Developing a Predictive Model to Predict the Progression of Alzheimer's Disease via a Neural Apoptotic Marker

Grant Proposal

Rishab Nair

Massachusetts Academy of Math and Science at WPI

Worcester, MA. USA

Executive Summary

This research initiative employs a two-phase approach to optimize Alzheimer's disease (AD) detection and progression. The initial phase targets the identification of unique markers associated with apoptotic neuronal cell death in AD, highlighting Cytochrome C and caspases as potential candidates. While Cytochrome C shows promise as a neuronal cell death indicator, its certainty is constrained by a limited time frame. Conversely, the complexities surrounding the activation of caspases—particularly executioner caspases—introduce uncertainties due to discrepancies between their activation and actual cell death. Contrary to these potential drawbacks, Cytochrome C and caspases still are widely used in existing studies to predict for general cell death, and provide opportunity to behave the same in neuronal cells.

In the second phase, the aim is to develop a predictive model using Artificial Intelligence (AI) methods, leveraging advancements in clinical bioinformatics to enable early detection and personalized treatment strategies. This research endeavor aims to optimize methods of AD detection and progression assessment utilizing AI-based predictive modeling via identifying specific neuronal cell death markers in Phase I. The ultimate objective is to provide experts with sophisticated tools for more informed, faster decision-making in managing Alzheimer's disease.

Keywords: Alzheimer's Disease, Apoptosis, Cytochrome C, Caspases, Artificial Intelligence, Clinical bioinformatics, Treatment

Developing a Predictive Model to Compute the Progression of Alzheimer's Disease via a Neural Apoptotic Marker

As of 2023, an estimated 6.7 million Americans aged 65 and older suffer from Alzheimer's. Within this population, Alzheimer's remains as the fifth-leading cause of death. By 2060, the population could grow to 13.8 million barring the development of medical breakthroughs to prevent, slow or cure the disease (Alzheimer's Association, 2023). Alzheimer's does not solely have such significance in just America, however; as of 2020, the disease has been classified as a world health concern (Breijyeh et al. 2020).

This project consists of two phases: determining a marker that can detect areas in the brain affected by Alzheimer's disease and using this marker to develop a predictive model to determine the progression of the disease so that medical professionals obtain a greater understanding of its behavior.

It is known that massive neuronal death in Alzheimer's disease is a result of apoptosis, or programmed cell death (Shimohama, 2000). Past studies on cell death and apoptosis have focused on potential markers that can detect the causation of apoptosis. However, these studies tend to center on general cells– not neurons. In its first phase, this project aims to close this research gap by determining a possible correlation of neuronal cells to existing cell death marker studies. The project aims to determine a specific cell death marker to detect affected areas in Alzheimer's.

Current methods of Alzheimer's diagnosis occur late into the progression of the disease, and detriment the effectiveness of therapeutic programs (Emory, 2011). Current therapeutic plans are also relatively generalized, although multiple different versions of the disease occur.

Therefore, in its second phase, this project seeks to mitigate these issues by creating a predictive model. This model aims for efficiency in detecting the disease in earlier stages, as well as prediction of its progression. Using this information, doctors can optimize and personalize the patients' therapeutic plans.

Cytochrome C: Potential Marker #1

To initiate apoptosis, the cell's mitochondria release several proteins into the cytosol. One such protein is Cytochrome C, a special type of protein normally found in the intermembrane space (IMS) of mitochondria (Cortese et al., 1995).

Upon cytosolic localization—or when the protein is released into the cytosol—Cytochrome C binds to apoptotic protease-activating factor 1 (APAF1). The binding then leads to the oligomerization of APAF1 and the formation of what is known as the apoptosome (Naniche et al., 2011). This product serves as the platform for activating what is known as caspases—a type of endoprotease—which causes the cell to die. The first activated caspase is known as initiator caspase 9, which then cleaves and consequently activates executioner caspases 3 and 7, leading to cell death (Naniche et al., 2011).

As Cytochrome C is generally viewed as the initiator for the apoptotic process, it is considered to be a good marker of cell death. The experiment that determined Cytochrome C's effectiveness concerned cell death in neurons in vitro and in vivo as well (Naniche et al., 2011), serving as another element to its potential benefit as a neuronal apoptotic marker. Cytochrome C does pose some drawbacks in terms of certainty, as the process has a chance of termination as a result of a time window between the cytosolic localization of the protein and the act of apoptosis itself (Naniche et al., 2011).

Caspases: Potential Marker #2

Caspases, a type of endoprotease, play a major role in the process of cell death. There are three major types of caspases- Initiator, Executioner, and Inflammation. Each caspase plays a different role in the grand scheme of cell death and have been found to hold connections to certain diseases (McIlwain et al., 2013).

Caspases have been broadly classified by known roles in apoptosis (3, 6, 7, 8, 9 in mammals) and inflammation (1, 4, 5, 12 in humans, 1, 11, 12 in mice) (McIlwain et al., 2013). Caspases involved in apoptosis, the focus of this project, are subclassified by their mechanism of action: initiator caspases (caspase-8 and -9) or executioner caspases (caspase-3, -6, and -7). Initiator caspases activate executioner caspases which subsequently coordinate activities to demolish key structural proteins and activate other enzymes—in other words, they act as the switch for apoptosis (McIlwain et al., 2013). As executioner caspases are the closest to the actual event of cell death, it is most likely that these compounds serve as the most accurate indication of apoptosis.

Unfortunately, the activation of executioner caspases may not necessarily mean apoptosis has occurred. A study was conducted within the HeLa cell line, where it was determined that high caspase-3 activity levels unsurprisingly killed all cells. In contrast, low levels kept the cells alive. Caspase-3 activity doses fit to kill median amounts of 15 to 30% of cells, however, still let 70 to 85% of cells live. With these doses, the rate, peak level, or the total amount of caspase-3 activity wasn't able to accurately predict cell death vs. survival (Nano et al., 2023). Therefore, caspase-3 may be faulty in its role as a marker. However, as this study was

conducted with HeLa cells, its conclusions may not adhere to the characteristics of neuronal cells.

Artificial Intelligence in the Medical Field

When conducting analysis of data such as brain scans of Cytochrome C and caspases, it is evident that manual analysis may not be ideal. Artificial Intelligence (AI) is a growing field, especially so in medicine and biotechnology. Machine Learning (ML) currently can be found to aid in disease diagnosis and precision medicine, or personalized medicine, and contributes to disease detection, diagnosis, and prediction; diseases such as metastatic breast cancer, atrial fibrillation, and urinary tract infection have benefited from ML computations (Bhardwaj et al., 2022). AI-based biomarkers assist in patient diagnosis, treatment response, and survival prediction as well, by taking its predictive elements into account. As AI grows as an area of research within the medical field, improvement of efficiency and reliability of epidemiological models will only grow clearer (Bhardwaj et al., 2022).

Clinical bioinformatics, a term that defines the usage of bioinformatics techniques for the identification of diseases, discovery of biomarkers, and therapy decision, is a newly emerging field, and the data used to make decisions in this field can, in turn, become very difficult for humans to manually handle. Mathematical modeling, via a computer model, must be used to alleviate this process. Predictive modeling has been proposed by researchers in the medical field, combining its elements with bioinformatics for improving current practices in disease identification, therapeutics, and prognostics (Pais, 2022). It has shown advantageous to unlock the full potential of high-throughput technologies as solutions for large population screening of multiple disorders and precision medicine (Pais, 2022). Of course, there is always a

certain degree of uncertainty associated with any prediction generated by any model, so predictive models should be seen as insights that enable clinicians to make better informed, supported decisions (Pais, 2022).

Table 1

Often used modelling techniques in clinical bioinformatics and their main characteristics.

Modelling Technique	Description	Application	Requirements
Statistical	Scoring and probability functions that assumes a distribution shape or behaviour.	Continuous Quantification	Data for parameter estimation. Depend on sample size.
Kinetic	Solving of systems of nonlinear differential equations. Do not assume any behaviour. Instead relies on rate laws of processes such as chemical reactions.	Binary Classification	Requires reported or estimated kinetic parameter. Do not depend on sample size.
Logical	Solving of logical equations based on predefined rules for each component. Assumes asynchronous or synchronous update schemes.	Binary Classification	Requires relational knowledge of its components. Do not depend on sample size.
Regression	Fitting of an assumed mathematical equation on data. Often are used models that describe a particular assumed data behaviour such as linear, polynomial, exponential, and logistic.	Binary Classification	Data for model fitting. Depend on sample size.
Random Forests	Supervised machine leaning algorithm based on averaging multiple generated decision trees.	Binary Classification	Data for model training and validation. Requires large datasets
Support Vector Machines	Supervised machine leaning algorithm based on clustering algorithms such as principal component analyses.	Binary Classification	Data for model training and validation. Requires large datasets
Neural Networks	Supervised machine leaning algorithm based on defining a set of neuron and layers as model components. Assumes all possible relational interactions between neurons.	Binary Classification	Data for model training and validation. Requires large datasets

Figure 1. Some of the most used modeling techniques in the medical setting (Pais, 2022).

Section II: Specific Aims

This proposal's objective is to determine a marker that can detect areas in the brain affected by Alzheimer's disease, and to use this marker to develop a predictive model to determine the diagnosis and progression of the disease.

The long-term goal is to increase the effectiveness of current diagnoses and therapeutic strategies for Alzheimer's Disease, where the central hypothesis of this proposal is to provide additional data for medical professionals to base their decisions on. The rationale is that with the data provided by the predictive model, medical professionals will be able to diagnose patients due to earlier indications of the disease via extensive evidence provided by the marker,

as well as create personalized therapeutic programs that are far more effective than existing, generalized medicines.

Specific Aim 1. Determine an apoptotic marker that is most prevalent and informative in those with Alzheimer's using existing studies and focused brain scans.

Specific Aim 2. Determine and generate a predictive model to compile data of brain scans highlighting this marker over a period of time.

Specific Aim 3. Code the determined model and optimize for a >90% accuracy.

The expected outcome of this work is a technology which paves the way for a potential cure for Alzheimer's via looking at the disease through its apoptotic factors rather than the beta-amyloid plaques, which are plaques in the brain arising from Alzheimer's disease and are generally used for Alzheimer's research, to create a new promising research angle.

Section III: Project Goals and Methodology

As of 2023, an estimated 6.7 million Americans aged 65 and older suffer from Alzheimer's. Within this population, Alzheimer's remains the fifth-leading cause of death. By 2060, the population could grow to 13.8 million barring the development of medical breakthroughs to prevent, slow or cure the disease (Alzheimer's Association, 2023). Alzheimer's does not solely have such significance in just America; as of 2020, the disease has been classified as a world health concern (Breijyeh et al. 2020).

Looking at the recent history of drug trials for Alzheimer's disease, one notices a cautious approach, resulting in a lagging medical industry in terms of treatments for the

problem. Current solutions employing the targeting of beta-amyloid plaques–abnormal protein clumps that accumulate in the brain of patients with Alzheimer's disease –in established Alzheimer's disease have been unsuccessful as it is thought that treating established disease may be too late (Waite, 2015). By aiming an approach at a different angle, such as the apoptotic markers mentioned before, there holds potential benefit of new, virtually untapped research methods. In the few studies that do employ caspase -3 as a general apoptotic marker, for example, there is some promise in its capability to signal cell death. However, it is also mentioned that apoptosis may occur in low levels, meaning questions still arise regarding its significance in relation to the overall trend of the disease (Stadelmann et al., 1999).

Methodology

Thorough background research is conducted using easily accessible sources to understand possible markers associated with apoptosis within Alzheimer's disease, such as Cytochrome C, executioner caspases -3, -7, or both–mentioned in the introduction. After working with an appropriate lab for support, functional Magnetic Resonance Imaging (fMRI) scans, the most common type of brain scan used for targeted imagery, will be used to determine the markers' significance and rate of activation in identifying the damaged brain regions. The next stage is gathering various brain scan datasets–from public online datasets–to examine and validate patterns from the potential markers. The procedure reiterates the fMRI step where the pattern for one potential marker doesn't hold strength.

Creating a predictive model involves teaming up with a mentor skilled in Artificial Intelligence (AI) and Machine Learning (ML). To gauge the path of Alzheimer's disease, this

model is finely tuned to highlight the identified marker patterns. Initially, a basic model is implemented, then consistently refined to boost accuracy in prediction.

Every step in this process calls for meticulous attention and careful consideration; each individual stride carries substantial weight, significantly impacting the overall outcome of the project. Maintaining a focused approach and exercising caution throughout this methodology is key to crafting a dependable predictive model for monitoring Alzheimer's disease progression.

Specific Aim #1

Determining the most prevalent and informative apoptotic marker in Alzheimer's is crucial to this project as the objective is to receive as much information of the progression of individuals with Alzheimer's disease. The methodology involves an extensive review and analysis of existing research literature focusing on apoptotic markers within Alzheimer's disease. This approach involves meticulous examination and comparison of various apoptotic markers, primarily Cytochrome C and caspases, elucidating their prevalence and informativeness in correlation with Alzheimer's pathology. The rationale behind this method lies in the critical need to determine an independent variable within the context of Alzheimer's disease for the subsequent predictive model. Given the complex nature of AD and its association with neuronal cell death through apoptosis, pinpointing the most prominent apoptotic marker holds significance in enhancing early detection, understanding disease progression, and potentially tailoring personalized therapeutic interventions. By undertaking a comprehensive review and analysis, the goal is to unveil an apoptotic marker that could serve as a crucial diagnostic and prognostic indicator in the management of Alzheimer's disease.

Potential Pitfalls and Alternative Strategies. If all researched potential markers fail, a combination factor of the individual markers will be used to determine the overall progression instead. The complex nature of the combination may in turn result in further difficulty in creating the predictive model; the model would then require multiple input variables to determine an outcome.

Specific Aim #2

The objective is to identify the most effective model capable of compiling and analyzing brain scan data, specifically emphasizing the progression of the marker's presence over a defined period within Alzheimer's disease. The methodology involves a comprehensive review of existing models used in neuroimaging and data analysis domains. Many data compilation models will be considered regarding their adaptability in tracking and highlighting the chosen specific marker for disease progression. Relying on available literature and expertise in computational neuroscience, the approach will focus on selecting models capable of not only collating brain scan data but also discerning patterns and variations in the marker's prevalence across different stages of Alzheimer's disease progression. The rationale behind this method lies in the urgency to identify a model that efficiently captures temporal changes in the marker's existence to reflect the prognosis of the disease, potentially allowing for earlier diagnostic capabilities and personalized treatment plans.

Potential Pitfalls and Alternative Strategies. As this part of the project is purely computational, no potential pitfalls are predicted.

Specific Aim #3

The objective involves implementing and refining the now selected primitive model. The focus will be on coding and fine-tuning the model to achieve an accuracy rate exceeding 90%. Meticulous design and model optimization will be conducted to effectively track and interpret temporal variations in the marker's prominence over time to the best extent in the allotted timeframe. This rationality behind this objective lies in its aim to enhance the model's predictive capacity, ensuring a high level of accuracy in outlining the progression of the disease, ultimately contributing to more precise disease monitoring and potential diagnostic advancements in Alzheimer's research.

Potential Pitfalls and Alternative Strategies. As this part of the project is purely computational as well, no potential pitfalls are predicted.

Section IV: Resources/Equipment

Firstly, access to a well-equipped laboratory facility is essential for conducting comprehensive analyses and experiments to identify apoptotic markers associated with Alzheimer's disease. Collaborating with experts and potential mentors possessing expertise in neurology is crucial for guidance and support. Utilizing functional Magnetic Resonance Imaging (fMRI) technology is crucial to confirm the effectiveness of current compounds—potentially Cytochrome C or executioner caspases -3 and -7 —in pinpointing affected brain regions. Obtaining diverse datasets consisting of multiple brain scans containing the identified compounds is necessary for detailed analysis. Lastly, engaging in mentorship to navigate the complexities of AI/ML models for simulating Alzheimer's disease progression is also crucial; this

collaboration will aid in selecting and constructing an optimal predictive model, refining it to enhance accuracy in predicting disease diagnosis and progression.

Section V: Ethical Considerations

Although helpful, it is crucial to remain vigilant of computational errors and potential bias of information originating from AI. It is imperative to remember that the final diagnosis should be conducted by experts. This project may seem to entertain the idea of having AI serve as a substitute for medical diagnoses; on the contrary, the goal of this project is solely to provide additional data for those with the ability to make decisions with an increased level of confidence.

Section VI: Timeline

1. Background

Using accessible sources, find a general understanding of what the marker may be (current standings: Cytochrome C, executioner caspases 3, 7, or both)

2. Equipment

Find a lab willing to help my search for the answer.

3. Process

Use fMRI scans(most probable) to confirm whether one of the current compounds is sufficient to determine the location of affected brain areas.

4. Data sets

Collect multiple different data scans of the compound in brain scans. This can potentially be found online for public usage.

5. Confirmation

Using data from 4, determine whether the pattern chosen in 3 is sufficient and consistent. If not, choose a different pattern and restart from 3.

6. Presenting the Solution

After a pattern is selected, choose an existing test (or if ability is given, create a new test entirely) that pinpoints this pattern.

7. AI Model to Simulate Alzheimer's Progression

    a) Find a mentor willing to help me with choosing and writing an optimal AI/ML model

    b) Work with the mentor and using the solution in 6 as the focus variable, create a predictive model to predict the progression of Alzheimer's disease

        i.) Find a model to base the algorithm off of

        ii.) Optimize model (increase success rate)

Section VIII: References

Alzheimer's Association (2023). 2023 Alzheimer's disease facts and figures. *Alzheimer's & dementia : the journal of the Alzheimer's Association, 19*(4), 1598–1695. https://doi.org/10.1002/alz.13016

Bhardwaj, A., Kishore, S., & Pandey, D. K. (2022). Artificial Intelligence in Biological Sciences. *Life (Basel, Switzerland)*, *12*(9), 1430. https://doi.org/10.3390/life12091430

Breijyeh, Z., & Karaman, R. (2020). Comprehensive Review on Alzheimer's Disease: Causes and Treatment. *Molecules (Basel, Switzerland)*, *25*(24), 5789. https://doi.org/10.3390/molecules25245789

McIlwain, D. R., Berger, T., & Mak, T. W. (2013). Caspase functions in cell death and disease. *Cold Spring Harbor perspectives in biology*, *5*(4), a008656. https://doi.org/10.1101/cshperspect.a008656

Naniche, N., Sau, D., & Pasinelli, P.. (2011). In Vivo and In Vitro Determination of Cell Death Markers in Neurons. *Methods in Molecular Biology*, *793*, 9–21. https://doi.org/10.1007/978-1-61779-328-8_2

Nano, M., Mondo, J. A., Harwood, J., Balasanyan, V., & Montell, D. J. (2023). Cell survival following direct executioner-caspase activation. *Proceedings of the National Academy of Sciences*, *120*(4), e2216531120. https://doi.org/10.1073/pnas.2216531120

Pais, R. J. (2022). Predictive Modelling in Clinical Bioinformatics: Key Concepts for Startups. *Biotech (Basel (Switzerland))*, *11*(3), 35. https://doi.org/10.3390/biotech11030035

Shimohama, S. (2000). Apoptosis in Alzheimer's disease--an update. *Apoptosis : an international*

journal on programmed cell death, *5*(1), 9–16. https://doi.org/10.1023/a:1009625323388

Stadelmann, C., Deckwerth, T. L., Srinivasan, A., Bancher, C., Brück, W., Jellinger, K., & Lassmann, H. (1999). Activation of caspase-3 in single neurons and autophagic granules of granulovacuolar degeneration in Alzheimer's disease. Evidence for apoptotic cell death. *The American journal of pathology*, *155*(5), 1459–1466. https://doi.org/10.1016/S0002-9440(10)65460-0

Waite, L. M. (2015). Treatment for Alzheimer's disease: has anything changed?. *Australian prescriber, 38*(2), 60–63. https://doi.org/10.18773/austprescr.2015.018