IMPROVING PREDICTIVE OUTCOMES
OF SOCIAL DETERMINANTS OF HEALTH (SDOH) PARAMETERS
WITH MACHINE LEARNING TECHNIQUES
IN HEALTHCARE ECOSYSTEM

by

Kapil Mohan, Executive MBA, PG Diploma, PMP, B. Tech

DISSERTATION
Presented to the Swiss School of Business and Management Geneva
In Partial Fulfilment
Of the Requirements
For the Degree

DOCTOR OF BUSINESS ADMINISTRATION

SWISS SCHOOL OF BUSINESS AND MANAGEMENT GENEVA

January, 2024

IMPROVING PREDICTIVE OUTCOMES
OF SOCIAL DETERMINANTS OF HEALTH (SDOH) PARAMETERS
WITH MACHINE LEARNING TECHNIQUES
IN HEALTHCARE ECOSYSTEM

by

Kapil Mohan

APPROVED BY

____
<Chair’s Name, Degree>, Chair

____
<Member’s Name, Degree>, Committee Member

____
<Member’s Name, Degree>, Committee Member

RECEIVED/APPROVED BY:

____
SSBM Representative

Dedication

This dissertation is dedicated to the pillars of my life, whose unwavering support and inspiration have made this journey possible.

To my parents, the bedrock of my existence, whose love and guidance have shaped the person I am today. Your sacrifices and unwavering belief in my potential have been the guiding light through every challenge.

To my wife and two daughters, the joys of my heart, especially to my younger daughter Suhani, whose inquisitive nature and resilience have inspired me more than words can express. Your laughter and boundless curiosity have been a constant reminder of the wonders that lie in exploring the unknown.

To my friends and colleagues, who have stood by me through thick and thin, providing laughter, solace, and invaluable perspective. Your camaraderie and support have been a source of strength and encouragement.

To my mentors, whose wisdom and insights have profoundly shaped my academic journey. Your guidance has not only enlightened my path in this research but has also instilled in me a passion for continuous learning and exploration.

This work is also a tribute to all who have walked with me on this path, directly or indirectly contributing to my growth and success. Your roles in my life's journey are deeply appreciated and forever cherished.

Acknowledgements

As I culminate this challenging yet rewarding journey of my Global Doctor of Business Administration, I find myself reflecting on the invaluable support and guidance that I have received. This accomplishment is not just a reflection of my efforts, but a testament to the encouragement and wisdom imparted by those around me.

Foremost, I express my deepest gratitude to Dr. Bhawna Nigam, whose mentorship has been a cornerstone of my academic and personal growth throughout this DBA program. Dr. Bhawna Nigam, your expertise in the field and your unwavering commitment to nurturing my potential have been instrumental in shaping my research and guiding me through complex challenges. Your insightful feedback, constructive criticism, and encouragement have been invaluable.

I extend my sincere thanks to SSBM and Upgrad, for offering and facilitating the GDBA program in India. This program has not only provided me with a rigorous academic platform but also a unique opportunity to delve into and contribute to the world of strategic chaos engineering. The resources, support, and learning environment fostered by these institutions have been pivotal in my research journey.

A special word of appreciation goes to the administrative and support staff at both SSBM and Upgrad. Your assistance in navigating the logistics and requirements of the program has allowed me to focus on my research and academic growth.

My journey would not have been the same without the intellectual stimulation and discussions provided by my peers and fellow researchers Saichinta, Kurt, Sriram, Nitin, Dr Tripti, Hardik, Shruti, Roshan, Vikhyat and Amisha. The collaborative environment and the diverse perspectives I encountered have enriched my experience and understanding, for which I am immensely grateful.

Finally, I would like to acknowledge the contributions of all those who have been part of my academic journey in ways big and small. Your support, whether in the form of advice, encouragement, or simply a listening ear, has been a source of strength and motivation.

ABSTRACT
IMPROVING PREDICTIVE OUTCOMES
OF SOCIAL DETERMINANTS OF HEALTH (SDOH) PARAMETERS
WITH MACHINE LEARNING TECHNIQUES
IN HEALTHCARE ECOSYSTEM

Kapil Mohan
2024

Dissertation Chair: <Chair's Name>
Co-Chair: <If applicable. Co-Chair's Name>

This dissertation examines transforming the healthcare system in the United States into a value-based care paradigm, explicitly focusing on the Social Determinants of Health (SDOH) and the obstacles impeding its implementation. The review of existing literature reveals on-going endeavors by the government and healthcare agencies to facilitate this transition by implementing new regulations, educating stakeholders, and centralizing providers. However, a significant trust deficit exists, which can be attributed to the need for more harmonization among stakeholders and the data related to SDOH, thus hindering the smooth adoption of value-based care.

The author emphasizes the significance of collaboration among members, providers, and payers to successfully implement value-based care. The importance of transparent disclosures, appropriate pricing of health plans, and objective claims assessment are identified as essential factors for attaining value-based care objectives. To address these challenges, the dissertation proposes the integration of community information as a means of establishing trust. Although specialized companies focusing on value-based care are emerging, concerns arise regarding patient trust and whether

financial incentives alone are sufficient to transition from a risk-free fee-for-service model.

Moreover, the literature review highlights the necessity of adopting a human-centric design approach to implement SDOH. Current models often fall short in addressing personalized healthcare programs and risk modelling, neglecting to consider the experiences of clinicians and consumers. To address this, the author introduces a novel disease-centric classification of SDOH to enhance the effectiveness of these models in clinical settings and disease management.

In conclusion, the dissertation examines innovative technological initiatives as potential solutions to the challenges of adopting value-based care. Prognostic risk modelling, data fragmentation, consumer activation, empowerment of physicians, and financial modelling tools are identified as key facilitators for adopting value-based care, underscoring the crucial role of technology in improving mechanisms for data capture and providing personalized insights.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER I:

INTRODUCTION

1.1 Introduction

Social determinants are the factors that influence the physical and social environment that a person lives in, the education a person receives and access to healthcare facilities. My study will focus on evaluating the impact of different social determinants from a clinical perspective. The study will collect data from public sources like Agency for Healthcare Research and Quality (AHRQ) databases for social determinants information in a local area and will also use health and demographic data collected by Center for Disease Control and Census Bureau, and is made available for research purposes. The study will also use patient RX orders from a pharmacy in US wherein the personally identifiable information (PII) have been removed.

Organizations like WHO, CDC, Robert Woods Foundation etc have come up with social determinants models. There have been assumptions that a particular social determinants model would have similar impact for every disease and during different stages of disease continuum eg, Air quality would have a same impact for hypertension and for cancer, and also during early stages of a disease (onset) versus mid stages and during advanced stages. The universality assumption needs to be assessed and if these are helping physicians use them in clinical settings which is the ultimate proof of their utility.

Value Based Care is an emerging trend in healthcare delivery in Unites States that was set as a strategic direction in Patient Protection and Affordable Care Act, 2010 to increase healthcare coverage, improve health outcomes and reduce healthcare cost burden on all stakeholders. There have been gradual steps taken towards it by policymakers and industry in terms of new regulations, Medicaid products and Accountable Care business models that focus on health outcomes, better awareness about Whole-person health among providers and general population, and technology enabled interoperability standards like FHIR.

1.2 Social Determinants of Health (SDOH)

SDOH are factors that affect the health of an individual and can be classified into four categories: social and economic conditions, physical environment, social relationships, and personal behaviours. WHO's study shows that social determinants can have a significant impact on how healthy people are, and how different types of determinants interact with each other to affect our well-being.

As per WHO studies, the SDOH have an important influence on health inequities - the unfair and avoidable differences in health status seen within and between countries. In countries at all levels of income, health and illness follow a social gradient: the lower the socioeconomic position, the worse the health. Research shows that the social determinants can be more important than health care or lifestyle choices in influencing health. For example, numerous studies suggest that SDOH account for between 30-55% of health outcomes. In addition, estimates show that the contribution of sectors outside health to population health outcomes exceeds the contribution from the health sector. These studies show that non-medical (environmental) factors can have a significant impact on how healthy people are, and how different types of determinants interact with each other to affect our well-being.

Some provisions of Patient Protection and Affordable Care Act (2010), reflect overwhelming evidence that to reduce healthcare costs, and to improve quality of care and population health, the Social Determinants of Health (SDOH) must be addressed. These policies include funding for partnership between public health agencies, community organizations, healthcare institutions, promotion of value-based payment models that incentivize integrated health and social care delivery, and support for Medicaid program innovations that directly address social needs as part of healthcare.

Improving health outcomes is a complex interplay between health system, community, and individual-level factors are increasingly seen as important to understanding patient health and identifying appropriate interventions. SDOH play a large role in bridging gaps in health inequity, improving population health outcomes and control healthcare costs. There are several efforts from different agencies and departments to categorize, measure and improve social determinants of health, namely,

• World Health Organization (WHO) Model: SDOH domains include Sociodemographic, Psychological, Behavioural, Individual-Level Social Relationships and Living Conditions, Neighbourhoods and Communities.

• PROGRESS+model: SDOH domains Place of residence, Race/ethnicity/culture/language, Occupation, Gender/sex, Religion, Education, Socioeconomic status, Social capital.

• Healthy People 2030 model by Department of Health and Human Services (DHHS): SDOH domains include Economic Stability, Education Access and Quality, Healthcare Access and Quality, Neighbourhood and Built Environment, Social and Community Context.

• Division of Nutrition, Physical Activity and Obesity (DNPAO) model by Centers for Disease Control (CDC): SDOH domains include Built environment, Community-clinical linkages, Food and nutrition security, Social connectedness, Tobacco-free policies.

• Factors Affecting Communities and Enabling Targeted Services (FACETS) model from Cantor et all (2017) using an open architecture: SDOH domains include Total population, Urban/rural classification, Racial diversity, Ethnic diversity

(Hispanic/non-Hispanic), US citizenship, Foreign vs native-born, Educational attainment, English proficiency, Poverty rate, Median household income ACS, Unemployment rate, Health insurance status, Respiratory Hazard Index, Access to healthy food, Distance to parks, Walkability score, Tobacco retailers/1000 population, Felony crime/1000 population, Gini index of inequality, Social Vulnerability Index, Housing violations/1000 units, Voter turnout.

• PLACES model by collaboration between CDC and the Robert Wood Johnson Foundation: SDOH domains include Unhealthy Behaviours, Health Outcomes, Prevention, Social Determinants, Race/Ethnicity.

1.3 Machine Learning Techniques

The use of AI has contributed to the development of SDOH, Eg. use of machine learning algorithms is helping researchers identify patterns in data about social determinants and prescribe precision medicine. However as per Melzer, (2022) and Lewis et al., (2021), there have been funding challenges for initiatives involving identifying and mitigating SDOH factors. This is primarily due to the uncertainty of outcomes and unavailability of enough data across communities organized in a fashion that can be used by scientists and policy makers for financial sustainability.

There have been some notable successes in using healthcare data for technology-based interventions. For example, machine learning algorithms have been used to analyse EHR data to identify patients at high risk of readmission or complications, allowing healthcare providers to intervene proactively. Wearable devices and mobile apps have also been used to collect data on patient behaviour and symptoms, providing valuable insights into how patients are managing their health outside of the clinic.

Improving health outcomes is a complex interplay between health system, community, and individual-level factors are increasingly seen as important to understanding patient health and identifying appropriate interventions. SDOH play a large role in bridging gaps in health inequity, improving population health outcomes and control healthcare costs.

Huang et al (2021) studied discharge disposition among heart failure and acute kidney injury in COVID-19 patients from 2 urban hospitals and proposed a new self-supervised machine earning framework, Deep significance clustering (DICE), that divided patients into clinically similar and risk-stratified subgroups that unsupervised clustering algorithms or supervised risk prediction algorithms could not generate. It also demonstrated the potential to apply DICE in heterogeneous populations, where having the same quantitative risk does not equate with having a similar clinical profile.

On the other side, Kasthurirathne et al (2017) conducted model experiments to evaluate the capacity for clinical, socioeconomic, and public health data sources to predict the need for various social service referrals among patients at a safety-net hospital. They integrated patient clinical data and community-level data representing patients' social determinants of health (SDOH) to build random forest decision models to predict the need for any, mental health, dietitian, social work, or other service referrals. They opined that the need for various social service referrals can be predicted with considerable accuracy using a wide range of readily available clinical and community data that measure socioeconomic and public health conditions. The also opined that use of SDOH did not result in significant performance improvements. Subsequently Jessica S Ancker and colleagues (2018) published a paper acknowledging the work of Kasthurirathne et al and highlighted possible reasons for differing results in their work while it is well established in medical science that population health is affected by socioeconomic status and other social determinants of health (SDOH). They submitted the reasoning for anomalies in previous work such as, a strong correlation of SDOH data existed with clinical data resulting in low impact on predictive power of SDOH (data diversity), insufficient variability in SDOH across patients (data diversity), community-level SDOH that was used could have been biased (data bias), and possible insufficiency of predictor variables (data diversity).

Lans et al (Aug 2022) did a systematic review to investigate whether prognostic ML models for orthopaedic surgery outcomes account for SDOH, and to what extent SDOH variables are included in the final models. Resources scanned were from PubMed, Embase and Cochrane for studies published up to 17 November 2020. Two reviewers independently extracted SDOH features using the PROGRESS+ framework. Across all studies, 96% (57/59) considered at least one PROGRESS+ factor during development. The most common factors were age (95%; 56/59) and gender/sex (96%; 57/59). Differential effect analyses, such as subgroup analysis, covariate adjustment, and baseline comparison, were rarely reported (10%; 6/59). Most models included age (92%; 54/59) and gender/sex (69%; 41/59) as final input variables. However, factors such as insurance status (7%; 4/59), marital status (7%; 4/59) and income (3%; 2/59) were seldom included. They concluded that the current level of reporting and consideration of SDOH during the development of prognostic ML models for orthopaedic outcomes is limited.

1.4 Healthcare Ecosystem

The Centers for Medicare and Medicaid Services (CMS) Health Equity Pillar defines health equity as the ecosystem where everyone can attain their highest level of health regardless of gender, disability, ethnicity, sexual orientation, geography, preferred language, socioeconomic status or factors that affect health outcomes and access to care. According to DNPAO, following factors influence Health Equity.

Social Determinants of Health - Differences in SDOH contribute to persistent chronic disease disparities among racial, ethnic, and socioeconomic groups as well as in different geographies and among people with different physical abilities.

Racism - It is a system of structures, policies, practices, and norms that assigns value and determines opportunity because of the way people look or the colour of their skin. This results in conditions that unfairly give advantages to some and disadvantages to others.

In recent years, the United States has made several efforts to promote health equity through laws and regulations like The Affordable Care Act (ACA): The Patient Protection and Affordable Care Act (2010), The National Partnership for Action to End Health Disparities (NPA), NYC Care program and others.

Inconsistent prevalence of diseases despite Health Equity efforts Despite Health equity efforts, diseases are not equivalently spread across communities. There are several reasons why the prevalence of diseases may vary across different communities. Some possible explanations include:

• Genetics: Certain diseases may be more prevalent in certain ethnic or racial groups due to genetic factors. For example, sickle cell anaemia is more common among African Americans and Hispanics, while cystic fibrosis is more common among Caucasians.

• Environmental factors: Different communities may be exposed to different environmental factors that can impact disease prevalence. For example, communities living near industrial sites may have higher rates of cancer due to exposure to pollutants.

• Access to healthcare: Communities with limited access to healthcare may have higher rates of certain diseases due to a lack of preventative care and early detection. This can be due to various factors like poverty, geographic isolation, and cultural barriers.

• Lifestyle factors: Lifestyle factors such as diet, exercise, and smoking habits can impact disease prevalence. For example, communities with a high prevalence of obesity may have higher rates of diabetes and heart disease.

These factors are interconnected and can influence each other. For example, limited access to healthcare can lead to delayed diagnoses and poor management of chronic conditions, which can worsen health outcomes.

1.5 Research Problem

Some studies have shown that SDOH impact a person’s wellness by upto 80%. The importance of getting a better understanding of the social determinants is increasing due to the fact that the environment where people live is changing, lifestyle changes are making people lives sedentary and thereby healthcare costs are increasing. There have been many models and prediction methods on how such changes affect a person’s health but none of them have seen a mass adoption. There have been data gaps in such models. Most of the social determinants data is either unavailable or not of good quality for usage in prediction models, or not accessible.

Organizations like WHO, CDC, Robert Woods Foundation etc have come up with social determinants models. There have been assumptions that a particular social determinants model would have similar impact for every disease and during different stages of disease continuum eg, Air quality would have a same impact for hypertension and for cancer, and also during early stages of a disease (onset) versus mid stages and during advanced stages. The universality assumption needs to be assessed and if these are helping physicians use them in clinical settings which is the ultimate proof of their utility.

One of the objectives of Value based care is to make Providers more accountable for health outcomes and thereby sharing the risk with Payers. In order to assess if SDOH models and assumptions contributed to ease for their utility in value based care, it is important to understand if such models are providing more confidence to physicians to assume higher risk.

1.6 Aim of Research

In order to achieve the objective of this research, there has been a selection of sub-objectives/aims that has been set:

1. Explore and develop a better understanding of how SDOH plays a role in people's health and the governmental and industry efforts in getting it adopted for accelerating transition of healthcare models to Value Based Care.

2. Assess the data challenges plaguing the collection, usage and adoption of SDOH factors.

3. Review the utility of technology interventions through sample validations.

4. Evaluate the prospect development of a holistic SDOH framework with a view to support Whole health practices.

1.7 Significance of study

The outcome of this research will be helpful and valuable to Providers and Payers as well as the healthcare industry overall. If a scientific methodology based on impact of societal factors on a person's health can be arrived at and adopted by different stakeholders in the ecosystem, it could be of immense benefit to patients and on health outcomes. It can also reduce Provider-Payer friction and develop trust that is a key ingredient to accelerated adoption of Value based care models in healthcare delivery.

1.7.2 Fast Healthcare Interoperability Resources (FHIR) standards are enabling healthcare data interoperability

Health Level Seven International (HL7) is a not-for-profit standard developing organization for providing a comprehensive framework and related standards for the exchange, integration, sharing, and retrieval of electronic health information that supports clinical practice and the management, delivery, and evaluation of health services. HL7 is supported by more than 1,600 members from over 50 countries, including healthcare providers, government stakeholders, payers, pharmaceutical companies, vendors/suppliers, and consulting firms.

The Fast Healthcare Interoperability Resources (FHIR) is a HL7 standard is a set of rules and specifications for exchanging electronic healthcare data. and advance interoperability. FHIR is composed of foundational, infrastructure, administrative, data exchange, and clinical reasoning capabilities. It facilitates representation and sharing of information among clinicians and organizations in a standard way, irrespective of how local EHRs represent or store the data. FHIR standards have a potential to improve EHR data sharing in clinical settings, promotes interoperability, reduces effort to implement new measures, improves alignment between clinical measures and clinical decision support systems.

1.8 Research Questions

Here the research is going to answer the following research questions:

1. How are industries contributing to the adoption of SDOH for the acceleration of Value-Based Care?

2. What indicators within the framework are most crucial for determining the progression of chronic diseases?

3. What role do individual behaviors and community factors play in the relationship between SDOH and Dementia and Alzheimer's prognosis?

4. Are there specific demographic or environmental factors that exacerbate the impact of SDOH more than others on Dementia and Alzheimer's prognosis?

This research aims to uncover the influence of Social Determinants of Health (SDOH) on individual health outcomes, investigate government initiatives for integrating SDOH into healthcare models, explore industry efforts in adopting SDOH for Value-Based Care, and identify challenges and opportunities in the incorporation of SDOH into healthcare systems. Research seeks to define key markers of Social Determinants of Health (SDOH) relevant to chronic disease progression, create a comprehensive impact-based framework integrating these markers, identify crucial indicators for assessing chronic disease progression, and evaluate how the proposed framework contributes to the management of chronic diseases for improved health outcomes.

Also focuses on identifying specific Social Determinants of Health (SDOH) markers associated with dementia, developing methods to quantify their impact, exploring demographic and environmental factors that may exacerbate this impact, and examining the implications of these findings for dementia prevention and management strategies. Finally, to uncover the Social Determinants of Health (SDOH) markers significantly contributing to the impact on Alzheimer's, analyze variations in this impact across different demographic groups, identify interventions targeting SDOH markers to mitigate the impact on Alzheimer's, and examine the role of individual behaviors and community factors in the relationship between SDOH and Alzheimer's outcomes.

CHAPTER II:

REVIEW OF LITERATURE

2.1 Healthcare Ecosystem in US

Healthcare delivery in the United States is a complex and multifaceted challenge. United States has some of the most advanced medical technologies and treatments in the world, and many Americans have access to high-quality healthcare services. However, there are also significant challenges and shortcomings in the healthcare system that affect access, cost, and quality of care.

Access to healthcare is a major issue in the United States. Millions of Americans are uninsured or underinsured, which means they may struggle to afford healthcare services or may delay seeking care until their condition worsens. Even for those with insurance, out-of-pocket costs for healthcare can be high, leading some people to skip or delay necessary medical care.

The cost of healthcare is also a significant concern. The United States spends more per capita on healthcare than any other country in the world, yet many people still lack access to affordable care. Rising healthcare costs have also contributed to the financial strain on individuals, families, and the economy.

Quality of care is another area of concern in the US healthcare system. Despite the advanced medical technologies and treatments available, there are significant disparities in health outcomes across different populations. In certain areas like maternal and infant mortality rates, US lags many other countries.

Overall, the state of healthcare delivery in the United States is characterized by a mix of strengths and weaknesses. While there are some innovative and high-quality healthcare providers and systems, there are also significant gaps in access, cost, and quality of care that need to be addressed.

2.1.1 Health Equity

The Centers for Medicare and Medicaid Services (CMS) Health Equity Pillar defines health equity as the ecosystem where everyone can attain their highest level of health regardless of gender, disability, ethnicity, sexual orientation, geography, preferred language, socioeconomic status or factors that affect health outcomes and access to care.

Within the CDC, DNPAO is part of the National Center for Chronic Disease Prevention and Health Promotion. According to DNPAO, following factors influence Health Equity:

• Social Determinants of Health - Differences in SDOH contribute to persistent chronic disease disparities among racial, ethnic, and socioeconomic groups as well as in different geographies and among people with different physical abilities.

• Racism - It is a system of structures, policies, practices, and norms that assigns value and determines opportunity because of the way people look or the colour of their skin. This results in conditions that unfairly give advantages to some and disadvantages to others. These advantages and disadvantages are passed down through generations. Racism, both interpersonal and systemic, limits the ability for some groups to build wealth by determining who owns land, buys houses, gets a quality education, and gets living wage jobs. Racism also affects access to quality health care (Healthy People 2030 report).

In recent years, the United States has made several efforts to promote health equity.

• The Affordable Care Act (ACA): The Patient Protection and Affordable Care Act (2010), also known as Obamacare, which into effect in 2010, expanded access to health insurance for millions of Americans, including those with pre-existing conditions and those who could not afford coverage. It also included provisions aimed at reducing health disparities, such as requiring insurance companies to cover preventive services without cost-sharing and investing in community health centres.

• The National Partnership for Action to End Health Disparities (NPA): The NPA is a national movement aimed at reducing health disparities and achieving health equity. It includes a partnership of more than 2,000 organizations and individuals, relying heavily on those on the front line who are actively engaged in minority health work at multiple levels, to address the social, economic, and environmental factors that contribute to health disparities.

• Many states and local governments have implemented health equity initiatives aimed at improving health outcomes for marginalized communities. For example, in 2017, the City of New York launched the NYC Care program, which provides affordable healthcare to uninsured and underinsured residents, with a particular focus on undocumented immigrants and other vulnerable populations.

• The COVID-19 pandemic highlighted the importance of addressing health equity, as it disproportionately impacted communities of colour and those with low-income. In response, the federal government launched several initiatives aimed at addressing health disparities related to COVID-19, such as investing in community-based testing and vaccination sites in underserved areas and expanding access to telehealth services.

Overall, there are ongoing efforts in the United States to promote health equity, but there is much work to be done to ensure that all individuals get the opportunity to achieve good health outcomes regardless of their background or socioeconomic status (Peltz et al., 2020). This includes things like:

1. Investing in public health infrastructure: Governments can invest in programs and policies that support public health, such as disease surveillance, vaccination programs, and health education campaigns.

2. Expanding access to healthcare: Governments can expand access to healthcare by increasing funding for programs like Medicaid, which provides health insurance to low-income individuals, or by investing in community health centers that serve underserved populations.

3. Addressing social determinants of health: Governments can address social determinants of health, such as poverty and lack of access to healthy food, by implementing policies and programs that support economic development, affordable housing, and healthy food access.

4. Advancing health equity in policies and programs: Governments can ensure that their policies and programs are designed to advance health equity by considering the impact on marginalized communities and working to reduce health disparities.

2.1.2. Inconsistent prevalence of diseases despite Health Equity efforts

Despite Health equity efforts, diseases are not equivalently spread across communities. There are several reasons why the prevalence of diseases may vary across different communities. Some possible explanations include:

• Genetics: Certain diseases may be more prevalent in certain ethnic or racial groups due to genetic factors. For example, sickle cell anaemia is more common among African Americans and Hispanics, while cystic fibrosis is more common among Caucasians.

• Environmental factors: Different communities may be exposed to different environmental factors that can impact disease prevalence. For example, communities living near industrial sites may have higher rates of cancer due to exposure to pollutants.

• Access to healthcare: Communities with limited access to healthcare may have higher rates of certain diseases due to a lack of preventative care and early detection. This can be due to various factors like poverty, geographic isolation, and cultural barriers.

• Lifestyle factors: Lifestyle factors such as diet, exercise, and smoking habits can impact disease prevalence. For example, communities with a high prevalence of obesity may have higher rates of diabetes and heart disease.

These factors are interconnected and can influence each other. For example, limited access to healthcare can lead to delayed diagnoses and poor management of chronic conditions, which can worsen health outcomes.

As per Clinical Knowledge Summaries report (NICE), in an examination of a biobank of more than 400 000 records in the United Kingdom with cardiorespiratory comorbidity, “Black” participants had more than two-fold higher risk for hospitalization even after adjustment for the Townsend Deprivation Index (a composite measure of socioeconomic deprivation (Martin A., 2022).

2.2 Social Determinants of Health (SDOH)

SDOH are factors that affect the health of an individual and can be classified into four categories: social and economic conditions, physical environment, social relationships, and personal behaviours. WHO's study shows that social determinants can have a significant impact on how healthy people are, and how different types of determinants interact with each other to affect our well-being.

As per WHO studies, the SDOH have an important influence on health inequities - the unfair and avoidable differences in health status seen within and between countries. In countries at all levels of income, health and illness follow a social gradient: the lower the socioeconomic position, the worse the health. Research shows that the social determinants can be more important than health care or lifestyle choices in influencing health. For example, numerous studies suggest that SDOH account for between 30-55% of health outcomes. In addition, estimates show that the contribution of sectors outside health to population health outcomes exceeds the contribution from the health sector. These studies show that non-medical (environmental) factors can have a significant impact on how healthy people are, and how different types of determinants interact with each other to affect our well-being (Sills et al., 2016).

Some provisions of Patient Protection and Affordable Care Act (2010), reflect overwhelming evidence that to reduce healthcare costs, and to improve quality of care and population health, the Social Determinants of Health (SDOH) must be addressed. These policies include funding for partnership between public health agencies, community organizations, healthcare institutions, promotion of value-based payment models that incentivize integrated health and social care delivery, and support for Medicaid program innovations that directly address social needs as part of healthcare.

2.2.1 World Health Organization (WHO) Model

The WHO works to address social determinants of health by compiling and disseminating evidence on what works to address these determinants to help build capacity and advocate for more action. As per their latest report, COVID-19 and the social determinants of health and health equity (2021), SDOH can be grouped into 5 domains:

Table 1
SDOH factors in WHO data model

S.No	SDOH domains	SDOH determinants	Data source
1	Sociodemographic	Sexual orientation	**Not specified**
2	Domains	Race/ethnicity	**Not specified**
3		**Country of origin/U.S. born or non-U.S. born**	**Not specified**
4		Education	**Not specified**
5		Employment	**Not specified**
6		Financial resource strain (Food and housing insecurity)	**Not specified**
7		Gender identity	**Not specified**
8	Psychological	Health literacy	**Not specified**
9	Domains	Stress	**Not specified**
10		Negative mood and affect (Depression, anxiety)	**Not specified**
11		Psychological assets (Conscientiousness, patient engagement/activation, optimism, self-efficacy)	**Not specified**
12		Negative mood and affect (Hostility and anger, hopelessness)	**Not specified**
13		Cognitive function in late life	**Not specified**
14		Psychological assets (Coping, positive affect, life satisfaction)	**Not specified**
15	Behavioural Domains	Dietary patterns	**Not specified**
16		Physical activity	**Not specified**
17		Tobacco use and exposure	**Not specified**
18		Alcohol use	**Not specified**
19		Abuse of other substances	**Not specified**

20		Sexual practices	**Not specified**
21		Exposure to firearms	**Not specified**
22		Risk-taking behaviours (Distractive driving and helmet use)	**Not specified**
23	Individual-Level Social Relationships	Social connections and social isolation	**Not specified**
24	and Living	Exposure to violence	**Not specified**
25	Conditions Domains	Social support (Emotional, instrumental, and other)	**Not specified**
26		Work conditions	**Not specified**
27		History of incarceration	**Not specified**
28		Military service	**Not specified**
29		Community and cultural norms (Health decision making)	**Not specified**
30	Neighbourhoods and Communities	Neighbourhood and community compositional characteristics (Socioeconomic and racial/ethnic characteristics)	**Not specified**
31		Neighbourhood and community contextual characteristics (Air pollution, allergens, other hazardous exposures, nutritious food options, transportation, parks, open spaces, healthcare and social services, educational and job opportunities)	**Not specified**

2.2.1 The GRAVITY project

The Gravity project is national public collaborative that started in 2017 for developing data standards to improve usage, sharing integration of social determinants of health (SDOH) information in clinical care and across disparate digital health and human service platforms. These national standards support the consistent use of the data across organizations, providers, and caregivers, and help to facilitate payment for social risk data collection and intervention activities such as referrals, counselling, and care coordination.

The Gravity Project divides each data set into four clinical processes – screening, diagnosis, goal setting, and interventions - while promoting individual privacy, safety, security, and accountability for patient records. It also provides processes and guidelines on how to record, document, and exchange such SDOH information. It invites entities to test the evolving terminology and data exchange standards with a goal to validate the use of Gravity-identified coded terminologies and their FHIR Implementations through real-world testing across clinical, social services, payer, and government electronic systems.

2.2.2 SDOH adoption efforts are usually designed by academicians and there is a need to bring Human Centric Design approaches in the game

In a technology-driven world, we should not lose sight of the human element. Human-centred design is the foundation for building simple and easy health care technology and programs. Data and analytics play a key role.

There are many classifications for SDOH data but most of the models are an attempt to address population health management requirements. However, these are missing the requirements of personalized health care programs and risk modelling, example none of the models could drill down to which SDOH factors are most important for which disease, that is most important for patients and physicians. These also do not consider clinician and consumer experiences that could help in adoption of the models. Neither do these take into account health outcomes, that is an important aspect to be considered.

Merging data with design can ensure human connection plays an integral role in care. Hence the author is proposing an alternate design of SDOH classification and its usage that is disease centric. This will improve the utility of such efforts and models in clinical settings and disease management.

2.2.3 Value chain of social determinants data is broken!

There are 1346 Z-codes at present but only a few are used by providers currently. Even after MR Stage 3 requirements, providers find it burdensome to record these in clinical information.

- These are inadequately classified for simplicity.

- There are too many for practical purposes – Can we find a new simpler classification method for Z-codes

Some Z-codes are mixed, eg below hence it is difficult to segregate them into rules unless a NLP model is applied.

Z0283 - Encounter for blood-alcohol and blood-drug test

Z6372 - Alcoholism and drug addiction in family

Medical error is the third leading cause of death after heart failure and cancer. Most the errors happen due to data unavailability or lower quality. Data Centric-AI has the potential to improve Data availability and accessibility challenges. Moreover, SDOH factors are interconnected and need fine tuning with co-relation based on empirical evidence.

2.3 Population Health Management

As the common knowledge goes, the value of preventing disease and disability is much greater than the investment required in public education, provider incentives, and public health infrastructure. However, only a tiny fraction of the US healthcare investment supports prevention and health promotion. Different States are implementing cost-effective and diverse strategies for prevention, early detection, and control of diseases like cancer, diabetes, heart disease & stroke, and arthritis. It appears that CDC, DHHS agencies and State Health Departments require substantial investments to be able to make a bigger impact.

Chronic disease affects health and quality of life and nearly 60% of adult Americans have at least one chronic disease. More than two-thirds of all deaths are caused by one or more of five chronic diseases: diabetes, cardiovascular disease, chronic obstructive pulmonary disease, stroke and cancer.

A recent Partnership to Fight Chronic Disease publication determined that treatment of the seven most common chronic diseases, coupled with productivity losses, will cost the U.S. economy $2 trillion dollars annually by 2030 ie $8,600 per person. The same analysis estimates that reductions in unhealthy behaviours could save 1,100,000 lives per year. Healthcare costs for people with a chronic condition is five times higher than those without such a condition on an average.

Improving prevention in chronic diseases would result in significant cost savings. Example, according to Commentary on Chronic Disease Prevention (2022), Medicare spending can be reduced by $14 billion by increasing the colorectal cancer screening rate to 70%.

2.3.1 Risky behaviors among the population are increasing

More and more people are being categorized as “high risk” for multiple chronic diseases in the ageing American population. It is important to recognize that the risk and impact is based on individual’s choices. Risky behaviours such as poor diet, lack of physical activity, tobacco use, and ignoring known risks like family history result in a significant increase in chronic conditions. Many people are routinely missing or ignoring their body’s warnings about the onset of chronic diseases. Sometimes they are unable to receive preventive care due to social or economic barriers. The outcome is poor collective health quality in US, resulting in much higher spending on healthcare (Huang et al., 2021).

2.4 Direct & Indirect costs to US economy

US spends $4.1 trillion dollars in health care costs. In the past 20 years, chronic disease has grown steadily, and today affects 50% of the U.S. population. 90% of this cost is spent for people with chronic and behavioural health conditions. As per data published by CDC and NIH, below is the Direct cost burden for different diseases in Unites States.

Table 2
Economic burden of Direct costs

Common ailments	Morbidity (% of US population affected)	Mortality rate	Economic burden of Direct costs (in Billion dollars)	Lifetime cost
Diabetes type 2	30 MN (Roughly 9% of population)	102,188 (2020 data)	$237 Bn (2021 data)	Not available
Alzheimer’s (Dementia)	6.5 Mn (Roughly 2% (10.7% with age 65 and older. Almost 2/3 are	134,242 (2020 data)	$321 Bn (2021 data)	$400K

	women)			
Cardiovascular (CVD)	55.9 MN (Roughly 4.9% of population)	696,692 annually (2020 data)	$229 Bn (2021 data)	Not available
Cerebrovascular (Stroke)	Roughly 10 MN (3.1% of population)	160,264 annually (2020 data)	$65 Bn (2021 data)	Not available
Cancer	Roughly 1.7 Mn (.025% in age groups under age 20, .35% among those aged 45–49, more than 1% in age groups 60 years and older)	609,360 (2022 data)	$200.7 Bn (2020 data)	$150K
Chronic Kidney diseases (CKD)	Roughly 6 Mn diagnosed cases. 90% do not know about the disease!	52,574 (2021 data)	$49 Bn (2021 data)	Not available
Parkinson's	0.22% men and 0.32% women	-	$29.6 Bn (2021 data)	Not available
Depression	15.6 MN (4.7% of population)	45,979 (2020 data)	$92 BN (2018 data)	Not available
Mental Health	39.8 MN (11.7% of population)	45,979 (2020 data)	$63-92 BN (2022 data)	Not available
Arthritis	59 MN (20% of US population)	NA	$140 Bn	Not available
Epilepsy	3.4 MN (Roughly 1.2% of US population)	NA	$8.6 BN	Not available
Tooth decay	80 MN (Roughly 25% of US population)	NA	Not available	Not available

Common chronic diseases are also a significant driver of productivity costs for the economy. Absenteeism also has a high impact on business. As per data published by CDC, NIH and several governmental and non-profit agencies, below is the Indirect cost burden for different diseases in United States.

Table 3
Economic burden of Indirect costs

Common ailments	Morbidity (% of US population affected)	Economic burden of Indirect costs (In Billion dollars)
Diabetes type 2	30 MN (Roughly 9% of population)	Roughly $340 BN
Alzheimer's (Dementia)	6.5 Mn (Roughly 2% (10.7% with age 65 and older. Almost 2/3 are women)	Roughly $260 Bn
Cardiovascular (CVD)	55.9 MN (Roughly 4.9% of population)	Roughly $118 BN
Cerebrovascular (Stroke)	Roughly 10 MN (3.1% of population)	Roughly $29 BN
Cancer	Roughly 1.7 Mn (.025% in age groups under age 20, .35% among those aged 45–49, more than 1% in age groups 60 years and older)	Roughly $470 BN
Chronic Kidney diseases (CKD)	Roughly 6 Mn diagnosed cases. 90% do not know about the disease!	Not available
Parkinson's	1.04 MN (0.22% men and 0.32% women)	Roughly $21 BN
Depression	15.6 MN (4.7% of	Roughly $144 BN

	population)	
Mental Health	39.8 MN (11.7% of population)	Roughly $140 BN
Arthritis	59 MN (20% of US population)	Roughly $164 BN
Epilepsy	3.4 MN (Roughly 1.2% of US population)	Roughly $7 BN
Tooth decay	80 MN (Roughly 25% of US population)	Roughly $45 BN

2.5 Value Based Care (VBC)

Value-based Care (VBC) emphasizes on quality of service and the patient experience with a goal to provide high-quality, cost-effective care for patients, while at the same time reducing administrative burdens on providers and payers. These models are becoming more popular, particularly in the United States. According to the HCPLAN Measurement Report (2021), the percentage of healthcare payments tied to VBC models increased from 23% in 2015 to 36% in 2018. The same survey also found that 43% of payments to Medicaid programs were tied to VBC models in 2018.

Kshirsagar et al (2022) did a study on VBC arrangements related to chronic kidney disease (CKD) care in population health management commercial market. They weighed the expected benefits through the lens of potentially addressing the adverse social determinants of health (SDOH) associated with kidney disease, along with potential risks for stakeholders. They opined that Kidney-VBC companies may be uniquely positioned to address health equity. The multidisciplinary care teams can help patients with relevant adverse SDOH, and provide education on optimal diet, access to healthy food areas, and KRT options.

Table 4
Kidney focused VBC companies

S.No	Name of Company (url)	Year Founded	Location(s)	***Current Partners***
1	Cricket Health (https://www.crickethealth.com/)	2015	San Francisco, Massachusetts, Texas	Fresenius Health Partners, Interwell Health, Cigna, Baylor Scott & White Health Plan
2	Monogram Health (https://www.monogramhealth.com/)	2019	Southern United States (Tennessee)	Cigna, Humana
3	Reach Kidney Care (https://www.reachkidneycare.org/)	2011	Southern United States, New Jersey, New York, Pennsylvania, Ohio, Missouri, Montana	Humana
4	Somatus (https://somatus.com/)	2016	Washington DC, East Coast, Southern United States	Anthem, BlueCross BlueShield Tennessee
5	Strive Health (https://www.strivehealth.com/)	2018	Missouri, Pennsylvania, North Carolina	Bon Secours Mercy Health, SSM Health, Humana, various BlueCross BlueShield health plans, NANI
6	Evergreen Nephrology (https://evergreennephrology.com/)	2021	Tennessee	RenalCare Associates, University of Pennsylvania, Colorado Kidney Care

Sanford (2000) has analysed the availability paediatric health systems require investments in identifying and mitigating SDOH in children. These investments may be channelled by changing Federal and State policy for utilizing Medicaid funds for non-medical interventions and introducing the Accountable Health Communities model to paediatrics. He has discussed the possibility of value-based payments to Accountable Care Organizations for these interventions.

CMS has more than 50 VBC models which it offers to Providers for getting into value-based contracts in the Medicaid and Medicare market. However, FFS models will likely continue to be used, particularly in Medicare programs, for the foreseeable future.

Overall, the trend is towards VBC as healthcare organizations seek to reduce costs while improving the quality of care. As proposed by Rahul Sharma (2020) and Chris Bethell (2021), SDOH parameters play a key role in determining VBC outcomes. However following challenges are afflicting the growth of this trend:

• Data sharing challenges: Private companies are generally not comfortable to share proprietary business practices without adequate regulation. They will likely opt to keep their performance on quality measures opaque as it reduces a company's advantage in the marketplace. Current regulations on data sharing practices may dampen overall innovation in VBC.

• Patient trust: Medical error is the third leading cause of death after heart failure and cancer. Kshirsagar et al (2022), in their study of the business model of several Kidney-VBC commercial companies, opined that the patient-provider relationship is already tenuous in underserved communities, where patients' prior experiences understandably lead to a distrust of the health care system. Explicitly tying medical care to profits, a key goal for the kidney-VBC companies and investors, could further erode the trust and may disrupt existing relationships between patients and providers. If poorly implemented, the teams could fragment care and confuse patients about who is directing care.

• Patients asking for Personalized care: As per a paper published by Ahmed et al (2020), Precision medicine can change the way we approach the traditional approach to medicine, which is symptom driven. To build a personalized healthcare model, we need to gather comprehensive patient and demographic information, which will enable the understanding of biological indicators and shifts in health, allowing earlier interventions. By integrating Electronic Health Records (EHR) with public health data, it will be possible to identify important patterns in disease progression for individual patients. In the study, they aimed to advance a new data-centric era of discovery in healthcare by analysing various published AI solutions, approaches, and perspectives (Motamedi et al., 2021).

2.5.1 Better understanding of Whole-person health

In the past 20 years, chronic disease has grown steadily, and today affects 50% of the U.S. population. Behavioural health affects 20% of Americans, causing significant and avoidable disability and death.

Mental Health America estimates that as many as 26 million people do not have access to the behavioural health resources and treatment they need. This limits their ability to live whole and productive lives. And even where it is available, the pressures of home, food and personal insecurity too often push health concerns into the background.

Gallo (2020), a professor in Mental Health, focuses his research on the intersection of physical and mental health. According to him, there's ample clinical and epidemiologic evidence that shows the risk for depression is higher among those who suffer from chronic illnesses. He opined in The American Journal of Geriatric Psychiatry that chronic illness and depression have a high correlation. People with such conditions and with depression have a higher risk of mortality than a person with similar ailment but without depression.

2.5.2 Shifting healthcare delivery from hospital settings to community care

CDC's Division of Nutrition, Physical Activity, and Obesity (DNPAO) works in prevention of chronic diseases by promoting good nutrition, regular physical activity, and a healthy weight. It has published Community Health Programs (2019) report on its website.

Racial and Ethnic Approaches to Community Health (REACH) is a community impact program run by DNPAO that was started in 1999. It has demonstrated that locally based and culturally tailored solutions can be effective in bridging the gaps in health that diverse communities in urban, rural, and tribal areas experience. As per the last report from REACH program for the period 2014-2018 that has been published on its website,

• Over 2.9 million people have better access to healthy foods and beverages.

• Over 322,000 people have benefited from smoke-free and tobacco-free interventions.

• Approximately 1.4 million people have more opportunities to be physically active.

• Over 830,000 people have access to local chronic disease programs that are linked to clinics.

Some examples of impact of REACH program are,

1. Partners in DeKalb County, Georgia, increased access to healthy foods for approximately 242,000 African Americans, selling more than 1,000 units of fruits and vegetables each week and reporting a 34% increase in consumption of fruits and vegetables among customers.

2. The REACH program at Creighton University partnered with the Omaha Housing Authority to create safer places for physical activity for over 330 residents of three low-income housing towers in Omaha, Nebraska.

3. In Orange County, California, the REACH program increased access to smoke free environments for more than 100,000 Asian American residents by increasing the number of commercial shopping plazas with voluntary smoke free policies.

4. Partners in Cuyahoga County, Ohio, developed the Produce Prescription Program for Hypertension to connect residents in need to healthy eating information and resources. Over 600 low-income patients with high blood pressure have better access to nutrition education and affordable produce.

5. The Toiyabe Indian Health Project increased the availability of healthy foods for over 3,000 American Indians in seven tribes and two tribal communities by increasing healthy food production in community gardens.

6. In Los Angeles, California, the Community Health Council (CHC), in collaboration with the African Americans Building a Legacy of Health program, worked to increase access to healthy and affordable food and beverages through efforts to change institutional practices and promote local investment. Specifically, these efforts have helped to leverage support from California's $200 million Fresh Food Financing Fund that seeks to eliminate food deserts and fight childhood obesity.

Partnerships to Improve Community Health (PICH) worked to make healthy living easier and more affordable where people live, learn, work, and play. To improve health and wellness in their communities, it focused on four risk factors:

• Tobacco use and exposure.

• Poor nutrition.

• Physical inactivity.

• Lack of access to opportunities for chronic disease prevention, risk reduction, and disease management.

Some examples of PICH community strategies include:

1. Protecting people from second-hand smoke exposure in indoor and outdoor spaces.

2. Promoting nutrition guidelines that encourage healthy food and beverage options in schools and worksites.

3. Increasing physical education classes so children have more physical activity opportunities each day.

4. Increasing the number of multi-disciplinary teams (i.e., physicians, pharmacists, community health workers) that help patients manage their chronic diseases.

2.5.3 Adoption of digital health records and technologies

Healthcare is a team effort, and much of the value derived from the healthcare delivery system results from the effective communication of information from one party to another and the ability of multiple parties to engage in interactive communication of information.

As per the eCQI Resource Center, an electronic health record (EHR) is a "longitudinal electronic record" of patient health information generated during their encounters in care delivery settings. It is also known as the electronic patient record, electronic medical record, or computerized patient record. EHRs include patient demographics, vital signs, problems, progress notes, medications, laboratory data, diagnoses and treatment, medications, allergies, immunizations, radiology images, laboratory results and past medical history.

The availability and accessibility of healthcare data in recent years has seen a significant increase due to the widespread adoption of electronic health records (EHRs). This data can be used by healthcare providers and researchers to identify patterns, trends, and potential health risks, as well as to develop and evaluate new interventions and treatments. In addition to EHRs, other sources of healthcare data that are becoming more widely available include wearable health monitors, mobile health apps, and telemedicine platforms. This has led to a wealth of data being generated in healthcare settings, including data from patient monitoring devices, lab tests, and clinical trials. These technologies are allowing patients to monitor their health and share data with their healthcare providers remotely, improving access to care and enabling real-time interventions. With fully functional EHRs,

• Information collected by the primary care provider is available emergency department clinician even if the patient is unconscious, eg, patient's life-threatening allergy information in EHR can be used to adjust emergency response appropriately.

• Duplicate lab tests can be avoided if the most recent ones are available in the record for the specialist.

• A patient can see the trend of the lab results by logging on to his/her own EHR record, which can help motivate him to take his medications and keep up with the lifestyle changes.

• The clinician's notes from the patient's hospital stay can facilitate discharge and follow-up care instructions, enabling the patient to move from expensive care setting to another one more smoothly.

However, there are still significant challenges to the widespread adoption and effective use of healthcare data for technology-based interventions. These include concerns around data privacy and security, as well as the need for standardized data formats and interoperability between different healthcare systems and devices. There are efforts underway to address these challenges and create a more cohesive and integrated healthcare data ecosystem, which is expected to drive further innovation and improve patient outcomes. Overall, the increasing availability and accessibility of healthcare data is creating new opportunities for technology-based interventions in healthcare(Stuckler, D. 2008).

Maysoun et al (2019) did a study with six EHR vendors in ambulatory and inpatient settings to understand on how to enable better data analytics in population health management space and what the vendors expect from any such software. As per a study, vendors recognized the need for more standardization of SDOH performance measures across various federal and state programs, better mapping of SDOH measures to multiple types of codes, development of more codes for all SDOH measures of interest, and interoperability of SDOH data. Vendors indicate they are actively seeking solutions to data standardization and interoperability challenges through internal product decisions and collaboration with policymakers.

2.5.4 Regulations involving capture and sharing of healthcare data are evolving

Meaningful Use Requirements regulation govern how and what data should be recorded in clinical settings. Meaningful Requirements (2017) reported that Centers for Medicare & Medicaid Services (CMS) has released the certification requirements for Stage 3 Meaningful Use regulations.

Table 5
Evolution of Meaningful Use regulation

Stage 1: Meaningful Use criteria focus on:	**Stage 2: Meaningful Use criteria focus on:**	**Stage 3: Meaningful Use criteria focus on:**
Electronically capturing health information in a standardized format	More rigorous health information exchange (HIE)	Improving quality, safety and efficiency leading to improved health outcomes
Using that information to	Increased requirements for	Decision support for

track key clinical conditions	e-prescribing and incorporating tab results	national high-priority conditions
Communicating that information for care coordination process	Electronic transmission of patient care summaries across multiple settings	Patient access to self-management tools
Initiating the reporting of clinical quality measures and public health information	More patient-controlled data	Access to comprehensive patient data through patient centred HIE
Using information to engage patients and their families in their care		Improving population health

MU Stage 3 became a law in 2017 and came into effect in 2018. It builds on the framework established in previous meaningful use stages and continues to promote EHR interoperability. The rule also focuses on providing more flexibility that simplifies the reporting requirements of the Medicare and Medicaid EHR Incentive Programs. There are eight major objectives proposed in the Stage 3 Meaningful Use rule.

Table 6
Meaningful Use Stage 3 requirements

S.No.	Objective	Description
1	Protect Patient Health Information	CMS focuses on the Health Insurance Portability and Accountability Act (HIPAA) and its aim to prevent the identification of patient health data. It introduces new technical, physical, and administrative safeguards that provide more strict and narrow requirements for keeping patient data safe and secure. Eg encryption of patient electronic health information, conducting risk analysis, to develop contingency plans and training programs, and

		ensuring physical safeguards like facility access controls and workstation security
2	Electronic Prescribing	E-prescribing is key in preventing medical errors and preventing illegal, fabricated prescriptions that are associated with drug abuse behaviours. Providers will continue to send more than 80 percent of their drug/treatment prescriptions electronically through certified EHR systems to prevent fraudulent prescribing and keep patient data secure, and to send more than 25 percent of hospital discharge medication orders through certified EHR systems.
3	Clinical Decision Support	It focus on improving performance on high-priority medical conditions by integrating clinical decision support tools and strategies to better patient safety and efficiency within the healthcare sector. CMS suggests eligible providers to implement five clinical decision support interventions that are adherent to at least four clinical quality measures at critical points in patient care, and to incorporate tools for drug-drug and drug-allergy interaction alerts for the entire EHR reporting period.
4	Computerized Provider Order Entry	CMS mandated that computerized provider order entry (CPOE) needs to be used for recording medication, laboratory, and radiology requests by eligible providers for more than 80 percent of medication orders, more than 60 percent of laboratory orders created will need to be recorded through the computerized order entry form, and to use CPOE for more than 60 percent of diagnostic imaging

		orders during the HER reporting period.
5	Patient Electronic Access to Health Information	CMS mandates eligible hospitals and providers to increase electronic access to patients to health data calls and their health information within 24 hours of its availability, increasing patient access to patient reminders, patient-specific education tools/resources, clinical summaries of medical appointments, and provide the ability to review and share health information with a third party
6	Coordination of Care through Patient Engagement	Patient engagement is focussed on increasing patient involvement in healthcare by changing prior behaviors among both providers and patients and improving health literacy among the patient population. CMS mandates healthcare professionals to utilize secure and private communication capabilities of certified EHR technology to work with patients or authorized caregivers regarding the patient's care, expanding the amount of options providers have to communicate with patients under the EHR Incentive Programs including the use of APIs.
7	Health Information Exchange	To reduce medical errors, improve clinical care decisions and coordination of care, physicians to provide summary of care records when transitioning patients among healthcare settings, accessing summary of care records during the first encounter with a new patient, and integrating summary of care records from other providers into their certified EHR technology.
8	Public Health and Clinical Data	Importance on the communication channels and asking providers to be actively engaged with public health

Registry Reporting	agencies or clinical data registries and to submit electronic public health data meaningfully through certified EHR systems.

However, as per Kshirsagar et al (2022), there are areas where regulations are missing: for example,

• VBC companies do not have many regulatory guardrails eg, they are not bound by regulations preventing “cherry-picking” or “lemon-dropping.”

• VBC companies are not accountable for connecting silos, reducing costs, and addressing adverse SDOH.

• Policies that require greater transparency of outcomes, quality, costs, engagement with local community resources, and implementation would improve interoperability and accountability.

• Requiring reinvestment of a percentage of profits into infrastructure and a minimum percentage of dually eligible (Medicare or Medicaid) enrolees could help address structural inequities.

2.5.5 Evolving financial reimbursement models towards value-based care

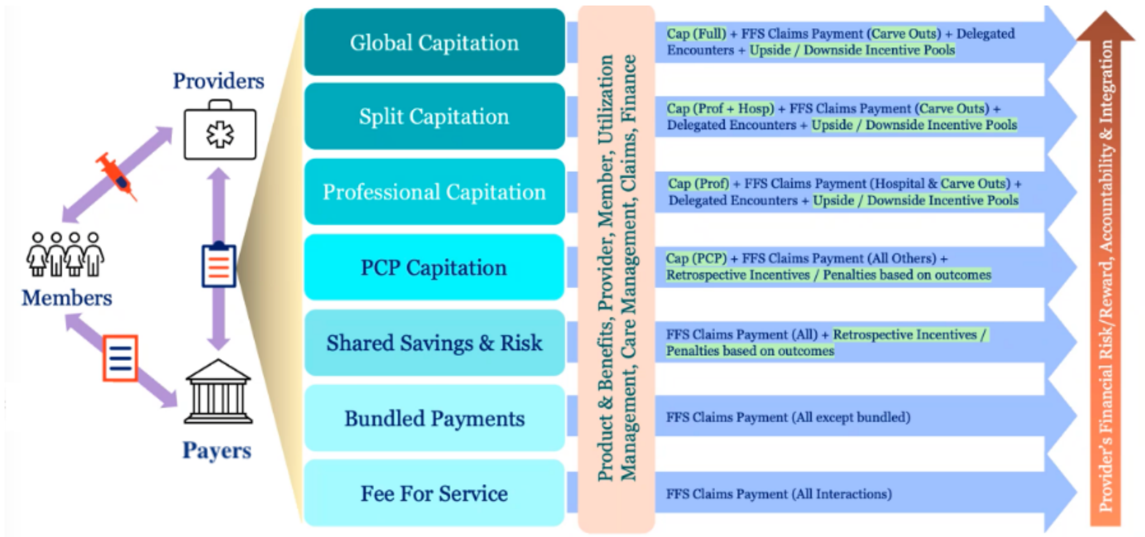


Figure 1 An online webinar by a Fortune 5 Healthcare services payer

The Medicare Access and CHIP Reauthorization Act MACRA) legislation was signed into law in 2015 and has been enhanced in 2020. It created the Quality Payment Program (QPP) for Medicare reimbursements to providers:

• Introduced the new Merit Based Incentive Payments System (MIPS) reporting system for clinicians under four categories wights as provides: Quality 45%, Cost 15%, Promoting Interoperability 25%, Improvement 15%

• Introduced MIPS Value Pathways framework that changes the way Medicare rewards clinicians for value over volume.

• Introduces penalties with performance lower than MIPS threshold.

• Gives bonus payments for exceptional performance beyond a threshold in eligible alternative payment models (APMs)

• Financial implications for MIPS eligible clinicians who choose not to report, with maximum penalty for not reporting rose to negative -9% (up from -7%) in 2020.

CMS has around 50 VBC models under MACRA for Medicare population, however the uptake of these plans needs to be increased. Similar reimbursement models need to be developed for Medicaid reimbursements.

2.5.6 Trust deficit due to unharmonized stakeholders and SDOH data is preventing adoption of value-based care

VBC objectives can be met if there is harmony between different stakeholders in the system. There is right disclosures, correct pricing of health plans, preventive and corrective prognosis during episodes of care without financial considerations, honest assessment of claims by payers.

VBC = Member + Provider + Payer.

Currently there are different interests for above stakeholders leading to a break in trust in the system. Example, a member may withhold prior medical history or comorbidities information in health plan disclosure to be in a position to reduce out of pocket expenses. Provider may have incorrect diagnosis by omission or commission in order to extend the inpatient procedure and inflate costs. Payer may decline a potential valid claim due to omission or commission of information.

Community information in US is available through various sources and systems of govt and non-profit agencies viz. American Census Bureau (ACS), Centre for Disease Control (CDC), Health and Human Services (DHHS).

Not all Qualitive health data is not interpreted into quantitively data for usage into anaylsis and prediction. Eg. A cross-sectional analysis of deaths in England found that relative mortality from COVID-19 was five times higher in households consisting of nine or more members. This could be a vital feature for improving SDOH.

To establish trust back, the system needs to be supported by a third party validation system ie community data;

In the new model, VBC = Member + Provider + Payer + Community information

There are specialized VBC companies being established that are willing to move the needle. However, there are questions that need to be answered,

• Do patients trust VBC companies due to the limited data and population health insights which is available in public domain?

• Are payments to VBC companies enough incentive to move them from Fee for Service model which is risk-free and with lower accountability?

2.5.7 Broadly, there is lack of reliable tools among patients and physicians to enable Connected Care

"Connected care is a way of pulling together all the aspects of what an individual can control or influence and connecting them to the full range of services available for their support. This includes their family, their employer, their community and their health system."- Rhonda Robinson Beale M.D. SVP Chief Medical Officer Mental Health Services, UnitedHealth Group

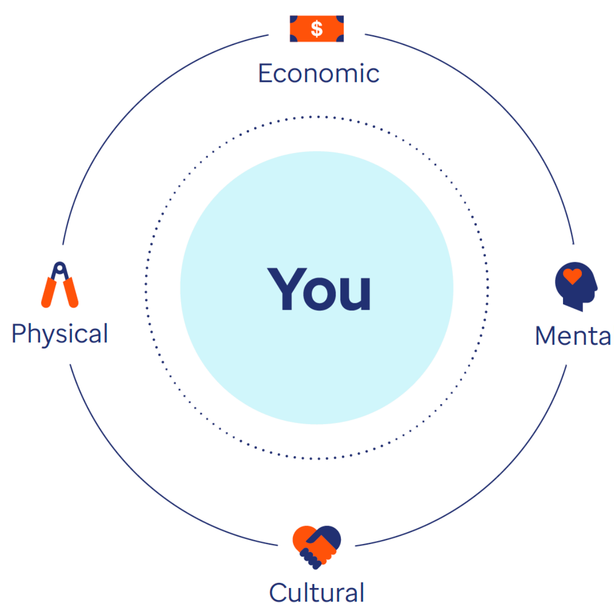


Figure 2 Connected care: Seeing the whole picture for whole person health

Based on literature research, there is a question mark on whether physicians and patients are equipped with adequate tools and insights to identify behavioural or mental health issues in addition to physical health (Lin et al., 2022).

2.5.8 Are innovative technology initiatives a way forward to address challenges impacting adoption of Value Based Care (VBC)?

In a recent event and competition in Optum (UHG company), following themes were identified as key innovation enablers for VBC.

- Prognostic Risk Modelling
- Fragmented and Ununiform Data
- Consumer Activation and Sustained Engagement
- Physician Empowerment and Engagement
- Financial Modelling Tools for VBC Adoption

Ideas were invited from 100000+ employees that generated 252 innovative ideas.

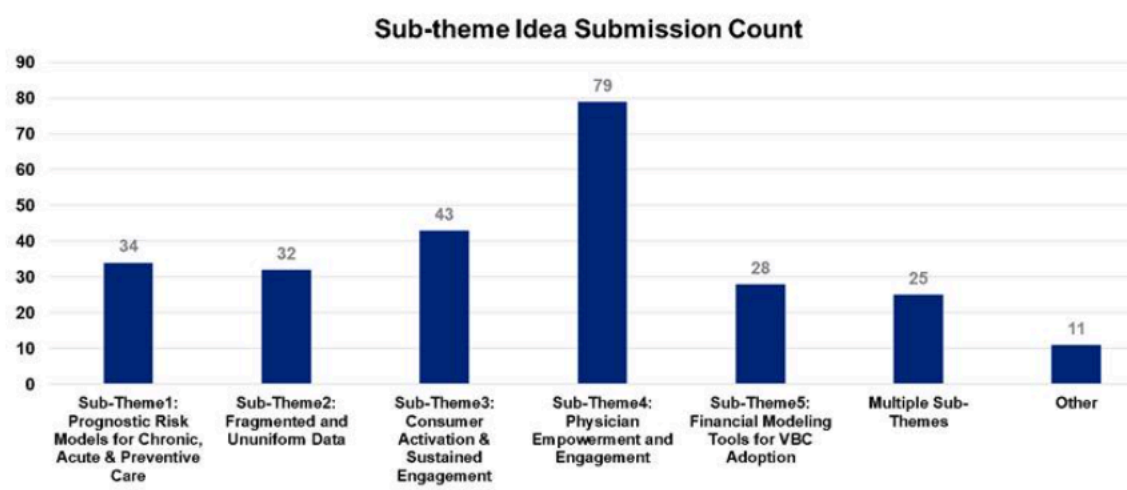


Figure 3 Ideation event by a Fortune 5 healthcare firm

31% of the ideas were related to Physician empowerment so that they can have a holistic view of a patient's physical and mental health, followed by 17% ideas that proposed that patients and their caregivers need to be activated and made aware of their holistic health and should be prodded to take timely prevention steps. This relates to Lin et al., (2021), assertion that Physicians are at the centre of value-based care arrangements.

The unavailability of holistic health information to physicians and patients is usually not available till later stages of disease lifecycle. At that time, physicians are not willing to take risk of capitated (fixed fee) arrangements hence this acts as a barrier to accelerating value-based care. Also, due to insufficient unavailability of social determinants information linked to a person, actionable insights using technology that are personalized for a patient, are usually not possible or are not very reliable. There is a need to improve data capture mechanisms and wherever there are operational barriers to capturing personalized data, can technology come to the rescue?

2.6 Machine Learning Techniques

2.6.1 Predictive and Prescriptive technologies are maturing leading to more meaningful interventions

The use of AI has contributed to the development of SDOH, Eg. use of machine learning algorithms is helping researchers identify patterns in data about social determinants and prescribe precision medicine. However as per Melzer (2022) and Tobin-Tyler et al (2021), there have been funding challenges for initiatives involving identifying and mitigating SDOH factors. This is primarily due to the uncertainty of outcomes and unavailability of enough data across communities organized in a fashion that can be used by scientists and policy makers for financial sustainability.

There have been some notable successes in using healthcare data for technology-based interventions. For example, machine learning algorithms have been used to analyse EHR data to identify patients at high risk of readmission or complications, allowing healthcare providers to intervene proactively. Wearable devices and mobile apps have also been used to collect data on patient behaviour and symptoms, providing valuable insights into how patients are managing their health outside of the clinic.

Improving health outcomes is a complex interplay between health system, community, and individual-level factors are increasingly seen as important to understanding patient health and identifying appropriate interventions. SDOH play a large role in bridging gaps in health inequity, improving population health outcomes and control healthcare costs.

Huang et al (2021) studied discharge disposition among heart failure and acute kidney injury in COVID-19 patients from 2 urban hospitals and proposed a new self-supervised machine earning framework, Deep significance clustering (DICE), that divided patients into clinically similar and risk-stratified subgroups that unsupervised clustering algorithms or supervised risk prediction algorithms could not generate. It also demonstrated the potential to apply DICE in heterogeneous populations, where having the same quantitative risk does not equate with having a similar clinical profile.

On the other side, Kasthurirathne et al (2017) conducted model experiments to evaluate the capacity for clinical, socioeconomic, and public health data sources to predict the need for various social service referrals among patients at a safety-net hospital. They integrated patient clinical data and community-level data representing patients' social determinants of health (SDOH) to build random forest decision models to predict the need for any, mental health, dietitian, social work, or other service referrals. They opined that the need for various social service referrals can be predicted with considerable accuracy using a wide range of readily available clinical and community data that measure socioeconomic and public health conditions. The also opined that use of SDOH did not result in significant performance improvements. Subsequently Jessica S Ancker and colleagues (2018) published a paper acknowledging the work of Kasthurirathne et al and highlighted possible reasons for differing results in their work while it is well established in medical science that population health is affected by socioeconomic status and other social determinants of health (SDOH). They submitted the reasoning for anomalies in previous work such as, a strong correlation of SDOH data existed with clinical data resulting in low impact on predictive power of SDOH (data diversity), insufficient variability in SDOH across patients (data diversity), community-level SDOH that was used could have been biased (data bias), and possible insufficiency of predictor variables (data diversity).

Amanda Lans et al (Aug 2022) did a systematic review to investigate whether prognostic ML models for orthopaedic surgery outcomes account for SDOH, and to what extent SDOH variables are included in the final models. Resources scanned were from PubMed, Embase and Cochrane for studies published up to 17 November 2020. Two reviewers independently extracted SDOH features using the PROGRESS+ framework. Across all studies, 96% (57/59) considered at least one PROGRESS+ factor during development. The most common factors were age (95%; 56/59) and gender/sex (96%; 57/59). Differential effect analyses, such as subgroup analysis, covariate adjustment, and baseline comparison, were rarely reported (10%; 6/59). Most models included age (92%; 54/59) and gender/sex (69%; 41/59) as final input variables. However, factors such as insurance status (7%; 4/59), marital status (7%; 4/59) and income (3%; 2/59) were seldom included. They concluded that the current level of reporting and consideration of

SDOH during the development of prognostic ML models for orthopaedic outcomes is limited.

2.6.2 Risks with using machine learning tools

A conventional ML lifecycle involves defining the problem, defining the scope and sources of data collection process, collect and prepare the data based on identified features, segment into train and test set, train the model iteratively by tuning hyperparameters, deploy the models, test them, receive feedback, fine tune features or hyperparameters till desired outcomes are achieved.

In this conventional ML lifecycle, focus has been on Training models or use latest models to achieve the ML outcomes. We could refer to this process as model-centric AI. It focuses on the computational models that are used to represent and solve problems. In this approach, models themselves are treated as first-class citizens and the focus is on understanding how they work and how they can be improved. Model-centric AIs are often used for tasks that require high precision, such as medical diagnosis, or tasks that are too expensive to hand over to humans, such as stock trading. This approach has been used in many fields including computer vision, natural language processing, speech recognition and machine translation. It has also been applied to other areas such as marketing, finance and healthcare.

Challenges with Model-centric AI?

There are different challenges that can be observed with a model-centric AI approach.

• Data Diversity: Generally, data is collected from human effort, and this may not have enough diversity eg. When we collect images of vehicle, we need the algorithms to be robust enough to understand different lighting conditions in the image, its orientation, size, grains in the image etc. Since the models are trained in the lab and are expected to perform in the real world, the data collection must be rigorous to be able to simulate varying real-life situations. We want the models to work in changing backgrounds and scenarios. The necessary data may not exist eg. if we were to go collect enough data for training process in the above use case, we would need to simulate multiple crashes in different conditions.

• Data bias: Another one is bias in the data set. It's hard to get a diverse data set when you have limited resources and time, which means that AI will be biased towards what it has learned from this data set. For large models we need large amounts of labelled data to achieve a certain level of accuracy. For complex use cases like self-driving cars, the amount of data that is required, and the accuracy and diversity of it is just impossible to get from the real world without having to physically simulate to generate the data. We also would not want to put humans or animals in harm's way to simulate certain situations like vehicle crashes.

• Overfitting: Overfitting is also one of the pitfalls of traditional machine learning due to lack of diversity in training data. It is when a model performs well with the training data but not with test data. The problem arises when a machine learns to replicate patterns that are taking place in the training data, but those same patterns are not present in the real world. eg. If we train facial recognition models on certain sections of people, then it may fail for other people in real world.

• Data depth: Most of the human generated data is in 2D applications while real-life application requires 3D modelling. Hence there are many cases where computational models need to make approximations that itself introduces precision error when applied to real world scenarios.

• Algorithmic inequity: This can be compared with inequities that occur when clinical trial participants are not representative of the patient population that ultimately receives the treatment. Obermeyer, Nissan et al (2021) found evidence of racial bias negatively impacting Black patients in a widely used algorithm for allocating additional health resources to patients with complex health needs. If ML model development studies do not utilize techniques to ensure algorithmic equity, they are at risk of unintended negative consequences, such as the perpetuation of health inequities.

There are tools like Prediction model study Risk Of Bias Tool (PROBAST) available to assess the risk of bias of a study. However these are not frequently used.

2.6.3 Data-centric AI approaches to address data availability and quality issues

AI systems typically involve two main ingredients – code and data.

The code reflects AI model or algorithm which is trained using the data. The conventional model-centric AI focuses on improving code to achieve better results given a fixed set of data. AI developers generally consider the training datasets from which their code is learning as a collection of ground-truth labels, and their AI model is made to fit that labeled training data. Thus, this approach generally assumes the training data as external from the AI development process.

On the other hand, data-centric AI aims to improve data quality to achieve better outcomes by treating code as an unchangeable entity. In other words, while model-centric AI deals with developing or improving the AI model or algorithm, data-centric AI deals with the labelling, augmenting, managing and curating of data. Data-centric AI may seem to be the pre-processing of data, however, it emphasizes an iterative AI life-cycle consisting of data collection, model training and analysing errors.

In model-centric AI, we spend relatively more time on optimizing an AI model whereas in data-centric AI, we spend rather more time on data quality improvement. In model-centric, we aim to find the most suitable AI model or an optimization technique for a given problem, whereas in data-centric we aim to find inconsistencies in the collected data for a given problem.

Nowadays, model-centric AI tends to optimize bigger AI models on large-scale datasets which therefore require large-scale datasets and lots of computing resources, whereas data-centric AI may require domain knowledge or experts to find inconsistencies in data.

Though most data-centric AI ideas already exist as conventional wisdom in the AI community, data-centric AI aims to build a systematic approach and the tools needed to facilitate this process.

Case Study: Data Augmentation with Generative Models

A competition on Data-centric AI was organized, Data-Centric AI Competition Submission Guide, 2021 to invite innovative approaches to validate if improving data quality helped improve model outcomes

The award winning entry by Motamedi et al. (2021), used a controlled RestNet50 architecture on the use case of identifying handwritten roman numerals for the experiment. They used a baseline dataset and validate the identification accuracy as 64%. The authors created auxiliary models to identify misclassified labels in the dataset and developed a dataset optimization pipeline that included Duplicate Detection and Elimination, Training Auxiliary Models, Dataset Investigation, Class Imbalance Resolution, and N-fold Cross-Validation.

They validated the accuracy of the model post the pipeline correction to be 69% which was 5% better than the baseline. Then they collected addition samples from a random set of users and synthesized additional 20% noisy samples using a GAN network with an architecture analogous to that of the DCGAN. The performance of RestNet50 network improved to 83% with additional human samples, and further to 84% using the synthesized samples.

Table 7
Data-Augmentation Approach results

Name	Training dataset size	Validation dataset size	Test dataset size	Test accuracy
Baseline dataset	2067	813	2420	64%
Pipeline optimized dataset	1370	500	2420	69%
Additional samples collected	8229	500	2420	83%
Additional synthesized noisy samples	9455	500	2420	84%

In a conventional ML lifecycle, focus has been on Training models or use latest models to achieve the ML outcomes. However it is observed that in many uses cases, latest model gives same or similar outcomes. On the other hand, improving data quality in terms of data diversity, completeness and quality improves the outcomes significantly.

2.7 New Era of Healthcare Ecosystem

2.7.1 What is required to transform the current state?

US Government agencies, non-governmental organizations and private healthcare corporations are putting in tremendous effort to address Health Equity, provide universal healthcare while keeping its costs in check, and improving improve health outcomes. For example, optimizing Per member per month cost (PMPM) cost for various diseases to reduce the economic burden of healthcare on US economy, bringing in accountability in healthcare services delivery, improve patient satisfaction through technology intervention at every stage of disease progression and management.

2.7.2 Investment in Preventive Health Services

Annual funding is needed to transform nation's public health infrastructure. A stable source of risk-factor and disease-prevention funding can better equip country's federal, state, local, and territorial public health agencies to coordinate together and to save lives. Good news is that funding for the "Preventive Health and Health Services Block Grant" and a new program for Public Health Infrastructure has been approved by the US Federal government recently.

2.7.3 Connected care for Whole person health

Connected care for whole person health is an approach to healthcare that focuses on treating the patient as a whole person, rather than just treating their individual symptoms or conditions. It involves integrating medical, behavioral, and social care to address the full range of a patient's health needs. These conditions cost us in quality of life, productivity and dollars.

• 90% of the US healthcare cost is spent for people with chronic and behavioral health conditions.

• Behavioral health affects 1 in 5 or 66 million Americans, causing significant and avoidable disability and death.

• Gallo (2020) estimates that as many as 26 million people do not have access to the behavioral health resources and treatment they need. This limits their ability to live whole and productive lives. And even where it is available, the pressures of home, food and personal insecurity too often push health concerns into the background.

The state of connected care for whole person health is rapidly evolving, driven by advances in technology, changes in healthcare delivery models, and increasing recognition of the importance of addressing social determinants of health. Some of the key trends in connected care for whole person health include:

• Care coordination: Connected care for whole person health requires coordination among multiple providers and care settings. Health information technology (HIT) systems can help facilitate this coordination, allowing providers to share information and work together more effectively.

• Population health management: Connected care for whole person health also involves managing the health of entire populations, not just individual patients. Population health management involves using data and analytics to identify and address health risks and disparities in specific patient populations.

• Social determinants of health: Addressing social determinants of health, such as poverty, housing insecurity, and food insecurity, is essential to improving overall health outcomes. Connected care for whole person health involves collaborating with community organizations and other stakeholders to address these social factors.

Overall, the state of connected care for whole person health is moving towards a more integrated, patient-centred, and data-driven approach to healthcare delivery. While there are still challenges to overcome, such as interoperability and data privacy concerns, the potential benefits of connected care for whole person health are significant.

2.7.4 Physician Empowerment is the key to improving health outcomes

EMR systems are referred by Physicians in clinical setting. They provide information related to current episode and to an extent, the medical history of a patient. Payer systems have a more complete data about the patient, family, neighborhood, and similar cohorts. Owing to the access to data, many payers have developed analytical and predictive systems for risk modelling. However, in current scenario, Provider's EMR systems are not very well integrated with Payer systems hence they are missing important insights and tools for clinical decision support.

As per a survey conducted by a Fortune 5 healthcare company, only about 51% of physicians are aware of the cost of treatments they select. Just 48% of physicians are comfortable discussing costs with their patients. Due to this information gap, Providers are generally hesitant to adopt value-based-care plans from Payers. This also leads to higher out-of-pocket expenses for patients. There is a need to bring transparency in terms of cost of care to both physicians and patients.

2.7.5 Consumer engagement (Activation) can reduce cost of healthcare

As per the HCPLAN Measurement Report (2021), Consumers are not yet fully engaged in payment reform efforts, with only 23% of respondents reporting that they had received information on payment reform from their healthcare providers.

As per a survey by a Fortune 5 Healthcare company, activated patients cost $1,987 less per patient annually which is a 31% difference in cost.

There is a need to involve patients and their care givers actively in disease management. The systems should be simplified and made easy to use by educated and less educated alike so that people can themselves understand their whole health and take timely action by either visiting screening centers, take preventive check-ups or ask for help.

2.7.6 Accelerate the migration of Care from Hospital settings to Community arrangements

US healthcare cost is increasing at a rate of 5-6% and this increase in unsustainable while the economy is growing at a rate of 2-3% per year. Even more, the increasing cost is not able to significantly improve health outcomes. There is a need to engage community arrangements more for preventive and treatment disease continuum. Community arrangements are setting that are tailored to the needs of people living in the area and adapted to environmental, cultural, and socio-economic context. People find it more comfortable and convenient reaching out to such facilities.

Some of the community arrangements are not registered in payer-networks eg As per Alzheimer’s Association report (2021), there is common knowledge that that alternative therapies like yoga clinic or music-based therapies are helpful in treatment for dementia and can help in managing such conditions better and at an optimal cost. However, these may not be covered well by payer networks raising out-of-pocket expenses for patients. While there is effort to bring such therapies under wellness clauses in insurance contracts but there is a greater and large-scale need to include alternative medicine and therapies in payer networks.

2.7.7 Innovative financial models are required to incentivize Providers taking on health outcome risks

Kshirsagar et al (2022) did a study related to VBC arrangements related to chronic kidney disease (CKD) care in population health management commercial market. They opined that even the payments introduced to improve equity (e.g., the Health Equity Incentive in the ESRD Treatment Choices) may not be enough to sway the companies to address adverse SDOH.

As per Dr Beale (2021), providers are at the centre of value-based care. Majority of them prefer to operate in fee-for-service low risk model. There is a need to introduce volume-based incentives in provider-payer agreements without the downside risks for providers. Payers will have to innovate on the current financial models to be able to influence transition of risk from payers to providers.

2.8 Summary

In the examination of existing literature, the researcher delves into the shift of the U.S. government and healthcare agencies towards a value-based care model, with a particular emphasis on the crucial role of providers in this transition. However, there are persistent challenges that arise primarily from a lack of trust, caused by the absence of harmonization among stakeholders and the presence of Social Determinants of Health (SDOH) data. The author points out that achieving the objectives of value-based care necessitates cooperation among members, providers, and payers, with a focus on open and honest communication, appropriate pricing of healthcare plans, and an impartial assessment of claims. The review highlights the importance of community information from various sources but also underscores the need for qualitative health data to be converted into quantitative data for analysis. To address the trust deficit, the author proposes the establishment of a mechanism for building trust through third-party validation using community data.

Moreover, the literature review emphasizes the need for an approach to the adoption of Social Determinants of Health (SDOH) that places humans at the center. Current models often lack personalization and fail to consider critical factors that are essential for the development of personalized healthcare programs and risk modeling. The author argues for a disease-focused classification of SDOH that takes into account the experiences of clinicians and consumers, with the aim of enhancing the usefulness of these models in clinical settings and disease management. The review also acknowledges challenges in the value chain of SDOH data, including the complexity of Z-codes and the need for a simpler classification method.

Furthermore, the literature review identifies a dearth of reliable tools for connected care among both patients and physicians. The concept of connected care involves integrating various aspects that individuals can control or influence, necessitating the availability of adequate tools and insights to identify not only physical health issues but also behavioral and mental health concerns. The review concludes by raising the question of whether innovative technological initiatives can address the challenges associated with the adoption of value-based care. Themes such as prognostic risk modeling, data fragmentation, consumer activation, physician empowerment, and financial modeling tools are highlighted as potential facilitators for the adoption of value-based care, with a particular focus on improving mechanisms for capturing data and leveraging technology to provide personalized insights.

CHAPTER III:
METHODOLOGY

3.1 Overview of the Research Problem

Some studies have shown that SDOH impact a person's wellness by upto 80% (ref). The importance of getting a better understanding of the social determinants is increasing due to the fact that the environment where people live is changing, lifestyle changes are making people lives sedentary and thereby healthcare costs are increasing. There have been many models and prediction methods on how such changes affect a person's health but none of them have seen a mass adoption. There have been data gaps in such models. Most of the social determinants data is either unavailable or not of good quality for usage in prediction models, or not accessible.

Organizations like WHO, CDC, Robert Woods Foundation etc have come up with social determinants models. There have been assumptions that a particular social determinants model would have similar impact for every disease and during different stages of disease continuum eg, Air quality would have a same impact for hypertension and for cancer, and also during early stages of a disease (onset) versus mid stages and during advanced stages. The universality assumption needs to be assessed and if these are helping physicians use them in clinical settings which is the ultimate proof of their utility.

One of the objectives of Value based care is to make Providers more accountable for health outcomes and thereby sharing the risk with Payers. In order to assess if SDOH models and assumptions contributed to ease for their utility in value based care, it is important to understand if such models are providing more confidence to physicians to assume higher risk.

3.2 Research Purpose and Questions

To understand how the U.S. healthcare system is changing to focus on providing better care for people. We want to find out how information about different aspects of people's lives, like where they live or their financial situation, affects their health and the care they receive. We'll look into the challenges that come with these changes, such as trust issues among different groups involved in healthcare. Our goal is to figure out how to make sure everyone—patients, healthcare providers, and insurance companies—works together smoothly for the benefit of everyone's health.

Additionally, we want to create a way to measure the impact of these changes on managing long-term health conditions, like Alzheimer's or dementia. We believe that by understanding more about the factors that contribute to these conditions, we can come up with better ways to prevent and manage them, leading to healthier outcomes for individuals.

In our investigation, we'll explore whether the tools and information currently available to doctors and patients are enough to identify not just physical health issues but also behavioral and mental health concerns. We want to make sure that everyone involved has the right tools and information to take care of all aspects of a person's health.

The proposed research aim to answer the following research questions

1. How are industries contributing to the adoption of SDOH for the acceleration of Value-Based Care?

2. What indicators within the framework are most crucial for determining the prognosis of chronic diseases?

3. What role do individual behaviors and community factors play in the relationship between SDOH and Dementia and Alzheimer's prognosis?

4. Are there specific demographic or environmental factors that exacerbate the impact of SDOH more than others on Dementia and Alzheimer's prognosis?

3.3 Research Objectives

1. Explore and develop a better understanding of how SDOH plays a role in people's health and the governmental and industry efforts in getting it adopted for accelerating transition of healthcare models to Value Based Care.

2. Assess the data challenges plaguing the collection, usage and adoption of SDOH factors.

3. Review the utility of technology interventions through sample validations.

4. Evaluate the prospect development of a holistic SDOH framework with a view to support Whole health practices.

3.4 Research Design

The proposed research design involves a comprehensive approach to investigate the changing landscape of the U.S. healthcare system, specifically focusing on the transition to a value-based care (VBC) model and the role of Social Determinants of Health (SDOH). To address the first objective, we plan to conduct a detailed literature review to analyze existing government initiatives and industry efforts aimed at implementing VBC. This will involve reviewing regulations, educational strategies, and the centralization of healthcare providers.

For the second objective, the research design includes the development of a framework for assessing the impact of SDOH on chronic diseases. This will involve a combination of qualitative and quantitative methods, including data analysis and synthesis to identify key SDOH markers and their relationship to chronic disease progression.

The third and fourth objectives focus on investigating the quantified impact of SDOH on dementia and Alzheimer’s, respectively. These will involve regression analyses to assess the statistical relationship between specific SDOH markers and the prevalence or progression of these health conditions.

To address the need for a human-centric design approach in SDOH adoption efforts (Objective 3.3), the research design will integrate qualitative methods such as interviews and surveys to capture the perspectives of stakeholders, including academicians, clinicians, and individuals affected by healthcare policies.

The challenges in the value chain of SDOH data (Objective 3.4) will be addressed through a combination of content analysis and classification methods. This will include reviewing existing Z-codes, evaluating their practicality, and proposing a simpler classification method.

For the fifth objective, which examines the lack of reliable tools for connected care, the research design will involve surveys and interviews with physicians and patients to assess the current state of tools and insights available for identifying both physical and mental health issues.

Finally, to investigate whether innovative technology initiatives are a way forward for VBC adoption (Objective 3.6), the research design includes a thematic analysis of ideas generated through events and competitions within a prominent healthcare firm, focusing on key innovation enablers such as prognostic risk modeling and fragmented data solutions.

This comprehensive research design aims to provide a thorough understanding of the complexities surrounding the adoption of VBC and the integration of SDOH in the

U.S. healthcare system. The combination of literature reviews, quantitative analyses, qualitative methods, and thematic analyses will contribute to a holistic examination of the research objectives.

3.4.1 Data Collection

The study will draw SDOH data from primarily the open database provided by Agency for Heathcare Research (AHRQ). This agency aggregates data from multiple sources, including US Census Tracts https://www.census.gov/data/developers/data-sets.html, Suite of Food Security Indicators - Datasets - "FAO catalog", Housing Density index https://earthexplorer.usgs.gov/, World Health Organization Air quality database 2022 (who.int), Centers for Disease and Control database, USGS Water-Quality Database, and Geospatial database https://www.arcgis.com/home/item.html?id=ab9400829f38405d9d2299ddeb3bb65d.

The FIPS county codes and US postal service Zipcodes do not have a direct mapping as one zipcode may span multiple counties in some cases. Hence the study will depend on a best case mapping database provided by open sources on Kaggle.

Primary data for Patient Health Records will be collected from around 1 to 2 mid-size and bigger payers, providers and pharmacies. This data is usually containing Personally identifiable information (PII) which is protected by Health Insurance Portability and Accountability Act of 1996 (HIPAA), Public Law 104-191. Hence the PII data will be removed or obfuscated prior to doing the analysis and experiments as per the guidelines of the law for compliance.

3.4.2 Explore SDOH's role in Value-Based Care

Exploring the Social Determinants of Health (SDOH) is extremely important in the concept of Value-Based Care (VBC), as it helps us gain a deeper understanding of the complex factors that influence health outcomes. SDOH covers a wide range of determinants, which are categorized into areas such as social and economic conditions, physical environment, social relationships, and personal behaviors. The World Health Organization (WHO) emphasizes the significance of SDOH and their impact on overall health and health disparities.

Research consistently shows that SDOH can have a greater impact on health outcomes than healthcare and lifestyle choices, accounting for 30-55% of the results. The socio-economic gradient is a significant aspect, as lower socio-economic positions are associated with poorer health. The dissertation explores various SDOH models proposed by reputable organizations like WHO, using factors such as race/ethnicity, education, employment, and psychological aspects to fully comprehend their influence.

Within the context of VBC, the dissertation examines the provisions of the Patient Protection and Affordable Care Act (2010), highlighting the crucial role of addressing SDOH in order to reduce healthcare costs, improve the quality of care, and enhance population health. The proposed research design takes a multifaceted approach to investigate the integration of SDOH in VBC adoption. It utilizes both qualitative and quantitative methods to assess the impact of SDOH on chronic diseases like dementia and Alzheimer's, utilizing regression analyses for in-depth exploration.

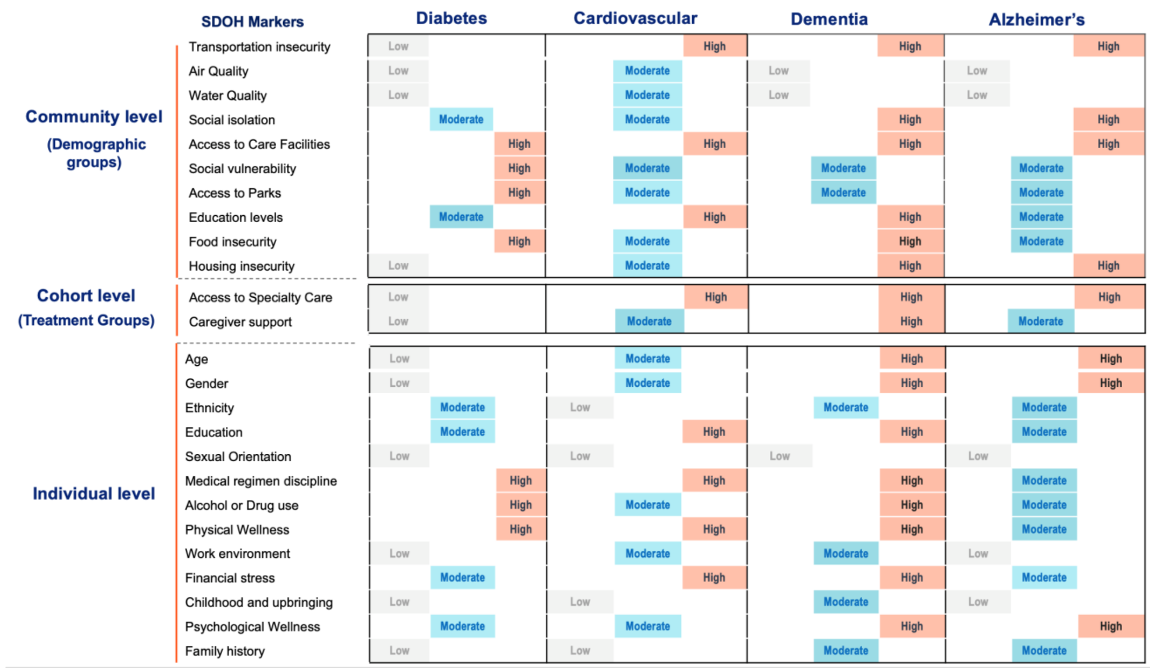
The study also expands its focus to innovative technology initiatives, exploring whether they offer viable solutions for the adoption of VBC. The research design acknowledges the challenges in the value chain of SDOH data, aiming to streamline classification methods and improve the reliability of tools for connected care.

Additionally, it recognizes the human-centric aspect of SDOH adoption, capturing the perspectives of stakeholders through qualitative methods.

By drawing insights from reputable models like PROGRESS+, Healthy People 2030, DNPAO, FACETS, PLACES, and HealthLandscape, the dissertation positions itself at the intersection of academia and practical application. It contextualizes SDOH in real-world scenarios, demonstrated through a comprehensive case study on COVID-19. The case study highlights the intricate connections between SDOH and COVID-19 infection and mortality rates, emphasizing the need for a holistic approach to healthcare that addresses social determinants.

In conclusion, the dissertation aims to uncover the complex relationship between SDOH and VBC, providing valuable insights for policymakers, healthcare professionals, and researchers. By bridging the gap between theoretical frameworks and practical implications, the study aims to make a meaningful contribution to the ongoing discussion on enhancing the effectiveness and inclusivity of healthcare systems through a thorough understanding of social determinants.

3.4.3 Framework To Determine Disease Prognosis Using SDOH Markers



Level	SDOH Markers	Diabetes	Cardiovascular	Dementia	Alzheimer's
Community level (Demographic groups)	Transportation insecurity	Low	High	High	High
	Air Quality	Low	Moderate	Low	Low
	Water Quality	Low	Moderate	Low	Low
	Social isolation	Moderate	Moderate	High	High
	Access to Care Facilities	High	High	High	High
	Social vulnerability	High	Moderate	Moderate	Moderate
	Access to Parks	High	Moderate	Moderate	Moderate
	Education levels	Moderate	High	High	Moderate
	Food insecurity	High	Moderate	High	Moderate
	Housing insecurity	Low	Moderate	High	High
Cohort level (Treatment Groups)	Access to Specialty Care	Low	High	High	High
	Caregiver support	Low	Moderate	High	Moderate
Individual level	Age	Low	Moderate	High	High
	Gender	Low	Moderate	High	High
	Ethnicity	Moderate	Low	Moderate	Moderate
	Education	Moderate	High	High	Moderate
	Sexual Orientation	Low	Low	Low	Low
	Medical regimen discipline	High	High	High	Moderate
	Alcohol or Drug use	High	Moderate	High	Moderate
	Physical Wellness	High	High	High	Moderate
	Work environment	Low	Moderate	Moderate	Low
	Financial stress	Moderate	High	High	Moderate
	Childhood and upbringing	Low	Low	Moderate	Low
	Psychological Wellness	Moderate	Moderate	High	High
	Family history	Low	Low	Moderate	Moderate

Figure 4 SDOH Framework

In figure 4, the SDOH framework is depicted, which includes three categories of SDOH markers: community level, cohort level, and individual level. These categories encompass various factors comprehensively, ensuring that health measures become more precise and accurate. Diseases are classified into different levels, namely low, moderate, and high.

The factors that impact an individual's health, known as SDOH, can be divided into four categories: social and economic conditions, physical environment, social relationships, and personal behaviors. Studies conducted by WHO demonstrate that these determinants have a significant influence on people's health and how they interact with each other to affect overall well-being. SDOH play a crucial role in health inequities, which are avoidable and unfair differences in health status observed within and between countries. Regardless of income level, health and illness follow a social gradient, with worse health outcomes associated with lower socioeconomic positions. Research indicates that social determinants are often more influential than healthcare or lifestyle choices when it comes to influencing health. Numerous studies suggest that SDOH account for 30-55% of health outcomes. Additionally, estimates show that factors outside the health sector make a greater contribution to population health outcomes than the health sector itself. These findings highlight the significant impact that non-medical factors have on people's health and how different determinants interact to shape overall well-being. The Patient Protection and Affordable Care Act (2010) includes provisions that recognize the overwhelming evidence supporting the need to address SDOH in order to reduce healthcare costs and improve quality of care and population health. These policies involve funding partnerships between public health agencies, community organizations, and healthcare institutions, as well as promoting value-based payment models that encourage integrated health and social care delivery. Furthermore, the act supports Medicaid program innovations that directly address social needs as part of healthcare.

Imagine we are embarking on a journey to construct a comprehensive roadmap that unravels the intricacies of disease development in individuals. Instead of solely focusing on medical aspects, we are delving deep into a myriad of factors that extend far beyond the confines of a doctor's office. These factors, known as Social Determinants of Health (SDOH) Markers as shown in the figure 4, encompass various facets of people's lives that possess the potential to significantly impact their overall well-being.

Individual Details: Let us commence this voyage by acquainting ourselves with each unique individual. How do we discern their age? What is their gender? Where do they originate from? To what extent have they pursued education? Essentially, we are immersing ourselves in the tapestry of their life story.

Lifestyle and Habits: Next, we delve into their daily routines and habits. Do they diligently adhere to their healthcare provider's guidance? Are they grappling with the excessive consumption of alcohol or illicit substances? How is their physical health? We meticulously assess whether they are actively nurturing their physical well-being.

Work and Environment: Subsequently, we contemplate the environment in which they spend a substantial portion of their day - their workplace. What kind of occupation do they engage in? Are they exposed to any detrimental elements? Scrutinizing their work environment enables us to ascertain if it fosters their overall health.

Financial Situation: Finances also play a pivotal role. Can they afford nutritious sustenance and adequate clothing? We meticulously calculate the proportion of their income allocated to essential living expenses. Recognizing that financial strain can exert an impact on health, we vigilantly monitor this aspect.

Early Years and Family Background: We transport ourselves back in time to their formative years - their childhood. What was their upbringing like? Did they grow up in a nurturing environment? We delve into the origins of their health-related habits.

Mind and Emotions: How resilient is their mental state? We evaluate their stress levels, memory retention, and decision-making abilities. Mental well-being constitutes an integral component of this intricate puzzle.

Family Tree Health: Lastly, we delve into the annals of their family history. What hereditary health conditions have been passed down through generations? Familiarizing ourselves with this knowledge enables us to comprehend if any genetic factors are at play.

Having assimilated all this information, we proceed to create a holistic panorama. By meticulously examining these Social Determinants of Health, our focus transcends mere medical symptoms. Instead, we endeavor to comprehend the entirety of their existence - where they reside, how they navigate their professional lives, their dietary choices, and much more.

This comprehensive overview empowers us to prognosticate the potential trajectory of diseases. For instance, if an individual inhabits an area with substandard air quality, it is likely to have a cumulative impact on their respiratory health over time.

Thus, the "Framework To Determine Disease Progression Using SDOH Markers" serves as a veritable detective tool. It allows us to perceive the complete narrative of an individual's life and harness that knowledge to anticipate and circumvent health issues. It is a testament to our commitment to treating the whole person, rather than merely addressing the symptoms they present at the doctor's office.

The Table 8 give the description of the individual markers involves in the SDOH framework proposed in the particular objective.

Table 8

Description of Maker involves in SDOH framework

S No.	SDOH Marker	Description	Metric
1	Transport Insecurity	Transport Insecurity SDOH marker is the ability of citizens to remain mobile using public transport so as to visit healthcare facilities, parks, nutrition, food, etc and remain social. It can be indirectly measured through Road Network Density & Transportation Nodes Service Area in the particular zip code.	ACS_PCT_WALK_2WORK_ZC (Percentage of workers walking to work (ages 16 and over))
2	Air Quality	Air Quality SDOH marker is the Annual average of Air Quality Index (AQI) levels of fine particulate	WUSTL_AVG_PM25 (Annual mean of Particulate Matter (PM2.5)

		matter in the particular zip code ie PM10, PM2.5 and NO2	concentration (µg/m3))
3	Water quality	Water Quality SDOH marker is the Annual average of Water Quality index (WQI) in the particular zip code ie Physical and Chemical characteristics of the water including pH, specific conductance, temperature, dissolved oxygen, and percent dissolved-oxygen saturation.	NEPHTN_PCT_ARSENIC_MCL_GREATER10 (Percentage of population in the county served by community water systems with yearly distribution of mean arsenic concentration >10)
4	Social Isolation	Social isolation SDOH marker is the Housing density in the particular zip code is an indirect measure of individual's social isolation - Measured through NDBI (Normalized Difference Built-up Index)	CEN_POPDENSITY_COUNTY (Population density (County))
5	Access to Care Facilities	Access to Care Facilities SDOH marker represents access to the Health Department Service Centers, including Tuberculosis Chest Centers, STD Services, STD/HIV Testing, Food Safety & Community Sanitation, Birth and Death Certificates, Early Intervention Offices, The Health Academy, Therapy Centers and Immunization Walk-In Centers.	POS_DIST_CLINIC_ZP (Distance in miles to the nearest health clinic (FQHC, RHC), calculated using population weighted ZIP centroids)
6	Social	Social Vulnerability SDOH marker	ACS_GINI_INDEX_ZC

	Vulnerability	is the Aggregated location health surveys conducted by community organizations like United Hospital Fund neighbourhoods. The health topics cover a number of areas including physical activity, diabetes, obesity, mental health, and sexual risk factors. It is measured through Social Vulnerability Index (SVI) in the particular zip code	(Gini index of income inequality (ZCTA level))
7	Access to Parks	Access to parks SDOH marker is the Proximity to Nearest Park in the particular zip code. It is Measured through Normalized Difference Vegetation index (NDVI)	Normalized Difference Vegetation index (NDVI)
8	Education Level	Education Level SDOH marker is the Average Literacy rate in the particular zip code. Individual literacy will be captured in Individual factors	ACS_PCT_COLLEGE_ASSOCIATE_DGR_ZC (Percentage of population with some college or associate's degree (ages 25 and over))
9	Food Insecurity	Food Insecurity SDOH marker is the threshold to classify 'severe' food insecurity corresponds to the severity associated with the item 'having not eaten for an entire day' on the global FIES scale. It is an indicator of lack of food access in the particular zip code.	ACS_PCT_HH_NO_FD_STMP_BLW_POV_ZC (Percentage of households not receiving food stamps/SNAP with income below the poverty level)

10	Financial Insecurity	Financial Insecurity SDOH marker is the lack of sufficient resources to afford basic life needs of family as prescribed by Federal Poverty Line (FPL).	ACS_PCT_PERSON_INC_BELOW99_ZC (Percentage of population with an income to poverty ratio under 1.00)
11	Housing Insecurity	Housing Insecurity SDOH marker is the lack of security in the particular zip code that is the result of high housing costs relative to income, poor housing quality, unstable neighbourhoods, overcrowding, uncomfortable conditions and homelessness. This is measured through Percentage of Total housing insecure households by American Housing Survey conducted by US Census Bureau	ACS_PCT_RENTER_HU_COST_30PCT_ZC (Percentage of renter-occupied housing units with rent equal to 30 percent or more of household income)

In the relentless pursuit to comprehend and tackle disparities in health, an all-encompassing framework has emerged, intricately weaving together a tapestry of Social Determinants of Health (SDOH) markers. These markers, each resembling a distinct thread, harmoniously intertwine to expose the landscape of community well-being. From the vibrant streets of the city to the air we inhale, this framework encompasses factors such as Transport Insecurity, meticulously scrutinizing the mobility of citizens through metrics like Walkability to work or to community facilities. It delves deep into the very air quality we breathe, with markers like PM2.5 and Air Quality Index (AQI) sourced from the WHO Ambient Air Quality Database.

The journey persists through the flow of water, thoroughly examining its drinking quality through the Arsenic presence in drinking water in the area, a measure of the life-sustaining blood of communities. Social isolation, a subtle yet profound determinant, is unveiled through the lens of housing density using the national database published by Federal Agency AHRQ. Access to essential care facilities becomes the focal point, meticulously analyzed through a walkability lens and spatial analysis of service centers using the national database published by Federal Agency AHRQ.

Advancing beyond physical landscapes, this framework takes into account the pulse of social vulnerability, aggregated from health surveys conducted by community organizations, expressed through the GINI index which is an accepted method to measure disparity in incomes in a particular area. It navigates the lush expanses of parks, assessing accessibility through the Normalized Difference Vegetation Index (NDVI). Literacy rates and educational landscapes are revealed using Census data, shedding light on the intellectual terrain.

Venturing deeper into individual factors, this framework encapsulates age, gender, ethnicity, education, sexual orientation, and various lifestyle markers. It delves into personal habits, medical discipline, substance use, physical well-being, work environments, financial stress, childhood experiences, psychological welfare, and family history. Each element paints a distinct stroke on the canvas of an individual's health journey.

The collective factors, such as Access to Specialty Care and Caregiver Support, emerge as guiding beacons in this constellation of health determinants. Access to advanced medical services is dissected through walkability analyses, while the availability of immediate support personnel becomes a pivotal consideration for patients.

In essence, this framework is a compelling narrative, a tale told through data and markers that transcend individual and community well-being. It empowers healthcare practitioners, policymakers, and communities alike to navigate the intricate web of social determinants, fostering a comprehensive approach to health that addresses not only the symptoms but the very essence of our lives.

3.4.4 To Investigate the Impact of SDOH Markers on Dementia

Dementia & Alzheimer's are similar but not the same even though the physical ailments and the conditions may seem similar, especially memory loss. Differences Between Dementia & Alzheimer’s are as below:

Dementia is an overall term that describes a wide range of symptoms associated with a decline in memory. It can also effect other thinking skills enough to reduce a person’s ability to perform everyday activities. Dementia is often incorrectly referred to as “senility” due to the widespread incorrect belief that serious mental decline is a normal part of aging.

Senile Dementia of the Alzheimer’s Type (SDAT), or simply called Alzheimer’s is a disease. It produces physical change in the brain. There is shrinking in some areas of the brain and widening in the others. This causes connections inside the brain to break and disrupt the brain’s electrical signals. Alzheimer’s disease accounts for 50 to 80 percent of the dementia cases and is the most type of dementia. Vascular dementia, which occurs after a stroke, is the second most common type. There are many other conditions such as thyroid problems and vitamin deficiencies that can cause symptoms of dementia. Some of these are reversible.

Table 9
Dementia diseases vs Alzheimer’s disease

	Dementia	Alzheimer’s

General Defintion	A brain related disorder caused by diseases and other conditions.	A type of Dementia but the most common type
Cause	Many, including Alzheimer's disease, stroke, thyroid issues, vitamin deficiencies, reactions to medicines, and brain tumors.	Unknown, but the "amyloid cascade hypothesis" is the most widely discussed and researched hypothesis today.
Duration	Permanent damage that comes in stages.	Average of 8 to 20 years.
Typical Age of Onset	65 years and older.	65 years but can occur as early as 30.
Symptomps	Issues with memory, focus and attention, visual perception, reasoning, judgement, and comprehension.	Difficulty remembering newly learned information. With advancement, disorientation, mood and behavior changes may occur.

SDOH mapping to the attributes in the AHRQ data layout taken table 10 provided a table with various Social Determinants of Health (SDOH) markers categorized into community level, cohort level, and individual level. Each marker is associated with specific attributes, descriptions, and data sources.

Table 10
SDOH data Mappings

S No.	SDOH Marker	Category	Attribute mapping in the sample data	AHRQ attribute description	Data source
1	Transportation insecurity	Community level SDOH marker	ACS_PCT_WALK_2WORK_ZC	Percentage of workers walking to work (ages 16 and over)	AHRQ Zip code file
2	Air Quality	Community level SDOH marker	WUSTL_AVG_PM25	Annual mean of Particulate Matter (PM2.5) concentration (µg/m3)	AHRQ Zip code file
3	Water Quality	Community level SDOH marker	NEPHTN_PCT_ARSENIC_MCL_GREATER10	Percentage of population in the county served by community water systems with yearly distribution of mean arsenic concentration >10	AHRQ County file
4	Social isolation	Community level SDOH marker	CEN_POPDENSITY_COUNTY	Population density (County)	AHRQ County file
5	Access to Care Facilities	Community level SDOH marker	POS_DIST_CLINIC_ZP	Distance in miles to the nearest health clinic (FQHC, RHC), calculated using population weighted ZIP centroids	AHRQ Zip code file
6	Social vulnerability	Community level SDOH marker	ACS_GINI_INDEX_ZC	Gini index of income inequality (ZCTA level)	AHRQ Zip code file
7	Access to Parks	Community level SDOH marker	Not available	Median distance to the nearest public park in the area	Not available
8	Education levels	Community level SDOH marker	ACS_PCT_COLLEGE_ASSOCIATE_DGR_ZC	Percentage of population with some college or associate's degree (ages 25 and over)	AHRQ Zip code file
9	Food insecurity	Community level SDOH marker	ACS_PCT_HH_NO_FD_STMP_BLW_POV_ZC	Percentage of households not receiving food stamps/SNAP with income below the poverty level	AHRQ Zip code file
10	Housing insecurity	Community level SDOH marker	ACS_PCT_RENTER_HU_COST_30PCT_ZC	Percentage of renter-occupied housing units with rent equal to 30 percent or more of household income	AHRQ Zip code file
11	Financial Insecurity	Community level SDOH marker	ACS_PCT_PERSON_INC_BELOW99_ZC	Percentage of population with an income to poverty ratio of under 1.00	AHRQ Zip code file
12	Access to Specialty Care	Cohort level SDOH Marker	AMFAR_MHFAC_RATE	Total number of community mental health care providers per 10,000 population	AHRQ County file

13	Caregiver support	Cohort level SDOH Marker	Not available	Average number of hours in a day that a caregiver is available to support the patient for performing daily activities	Not available
14	Age	Individual level SDOH marker	Age	Age of the patient as on Oct 2023	EHR file
15	Gender	Individual level SDOH marker	Gender	Gender of the patient as self-disclosed	EHR file
16	Ethnicity	Individual level SDOH marker	Not available	Race or Ethnicity of the patient	Not available
17	Education	Individual level SDOH marker	Not available	Highest education level attained by the patient	Not available
18	Sexual Orientation	Individual level SDOH marker	Not available	Sexual preferences of the patient	Not available
19	Medical regimen discipline	Individual level SDOH marker	Not available	The percentage of times in a month that patient has been able to comply with prescribed medication or therapy. It is manually determined by a practitioner during screening	Not available
20	Alcohol or Drug use	Individual level SDOH marker	Not available	The average number of times in a month that patient has consumed alcohol above prescribed safe limits or taken non-prescribed medication/drugs. It is manually determined by a practitioner during screening	Not available
21	Physical Wellness	Individual level SDOH marker	Not available	Measured using Body Mass Index (BMI)	Not available
22	Work environment	Individual level SDOH marker	Not available	Stress or toxic environment at work place. It is manually determined by a practitioner during screening	Not available
23	Financial stress	Individual level SDOH marker	Not available	Percentage of household income spent on cost of living expenses.	Not available
24	Childhood and upbringing	Individual level SDOH marker	Not available	Abusive or toxic environment at home during the patient growing age of 0-18 years. It is manually determined by a practitioner during screening	Not available
25	Psychological Wellness	Individual level SDOH marker	Not available	Psychological stress levels of the patient. It is manually determined by a practitioner during screening	Not available
26	Family history	Individual level SDOH marker	Not available	Record of family history of ailments. It is manually determined by a practitioner during screening	Not available

3.4.4.1 Agency for Healthcare Research and Quality (AHRQ)

The Agency for Healthcare Research and Quality (AHRQ) is the lead Federal agency with the responsibility of improving the safety and quality of healthcare for all Americans. AHRQ develops the knowledge, tools, and data needed to improve the healthcare system and help consumers, healthcare professionals, and policymakers to make informed health decisions.

AHRQ is working to tackle some of the health care system's greatest challenges, including Improving care for people with multiple chronic conditions and incorporating the latest research findings into electronic health records to facilitate clinical decision making. AHRQ Quality Indicators (AHRQ QIs) are standardized, evidence-based measures of healthcare quality that can be used with readily available hospital inpatient administrative data to measure and track clinical performance and outcomes. AHRQ QIs provide healthcare decision makers, such as program managers, researchers, and others at the Federal, State, and local levels, with tools to assess their data, highlight potential quality concerns, identify areas for further study and investigation, and track changes over time.

Area-level QIs can be used as a "screening tool" to help flag potential healthcare access problems or concerns about population health. They can also help public health agencies, State data organizations, healthcare systems, and others interested in improving healthcare quality in their communities to identify and investigate communities that may need interventions. AHRQ QIs are available via free software distributed by AHRQ.

The AHRQ SDOH Database is an Environmental scan of public social determinants of health and provides a one-stop source for data to analyze characteristics of communities across the United States across multiple domains. The purpose of SDOH Database are:

- Make community-level SDOH data easier to use in analyses to inform decisions to improve health outcomes.
- Account for health differences across areas and identify effective interventions tailored to populations served.

- Support Direct analysis using the database but can also be linked with other data sources to conduct more detailed analyses.

AHRQ Database spans multiple years and three geographic levels – County level, ZIP Code level and Tract level. It draws from 44 different data sources, including over 17,000 variables across all geographic levels and years shown in figure 5.

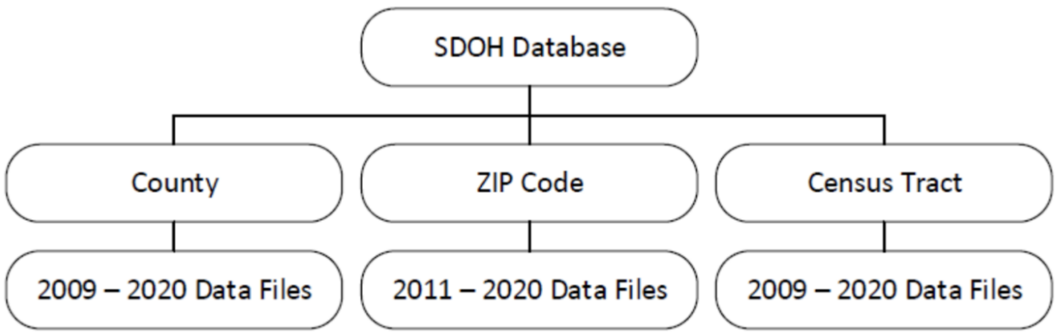


Figure 5 SDOH Databases geographic levels and years

AHRQ SDOH database provides Community-Level SDOH Variables Organized by Domains and Topics. As in below figure 6.

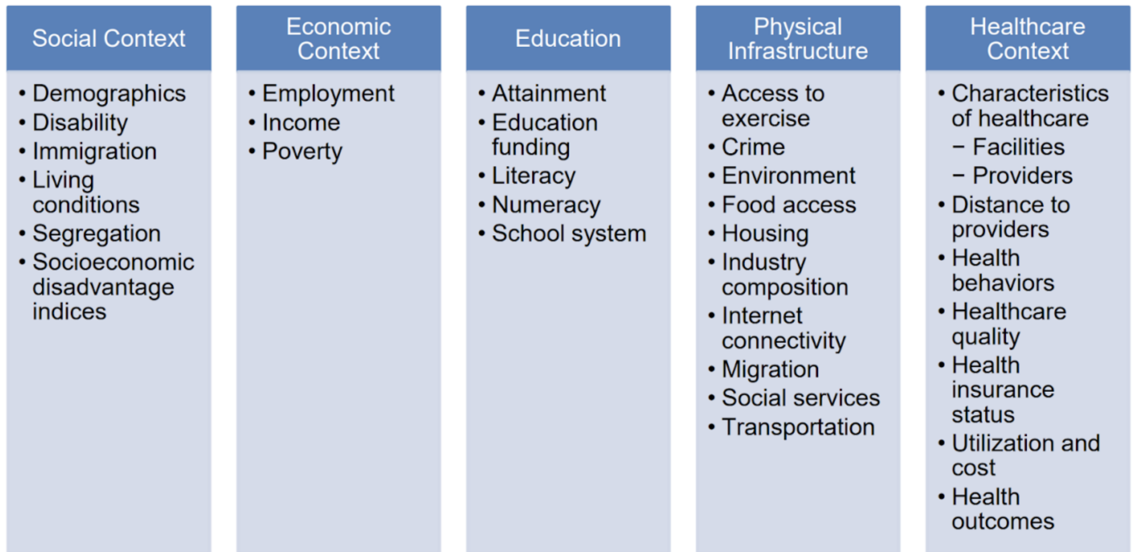


Figure 6 Community-Level SDOH Variables Organized by Domains and Topics

In Next is the distribution of variables in the SDOH Database by Domain and Geographic Levels in figure 7.

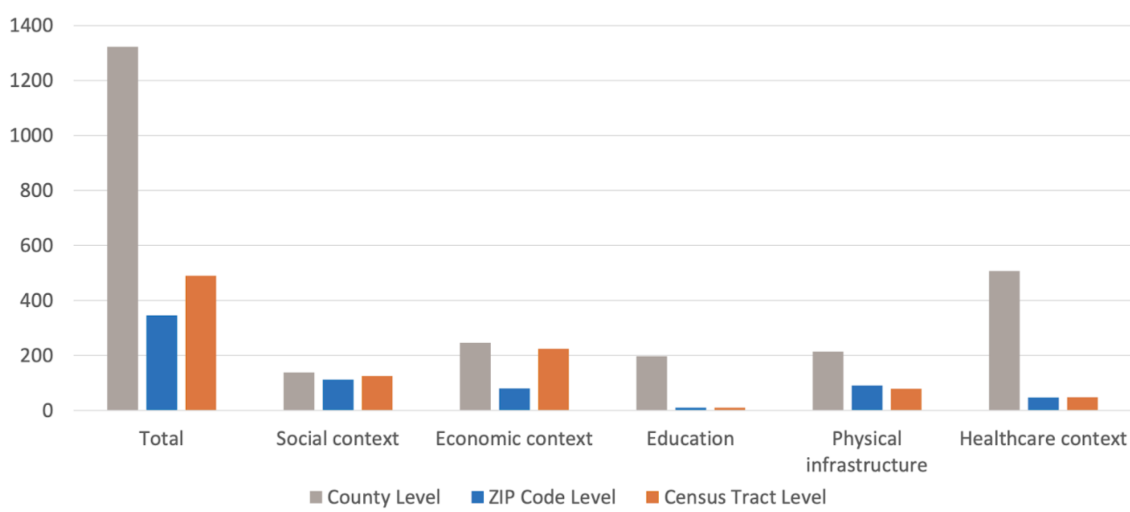


Figure 7 Distribution of variables in the SDOH Database

For the SDOH database, ZIP Codes have been linked to ZCTAs using a 1:1 match when available, and on a spatial join when not available. Overall, roughly 76.4% of ZIP Codes match directly to a ZCTA, and 23.4% were matched using a spatial join.

Some of the data sources for AHRQ SDOH Database are as below in figure 8.

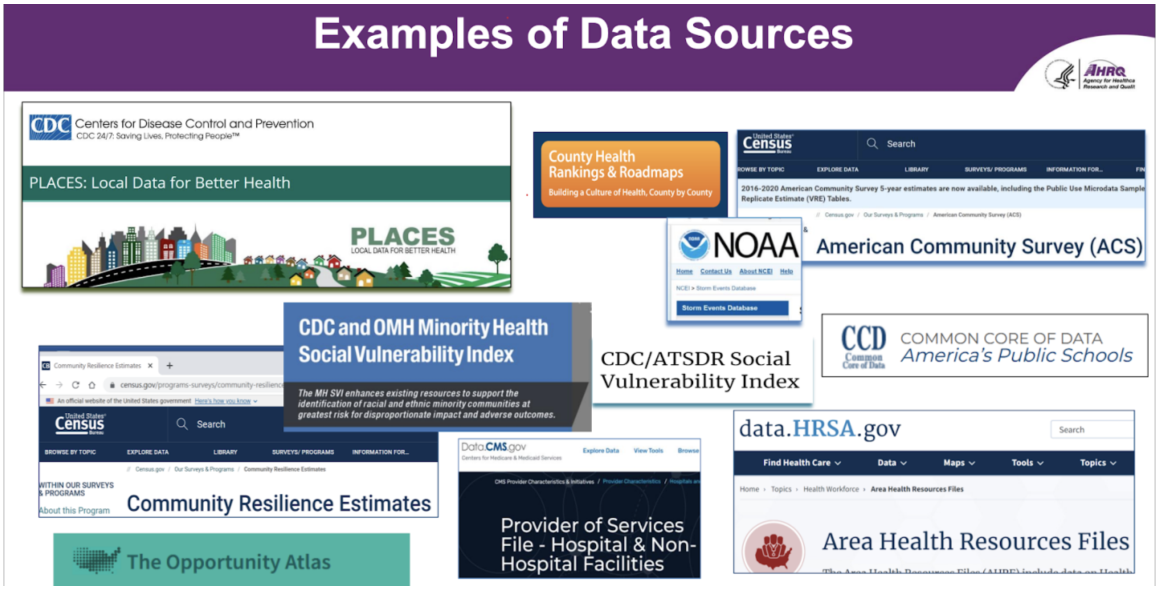


Figure 8 data sources for AHRQ SDOH Database

Below data files from AHRQ have been used for SDOH correlation analysis:

1. County File from AHRQ

https://www.ahrq.gov/sites/default/files/wysiwyg/sdoh/SDOH_2020_COUNTY_1_0.xlsx

2. Zip code File from AHRQ

https://www.ahrq.gov/sites/default/files/wysiwyg/sdoh/SDOH_2020_ZIPCODE_1_0.xlsx

Next going to describe the data preparation and mapping generation for the dementia disease and its impact on the SDOH markers.

3.4.4.2 Employee Health Records (EHR)

Employee Health Records (EHR) is a comprehensive set of health records of patients. The below data has been sourced from obfuscated database of Pharmacy Orders of online pharmacy company with consent. Its included in table 11 below.

Table 11
EHR Attributes and Description

Attribute	**Attribute Description**
Patient Key	Patient identifier (obfuscated)
AGE	Age of the patient
GENDER	Gender of the patient
Diagnosis_Code	ICD-10-CM diagnosis code for the patient
Posting_Date	The Data and time when pharmacy order was received
DIAGCODE_Long_Description	Description for ICD-10-CM diagnosis code for the patient
Address ZIP code	Postal Zip code where the patient resides in US
City	City where the patient resides in US
State	State where the patient resides in US
Age Group	Age group of the person
Posting Year	The year when patient ordered medicines
Zipcode	Postal zipcode of the area where patient lives
Zipcode Extn	Postal zipcode extension of the area where patient lives
County	FIPS county name where patient lives
FIPS code	FIPS county code where patient lives

3.4.4.2 Data Preparation

Below Figure shows the flow chart for the particular objective in which SDOH marker impact is generated for the demetia disease. The process begins by removing patient names and confidential identifiers from the dataset to protect privacy. Next, the zip codes are separated from the patient addresses for mapping purposes. Cleaning of medical codes and descriptions is done by removing dots and correcting errors. Data characteristics, such as the number of patients, records, and details about cities and zip codes, are then determined. The patient population is divided into age groups based on their ages as of October 2023. Posting year and month are derived from posting timestamps to understand patient diagnosis history. Hypothesis tests are conducted on individual Social Determinants of Health (SDOH) factors, such as Gender, using CHI-SQUARE analysis to find correlations with disease prognosis. The flowchart also includes steps for testing community SDOH factors' impact on disease progression, filtering for a length of stay greater than or equal to two years. The results of the analysis are presented, concluding the flowchart.

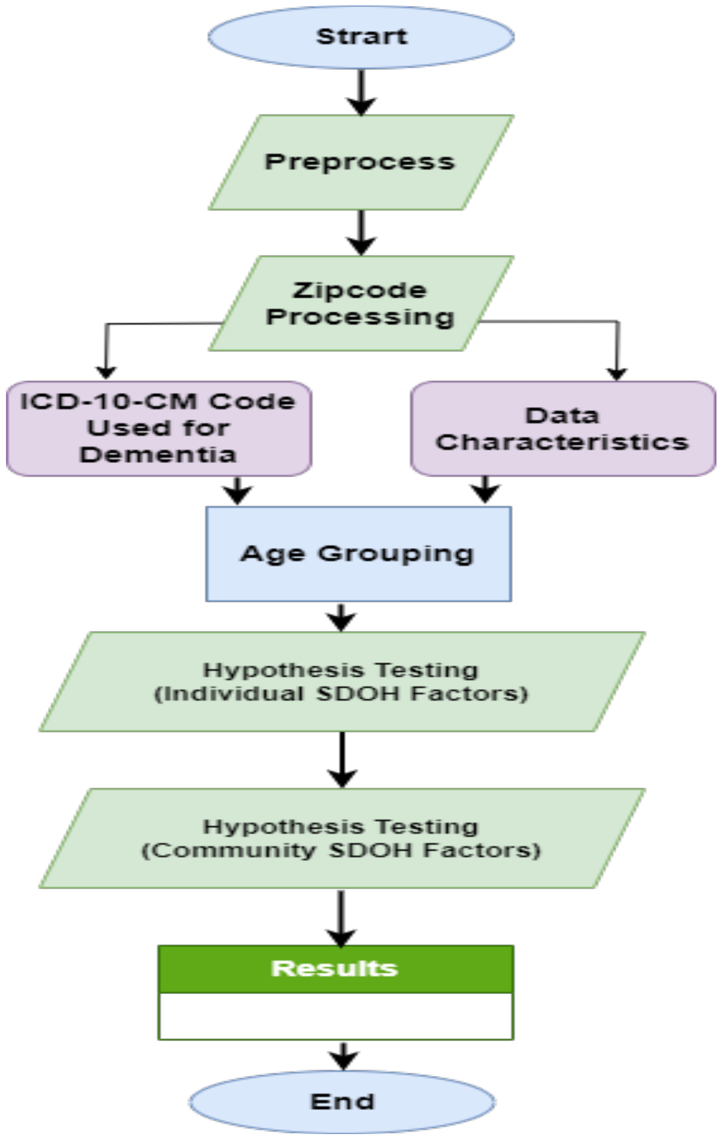


Figure 9 Steps to prepare data for dementia disease

1. Patient names have been removed from the dataset to protect the personally identifying information (PII) and confidential information.

2. Patient identifiers like MRN have been removed from the dataset as this would be confidential information that needs to be obfuscated.

3. Separated Zipcode and Zipcode extensions from the Address zipcode field of the patients so that these can be mapped to the community data collected for these locations.

Table 12

Zipcode Data City Wise

City	Zipcode	Total
Henderson	89002	206
	89011	49
	89012	235
	89014	96
	89015	160
	89016	90
	89044	427
	89052	299
	89074	706
Las Vegas	89101	223
	89102	68
	89103	282
	89104	88
	89106	20
	89107	411
	89108	915
	89110	441
	89113	63
	89115	80
	89117	730
	89118	356
	89119	523
	89120	30
	89121	859
	89122	236
	89123	36
	89128	206
	89129	943
	89130	775
	89131	207
	89134	335
	89135	726
	89139	173
	89141	114
	89142	71
	89143	8

	89144	193
	89145	202
	89146	478
	89147	939
	89148	83
	89149	324
	89156	40
	89166	143
	89169	11
	89183	3
North Las Vegas	89030	240
	89031	777
	89032	565
	89081	234
	89084	156
	89086	31
Pahrump	89048	30
	89061	24
RENO	89506	14
	89523	48
Grand Total		**15722**

4. ICD-10-CM code data quality has been improved by removing dots(.) from the codes to represent it in a standard format as developed and maintained by CDC's National Center for Health Statistics under authorization by the WHO.

5. ICD-10-CM long description data quality has been improved by correcting typographical and grammatical errors that usually happens where these are entered manually in the system by medical coding analysts. Final list of ICD-10-CM codes and their descriptions is as below that has been used in the analysis and experiments.

Table 13
ICD-10-CM code specification for Dementia Disease

ICD_Code	ICD_CODE_Long_Description
G300	Alzheimer's disease with early onset
G301	Alzheimer's disease with late onset

G308	Alzheimer's disease, Other
G309	Alzheimer's disease, unspecified
G3101	Frontotemporal dementia/Pick's disease
G3109	Frontotemporal dementia, Other
G311	Senile degeneration of brain, not elsewhere classified
G312	Degeneration of nervous system due to alcohol
G3183	Mixed Lewy body & vascular dementia with behavioural disturbances
G3184	Mild cognitive impairment, so stated
G3185	**Frontotemporal dementia/Corticobasal degeneration (CBD)**
G3189	Degenerative diseases of nervous system, Other specified
G319	**Degenerative disease of nervous system, unspecified**

6. Data characteristics are as below:

- Count of distinct patients - 5203
- Count of patient records – 15722 records
- Data represents patients from five cities of Nevada state representing 56 zip codes.

7. Divide the patient population into age group as below, based on their age as on Oct 2023.

Table 14
Patients age category for Dementia Disease

Age Group
<=50 years
51-60 years
61-70 years
71-80 years
81-90 years
90+ years

8. Derive Posting Year and posting month from posting timestamp to understand patient diagnosis history.

Table 15
Transportation Insecurity category

WalkToWork Groups
<=1 mile
1-3 miles
>3 miles

9. For Individual SDOH factors (Gender), do Hypothesis test to determine the extent of dependence of these factors on chronic disease prognosis. CHI-SQUARE analysis has been used for categorical data (eg AQI, distance to health facility) to determine their correlation to diseases prognosis. Find out Chi square value, p-value.

Table 16
Social isolation category

Population Density Group
<=200
200-20000
20000-50000
>50000

10. Divide the DistanceToCare data from SDOH database into groups as below,.

Table 17
Access To Care category

DistanceToCare Group
<=1 mile
1-5 miles

5-20 miles
>20 miles

11. Divide the BelowFPL data from SDOH database into groups as below.

Table 18
Social Vulnerability category

BelowFPL Group
<=20%
20-50%
50-80%
>80%

12. Divide the CollegeDegree data from SDOH database into groups as below.

Table 19
Education Levels category

CollegeDegree Group
<=20%
20-30%
30-50%
>50%

13. Divide the NoFoodStamps data from SDOH database into groups as below.

Table 20
Food Insecurity category

NoFoodStamps Group
<=10%
10-20%
20-50%
>50%

14. Divide the HighRent data from SDOH database into groups as below.

Table 21
Housing Insecurity category

HighRent Group
<=30%
30-50%
50-70%
>70%

15. Divide the Income Disparity data from SDOH database into groups as below.

Table 22
Financial Insecurity category

IncomeDisparity Group
<=20%
20-40%
40-60%
>60%

16. Divide the PM2.5 data from SDOH database into groups as below.

Table 23
Air Quality category

PM2.5 Group
<=12
12-30
30-80
>80

17. Divide the Water Quality data from SDOH database into groups as below.

Table 24
Water Quality category

WaterQuality Group
<=10
10-30
30-50
>50

18. Divide the AccessToSpecialtyCare data from SDOH database into groups as below.

Table 25
Access To Specialty Care category

Access To Specialty Care Group
<=2 per 10000 people
2-5 per 10000 people
5-10 per 10000 people
>10 per 10000 people

19. Derive Posting Year and posting month from posting timestamp to understand patient diagnosis history.

20. For Individual SDOH factors (Gender), do Hypothesis test to determine the extent of dependence of these factors on chronic disease prognosis. CHI-SQUARE analysis has been used for categorical data (eg AQI, distance to health facility) to determine their correlation to diseases prognosis. Find out Chi square value, p-value.

3.4.5 To Investigate the Impact of SDOH Markers on Alzheimer’s

To examine the influence of Social Determinants of Health (SDOH) markers on Alzheimer's disease, it is imperative to comprehend the distinctive attributes of Alzheimer's in comparison to other forms of dementia. Alzheimer's, accounting for 50 to

80 percent of dementia cases, is a distinct category of dementia distinguished by physical changes in the brain. These changes encompass both shrinkage in certain regions and expansion in others, resulting in disrupted electrical signals. The etiology of Alzheimer's remains incompletely understood, yet the "amyloid cascade hypothesis" is widely deliberated and researched. The duration of Alzheimer's is characterized by progressive and irreversible damage occurring in stages, with an average span lasting from 8 to 20 years. Symptoms manifest as difficulties in retaining newly acquired information, disorientation, as well as mood and behavioral alterations as the disease advances.

When investigating the influence of SDOH markers on Alzheimer's, the process of data preparation and mapping generation adheres to a methodical approach. To safeguard privacy, patient names and confidential identifiers are eliminated from the dataset, while patient addresses undergo processing to separate zip codes for mapping purposes. Furthermore, medical codes, specifically ICD-10-CM codes, undergo quality enhancement by standardizing their format and rectifying errors in their descriptions. Subsequently, the dataset, which represents patients from five cities in Nevada encompassing 56 zip codes, is characterized by counting distinct patients and patient records. The patient population is then segregated into age groups based on their ages as of October 2023, and the posting year and month are derived from timestamps to gain insight into patient diagnosis history.

To evaluate the impact of individual SDOH factors on the prognosis of Alzheimer's, hypothesis tests, particularly CHI-SQUARE analysis, are conducted. This entails exploring the correlation between SDOH factors such as gender and categorical data including Air Quality Index (AQI) and distance to health facilities with the prognosis of Alzheimer's. Chi-square values and p-values are utilized to quantify the degree of dependence. Additionally, community-level SDOH factors are examined by analyzing the length of stay, with a specific focus on cases where the duration of stay is equal to or exceeds two years, in order to measure the impact on disease progression.

In conclusion, the investigation into the influence of SDOH markers on Alzheimer's necessitates meticulous data preparation, characterization, and statistical analyses to comprehend the intricate interplay between social determinants and the advancement of this particular form of dementia.

3.4.5.1 Data Preparation

The flowchart below illustrates the process aimed at understanding the impact of Social Determinants of Health (SDOH) markers on Alzheimer's disease. The initial step involves safeguarding patient privacy by removing names and confidential identifiers from the dataset. Subsequently, patient addresses are processed to extract zip codes for mapping purposes. The medical codes and descriptions undergo a cleaning process, involving the removal of dots and correction of errors. Key data characteristics, including the count of patients, records, and details about cities and zip codes, are then determined.

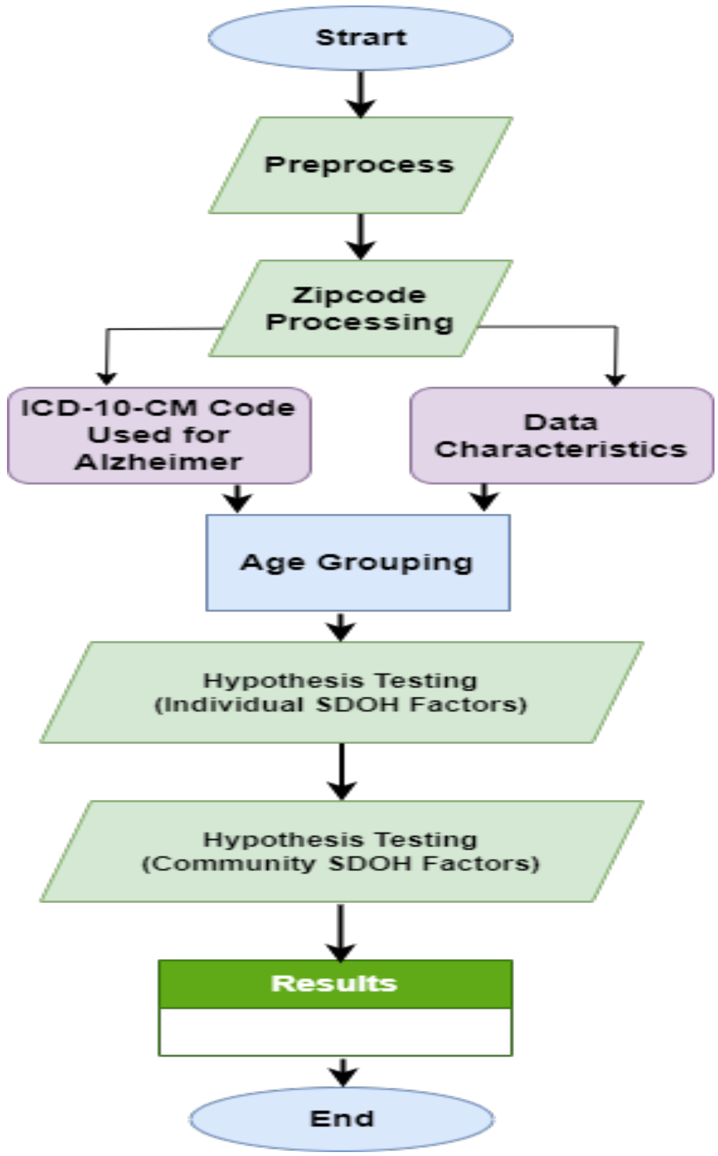


Figure 10 Steps to prepare data for Alzheimer's disease

1. Patient names have been removed from the dataset to protect the personally identifying information (PII) and confidential information .
2. Patient identifiers like MRN have been removed from the dataset as this would be confidential information that needs to be obfuscated.

3. Separated Zipcode and Zipcode extensions from the Address zipcode field of the patients so that these can be mapped to the community data collected for these locations.

4. ICD-10-CM code data quality has been improved by removing dots(.) from the codes to represent it in a standard format as developed and maintained by CDC's National Center for Health Statistics under authorization by the WHO.

5. ICD-10-CM long description data quality has been improved by correcting typographical and grammatical errors that usually happens where these are entered manually in the system by medical coding analysts. Final list of ICD-10-CM codes and their descriptions is as below that has been used in the analysis and experiments.

Table 26
ICD-10-CM code specification for Alzheimer's Disease

ICD_Code	ICD_CODE_Long_Description
G300	Alzheimer's disease with early onset
G301	Alzheimer's disease with late onset
G308	Alzheimer's disease, Other
G309	Alzheimer's disease, unspecified

6. Data characteristics are as below:

- Count of distinct patients - 5203
- Count of patient records – 15722 records
- Data represents patients from five cities of Nevada state representing 56 zip codes.

7. Divide the patient population into age group as below, based on their age as on Oct 2023.

Table 27
Patients age category for Alzheimer's Disease

Age Group
<=50 years
51-60 years
61-70 years
71-80 years
81-90 years
90+ years

8. Divide the WalkToWork data from SDOH database into groups as below.

Table 28
Transportation Insecurity category

Walk To Work Groups
<=1 mile
1-3 miles
>3 miles

9. Divide the Population Density data from SDOH database into groups as below,.

Table 29
Social isolation category

Population Density Group
<=200
200-20000
20000-50000
>50000

10. Divide the DistanceToCare data from SDOH database into groups as below.

Table 30
Access To Care category

Distance To Care Group

<=1 mile
1-5 miles
5-20 miles
>20 miles

11. Divide the BelowFPL data from SDOH database into groups as below,.

Table 31
Social Vulnerability category

BelowFPL Group
<=20%
20-50%
50-80%
>80%

12. Divide the CollegeDegree data from SDOH database into groups as below,.

Table 32
Education Levels category

CollegeDegree Group
<=20%
20-30%
30-50%
>50%

13. Divide the NoFoodStamps data from SDOH database into groups as below,.

Table 33
Food Insecurity category

NoFoodStamps Group
<=10%
10-20%
20-50%

>50%

14. Divide the HighRent data from SDOH database into groups as below,.

Table 34
Housing Insecurity category

HighRent Group
<=30%
30-50%
50-70%
>70%

15. Divide the IncomeDisparity data from SDOH database into groups as below,.

Table 35
Financial Insecurity category

IncomeDisparity Group
<=20%
20-40%
40-60%
>60%

16. Divide the PM2.5 data from SDOH database into groups as below,.

Table 36
Air Quality category

PM2.5 Group
<=12
12-30
30-80
>80

17. Divide the WaterQuality data from SDOH database into groups as below,.

Table 37
Water Quality category

WaterQuality Group
<=10
10-30
30-50
>50

18. Divide the AccessToSpecialtyCare data from SDOH database into groups as below,.

Table 38
Access To Specialty Care category

AccessToSpecialtyCare Group
<=2 per 10000 people
2-5 per 10000 people
5-10 per 10000 people
>10 per 10000 people

19. Derive Posting Year and posting month from posting timestamp to understand patient diagnosis history.
20. For Individual SDOH factors (Gender), do Hypothesis test to determine the extent of dependence of these factors on chronic disease prognosis. CHI-SQUARE analysis has been used for categorical data (eg AQI, distance to health facility) to determine their correlation to diseases prognosis. Find out Chi square value, p-value.

3.5 Research Design Limitations

1. EHR data does not include Race- The racial background of an individual is often regarded as confidential information, and patients may choose not to disclose this information during their medical treatments. Consequently, Electronic Health Records

(EHR) commonly do not capture data related to an individual's race. Nonetheless, medical research indicates that certain diseases exhibit varying prevalence among different racial groups. Despite recognizing this potential correlation, conducting thorough correlation analysis has proven challenging within the current context.

2. EHR data does not include Sexual orientation- The sexual orientation of an individual is considered sensitive information, and patients may opt not to disclose this aspect during their medical treatments. As a result, Electronic Health Records (EHR) typically do not include data related to the sexual orientation of patients. However, research in the medical field acknowledges that certain health issues may have varying prevalence among different sexual orientations. Despite this recognition, conducting comprehensive correlation analysis with sexual orientation data poses challenges within the existing framework.

3. Most EHR data is collected by addresses ie City, State, Zipcode while the Census tracts are available by FIPS code. There is no direct mapping between the two. Zip codes are collections of mail delivery routes used by USPS to expedite mail delivery and are organized for postal delivery routes efficiency. FIPS codes are organized by county lines as defined us Federal government and are assigned in alphabetic order to counties counting by odd numbers. The first county in a state alphabetically is always 001, the second county is always 003, the third is 005 and so on. Hence ZIP Codes and cities often cross one or more county lines. A particular ZIP Code or city may yield up to 5 correct FIPS Codes. Likewise a city may cross multiple county lines . This results in multiple valid FIPS Codes for a single ZIP. Hence the SDOH data which is available for FIPS (Counties) may not be an exact representation of where the person lines. But still it is a close representation and can be used for research purposes.

4. Most of the measures/values of Community SDOH markers are taken from AHRQ files. These standardized national level measure has been considered suitable for qualitative research. However there could be different or higher quality measures from local agencies or private organizations that could improve the accuracy of correlation analysis but have not been considered.

5. SDOH data from AHRQ is from 2020. It is assumed that these factors have not changed much significantly in 2021, 2022 and 2023 and hence can be used for correlating with EHR data.

3.6 Conclusion

Dementia & Alzheimer's are similar but not the same even though the physical ailments and the conditions may seem similar, especially memory loss. Dementia is an overall term that describes a wide range of symptoms associated with a decline in memory. It can also effect other thinking skills enough to reduce a person's ability to perform everyday activities. Dementia is often incorrectly referred to as "senility" due to the widespread incorrect belief that serious mental decline is a normal part of aging.

Senile Dementia of the Alzheimer's Type (SDAT), or simply called Alzheimer's is a disease. It produces physical change in the brain. There is shrinking in some areas of the brain and widening in the others. This causes connections inside the brain to break and disrupt the brain's electrical signals. Alzheimer's disease accounts for 50 to 80 percent of the dementia cases and is the most type of dementia. Vascular dementia, which occurs after a stroke, is the second most common type. There are many other conditions such as thyroid problems and vitamin deficiencies that can cause symptoms of dementia. Some of these are reversible. Ozone is a strong oxidizing pollutant that can cause nerve damage by inducing the release of free radicals, activating the production of inflammatory cytokines, and damaging the integrity of the blood–brain barrier.

Medications for Alzheimer's disease : Cholinesterase inhibitors & Memantine Three commonly prescribed cholinesterase inhibitors are:

- **Donepezil (Aricept)** is approved to treat all stages of the disease. It's taken once a day as a pill.
- **Galantamine (Razadyne)** is approved to treat mild to moderate Alzheimer's. It's taken as a pill once a day or as an extended-release capsule twice a day.
- **Rivastigmine (Exelon)** is approved for mild to moderate Alzheimer's disease. It's taken as a pill. A skin patch is available that can also be used to treat severe Alzheimer's disease.
- **Aducanumab :** intravenous infusion therapy -Approved only for patients with mild cognitive impairment and mild dementia due to Alzheimer's disease.
- **Lecanemab is given as an IV infusion every two weeks**

CHAPTER IV:

RESULTS

4.1 Results Investigated on the Impact of SDOH Markers on Dementia prognosis

Dementia is an overall term that describes a wide range of symptoms associated with a decline in cognitive functioning that includes memory loss. It can also effect other thinking skills enough and ability to reason to reduce a person's ability to perform everyday activities. Some people with dementia cannot control their emotions, and their personalities may also change. Dementia is often incorrectly referred to as "senility" due to the widespread incorrect belief that serious mental decline is a normal part of aging.

4.1.1 Dementia prognosis with AGE Groups

Data distribution for age groups.

Table 39
Dementia Prognosis Data distribution for age groups

Diagnosis_Code	<=50	51-60	61-70	71-80	81-90	90+
G300	2	1	32	49	70	22
G301	0	0	2	26	56	28
G308	0	1	3	9	13	3
G309	7	17	104	435	791	331
G3101	0	0	1	4	3	0
G3109	0	0	5	7	3	0
G311	0	0	0	11	10	4
G312	1	0	1	3	0	0
G3183	0	4	7	25	27	6
G3184	27	51	167	366	352	85

G3185	0	0	1	1	0	0
G3189	1	2	3	5	4	1
G319	13	33	235	742	794	196

Below the chart illustres the dementia prognosis data distribution among the different age groups.

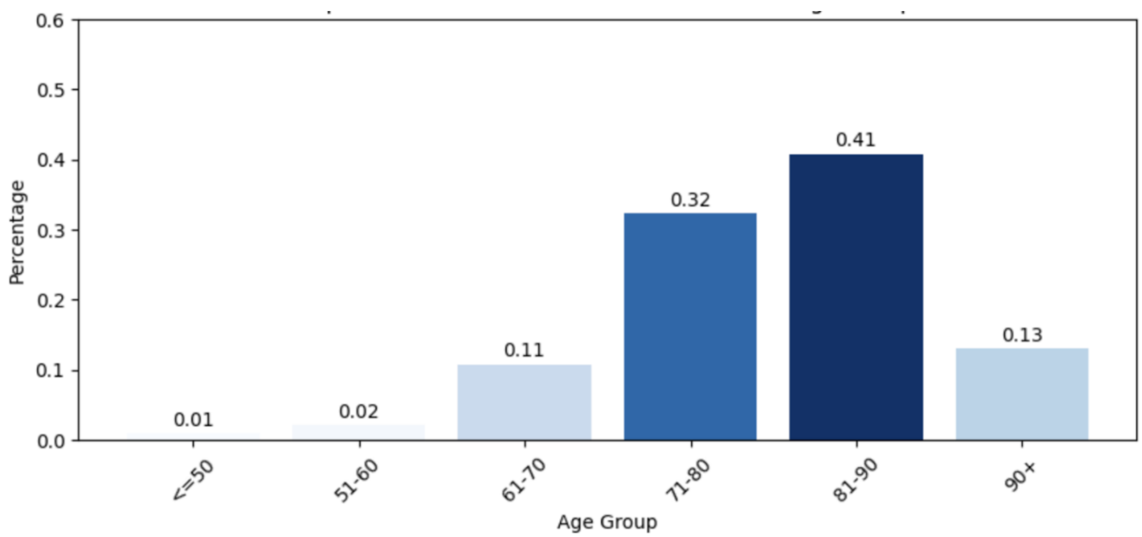


Figure 11 Proportion of Dementia Patients in Different Age Groups

Data for the testing purpose

- Significance level = 0.3
- Very significant level = 0.05
- Null Hypothesis (H0) – Dementia outcomes are independent of Age groups.
- Alternate Hypothesis (H1) – Dementia outcomes are dependent on Age groups.

Python program to perform the Chi-square test.

#Chi-Square Test of Independence to determine if there is a statistically significant relationship between the two categorical variables.

from scipy.stats import chi2_contingency

```
contingency_table =
pd.crosstab(diagnosis_df['Diagnosis_Code'],diagnosis_df['Age Group'])
chi2,p,dof,expected= chi2_contingency(contingency_table)
print(f"Chi-square value: {chi2}")
print(f"P-value: {p}")
print(f"Degrees of freedom: {dof}")
```

Generated values for parameters as output.

- Chi-square value: 415.47825576203945
- P-value: 1.2777465700857161e-54
- Degrees of freedom: 60

Chi-Square Value (chi2 = 415):The chi-square value is a measure of how much the observed counts deviate from the expected counts if there were no association between the variables. In our case, a chi-square value of 415 indicates a substantial difference between observed and expected frequencies.

P-value (pvalue = 1.2777465700857161e-54):The p-value is the probability of observing a chi-square statistic as extreme as the one calculated, assuming the null hypothesis (no association) is true. In our case, the very small p-value)) suggests strong evidence against the null hypothesis. It indicates that the association between the variables is likely not due to random chance.

Degrees of Freedom (dof = 60):Degrees of freedom represent the number of values in the final calculation of a statistic that are free to vary. In our case, with 60 degrees of freedom, the critical value for determining statistical significance is based on the chi-square distribution with 60 degrees of freedom.

Since p-value is lesser than Significance level (.01), Null hypothesis (H0) is Rejected.

Since p-value is much lesser than Very Significant level (.001), there is a **HIGH** correlation between Age groups and Dementia diagnosis.

4.1.2 Dementia prognosis with GENDER

Data distribution for GENDER

Table 40
Data distribution for Gender

Diagnosis_Code	F	M
G300	115	61
G301	70	42
G308	20	9
G309	1046	639
G3101	6	2
G3109	5	10
G311	20	5
G312	1	4
G3183	24	45
G3184	599	449
G3185	2	0
G3189	6	10
G319	1113	900

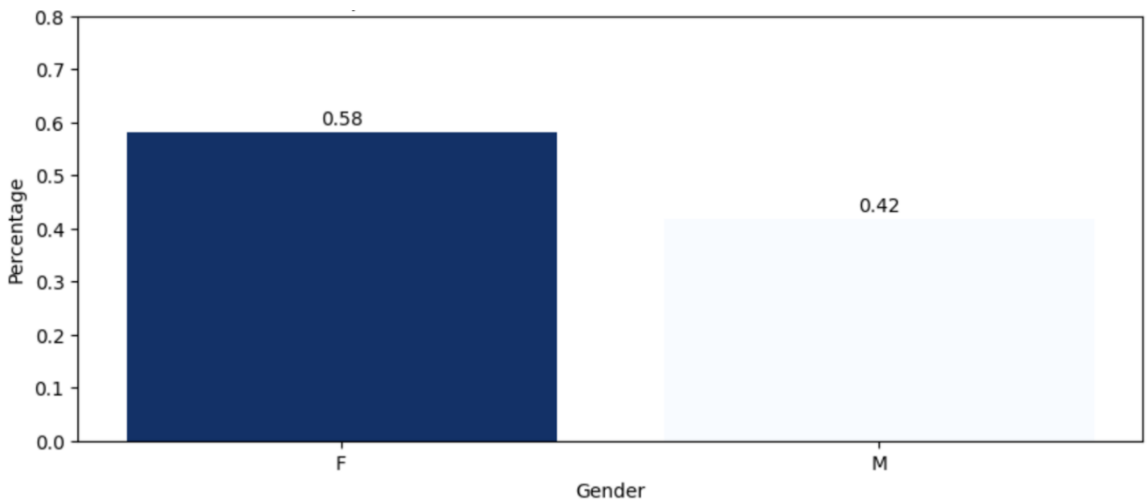


Figure 12 Proportion of Dementia Patients in Different Gender

Data for the testing purpose

- Significance level = 0.3
- Very significant level = 0.05

Null Hypothesis (H0) – Dementia outcomes are independent of Gender

Alternate Hypothesis (H1) – Dementia outcomes are dependent on Gender

Generated values for parameters as output.

- Chi-square value: 56.22915113353555
- P-value: 1.0885164235114124e-07
- Degrees of freedom: 12

The impact of this markers Since p-value is lesser than Significance level, Null hypothesis (H0) is Rejected.

Since p-value is much lesser than Very Significant level, there is a **HIGH** correlation between GENDER and Dementia diagnosis.

4.1.3 Dementia prognosis with Transport Insecurity

Data distribution for Distance of Walk to Work

Table 41
Data distribution for Distance of Walk to Work

Diagnosis_Code	<=1 Mile	1-3 Miles	>3 Miles
G300	64	102	10
G301	48	55	9
G308	16	10	3
G309	638	928	121
G3101	3	5	0
G3109	5	9	1
G311	9	14	2
G312	2	2	1
G3183			
G3184	421	552	75
G3185	1	1	0
G3189	8	7	1
G319	729	1128	156

The Table 19 contains the distance for diffenret diagnosis codes/ ICD code in the three specified ranges in each column.

Figure 13 Proportion of Dementia Patients in Different Walk2Work Groups (Transportation Insecurity)

Significance level = 0.3

Very significant level = 0.05

Null Hypothesis (H0) – Dementia outcomes are independent of Transportation Insecurity

Alternate Hypothesis (H1) – Dementia outcomes are dependent on Transportation Insecurity

Generated values for parameters as output.

- Chi-square value: 17.183634268882496
- P-value: 0.8407250277318131
- Degrees of freedom: 24

Since p-value is higher than Significance level, Null hypothesis (H0) is Accepted. It is deducted that there is a **LOW or negligible** correlation between GENDER and Dementia prognosis.

4.1.4 Dementia prognosis with AIR Quality

Data distribution for Air Quality (PM 2.5)

Table 42
Data distribution for Air Quality

Diagnosis_Code	<=12
G300	176
G301	112
G308	29
G309	1685
G3101	8
G3109	15
G311	25
G312	5
G3183	0
G3184	4108
G3185	2
G3189	16
G319	2013

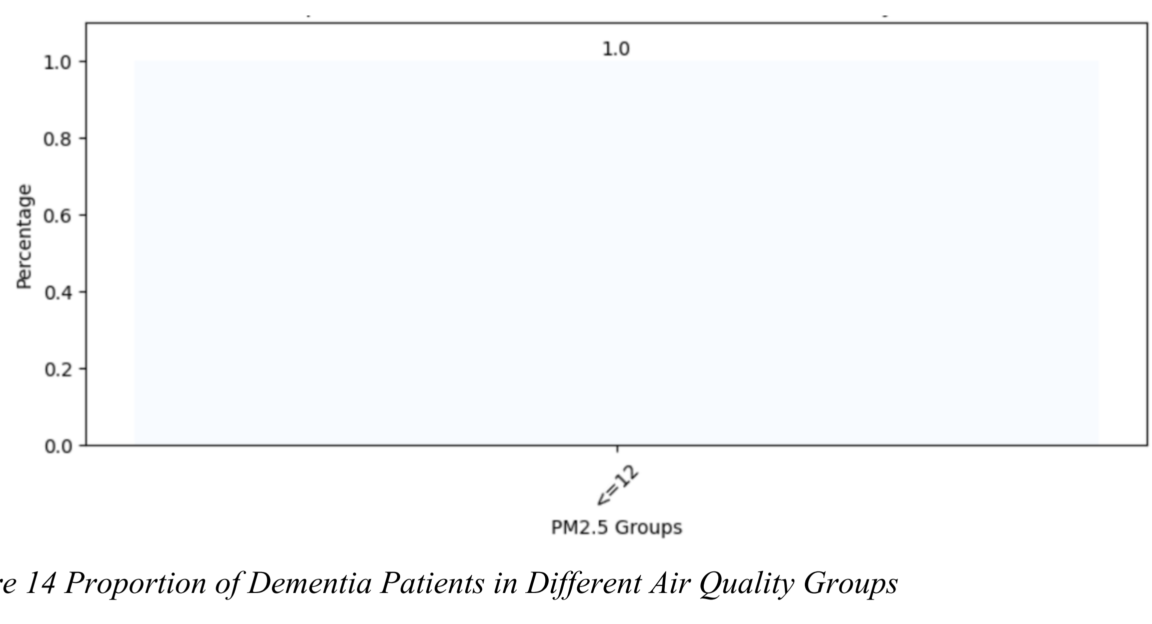


Figure 14 Proportion of Dementia Patients in Different Air Quality Groups

- Significance level = 0.3
- Very significant level = 0.05

Null Hypothesis (H0) – Dementia outcomes are independent of Air Quality

Alternate Hypothesis (H1) – Dementia outcomes are dependent on Air Quality

Generated values for parameters as output.

- Chi-square value: 0.0
- P-value: 0
- Degrees of freedom: 0

Since p-value is higher than Significance level, Null hypothesis (H0) is Accepted. It is deducted that there is a **LOW or negligible** correlation between Air Quality and Dementia prognosis.

4.1.5 Dementia prognosis with Social Isolation

Data distribution for Social isolation (Population Density)

Table 43
Data distribution for Social isolation

Diagnosis_Code	**PopulationDensity Groups <=500**
G300	176
G301	112
G308	29
G309	1685
G3101	8
G3109	15
G311	25
G312	5
G3183	0

G3184	4108
G3185	2
G3189	16
G319	2013

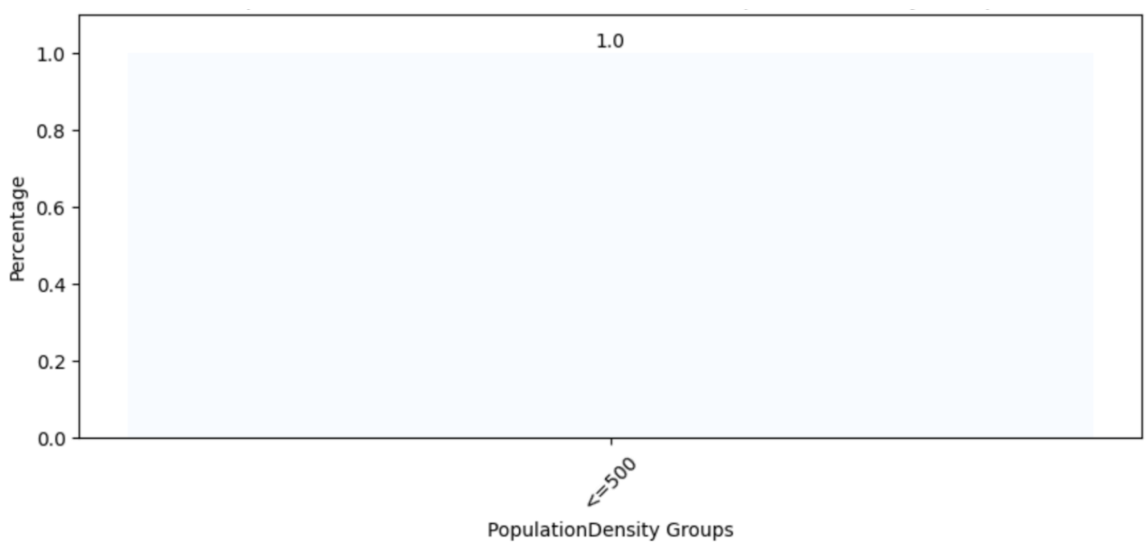


Figure 15 Proportion of Dementia Patients in Different Population Density Groups

- Significance level = 0.3
- Very significant level = 0.05

Null Hypothesis (H0) – Dementia outcomes are independent of Social isolation

Alternate Hypothesis (H1) – Dementia outcomes are dependent on Social isolation

Generated values for parameters as output.

- Chi-square value: 0.0
- P-value: 0
- Degrees of freedom: 0

Since p-value is higher than Significance level, Null hypothesis (H0) is Accepted. It is deducted that there is a **LOW or negligible** correlation between Social Isolation and Dementia prognosis.

4.1.6 Dementia prognosis with Access to Care Facilities

Data distribution for Access to Care (Distance to Care)

Table 44
Data distribution for Access to Care

Diagnosis_Code	<=1 mile	1-5 miles	5-20 miles	20+ miles
G300	20	97	59	0
G301	18	62	32	0
G308	4	17	8	0
G309	235	939	504	7
G3101	0	5	3	0
G3109	2	12	1	0
G311	3	14	7	1
G312	1	1	2	1
G3183	0	0	0	0
G3184	150	550	346	2
G3185	0	1	1	0
G3189	1	8	7	0
G319	289	1114	600	10

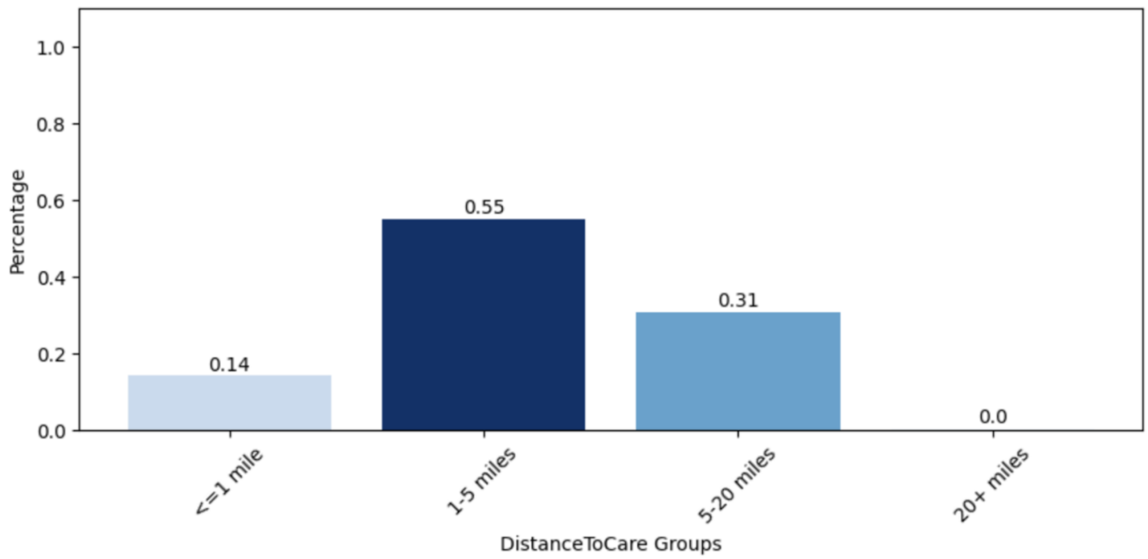


Figure 16 Proportion of Dementia Patients in Different Distance To Care Groups

- Significance level = 0.3
- Very significant level = 0.05

Null Hypothesis (H0) – Dementia outcomes are independent of Access to Care

Alternate Hypothesis (H1) – Dementia outcomes are dependent on Access to Care

Generated values for parameters as output.

- Chi-square value: 74.89112275623064
- P-value: 0.00015210618141466904
- Degrees of freedom: 36

Since p-value is lesser than Significance level, Null hypothesis (H0) is Rejected.

Since p-value is much lesser than Very Significant level, it is deducted that there is a **HIGH** correlation between Access to Care and Dementia prognosis.

4.1.7 Dementia prognosis with Social Vulnerability

Data distribution for Social Vulnerability (Income Disparity measured through GINI Index)

Table 45
Data distribution for Social Vulnerability

Diagnosis_Code	20-40%	40-60%
G300	30	146
G301	29	83
G308	5	24
G309	363	1322
G3101	3	5
G3109	3	12
G311	6	19
G312	0	5
G3183	0	0
G3184	239	809
G3185	1	1
G3189	4	12
G319	411	1602

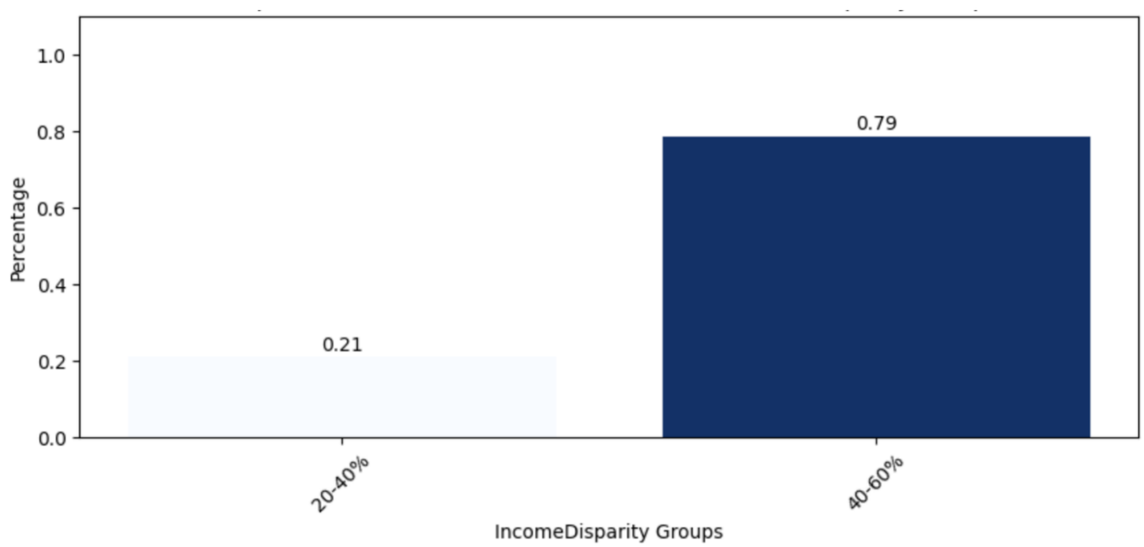


Figure 17 Proportion of Dementia Patients in Different Income Disparity Groups

- Significance level = 0.3
- Very significant level = 0.05

Null Hypothesis (H0) – Dementia outcomes are independent of Social Vulnerability

Alternate Hypothesis (H1) – Dementia outcomes are dependent on Social Vulnerability

Generated values for parameters as output.

- Chi-square value: 11.769210003812553
- P-value: 0.46438828368552343
- Degrees of freedom: 12

Since p-value is lesser than Significance level, Null hypothesis (H0) is Rejected.

Since p-value is higher than Significance level, Null hypothesis (H0) is Accepted. It is deducted that there is a **LOW or negligible** correlation between Social Vulnerability and Dementia prognosis.

4.1.8 Dementia prognosis with Education Levels

Data distribution for Education Levels (Income Disparity measured through GINI Index)

Table 46
Data distribution for Education Levels

Diagnosis_Code	<=20%	20-30%	30-50%
G300	0	41	135
G301	2	20	90
G308	1	3	25
G309	30	293	1362
G3101	0	1	7
G3109	0	6	9
G311	0	3	22
G312	1	0	4
G3183	0	0	0
G3184	10	189	849
G3185	0	0	2
G3189	0	3	13
G319	27	366	1620

Table 46 contains the data related to the education level taken from the GINI Index, and framed it in three level of the percentage ranges.

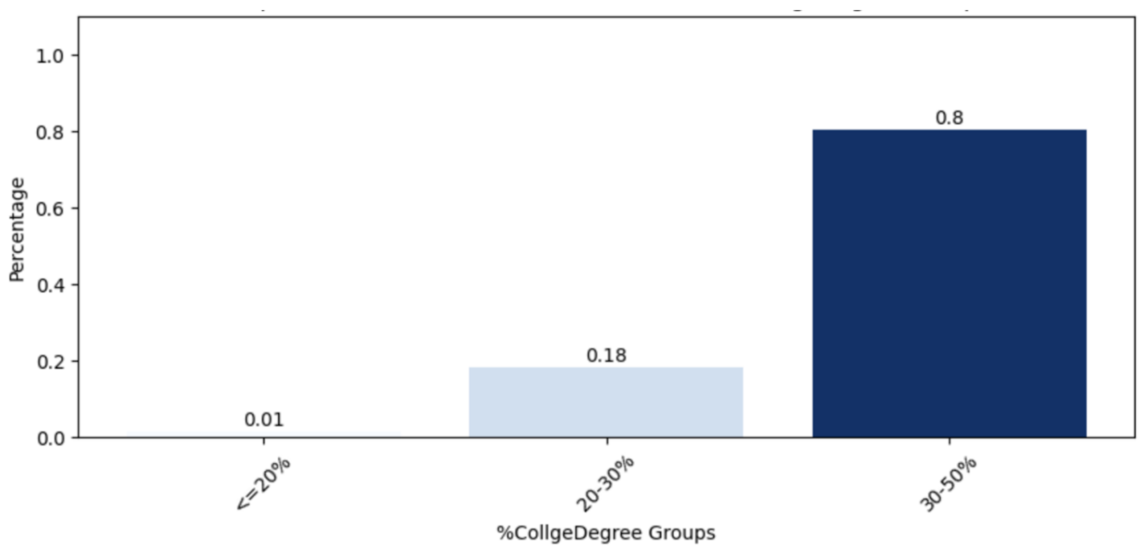


Figure 18 Proportion of Dementia Patients in with College Degree

Data contain for the analysis initially

- Significance level = 0.3
- Very significant level = 0.05

Null Hypothesis (H0) – Dementia outcomes are independent of Education levels.

Alternate Hypothesis (H1) – Dementia outcomes are dependent on Education levels.

Generated values for parameters as output.

- Chi-square value: 41.46847569334546
- P-value: 0.014789299546059408
- Degrees of freedom: 24

Since p-value is lesser than Significance level, Null hypothesis (H0) is Rejected.

Since p-value is lesser than Very Significant level, it is deducted that there is a **HIGH** correlation between Education Levels and Dementia prognosis.

4.1.9 Dementia prognosis with Food Insecurity

Data distribution for Food Insecurity (People below poverty line and not on food stamps)

Table 47
Data distribution for Food Insecurity

Diagnosis_Code	<=10%	10-20%
G300	145	31
G301	92	20
G308	24	5
G309	1366	319
G3101	8	0
G3109	9	6
G311	21	4
G312	4	1
G3183	0	0
G3184	870	178
G3185	2	0
G3189	13	3
G319	1647	366

The above table 47 shows the data for the distribution for Food Insecurity and made in no food group stams.

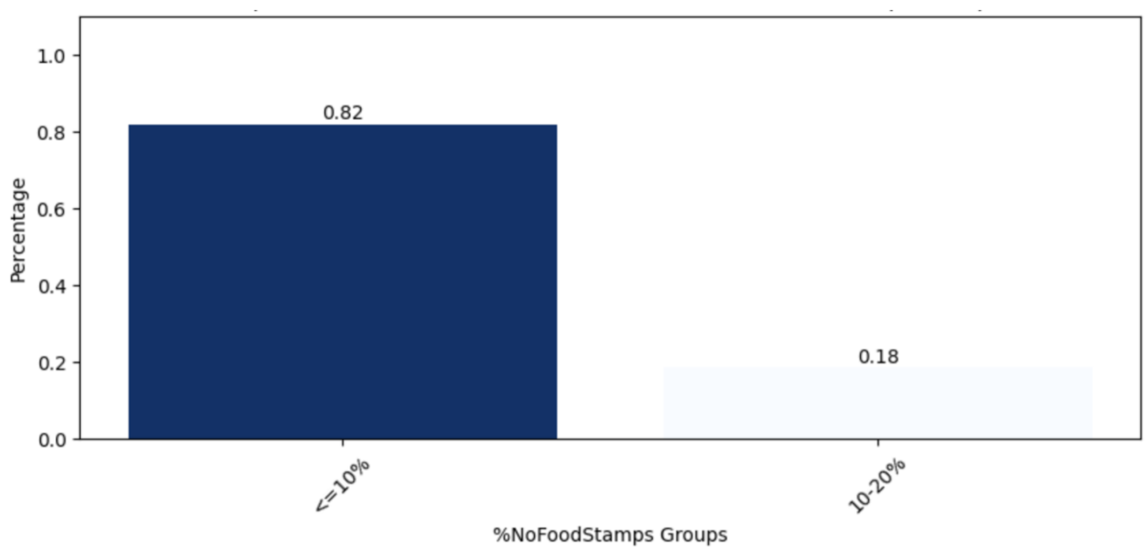


Figure 19 Proportion of Dementia Patients with Food Insecurity

Data contain for the analysis initially

- Significance level = 0.3
- Very significant level = 0.05

Null Hypothesis (H0) – Dementia outcomes are independent of Food Insecurity

Alternate Hypothesis (H1) – Dementia outcomes are dependent on Food Insecurity.

Parameters data generated are:

- Chi-square value: 14.09633099328138
- P-value: 0.2945991127050013
- Degrees of freedom: 12

Since p-value is lesser than Significance level, Null hypothesis (H0) is Rejected.

Since p-value is higher than Very Significant level, it is deducted that there is a **MODERATE** correlation between Food Insecurity and Dementia prognosis.

4.1.10 Dementia prognosis with Housing Insecurity

Data distribution for Housing Insecurity (People who spend >30% income on rent)

Table 48
Data distribution for Housing Insecurity

Diagnosis_Code	30-50%	50-70%
G300	79	97
G301	49	63
G308	16	13
G309	702	983
G3101	4	4
G3109	3	12
G311	6	19
G312	2	3
G3183	0	0
G3184	456	592
G3185	2	0
G3189	8	8
G319	797	1216

The above table 48 shows the data distribution for Housing Insecurity.

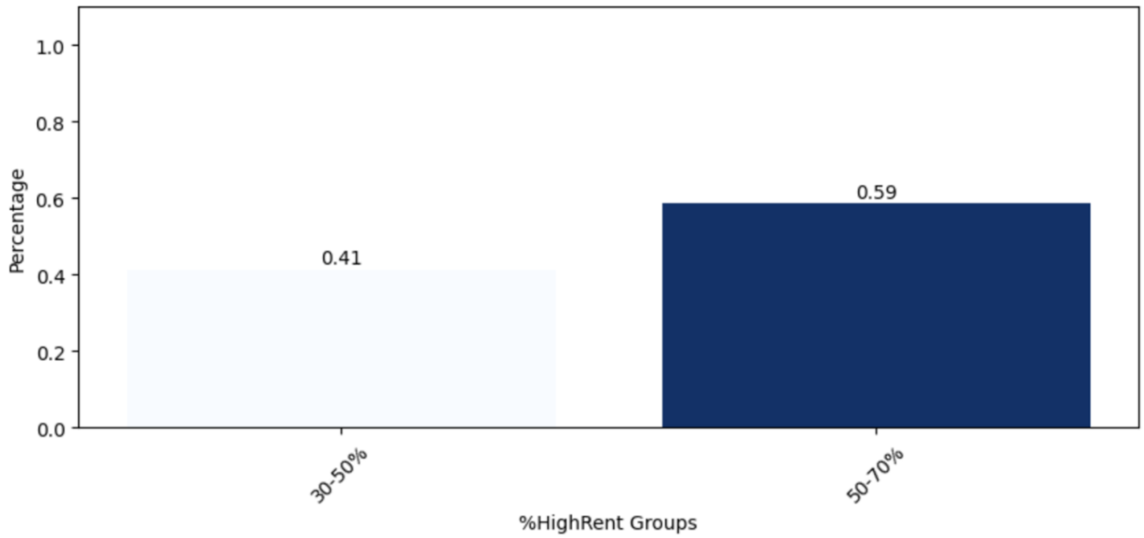


Figure 20 Proportion of Dementia Patients with Housing Insecurity

Data contain for the analysis initially

- Significance level = 0.3
- Very significant level = 0.05

Null Hypothesis (H0) – Dementia outcomes are independent of Housing Insecurity.

Alternate Hypothesis (H1) – Dementia outcomes are dependent on Housing Insecurity.

Parameters data generated are:

- Chi-square value: 18.838097627984773
- P-value: 0.09251125246398216
- Degrees of freedom: 12

Since p-value is lesser than Significance level, Null hypothesis (H0) is Rejected.

Since p-value is higher than Very Significant level, it is deducted that there is a **MODERATE** correlation between Housing Insecurity and Dementia prognosis.

4.1.11 Dementia prognosis with Financial Insecurity

Data distribution for Financial Insecurity (% of people whose income is below poverty line)

Table 49
Data distribution for Financial Insecurity

Diagnosis_Code	<=20%	20-50%
G300	161	15
G301	103	9
G308	25	4
G309	1502	183
G3101	8	0
G3109	14	1
G311	23	2
G312	4	1
G3183	0	0
G3184	950	98
G3185	2	0
G3189	15	1
G319	1812	201

The table 49 specifies the data that is below the FLP groups in two percentage categories.

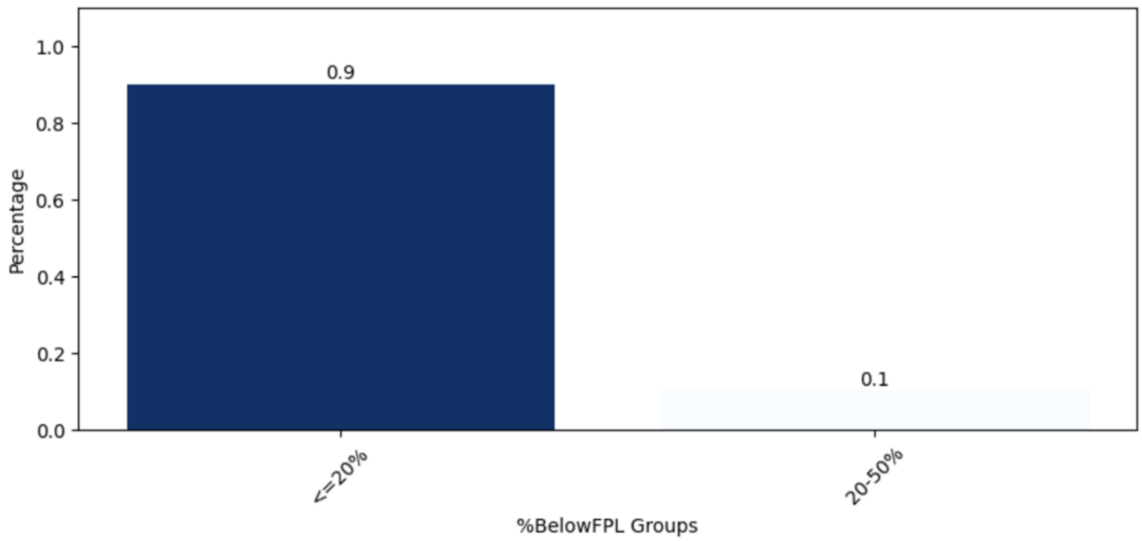


Figure 21 Proportion of Dementia Patients with Financial Insecurity

Data contain for the analysis initially

- Significance level = 0.3
- Very significant level = 0.05

Null Hypothesis (H0) – Dementia outcomes are independent of Housing Insecurity.

Alternate Hypothesis (H1) – Dementia outcomes are dependent on Housing Insecurity.

Parameters data generated are:

- Chi-square value: 9.4452520419399
- P-value: 0.6645026452657894
- Degrees of freedom: 12

Since p-value is higher than Significance level, Null hypothesis (H0) is Accepted. It is deducted that there is a **LOW or negligible** correlation between Housing Insecurity and Dementia prognosis.

4.1.12 Dementia prognosis with Access to Specialty Care

Data distribution for Access to Specialty Care (Community Mental health facilities)

Table 50
Data distribution for Access to Specialty Care

Diagnosis_Code	**<=2 per 10000 people**
G300	176
G301	112
G308	29
G309	1685
G3101	8
G3109	15
G311	25
G312	5
G3183	0
G3184	1048
G3185	2
G3189	16
G319	2013

Table 50 above contains the data for Access to Specialty Care (Community Mental health facilities)

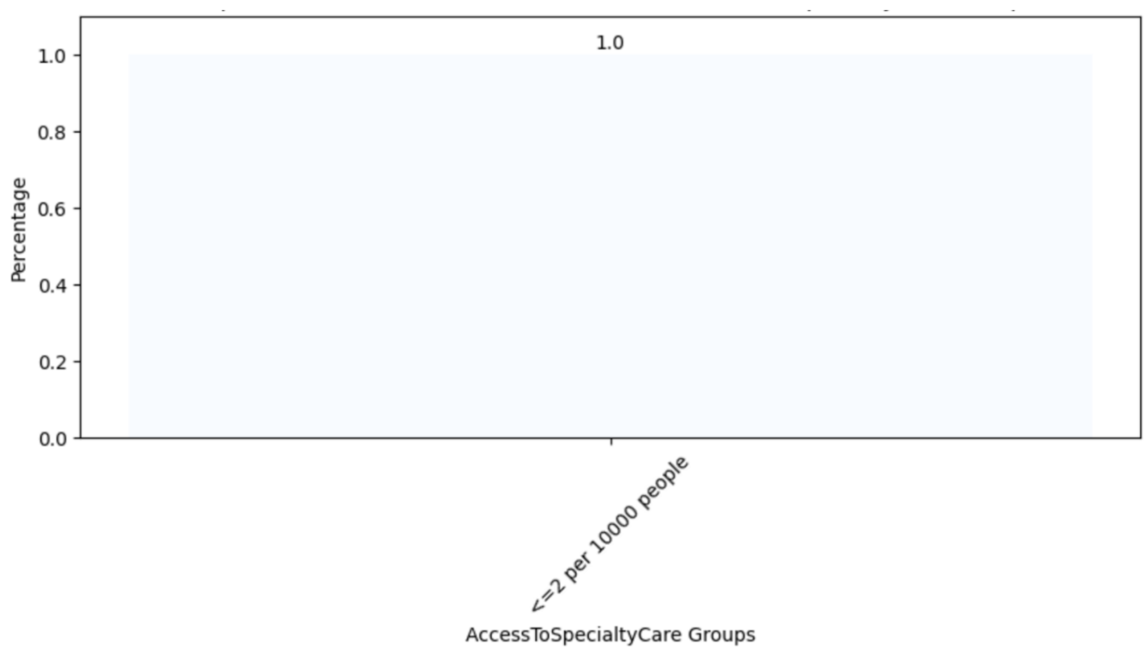


Figure 22 Proportion of Dementia Patients with Access to Specialty Care

Data contain for the analysis initially

- Significance level = 0.3
- Very significant level = 0.05

Null Hypothesis (H0) – Dementia outcomes are independent of Access To Specialty Care

Alternate Hypothesis (H1) – Dementia outcomes are dependent on Access To Specialty Care

Parameters data generated are:

- Chi-square value: 0.0
- P-value: 1.0
- Degrees of freedom: 0

Since p-value is higher than Significance level, Null hypothesis (H0) is Accepted. It is deducted that there is a **LOW or negligible** correlation between Access To Specialty Care and Dementia prognosis.

4.2 Results Investigated on Impact of SDOH Markers on Alzheimer's prognosis

Senile Alzheimer's of the Alzheimer Type (SDAT), or simply called Alzheimer's is a disease. It produces physical change in the brain. There is shrinking in some areas of the brain and widening in the others. This causes connections inside the brain to break and disrupt the brain's electrical signals. Alzheimer's disease accounts for 50 to 80 percent of the Alzheimer's cases and is the most type of Alzheimer's. Vascular Alzheimer's, which occurs after a stroke, is the second most common type. There are many other conditions such as thyroid problems and vitamin deficiencies that can cause symptoms of Alzheimer's. Some of these are reversible.

4.2.1 Alzheimer's prognosis with AGE Groups

Data distribution for age groups

Table 51
Data distribution for age groups for Alzheimer's

Diagnosis_Code	<=50	51-60	61-70	71-80	81-90	90+
G300	2	1	32	49	70	22
G301	0	0	2	26	56	28
G308	0	1	3	9	13	3
G309	7	17	104	435	791	331

The table 51 above shows the data for the different age groups for the case of Alzheimer's.

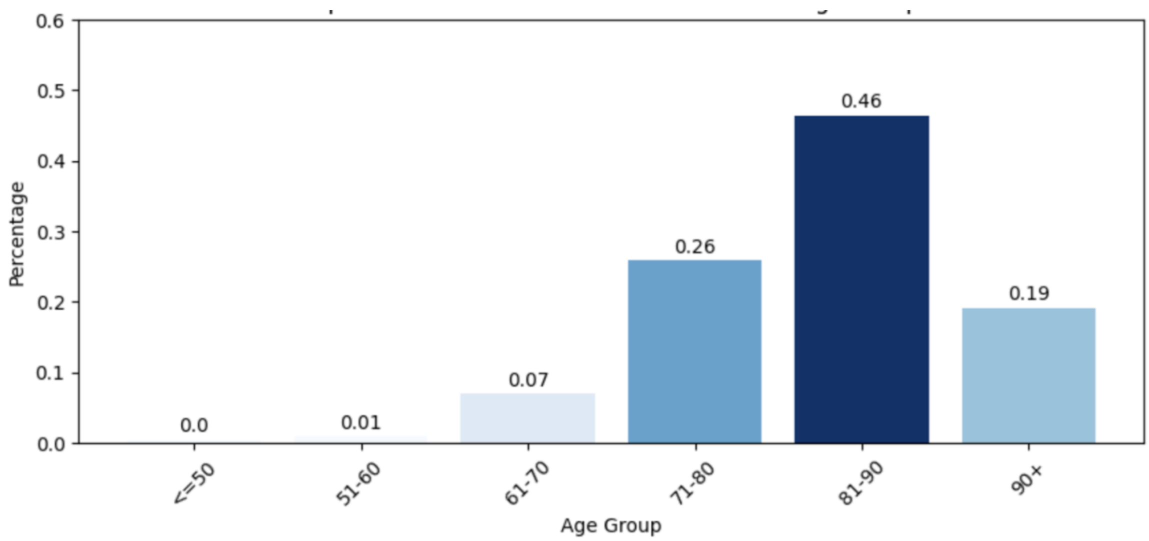


Figure 23 Proportion of Alzheimer's Patients in Different Age Groups

Data contain for the analysis initially

- Significance level = 0.3
- Very significant level = 0.05

Null Hypothesis (H0) – Alzheimer's outcomes are independent of Age groups

Alternate Hypothesis (H1) – Alzheimer's outcomes are dependent on Age groups

Parameters data generated are:

- Chi-square value: 53.89842825510161
- P-value: 2.731505467085729e-06
- Degrees of freedom: 15

Since p-value is lesser than Significance level, Null hypothesis (H0) is Rejected.

Since p-value is much lesser than Very Significant level, there is a **HIGH** correlation between Age groups and Alzheimer's prognosis.

4.2.2 Alzheimer's prognosis with GENDER

Data distribution for gender

Table 52

Data distribution for gender in Alzheimer's

Diagnosis_Code	F	M
G300	115	61
G301	70	42
G308	20	9
G309	1046	639

The table 52 depicts the gender data distribution for the Alzheimer's disease in ICD codes for male and female categories.

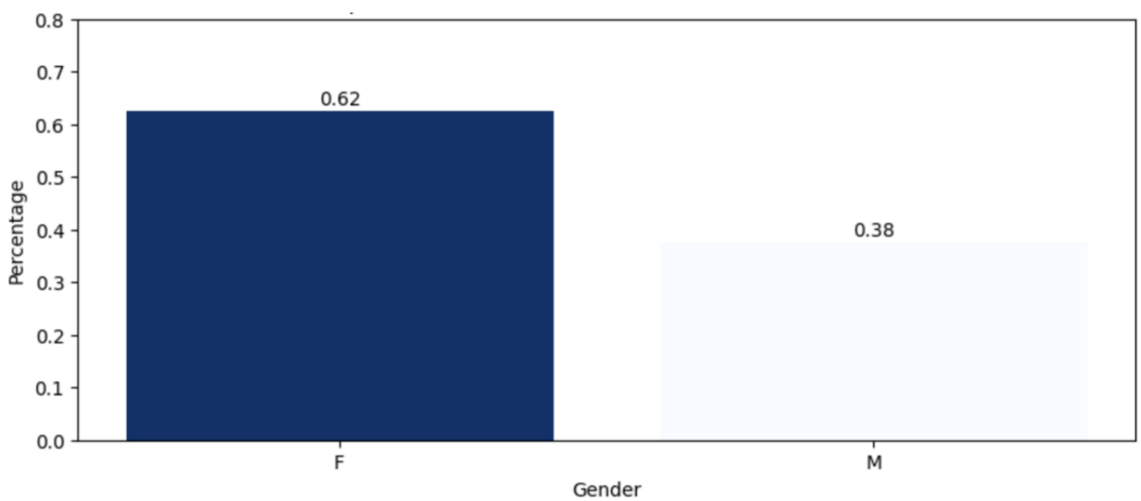


Figure 24 Proportion of Alzheimer's Patients in Different Gender

Data contain for the analysis initially

- Significance level = 0.3
- Very significant level = 0.05

Null Hypothesis (H0) – Alzheimer's outcomes are independent of Gender

Alternate Hypothesis (H1) – Alzheimer's outcomes are dependent on Gender

Parameters data generated are:

- Chi-square value: 1.251548156191844
- P-value: 0.7406693057008342

- Degrees of freedom: 3

Since p-value is higher than Significance level, Null hypothesis (H0) is Accepted. It is deducted that there is a **LOW or negligible** correlation between Gender and Alzheimer's prognosis.

4.2.3 Alzheimer's prognosis with Transportation Insecurity

Data distribution for Distance of Walk to Work

Table 53
Data distribution for Distance of Walk

Diagnosis_Code	<=1 Mile	1-3 Miles	>3 Miles
G300	64	102	10
G301	48	55	9
G308	16	10	3
G309	638	928	121

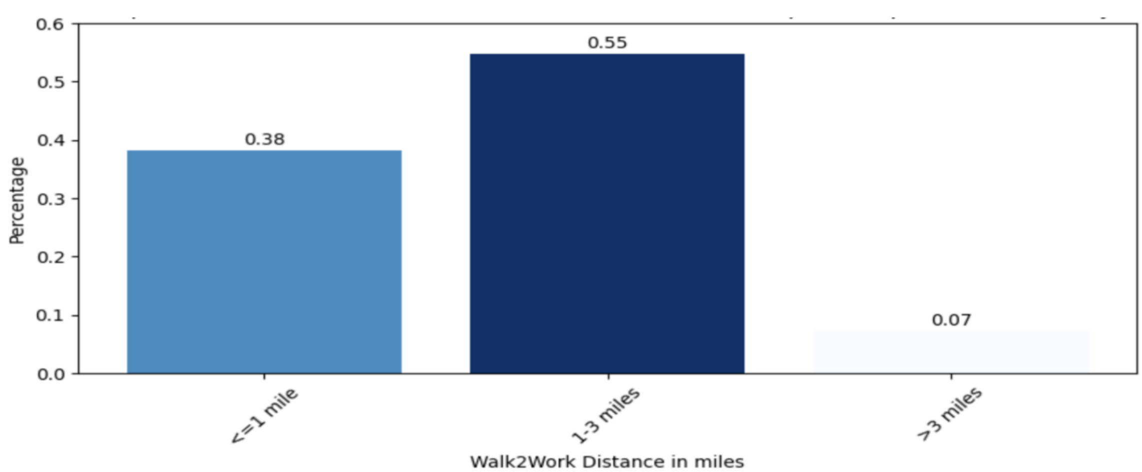


Figure 25 Proportion of Alzheimer's Patients in Different Walk2Work Groups (Transportation Insecurity)

Data contain for the analysis initially

- Significance level = 0.3

- Very significant level = 0.05

Null Hypothesis (H0) – Alzheimer's outcomes are independent of Transportation Insecurity

Alternate Hypothesis (H1) – Alzheimer's outcomes are dependent on Transportation Insecurity

Parameters data generated are:

- Chi-square value: 7.33415508650906
- P-value: 0.2910468670611926
- Degrees of freedom: 6

Since p-value is lesser than Significance level, Null hypothesis (H0) is Rejected.

Since p-value is higher than Very Significant level, it is deducted that there is a **MODERATE** correlation between Transportation Insecurity and Alzheimer's prognosis.

4.2.4 Alzheimer's prognosis with AIR Quality

Data distribution for Air Quality (PM 2.5)

Table 54
Data distribution for Air Quality Alzheimer's

Diagnosis_Code	<=12
G300	176
G301	112
G308	29
G309	1685

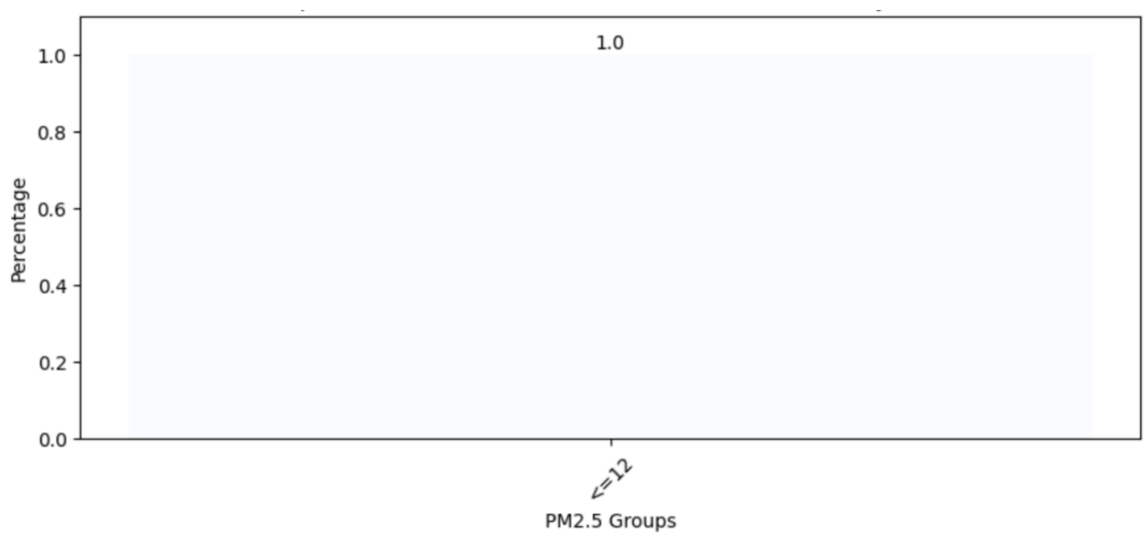


Figure 26 Proportion of Alzheimer's Patients in Different Air Quality Groups

Data contain for the analysis initially

- Significance level = 0.3
- Very significant level = 0.05

Null Hypothesis (H0) – Alzheimer's outcomes are independent of Air Quality

Alternate Hypothesis (H1) – Alzheimer's outcomes are dependent on Air Quality

Parameters data generated are:

- Chi-square value: 0.0
- P-value: 1.0
- Degrees of freedom: 0

Since p-value is higher than Significance level, Null hypothesis (H0) is Accepted. It is deducted that there is a **LOW or negligible** correlation between Air Quality and Alzheimer's prognosis.

4.2.5 Alzheimer's prognosis with Social Isolation

Data distribution for Social isolation (Population Density)

Table 55
Data distribution for Social isolation

Diagnosis_Code	<=500
G300	176
G301	112
G308	29
G309	1685

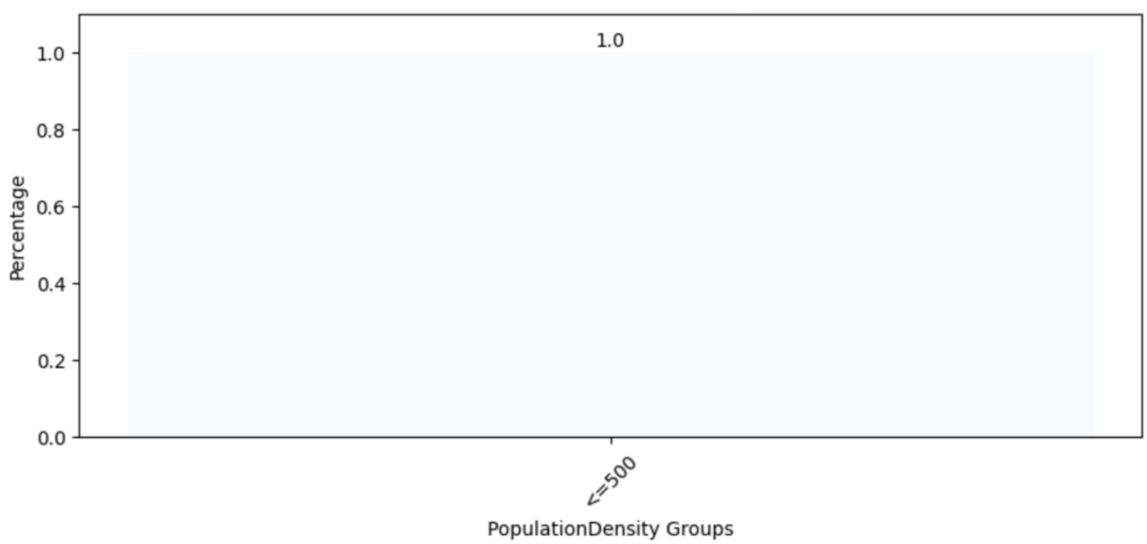


Figure 27 Proportion of Alzheimer's Patients in Different Population Density Groups

Data contain for the analysis initially

- Significance level = 0.3
- Very significant level = 0.05

Null Hypothesis (H0) – Alzheimer's outcomes are independent of Social isolation

Alternate Hypothesis (H1) – Alzheimer's outcomes are dependent on Social Isolation

Parameters data generated are:

- Chi-square value: 0.0
- P-value: 1.0
- Degrees of freedom: 0

Since p-value is higher than Significance level, Null hypothesis (H0) is Accepted. It is deducted that there is a **LOW or negligible** correlation between Social Isolation and Alzheimer's prognosis.

4.2.6 Alzheimer's prognosis with Access to Care Facilities

Data distribution for Access to Care (Distance to Care)

Table 56
Data distribution for Access to Care (Distance to Care)

Diagnosis_Code	<=1 mile	1-5 miles	5-20 miles	20+ miles
G300	20	97	59	0
G301	18	62	32	0
G308	4	17	8	0
G309	235	939	504	7

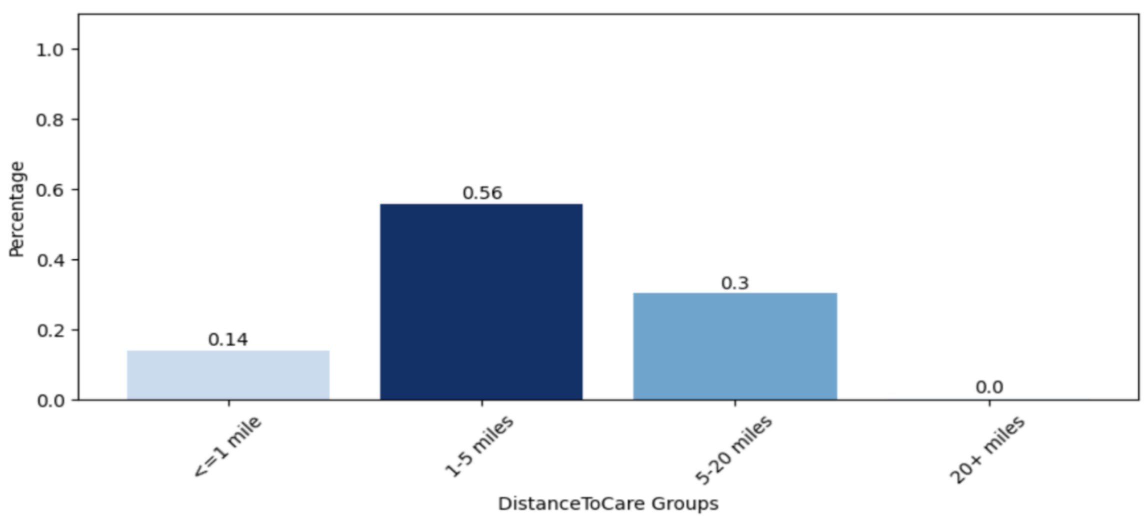


Figure 28 Proportion of Alzheimer's Patients in Different Distance To Care Groups

Data contain for the analysis initially

- Significance level = 0.3
- Very significant level = 0.05

Null Hypothesis (H0) – Alzheimer’s outcomes are independent of Access To Care

Alternate Hypothesis (H1) – Alzheimer’s outcomes are dependent on Access To Care

Parameters data generated are:

- Chi-square value: 3.4237436580338203
- P-value: 0.9451057855602022
- Degrees of freedom: 9

Since p-value is higher than Significance level, Null hypothesis (H0) is Accepted. It is deducted that there is a **LOW or negligible** correlation between Access To Care and Alzheimer’s prognosis.

4.2.7 Alzheimer’s prognosis with Education Levels

Data distribution for Education Levels (Income Disparity measured through GINI Index)

Table 57
Data distribution for Education Levels

Diagnosis_Code	**<=20%**	**20-30%**	**30-50%**
G300	0	41	135
G301	2	20	90
G308	1	3	25
G309	30	293	1362

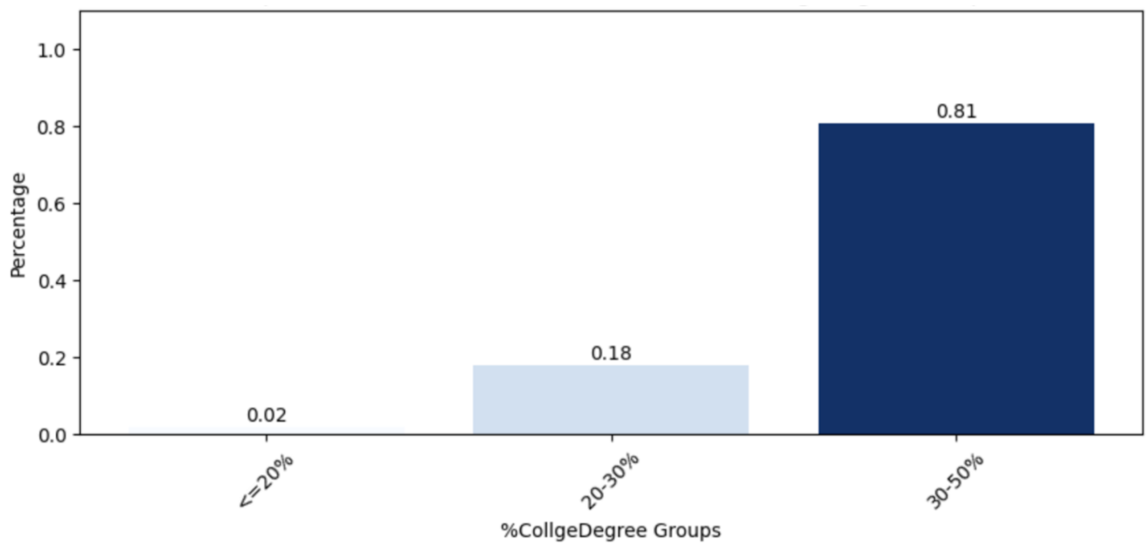


Figure 29 Proportion of Alzheimer's Patients in with College Degree

Data contain for the analysis initially

- Significance level = 0.3
- Very significant level = 0.05

Null Hypothesis (H0) – Alzheimer's outcomes are independent of Education Levels

Alternate Hypothesis (H1) – Alzheimer's outcomes are dependent on Education Levels

Parameters data generated are:

- Chi-square value: 8.16088916811134
- P-value: 0.22655158698248637
- Degrees of freedom: 6

Since p-value is lesser than Significance level, Null hypothesis (H0) is Rejected.

Since p-value is higher than Very Significant level, it is deducted that there is a **MODERATE** correlation between Education Levels and Alzheimer's prognosis.

4.2.8 Alzheimer's prognosis with Food Insecurity

Data distribution for Food Insecurity (People below poverty line and not on food stamps)

Table 58
Data distribution for Food Insecurity

Diagnosis_Code	<=10%	10-20%
G300	145	31
G301	92	20
G308	24	5
G309	1366	319

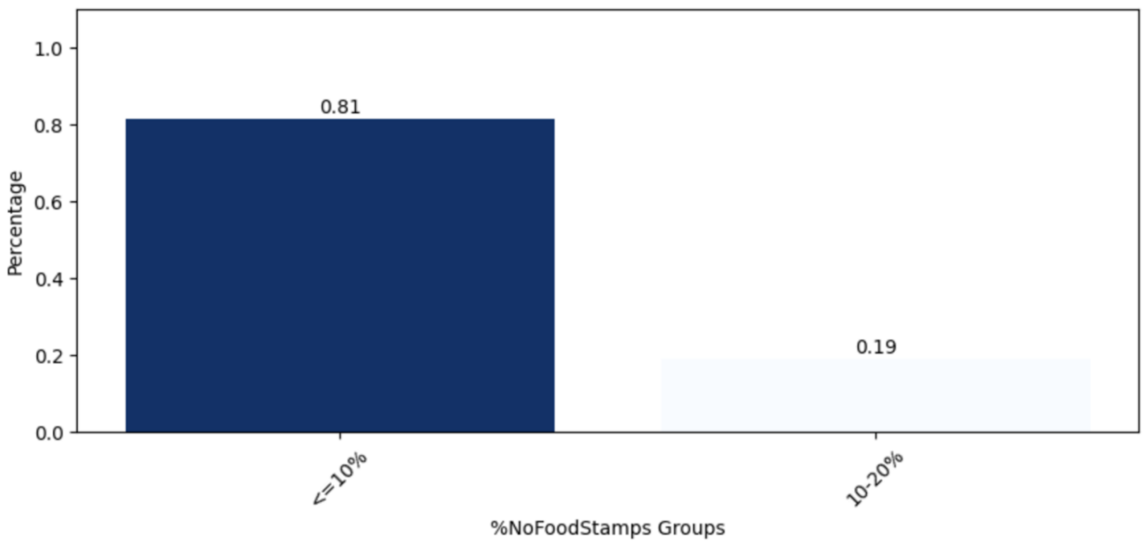


Figure 30 Proportion of Alzheimer's Patients with Food Insecurity

Data contain for the analysis initially

- Significance level = 0.3
- Very significant level = 0.05

Null Hypothesis (H0) – Alzheimer's outcomes are independent of Food Insecurity

Alternate Hypothesis (H1) – Alzheimer's outcomes are dependent on on Food Insecurity

Parameters data generated are:

- Chi-square value: 0.28741311373169864

- P-value: 0.9623780869046944
- Degrees of freedom: 3

Since p-value is higher than Significance level, Null hypothesis (H0) is Accepted. It is deducted that there is a **LOW or negligible** correlation between Food Insecurity and Alzheimer's prognosis.

4.2.9 Alzheimer's prognosis with Housing Insecurity

Data distribution for Housing Insecurity (People who spend >30% income on rent)

Table 59
Data distribution for Housing Insecurity

Diagnosis_Code	30-50%	50-70%
G300	79	97
G301	49	63
G308	16	13
G309	702	983

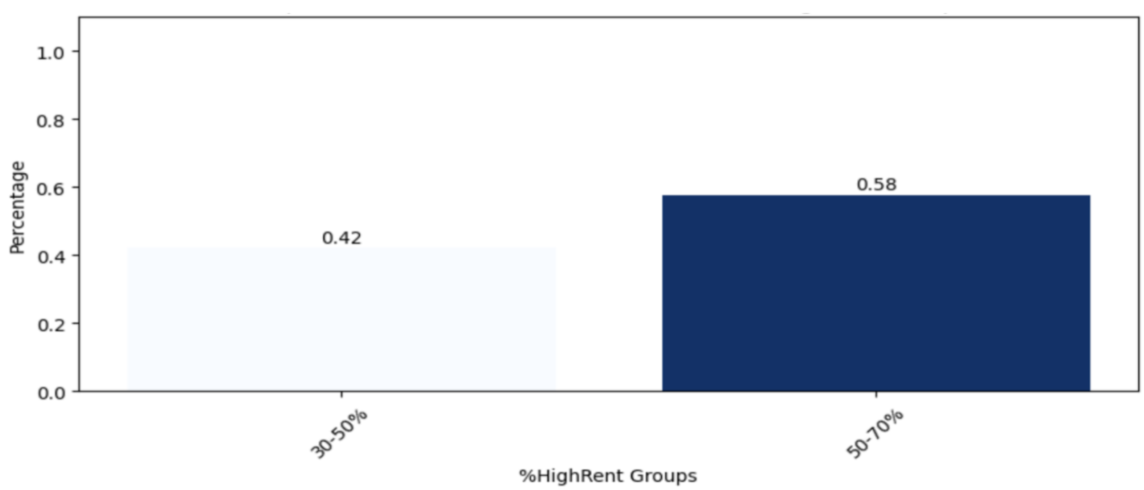


Figure 31 Proportion of Alzheimer's Patients with Housing Insecurity

Data contain for the analysis initially

- Significance level = 0.3
- Very significant level = 0.05

Null Hypothesis (H0) – Alzheimer's outcomes are independent of Housing Insecurity

Alternate Hypothesis (H1) – Alzheimer's outcomes are dependent on on Housing Insecurity

Parameters data generated are:

- Chi-square value: 2.8281952116880666
- P-value: 0.41887949062898333
- Degrees of freedom: 3

Since p-value is higher than Significance level, Null hypothesis (H0) is Accepted. It is deducted that there is a **LOW or negligible** correlation between Housing Insecurity and Alzheimer's prognosis.

4.2.10 Alzheimer's prognosis with Financial Insecurity

Data distribution for Financial Insecurity (% of people whose income is below poverty line)

Table 60
Data distribution for Financial Insecurity

Diagnosis_Code	**<=20%**	**20-50%**
G300	161	15
G301	103	9
G308	25	4
G309	1502	183

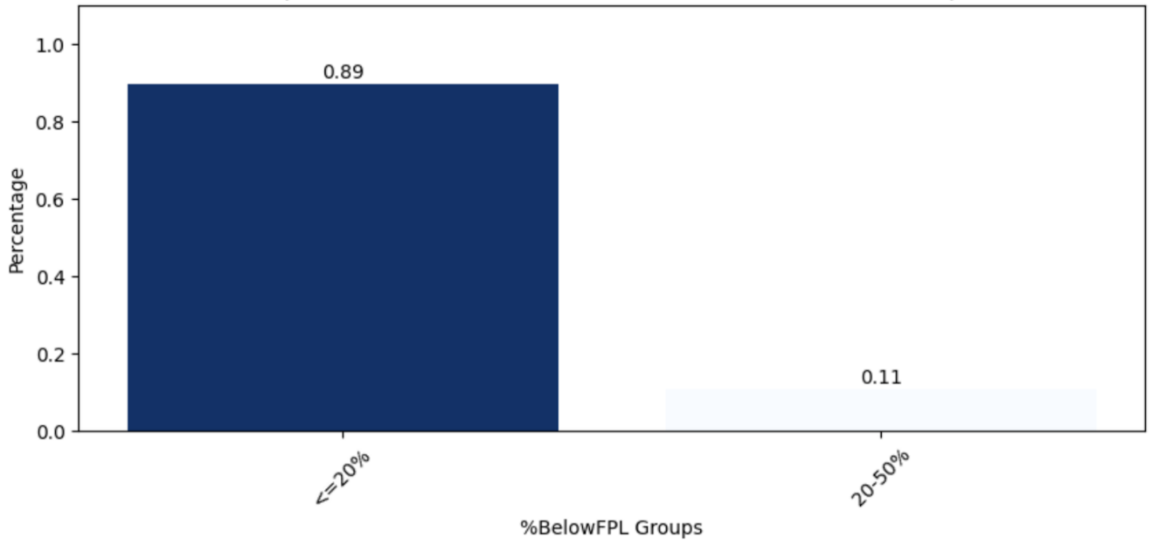


Figure 32 Proportion of Alzheimer's Patients with Financial Insecurity

Data contain for the analysis initially

- Significance level = 0.3
- Very significant level = 0.05

Null Hypothesis (H0) – Alzheimer's outcomes are independent of Financial Insecurity

Alternate Hypothesis (H1) – Alzheimer's outcomes are dependent on on Financial Insecurity

Parameters data generated are:

- Chi-square value: 2.013682320775459
- P-value: 0.5695717498902233
- Degrees of freedom: 3

Since p-value is higher than Significance level, Null hypothesis (H0) is Accepted. It is deducted that there is a **LOW or negligible** correlation between Financial Insecurity and Alzheimer's prognosis.

4.2.11 Alzheimer's prognosis with Access to Specialty Care

Data distribution for Access to Specialty Care (Community Mental health facilities)

Table 61

Data distribution for Access to Specialty Care

Diagnosis_Code	<=2 per 10000 people
G300	176
G301	112
G308	29
G309	1685

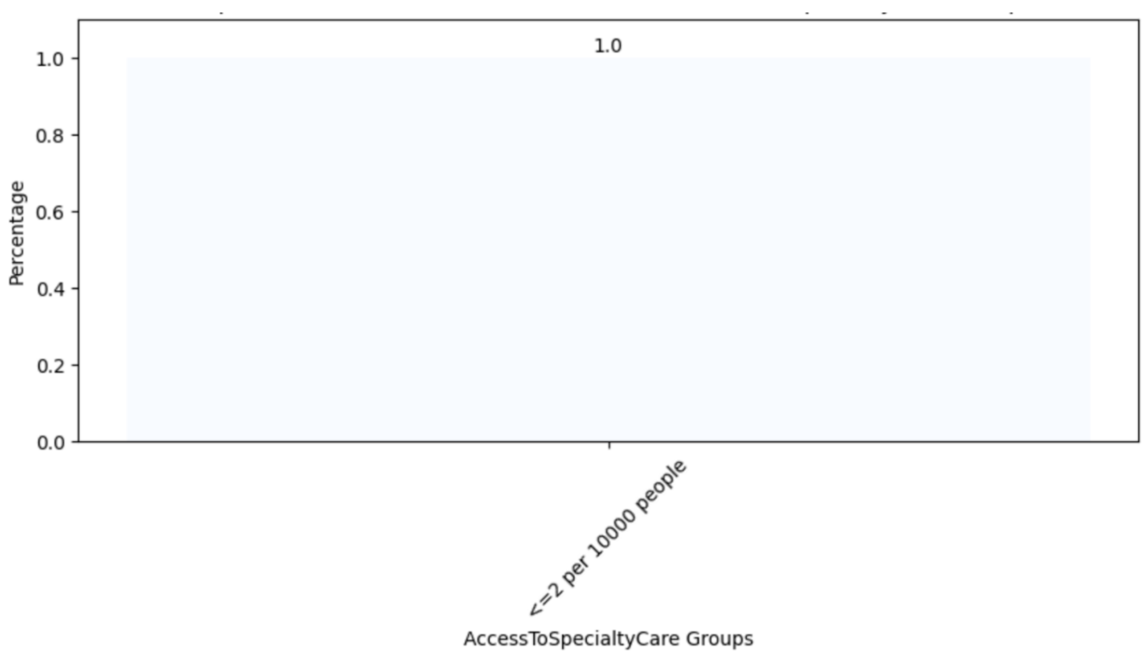


Figure 33 Proportion of Alzheimer's Patients with Access to Specialty Care

Data contain for the analysis initially

- Significance level = 0.3
- Very significant level = 0.05

Null Hypothesis (H0) – Alzheimer's outcomes are independent of Access to Specialty Care

Alternate Hypothesis (H1) – Alzheimer's outcomes are dependent on Access to Specialty Care

Parameters data generated are:

- Chi-square value: 0.0
- P-value: 1.0
- Degrees of freedom: 0

Since p-value is higher than Significance level, Null hypothesis (H0) is Accepted. It is deducted that there is a **LOW or negligible** correlation between Access to Specialty Care and Alzheimer's prognosis.

4.3 Summary of Findings

Following table 39 is the summary of Chi-Square test to determine correlation of SDOH factors on Dementia prognosis.

Table 62
Summary of Chi-Square test to determine correlation of SDOH factors on Dementia prognosis

S No	SDOH marker	p-Value	Chi Square value	Degrees of Freedom
1	Age	1.2777465700857161e-54	415.47825576203945	60
2	Gender	1.0885164235114124e-07	56.22915113353555	12
3	Transport Insecurity	0.8407250277318131	17.183634268882496	24
4	Air Quality	1.0	0.0	0
5	Social Isolation	1.0	0.0	0

6	Access to Care	0.00015210618141466904	74.89112275623064	36
7	Social Vulnerability	0.46438828368552343	11.769210003812553	12
8	Education Levels	0.014789299546059408	41.46847569334546	24
9	Food Insecurity	0.2945991127050013	14.09633099328138	12
10	Housing Insecurity	0.09251125246398216	18.838097627984773	12
11	Financial Insecurity	0.6645026452657894	9.4452520419399	12
12	Access To Specialty Care	1.0	0.0	0

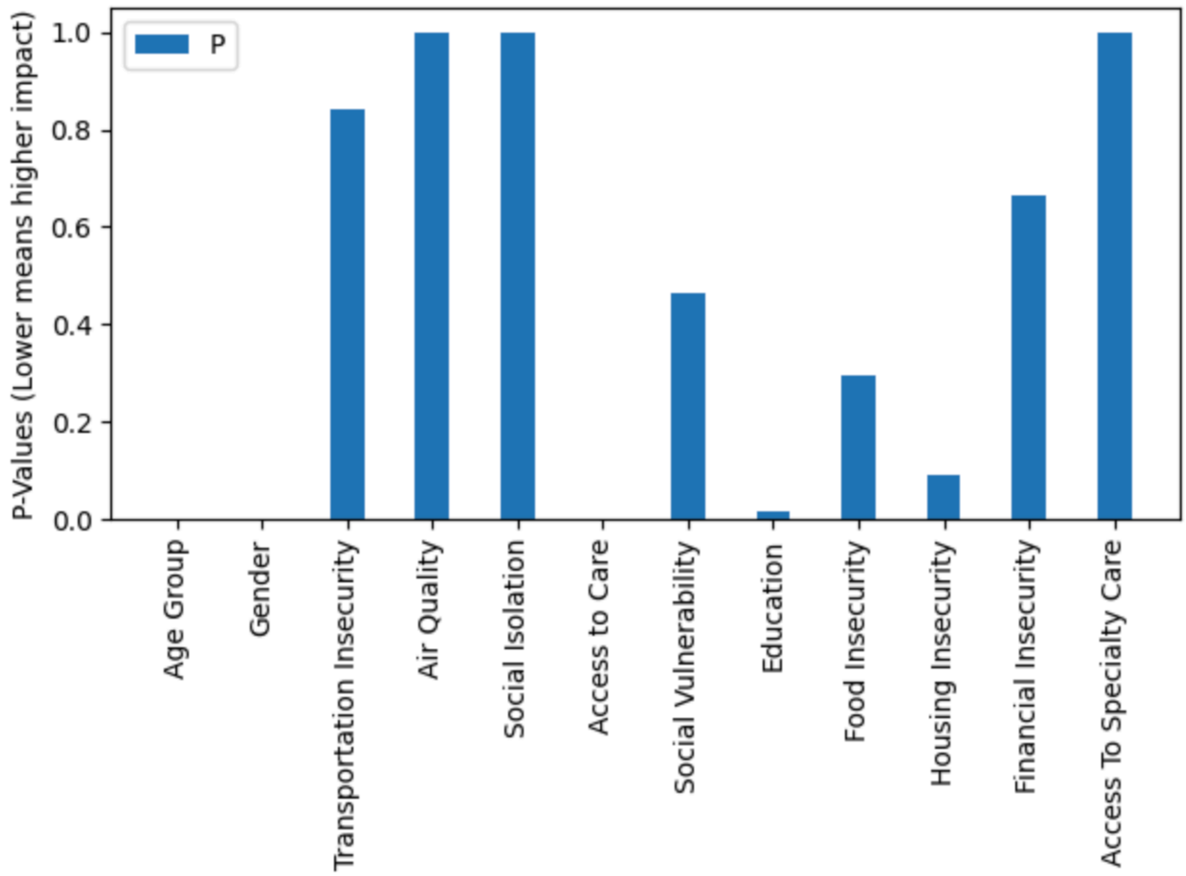


Figure 34 Impact of different SDOH marker on Dementia prognosis

Graph: Impact of different SDOH marker on Dementia prognosis

Following table 40 is the summary of Chi-Square test to determine correlation of SDOH factors on Alzheimer's prognosis.

Table 63
Summary of Chi-Square test to determine correlation of SDOH factors on Alzheimer's prognosis

S No	SDOH marker	p-Value	Chi Square value	Degrees of Freedom
1	Age	2.731505467085729e-06	53.89842825510161	15
2	Gender	0.7406693057008342	1.251548156191844	3

3	Transport Insecurity	0.2910468670611926	7.33415508650906	6
4	Air Quality	1.0	0.0	0
5	Social Isolation	1.0	0.0	0
6	Access to Care	0.9451057855602022	3.4237436580338203	9
7	Social Vulnerability	0.3018657528349465	3.649646759371645	3
8	Education Levels	0.22655158698248637	8.16088916811134	6
9	Food Insecurity	0.9623780869046944	0.28741311373169864	3
10	Housing Insecurity	0.41887949062898333	2.8281952116880666	3
11	Financial Insecurity	0.5695717498902233	2.013682320775459	3
12	Access To Specialty Care	1.0	0.0	0

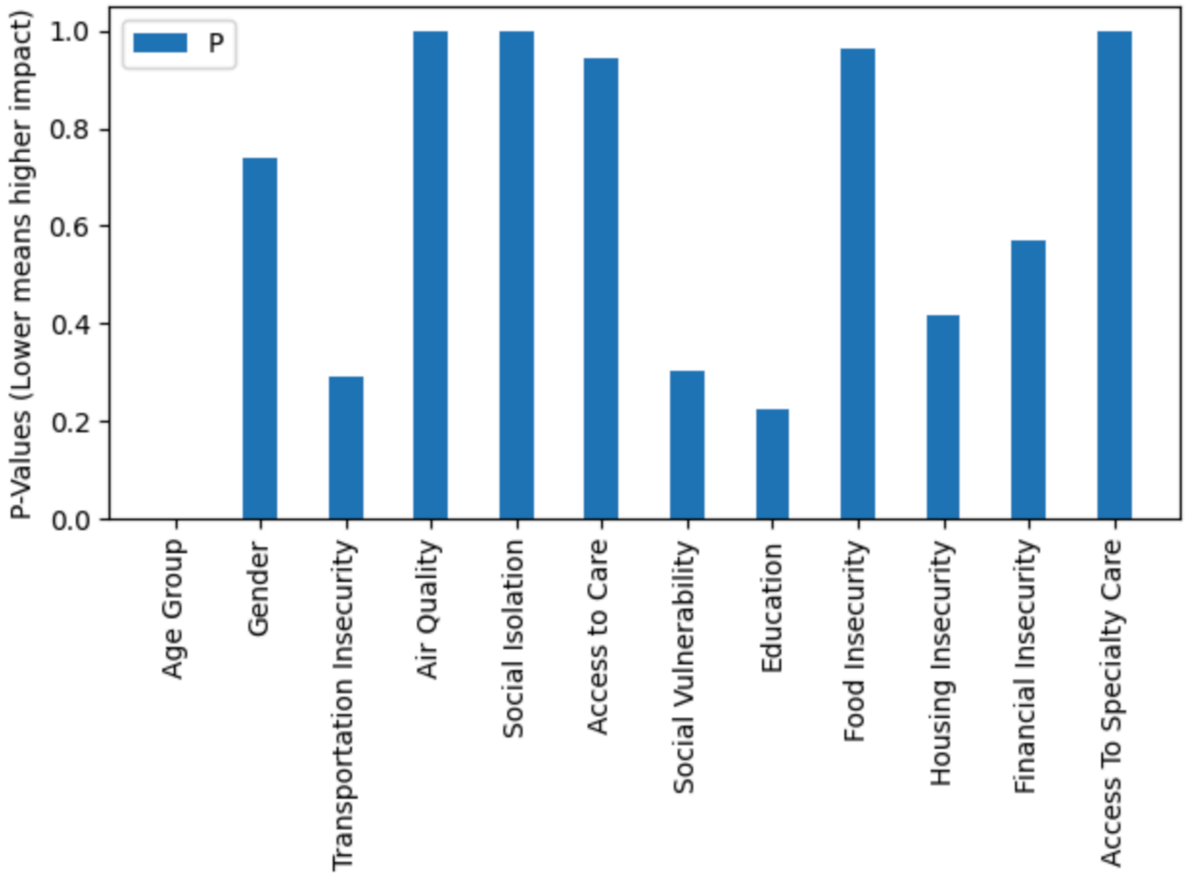


Figure 35 Impact of different SDOH marker on Alzheimer's prognosis

Graph: Impact of different SDOH marker on Alzheimer's prognosis

Significance level of 0.05 p-value will be used to distinguish between High impact and Moderate impact. A value lesser than Significance level means High Impact.

Significance level of 0.3 p-value will be used to distinguish between Moderate impact and Low impact. A value lesser than Significance level means Moderate Impact.

CHAPTER V:

DISCUSSION

5.1 Discussion of Results

Dementia & Alzheimer's are similar but not the same even though the physical ailments and the conditions may seem similar, especially memory loss. Dementia is an overall term that describes a wide range of symptoms associated with a decline in memory. It can also effect other thinking skills enough to reduce a person’s ability to perform everyday activities. Dementia is often incorrectly referred to as “senility” due to the widespread incorrect belief that serious mental decline is a normal part of aging.

Senile Dementia of the Alzheimer’s Type (SDAT), or simply called Alzheimer’s is a disease. It produces physical change in the brain. There is shrinking in some areas of the brain and widening in the others. This causes connections inside the brain to break and disrupt the brain’s electrical signals. Alzheimer’s disease accounts for 50 to 80 percent of the dementia cases and is the most type of dementia. Vascular dementia, which occurs after a stroke, is the second most common type. There are many other conditions such as thyroid problems and vitamin deficiencies that can cause symptoms of dementia. Some of these are reversible.

Following is the heatmap of Chi-Square test that was carried out for the 12 SDOH markers on two diseases viz Dementia and Alzheimer’s. Shown in below figure 34.

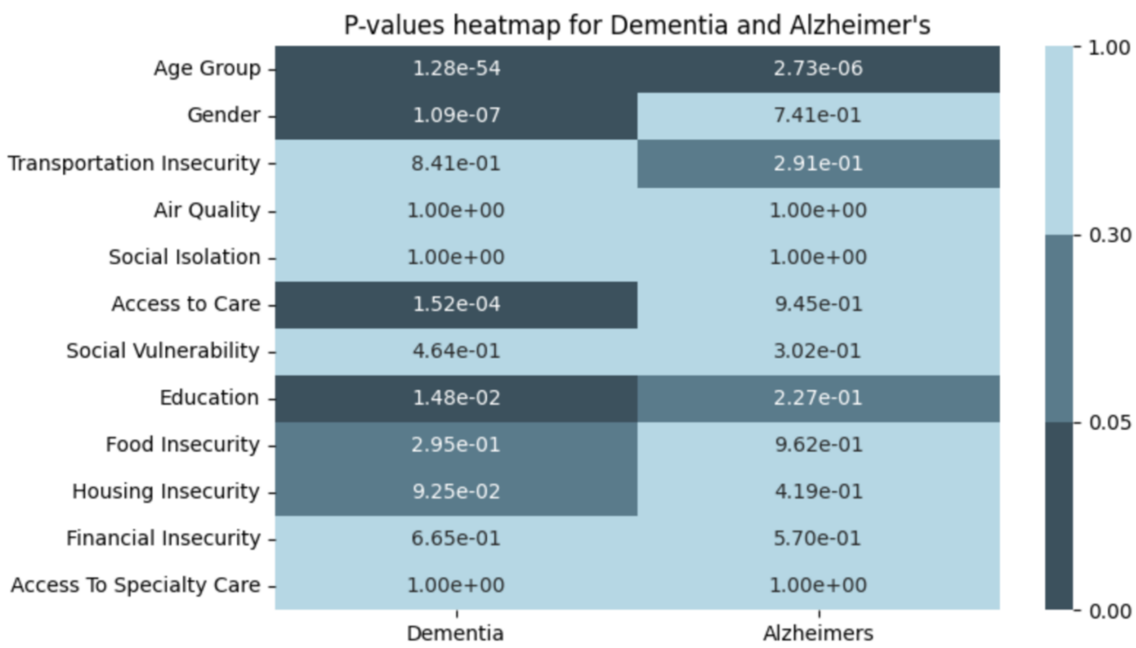


Figure 36 Impact of different SDOH marker on Dementia and Alzheimer's prognosis

Significance level of 0.05 p-value will be used to distinguish between High impact and Moderate impact. A value lesser than this Significance level means HIGH Impact of the SDOH marker on disease prognosis.

Significance level of 0.3 p-value will be used to distinguish between Moderate impact and Low impact. A value lesser than this Significance level means MODERATE Impact of the SDOH marker on disease prognosis.

A p-value of greater than 0.3 means LOW Impact of the SDOH marker on disease prognosis.

It is observed that SDOH markers have a higher influence on Dementia prognosis as compared to Alzheimer's. This conforms to the medical literature which state that dementia is caused by imbalance in nervous system and could be affected by genetic composition combined with external factors, while Alzheimer's is more of a genetic disease due to damage of cells in nervous system in higher ages, typically 60+ years

5.2 Discussion of Dementia Disease Impacts

Based on Chi-Square test, following is the summary results in order of their extent of impact of different SDOH factors on Dementia prognosis given in the table 41.

Table 64
Influence of SDOH markers on Dementia prognosis

S No	SDOH marker	P-Value from Chi-Square test on the data sample	Impact of SDOH marker on Dementia prognosis (inference)	Comment
1	Age	1.2777465700857161e-54	High	
2	Gender	1.0885164235114124e-07	High	
3	Access to Care	0.00015210618141466904	High	
4	Education Levels	0.014789299546059408	High	
5	Housing Insecurity	0.09251125246398216	Moderate	
6	Food Insecurity	0.2945991127050013	Moderate	
7	Social Vulnerability	0.46438828368552343	Low	
8	Financial Insecurity	0.6645026452657894	Low	
9	Transport Insecurity	0.8407250277318131	Low	
10	Air	1.0	Could not be	The dispersion

	Quality		determined	in data sample was not enough to determine an impact
11	Social Isolation	1.0	Could not be determined	The dispersion in data sample was not enough to determine an impact
12	Access To Specialty Care	1.0	Could not be determined	The dispersion in data sample was not enough to determine an impact

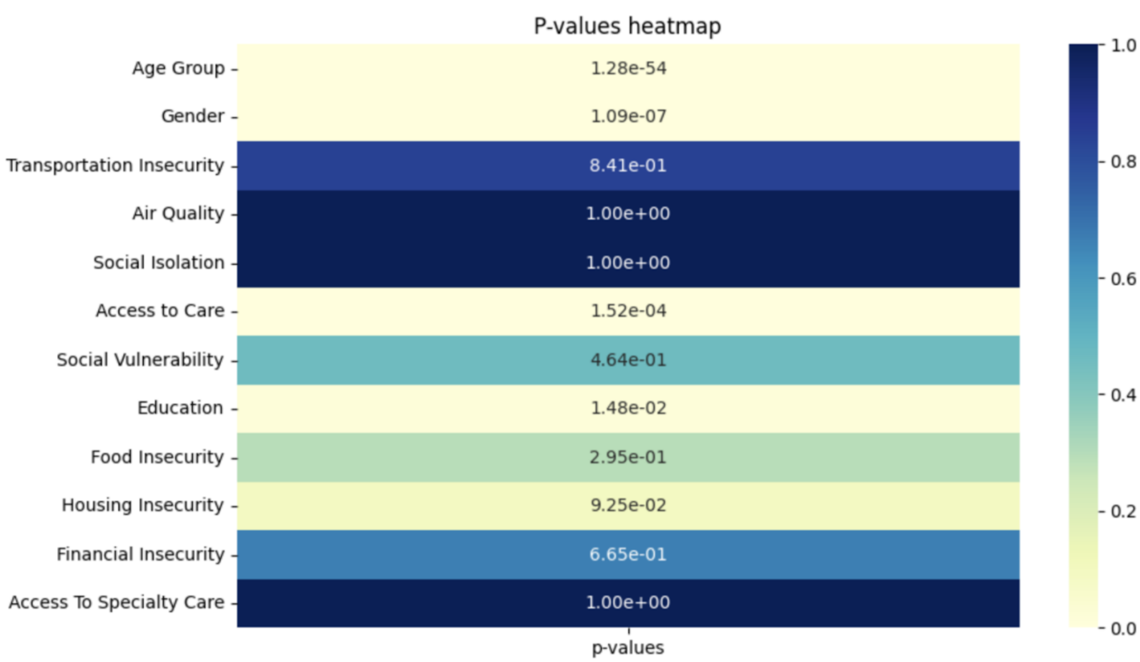


Figure 37 Chi-Square test of different SDOH markers on Dementia

P-value of zero (Yellow color) shows complete correlation between SDOH marker and Dementia prognosis.

P-value of one (Blue color) shows no correlation between SDOH marker and Dementia prognosis.

98% of dementia cases are determined after the age of 60 years. Hence the disease occurs later in life and the probability increases incrementally in the age groups of 7i-80 years and 80+ years. This shows a HIGH correlation of Age to Dementia prognosis with advancing age in patients. Age correlation for Dementia is 390% higher when Age >=70 years and 600% higher when Age >=80

58% of dementia cases are with Females. Hence the disease is more common in female group. Gender correlation for Dementia is 50% higher when Gender is 'Female' (Females have 50% higher probability of Dementia prognosis as compared to Males)

Based on the above analysis, it can be deduced that Age, Gender, Education Levels and Access to Care (primary care) have the highest influence on Dementia prognosis.

Following that, it can be deduced that Housing Insecurity and Food Insecurity have the next level of influence on Dementia prognosis.

It can also be deduced that Social Vulnerability, Financial Insecurity and Transport Insecurity have the lowest level of influence on Dementia prognosis.

Below table 42 summarizes the impact of different SDOH markers on Dementia that was derived from literature review (theoretical) as compared to what was observed from the sample dataset using Chi-Square test.

Table 65
Influence of SDOH markers on Dementia prognosis

S No	SDOH marker	Impact based framework (based on Literature review)	Inference from Chi-Square test	Comment
1	Age	High	High	
2	Gender	High	High	
3	Access to Care	High	High	
4	Education Levels	High	High	
5	Housing Insecurity	High	Moderate	
6	Food Insecurity	High	Moderate	
7	Social Vulnerability	Moderate	Low	
8	Financial Insecurity	High	Low	
9	Transportation Insecurity	High	Low	
10	Air Quality	Low	Could not be determined	The dispersion in data sample was not enough to determine an impact
11	Social Isolation	High	Could not be determined	The dispersion in data sample was not enough to determine an impact

12	Access To Specialty Care	High	Could not be determined	The dispersion in data sample was not enough to determine an impact

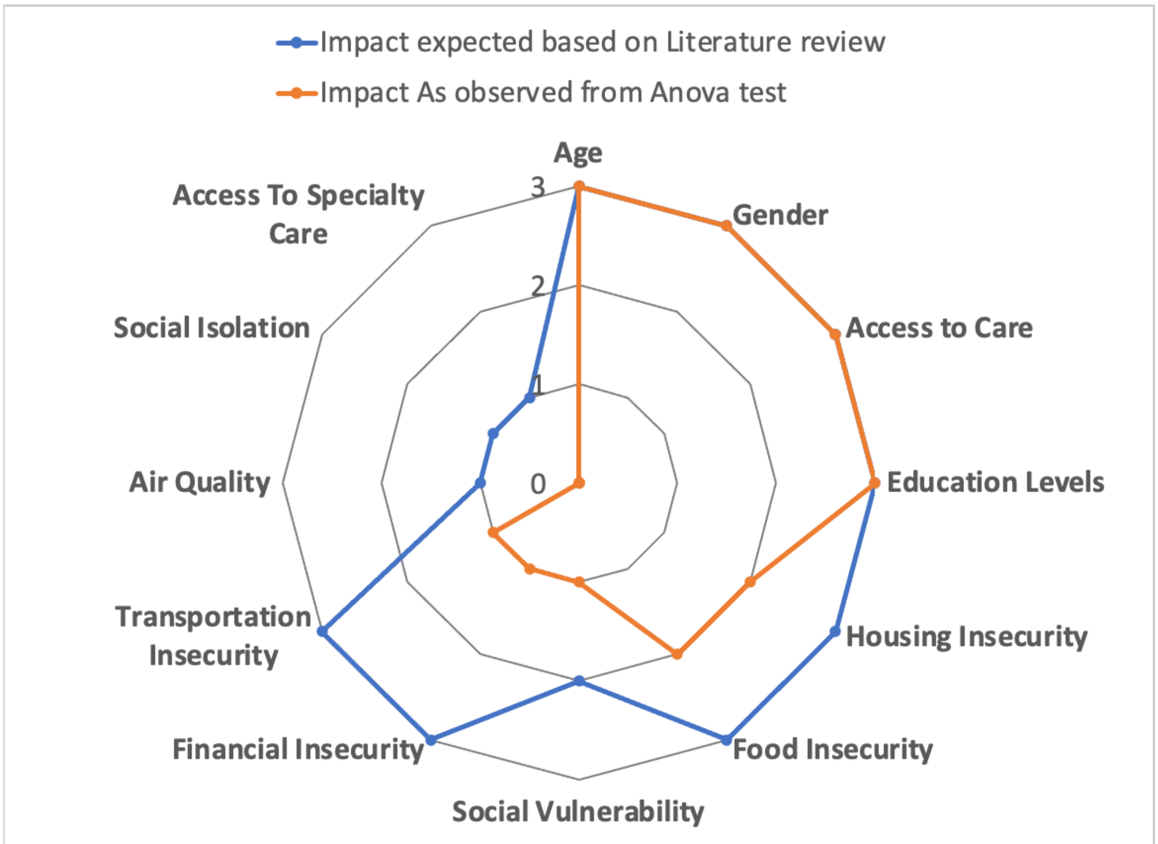


Figure 38 Impact assessment of SDOH markers as per Literature research versus as observed from the sample dataset

There is a certain level of variation observed in the impact of SDOH markers on Dementia prognosis that was determined from literature review versus what has been determined from sample data using Chi-Square test. This is most likely due to missing data in the sample dataset. The sample data was collected for confirmed Dementia cases and did not include negative outcomes and hence there would also be a data insufficiency in the sample data.

5.2 Discussion of Alzheimer's Disease Impacts

Based on Chi-Square test, following is the summary results in order of their extent of impact of different SDOH factors on Alzheimer's prognosis.

Table 66
Impact of different SDOH factors on Alzheimer's prognosis

S No	SDOH marker	P-Value from Chi-Square test on the data sample	Impact of SDOH marker on Alzheimer's prognosis (inference)	Comment
1	Age	2.731505467085729e-06	High	
2	Financial Insecurity	0.22655158698248637	Moderate	
3	Access to Care	0.2910468670611926	Moderate	
4	Social Vulnerability	0.3018657528349465	Moderate	
5	Air Quality	0.41887949062898333	Low	
6	Social Isolation	0.5695717498902233	Low	
7	Gender	0.7406693057008342	Low	
8	Food Insecurity	0.9451057855602022	Low	
9	Transport Insecurity	0.9623780869046944	Low	
10	Housing	1.0	Could not be	The dispersion in

	Insecurity		determined	data sample was not enough to determine an impact
11	Education Levels	1.0	Could not be determined	The dispersion in data sample was not enough to determine an impact
12	Access To Specialty Care	1.0	Could not be determined	The dispersion in data sample was not enough to determine an impact

P-value of zero (Yellow color) shows complete correlation between SDOH marker and Dementia prognosis.

P-value of one (Blue color) shows no correlation between SDOH marker and Dementia prognosis.

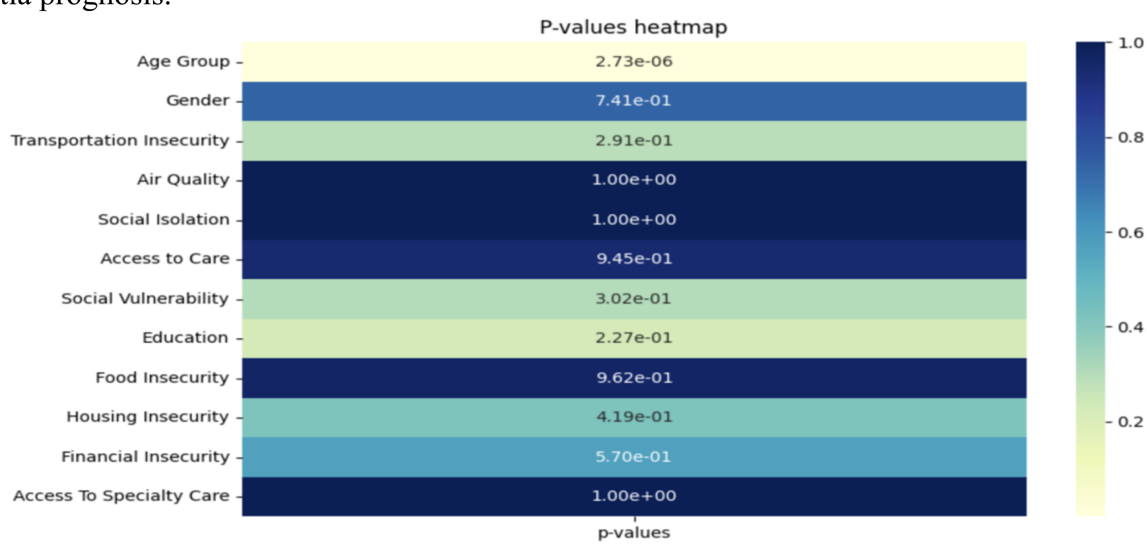


Figure 39 Chi-Square test of different SDOH markers on Alzheimer's

99% of Alzheimer’s cases are determined after the age of 60 years. Hence the disease occurs later in life and the probability increases incrementally in the age groups of 71-80 years and 80+ years. As compared to Dementia patients, Age correlation for Alzheimer’s becomes 10% higher when Age >=80 years and 33% higher when Age >=90. This shows a HIGH correlation of Age to Alzheimer’s prognosis with advancing age in patients

58% of Alzheimer’s cases are with Females. Hence the disease is more common in female group. As compared to Dementia patients, Gender correlation for Alzheimer’s becomes 8% higher when Gender is ‘Female’

Based on the above analysis, it can be deduced that Age, Gender, Education Levels and Access to Care (primary care) have the highest influence on Dementia prognosis.

Following that, it can be deduced that Housing Insecurity and Food Insecurity have the next level of influence on Dementia prognosis.

It can also be deduced that Social Vulnerability, Financial Insecurity and Transport Insecurity have the lowest level of influence on Dementia prognosis.

Following table summarizes the impact of different SDOH markers on Alzheimer’s that was derived from literature review (theoretical) as compared to what was observed from the sample dataset using Chi-Square test.

Table 67

Influence of SDOH markers on Alzheimer’s prognosis

S No	SDOH marker	Impact based framework (based on Literature review)	Inference from Chi-Square test	Comment
1	Age	High	High	
2	Gender	High	Low	

3	Access to Care	High	Moderate	
4	Housing Insecurity	High	Could not be determined	The dispersion in data sample was not enough to determine an impact
5	Transportation Insecurity	High	Low	
6	Social Isolation	High	Low	The dispersion in data sample was not enough to determine an impact
7	Access To Specialty Care	High	Could not be determined	The dispersion in data sample was not enough to determine an impact
8	Education Levels	Moderate	Could not be determined	
9	Food Insecurity	Moderate	Low	
10	Social Vulnerability	Moderate	Moderate	
11	Financial Insecurity	Moderate	Moderate	
12	Air Quality	Low	Low	

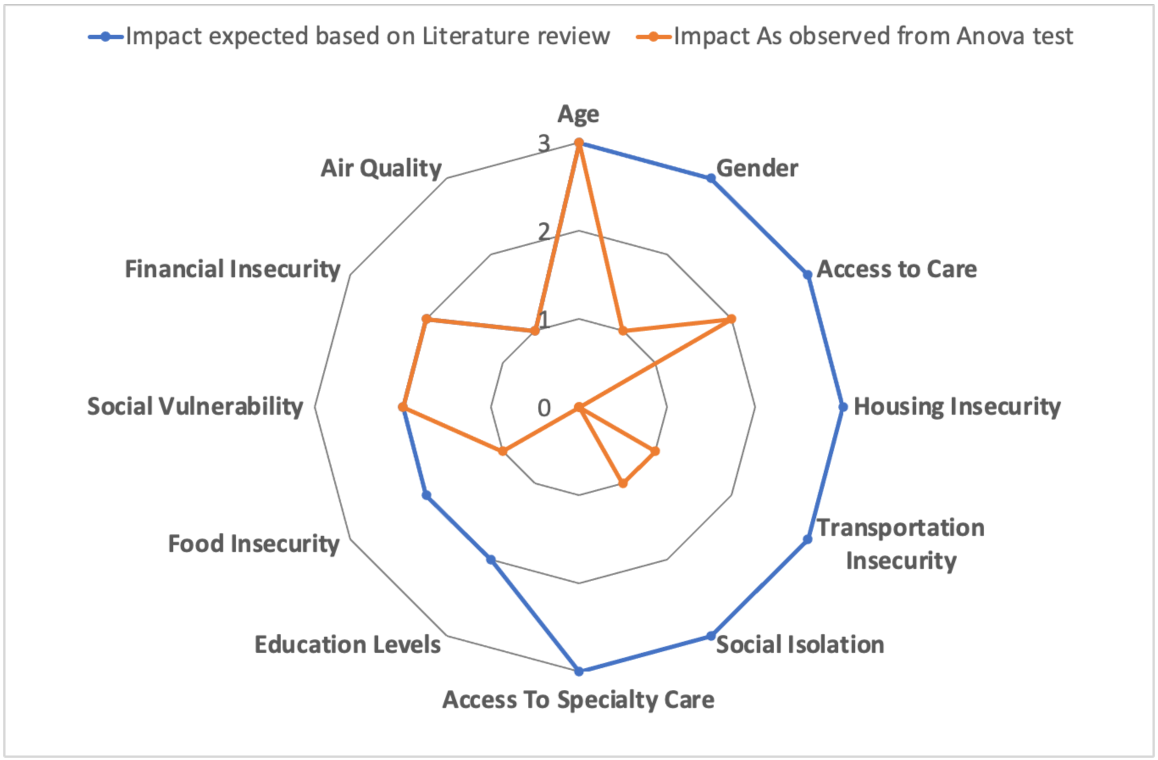


Figure 40 Impact assessment of SDOH markers as per Literature research versus as observed from the sample dataset

There is a high level of variation observed in the impact of SDOH markers on Alzheimer's prognosis that was determined from literature review versus what has been determined from sample data using Chi-Square test. This is most likely due to missing data in the sample dataset. The sample data was collected for confirmed Alzheimer's cases and did not include negative outcomes and hence there would also be a data insufficiency in the sample data.

CHAPTER VI:

SUMMARY, IMPLICATIONS, AND RECOMMENDATIONS

6.1 Summary

Social determinants have a big role in an individual's health and wellness. This is rather well known through traditional experiences and is often used by care providers and patients in different situations of healthcare continuum. However most of the applications are based on prior knowledge and experience and not necessarily data and analytics based.

Many public and private organizations like WHO, CDC, NIH, CMS, US Census Bureau, non-government organizations, and industry have made good progress in collecting demographic data and publishing it publicly for research. The new SDOH framework proposed in this research draws many salient points from these efforts, combines it with a novel disease-focussed approach of SDOH application, and goes very granular at a zip code and community level to understand the patient situation better. This framework is data-based, pragmatic, evidence-based and personalized which can transform healthcare practices with further research and innovations in its application and use cases.

New popular healthcare paradigms propose adoption of Whole health practices to improve health outcomes and reduce cost of healthcare. These are some of the building blocks of Value Based Care models which US government is trying to push to increase health equity, provide universal healthcare and still reduce healthcare cost burden. Hence a judicious implementation can significantly impact the adoption of the Value Based Care in US.

6.2 Implications

A comprehensive disease prognosis and management plan should ideally involve a combination of clinical factors and non-clinical factors (social determinants) to understand the health situation of an individual. In absence of such an understanding, Providers are unable to take the additional legal and financial risks that are part of Value based care ecosystem. Developing a comprehensive understanding of Whole health practices using real world data, statistics and insights could help in accelerating the adoption of Value Based Care as per the guidance provided in Affordable Care Act (ACA) 2010.

A comprehensive understanding of Whole health is mostly missing at the point of care for Care providers. Providers have a very little time during the patient encounter to draw up a complete understanding of patient situation. These professionals try to make up for this gap through regular screenings which tends to be a manual process, may be incomplete, prone to errors and costly. Insights derived from SDOH framework in the backend (eg prognosis risk models, automated recommendations etc) may be integrated into clinical decision support systems to improve health outcomes for the patients and providers.

Personalized healthcare is the next frontier for healthcare industry. There is a lot of research happening in government and private organizations to understand the complex human body ecosystem, how every human body is different from the other, and how each body reacts to external stimuli like the social determinants. Different social determinants have a different impact on different individuals. Further research on these factors and adding more factors could provide a framework that could be used in personalized healthcare and wellness.

6.3 Recommendations for Future Research

Current study focused on social determinants data as provided by Agency for Healthcare Research and Quality (AHRQ) for the state of Nevada. While this a credible agency and operated at a national level in US, but there are also state and local agencies that collect similar data and sometimes even more attributes that are not covered by AHRQ. Some of the examples are Open Street Maps (https://extract.bbbike.org/), Suite of Food Security Indicators - Datasets - "FAO catalog", Housing Density index (https://earthexplorer.usgs.gov/), World Health Organization Air quality database 2022 (who.int), USGS Water-Quality Database, and Geospatial database (https://www.arcgis.com/) among others. Using data from local agencies along with AHRQ could improve the granularity and quality of data, and hence improve the correlation analysis.

There are data gaps in SDOH as many agencies are collecting this data in siloed manner. There need to be consistency in the measurement as well as the metric they follow. Still there are missing data for many zipcode locations that allows loss of granularity and hence the impact of SDOH markers cannot be determined precisely. Synthetic data augmentation techniques can help in largely addressing the gaps. Eg Water quality is not consistently and regularly measured in all counties of US, while it is common knowledge that sustained exposure to arsenic presence in drinking water can cause stenting, cognitive decline, cancer, and various other diseases. New technologies like Generative AI can be utilized to create synthetic data for such markers by assessing a correlation between similar regions, geography, and supply systems as that of the principal location.

In this research we covered a detailed assessment of SDOH markers on two diseases viz Dementia and Alzheimer's. However, the SDOH framework is generic enough and further analysis and research on other common chronic diseases like cancer, cardiovascular, diabetes among others. Such scientific research could provide evidence for certain beliefs eg impact of a person's lifestyle on dementia prognosis, and even help disperse some myths eg. Race has a predominant role in dementia prognosis as compared to other social determinants.

Current research focused on 12 SDOH markers within the scope as per the proposed framework. However human body is complex, and every human body is different from the other. Different social determinants have a different impact on different individuals. Further research on these factors and adding more factors could provide a framework that could be used in personalized healthcare and wellness.

An informal survey from the Providers indicate that they are aware of some of the correlations of social determinants as common knowledge and use these factors for disease diagnosis, prediction, and management at the point of care. However, structured data-based evidence for some of the markers is not available along with their extent of impact on different diseases for different people. Hence, they are unable to use the same for taking on financial risks in the Value based care system. Formal survey-based research across a cross section of Providers would help provide inputs and feedback from the clinical settings on what is missing in the holistic adoption of non-clinical factors (social determinants) along with clinical factors in a comprehensive prognosis and management plan.

Much of research and analytics on SDOH happens in backend systems and available to claim adjudicators, health plan underwriters, policy makers etc. However, it has been difficult to make these insights available to providers at the point of care due to inconsistency across industry of the definition, understanding and interpretation of these models. There is potential scope to research and develop SDOH ontology that is universally accepted. This can significantly improve the integration of these insights into EHR and clinical decision support systems thereby improving health outcomes.

There have been assumptions that a particular social determinants model would have similar impact for every disease and during different stages of disease continuum eg, Air quality would have a same impact for hypertension and for cancer, and also during early stages of a disease (onset) versus mid stages and during advanced stages. The universality assumption needs to be assessed if these are helping physicians use them in clinical settings. This needs to be researched further which will provide the ultimate proof of their utility.

6.4 Conclusion

Further research is recommended on this subject and similar topics as highlighted in the recommendations section.

I sincerely believe that the approach mentioned in this research can help people live healthier lives and can help make the health system work better for everyone.

In conclusion, the research underscores the pivotal role of social determinants of health (SDOH) in shaping an individual's well-being, emphasizing the need for data-driven and analytics-based applications to enhance healthcare practices. The proposed SDOH framework, grounded in a disease-focused approach, offers a granular understanding at the zip code and community levels, promoting personalized and evidence-based healthcare practices. The implications of this framework extend to the adoption of Whole Health practices and Value-Based Care models, aligning with the Affordable Care Act's guidance. The study highlights the importance of integrating non-clinical factors, particularly SDOH, into comprehensive disease prognosis and management plans. Recommendations for future research include leveraging data from local agencies, addressing data gaps through synthetic data augmentation, expanding the analysis to other chronic diseases, and exploring the potential of SDOH ontology for universal acceptance. The research calls for a collaborative effort to bridge gaps in understanding non-clinical factors in healthcare, providing a foundation for informed decision-making and improved health outcomes.

REFERENCES

Alzheimers's Association. Facts and Figures; 2021; Published online at https://www.alz.org/alzheimers-dementia/facts-figures

Ahmed, Z., Mohamed, K., Zeeshan, S. and Dong, X., 2020. Artificial intelligence with multi-functional machine learning platform development for better healthcare and precision medicine. Database, 2020, p.baaa010.

Agency for Healthcare Research and Quality (AHRQ). Social Determinants of Health Database. Published online at https://www.ahrq.gov/sdoh/data-analytics/sdoh-data.html

American Cancer Society. The Costs of Cancer 2020 Edition. Published online at https://www.fightcancer.org/sites/default/files/National%20Documents/Costs-of-Cancer-2020-10222020.pdf

American Psychiatric Association. The Economic Cost of Depression is Increasing; Direct Costs are Only a Small Part. Published online at https://www.psychiatry.org/News-room/APA-Blogs/The-Economic-Cost-of-Depression-is-Increasing

Cairns, C. and Kang, K., 2022. National hospital ambulatory medical care survey: 2019 emergency department summary tables; DOI: https://dx.doi.org/10.15620/cdc:121911.

Centers for Disease Control and Prevention (CDC). Social Determinants of Health (SDOH) and PLACES Data; Published online at https://www.cdc.gov/places/social-determinants-of-health-and-places-data/index.html

Centers for Disease Control and Prevention (CDC). ICD codes. 2021. Published online at https://www.cdc.gov/nchs/icd/icd-10-cm.htm

Centers for Medicare and Medicaid Services. MACRA. 2015. Published online at https://www.cms.gov/Medicare/Quality-Initiatives-Patient-Assessment-Instruments/Value-Based-Programs/MACRA-MIPS-and-APMs/MACRA-MIPS-and-APMs

Chris Bethell. How Health Plans Can Use SDoH Data to Succeed in Value-Based Care. https://www.spectramedix.com, Feb 2020

Division for Heart Disease and Stroke Prevention (DHDSP) of Centers for Disease Control and Prevention (CDC); Heart Disease Facts; Published 2021; https://www.cdc.gov/heartdisease/facts.htm

Division of National Institute on Aging (DNIA) of National Institute of Health (NIH). Alzheimer's Disease Fact Sheet; Published 2021; 'https://www.nia.nih.gov › alzheimers-disease-fact-sheet'

Division of Neurological Disorders and Stroke (DNDS) of National Institute of Health (NIH), US, Publication No. 15-5595; Parkinson's Disease: Challenges, Progress, and Promise; Published September 30, 2015; https://www.ninds.nih.gov/current-research/focus-disorders/focus-parkinsons-disease-research/parkinsons-disease-challenges-progress-and-promise

Drachenberg J. Patient-first language can affect patient care. Relias Media. September 29, 2021.

EHR Intelligence. Meaningful Requirements; 2017. Published online at https://ehrintelligence.com/news/top-8-goals-of-stage-3-meaningful-use-proposed-rule

Lin, E.H., Katon, W., Von Korff, M., Rutter, C., Simon, G.E., Oliver, M., Ciechanowski, P., Ludman, E.J., Bush, T. and Young, B., 2004. Relationship of depression and diabetes self-care, medication adherence, and preventive care. Diabetes care, 27(9), pp.2154-2160.

Espinosa, O., Coffee-Borden, M.P.P., Brandon Coffee-Borden, M.P.P., Bakos, A. and Nweke, O., 2016. Implementation of the National Partnership for Action to End Health Disparities: a three-year retrospective. Journal of Health Disparities Research and Practice, 9(6), p.3.

Strickland, E., 2022. Are you still using real data to train your AI. IEEE Spectrum.

eCQI Resource Center. EMR versus EHR Difference, Published online at https://www.healthit.gov/buzz-blog/electronic-health-and-medical-records/emr-vs-ehr-difference

eCQI Resource Center, FHIR - Fast Healthcare Interoperability Resources. Published online at https://ecqi.healthit.gov/fhir

Freij, M., Dullabh, P., Lewis, S., Smith, S.R., Hovey, L. and Dhopeshwarkar, R., 2019. Incorporating social determinants of health in electronic health records: qualitative study of current practices among top vendors. JMIR medical informatics, 7(2), p.e13849.

Fernandez G.; The intersection of mental health and chronic disease. Johns Hopkins Bloomberg School of Public Health. December 16, 2021. https://publichealth.jhu.edu/2021/the-intersection-of-mental-health-and-chronic-disease

Health Care Payment Learning & Action Network; HCPLAN Measurement Report; 2021; Milbank Memorial Fund. Published online at https://hcp-lan.org/apm-measurement-effort/2020-2021-apm/

Healthy People 2030 report from Department of Health and Human Services (DHSS). Social determinants of health and health equity. 2021. Published online at https://health.gov/healthypeople/priority-areas/social-determinants-health

Holman, H.R., 2020. The relation of the chronic disease epidemic to the health care crisis. ACR open rheumatology, 2(3), pp.167-173.

Holzmeyer, C., 2021. Beyond 'AI for Social Good'(AI4SG): social transformations—not tech-fixes—for health equity. Interdisciplinary Science Reviews, 46(1-2), pp.94-125.

Houlihan, J. and Leffler, S., 2019. Assessing and addressing social determinants of health: a key competency for succeeding in value-based care. Primary Care: Clinics in Office Practice, 46(4), pp.561-574.

Institute of Medicine. Capturing Social and Behavioral Domains and Measures in Electronic Health Records: Phase 2. 2014. Washington, DC: The National Academies Press. https://doi.org/10.17226/18951.

International Social Security Association. (2023). Information and communication technology- Guideline 91. Electronic health record system. Retrieved March 15, 2023, from https://ww1.issa.int/guidelines/ict/180156

Joseph J. Gallo, M.D., M.P.H. ; Multimorbidity and Mental Health; Published February 10, 2017; American Association for Geriatric Psychiatry Volume 25 Issue 5; DOI:https://doi.org/10.1016/j.jagp.2017.02.007

Jessica S Ancker, Min-Hyung Kim, Yiye Zhang, Yongkang Zhang, Jyotishman Pathak. The potential value of social determinants of health in predicting health outcomes, Journal of the American Medical Informatics Association, Volume 25, Issue 8, August 2018, Pages 1109–1110

Kasthurirathne, S.N., Vest, J.R., Menachemi, N., Halverson, P.K. and Grannis, S.J., 2018. Assessing the capacity of social determinants of health data to augment predictive models identifying patients in need of wraparound social services. Journal of the American Medical Informatics Association, 25(1), pp.47-53.

Lans, A., Kanbier, L.N., Bernstein, D.N., Groot, O.Q., Ogink, P.T., Tobert, D.G., Verlaan, J.J. and Schwab, J.H., 2023. Social determinants of health in prognostic machine learning models for orthopaedic outcomes: a systematic review. Journal of Evaluation in Clinical Practice, 29(2), pp.292-299.

Lewis, G. et al. (2013). 'How Health Systems Could Avert "Triple Fail" Events That Are Harmful, Are Costly, And Result In Poor Patient Satisfaction', Health Affairs, 32(4), pp. 669–676.

Lin, E., Dave, G. and Kshirsagar, A.V., 2022. The New Kidney-Focused Companies: A Privatized Approach to Value-Based Care and Addressing Social Determinants of Health. Journal of the American Society of Nephrology.

Martin, A., 2022. Payor Influence on Social Determinants of Health (Doctoral dissertation, The College of St. Scholastica).

Mental Health America. Access to Care Cata 2022. Published online at https://mhanational.org/issues/2022/mental-health-america-access-care-data

MD Interactive. MIPS 2020 Final Rule Now Available - What Does it Mean for You? Nov 2019. Published online at https://mdinteractive.com/mips-blog/mips-2020-final-rule-now-available-what-does-it-mean-you

Motamedi, M., Sakharnykh, N. and Kaldewey, T., 2021. A data-centric approach for training deep neural networks with less data. arXiv preprint arXiv:2110.03613.

Melzer, S.M., 2022. Addressing social determinants of health in pediatric health systems: balancing mission and financial sustainability. Current Opinion in Pediatrics, 34(1), pp.8-13.

Cantor, M.N., Chandras, R. and Pulgarin, C., 2018. FACETS: using open data to measure community social determinants of health. Journal of the American Medical Informatics Association, 25(4), pp.419-422.

National Center for Health Statistics (NCHS) of Centers for Disease Control and Prevention (CDC); Cerebrovascular Disease or Stroke; Published 2021; https://www.cdc.gov/nchs/fastats/stroke.htm

National Center for Chronic Disease Prevention and Health Promotion (NCCDPHP) of Centers for Disease Control and Prevention (CDC). Health and economic costs of chronic diseases. Published online at https://www.cdc.gov/chronicdisease/about/costs/index.htm

National Center for Chronic Disease Prevention and Health Promotion (NCCDPHP) of Centers for Disease Control and Prevention (CDC). Community Health Programs. Published online at https://www.cdc.gov/nccdphp/dch/programs/index.htm

National Library of Medicine of National Institute of Health (NIH). Healthcare Costs and Absenteeism among Caregivers of Adults with Partial-Onset Seizures: Analysis of Claims from an Employer Database. Published online at https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6306097/

National Institute of Mental Health. Mental health. January 2022. Published online at https://www.nimh.nih.gov/

National Institute of Health and Care Excellence (NICE), UK; Clinical Knowledge Summaries (CKS); https://cks.nice.org.uk/topics/parkinsons-disease/background-information/prevalence

National Association of Chronic Disease Directors. Commentary on Chronic Disease Prevention in 2022. Published online at https://chronicdisease.org/wp-content/uploads/2022/04/FS_ChronicDiseaseCommentary2022FINAL.pdf

Obermeyer, Z., Nissan, R., Stern, M., Eaneff, S., Bembeneck, E.J. and Mullainathan, S., 2021. Algorithmic bias playbook. Center for Applied AI at Chicago Booth.

Optum (UHG company). Connected Care – Whole picture for holistic health. 2021. Published online at https://www.optum.com/business/insights/advisory-services/page.hub.connected-care-whole-picture.html

Optum (UHG company). Connected Care. 2021. Published online at https://www.optum.com/business/insights/advisory-services/page.hub.success-stories-connected-care.html

Optum (UHG company). Fusing data with human-centered design. 2021. Published online at https://www.optum.com/business/insights/c-suite/page.fusing-data-human-centered-design.download.html

Optum (UHG company). A CMO toolkit. July 2021.

Peltz, A., Rogers, S. and Garg, A., 2020. An equity lens for identifying and addressing social needs within pediatric value-based care. Pediatrics, 146(4).

Rahul Sharma. The Crucial Role of SDOH Data in Value-Based Care. https://hitconsultant.net/2021/06/08/sdoh-value-based-care-role/, Oct 2021

Robert Graham Center, Health Landscape. Community Data. 2021. Published online at https://healthlandscape.org/community-data-and-research/

Santo L, Okeyode T. National Ambulatory Medical Care Survey: 2018 National Summary Tables; Available from: https://www.cdc.gov/nchs/data/ahcd/namcs_summary/ 2018-namcs-web-tables-508.pdf

Sills, M.R., Hall, M., Colvin, J.D., Macy, M.L., Cutler, G.J., Bettenhausen, J.L., Morse, R.B., Auger, K.A., Raphael, J.L., Gottlieb, L.M. and Fieldston, E.S., 2016. Association of social determinants with children's hospitals' preventable readmissions performance. JAMA pediatrics, 170(4), pp.350-358.

Stuckler, D., 2008. Population causes and consequences of leading chronic diseases: a comparative analysis of prevailing explanations. The Milbank Quarterly, 86(2), pp.273-326.

The Lancet. 40% of dementia cases could be prevented or delayed by targeting 12 risk factors throughout life. 30 July 2020. Published online at https://www.alzheimers.org.uk/news/2020-07-30/lancet-40-dementia-cases-could-be-prevented-or-delayed-targeting-12-risk-factors

Weiner, S., 2019. Addressing the escalating psychiatrist shortage. Association of American Medical Colleges AAMCNews.

Huang, Y., Liu, Y., Steel, P.A., Axsom, K.M., Lee, J.R., Tummalapalli, S.L., Wang, F., Pathak, J., Subramanian, L. and Zhang, Y., 2021. Deep significance clustering: a novel approach for identifying risk-stratified and predictive patient subgroups. Journal of the American Medical Informatics Association, 28(12), pp.2641-2653.