Overview of Cerebrospinal Fluid, Glymphatic System, and Diagnosis of Multiple Sclerosis Through Machine Learning Extrapolation

Ethan Ngo

Abstract:

The glymphatic system is a newly discovered process in the subarachnoid space that regulates metabolic clearance in the brain. It will revolutionize the scientific understanding of sleep, cognitive function, and CNS disorders. In this article, the evolutionary origins and roles of cerebrospinal fluid will be analyzed to learn more about this clearance system. Moreover, this article will review and discuss the entirety of the glymphatic system, specifically the processes that occur in the perivascular space. Certain glial cells and proteins will be highlighted, such as astrocytes and AQP4, respectively, to further explain fluid transport and entry into the lymphatic network. The purpose of sleep and the accumulation of harmful CSF by-products will be closely scrutinized in relation to glymphatic impairment. CNS disorders have recently been associated with these neurobiological problems. Current research provides the use of CSF biomarker information to diagnose and identify neurodegenerative diseases. The new rise in deep learning, artificial neural networks, has innovated the possibilities of machine learning. Artificial intelligence algorithms, such as Random Forest, will be utilized to extrapolate data on prevalent neurobiological factors of multiple sclerosis (MS) patients. The program will also demonstrate machine learning capabilities by predicting MS diagnosis in both males and females. The new age of neurotechnology will feature more tangible and limitless roles in biomarker identification of CNS disorders. The aim of this article is to explore the glymphatic system and cerebrospinal fluid with future technological innovations.

Introduction

In the central nervous system (CNS), the cerebrospinal fluid (CSF) regulates and circulates the subarachnoid space, located between the arachnoid and pia mater. Composed primarily of blood plasma, it travels with the interstitial fluid (ISF) through hydrostatic and osmotic exchange. CSF contains blood-derived proteins, albumin, and immunoglobulins for metabolic recovery and protection of the brain. It is produced in various channels of the ventricular system and then released into the meninges; the layer of dura, pia, and arachnoid mater that encloses the brain (Di Terlizzi & Platt, 2006). From this space, CSF provides nutrients for neurons and is heavily involved in metabolic disposal through the glymphatic system (Benveniste et al., 2017). As CSF travels throughout the brain, it accumulates harmful by-products that can lead to neurodegenerative diseases. The newly discovered CNS ‘waste clearance’ system, the glymphatic system, has been recently identified in several disorders, one of which is Multiple Sclerosis (MS). Abnormalities and impairment in glymphatic function have led to MS demyelination and acute symptoms in progressive stages of clinical disability. Defects in the glymphatic system mediate underlying pathological mechanisms in multiple sclerosis and other neurodegenerative ailments (Carotenuto et al., 2022). The modern innovation of machine learning in healthcare and data analysis of CSF biomarkers can aid with the diagnosis of CNS disorders and help procure solutions.

Evolutionary Origins of CSF in Ancestral Lineage

CSF descends from a deuterostome lineage of the ancestral organisms Abulacracia and Chordata. The original function of CSF was to contribute to a balanced chemical environment for nonsynaptic signal communication. However, as evolution occurred, dendritic processes were formed in CSF brain ventricles and led to the formation of terminals (stereocilia). These passageways allowed for CSF-contacting neurons to enter synaptic zones in the brain. Development of CSF influx further continued after the massive clade formed between the deuterostome ancestors (Bueno & Garcia-Fernàndez, 2016).

The Abulacracia family descended into invertebrates. The derived species have an absence of CNS and possess minimal cognitive capability. Meanwhile, the Chordata organism evolved into Urochadata, Cephalachorta, and Vertebrata—the Vertebrata species, the last of the evolutionary line, includes fish, amphibians, reptiles, birds, and mammals. In other words, any organism with a backbone, the Vertebrata species, contains CSF to help protect its brain and spinal cord.

The Vertebrata species has neuronal communication and provides nutrients to cells through CSF barrier transfer systems. The neuronal expansion led to the distribution of CSF functions and the creation of emergent properties. Due to the development of neuron complexity, CNS neurons do not directly detect environmental conditions; however, they synapse with other neurons in sensory organs. Furthermore, the evolution of CSF led to the formation of layers of neurons, further advancing the neocortex.

Ventricular System: Synthesis and Circulation of CSF

The ventricular system, located throughout the brain parenchyma, is a series of connected cavities that partakes in the synthesis and excretion of the cerebrospinal fluid to the subarachnoid space and the spinal cord (Shenoy & Lui, 2023). The subarachnoid space lies between two layers of the meninges: the pia mater and the arachnoid mater. CSF is produced by a densely folded vascularized structure known as the choroid plexus. The choroid plexus is a group of epithelial cells that lies on the basolateral membrane, allowing it to connect to the surrounding vessel walls. It is facilitated through ion transport channels, which utilize an osmotic gradient to provide water and nutrients from fenestrated blood capillaries. At the apical membrane, motile cilia help move CSF through the ventricular system (Javed et al., 2023).

Once CSF is produced, it travels from the lateral ventricle to the third ventricle by the interventricular foramen. The fluid then enters the cerebral aqueduct and flows into the fourth ventricle. Each ventricle contains a choroid plexus, allowing for a convective flow and pressurized pathway. At this point, CSF can exit from four passageways: right aperture, median aperture, left aperture, and central canal brain stem. Finally, CSF circulates the subarachnoid space, providing nutrients and protection for the brain (Khasawneh et al., 2018).

Function and Purpose of CSF

There are three primary purposes of CSF:

- **Protection of the brain and spinal cord**

During head and neck injuries, CSF acts as a shock absorption and insulation between the brain and skull; thus, preventing intracranial pressure (Telano & Baker, 2023). This is to prevent damage to the functional tissue and maintain cognitive ability.

- **Nutrients for brain parenchyma**

CSF provides nutrients to neuron cells and glial cells. These nutrients include glucose, protein, and minerals. Moreover, CSF helps provide an ionically balanced environment in order to promote normal neuron-synaptic exchange.

- **Metabolic Waste Removal**

CSF utilizes the glymphatic system in order to irrigate the unhealthy by-products in CSF (Veening & Barendregt, 2010). Sinuses are another route for CSF to enter the lymphatic system and be refreshed.

The Glymphatic System

- **Overview**

The glymphatic system is the brain's metabolic waste disposal. It utilizes the convective flow of CSF and interstitial fluid (ISF), also known as the glymphatic influx, in order to clear waste products and reduce the amount of toxic by-products. The glymphatic system utilizes capillaries, astrocytes, and sinuses to irrigate CSF (Benveniste et al., 2017). If not regulated correctly, by-products can be accumulated within the CSF, creating fatal repercussions. Malfunctions in the glymphatic system can mediate the progression of several neurodegenerative diseases.

- **CSF Process in Glymphatic System (Perivascular Space)**

At the cortical surface of the brain, CSF travels along the pial artery. The fluid is facilitated by smooth muscle cells and an influx gradient. The pial artery then penetrates the brain parenchyma, allowing for the creation of the perivascular space, otherwise known as the Virchow-Robin space (Figure 1). In this periarterial region, the arterial pulsation and the convective influx help drive CSF down the basal lamina, a sheet of extracellular matrix (Jessen et al., 2015).

The basal lamina continues to enter the parenchyma, eventually diminishing into a capillary region. At the base of the basal lamina, a tight junction binds the vascular endfeet of

astrocytes (Figure 1). Astrocytes are glial cells in the CNS that promote metabolic and neuroprotective functions.

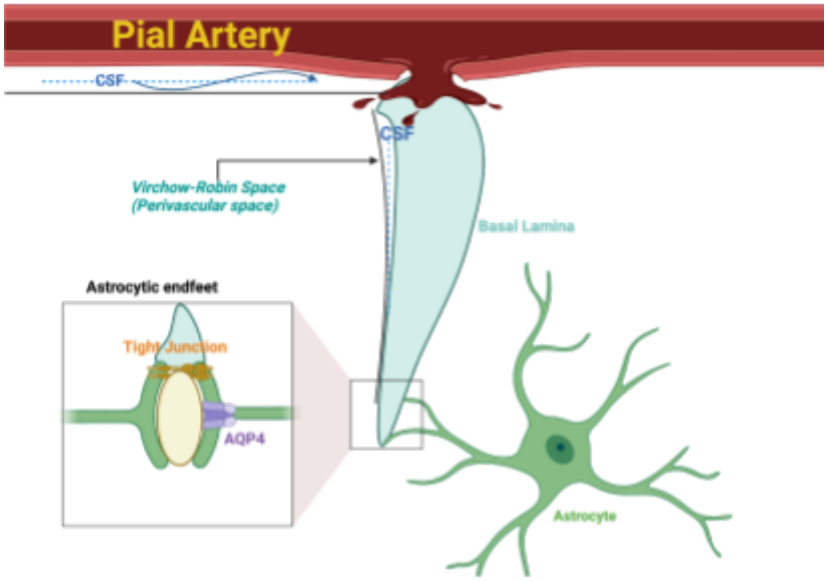

Figure 1: Glymphatic Process (Perivascular Space)

In the astrocytic endfeet, a protein channel, aquaporin-4, facilitates the interstitial (ISF) and CSF exchange (Mestre et al., 2020). AQP4 then excretes the CSF-ISF fluid into the neuronal tissue, generating a glymphatic influx. The overall pressure gradient of CSF generates a convective flow of ISF into the perivenous space, a fluid-filled structure traveling along veins in the perivascular space (Figure 2). In this region, interstitial fluid is collected and enters the cervical lymph system, cleansing off harmful by-products (Jessen et al., 2015).

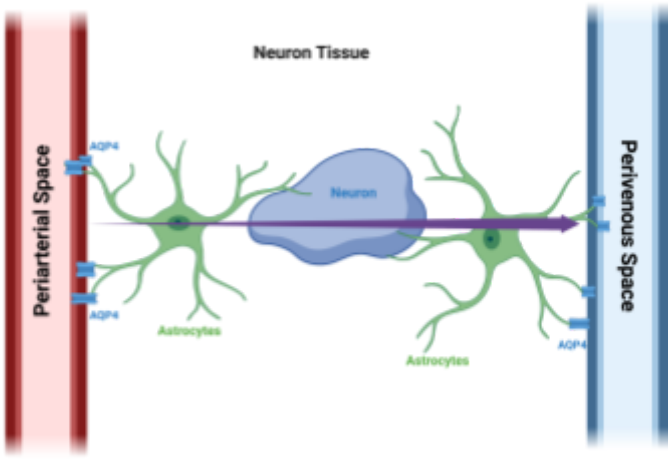

Figure 2: Astrocytic Polarization in Periarterial to Perivenous Space

Overall, the glymphatic system is a neurovascular process involving constant CSF-ISF exchange. Glymphatic activity is heavily reliant on astrocytes and AQP4. These glial bodies are the main facilitators of glymphatic influx and convective fluid transportation.

❖ Role AQP4

Aquaporin-4 channels are proteins that constantly exchange interstitial and cerebrospinal fluid. Expressed by glial cells, AQP4 plays an essential role in draining CSF-ISF fluid into perivascular pathways (Gao et al., 2023). These protein channels produce a polarity and flow into the neuron tissue, exchanging CSF across the subarachnoid space (Jessen et al., 2015).

AQP4 often decreases in functionality as an individual ages. For instance, AQP4 utilizes osmotic pressure to produce a current. However, these aquaporin channels degrade over time. The polarization and concentration decrease significantly, leading to a dispersal of influx gradients. Once this occurs, CSF-ISF exchange becomes sporadic, and the glymphatic system becomes impaired (Jessen et al., 2015). The disruption of CSF-ISF fluid transportation from the periarterial to the perivenous space greatly increases B-amyloid protein and other harmful products (Iliff & Simon, 2019). Without clear irrigation from the lymphatic system, neurological diseases like Alzheimer's tend to rise more frequently.

❖ Sleep

The glymphatic system is optimal during sleep. The system constantly detoxifies the brain yet undergoes certain cycles of efficient production. In the daytime, it remains generally inactive. However, during certain stages of sleep, like REM, chemicals, and regulations regarding homeostasis help drive CSF irrigation throughout the brain.

The glymphatic system is also regulated through a chief modulator, norepinephrine. During the daytime, this chemical is constantly released into the bloodstream to activate cognitive function, attention, and arousal (Hussain et al., 2023). Thus, during certain stages of sleep, the central acuity of norepinephrine dramatically decreases. The arousal of the individual minimizes and, therefore, expands the extracellular and interstitial space. Once this occurs, interstitial fluid and CSF influx increase due to less resistance of the surrounding walls (Reddy & Van Der Werf, 2020). Greater volume of ISF and CSF leads to more metabolic waste disposal, therefore contributing to the rise in glymphatic function during sleep.

Moreover, the glymphatic system functions in tandem with respiratory signals and intracranial oscillations. A lower frequency of events in the respiration system causes stable blood pressure levels, coinciding with an established heart rate. The most efficient stage, the greatest amyloid-beta removal, of the glymphatic system occurs at N3, in which arterial pulsations are minimal, and the drive for CSF maintains relatively high quantities (Reddy & Van Der Werf, 2020). At this stage, copious amounts of amyloid beta are removed and detoxified. This process is extremely delicate. Even minimal changes to the arterial pulsations decrease glymphatic influx and lessen toxin disposal.

❖ **Sinuses: Lymphatic meningeal vessels**

Moreover, the glymphatic system has an alternative route for CSF irrigation. In the meningeal layers, there are several air-filled cavities that are known as sinuses. The most commonly used are the superior sagittal sinus and the transverse sagittal sinus. These sinuses contain arachnoid granulations, which sustain pathways into the lymph meningeal vessels of the lymphatic systems (Letchuman & Donohoe, 2023). The CSF is then rinsed throughout the brain and exits the sinuses, where it is replenished and cleansed.

Impairment to Glymphatic System (CSF Biomarkers)

As CSF travels throughout the ventricular system, it collects unwanted by-products that lead to severe damage to neuron cells. If not treated carefully, these products can cause serious neurodegenerative, thus leading to impairments in mental acuteness and motion capabilities (Carotenuto et al., 2022).

In order to diagnose neurodegeneration, scientists generally use CSF biomarkers to determine if the disorder is present in the patients (Mattsson-Carlgren et al., 2022). There are three main CSF biomarkers:

I. B-amyloid
II. Tau Protein
III. $A\beta_{42}$

Common Technologies Used in CSF Biomarker Tracing

I. Mass Spectrometry

Innovative technology that is used to determine mass. Commonly used in proteomics in the identification of proteins and classification of chemical structures. In order for mass spectrometry to function, the source or protein must be ionized, containing charged particles. With an analyzer, the technology can split the molecules from the mass. This allows the detection of ionized atoms (Parker et al., 2010) in the protein. The device proceeds to amplify the result and retrieve quantifiable signals from the excavated charged particle.

Moreover, mass spectrometry utilizes two different performance outputs: unbiased vs. targeted. Target mass spectrometry is more high-throughput and sensitive. It coincides in tandem with parallel reaction monitoring (Bharucha et al., 2019). On the contrary, an unbiased approach is the proteomic study of novel biomarkers and large-scale protein identification. The holistic review provides distinction and expression divergence, leading to the categorization of study cohorts.

II. ELISA

Enzyme-Linked Immunosorbent Assay, ELISA, is an analytics biochemical assay that utilizes antibodies and ligand signals to detect the immunological response of a particular antigen. ELISA binds the antibody toward the foreign object and monitors the change in gene expression. In general, technology employs a major role in the diagnosis of immunological-compromised diseases.

For instance, ELISA can be used in the diagnosis of Parkinson's disease. Tyrosine hydroxylase, an oxygenase, is heavily associated with the pathogenesis of various disorders in the brain. With the targeted approach, the assay can facilitate TH expression with neuromodulatory agents, antibodies, to develop a measurable analysis of developmental changes (Fauss et al., 2013). Other diseases that ELISA can detect include, but are not limited to, HIV and neurocysticercosis (Rosas et al., 1986).

III. Two-photon microscopy

Two-photon microscopy is an imaging technique that utilizes fluorescent lighting to trace chemical biomarkers. First, tracers are injected into the *vivo* sections of the brain parenchyma, thus creating a depictive graph. An invasive procedure is required in order for the two-photon microscopy to function. This includes the opening of a cranial window and immense skinning of the skull. However, the amount of exposure and field of view is extremely narrow. Once the confocal microscopy is placed in the aperture, two photons are fired into the targeted area to visualize the fluorescent tracers on biomarkers.

CSF Neurodegenerative Diseases

CNS neurodegenerative diseases are heavily influenced by impairments of CSF flow and the glymphatic system (Simon & Iliff, 2016). The five most common disorders caused by CSF-related factors include:

- Alzheimer's
- **Multiple Sclerosis (MS)**
- Hydrocephalus
- Parkinson Disease
- Amyotrophic Lateral Sclerosis (ALS)

Project: Machine Learning Extrapolation of Biomarkers in Multiple Sclerosis

Introduction of Machine Learning

Machine learning is the use and development of computer systems. New technology has allowed for machine learning to draw and analyze data, allowing for inferences and identification of trends. Essentially, machine learning adapts by continuously learning test sets, eliminating the need for explicit instructions (Baştanlar & Özuysal, 2014).

Role of Machine Learning and Artificial Intelligence

The development of deep learning, DL, has expanded the medical limitations that once bounded artificial intelligence. DL has opened a new gate of complex technologies and innovation of practical uses. Before, algorithms possessed only three to five layers of neural networks. With this new type of machine learning, AI has expanded to ten layers, implicating millions more artificial neurons and capacities (Bohr & Memarzadeh, 2020). With this rise, machine learning became impossible to deny in applications of healthcare.

AI has demonstrated value in areas of administrative offices, robotic surgeries, and clinical image analysis. Precision medicine, such as omics-based tests, has used machine learning programs to generate correlations from population pools and predict treatments. The process of drug discovery and molecular target identification, like SMILES (simplified molecular input line entry system), has been simplified (Bohr & Memarzadeh, 2020). At Johns Hopkins University, smart tissue autonomous robots, otherwise known as STAR, have outperformed surgical doctors in procedures like bowel anastomosis (Graham, 2022).

Moreover, the new world of deep learning and medical scanning recognition augments a new reality of accurate data collection. Influenced by substantial DL techniques, convolutional neural networks (CNNs) have promised a new interpretation of radiological images in the human visual cortex. In general, analysis through AI can implicate a flexible structure for healthcare systems.

Introduction to Multiple Sclerosis

Multiple sclerosis, otherwise known as MS, is an autoimmune inflammatory disease that attacks neuronal tissue and leads to cognitive dysfunction. The immune system targets the myelin sheath on the axon of the neuron. Neuronal communication becomes impaired and causes a cascade through the brain stem (Barkhane et al., 2022).

The causes of MS are a combination of genetic inheritance and environmental factors. In recent studies, individuals who have relatives with MS have susceptibility to the disorder (NINDS, 2023). Females have a higher chance of MS, 3:1 ratio, than men due to their abrupt hormonal imbalances during puberty, pregnancy, puerperium, and menopause. However, testosterone in men leads to more serious cases of this condition (Ysrraelit & Correale, 2019).

Additionally, MS tends to target people with a lack of exposure to sunlight, signifying a deficiency in vitamin D (NINDS, 2023).

MS can be caused by the misidentification of “foreign” objects in the body. If the blood-brain barrier is damaged, white-blood cells will flood into the brain stem. Thus, the immune system will misinterpret myelin sheaths as “foreign” and utilize T-cells to break down these structures (Miron, 2019).

There are four main types of MS:

I. Relapse-remitting MS
    A. Periods of recovery and activity of symptoms
    B. The duration of the attack varies
    C. The most common form of the disease
II. Secondary-progressive MS
    A. Relapse of MS attacks
    B. Gradual degradation of symptoms
III. Primary-progressive MS
    A. Progressive symptoms
    B. No form of relapse
IV. Progressive-relapsing MS
    A. Acute relapses
    B. Progressive symptoms
    C. The rarest form of the disease

Extensive Biological Oversive of MS

Oligodendrocytes are organisms that synthesize myelin and protect the axonal pathway of a neuron. In MS, rogue T-cells, a form of lymphocytes, first bind onto the myelin sheath. This binding causes the release of cytokines or small signaling proteins, including IL-1, IL-6, TNF-$\propto$, INF-γ. These cytokines facilitate the dilution of the blood vessels (Figure 3).

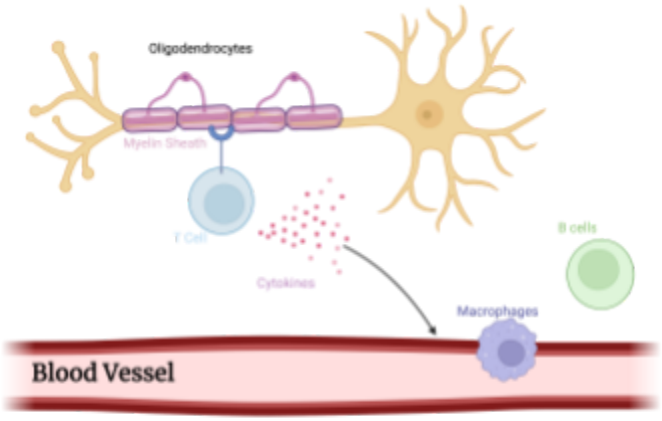

Figure 3: Stage I of Myelin Deterioration in MS

Once the permeable barrier between the blood vessel and the brain dilutes, an inflammatory response will occur. This will allow for macrophages and B-cells to enter the brain. Macrophages are digestive proteins that will engulf microorganisms to clean out dead cells and toxic products. B-cells are lymphocytes that help facilitate this action by producing antibodies. These antibodies bind to the microorganism and neutralize it. Macrophages will use these antibodies as a signal and detect the defective cell. In MS, the cell is not “defective” but rather caused by mis-stimulation of rogue T cells.

Macrophages will then engulf the oligodendrocytes and lead to demyelination. The sheath will then decay and lead to an exposed axon, therefore causing interference in neuron signals. The disintegrated sheaths will form plagues, also known as sclera (Figure 4). Thus, cognitive function and motile ability become impaired through the progressive decay of signal transduction of neurons.

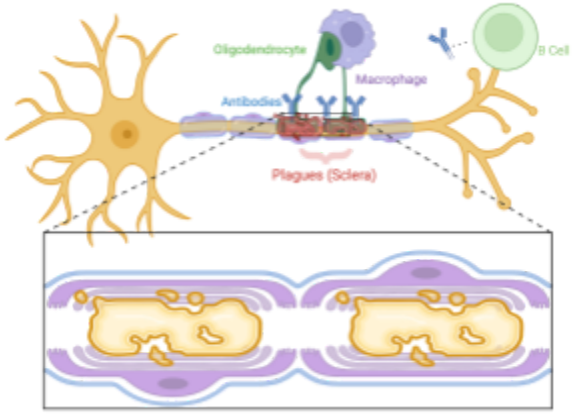

Figure 4: Stage II of Myelin Deterioration in MS

Oligoclonal Bands

Oligoclonal bands (OGBs) are immunoglobulin proteins that surface in the CSF due to an inflammatory response. These bands are key detectors in CNS diseases. In MS, when the T-cells begin to attach to the myelin sheath, they send cytokines, or inflammatory markers, that help B-cells pass through the blood-brain membrane. Oligoclonal bands are then generated and sent to the CNS (Neilson et al., 2022).

However, oligoclonal bands are present in different neurological diseases as well. OGBs are present in multiple sclerosis, herpes simplex encephalitis, and subacute sclerosing panencephalitis. In some exceptions, there are positive bands in Alzheimer's, amyotrophic lateral sclerosis, and polyneuropathy (Chu et al., 1983). Essentially, OGBs help in the diagnosis of a CNS immunologically compromised disease but do not indicate the specific condition.

Introduction to Dataset and Machine Learning Proposal

In a cohort case study conducted on Mexican mestizo patients, scientists analyzed the presence of CSF oligoclonal bands, as well as auditory, sensory, and motor reactions to various simulations. These indicators were collected in a shotgun assay, providing an accurate diagnosis of CNS disorder: multiple sclerosis. Two test groups were held to record data on patients' backgrounds and physical exams (Benjamin Pineda, 2023).

The first group, CDMS, were clinically diagnosed patients of MS and documented their score in the Expanded Disability Status Score (EDSS), a quantifiable scale on monitoring the disabling effects of MS in the patient. The second group, non-CDMS, were non-diagnosed patients of MS. Their symptoms were recorded; however, no values/documentation of their EDSS score were available.

In the project, machine learning will identify and diagnose non-CDMS patients. By using artificial intelligence, a trainable and testable algorithm will be able to read the code and provide an accurate prognosis. Furthermore, the AI will be able to identify neurobiological factors and their importance to the diagnosis of MS in both female and male patients. The general consensus involves machine learning, its application and purpose in healthcare, and the demonstration of the new innovative age of biological artificial intelligence.

Data Pre-Processing

To first analyze the dataset, several packages and Python libraries were imported. For example, numpy and pandas were used to store numerical information and allow modification of the data to fit in an array. The public dataset collected by the cohort study was uploaded and stored in a pandas data frame. This is a standard process in ensuring advanced mathematical computations in a relational way. Moreover, seaborn and matplotlib.pyplot were used to display the data visually throughout the code. Matplot was primarily utilized in illustrating the variances of MS relative importance, which will later be explained.

Before the program was able to run properly, several data pre-processing steps were required. Several columns were dropped: initial EDSS, age, breastfeeding, and schooling. These columns were used for background information on the patient and had no relevant importance in the diagnosis of MS. To predict non-CDMS patients accurately, the program had to solely focus on neurobiological components.

Due to multiple sclerosis's changes in behavior in different sexes, the dataset was copied and segmented into two parts. This was fundamental; allowed for the biological relative importance to be scaled differently in males and other factors to be highlighted in females. Moreover, by splitting the code, the predicted EDSS can be compared more precisely since males have a tendency to have a higher score on the disability scale.

Relative importance: Interpretation

Using scikit-learn, a machine learning tool for predictive data analysis, the female data set was split into a train and test set. The StandardScalar() method was then imported and implemented into the code to restructure the dataset's features. The deep learning algorithm, Random Forest, created a decision tree based on the test and train set. This allowed it to determine trends in certain factors that have the most effect on the diagnosis of MS.

- **Females**

Interestingly, the machine learning algorithm revealed that the most important factor in identifying MS for females was the periventricular MRI (Figure 5). This MRI locates periventricular white matter, a network of nerve fibers, in the frontal, parietal, temporal, or occipital part of the brain.

Studies demonstrate that the periventricular white matter in the cerebellum controls sensory information. MS demyelination in the cerebellum can cause lesions, abnormal or damaged areas, in the spinal cord, emphasizing destabilizing motor functions in the brainstem. Therefore, women have common MS symptoms of dizziness and loss of balance (Filippi et al., 2019). Periventricular white matter lesions can also lead to higher chances of neurodegenerative diseases like dementia and Alzheimer's. Most patients with MS also experience pseudobulbar palsy, or trouble in upper motor movement (Casini, 2013).

The initial symptoms have less of an effect on MS diagnosis compared to males. This is because women generally have a relapse-remitting MS, a mild form of the disorder.

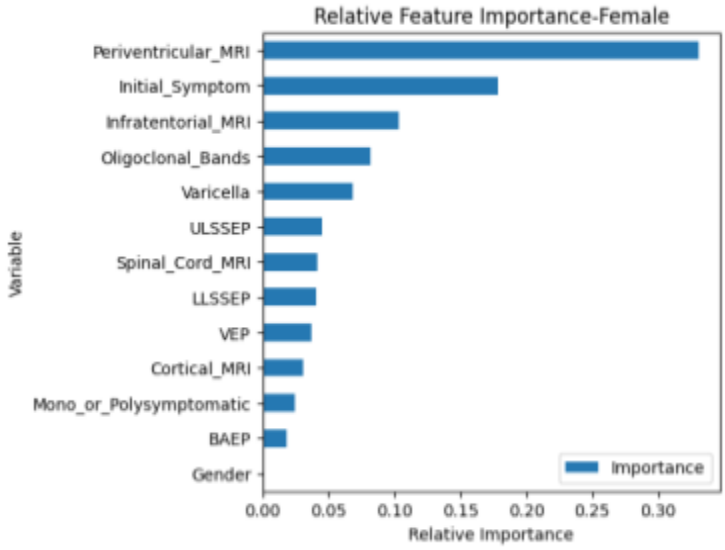

Figure 5: Relative Feature Importance in Females

❖ Males

The same process of splitting the dataset and training it through Random Forest was utilized for the male test case. In comparison to the females, initial symptoms and the infratentorial MRI had more significance in the diagnosis of MS (Figure 6).

The infratentorial white matter located in the cerebellum facilitates muscle control, movement, tremor, memory, and balance. Located in the brain stem, the infratentorial white matter affects language processing and motor instability. As lesions in the spinal cord arise, individuals experience a progressive disability. This could be a possible justification for men having greater neurodegeneration and primary-progressive MS (Dekker et al., 2020).

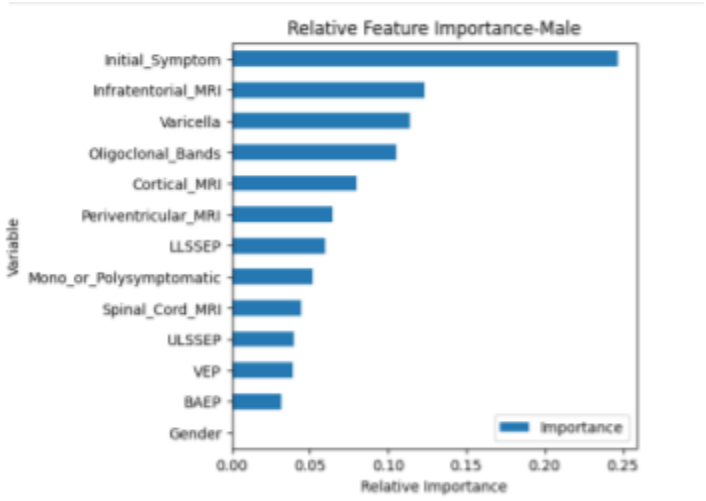

Figure 6: Relative Feature Importance in Males

Although men have a lower rate of acquiring multiple sclerosis in the first place, the symptoms that they exhibit are much more acute than women. This correlates to machine learning's prediction of initial symptoms playing a vital role in a patient's score on the EDSS scale.

Clearly, machine learning is efficient in identifying the biological importance of CNS disorders. The artificial intelligence was able to track trends and use a training set to calculate varying MS indicators in different sexes. It was able to reveal key qualities, such as the periventricular and intraventricular white matter, that were not evidently noticeable.

Machine Learning Discoveries

As showcased in the program, machine learning abstraction can lead to multiple discoveries. For example, many studies indicated that the largest MS indicator was oligoclonal bands. Although this biological feature ranked high in both females and males, it was not the primary biomarker. Despite CSF being hypothesized as the major contributor to MS, the data implies white matter lesions play a much more significant role.

Moreover, artificial intelligence detected another component of multiple sclerosis: varicella. Varicella, also commonly known as chickenpox, has a "causal effect" on MS. Current research is being conducted to elucidate this unknown connection in order to prevent MS progression (Zhu et al., 2023).

Diagnosis of Machine Learning in Patients

Artificial intelligence can also predict the prognosis of a patient's condition if given a trained dataset. Random forest's built-in method, forest.predict(), used the test set as a parameter to predict MS in both CDMS and non-CDMS patients. The algorithm had an 85% accuracy in

correctly identifying multiple sclerosis in non-CDMS individuals (Figure 7). With larger datasets and more advanced deep learning algorithms, predictions can become more precise. The results are displayed down below:

	precision	recall	f1-score	support
CDMS	0.62	0.79	0.70	19
Non-CDMS	0.85	0.72	0.78	32
accuracy			0.75	51
macro avg	0.74	0.75	0.74	51
weighted avg	0.77	0.75	0.75	51

Figure 7: Random Forest Prediction on MS Diagnosis

Conclusion: Benefits of Machine Learning in Healthcare

Machine learning has multiple positive implications for health care. In the diagnosis of serious health conditions, it eliminates human error like subjective testing. In most cases, scientists have bias over symptoms and possible conditions. Trends demonstrate that humans naturally gravitate toward certain prognoses based on recency and their accustomed knowledge of the subject. With machine learning, data is empirically analyzed, and diagnosis is a direct reflection of these results (Habehh & Gohel, 2021). Moreover, machine learning can dramatically improve the deliverability of solutions among a variety of patients. Data analysis appeared to be more scalable and accurate, reaching wider demographics in different communities. With more accessible solutions, resources can be expended carefully and efficiently, an alternative pathway in receiving medical aid (Davenport & Kalakota, 2019). The prescriptive purposes of machine learning foreshadow the new modern era of technology. Algorithms can be trained to plan an assortment of cognitive behavioral therapies, unconventional and personalized toward the patient. Health care will be revolutionized by the new world of biotechnology.

Proposal:

Future approaches to innovation in artificial intelligence and machine learning will prosper scientific revelations on cerebrospinal fluid and glymphatic systems. The intrinsic dive into this newly-discovered process will lead to an established discovery of neurodegenerative linkages and neurobiological metabolic recovery. With deep learning algorithms, biomarker identification of cerebrospinal fluid can extrapolate more data on its roles and functions in the brain. The exploration of CSF in the glymphatic structure will pioneer research on neuronal cognition and CNS diseases. The establishment of neurotechnology will promote scalable solutions and manifest in a modern blueprint for healthcare systems.

Methods Section

To view the full code, please visit EthanNgo_MSMachineLearning.ipynb in Github public repository

Figure Legend

Fig. 1. **Glymphatic Process (Perivascular Space)**

The picture demonstrates the processes that are involved with the glymphatic system. CSF travels from the perivenous space, through the basal lamina, and into the astrocytic endfeet. The tight junction facilitates the CSF transportation from the extracellular sheet into the subarachnoid space. Then, CSF and ISF are pumped into the neuron tissue, eventually leading down to the perivenous space. *Created with BioRender.com*

Fig. 2. **Astrocytic Polarization in Periarterial to Perivenous Space**

AQP4 channels are gated at the astrocytic endfeet. The astrocytes attach to both the periarterial and the perivenous walls. When CSF is pumped into the subarachnoid space and neuron tissue, the fluid follows a pressure gradient produced by these water channels. However, as aging occurs, this pressure gradient degrades because of the deterioration of AQP4. *Created with BioRender.com*

Fig. 3. **Stage I of Myelin Deterioration in MS**

T-cells first bind onto the myelin sheath of the oligodendrocytes. Cytokine chemicals then enter the bloodstream to signal macrophages and B-cells emerge. This is the first stage in the breakdown of myelin. *Created with BioRender.com*

Fig. 4. **Stage II of Myelin Deterioration in MS**

B-cells send antibodies onto the myelin sheath. This alerts the macrophages to attach to the oligodendrocytes. Gradually, the myelin sheath decays and forms into plagues, also known as the sclera. *Created with BioRender.com*

Fig. 5. **Relative Feature Importance in Females**

Random forest predicts the important factors that contribute to MS in females. The most notable contributor was periventricular MRI; illustrating the role of periventricular lesions in MS symptoms.

Fig. 6. **Relative Feature Importance in Males**

Random forest predicts the important factors that contribute to MS in males. The most notable contributors were initial symptoms and infratentorial MRI, illustrating the severe neurodegeneration by infratentorial lesions.

Fig. 7. **Random Forest Prediction on MS Diagnosis**

Deep learning predicted diagnosis of MS in non-CDMS patients with an 85% proficiency. However, in CDMS patients, random forest predicted with 62% proficiency. Greater volume of data and advanced learning models will have more accurate prognosis.

References

1. Barkhane, Z., Elmadi, J., Satish Kumar, L., Pugalenthi, L. S., Ahmad, M., & Reddy, S. (2022). Multiple Sclerosis and Autoimmunity: A Veiled Relationship. *Cureus*. https://doi.org/10.7759/cureus.24294
2. Başţanlar, Y., & Özuysal, M. (2014). Introduction to Machine Learning. In M. Yousef & J. Allmer (Eds.), *miRNomics: MicroRNA Biology and Computational Analysis* (Vol. 1107, pp. 105–128). Humana Press. https://doi.org/10.1007/978-1-62703-748-8_7
3. Benjamin Pineda. (2023). *Conversion predictors of Clinically Isolated Syndrome to Multiple Sclerosis in Mexican patients: A prospective study.* [dataset]. Mendeley. https://doi.org/10.17632/8WK5HJX7X2.1
4. Benveniste, H., Lee, H., & Volkow, N. D. (2017). The Glymphatic Pathway: Waste Removal from the CNS via Cerebrospinal Fluid Transport. *The Neuroscientist*, *23*(5), 454–465. https://doi.org/10.1177/1073858417691030
5. Bharucha, T., Gangadharan, B., Kumar, A., De Lamballerie, X., Newton, P. N., Winterberg, M., Dubot-Pérès, A., & Zitzmann, N. (2019). Mass spectrometry-based proteomic techniques to identify cerebrospinal fluid biomarkers for diagnosing suspected central nervous system infections. A systematic review. *Journal of Infection*, *79*(5), 407–418. https://doi.org/10.1016/j.jinf.2019.08.005
6. Bohr, A., & Memarzadeh, K. (2020). The rise of artificial intelligence in healthcare applications. In *Artificial Intelligence in Healthcare* (pp. 25–60). Elsevier. https://doi.org/10.1016/B978-0-12-818438-7.00002-2
7. Bueno, D., & Garcia-Fernàndez, J. (2016). Evolutionary development of embryonic cerebrospinal fluid composition and regulation: An open research field with implications for brain development and function. *Fluids and Barriers of the CNS*, *13*, 5. https://doi.org/10.1186/s12987-016-0029-y
8. Carotenuto, A., Cacciaguerra, L., Pagani, E., Preziosa, P., Filippi, M., & Rocca, M. A. (2022). Glymphatic system impairment in multiple sclerosis: Relation with brain damage and disability. *Brain*, *145*(8), 2785–2795. https://doi.org/10.1093/brain/awab454
9. Casini, G. (2013). Are Periventricular Lesions Specific for Multiple Sclerosis? *Journal of Neurology & Neurophysiology*, *04*(02). https://doi.org/10.4172/2155-9562.1000150
10. Chu, A. B., Sever, J. L., Madden, D. L., Iivanainen, M., Leon, M., Wallen, W., Brooks, B. R., Lee, Y. J., & Houff, S. (1983). Oligoclonal IgG bands in cerebrospinal fluid in various neurological diseases. *Annals of Neurology*, *13*(4), 434–439. https://doi.org/10.1002/ana.410130410
11. Davenport, T., & Kalakota, R. (2019). The potential for artificial intelligence in healthcare. *Future Healthcare Journal*, *6*(2), 94–98. https://doi.org/10.7861/futurehosp.6-2-94
12. Dekker, I., Sombekke, M. H., Balk, L. J., Moraal, B., Geurts, J. J., Barkhof, F., Uitdehaag, B. M., Killestein, J., & Wattjes, M. P. (2020). Infratentorial and spinal cord lesions: Cumulative predictors of long-term disability? *Multiple Sclerosis Journal*, *26*(11), 1381–1391. https://doi.org/10.1177/1352458519864933
13. Di Terlizzi, R., & Platt, S. (2006). The function, composition and analysis of cerebrospinal fluid in companion animals: Part I – Function and composition. *The Veterinary Journal*, *172*(3), 422–431. https://doi.org/10.1016/j.tvjl.2005.07.021
14. Fauss, D., Motter, R., Dofiles, L., Rodrigues, M. A. V., You, M., Diep, L., Yang, Y., Seto, P., Tanaka, K., Baker, J., & Bergeron, M. (2013). Development of an enzyme-linked

immunosorbent assay (ELISA) to measure the level of tyrosine hydroxylase protein in brain tissue from Parkinson's disease models. *Journal of Neuroscience Methods*, *215*(2), 245–257. https://doi.org/10.1016/j.jneumeth.2013.03.012
15. Filippi, M., Preziosa, P., Banwell, B. L., Barkhof, F., Ciccarelli, O., De Stefano, N., Geurts, J. J. G., Paul, F., Reich, D. S., Toosy, A. T., Traboulsee, A., Wattjes, M. P., Yousry, T. A., Gass, A., Lubetzki, C., Weinshenker, B. G., & Rocca, M. A. (2019). Assessment of lesions on magnetic resonance imaging in multiple sclerosis: Practical guidelines. *Brain*, *142*(7), 1858–1875. https://doi.org/10.1093/brain/awz144
16. Gao, Y., Liu, K., & Zhu, J. (2023). Glymphatic system: An emerging therapeutic approach for neurological disorders. *Frontiers in Molecular Neuroscience*, *16*, 1138769. https://doi.org/10.3389/fnmol.2023.1138769
17. Graham, C. (2022). Robot Performs First Laparoscopic Surgery Without Human Help. *Johns Hopkins University: The Hub*. https://hub.jhu.edu/2022/01/26/star-robot-performs-intestinal-surgery/
18. Habehh, H., & Gohel, S. (2021). Machine Learning in Healthcare. *Current Genomics*, *22*(4), 291–300. https://doi.org/10.2174/1389202922666210705124359
19. Hussain, L. S., Reddy, V., & Maani, C. V. (2023). Physiology, Noradrenergic Synapse. In *StatPearls*. StatPearls Publishing. http://www.ncbi.nlm.nih.gov/books/NBK540977/
20. Iliff, J., & Simon, M. (2019). CrossTalk proposal: The glymphatic system supports convective exchange of cerebrospinal fluid and brain interstitial fluid that is mediated by perivascular aquaporin-4. *The Journal of Physiology*, *597*(17), 4417–4419. https://doi.org/10.1113/JP277635
21. Javed, K., Reddy, V., & Lui, F. (2023). Neuroanatomy, Choroid Plexus. In *StatPearls*. StatPearls Publishing. http://www.ncbi.nlm.nih.gov/books/NBK538156/
22. Jessen, N. A., Munk, A. S. F., Lundgaard, I., & Nedergaard, M. (2015). The Glymphatic System: A Beginner's Guide. *Neurochemical Research*, *40*(12), 2583–2599. https://doi.org/10.1007/s11064-015-1581-6
23. Khasawneh, A. H., Garling, R. J., & Harris, C. A. (2018). Cerebrospinal fluid circulation: What do we know and how do we know it? *Brain Circulation*, *4*(1), 14–18. https://doi.org/10.4103/bc.bc_3_18
24. Letchuman, V., & Donohoe, C. (2023). Neuroanatomy, Superior Sagittal Sinus. In *StatPearls*. StatPearls Publishing. http://www.ncbi.nlm.nih.gov/books/NBK546615/
25. Mattsson-Carlgren, N., Grinberg, L. T., Boxer, A., Ossenkoppele, R., Jonsson, M., Seeley, W., Ehrenberg, A., Spina, S., Janelidze, S., Rojas-Martinex, J., Rosen, H., La Joie, R., Lesman-Segev, O., Iaccarino, L., Kollmorgen, G., Ljubenkov, P., Eichenlaub, U., Gorno-Tempini, M. L., Miller, B., … Rabinovici, G. D. (2022). Cerebrospinal Fluid Biomarkers in Autopsy-Confirmed Alzheimer Disease and Frontotemporal Lobar Degeneration. *Neurology*, *98*(11), e1137–e1150. https://doi.org/10.1212/WNL.0000000000200040
26. Mestre, H., Mori, Y., & Nedergaard, M. (2020). The Brain's Glymphatic System: Current Controversies. *Trends in Neurosciences*, *43*(7), 458–466. https://doi.org/10.1016/j.tins.2020.04.003
27. Miron, V. (2019). Immune system and MS. *Multiple Sclerosis Society UK*. https://www.mssociety.org.uk/research/explore-our-research/emerging-research-and-treatments/immunomodulation#:~:text=In%20MS%2C%20the%20immune%20system,properly%20and%20may%20eventually%20die

28. Neilson, L. E., Hollen, C., Hiller, A., & Wooliscroft, L. (2022). Oligoclonal Bands in Multiple System Atrophy: Case Report and Proposed Mechanisms of Immunogenicity. *Frontiers in Neuroscience*, *16*, 852939. https://doi.org/10.3389/fnins.2022.852939
29. NINDS. (2023). National Institute of Neurological Disorders and Stroke: Multiple Sclerosis. *U.S. Department of Health and Human Services*. https://www.ninds.nih.gov/health-information/disorders/multiple-sclerosis#:~:text=Research%20shows%20that%20genetic%20vulnerabilities%20combined%20with%20environmental%20factors%20may%20cause%20MS.&text=MS%20itself%20is%20not%20inherited,relative%20who%20also%20have%20MS.
30. Parker, C. E., Warren, M. R., & Mocanu, V. (2010). Mass Spectrometry for Proteomics. In O. Alzate (Ed.), *Neuroproteomics*. CRC Press/Taylor & Francis. http://www.ncbi.nlm.nih.gov/books/NBK56011/
31. Reddy, O. C., & Van Der Werf, Y. D. (2020). The Sleeping Brain: Harnessing the Power of the Glymphatic System through Lifestyle Choices. *Brain Sciences*, *10*(11), 868. https://doi.org/10.3390/brainsci10110868
32. Rosas, N., Sotelo, J., & Nieto, D. (1986). ELISA in the Diagnosis of Neurocysticercosis. *Archives of Neurology*, *43*(4), 353–356. https://doi.org/10.1001/archneur.1986.00520040039016
33. Shenoy, S. S., & Lui, F. (2023). Neuroanatomy, Ventricular System. In *StatPearls*. StatPearls Publishing. http://www.ncbi.nlm.nih.gov/books/NBK532932/
34. Simon, M. J., & Iliff, J. J. (2016). Regulation of cerebrospinal fluid (CSF) flow in neurodegenerative, neurovascular and neuroinflammatory disease. *Biochimica et Biophysica Acta (BBA) - Molecular Basis of Disease*, *1862*(3), 442–451. https://doi.org/10.1016/j.bbadis.2015.10.014
35. Telano, L. N., & Baker, S. (2023). Physiology, Cerebral Spinal Fluid. In *StatPearls*. StatPearls Publishing. http://www.ncbi.nlm.nih.gov/books/NBK519007/
36. Veening, J. G., & Barendregt, H. P. (2010). The regulation of brain states by neuroactive substances distributed via the cerebrospinal fluid; a review. *Cerebrospinal Fluid Research*, *7*(1), 1. https://doi.org/10.1186/1743-8454-7-1
37. Ysrraelit, M. C., & Correale, J. (2019). Impact of sex hormones on immune function and multiple sclerosis development. *Immunology*, *156*(1), 9–22. https://doi.org/10.1111/imm.13004
38. Zhu, G., Zhou, S., Xu, Y., Gao, R., Zhang, M., Zeng, Q., Su, W., & Wang, R. (2023). Chickenpox and multiple sclerosis: A Mendelian randomization study. *Journal of Medical Virology*, *95*(1), e28315. https://doi.org/10.1002/jmv.28315