusiform Cortex
Occipital Fusiform Gyrus
Frontal Opercular Cortex
Central Opercular Cortex
Parietal Opercular Cortex
Planum Polare
Heschl's Gyrus (includes H1 and H2)
Planum Temporale
Supracalcarine Cortex
Occipital Pole
ROIs

(d)


Patient group mean CBF distribution across ROIs (SCORE outlier cleaning)
mean CBF value
250
200
150
100
50
0
−50
Frontal Pole
Insular Cortex
Superior Frontal Gyrus
Middle Frontal Gyrus
Inferior Frontal Gyrus, pars triangularis
Inferior Frontal Gyrus, pars opercularis
Precentral Gyrus
Temporal Pole
Superior Temporal Gyrus, anterior division
Superior Temporal Gyrus, posterior division
Middle Temporal Gyrus, anterior division
Middle Temporal Gyrus, posterior division
Middle Temporal Gyrus, temporooccipital part
Inferior Temporal Gyrus, posterior division
Inferior Temporal Gyrus, temporooccipital part
Postcentral Gyrus
Superior Parietal Lobule
Supramarginal Gyrus, anterior division
Supramarginal Gyrus, posterior division
Angular Gyrus
Lateral Occipital Cortex, superior division
Lateral Occipital Cortex, inferior division
Intracalcarine Cortex
Frontal Medial Cortex
Juxtapositional Lobule Cortex (formerly Supplementary Motor Cortex)
Subcallosal Cortex
Paracingulate Gyrus
Cingulate Gyrus, anterior division
Cingulate Gyrus, posterior division
Precuneous Cortex
Cuneal Cortex
Frontal Orbital Cortex
Parahippocampal Gyrus, anterior division
Parahippocampal Gyrus, posterior division
Lingual Gyrus
Temporal Fusiform Cortex, posterior division
Temporal Occipital Fusiform Cortex
Occipital Fusiform Gyrus
Frontal Operculum Cortex
Central Opercular Cortex
Parietal Operculum Cortex
Planum Polare
Heschl's Gyrus (includes H1 and H2)
Planum Temporale
Supracalcarine Cortex
Occipital Pole
ROIs

(e)


Control group mean CBF distribution across ROIs (PAOCSL outlier cleaning)
mean CBF value
100
50
0
−50
−100
ROIs
Frontal Pole
Insular Cortex
Superior Frontal Gyrus
Middle Frontal Gyrus
Inferior Frontal Gyrus, pars triangularis
Inferior Frontal Gyrus, pars opercularis
Precentral Gyrus
Temporal Pole
Superior Temporal Gyrus, anterior division
Superior Temporal Gyrus, posterior division
Middle Temporal Gyrus, anterior division
Middle Temporal Gyrus, posterior division
Middle Temporal Gyrus, temporooccipital part
Inferior Temporal Gyrus, posterior division
Inferior Temporal Gyrus, temporooccipital part
Postcentral Gyrus
Superior Parietal Lobule
Supramarginal Gyrus, anterior division
Supramarginal Gyrus, posterior division
Angular Gyrus
Lateral Occipital Cortex, superior division
Lateral Occipital Cortex, inferior division
Intracalcarine Cortex
Frontal Medial Cortex
Juxtapositional Lobule Cortex (formerly Supplementary Motor Cortex)
Subcallosal Cortex
Paracingulate Gyrus
Cingulate Gyrus, anterior division
Cingulate Gyrus, posterior division
Precuneous Cortex
Cuneal Cortex
Frontal Orbital Cortex
Parahippocampal Gyrus, anterior division
Parahippocampal Gyrus, posterior division
Lingual Gyrus
Temporal Fusiform Cortex, posterior division
Temporal Occipital Fusiform Cortex
Occipital Fusiform Gyrus
Frontal Opercullum Cortex
Central Opercular Cortex
Parietal Operculum Cortex
Planum Polare
Heschl's Gyrus (includes H1 and H2)
Planum Temporale
Supracalcarine Cortex
Occipital Pole

(f)

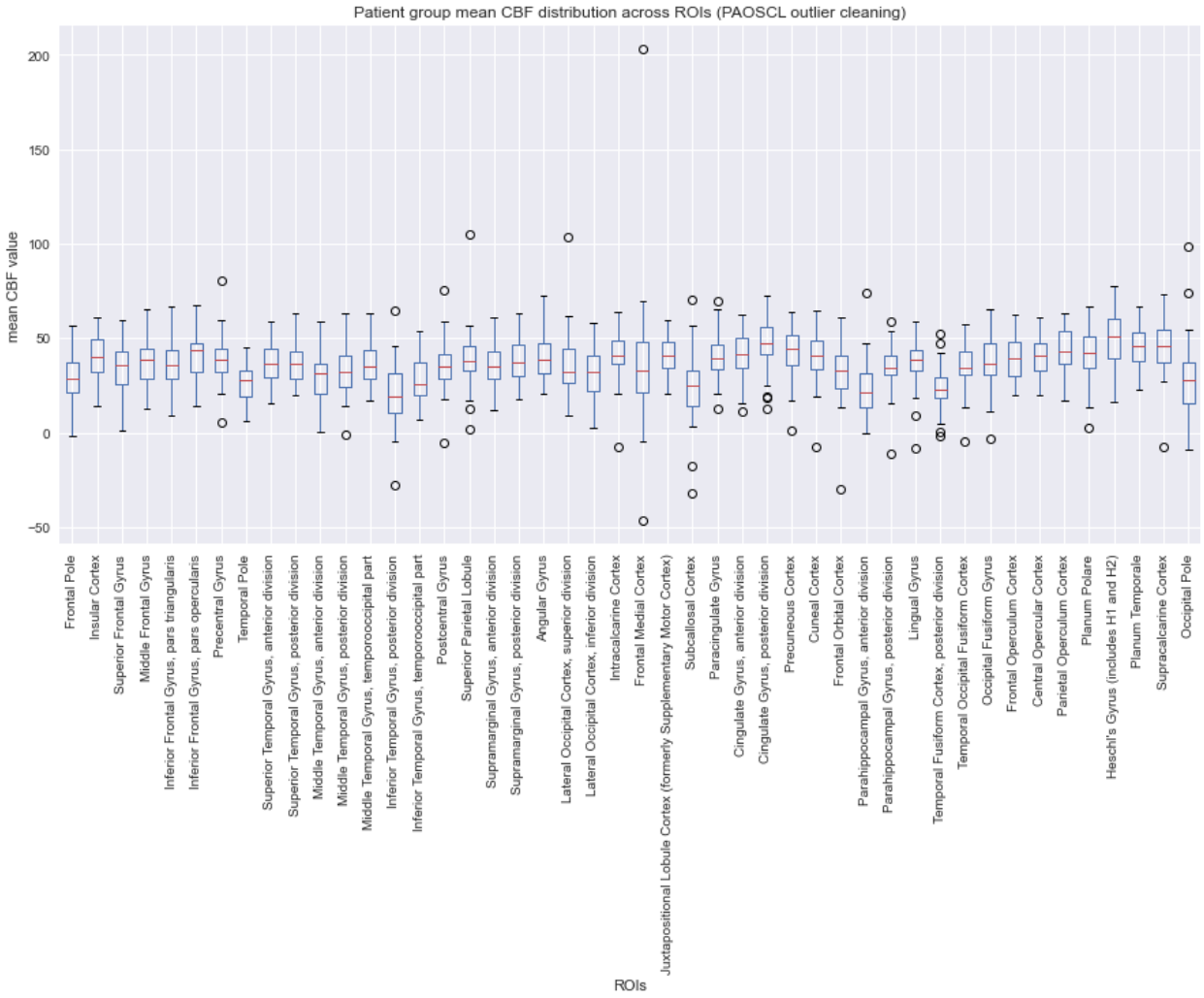
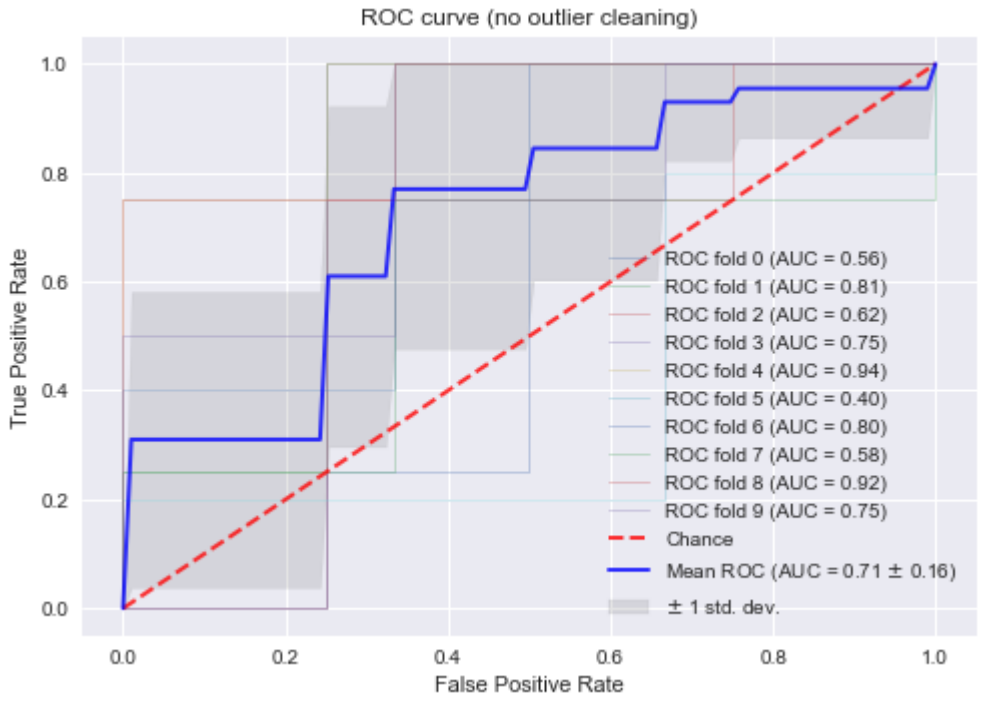
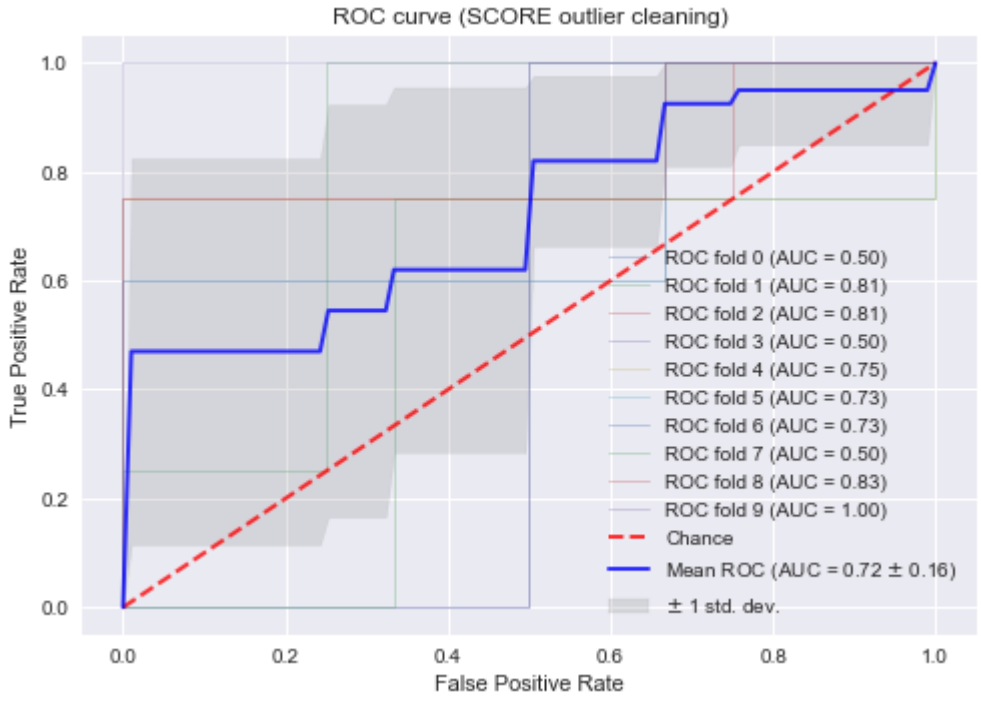

Patient group mean CBF distribution across ROIs (PAOSCL outlier cleaning)
mean CBF value
200
150
100
50
0
−50
Frontal Pole
Insular Cortex
Superior Frontal Gyrus
Middle Frontal Gyrus
Inferior Frontal Gyrus, pars triangularis
Inferior Frontal Gyrus, pars opercularis
Precentral Gyrus
Temporal Pole
Superior Temporal Gyrus, anterior division
Superior Temporal Gyrus, posterior division
Middle Temporal Gyrus, anterior division
Middle Temporal Gyrus, posterior division
Middle Temporal Gyrus, temporooccipital part
Inferior Temporal Gyrus, posterior division
Inferior Temporal Gyrus, temporooccipital part
Postcentral Gyrus
Superior Parietal Lobule
Supramarginal Gyrus, anterior division
Supramarginal Gyrus, posterior division
Angular Gyrus
Lateral Occipital Cortex, superior division
Lateral Occipital Cortex, inferior division
Intracalcarine Cortex
Frontal Medial Cortex
Juxtapositional Lobule Cortex (formerly Supplementary Motor Cortex)
Subcallosal Cortex
Paracingulate Gyrus
Cingulate Gyrus, anterior division
Cingulate Gyrus, posterior division
Precuneous Cortex
Cuneal Cortex
Frontal Orbital Cortex
Parahippocampal Gyrus, anterior division
Parahippocampal Gyrus, posterior division
Lingual Gyrus
Temporal Fusiform Cortex, posterior division
Temporal Occipital Fusiform Cortex
Occipital Fusiform Gyrus
Frontal Operculum Cortex
Central Opercular Cortex
Parietal Operculum Cortex
Planum Polare
Heschl's Gyrus (includes H1 and H2)
Planum Temporale
Supracalcarine Cortex
Occipital Pole
ROIs

Figure 3. ROC curves evaluating classifier performance.
ROC curves for a single 10-fold cross validation, linear kernel SVM classifier trained on data with (a) no outlier cleaning, (b) SCORE outlier cleaning, and (c) PAOCSL outlier cleaning.

(a)



(b)


ROC curve (SCORE outlier cleaning)
True Positive Rate
False Positive Rate
0.0
0.2
0.4
0.6
0.8
1.0
ROC fold 0 (AUC = 0.50)
ROC fold 1 (AUC = 0.81)
ROC fold 2 (AUC = 0.81)
ROC fold 3 (AUC = 0.50)
ROC fold 4 (AUC = 0.75)
ROC fold 5 (AUC = 0.73)
ROC fold 6 (AUC = 0.73)
ROC fold 7 (AUC = 0.50)
ROC fold 8 (AUC = 0.83)
ROC fold 9 (AUC = 1.00)
Chance
Mean ROC (AUC = 0.72 ± 0.16)
± 1 std. dev.

(c)

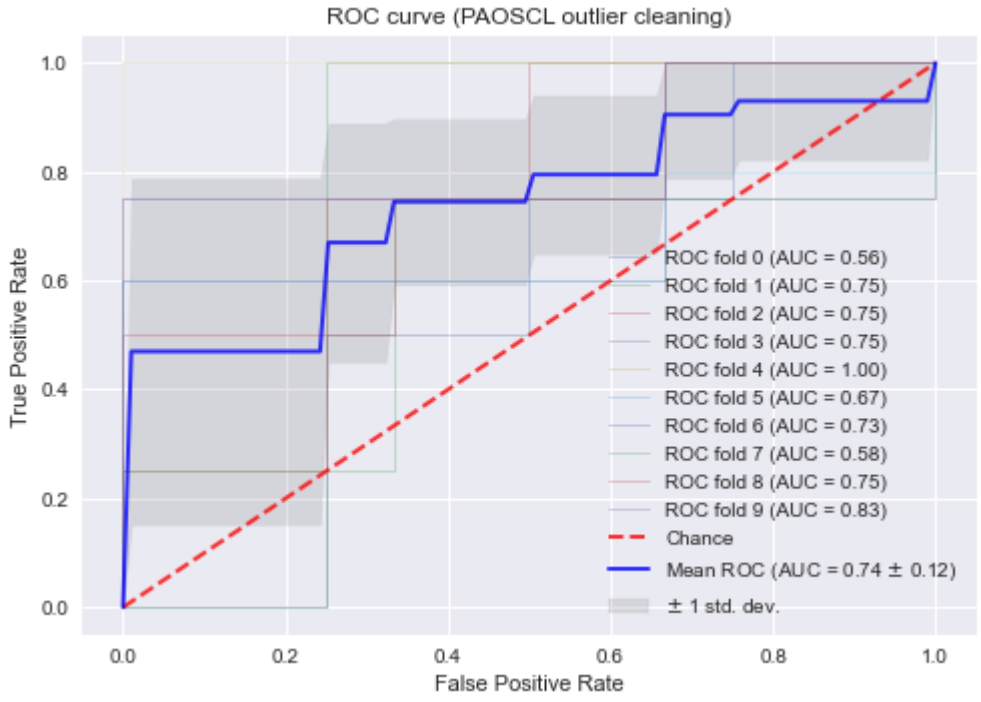
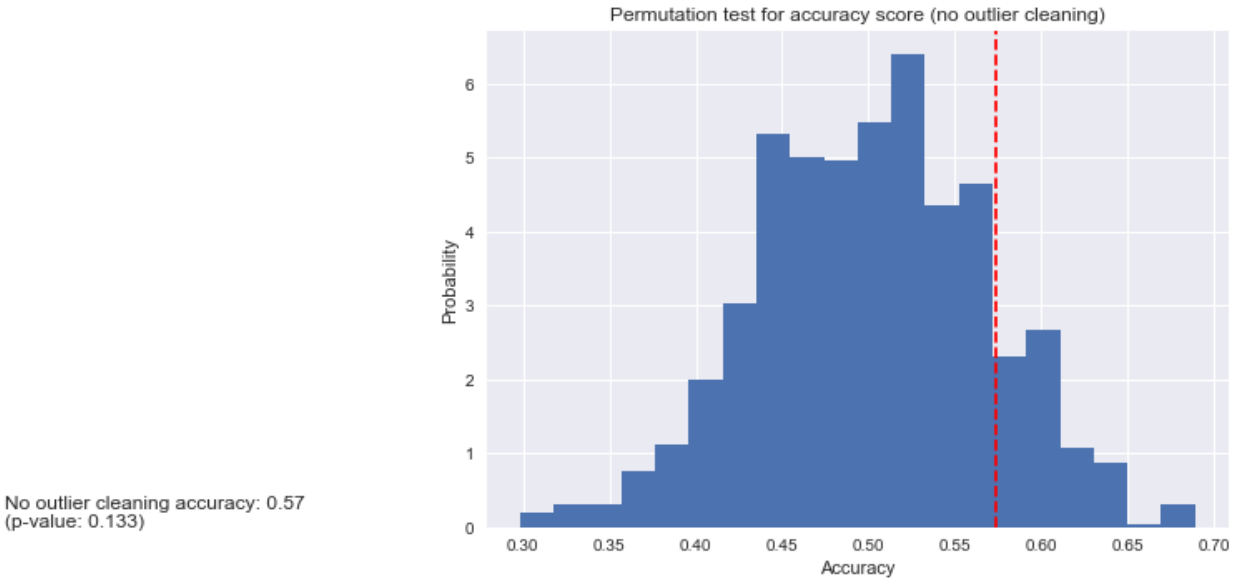
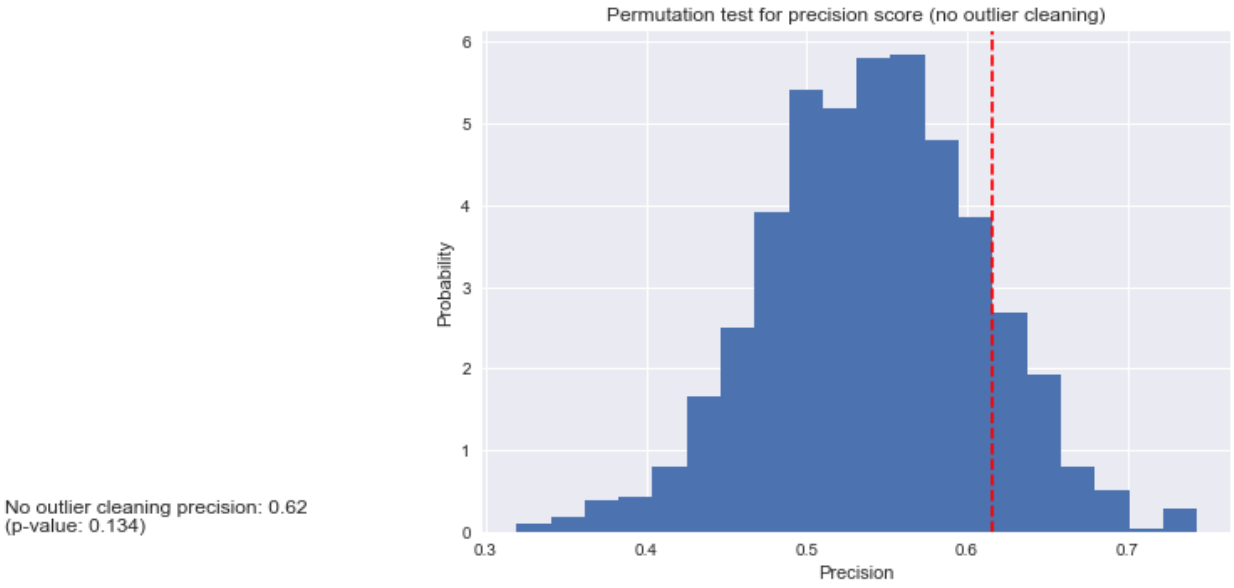

ROC curve (PAOSCL outlier cleaning)
True Positive Rate
False Positive Rate
0.0
0.2
0.4
0.6
0.8
1.0
ROC fold 0 (AUC = 0.56)
ROC fold 1 (AUC = 0.75)
ROC fold 2 (AUC = 0.75)
ROC fold 3 (AUC = 0.75)
ROC fold 4 (AUC = 1.00)
ROC fold 5 (AUC = 0.67)
ROC fold 6 (AUC = 0.73)
ROC fold 7 (AUC = 0.58)
ROC fold 8 (AUC = 0.75)
ROC fold 9 (AUC = 0.83)
Chance
Mean ROC (AUC = 0.74 ± 0.12)
± 1 std. dev.

Figure 4. Permutation-based *p*-value tests for accuracy, precision, and recall. Permutation-based *p*-value tests for (a) accuracy, (b) precision and (c) recall with no outlier cleaning, (d) accuracy, (e) precision and (f) recall with SCORE outlier cleaning, (g) accuracy, (h) precision and (i) recall with PAOCSL outlier cleaning. The red line represents the mean performance metric over 1,000 iterations of model training on the unpermuted data. The blue histogram represents the null distribution for each performance metric over 1,000 iterations of model training on permuted data.

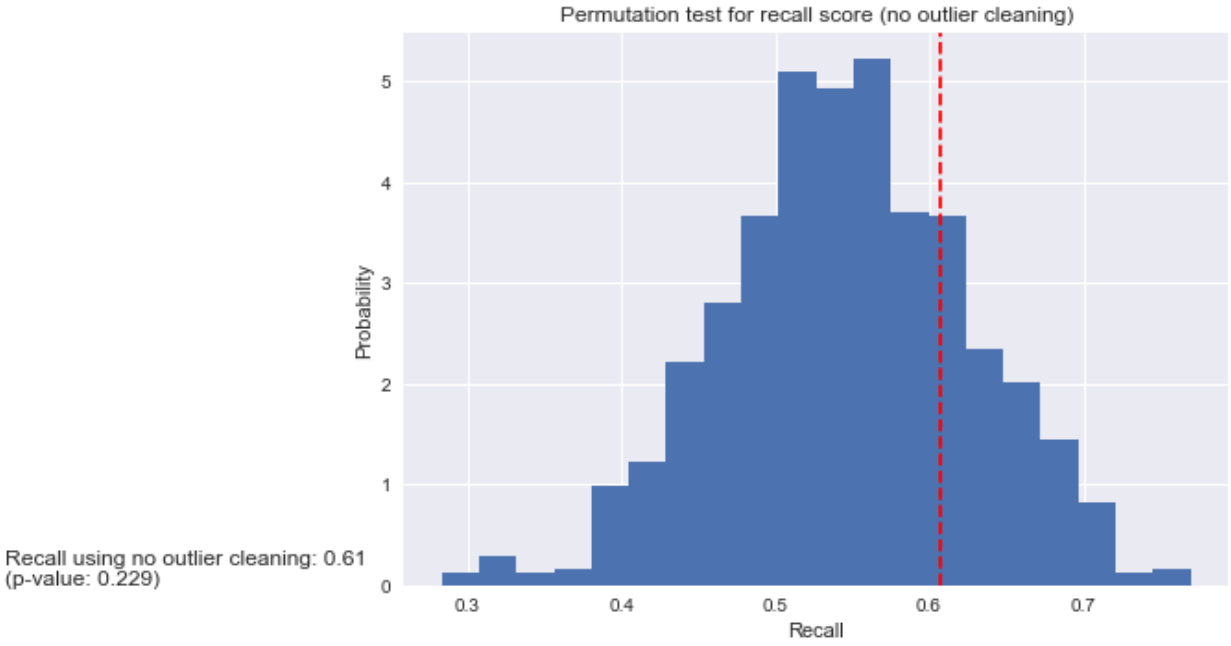
(a)



(b)

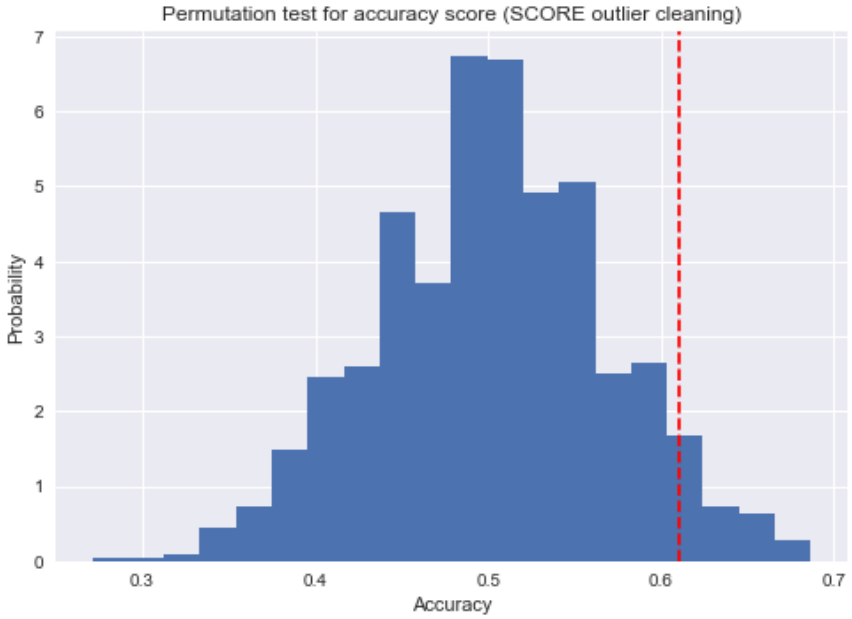


(c)


Permutation test for recall score (no outlier cleaning)
Probability
Recall
0
1
2
3
4
5
0.3
0.4
0.5
0.6
0.7
Recall using no outlier cleaning: 0.61
(p-value: 0.229)

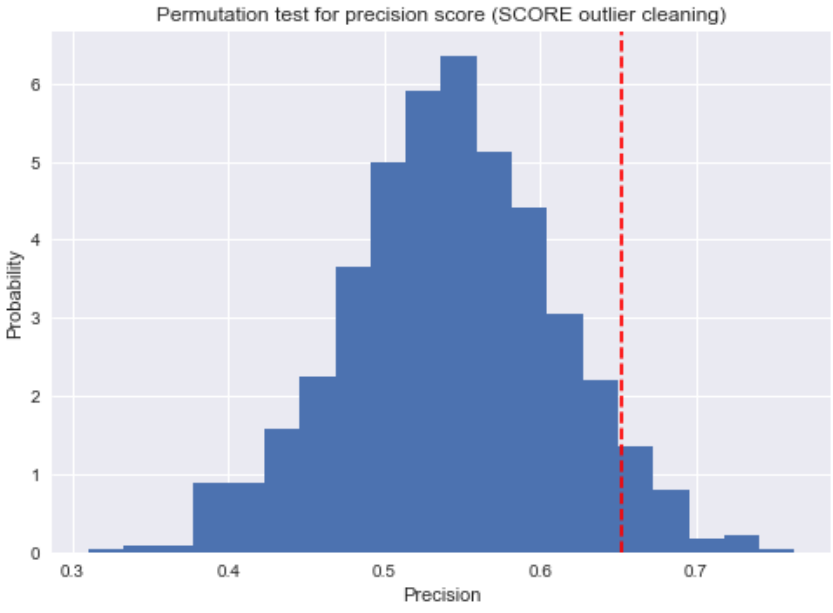
(d)

SCORE outlier cleaning accuracy: 0.61
(p-value: 0.060)



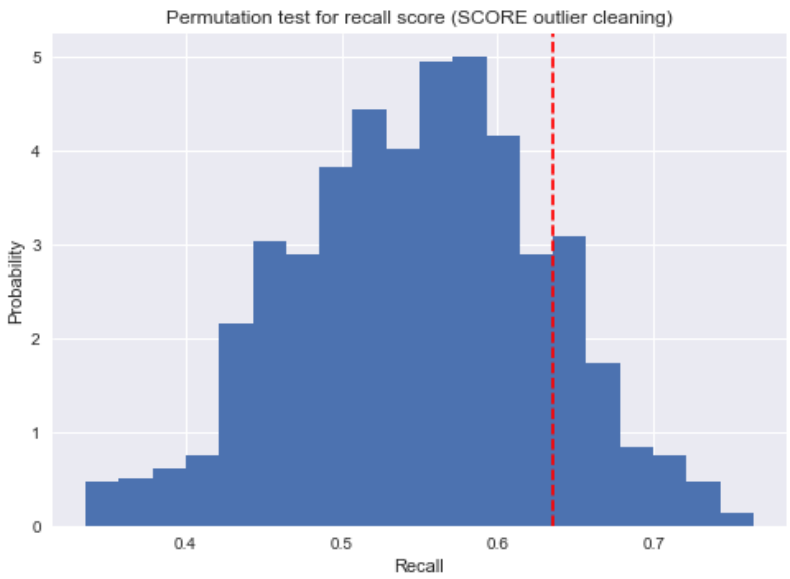
(e)

SCORE outlier cleaning precision: 0.65
(p-value: 0.057)

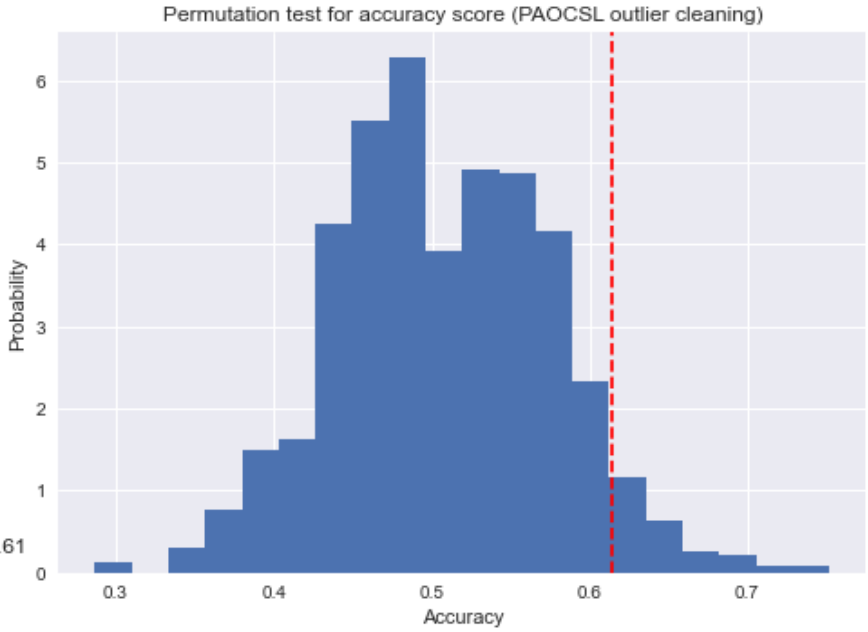


(f)

SCORE outlier cleaning recall: 0.64
(p-value: 0.151)

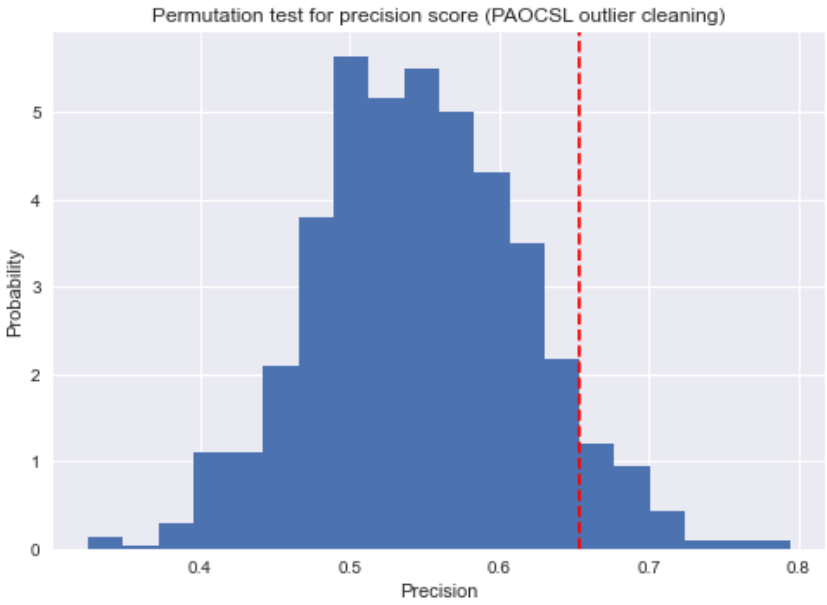


(g)



PAOCSL outlier cleaning accuracy: 0.61
(p-value: 0.055)

(h)



PAOCSL outlier cleaning precision: 0.65
(p-value: 0.067)

(i)

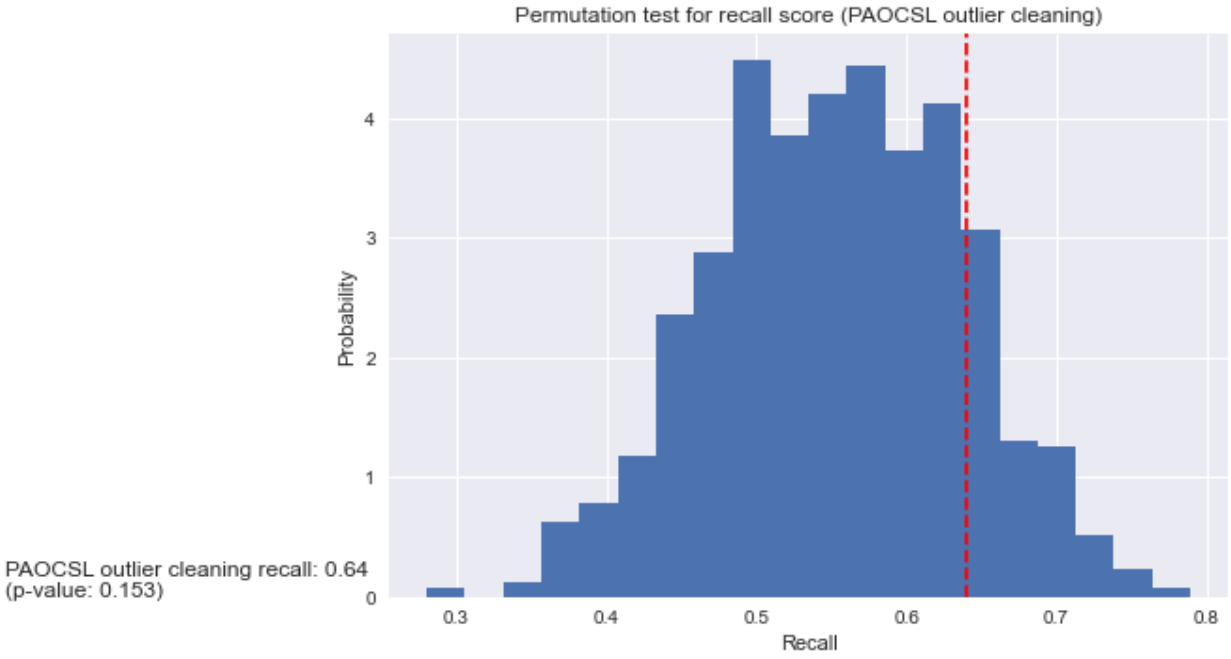
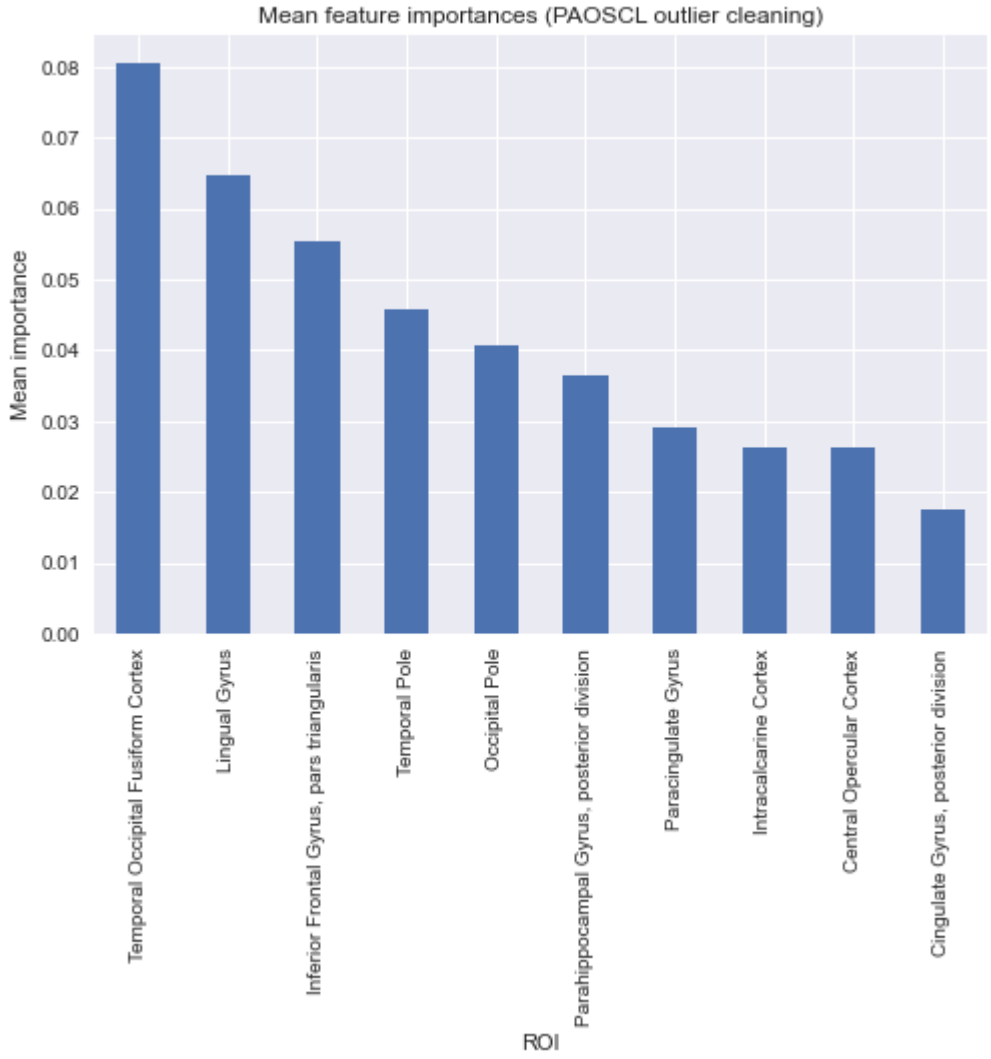
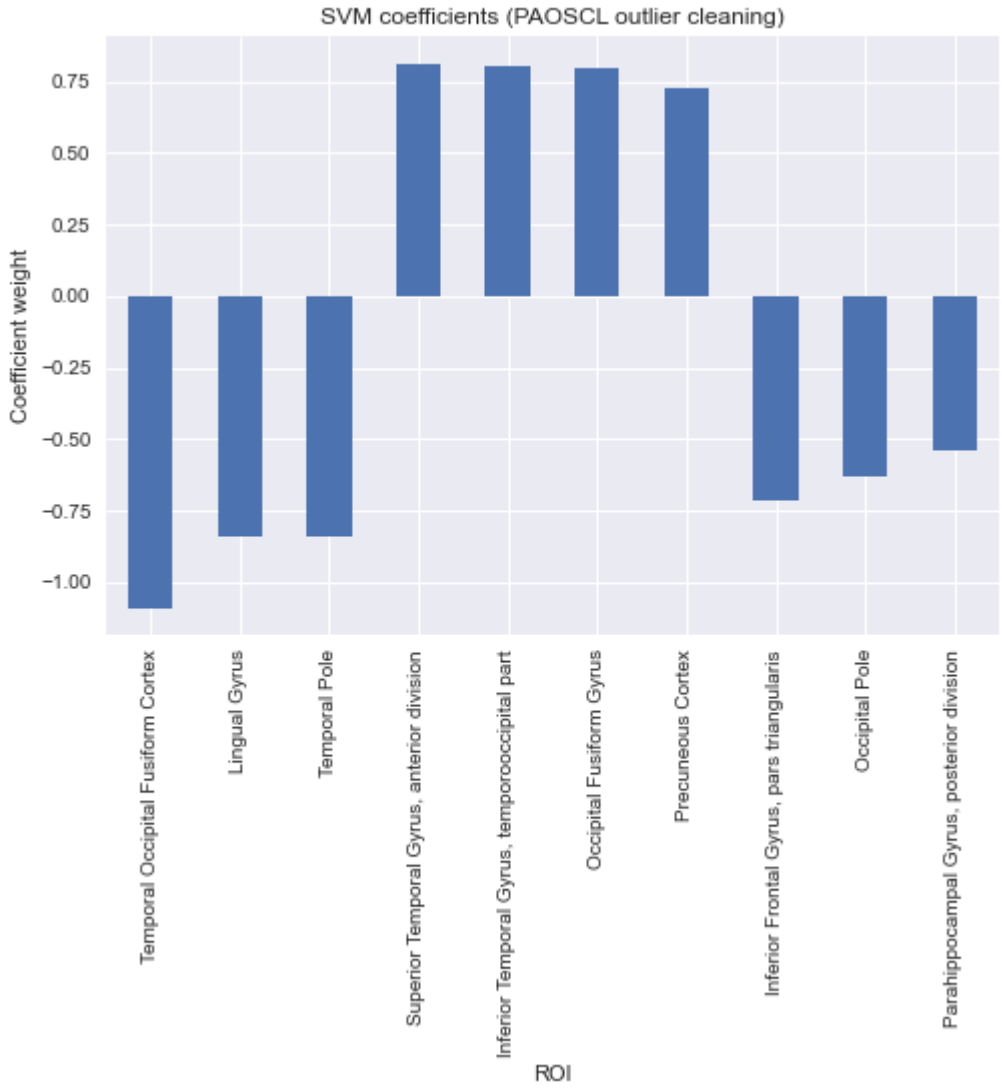

Permutation test for recall score (PAOCSL outlier cleaning)
PAOCSL outlier cleaning recall: 0.64
(p-value: 0.153)
Probability
Recall
0
1
2
3
4
0.3
0.4
0.5
0.6
0.7
0.8

Figure 5. Most predictive ROIs.
Top ten most predictive ROIs returned from the classifier trained on PAOCSL data determined by (a) the permutation importance method and (b) the SVM coefficients (a positive sign contributes to the prediction of TBI patients, whereas a negative sign contributes to the prediction of control subjects).

(a)



(b)


SVM coefficients (PAOSCL outlier cleaning)
Coefficient weight
0.75
0.50
0.25
0.00
−0.25
−0.50
−0.75
−1.00
Temporal Occipital Fusiform Cortex
Lingual Gyrus
Temporal Pole
Superior Temporal Gyrus, anterior division
Inferior Temporal Gyrus, temporooccipital part
Occipital Fusiform Gyrus
Precuneous Cortex
Inferior Frontal Gyrus, pars triangularis
Occipital Pole
Parahippocampal Gyrus, posterior division
ROI

4. Discussion

The current study's main findings show that an SVM classifier trained on ASL-generated CBF maps cannot significantly differentiate between TBI patients versus healthy controls. Moreover, there was no obvious difference in performances across the SVM classifiers trained on no outlier cleaned, SCORE outlier cleaned, or PAOCSL outlier cleaned CBF data. Nevertheless, the overall analysis does provide possible indications there may be discriminative power in a machine learning model trained on CBF data as the models trained on SCORE and PAOCSL outlier cleaned data had higher than chance performance that approached significance.

Several methods of assessing classifier performance were examined to evaluate the discriminative power of each classifier. A preliminary investigation of the SVM performances showed a single iteration of a classifier trained on the PAOCSL outlier cleaned data performed best on 10-fold cross validation, with a mean AUC of 0.74 (Figure 3c). The AUC is a measure of the true positive rate as compared to the false positive rate and thus provides information on the classifier's degree of misclassifications (Rakotomamonjy, 2004). A classifier correctly classifies all samples when AUC is equal to 1 and a classifier that predicts at random has an AUC of 0.5 (Rakotomamonjy, 2004), hence an AUC of 0.74, as is the case of the SVM trained on the PAOCSL outlier cleaned data, is reasonably better than chance performance. Both the additional SVMs trained on no outlier cleaned data, as well as the SCORE outlier cleaned data, also performed better than chance, with AUCs of 0.71 and 0.72, respectively (Figures 3a & 3b). Overton, et al. (2020) report similar "excellent" AUC findings in their study using a 10-fold cross-validated SVM trained on ASL-generated CBF data in psychosis spectrum classification in youths.

Upon deeper examination, however, a noticeable variance in model accuracies across the 10-fold cross validation is apparent. This may be due in part to sample size, as previous research in machine learning validation has questioned the validity of K-fold cross validation, reporting higher

variability in classifier performances trained on small sample sizes (Vabalas et al., 2019). Furthermore, even if a classifier has good performance, it does not mean the data has meaningful signal (Ojala & Garriga, 2010). Therefore, it was imperative to validate the performance of each classifier with an approach that ascertains whether the classifiers have learned any significant predictive patterns in the data (Ojala & Garriga, 2010).

A permutation-based *p*-value test was used to assess if each of the three classifier performances significantly differed when features and classes were independent, that is, when there is no difference between TBI patients and healthy controls in terms of CBF, versus when they are dependent. The results of the permutation-based *p*-value test revealed a near, but not quite significant, highest mean accuracy of 0.61 (*p*-value = 0.055) for the SVM classifier trained on the PAOCSL outlier cleaned data. While this performance appears better than an SVM classifier trained on data with no outlier cleaning (mean accuracy = 0.58, *p*-value = 0.130), it is not as apparent whether the SVM classifier trained on PAOCSL outlier cleaned data outperformed the SVM classifier trained on SCORE outlier cleaned data (mean accuracy = 0.61, *p*-value = 0.06), given the accuracies are the same, with the SCORE classifier having a slightly higher *p*-value. A more sophisticated statistical approach, such as ANOVA, would theoretically be required to make such a claim, assuming the classifiers had yielded statistically significant results from the permutation-based *p*-value classifier performance test. An additional caveat, even in the case that the results had been statistically significant, is that the mean precision was greater than both the mean accuracy and mean recall for each of the three SVM classifiers. It's important to note that precision is the proportion of predicted positive cases (TBI patients in the case of this study) that are correctly identified. While a relevant metric, it is arguably not the primary metric used in medical contexts (Powers, 2020). Furthermore, accuracy is also not always a good performance metric as it is sensitive to unbalanced data (Rakotomamonjy, 2004). Given the aim of the classifier

is to identify TBI patients from healthy control subjects to potentially target them for TBI treatment modalities, a more useful metric of model performance would prioritize recall (sensitivity) over accuracy, precision, and likewise, specificity, allowing the classifier to err on the side of catching more TBI patients rather than falsely classify them as healthy.

An additional aim of this study was to identify the most important ROIs when classifying TBI patients from healthy control subjects. Insofar as the results of the SVM classification approached significance, this investigation into ROI importance was important because it will lead to additional research on what approaches will improve classification. Since the SVM model trained on PAOCSL data reached a near significant mean accuracy, this model was selected to investigate the ROIs most predictive of TBI via a permutation feature importance method. These results were then compared to the SVM coefficients which can also be used as a proxy of feature importance, given a linear kernel (Shah et al., 2016). This revealed several overlapping ROIs between the two methods, including the temporal occipital fusiform cortex, the lingual gyrus, the temporal pole, the inferior frontal gyrus (pars triangularis), the occipital pole, and the parahippocampal gyrus (posterior division).

These ROIs have thought-provoking functions that may shed light on how they relate to TBI outcomes. For example, the temporal occipital fusiform cortex is a critical area for visual recognition, relating especially to faces (Kamps et al., 2019). The function of the lingual gyrus is also related to visual recognition, in particular to word processing (Mechelli et al., 2000). The temporal pole has been indicated in social and emotional processing (Olson et al., 2007). The inferior frontal gyrus, specifically the pars triangularis portion in the left hemisphere, is part of Broca's area and is important in speech production (Foundas et al., 1996) and grammatical processing. The occipital pole plays an important role in the representation of the field of vision in the visual cortex (McFadzean et al., 1994). Lastly, the posterior division of the parahippocampal

gyrus has been shown to be involved in memory-related neural structures in the human brain (Thangavel et al., 2008). Previous research using the same participant pool as this current study indicated the TBI patients had reduced performance on several neuropsychological measures in cognitive domains such as mental processing speed, verbal learning, and executive functioning (Ware et al., 2020). While it's important to be cautious in proposing only a speculative connection of the important ROIs to the previously mentioned cognitive domains, the functionality of the most important ROIs returned from both the feature importance permutation method and SVM coefficient analysis generally coincide with cognitive deficits observed in the TBI patient group. Other past research has also linked sustained-attention and working-memory deficits in TBI patients with superior occipital cortex and superior temporal cortex hypoperfusion (Kim et al., 2012), both regions sharing overlap with the important ROIs found in this study. Yet another study found that CBF was reduced in related areas in the temporal and limbic lobes in early stage mild TBI and was correlated with cognitive impairment associated with TBI (Peng et al., 2016). The feature importance permutation method and the analysis of the SVM coefficients in this current study also revealed additional important ROIs, such as the paracingulate gyrus and the precuneus cortex, respectively. Areas in the cingulate cortex, as well as the precuneus, have likewise been shown to have hypoperfusion in TBI patients (Kim et al., 2010). Moreover, other research also hypothesizes that TBI causes neuropathological alterations in the brain that are similar to those seen in Alzheimer's disease (Peng et al., 2016) and another ASL imaging study has shown hypoperfusion in the posterior cingulate cortex as well as the precuneus in patients who later develop Alzheimer's disease (Xekardaki et al., 2015). Taken overall, the findings from past research into the neuroanatomical changes seen in CBF as it relates to TBI and the subsequent impacts on cognition, are in line with the most important ROIs found in this study.

4.1. Limitations of the current study

The current study has some limitations, several of which may relate to the only near significant findings. As such, this study should be considered an initial exploration of potential future work in the yet uncharted field of machine learning in detecting TBI neuropathology and predicting TBI outcomes using CBF data.

While the small sample size in this study is certainly an important limitation in terms of generalizability, it likely is not a main factor in the only slightly better than chance SVM classifier performances observed here. In fact, Abdelrahman et al. (2020) obtained a 90.5% accuracy score using a five-fold cross validation SVM classifier, distinguishing TBI patients from healthy controls using DTI data. They further validated their results on a 1,000 iteration permutation test using only a sample size of 52 participants. Furthermore, Vabalas et al. (2019) found that small sample size is actually associated with higher reported classification accuracy of several neuroimaging studies using machine learning. As mentioned previously, this could be due to biased performance estimates inherent in the K-fold cross validation methodology. Hence, several works recommend using permutation-based *p*-values for assessing the robustness of a classifier (Golland & Fischl, 2003; Golland et al., 2005; Hsing et al., 2003; Molinaro et al., 2005). As Ojala and Garriga (2010) point out, this permutation test ascertains whether there is a predictive underlying structure (or signal) in the data and whether the classifier can learn this structure. However, they also point out that while a high *p*-value can suggest that the data contains no structure, it may also suggest that the classifier was simply unable to effectively use the structure in the data. This idea highlights two possible limitations of the current study — the data may contain no significant discriminative signal in predicting TBI patients, and/or the SVM classifier, due to algorithmic considerations, cannot discern any discriminative signal from the CBF data.

In the case that the data contains no signal, there are several considerations to take into

account. First and foremost, CBF may intrinsically not have any meaningful signal that predicts TBI. This seems unlikely, however, given the large body of research and literature that supports that TBI is often characterized by decreases in CBF (Graves & Kreipke, 2015). Another consideration may be the dimensionality reduction and feature selection techniques used in this study. An often-standard approach in this regard, given the voxel-based and therefore high variable count nature of neuroimaging data, is to use prior neuroscientific knowledge, such as probabilistic atlases like the Harvard-Oxford atlas used here, to create ROIs (Craddock et al., 2012). However, several other brain parcellation atlases exist that are, perhaps, better suited to represent the CBF signal in the data used in this study. In addition to atlas-derived methods to define ROI-based features, data-driven methods such as Principal Component Analysis (PCA) collapse the data in such a way to create fewer, but more informative, variables (i.e., features). While such an approach may better represent the structure in data, the resulting features from data-driven methods can be harder to interpret when compared to more transparent ROI analyses (Smith & Nichols, 2018). Another caveat pertinent to the realm of feature selection concerns the methods of ROI feature importance mentioned previously — a limitation exists in that the important features identified are only important to the SVM classifier used and do not speak to the intrinsic predictive power of any particular feature by itself (Scikit-learn developers, 2020c). That is to say, depending on the machine learning algorithm used, a different set of feature importances may result. Furthermore, it's relevant to note two ROIs, the temporal fusiform cortex (anterior division) and the inferior temporal gyrus (anterior division), were discarded from analysis in this study due to high null values and therefore lack of data for those regions. It's possible CBF data from those regions may have had predictive power in the classification of TBI patients versus healthy controls.

Yet another limitation related to the CBF data relates to data quality. The ASL technique used in this study is more susceptible to artifacts than background-suppressed ASL techniques that

are now becoming more widely adopted (Ware et al., 2020). These artifacts, as well as motion artifacts, and ultimately the very low SNR inherent in ASL imaging, may have obscured any clear signal in the resulting CBF data (Li et al., 2018). It's also possible the outlier cleaning algorithms examined in this study, SCORE and PAOCSL, are insufficient at detecting outliers in this data. For example, the SCORE outlier cleaning algorithm operates by removing entire volumes of data which could result in significant data loss and accordingly may affect any predictive signal in the CBF maps. On the other hand, while the PAOCSL outlier cleaning algorithm retains more data, it relies on a reference CBF map based on prior knowledge about CBF distributions; this method may make assumptions not appropriate for this particular data.

A final limitation to consider is the usage of an SVM in the machine learning component of this study. While SVM models have become quite widely used in neuroimaging studies, such as those using fMRI data (Gaonkar & Davatzikos, 2013), a literature search reveals the use of SVM with CBF data is less common. As mentioned previously, the only slightly better-than-chance SVM classifier performances observed here may be due to the classifier not being able to learn from the structure in the data. In particular, the SVM algorithm operates by determining the maximum margin to separate classes by finding the optimal hyperplane (Shah et al., 2016). In addition to the features, SVMs include a hyperparameter C which influences the hyperplane margin and therefore misclassification rate (Gaonkar & Davatzikos, 2013). A limitation of this study is this hyperparameter was left at the default scikit-learn value of 1.0 and perhaps hyperparameter tuning would improve the SVM performance. Furthermore, a linear kernel was implemented for sake of interpretability of the coefficients, and this assumes a linear relationship in the CBF data (Shah et al., 2016). However, it is possible the predictive signal in CBF data is simply not linear.

4.2. Future studies

There are several promising avenues to consider for future research as it relates to this

study. A meaningful area for further study includes extending the machine learning methodology here to examine longitudinal CBF changes in TBI patients at 6 and 12 months post-injury, in addition to the 3 months post-injury time frame under investigation in this study. This would involve greater classifier complexity, however understanding ramifications of long-term CBF changes for TBI patient outcomes would allow for more specific, targeted treatment plans for individuals with longer term cognitive impairment prognoses.

The results of this study could also be expanded on by combining different neuroimaging modalities to improve classification performance of TBI patients from healthy controls. For example, ASL CBF maps could be combined with DTI white matter structural data, and/or MWI myelin data, to generate more robust data. Another group has in fact combined the application of DTI, Diffusion Kurtosis Imaging (DKI), and ASL data to predict mild TBI patient outcomes post-injury, and while they did not employ any machine learning techniques, their results showed potential enhancement in the identification of patients with mild TBI (Grossman et al., 2013). With regard to a multimodal neuroimaging machine learning precedence, albeit not in the field of TBI research, Overton et al. (2020) have successfully trained an SVM on a concatenated data set of both ASL and BOLD fMRI data in predicting psychosis spectrum in youths.

Another particularly exciting area of future work is the application of deep learning. Deep learning is an emerging machine learning technique gaining traction in analyzing high dimensional spaces characteristic in neuroimaging data (Thomas et al., 2019). The basic premise behind deep learning is that it leverages representation-learning methods, with the input data, and each subsequent representation layer, being transformed by simple, yet non-linear, functions, allowing deep learning methods to learn complex patterns in the data (Thomas et al., 2019). An especially useful aspect of deep learning is that it can also be used to identify signals from noisy data, transforming the data in such a way that the irrelevant signal is reduced, and the relevant signal is

intensified, thereby detecting latent relationships in the data (Thomas et al., 2019). In fact, research into using a convolutional neural network (CNN), a popular deep learning architecture used in medical imaging processing, to denoise ASL data, has shown improved SNR measures compared to current methods (Xie et al., 2020). Another area of potential future research involves using deep learning for anatomical segmentation and cortical parcellation, having the ability to generate ROIs from within the structure of the data while mapping to known whole-brain regions (Henschel et al., 2020). Moreover, deep learning can be leveraged for automatic feature learning from data and classification at the same time, which Plis et al. (2014) have successfully demonstrated, classifying both schizophrenic and Huntington disease patients using a deep belief network trained on MRI data. In summary, deep learning may be a potential solution to several of the limitations faced by this study.

5. Conclusion

Previous research has shown that CBF may provide a quantitative measure of structural and functional changes following TBI (Kim et al., 2010); however, the specific relationship between CBF measures and TBI remains ambiguous (Ware et al., 2020). Hence more research is necessary for meaningful understanding of changes in CBF as it relates to TBI. This study investigated the predictive performance of several SVM models trained on CBF data. The principal aim was to classify TBI patients from healthy control subjects and consequently identify the most predictive ROIs as they relate to CBF and TBI. Given the well-established low SNR in ASL-generated CBF maps (Dolui et al., 2017, Li et al., 2018), the CBF data was approached with three different outlier cleaning methods: no outlier cleaning, as well as two additional, sophisticated adaptive algorithms, SCORE and PAOCSL. Since previous work has shown the PAOCSL algorithm to have the greatest increase in SNR as compared to the SCORE algorithm (Li et al., 2018), an important additional

aim was to evaluate the performance of an SVM classifier trained on PAOCSL cleaned CBF data. While none of the SVM classifiers were shown to perform significantly better than random guessing, using more sophisticated outlier cleaning methods tended to improve the classification accuracy. The results of this study may be used to guide more pointed future research into the use of machine learning to predict TBI from ASL-generated CBF data.

Appendix

Appendix Table 1. ROIs by total voxel size.
In addition to total voxels for each ROI, mean CBF differences between control and patient subjects are shown for each outlier cleaning method.

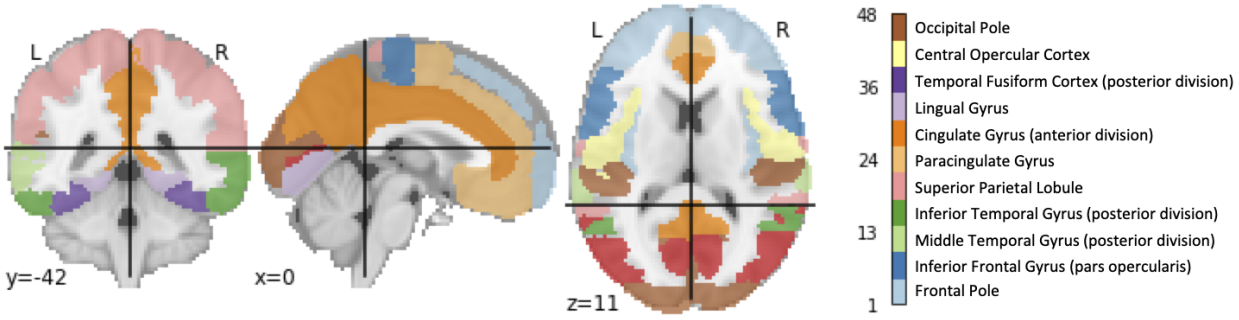
Control – Patient

ROIs	***Total Voxels***	***No outlier cleaning***	***SCORE***	***PAOCSL***
Frontal Pole	15,397	10.58	4.70	6.13
Lateral Occipital Cortex, superior division	9,779	0.84	-0.96	-0.22
Precentral Gyrus	8,573	1.20	0.75	1.53
Postcentral Gyrus	6,895	-0.15	-0.57	0.68
Precuneous Cortex	5,623	3.96	1.49	3.54
Middle Frontal Gyrus	5,316	5.05	5.32	5.59
Occipital Pole	5,276	9.00	3.70	5.20
Superior Frontal Gyrus	5,080	3.11	3.80	3.84
Temporal Pole	4,711	6.41	4.45	5.27
Lateral Occipital Cortex, inferior division	4,089	6.37	3.82	4.56
Lingual Gyrus	3,381	10.50	7.95	9.34
Frontal Orbital Cortex	3,148	5.85	5.13	7.02
Paracingulate Gyrus	2,944	8.66	6.80	7.69

Superior Parietal Lobule	2,908	-0.79	4.02	-0.12
Cingulate Gyrus, anterior division	2,592	8.89	7.10	7.90
Middle Temporal Gyrus, posterior division	2,525	3.95	3.71	4.37
Cingulate Gyrus, posterior division	2,412	7.88	5.17	7.34
Angular Gyrus	2,409	4.24	2.89	3.56
Insular Cortex	2,341	9.19	8.07	8.60
Supramarginal Gyrus, posterior division	2,259	3.58	2.38	3.17
Middle Temporal Gyrus, temporooccipital part	2,004	5.45	4.29	4.88
Inferior Temporal Gyrus, posterior division	1,942	2.40	2.20	3.31
Central Opercular Cortex	1,886	7.28	6.21	7.14
Superior Temporal Gyrus, posterior division	1,830	4.89	3.75	4.89
Occipital Fusiform Gyrus	1,806	6.57	5.51	6.87
Supramarginal Gyrus, anterior division	1,742	2.34	1.03	2.43
Temporal Fusiform Cortex, posterior division	1,594	2.26	1.11	3.64
Juxtapositional Lobule Cortex (formerly Supplementary Motor Cortex)	1,484	1.86	1.82	2.78
Inferior Temporal Gyrus, temporooccipital part	1,470	2.22	2.78	2.98
Temporal Occipital Fusiform Cortex	1,469	7.71	6.42	8.36
Intracalcarine Cortex	1,401	9.52	7.70	8.27

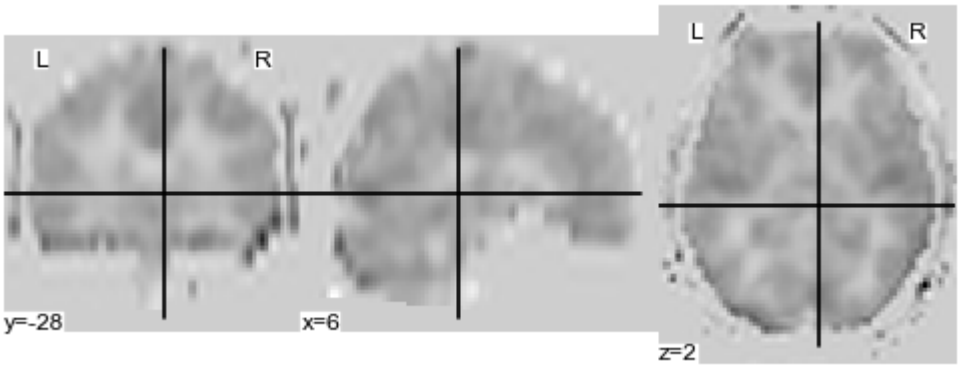
Inferior Frontal Gyrus, pars opercularis	1,384	6.77	5.62	8.90
Parahippocampal Gyrus, anterior division	1,248	0.73	-0.04	2.54
Cuneal Cortex	1,227	5.02	3.14	4.26
Subcallosal Cortex	1,142	6.11	5.11	6.36
Parietal Operculum Cortex	1,119	7.10	5.02	6.92
Inferior Frontal Gyrus, pars triangularis	1,103	9.22	5.62	6.29
Frontal Medial Cortex	976	5.05	2.19	4.50
Planum Temporale	952	6.61	4.54	6.62
Middle Temporal Gyrus, anterior division	848	4.05	2.91	4.03
Planum Polare	749	7.92	6.70	7.46
Parahippocampal Gyrus, posterior division	710	8.72	6.92	8.51
Frontal Operculum Cortex	687	8.32	7.94	8.30
Heschl's Gyrus (includes H1 and H2)	604	8.94	6.89	8.87
Superior Temporal Gyrus, anterior division	521	2.46	1.14	2.26
Supracalcarine Cortex	261	9.02	7.12	8.14
Discarded ROIs	Total Voxels	No outlier cleaning	SCORE	PAOCSL
Inferior Temporal Gyrus, anterior division	647	N/A	N/A	N/A
Temporal Fusiform Cortex, anterior division	610	N/A	N/A	N/A

Appendix Figure 1. Harvard-Oxford reference atlas used to generate ROIs.
Frontal, midsagittal, and horizontal views of the brain plotted with the visible ROIs in each view as generated using the Harvard-Oxford atlas in nilearn.

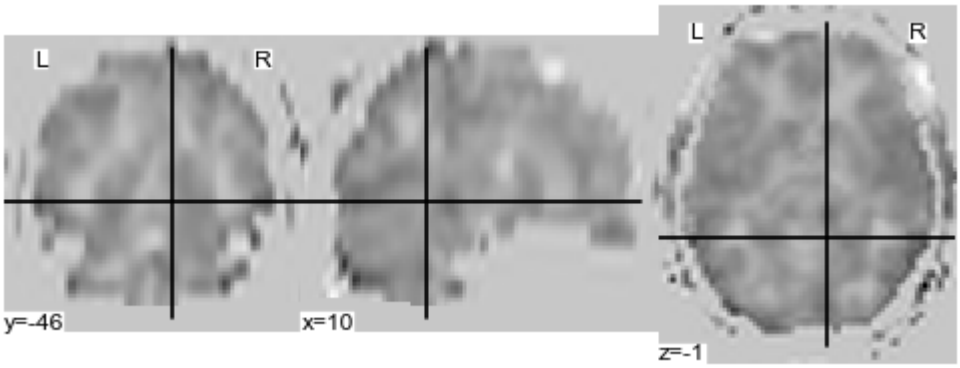


Appendix Figure 2. Example TBI patient misclassified as a control subject.
Frontal, sagittal, and horizontal views of a TBI patient misclassified as a healthy control subject by the SVM model trained on data from (a) no outlier cleaning, (b) SCORE outlier cleaning, and (c) PAOCSL outlier cleaning.

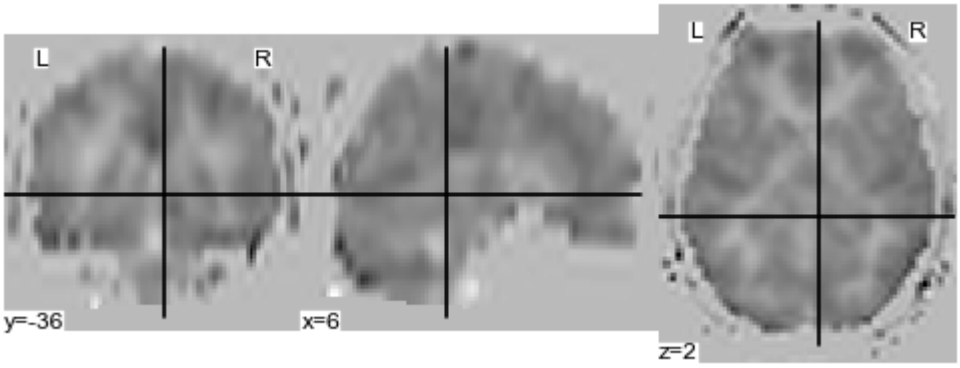
(a)



(b)



(c)



References

Abdelrahman, H. A. F., Ubukata, S., Ueda, K., Fujimoto, G., Oishi, N., Aso, T., & Murai, T. (2020). Machine learning classification of traumatic brain injury patients and healthy controls using multiple indices of diffusion tensor imaging. doi: 10.21203/rs.3.rs-42800/v1

Abraham, A., Pedregosa, F., Eickenberg, M., Gervais, P., Mueller, A., Kossaifi, J., ... & Varoquaux, G. (2014). Machine learning for neuroimaging with scikit-learn. *Frontiers in neuroinformatics, 8*, 14. doi: 10.3389/fninf.2014.00014

Aguirre, G. K., Detre, J. A., Zarahn, E., & Alsop, D. C. (2002). Experimental design and the relative sensitivity of BOLD and perfusion fMRI. *NeuroImage, 15(3)*, 488–500. doi: 10.1006/nimg.2001.0990

Akbani, R., Kwek, S., & Japkowicz, N. (2004, September). Applying support vector machines to imbalanced datasets. In *European conference on machine learning* (pp. 39-50). Springer, Berlin, Heidelberg. doi: 10.1007/978-3-540-30115-8_7

Alsop, D. C., & Detre, J. A. (1996). Reduced transit-time sensitivity in noninvasive magnetic resonance imaging of human cerebral blood flow. *Journal of Cerebral Blood Flow & Metabolism*, *16*(6), 1236-1249. doi: 10.1097/00004647-199611000-00019

Barclay, L., Zemcov, A., Reichert, W., & Blass, J. P. (1985). Cerebral blood flow decrements in chronic head injury syndrome. *Biological psychiatry, 20*(2), 146–157. doi: 10.1016/0006-3223(85)90074-5

Bentourkia, M., Bol, A., Ivanoiu, A., Labar, D., Sibomana, M., Coppens, A., Michel, C., Cosnard, G., & De Volder, A. G. (2000). Comparison of regional cerebral blood flow and glucose metabolism in the normal brain: effect of aging. *Journal of the neurological sciences, 181*(1-2), 19–28. doi: 10.1016/s0022-510x(00)00396-8

Betrus, C., & Kreipke, C. W. (2013). Historical perspectives in understanding traumatic brain

injury and in situating disruption in CBF in the pathotrajectory of head trauma. In *Cerebral Blood Flow, Metabolism, and Head Trauma* (pp. 1-27). Springer, New York, NY. doi: 10.1007/978-1-4614-4148-9_1

Blennow, K., Brody, D. L., Kochanek, P. M., Levin, H., McKee, A., Ribbers, G. M., ... & Zetterberg, H. (2016). Traumatic brain injuries. *Nature reviews Disease primers*, *2*(1), 1-19. doi: 10.1038/nrdp.2016.84

Botteri, M., Bandera, E., Minelli, C., & Latronico, N. (2008). Cerebral blood flow thresholds for cerebral ischemia in traumatic brain injury. A systematic review. *Critical care medicine, 36*(11), 3089–3092. doi: 10.1097/CCM.0b013e31818bd7df

Breiman, L. (2001). Random forests. *Machine learning, 45*(1), 5-32. doi: 10.1023/A:1010933404324

Bron, E. E., Steketee, R. M., Houston, G. C., Oliver, R. A., Achterberg, H. C., Loog, M., ... & Alzheimer's Disease Neuroimaging Initiative. (2014). Diagnostic classification of arterial spin labeling and structural MRI in presenile early stage dementia. *Human brain mapping, 35*(9), 4916-4931. doi: 10.1002/hbm.22522

Brown, G. G., Clark, C., & Liu, T. T. (2007). Measurement of cerebral perfusion with arterial spin labeling: Part 2. Applications. *Journal of the International Neuropsychological Society: JINS, 13*(3), 526–538. doi: 10.1017/S1355617707070634

Buxton, R. B., Frank, L. R., Wong, E. C., Siewert, B., Warach, S., & Edelman, R. R. (1998). A general kinetic model for quantitative perfusion imaging with arterial spin labeling. *Magnetic resonance in medicine, 40*(3), 383–396. doi: 10.1002/mrm.1910400308

Carroll, T. J., Horowitz, S., Shin, W., Mouannes, J., Sawlani, R., Ali, S., ... & Futterer, S. (2008). Quantification of cerebral perfusion using the “bookend technique”: an evaluation in CNS tumors. *Magnetic resonance imaging*, *26*(10), 1352-1359. doi: 10.1016/j.mri.2008.04.010

Choi, J. Y., Hart, T., Whyte, J., Rabinowitz, A. R., Oh, S. H., Lee, J., & Kim, J. J. (2019). Myelin water imaging of moderate to severe diffuse traumatic brain injury. *NeuroImage: Clinical, 22*, 101785. doi: 10.1016/j.nicl.2019.101785

Cole, J. H., Leech, R., Sharp, D. J., & Alzheimer's Disease Neuroimaging Initiative (2015). Prediction of brain age suggests accelerated atrophy after traumatic brain injury. *Annals of neurology, 77*(4), 571–581. doi: 10.1002/ana.24367

Collij, L. E., Heeman, F., Kuijer, J. P., Ossenkoppele, R., Benedictus, M. R., Möller, C., ... & Wink, A. M. (2016). Application of machine learning to arterial spin labeling in mild cognitive impairment and Alzheimer disease. *Radiology, 281*(3), 865-875. doi: 10.1148/radiol.2016152703

Cox, D. D., & Savoy, R. L. (2003). Functional magnetic resonance imaging (fMRI) "brain reading": detecting and classifying distributed patterns of fMRI activity in human visual cortex. *NeuroImage, 19*(2), 261–270. doi: 10.1016/s1053-8119(03)00049-1

Craddock, R. C., James, G. A., Holtzheimer III, P. E., Hu, X. P., & Mayberg, H. S. (2012). A whole brain fMRI atlas generated via spatially constrained spectral clustering. *Human brain mapping, 33*(8), 1914-1928. doi: 10.1002/hbm.21333

Davatzikos, C., Ruparel, K., Fan, Y., Shen, D. G., Acharyya, M., Loughead, J. W., ... & Langleben, D. D. (2005). Classifying spatial patterns of brain activity with machine learning methods: application to lie detection. *Neuroimage, 28*(3), 663-668. doi: 10.1016/j.neuroimage.2005.08.009

de Rooij, R., & Kuhl, E. (2018). Physical Biology of Axonal Damage. *Frontiers in cellular neuroscience, 12*, 144. doi: 10.3389/fncel.2018.00144

Detre, J. A., Leigh, J. S., Williams, D. S., & Koretsky, A. P. (1992). Perfusion imaging. *Magnetic resonance in medicine, 23*(1), 37–45. doi: 10.1002/mrm.1910230106

Detre, J. A., Wang, J., Wang, Z., & Rao, H. (2009). Arterial spin-labeled perfusion MRI in basic and clinical neuroscience. *Current opinion in neurology, 22*(4), 348–355. doi: 10.1097/WCO.0b013e32832d9505

Detre, J. A., Zhang, W., Roberts, D. A., Silva, A. C., Williams, D. S., Grandis, D. J., ... & Leigh, J. S. (1994). Tissue specific perfusion imaging using arterial spin labeling. *NMR in Biomedicine, 7*(1-2), 75-82. doi: 10.1002/nbm.1940070112

Dolui, S., Wang, Z., Shinohara, R.T., Wolk, D.A., & Detre, J.A., (2017). Structural Correlation-based Outlier Rejection (SCORE) algorithm for arterial spin labeling time series. *Journal of Magnetic Resonance Imaging, 45*(6), 1786–1797. doi: 10.1002/jmri.25436

Duong, T., Kim, D., Ugurbil, K., & Kim, S.G. (2001). Localized cerebral blood flow response at submillimeter columnar resolution. *Proceedings of the National Academy of Sciences, 98*, 10904-10909. doi: 10.1073/pnas.191101098

Eickhoff, S. B., Constable, R. T., & Yeo, B. T. (2018). Topographic organization of the cerebral cortex and brain cartography. *Neuroimage, 170*, 332-347. doi: 10.1016/j.neuroimage.2017.02.018

Foundas, A. L., Leonard, C. M., Gilmore, R. L., Fennell, E. B., & Heilman, K. M. (1996). Pars triangularis asymmetry and language dominance. *Proceedings of the National Academy of Sciences, 93*(2), 719–722. doi: 10.1073/pnas.93.2.719

Fridley, J., Robertson, C., & Gopinath, S. (2015). Quantitative lobar cerebral blood flow for outcome prediction after traumatic brain injury. *Journal of neurotrauma, 32*(2), 75–82. doi: 10.1089/neu.2014.3350

Friston, K. J., Josephs, O., Zarahn, E., Holmes, A. P., Rouquette, S., & Poline, J. (2000). To smooth or not to smooth? Bias and efficiency in fMRI time-series analysis. NeuroImage, 12(2), 196–208. https://doi.org/10.1006/nimg.2000.0609

Gaonkar, B., & Davatzikos, C., (2013). Analytic estimation of statistical significance maps for support vector machine based multi-variate image analysis and classification. *Neuroimage, 78*, 270–283. doi: 10.1016/j.neuroimage.2013.03.066

Golding, E. (2002). Sequelae following traumatic brain injury. The cerebrovascular perspective. *Brain research reviews, 38*(3), 377–388. doi: 10.1016/S0165-0173(02)00141-8

Golding, E. M., Robertson, C. S., & Bryan, R. M. (1999). The consequences of traumatic brain injury on cerebral blood flow and autoregulation: a review. *Clinical and experimental hypertension, 21*(4), 299-332. doi: 10.3109/10641969909068668

Golland, P., & Fischl, B. (2003). Permutation tests for classification: towards statistical significance in image-based studies. In *Biennial international conference on information processing in medical imaging* (pp. 330-341). Springer, Berlin, Heidelberg. doi: 10.1007/978-3-540-45087-0_28

Golland, P., Liang, F., Mukherjee, S., & Panchenko, D. (2005). Permutation tests for classification. In *International conference on computational learning theory* (pp. 501-515). Springer, Berlin, Heidelberg. doi: 10.1007/11503415_34

Graves, J. C., & Kreipke, C. W. (2015). Endothelin, Cerebral Blood Flow, and Traumatic Brain Injury. *Brain Neurotrauma: Molecular, Neuropsychological, and Rehabilitation Aspects.*

Grossman, E. J., Jensen, J. H., Babb, J. S., Chen, Q., Tabesh, A., Fieremans, E., ... & Grossman, R. I. (2013). Cognitive impairment in mild traumatic brain injury: a longitudinal diffusional kurtosis and perfusion imaging study. *American Journal of Neuroradiology, 34*(5), 951-957. doi: 10.3174/ajnr.A3358

Grover, V. P., Tognarelli, J. M., Crossey, M. M., Cox, I. J., Taylor-Robinson, S. D., & McPhail, M. J. (2015). Magnetic Resonance Imaging: Principles and Techniques: Lessons for Clinicians. *Journal of clinical and experimental hepatology, 5*(3), 246–255. doi:

10.1016/j.jceh.2015.08.001

Guyon, I., Weston, J., Barnhill, S., & Vapnik, V. (2002). Gene selection for cancer classification using support vector machines. *Machine learning, 46*(1), 389-422. doi: 10.1023/A:1012487302797

Harris, T. C., de Rooij, R., & Kuhl, E. (2019). The shrinking brain: cerebral atrophy following traumatic brain injury. *Annals of biomedical engineering, 47*(9), 1941–1959. doi: 10.1007/s10439-018-02148-2

Hawkins, D. M. (2004). The problem of overfitting. *Journal of chemical information and computer sciences, 44*(1), 1-12. doi: 10.1021/ci0342472

Henschel, L., Conjeti, S., Estrada, S., Diers, K., Fischl, B., & Reuter, M. (2020). Fastsurfer-a fast and accurate deep learning based neuroimaging pipeline. *NeuroImage, 219*, 117012. doi: 10.1016/j.neuroimage.2020.117012

Hlatky, R., Contant, C. F., Diaz-Marchan, P., Valadka, A. B., & Robertson, C. S. (2004). Significance of a reduced cerebral blood flow during the first 12 hours after traumatic brain injury. *Neurocritical care, 1*(1), 69-83. doi: 10.1385/NCC:1:1:69

Honda, M., Ichibayashi, R., Yokomuro, H., Yoshihara, K., Masuda, H., Haga, D., Seiki, Y., Kudoh, C., & Kishi, T. (2016). Early Cerebral Circulation Disturbance in Patients Suffering from Severe Traumatic Brain Injury (TBI): A Xenon CT and Perfusion CT Study. *Neurologia medico-chirurgica, 56*(8), 501–509. doi: 10.2176/nmc.oa.2015-0341

Houle, M. E., Kriegel, H. P., Kröger, P., Schubert, E., & Zimek, A. (2010, June). Can shared-neighbor distances defeat the curse of dimensionality?. In *International Conference on Scientific and Statistical Database Management* (pp. 482-500). Springer, Berlin, Heidelberg. doi: 10.1007/978-3-642-13818-8_34

Hsing, T., Attoor, S., & Dougherty, E. (2003). Relation between permutation-test p values and

classifier error estimates. *Machine Learning, 52*(1), 11-30. doi: 10.1023/A:1023985022691

Hua, X., Leow, A. D., Parikshak, N., Lee, S., Chiang, M. C., Toga, A. W., ... & Alzheimer's Disease Neuroimaging Initiative. (2008). Tensor-based morphometry as a neuroimaging biomarker for Alzheimer's disease: an MRI study of 676 AD, MCI, and normal subjects. *Neuroimage, 43*(3), 458-469. doi: 10.1016/j.neuroimage.2008.07.013

Hughes, G. (1968). On the mean accuracy of statistical pattern recognizers. *IEEE transactions on information theory, 14*(1), 55-63. doi: 10.1109/TIT.1968.1054102

Humphreys, I., Wood, R. L., Phillips, C. J., & Macey, S. (2013). The costs of traumatic brain injury: a literature review. *ClinicoEconomics and outcomes research: CEOR*, *5*, 281. doi: 10.2147/CEOR.S44625

Hyder, A. A., Wunderlich, C. A., Puvanachandra, P., Gururaj, G., & Kobusingye, O. C. (2007). The impact of traumatic brain injuries: a global perspective. *NeuroRehabilitation*, *22*(5), 341-353. doi: 10.3233/NRE-2007-22502

Joris, P. J., Mensink, R. P., Adam, T. C., & Liu, T. T. (2018). Cerebral blood flow measurements in adults: a review on the effects of dietary factors and exercise. *Nutrients, 10*(5), 530. doi: 10.3390/nu10050530

Kallakuri, S., Bandaru, S., Zakaria, N., Shen, Y., Kou, Z., Zhang, L., ... & Cavanaugh, J. M. (2015). Traumatic brain injury by a closed head injury device induces cerebral blood flow changes and microhemorrhages. *Journal of clinical imaging science, 5*. doi: 10.4103/2156-7514.166354

Kamps, F. S., Morris, E. J., & Dilks, D. D. (2019). A face is more than just the eyes, nose, and mouth: fMRI evidence that face-selective cortex represents external features. *Neuroimage, 184*, 90-100. doi: 10.1016/j.neuroimage.2018.09.027

Kemeny, S., Ye, F. Q., Birn, R., & Braun, A. R. (2005). Comparison of continuous overt speech

fMRI using BOLD and arterial spin labeling. *Human brain mapping, 24*(3), 173–183. doi: 10.1002/hbm.20078

Kenney, K., Amyot, F., Haber, M., Pronger, A., Bogoslovsky, T., Moore, C., & Diaz-Arrastia, R. (2016). Cerebral vascular injury in traumatic brain injury. *Experimental neurology, 275*, 353-366. doi: 10.1016/j.expneurol.2015.05.019

Kim, J., Whyte, J., Patel, S., Avants, B., Europa, E., Wang, J., Slattery, J., Gee, J. C., Coslett, H. B., & Detre, J. A. (2010). Resting cerebral blood flow alterations in chronic traumatic brain injury: an arterial spin labeling perfusion FMRI study. *Journal of neurotrauma, 27*(8), 1399–1411. doi: 10.1089/neu.2009.1215

Kim, J., Whyte, J., Patel, S., Europa, E., Slattery, J., Coslett, H. B., & Detre, J. A. (2012). A perfusion fMRI study of the neural correlates of sustained-attention and working-memory deficits in chronic traumatic brain injury. *Neurorehabilitation and Neural Repair, 26*(7), 870-880. doi: 10.1177/1545968311434553

Kim, J., Whyte, J., Wang, J., Rao, H., Tang, K. Z., & Detre, J. A. (2006). Continuous ASL perfusion fMRI investigation of higher cognition: quantification of tonic CBF changes during sustained attention and working memory tasks. *NeuroImage, 31*(1), 376–385. doi: 10.1016/j.neuroimage.2005.11.035

Kwong, K. K., Belliveau, J. W., Chesler, D. A., Goldberg, I. E., Weisskoff, R. M., Poncelet, B. P., Kennedy, D. N., Hoppel, B. E., Cohen, M. S., & Turner, R. (1992). Dynamic magnetic resonance imaging of human brain activity during primary sensory stimulation. *Proceedings of the National Academy of Sciences, 89(*12), 5675–5679. doi: 10.1073/pnas.89.12.5675

LaConte, S., Strother, S., Cherkassky, V., Anderson, J., & Hu, X. (2005). Support vector machines for temporal classification of block design fMRI data. *NeuroImage, 26*(2), 317–329. doi:

10.1016/j.neuroimage.2005.01.048

Langlois, J. A., Rutland-Brown, W., & Wald, M. M. (2006). The epidemiology and impact of traumatic brain injury: a brief overview. *The Journal of head trauma rehabilitation, 21(5)*, 375–378. doi: 10.1097/00001199-200609000-00001

Li, Y., Dolui, S., Xie, D., & Wang, Z. (2018). Priors-guided slice-wise adaptive outlier cleaning for arterial spin labeling perfusion MRI. *Journal of neuroscience methods, 307*, 248–253. doi: 10.1016/j.jneumeth.2018.06.007

Liu, T. T., & Brown, G. G. (2007). Measurement of cerebral perfusion with arterial spin labeling: Part 1. Methods. *Journal of the International Neuropsychological Society: JINS, 13*(3), 517–525. doi: 10.1017/S1355617707070646

Luh, W. M., Wong, E. C., Bandettini, P. A., Ward, B. D., & Hyde, J. S. (2000). Comparison of simultaneously measured perfusion and BOLD signal increases during brain activation with T1-based tissue identification. *Magnetic resonance in medicine, 44*(1), 137–143. doi: 10.1002/1522-2594(200007)44:1<137::AID-MRM20>3.0.CO;2-R

Maas, A. I., Menon, D. K., Adelson, P. D., Andelic, N., Bell, M. J., Belli, A., ... & Francony, G. (2017). Traumatic brain injury: integrated approaches to improve prevention, clinical care, and research. *The Lancet Neurology, 16*(12), 987-1048. doi: 10.1016/S1474-4422(17)30371-X

Mandeville, J. B., Marota, J. J., Ayata, C., Moskowitz, M. A., Weisskoff, R. M., & Rosen, B. R. (1999). MRI measurement of the temporal evolution of relative CMRO(2) during rat forepaw stimulation. *Magnetic resonance in medicine, 42*(5), 944–951. doi: 10.1002/(SICI)1522-2594(199911)42:5<944::AID-MRM15>3.0.CO;2-W

Marion, D. W., Darby, J., & Yonas, H. (1991). Acute regional cerebral blood flow changes caused by severe head injuries. *Journal of neurosurgery, 74*(3), 407–414. doi:

10.3171/jns.1991.74.3.0407

Martin, N. A., Patwardhan, R. V., Alexander, M. J., Africk, C. Z., Lee, J. H., Shalmon, E., ... & Becker, D. P. (1997). Characterization of cerebral hemodynamic phases following severe head trauma: hypoperfusion, hyperemia, and vasospasm. *Journal of neurosurgery, 87*(1), 9-19. doi: 10.3171/jns.1997.87.1.0009

McFadzean, R., Brosnahan, D., Hadley, D., & Mutlukan, E. (1994). Representation of the visual field in the occipital striate cortex. *British Journal of Ophthalmology, 78*(3), 185-190.

McKinney, W. (2010). Data structures for statistical computing in python. In *Proceedings of the 9th Python in Science Conference* (Vol. 445, pp. 51–56). doi: 10.25080/Majora-92bf1922-00a

Mechelli, A., Humphreys, G. W., Mayall, K., Olson, A., & Price, C. J. (2000). Differential effects of word length and visual contrast in the fusiform and lingual gyri during. *Proceedings of the Royal Society of London. Series B: Biological Sciences, 267*(1455), 1909-1913. doi: 10.1098/rspb.2000.1229

Meythaler, J. M., Peduzzi, J. D., Eleftheriou, E., & Novack, T. A. (2001). Current concepts: diffuse axonal injury-associated traumatic brain injury. *Archives of physical medicine and rehabilitation, 82*(10), 1461–1471. doi: 10.1053/apmr.2001.25137

Miller, K. L., Luh, W. M., Liu, T. T., Martinez, A., Obata, T., Wong, E. C., Frank, L. R., & Buxton, R. B. (2001). Nonlinear temporal dynamics of the cerebral blood flow response. *Human brain mapping, 13*(1), 1–12. doi: 10.1002/hbm.1020

Mitchell, T. M., Hutchinson, R., Niculescu, R. S., Pereira, F., Wang, X., Just, M., & Newman, S. (2004). Learning to decode cognitive states from brain images. *Machine learning, 57*(1), 145-175. doi: 10.1023/B:MACH.0000035475.85309.1b

Molinaro, A. M., Simon, R., & Pfeiffer, R. M. (2005). Prediction error estimation: a comparison of

resampling methods. *Bioinformatics, 21*(15), 3301-3307. doi: 10.1093/bioinformatics/bti499

Mourão-Miranda, J., Bokde, A. L., Born, C., Hampel, H., & Stetter, M. (2005). Classifying brain states and determining the discriminating activation patterns: Support Vector Machine on functional MRI data. *NeuroImage, 28*(4), 980–995. doi: 10.1016/j.neuroimage.2005.06.070

Noble, W. S. (2006). What is a support vector machine?. *Nature biotechnology, 24*(12), 1565-1567. doi: 10.1038/nbt1206-1565

Ogawa, S., Menon, R. S., Tank, D. W., Kim, S. G., Merkle, H., Ellermann, J. M., & Ugurbil, K. (1993). Functional brain mapping by blood oxygenation level-dependent contrast magnetic resonance imaging. A comparison of signal characteristics with a biophysical model. *Biophysical journal, 64*(3), 803–812. doi: 10.1016/S0006-3495(93)81441-3

Ojala, M., & Garriga, G. C. (2010). Permutation tests for studying classifier performance. *Journal of Machine Learning Research, 11*(6).

Olson, I. R., Plotzker, A., & Ezzyat, Y. (2007). The enigmatic temporal pole: a review of findings on social and emotional processing. *Brain, 130*(7), 1718-1731. doi: 10.1093/brain/awm052

Overton, D.J., Bhagwat, N., Viviano, J.D., Jacobs, G.R., & Voineskos, A.N. (2020). Identifying psychosis spectrum youth using support vector machines and cerebral blood perfusion as measured by arterial spin labeled fMRI. *NeuroImage: Clinical, 27*, 102304. doi: 10.1016/j.nicl.2020.102304

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., … (2011). Scikit-learn: Machine learning in Python. *the Journal of machine Learning research, 12*, 2825–2830.

Peng, S. P., Li, Y. N., Liu, J., Wang, Z. Y., Zhang, Z. S., Zhou, S. K., Tao, F. X., & Zhang, Z. X. (2016). Pulsed arterial spin labeling effectively and dynamically observes changes in

cerebral blood flow after mild traumatic brain injury. *Neural regeneration research, 11*(2), 257. doi: 10.4103/1673-5374.177733

Petcharunpaisan, S., Ramalho, J., & Castillo, M. (2010). Arterial spin labeling in neuroimaging. *World journal of radiology, 2*(10), 384–398. doi: 10.4329/wjr.v2.i10.384

Petersen, E. T., Zimine, I., Ho, Y. C., & Golay, X. (2006). Non-invasive measurement of perfusion: a critical review of arterial spin labelling techniques. *The British journal of radiology, 79*(944), 688–701. doi: 10.1259/bjr/67705974

Petrella, J. R., & Provenzale, J. M. (2000). MR perfusion imaging of the brain: techniques and applications. AJR. *American journal of roentgenology, 175*(1), 207–219. doi: 10.2214/ajr.175.1.1750207

Plis, S. M., Hjelm, D. R., Salakhutdinov, R., Allen, E. A., Bockholt, H. J., Long, J. D., ... & Calhoun, V. D. (2014). Deep learning for neuroimaging: a validation study. *Frontiers in neuroscience, 8*, 229. doi: 10.3389/fnins.2014.00229

Pollock, J. M., Tan, H., Kraft, R. A., Whitlow, C. T., Burdette, J. H., & Maldjian, J. A. (2009). Arterial spin-labeled MR perfusion imaging: clinical applications. *Magnetic resonance imaging clinics of North America, 17*(2), 315–338. doi: 10.1016/j.mric.2009.01.008

Ponsford. (2013). Factors contributing to outcome following traumatic brain injury. *NeuroRehabilitation, 32*(4). doi: 10.3233/NRE-130904

Powers, D. M. (2020). Evaluation: from precision, recall and F-measure to ROC, informedness, markedness and correlation. *arXiv preprint arXiv:2010.16061.*

Proisy, M., Mitra, S., Uria-Avellana, C., Sokolska, M., Robertson, N. J., Le Jeune, F., & Ferré, J. C. (2016). Brain perfusion imaging in neonates: an overview. American Journal of *Neuroradiology, 37*(10), 1766-1773. doi: 10.3174/ajnr.A4778

Rabinowitz, A. R., Hart, T., Whyte, J., & Kim, J. (2018). Neuropsychological Recovery

Trajectories in Moderate to Severe Traumatic Brain Injury: Influence of Patient Characteristics and Diffuse Axonal Injury. *Journal of the International Neuropsychological Society: JINS, 24*(3), 237. doi: 10.1017/S1355617717000996

Rakotomamonjy, A. (2004, August). Optimizing Area Under Roc Curve with SVMs. In *ROCAI* (pp. 71-80).

Rodríguez-Pérez, R., Vogt, M., & Bajorath, J. (2017). Support vector machine classification and regression prioritize different structural features for binary compound activity and potency value prediction. *ACS omega, 2*(10), 6371-6379. doi: 10.1021/acsomega.7b01079

Sandsmark, D. K., Bashir, A., Wellington, C. L., & Diaz-Arrastia, R. (2019). Cerebral Microvascular Injury: A Potentially Treatable Endophenotype of Traumatic Brain Injury-Induced Neurodegeneration. *Neuron, 103*(3), 367–379. doi: 10.1016/j.neuron.2019.06.002

Schölkopf, B., Smola, A. J., & Bach, F. (2001). Learning with kernels: support vector machines, regularization, optimization, and beyond. MIT press.

Scikit-learn developers. (2020). *Cross-validation: evaluating estimator performance*. scikit-learn 0.24.1. https://scikit-learn.org/stable/modules/cross_validation.html

Scikit-learn developers. (2020). *Importance of Feature Scaling*. scikit-learn 0.24.1. https://scikit-learn.org/stable/auto_examples/preprocessing/plot_scaling_importance.html

Scikit-learn developers. (2020). *Permutation feature importance*. scikit-learn 0.24.1. https://scikit-learn.org/stable/modules/permutation_importance.html

Shah, Y. S., Hernandez-Garcia, L., Jahanian, H., & Peltier, S. J. (2016). Support vector machine classification of arterial volume-weighted arterial spin tagging images. *Brain and behavior, 6*(12), e00549. doi: 10.1002/brb3.549

Smith, S. M., & Nichols, T. E. (2018). Statistical challenges in “big data” human neuroimaging. *Neuron, 97*(2), 263-268. doi: 10.1016/j.neuron.2017.12.018

Tan, H., Maldjian, J. A., Pollock, J. M., Burdette, J. H., Yang, L. Y., Deibler, A. R., & Kraft, R. A. (2009). A fast, effective filtering method for improving clinical pulsed arterial spin labeling MRI. *Journal of magnetic resonance imaging: JMRI, 29*(5), 1134–1139. doi: 10.1002/jmri.21721

Thangavel, R., Van Hoesen, G. W., & Zaheer, A. (2008). Posterior parahippocampal gyrus pathology in Alzheimer's disease. *Neuroscience, 154*(2), 667–676. doi: 10.1016/j.neuroscience.2008.03.077

Thomas, A. W., Heekeren, H. R., Müller, K. R., & Samek, W. (2019). Analyzing neuroimaging data through recurrent deep learning models. *Frontiers in neuroscience, 13*, 1321. doi: 10.3389/fnins.2019.01321

Tjandra, T., Brooks, J. C., Figueiredo, P., Wise, R., Matthews, P. M., & Tracey, I. (2005). Quantitative assessment of the reproducibility of functional activation measured with BOLD and MR perfusion imaging: implications for clinical trial design. *NeuroImage, 27*(2), 393–401. doi: 10.1016/j.neuroimage.2005.04.021

Torre, J. C. de la, Olmo, A. del, & Valles, S. (2019). Can mild cognitive impairment be stabilized by showering brain mitochondria with laser photons?. *Neuropharmacology*. doi: 10.1016/j.neuropharm.2019.107841

Vabalas, A., Gowen, E., Poliakoff, E., & Casson, A. J. (2019). Machine learning algorithm validation with a limited sample size. *PloS one, 14*(11), e0224365. doi: 10.1371/journal.pone.0224365

Wang, J., Aguirre, G. K., Kimberg, D. Y., Roc, A. C., Li, L., & Detre, J. A. (2003). Arterial spin labeling perfusion fMRI with very low task frequency. *Magnetic resonance in medicine, 49*(5), 796–802. doi: 10.1002/mrm.10437

Wang, J., Li, L., Roc, A. C., Alsop, D. C., Tang, K., Butler, N. S., ... & Detre, J. A. (2004).

Reduced susceptibility effects in perfusion fMRI with single-shot spin-echo EPI acquisitions at 1.5 Tesla. *Magnetic resonance imaging, 22*(1), 1-7. doi: 10.1016/S0730-725X(03)00210-8

Wang, M. L., & Li, W. B. (2016). Cognitive impairment after traumatic brain injury: The role of MRI and possible pathological basis. *Journal of the neurological sciences, 370*, 244–250. doi: 10.1016/j.jns.2016.09.049

Wang, X., Hutchinson, R., & Mitchell, T. M. (2003). Training fMRI classifiers to detect cognitive states across multiple human subjects. *NIPS03, 16*.

Wang, Z., Aguirre, G. K., Rao, H., Wang, J., Fernández-Seara, M. A., Childress, A. R., & Detre, J. A. (2008). Empirical optimization of ASL data analysis using an ASL data processing toolbox: ASLtbx. *Magnetic resonance imaging, 26*(2), 261–269. doi: 10.1016/j.mri.2007.07.003

Wang, Z., Childress, A. R., & Detre, J. A. (2006, August). Boost up the detection sensitivity of ASL perfusion fMRI through support vector machine. In *2006 International Conference of the IEEE Engineering in Medicine and Biology Society* (pp. 1006-1009). IEEE. doi: 10.1109/IEMBS.2006.260382

Wang, Z., Das, S. R., Xie, S. X., Arnold, S. E., Detre, J. A., Wolk, D. A., & Alzheimer's Disease Neuroimaging Initiative (2013). Arterial spin labeled MRI in prodromal Alzheimer's disease: A multi-site study. *NeuroImage. Clinical, 2,* 630–636. doi: 10.1016/j.nicl.2013.04.014

Ware, J. B., Dolui, S., Duda, J., Gaggi, N., Choi, R., Detre, J., ... & Kim, J. J. (2020). Relationship of Cerebral Blood Flow to Cognitive Function and Recovery in Early Chronic Traumatic Brain Injury. *Journal of neurotrauma, 37*(20), 2180-2187. doi: 10.1089/neu.2020.7031

Ware, J. B., Hart, T., Whyte, J., Rabinowitz, A., Detre, J. A., & Kim, J. (2017). Inter-subject

variability of axonal injury in diffuse traumatic brain injury. J*ournal of neurotrauma, 34*(14), 2243-2253. doi: 10.1089/neu.2016.4817

Williams, D. S., Detre, J. A., Leigh, J. S., & Koretsky, A. P. (1992). Magnetic resonance imaging of perfusion using spin inversion of arterial water. *Proceedings of the National Academy of Sciences, 89*(1), 212–216. doi: 10.1073/pnas.89.1.212

Wolf, R. L., & Detre, J. A. (2007). Clinical neuroimaging using arterial spin-labeled perfusion magnetic resonance imaging. *Neurotherapeutics, 4*(3), 346-359. doi: 10.1016/j.nurt.2007.04.005

Xekardaki, A., Rodriguez, C., Montandon, M. L., Toma, S., Tombeur, E., Herrmann, F. R., ... & Haller, S. (2015). Arterial spin labeling may contribute to the prediction of cognitive deterioration in healthy elderly individuals. *Radiology, 274*(2), 490-499. doi: 10.1148/radiol.14140680

Xie, D., Li, Y., Yang, H., Bai, L., Wang, T., Zhou, F., ... & Wang, Z. (2020). Denoising arterial spin labeling perfusion MRI with deep machine learning. *Magnetic resonance imaging, 68*, 95-105. doi: 10.1016/j.mri.2020.01.005

Zarahn, E., Aguirre, G. K., & D'Esposito, M. (1997). Empirical analyses of BOLD fMRI statistics. *Neuroimage, 5*(3), 179-197. doi: 10.1006/nimg.1997.0263