The Impact of Lifestyle Interventions

On Delaying Alzheimer's Onset

A THESIS

Presented to

The Faculty of the Department of Economics and Business The

Colorado College

By

Mahnoor Rehman

December 2024

The Impact of Lifestyle Interventions

On Delaying Alzheimer's Onset

Mahnoor Rehman

December 2024

Mathematical Economics

Abstract

Alzheimer's Disease (AD) is a progressive neurodegenerative disorder that poses significant social, economic, and health burdens on individuals and the healthcare system. This thesis explores the role of lifestyle interventions, specifically diet, in delaying the onset of AD and reducing its socioeconomic toll. Using the Cox Proportional Hazards Model, the analysis integrates data from two distinct sources to examine the relationship between health indicators and the risk of AD onset.

ACKNOWLEDGEMENTS

The completion of my thesis would not have been possible without the time and support of many individuals. I am deeply grateful to Professor Esther Redmount, my thesis advisor, for her invaluable guidance and encouragement throughout this journey. Her inspiration and assistance have been instrumental in helping me explore such an important yet challenging research topic.

ON MY HONOR, I HAVE NEITHER GIVEN NOR

RECEIVED UNAUTHORIZED AID ON THIS THESIS

Mahnoor Rehman

Signature

TABLE OF CONTENTS

Introduction

Alzheimer's Disease (AD) is a form of dementia that has an impact on a person's cognitive abilities of thinking and memory.[1] AD is a progressive condition that affects multiple regions of the brain, with its impact worsening over time. Alzheimer's disease caused upto 120,000 deaths in America in 2022, with a mortality rate of 28.9 deaths per 100,000 people, according to Centers for Disease Control and Prevention.[2]Alzheimer's onset is of two types: Early-onset and Late onset. For Early-onset, the symptoms start developing as early as 30's or before the age of 65 and Late-onset AD symptoms start appearing in mid-60s or later.[3] In the early stages of Alzhiemers, patients experience mild loss of memory, poor judgment and struggle to maintain conversations.[4] However, at later stages Alzheimer's leads towards complete memory loss, disorientation, inability to recognize time and place and suspicion about family, friends and professional caretakers. In its most severe case, patients also experience difficulty speaking, swallowing and walking.[5]Most people require hospice care at this stage.

Alzhiemer's is one of the costliest diseases of our time with high levels of direct and indirect costs. Healthy changes in lifestyle can play a role in reducing the risk of AD to a certain extent.

---

[1] "What Is Alzheimer's?" *Alzheimer's Disease and Dementia*, www.alz.org/alzheimers-dementia/what-is-alzheimers.
[2] *Alzheimer's Death Rates Are Highest in These 10 States | Healthiest Communities Health News | U.S. News*, www.usnews.com/news/healthiest-communities/slideshows/alzheimers-death-rates-are-highest-in-these-10-states.
[3] *What Are the Signs of Alzheimer's Disease? | National Institute on Aging*, www.nia.nih.gov/health/alzheimers-symptoms-and-diagnosis/what-are-signs-alzheimers-disease.
[4] "10 Early Signs and Symptoms of Alzheimer's and Dementia." *Alzheimer's Disease and Dementia*, www.alz.org/alzheimers-dementia/10_signs.
[5] *What Are the Signs of Alzheimer's Disease? | National Institute on Aging*, www.nia.nih.gov/health/alzheimers-symptoms-and-diagnosis/what-are-signs-alzheimers-disease.

These changes can be encouraged through extensive public healthcare interventions by the government, NGOs and healthcare campaigns. This thesis will focus on how lifestyle changes, especially those in diet and physical activity can impact the onset of AD.

Costs Associated with Alzheimer's Disease

Alzheimer's is a major contributor to healthcare costs, especially in populations above the age of 65. Alzheimer's is the most expensive disease in the United States following diabetes and heart Strokes. By 2060, the number of Americans with AD is projected to rise to nearly 13.8 million according to the Alzheimer's Association.[6] As of 2023, 6.7 million Americans age 65 and older have AD.[7]

Total payments in 2023 for health care, long-term care and hospice services for people aged above 65 with dementia are estimated to be $345 billion.[8] AD is projected to account for in excess of $350 billion or 8 percent of total US health care spending. Unlike other medical care, it is important to note the difference between direct and indirects costs for AD care. Direct costs refers to economic expenditure on patients in terms of medical and nonmedical care with the measurement unit of healthcare re- sources as well as social care services. The indirect costs refer to unpaid AD caregivers. According to the Alzheimer Association, in 2023, more than 18 billion people provided unpaid care to patients at the economic value of $350 Billion.[9]

[6] *2023 Alzheimer's Disease Facts and Figures - 2023 - Alzheimer's & Dementia - Wiley Online Library*, alz-journals.onlinelibrary.wiley.com/doi/10.1002/alz.13016.
[7] *2023 Alzheimer's Disease Facts and Figures - 2023 - Alzheimer's & Dementia - Wiley Online Library*,
[8] *2023 Alzheimer's Disease Facts and Figures - 2023 - Alzheimer's & Dementia - Wiley Online Library*.
[9] *2023 Alzheimer's Disease Facts and Figures - 2023 - Alzheimer's & Dementia - Wiley Online Library*,

AD does not have a cure so far but there are three available treatments so far: aducanumab, donanemab, and lecanemab. These drugs are impactful in the middle stage of the disease and are known to reduce some cognitive and behavioral symptoms.[10] Between 1995 and 2021, all of the private expenditures on clinical stage AD Research and Development were estimated at $42.5 billion.[11] The drug development process for AD is extremely challenging with a high failure rate. Due to the rising trend of AD in the aging population in addition to the high failure rates of AD drug development, the costs of medications are relatively higher for AD.

At the same, the US does not have sufficient workforce to support the number of AD patients in the country. 55 % of primary care physicians (PCPs) caregivers for people living with Alzheimer's report there are not enough dementia care specialists in their communities. In order to catch up with the growing need of AD medical care, the number of geriatricians will need to be almost tripled to effectively care for the number of people projected to have AD by 2050. Furthermore in the next five years, more than a million additional direct care workers will be needed to care for the growing AD population.[12]

It is estimated that the health care costs for AD patients will reach $360 billion without unpaid labor value of caregivers. People living with Alzheimer's or other dementias have twice as many hospital stays as well as a nurse facility stays per year as other older people. Medicare is

[10] *How Is Alzheimer's Disease Treated? | National Institute on Aging*, www.nia.nih.gov/health/alzheimers-treatment/how-alzheimers-disease-treated.
[11] Cummings, Jeffrey L, et al. "The Costs of Developing Treatments for Alzheimer's Disease: A Retrospective Exploration." *Alzheimer's & Dementia : The Journal of the Alzheimer's Association*, U.S. National Library of Medicine, Mar. 2022.
[12] *2023 Alzheimer's Disease Facts and Figures - 2023 - Alzheimer's & Dementia - Wiley Online Library*, alz-journals.onlinelibrary.wiley.com/doi/10.1002/alz.13016.

expected to cover around 64% of the total however out of pocket is expected to be $91 billion. These numbers are to increase to nearly $1 trillion in 2050.[13]

Lifestyle Interventions

Although AD does not have a cure so far, there are prevention tactics that can be implemented to delay the onset and progression in the populations. There are multiple risk factors including unhealthy diet, lack of exercise and physical activity, smoking, diabetes, and emotional stress, which can trigger the onset of Alzheimer's disease. However, adapting a lifestyle that avoids these risk factors is helpful in delaying AD symptoms. There is scientific evidence that a healthy diet which includes a substantial amount of vegetables reduces the risk of Alzheimers by 38%. Similarly, Omega-3 fatty acids are known to reduce the risk by 60%.[14] The Mediterranean diet is considered to be one of the most effective diet plans to prevent dementia.

There is extensive research on lifestyle approaches to target the AD pathological pathways, including glucose hypometabolism, inflammation and oxidative stress. The Finnish Geriatric Intervention Study to Prevent Cognitive Impairment and Disability (FINGER) conducted a controlled trial that showed a multi-domain approach including nutrition, physical activity, and cognitive training reduces the risk of dementia in the aging population.[15] Another study published in the journal Alzheimer's Research & Therapy, conducted a trial on a total of 51 adults aged 45 – 90 years. Here as well, the multi-domain intervention of diet, exercise, cognitive

[13] *2023 Alzheimer's Disease Facts and Figures - 2023 - Alzheimer's & Dementia - Wiley Online Library*, alz-journals.onlinelibrary.wiley.com/doi/10.1002/alz.13016.
[14] Stefaniak, Oliwia, et al. "Diet in the Prevention of Alzheimer's Disease: Current Knowledge and Future Research Requirements." *Nutrients*, U.S. National Library of Medicine, 30 Oct. 2022, www.ncbi.nlm.nih.gov/pmc/articles/PMC9656789/.
[15]Stefaniak, Oliwia, et al. "Diet in the Prevention of Alzheimer's Disease: Current Knowledge and Future Research Requirements." *Nutrients*, U.S.

training, and vascular risk monitoring showed significant improvement in cognition and function for the lifestyle intervention group as compared to the control group in the span of 20 weeks.[16]

Another risk factor for AD onset is smoking tobacco which is known to increase the risk of both vascular and Alzheimer dementia. According to the 2014 World Alzheimer Report, there were four studies conducted out of which two showed a positive correlation between cigarettes smoked and dementia risk.[17] According to the World Health Organization, 14% of AD cases worldwide are linked with smoking.[18] Exercise is another significant factor that impacts the onset and progression of dementia. As compared to people who do not exercise on a regular basis, people who do are 20% less likely to experience dementia.[19] Most of the studies explored aerobic exercises among middle age and elderly people to see the impact on the population that is most vulnerable to risks. Living a healthy lifestyle is one of those tactics. This includes avoiding smoking and drinking, physical activity, healthy diet, sleep duration. Etc.

Delaying onset of AD has socio economic benefits due to lowering healthcare costs. A 1-year delay reduces formal costs in 2030 by $70 billion and informal costs by $43 billion, which are a 17% reduction in formal costs and a 20% reduction in informal costs.[20] Hence, we see that there is a need for AD onset delay in the United States and public interventions through governments

---

[16] Dr. Sanchari Sinha Dutta, Ph.D. “Can Lifestyles Changes Reduce Alzheimer’s Risk?” *News*, 11 June 2024,
[17]*World Alzheimer Report 2014 - Dementia and Risk*, www.alzint.org/u/WorldAlzheimerReport2014.pdf. Accessed 14 Dec. 2024.
[18] *Tobacco & Dementia*, iris.who.int/bitstream/handle/10665/128041/WHO_NMH_PND_CIC_TKS_14.1_eng.pdf.
[19] “Physical Activity and the Risk of Dementia.” *Alzheimer’s Society*, www.alzheimers.org.uk.
[20]Zissimopoulos, Julie, et al. “The Value of Delaying Alzheimer’s Disease Onset.” *Forum for Health Economics & Policy*, U.S. National Library of Medicine, Nov. 2014, www.ncbi.nlm.nih.gov/pmc/articles/PMC4851168/.

and organizations can be helpful in doing so. For the rest of the thesis, I will be doing an exploration of lifestyle patterns including diet and exercise to see how it might affect the onset of AD and potentially delay it in the elder population

Literature Review

For a long time, researchers have been investigating Alzheimer's Disease (AD), especially regarding the growing economic burden of the disease. Due to the rapidly increasing rate of AD among the population, there is an urgent need to accelerate research and development methods for sustainable and cost effective treatments as well as strategies for prevention and delaying disease progression. Diagnosis timing plays a crucial role in estimating the costs paid throughout the diagnosis and treatment of dementia. There has been extensive research on the effects of AD diagnosis timing and how various lifestyle interventions including medications and treatments might affect that.[21] A constant debate is prevalent within the clinical and public health communities to find out the appropriate time to begin assessing and supporting patients who are showing early symptoms or might have the risk of Dementia.[22] The development of disease modifying therapies (DMTs) are also found to be more useful for patients who are diagnosed at earlier rather than a later stage.[23] In a modeling framework that assessed cost effectiveness of early diagnosis, it was found that when modeling a treatment for AD, the benefits are substantially increased when there is extended time added for the patient in the predementia state.[24]

[21] Barnett, Jennifer H et al. "Early intervention in Alzheimer's disease: a health economic study of the effects of diagnostic timing." *BMC neurology* vol. 14 101. 7 May. 2014.
[22] Barnett, Jennifer H et al. "Early intervention in Alzheimer's disease"
[23] Barnett, Jennifer H et al. "Early intervention in Alzheimer's disease"
[24] Green, Colin, et al. "Assessing Cost-Effectiveness of Early Intervention in Alzheimer's Disease: An Open-Source Modeling Framework." *Alzheimer's & Dementia*, vol. 15, no. 10, 2019, pp. 1309-1321.

Similar research has been conducted to identify the most optimal and cost-effective treatments for Alzheimer's Disease (AD) patients. In a systematic review of articles published between 1st January 2007 and 15th July 2010, dealing with the pharmacoeconomic factors associated with AD medications, treatments with Cholinesterase inhibitors (CEIs) or memantine were found to be most cost-effective for AD patients. [25] There is pressure to improve medical management for AD medicare due to the immeasurable socioeconomic burden of the disease on caregivers that are usually unpaid family members. One way to move forward is through disease management programs that involve multidisciplinary teams of practitioners, in cases of diabetes and heart diseases.[26] Such programmes can be equipped to identify and treat AD patients which will improve the overall medical and economic results for the future. [27]

Additionally, sizable research is being carried out to understand the impact of lifestyle interventions such as education, cognitive training, psychological therapy, physical exercise, healthy diet on delaying and/ or preventing depression. Foods that have high levels of antioxidant and anti-inflammatory properties, are known to regulate the immune system and positively affect the delay in progression of cognitive impairment and AD.[28] These dietary habits may assist in the prevention of cognitive decline which can further prevent and delay the progression of AD symptoms.[29]

---

[25] Fillit, H M. "The pharmacoeconomics of Alzheimer's disease." *The American journal of managed care* vol. 6,22 Suppl (2000): S1139-44; discussion S1145-8.2 (2012).

[26] Pouryamout, L., Dams, J., Wasem, J. *et al.* "Economic Evaluation of Treatment Options in Patients with Alzheimer's Disease." *Drugs* **72**, 789–802 (2012).

[27] Pouryamout, L. *et al.* "Economic Evaluation of Treatment Options in Patients with Alzheimer's Disease."

[28] McGrattan, Andrea M et al. "Diet and Inflammation in Cognitive Ageing and Alzheimer's Disease." *Current nutrition reports* vol. 8,2 (2019).

[29] Canevelli, Marco et al. "Sundowning in Dementia: Clinical Relevance, Pathophysiological Determinants, and Therapeutic Approaches." *Frontiers in medicine vol.* 3 73. 27 Dec. 2016.

Physical exercise is also another important factor to consider when treating or preventing different types of dementia. In the 1970s, a study was published that showed middle-aged sports practitioners respond better in cognitive tasks compared to sedentary aged-matched subjects.[30] Another study found that 1 year of moderate-intensity exercise increased the size of the hippocampus, a region of the brain associated with memory, in healthy older individuals.[31] There have also been studies conducted on the benefits of strength training among the middle aged population to delay the symptoms of AD. Although research on the positive impacts of physical activity, including resistance training, remains limited and requires further exploration, there is still significant evidence to support the long-term benefits of physical activity in reducing the risk of dementia, particularly for individuals in presymptomatic and predementia stages.[32]

Cognitive Behavioral Therapy (CBT) has been quite effective in addressing mental health conditions like depression and anxiety. CBT is a form of talking therapy aimed to understand the interconnections between thoughts, feelings and behaviors to make positive life changes.[33] Dr Josh Stott from University College London is currently researching CBT as a way to help people in their early stages of dementia, who are also experiencing anxiety and depression.[34] Furthermore, A Cochrane reviewed data from 33 cognitive therapy studies which showed that people completing cognitive therapy may show some benefits in overall cognition including

[30] Spirduso, W W, and P Clifford. "Replication of age and physical activity effects on reaction and movement time." *Journal of gerontology* vol. 33,1 (1978): 26-30.

[31] Erickson, Kirk I et al. "Exercise training increases size of hippocampus and improves memory." *Proceedings of the National Academy of Sciences of the United States of America* vol. 108,7 (2011): 3017-22.

[32] De la Rosa, Adrian et al. "Physical exercise in the prevention and treatment of Alzheimer's disease." *Journal of sport and health science* vol. 9,5 (2020): 394-404.

[33] "Cognitive Behavioural Therapy (CBT) for Dementia." Alzheimer's Society.

[34] "Cognitive Behavioural Therapy (CBT) for Dementia." Alzheimer's Society.

cognitive abilities such as verbal fluency which is something that AD patients struggle with in the middle and late stages. [35]

Survival Analysis

Survival analysis is an analytical method that focuses on time-to-event data.[36] When it comes to health economic evaluation, models are usually dependent on estimates of how much time it takes to achieve specific outcomes. These outcomes might be the time of first diagnosis, disease progression, or time to death.[37] Survival analysis is valuable for analyzing the distribution of survival times, such as life expectancy in cases of chronic diseases like Cancer, Alzheimer's, etc. The methods of survival analysis are widely used as they also help to find the estimates of mean survival benefit for individuals.[38] There are two key features of survival analysis, known as survival and hazard function. The survival function *s(t)* represents the probability that an individual survives beyond a certain time t.[39] It is significant because it provides probability of survival for different values of *t*. The hazard function $h(t)$ represents the instant risk of an event such as death or disease progression occurring at a specific time, given that the individual has survived up to that point.[40] Essentially for the survivor model, the hazard function $h(t)$ represents the occurrence rate of the event, while the survival function *s(t)* represents the non occurrence

[35] Bahar-Fuchs, Alex et al. "Cognitive Training for People with Mild to Moderate Dementia: A Cochrane Review." *BJPsych Advances* 26.2 (2020): 66–66.
[36] "Survival Analysis." *YHEC*, 28 Jan. 2021, https://yhec.co.uk/glossary/survival-analysis/.
[37] Authors, et al. Extrapolating Clinical Evidence Within Economic Evaluations: CADTH Methods and Guidelines. Canadian Agency for Drugs and Technologies in Health, May 2023.
[38] Authors, et al."Extrapolating Clinical Evidence Within Economic Evaluation"
[39] Clark, T G et al. "Survival analysis part I: basic concepts and first analyses." *British journal of cancer* vol. 89,2 (2003): 232-8. doi:10.1038/sj.bjc.6601118
[40] Clark, T G et al. "Survival analysis part 1"

rate of the event.[41] The survivor data is rarely normally distributed. Usually, it is skewed and consists of a larger sequence of earlier events as compared to later events.

A survival analysis helps to assess the likelihood of a patient surviving for a specific amount of time given certain interventions, we can employ it as a model to estimate the benefits of delaying onset of Alzheimer's through public health interventions. The analysis can be utilized to estimate the time until the onset of AD, which would be the "event" for the model. Groups of individuals receiving public health interventions such as diet, exercise, and cognitive therapy compared to the control group who are not receiving interventions.

The Cox Proportional Hazards Model, also known as the Cox regression model, is one of the most widely used statistical methods for analyzing survival data. It allows for the estimation of the effect of several variables on the time of a particular event. The hazard function describes the instantaneous rate at which the event occurs, provided that it has not yet occurred. Essentially, the Cox model will estimate how covariates such as the lifestyle interventions in this case can impact the risk of the onset of AD. It is widely used in medical, biological, and social sciences to analyze time-to-event data, such as the time until death, disease onset, or failure of a machine component. The model assumes that the effect of the predictor variables on the hazard will stay constant over time. Unlike other survival models, the Cox model does not require a specific baseline hazard function.

[41] Clark, T G et al. "Survival analysis part 1"

Data Analysis

Initially, I planned on utilizing data from all three factors of diet, exercise and cognitive training, however, I struggled to find reliable and relevant data for this thesis. Most of the data was private due to the high confidentiality status. I found a study that examines the health status of 226 elderly individuals using dietary, physical, and lifestyle data collected in 2000.

[42]In the dataset, key demographic variables included gender, age at the time of the interview, and living situation, etc. Dietary habits were evaluated through the frequency at which various food items, including tea, coffee, meat, fish, raw fruits, cooked vegetables, and chocolate, were consumed. This provided a detailed view of individual nutritional patterns. Physical indicators such as height and weight were used to calculate the Body Mass Index (BMI), which is a widely used measure of overall health conditions. Additionally, the dataset included information on cooking practices, specifically the type of fat used in meal preparation, ranging from butter and olive oil to less healthy options like goose fat.

I applied a classification framework to determine whether subjects could be categorized as "Healthy" or "Not Healthy" based on specific criteria. Later I had to adjust the criteria to make it more user friendly and balanced for further analysis. Health status was determined by looking at a combination of dietary, physical, and lifestyle factors. Dietary health was assessed based on the regular consumption of raw fruits, cooked vegetables, and fish. Individuals consuming these foods at least four times per week were considered to have a healthy diet. Physical health was evaluated using the Body Mass Index (BMI), with individuals falling within the normal range of 18.5 to 24.9 considered physically healthy. Lifestyle activity was measured by the regular intake of tea or coffee, and people who were consuming the drinks at least three times were seen as more active and thus healthier. Individuals who met all three criteria were classified as "Healthy." Those who did not meet one or more of these criteria were categorized as "Not

[42] *Pierre Lafaye de Micheaux*, www.biostatisticien.eu/springeR/jeuxDonnees4.html. Accessed 3 Dec. 2024.

Healthy." This classification was based on stringent health criteria, hence it identified only one individual as "Healthy," while the remaining 225 individuals were classified as "Not Healthy."

To achieve a more manageable and balanced distribution of the data, I adjusted the health criteria to be less stringent. The thresholds for fruit, vegetable, and fish consumption were lowered. Individuals who were consuming these items at least three times per week were now classified as healthy. Similarly, changes were made to the BMI threshold. The acceptable range for Body Mass Index (BMI) was also made larger to include values between 18.0 and 27.0. Furthermore, the activity criteria was changed so that the individuals who consumed tea or coffee occasionally were also considered, with a minimum threshold of at least two times per week. These adjustments helped to reflect better on the diversity of health habits and physical indicators within the population and led to a more balanced categorization of health status. After implying these changes in the thresholds for dietary, physical, and lifestyle factors, 41 individuals were categorized as "Healthy," which represents almost 18% of the population. The remaining 185 individuals were still classified as "Not Healthy." This adjustment provided a more balanced depiction of health within the population.

The second data source utilized in this study is the Alzheimer's Disease and Healthy Aging Data, from the Centers for Disease Control and Prevention (CDC).[43] Most of the information in this dataset comes from the Behavioral Risk Factor Surveillance System (BRFSS), collected between 2015 and 2022. To simplify the analysis and address challenges with data confinement, I focused specifically on data of Alzheimer's patients for the year 2022. From this subset, I randomized 80 values to create a sample that can be easily managed for data analysis. The table below shows the descriptive statistics from the data:

[43] "Alzheimer's Disease and Healthy Aging Data." *Centers for Disease Control and Prevention*, Centers for Disease Control and Prevention, chronicdata.cdc.gov/Healthy-Aging/Alzheimer-s-Disease-and-Healthy-Aging-Data/hfr9-rurv/about_data. Accessed 3 Dec. 2024.

Variable	Mean	Standard Deviation	Minimum	Median	Maximum
Age	74.48	6.01	65	74	91
BMI	24.64	3.53	16.33	24.62	36.98
Tea Consumption	0.71	1.45	0	0	10
Coffee Consumption	1.62	1.25	0	2	5
Fish Consumption	2.62	0.92	0	3	5
Raw Fruit	4.49	1.09	0	5	5
Cooked Vegetables	4.38	1.01	0	5	5

Table 1: Descriptive Statistics of Key Variables

To establish a connection between the two datasets, an analytical framework using the Cox Proportional Hazards Model was used. This model enables the integration of health status indicators from the first dataset with Alzheimer's onset data from the CDC dataset. The hazard function in the Cox Proportional Hazards Model is defined as

$$h(t|X)=h0(t)\cdot\exp(\beta1\cdot\text{Health Status}+\beta2\cdot\text{Age}+\beta3\cdot\text{Dietary Factors}+\beta4\cdot\text{BMI}+\beta5\cdot\text{Lifestyle Factors})$$

In this equation, $h(t|X)$ represents the risk of Alzheimer's onset at time t, dependent on a set of covariates (X). The baseline hazard, h0(t), shows the risk for individuals with baseline characteristics, such as those without significant health interventions or average health indicators. The covariates include health status, age, dietary factors, BMI, and lifestyle habits. The main aim of these variables is to build a comprehensive model for analyzing health outcomes.

The health status variable is derived from the first dataset as a critical link between the two datasets. It is a binary indicator that classifies individuals as "Healthy" or "Not Healthy" based on dietary, physical, and lifestyle criteria. This variable directly modifies the baseline hazard for Alzheimer's onset, with healthier

individuals expected to have a lower risk. Age, a shared variable across both datasets, functions as the time-to-event variable in the CDC dataset, representing the time until Alzheimer's onset. These variables are integrated as covariates in the Cox model to estimate their impact on the hazard of Alzheimer's onset. For instance, increased consumption of fruits and vegetables or a BMI within the normal range may reduce the hazard of Alzheimer's onset, hence linking dietary patterns to long-term cognitive health.

My first step in the analysis was to preprocess the Lifestyle dataset and prepare it for survival analysis. The covariates included dietary factors, physical factors, and lifestyle factors). I ran my code on python with the lifelines library. I used Cox model to assess how dietary, physical, and lifestyle factors influence the likelihood of being classified as "Healthy." For this analysis, I treated health status as the binary event variable.

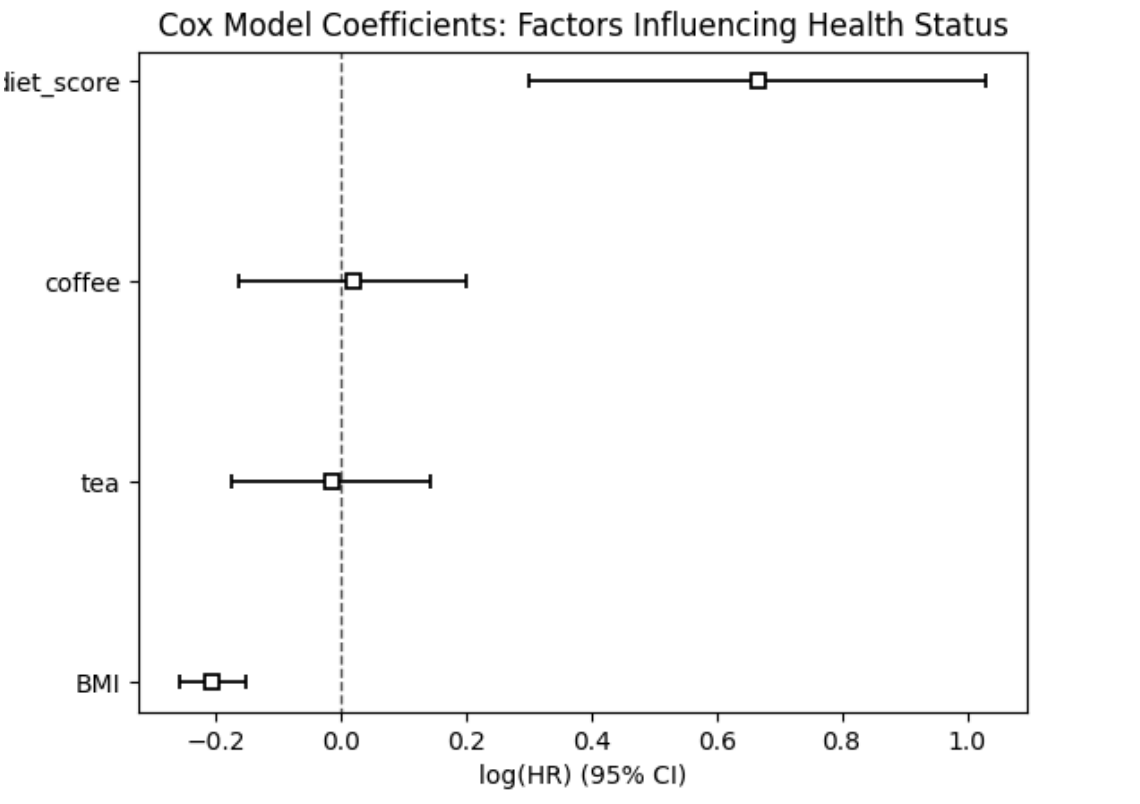

Fig 1 Cox Model

The hazard ratios (HR) and their logarithmic values (log(HR) provide insights into the impact of different covariates such as (diet_score, coffee, tea, and BMI) on the likelihood of being classified as "Healthy."or “Not Healthy”. The error bars represent the 95% confidence intervals (CI), which indicate the statistical reliability of the results. If a CI includes zero, the variable's effect is considered statistically insignificant.

Diet_score showed a significantly positive log HR at 0.669, indicating that higher dietary scores, due to consumption of healthy foods such as fruits, vegetables, and fish, increases the chances of being classified as "Healthy." The confidence interval for diet_score does not include zero which confirms its statistical significance and focuses on its positive impact. Coffee consumption also has a positive log HR at 0.019, indicating a higher chance of being classified as "Healthy." However, its effect is less as compared to diet_score. This finding remains statistically significant, as its confidence interval does not cross zero. This shows that coffee could have health benefits, however its impact is not as big as diet. On the other hand, tea consumption has a log HR close to zero at -0015, implying a negligible effect on being healthy. This variable is not statistically significant as its confidence interval includes zero. Similarly, BMI shows a slightly negative log HR at -0.204, suggesting that individuals with BMI values outside the healthy range are less likely to be classified as "Healthy." However, its confidence interval also includes zero, making BMI’s impact statistically insignificant in this analysis.

In the next step, I integrated the health status and age variables into the CDC dataset to analyze Alzheimer’s onset. My aim was to estimate the impact of these factors on the hazard of Alzheimer's Onset

using the Cox model.

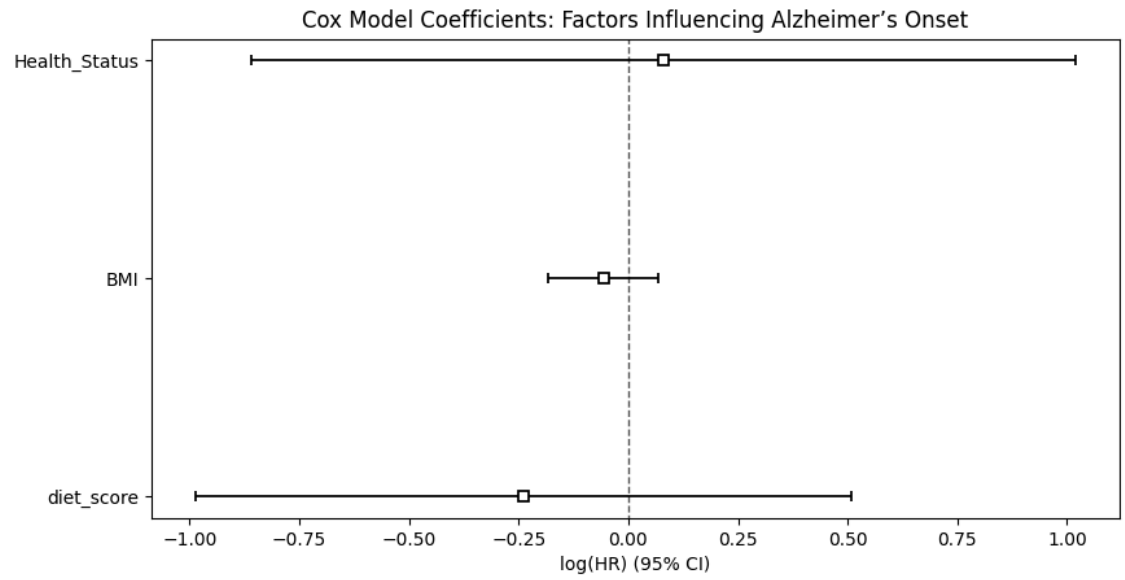

Fig 2 Cox Model

Health_Status emerges as a significant protective factor in this model. Its log HR, at -0.270, highlights that individuals classified as "Healthy" have less risk of Alzheimer's onset compared to those not classified as healthy. The confidence interval extends far below zero, which shows the significance of this effect.. Similarly, Diet_Score demonstrates a negative coefficient, also close to -1, suggesting that better dietary habits reduce risk of Alzheimer's onset. While the confidence interval is relatively wide, it does not include zero, indicating statistical significance. This finding reinforces the idea that healthy dietary patterns protect elderly people against cognitive decline.

On the other hand, BMI shows a coefficient at -0.0231 with a narrow confidence interval. The Z-score for the BMI variable highlights its significant difference from the null hypothesis, despite the small hazard ratio. This shows that while the effect size of BMI on Alzheimer's onset is moderate, it still remains statistically important. The narrow confidence interval around the coefficient also sheds light on its relevance. Although the confidence interval includes zero, the statistical significance of the Z-score

indicates that BMI cannot be disregarded as an insignificant factor. The negative association suggests that having a BMI score closer to the "healthy" range (18.0 – 27.0) contributes to lowering the risk of Alzheimer's onset, even if its impact is not as pronounced as other variables such as diet score. The topic requires more research and extensive data values to further investigate the impact of BMI score.

The strong protective effects of Health_Status and Diet_Score suggest insights for public health initiatives. Promoting overall well-being and encouraging healthier dietary patterns could reduce Alzheimer's risk. However, the lack of significance for BMI raises questions and highlights the need for additional research to explore this inconsistency.

Limitations

The analysis faces many different limitations related to data availability and accessibility. Much of the relevant data on Alzheimer's Disease (AD) and its associated risk factors, such as dietary habits, physical activity, and cognitive training, is present in private datasets or bound by confidentiality agreements, so only a professional medical researcher can get access to them. Hence this study was restricted only to publicly available datasets and previous research, which may not represent the full extent of variables influencing AD onset. For example, the Bordeaux dataset, used to classify health status, relied on a relatively small sample size and self-reported metrics, which can introduce biases such as inaccuracies in reporting. The adjustments that were made to the dataset in the health classification criteria to achieve a balanced distribution may have diluted the precision of the findings due to less strict thresholds.

Another limitation was in the integration of the two datasets. While the Cox Proportional Hazards Model allowed for a meaningful analysis of the combined data, inconsistencies in variables such as timeframes and health indicators created challenges. The CDC data used for

Alzheimer's onset analysis was also randomized that might have affected its statistical power and external validity. Furthermore, the difference in the effects of tea and coffee seems to be the result of inconsistency and randomization, which can probably be fixed with a more accurate and larger data set.

Conclusion

This thesis highlights the role of lifestyle interventions in delaying the onset of Alzheimer's Disease (AD), further offering insight into healthier habits that can mitigate the socioeconomic impact of AD. I used the Cox Proportional Hazards Model, which confirms the effects of dietary health and overall well-being on reducing the risk of AD onset. Dietary factors such as regular consumption of fruits, vegetables, and fish emerged as significant contributors to lowering the risk of AD, which encourages balanced nutritional habits in at-risk populations.

While dietary patterns and overall health status showed strong protective effects, the absence of a significant correlation for BMI highlights the complexity of AD's multifactorial nature. This finding can use further investigation into the biological, lifestyle, and environmental factors that influence AD progression. Expanding the scope of research to include larger and more diverse datasets could help in clarifying the correlations more which would further help in creating sustainable health strategies. .

Future research should focus on the limitations of data accessibility and can look at how lifestyle changes engage with genetic predispositions, socio-economic factors, and existing health conditions. Moreover, utilizing t advanced statistical models and longitudinal datasets could provide more accurate predictions and valuable insights.

In conclusion, this study looks at lifestyle interventions in reducing the growing burden of Alzheimer's Disease on patients. By using preventative measures such as a healthy diet, society can not only enhance the quality of life for millions of individuals but also assist the healthcare systems in the current state of economy. With efforts and multidisciplinary collaboration, we can imagine a future where Alzheimer's Disease is no longer an inevitability for aging populations but a manageable and preventable disease.

Appendix

covariate	coef	exp(coef)	se(coef)	coef lower 95%	coef upper 95%	exp(coef) lower 95%	exp(coef) upper 95%	cmp to	z	p	-log2(p)
diet_score	0.665910	1.946261	0.186407	0.300559	1.031262	1.350614	2.804602	0.0	3.572343	3.538012e-04	11.464774
BMI	-0.204297	0.815220	0.027635	-0.258461	-0.150134	0.772239	0.860593	0.0	-7.392714	1.438619e-13	42.660380
tea	-0.015413	0.984705	0.080991	-0.174154	0.143327	0.840168	1.154107	0.0	-0.190310	8.490662e-01	0.236051
coffee	0.019069	1.019252	0.092597	-0.162416	0.200555	0.850087	1.222081	0.0	0.205941	8.368373e-01	0.256981

Table 2: Cox Proportional Hazard Model for Predicting “Healthy” Status of Patients

covariate	coef	exp(coef)	se(coef)	coef lower 95%	...	cmp to	z	p	-log2(p)
					...				
Health_Status	-0.270452	0.763035	0.535798	-1.320596	...	0.0	-0.504765	0.613724	0.704338
BMI	-0.023164	0.977102	0.066640	-0.153775	...	0.0	-0.347603	0.728138	0.457716
diet_score	-0.378708	0.684745	0.396510	-1.155855	...	0.0	-0.955103	0.339526	1.558408

[3 rows x 11 columns]

Table 3: Cox Proportional Hazard Model for Predicting Alzheimer’s Onset