Aging, Dementia and TBI Data Analysis

Introduction

With an aging population that continues to grow, we are seeing more and more people having neurodegenerative diseases, such as Dementia and Alzheimer's disease. These diseases are caused by damage to nerve cells (neurons) in the brain. About 11% of the US population over the age of 65 are believed to be affected by dementia. The incidence of dementia onset is estimated to double every five years to about 40% to 60% after age 90. This significantly increases demands on the public health system and on medical and social services, causing large economic burden and healthcare challenges.

As a group of data scientists who want to analyze real world data to get useful insights, we have worked on datasets from the Adult Changes in Thought (ACT) study (http://aging.brain-map.org/overview/home). The ACT study is a longitudinal population-based prospective cohort study of brain aging and incident dementia in the Seattle metropolitan area. We are interested in understanding the relationship between cognition, brain pathology and injury in the aged brain.

We analyzed the donors' de-identified clinical information and tissue samples' metrics from different imaging technologies such as Histology and immunohistochemistry (IHC) and In situ hybridization (ISH). We built a supervised learning model that predicts whether a person has dementia or no dementia using demographic information and certain metrics from tests. We also drew useful insights from the visualization we built.

Data Overview

There are a total of 107 participants (63 males and 44 females) with a wide range of educational backgrounds. The cohort is quite old, ranging from 77 to 102 years old at time of death. Median age is 90. (Figure 1)

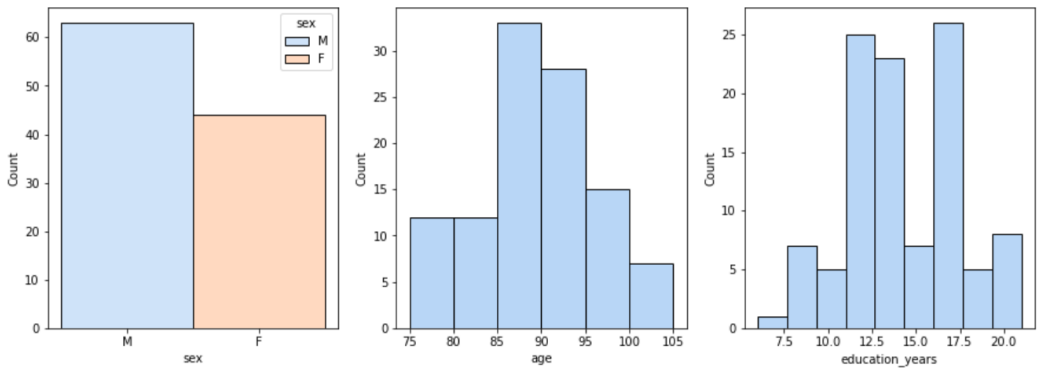


Figure 1

Around half of the donors were diagnosed with dementia. According to DSM-IV (Diagnostic and statistical manual of mental disorders) clinical diagnosis, 30 are diagnosed with AD, 12 with dementia of multiple etiologies, and four with vascular dementia. (Figure 2)

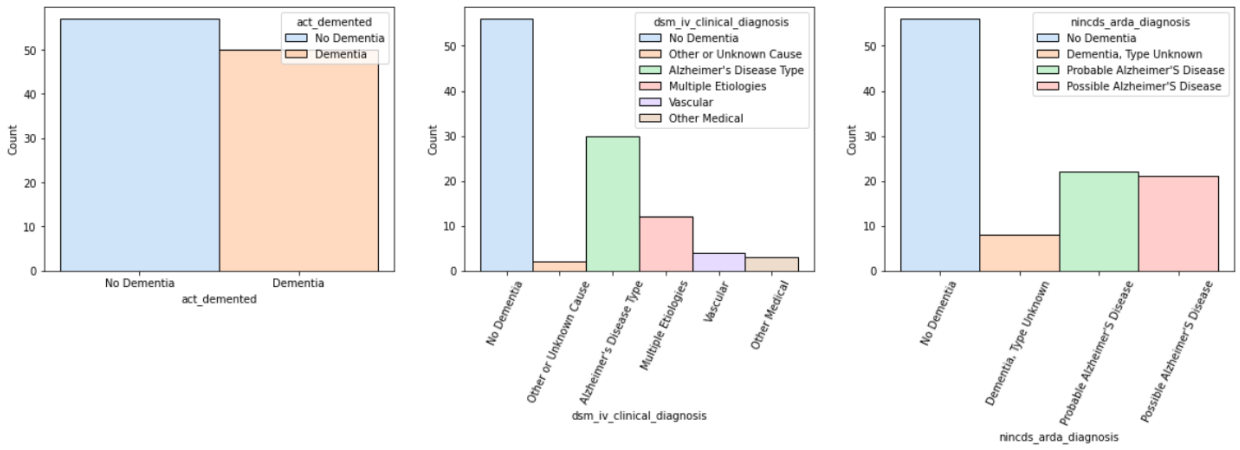


Figure 2

Dementia is an overall term for a particular group of symptoms. The characteristic symptoms of dementia are difficulties with memory, language, problem-solving and other thinking skills. Dementia has several causes that reflect specific changes in the brain: Alzheimer's disease, Parkinson's disease dementia and Lewy body disease.

Approximately 2/3 of dementia cases are diagnosed as Alzheimer's disease (AD), estimated to reach 13.8 million cases by 2050. The brain changes of Alzheimer's disease include the accumulation of the abnormal proteins beta-amyloid (A**β** plaques) outside neurons and twisted strands of the protein tau (tangles) inside neurons in the brain. These changes are accompanied by the death of neurons and damage to brain tissue.

For each donor, tissues were collected from four brain regions known to show neurodegeneration and pathology as a result of AD and Lewy body disease. They are frontal white matter(FWM), temporal cortex (TCx), parietal cortex(PCx), hippocampus(HIP).

The cerebrum represents one of the largest regions of the brain. The cerebral cortex serves as the outer layer of the cerebrum and it consists of mostly of gray matter whereas white matter lies in the center. Four lobes make up the cerebral cortex: the frontal lobe, the parietal lobe, the temporal lobe, and the occipital lobe. Each lobe has a distinct function.

TODO: decide which image to use.

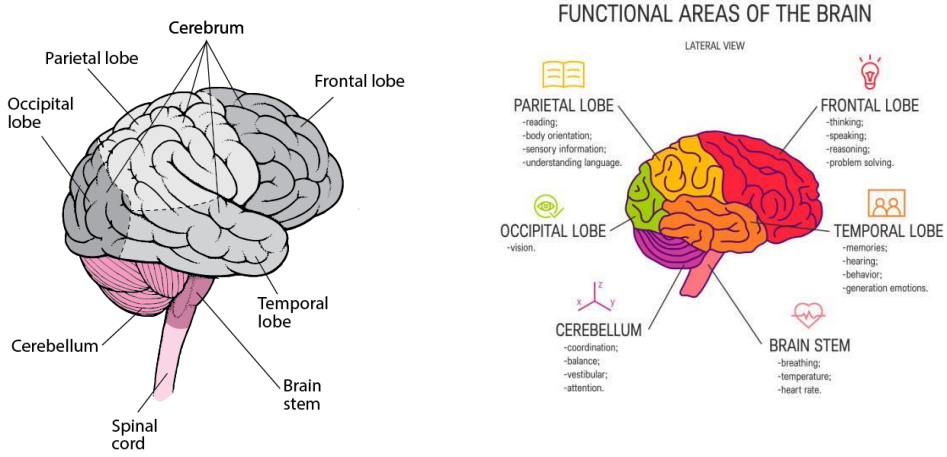


Figure 3

Immunohistochemistry (IHC) was used on both fresh frozen and formalin fixed paraffin embedded (FFPE) tissue to stain and quantify proteins marking dementia-related pathologic findings, including pTau, Aβ, α-synuclein (Lewy bodies), and pTDP-43, as well as microglia (IBA1) and astrocytes (GFAP).

Immunohistochemistry (IHC) an imaging technique to visualize antigens in cells. Labeled antibodies bind to target antigens in the cell to image the distribution and localization of specific proteins of interest.

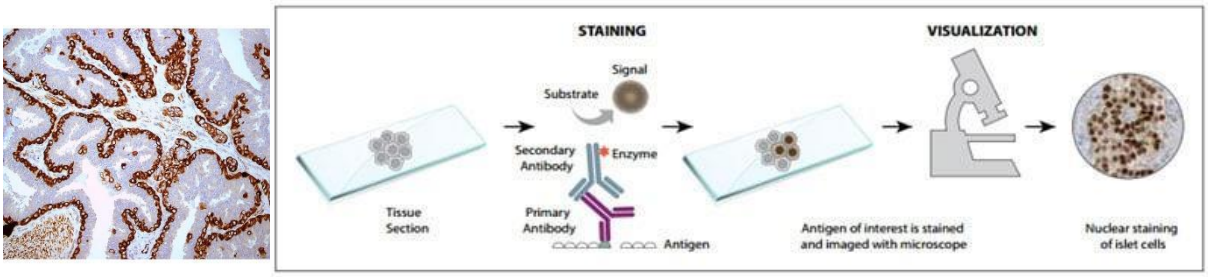
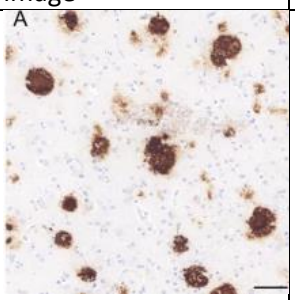
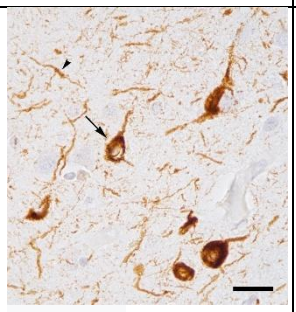
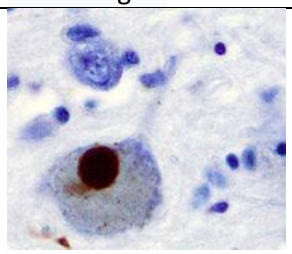


Figure 4

Here is a table which provides more details about the proteins we are interested in.

protein	description	measure	image	location
beta-amyloid (A**β**)	Aβ is formed from the breakdown of a larger protein, called amyloid precursor protein(APP) and is toxic. Abnormal levels of this naturally occurring protein clump together to form plaques that collect between neurons and disrupt cell function.	CERAD (Consortium to Establish a Registry for Alzheimer's Disease) score: 0 (none) to 3 (severe)	 Aβ plaques in an AD patient with an	between neurons

	Cerebrospinal fluid (CSF) levels of amyloid-beta 42(Aβ 42) serve as an excellent marker for brain amyloid as detected by the amyloid tracer, Pittsburgh compound B (PIB).		autosomal dominant pattern (ADAD) caused by mutations in the amyloid precursor protein (APP) or in the two presenilin (PSEN1 and PSEN2) genes.	
pTau	Tau normally binds to microtubules and assists with their self-assembly, formation and stabilization. However, when tau is hyperphosphorylated, it is unable to bind and the microtubules become unstable and begin disintegrating. The unbound tau clumps together in formations called neurofibrillary tangles (NFT).	Braak stage: 0 (none) to 6 (extensive neocortical NFTs)	 Abnormal accumulation of tau protein, which constitutes NFT, in neuronal cell bodies (arrow) and neuronal extensions (arrowhead) in the neocortex of a patient who died with AD at Braak stage 6	inside neurons
alpha-synuclein (α-synuclein)	Alpha-synuclein aggregates forming Lewy bodies, a spherical masses in the cytoplasm that displace other cell components, restricting the mobility of synaptic vesicles, consequently attenuating synaptic vesicle recycling and neurotransmitter release		 Positive Alpha-Synuclein staining of a Lewy body in a patient with Parkinson's disease	inside neurons

Braak Staging is the method used to classify the degree of pathology in Parkinson's and Alzheimer's disease. For AD patient, Braak stages 1 and 2 are used when neurofibrillary tangle involvement is confined mainly to the entorhinal cortex, stages 3 and 4 when there is also involvement of limbic regions such as the hippocampus, and 5 and 6 when there is extensive neocortical involvement. In the following image (Figure 5), yellow represents the origin of Parkinson's pathology. Pink/purple represent Stages 1

and 2. Blue represents Stages 3 and 4. Orange represents Stage 5. Yellow represents full neocortex engagement and Stage 6.

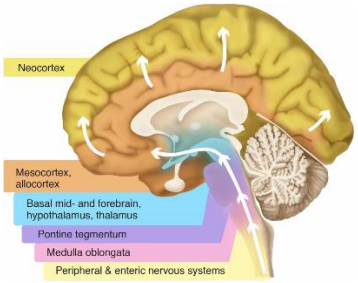


Figure 5

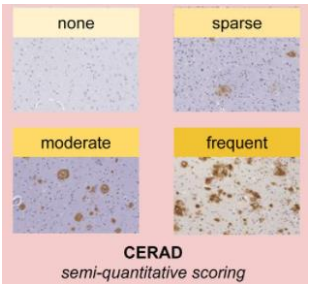


Figure 6

CERAD score is a measure of neurotic plaques which is on a 4-point scale (0 to 3): normal, mild, moderate, severe impairment based on semiquantitative estimates of neurotic plaque density as recommended by the Consortium to Establish a Registry for Alzheimer's Disease (CERAD). Figure 6 shows the IHC stained plaques.

Data Analysis on the donor information

We performed data analysis on the donor information. The following table shows the summary of demographics for donors with and without dementia.

Attribute	Dementia		No Dementia		
	Mean	SD	Mean	SD	p-value
Age at death (years)	90.3	6.5	88.6	7.0	0.21
Education (years)	13.5	3.2	14.7	3.2	0.07
Number of TBIs	0.6	0.7	0.6	0.7	0.87
Age at first TBI	21.1	30.2	20.6	30.7	0.94
Braak stage	4.1	1.7	2.8	1.5	8.7E-05
NIA Reagan	1.9	0.9	1.4	0.7	1.2E-03
CERAD score	1.8	1.2	1.2	0.9	0.01
	Count		Count		
APOE ε4 alleles	13 Yes/32 No		7 Yes/48 No		
Gender	27 M/23 F		36 M/21 F		

Let's take a closer look at the relationship between different demographics information. Figure 3 shows the boxplots for Age and CERAD score, Age and Braak Stage. It is obvious that older donors are likely to have higher CERAD score and Braak stage.

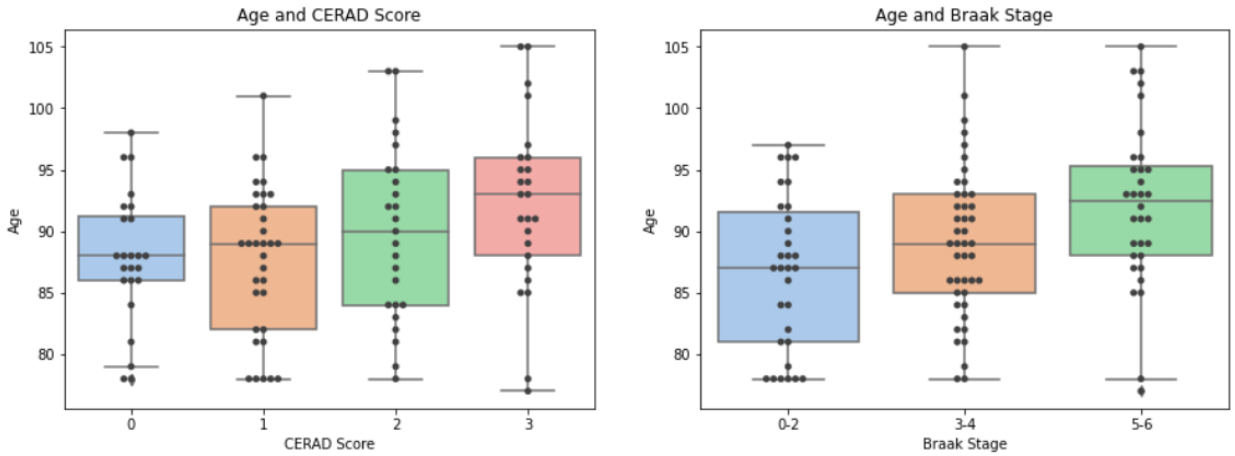


Figure 3

CERAD score and Braak stage are generally higher in donors with dementia compared to donors without dementia, as expected (Figure 4).

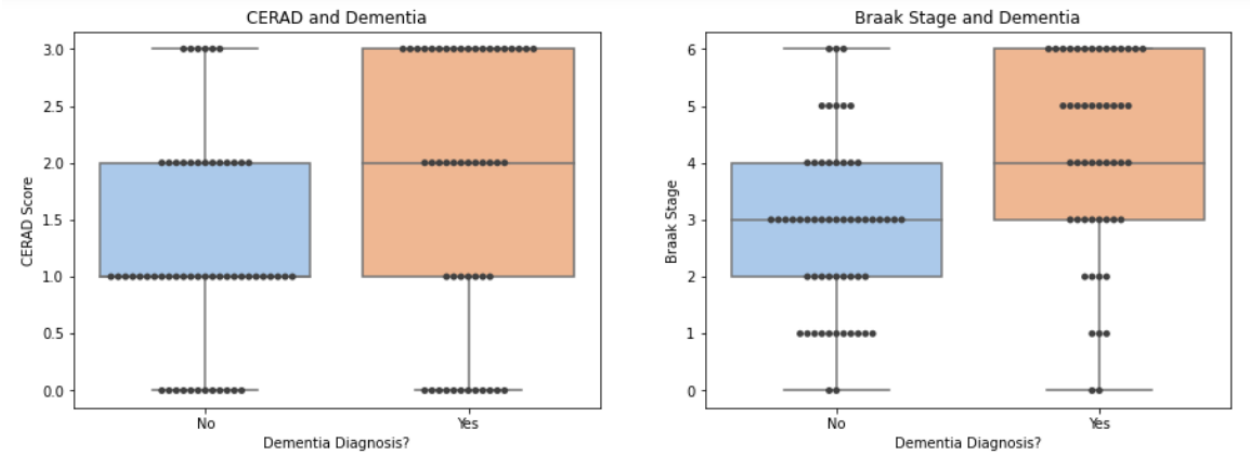


Figure 4

The following Sankey diagram depicts the relationship between brain region, braak stage and dementia status. The left column shows the four brain regions, the middle column shows the 7 braak stage number. The right column shows the two dementia status. The width of each arrow between columns and the height of each bar of the column are based on the quantity. We can observe that samples are evenly distributed among 4 brain regions. Very few samples have braak stage of 0. More samples have braak stage of 3 than any other stage. Most samples with braak stage of 5 or 6 have dementia and most samples with braak stage of 1 or 2 or 3 have no dementia. Almost half of the samples with braak stage of 4 have dementia.

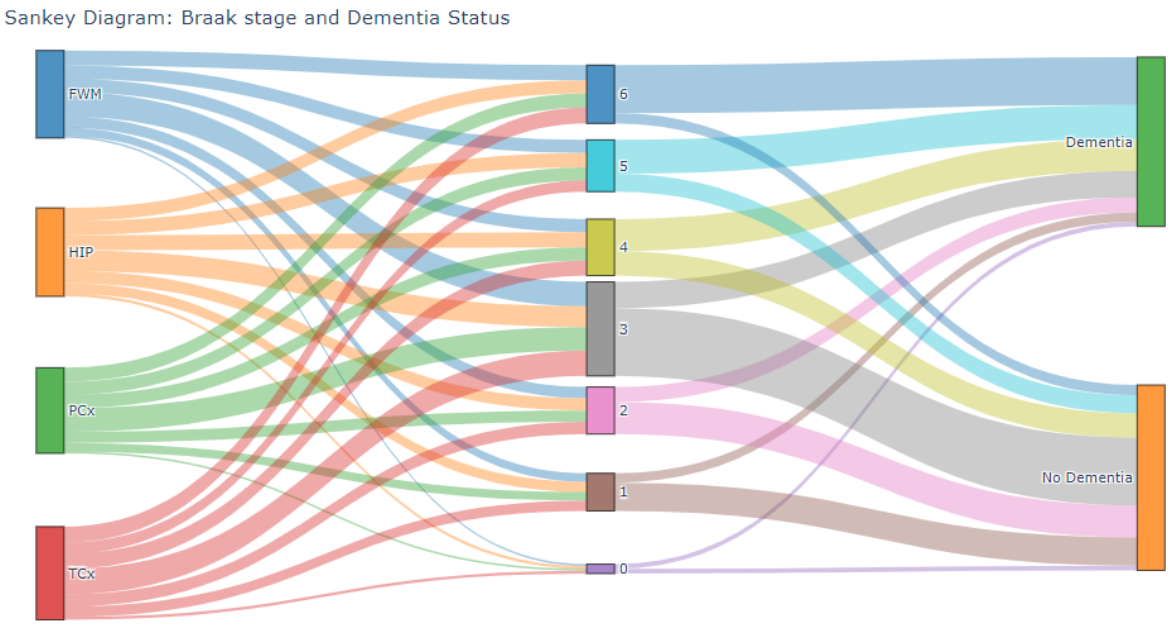


Figure 5 Braak Stage and Dementia Status

Our dataset includes 53 participants reported 1-3 lifetime TBI (Traumatic Brain Injury) with loss of consciousness, along with 54 individuals matched for age, sex, and year of death who did not report a TBI with loss of consciousness.

The following is another Sankey diagram that depicts the relationship between brain region, number of TBI and dementia status. As the samples with no TBI or 1 TBI are almost evenly distributed between the two dementia status, we don't really see any strong correlation between the number of TBI and dementia status.

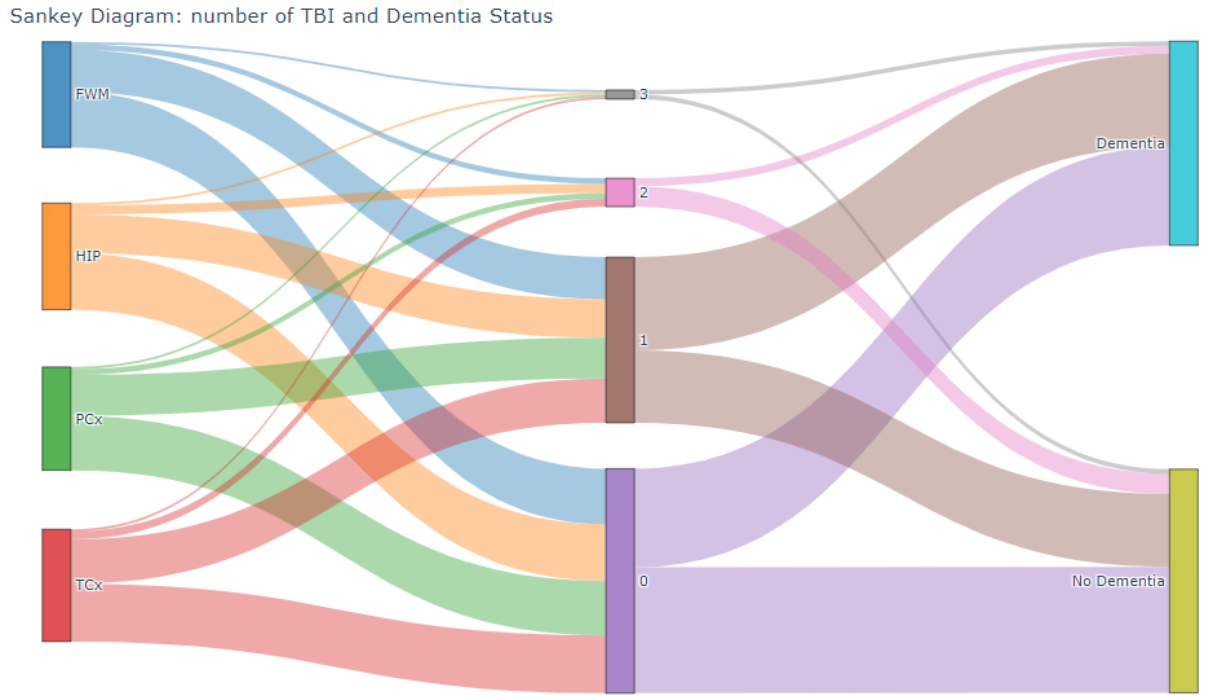


Figure 6

65% of the *APOE* ε4-positive participants had dementia while 40% of the *APOE*ε4 negative participants had dementia, consistent with the role of this gene as a strong risk factor for AD. Inheriting APOE4 does not mean a person will develop the disease, it does increase the risk for dementia. The APOE protein helps carry cholesterol and other types of fat in the bloodstream. Recent studies suggest that problems with brain cells' ability to process fats, or lipids, may play a key role in Alzheimer's and related diseases.

Data Analysis on the Protein and Pathology information

The protein and pathology metrics for tissue specimens were collected using traditional antibody based IHC methods complemented by Luminex quantification and isoprostane quantification.

Tau protein have roles primarily in maintaining the stability of microtubules in axons and are abundant in the neurons of the central nervous system (CNS). However, when tau is hyperphosphorylated, it is called pTau which is unable to bind and the microtubules become unstable and begin disintegrating. The unbound tau clumps together in formations called neurofibrillary tangles (NFT).

Phosphorylation-dependent anti-tau antibodies such as monoclonal antibody AT8 is widely used to recognize pTau. We can see from our dataset that the AT8 IHC and pTau Luminex level are strongly correlated (Pearson correlation coefficient is 0.68).

We calculated Pearson correlation coefficient between all the protein and pathology qualifications and all the attributes of donors and find that the following correlations are moderate and strong ($r > 0.4$).

Donor attribute	Protein/pathology quantification	correlation
CERAD	ihc_a_beta	0.43
CERAD	ihc_at8_ffpe	0.41

CERAD	ab42_pg_per_mg	0.63
BRAAK	ihc_at8_ffpe	0.46
BRAAK	ab42_pg_per_mg	0.46
NIA Reagan	ihc_at8_ffpe	0.44
NIA Reagan	ihc_a_beta	0.4
NIA Reagan	ab42_pg_per_mg	0.58

It seems that AT8 IHC, Aβ IHC and Aβ42 are moderate correlated with all CERAD score, Braak Stage and NIA Reagan. As criteria for NIA Reagon diagnosis relies on both neurofibrillary tangles (Braak) and neuritic plaques (CERAD), we didn't include RIA Reagon in our analysis. We focused our analysis on protein molecular quantification of Aβ and pTau.

Explore Aβ and pTau correlations in different brain regions and different braak stages

Some research studies suggest that formation of plaques and tau phosphorylation might be linked to each other. We analyzed the correlation of Aβ and pTau for the whole dataset and each brain region. From Figure 7, we saw that Aβ and pTau have week correlation (Pearson correlation coefficient is 0.33) overall for all brain regions. But they have moderate correlations in hippocampus and cortex while having no or very weak correlation in frontal white matters. Also the correlations in later Braak stages(5 and 6) are much stronger than early stages (0-4).

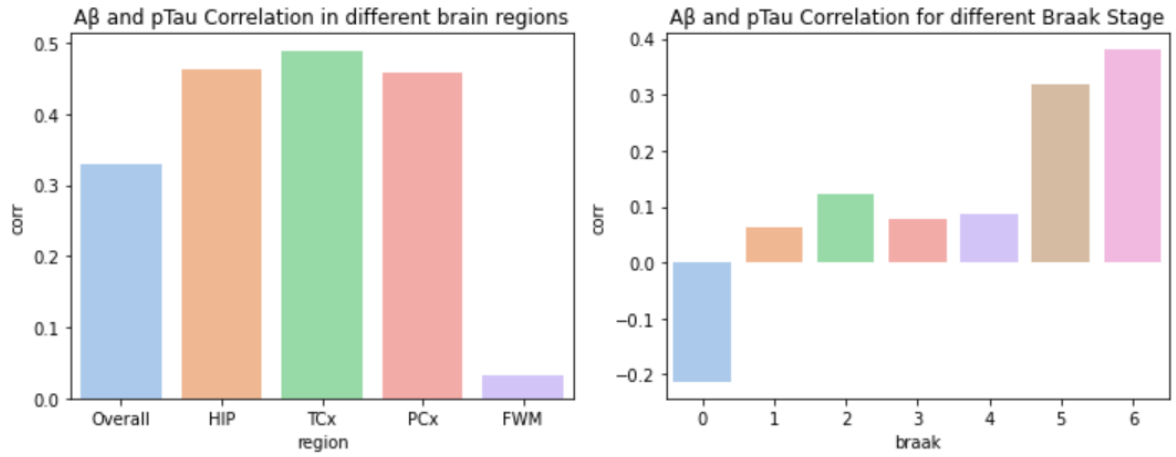


Figure 7

Our analysis reveals pathological tau (AT8 IHC and pTau Luminex) tended to be higher in hippocampus while Aβ (Aβ IHC and Aβ42 Luminex) is higher in cortex, consistent with known AD pathological distributions and progression just as shown in Figure 8

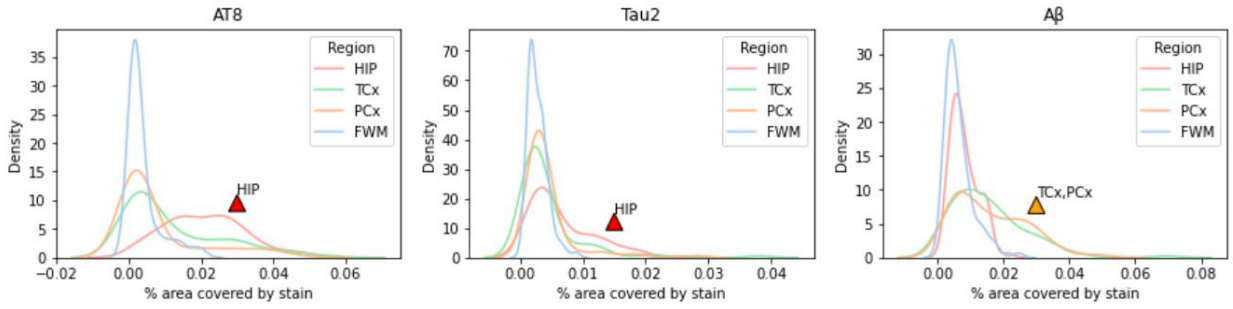

AT8
Tau2
Aβ
Region
HIP
TCx
PCx
FWM
Density
% area covered by stain
HIP
HIP
TCx,PCx

Figure 8