2019

Essays in Applied Microeconomics

Sidketa Ida Fofana

Essays in Applied Economics

by

Sidketa Ida Fofana

A Dissertation

Submitted in Partial Fulfillment of the

Requirements for the Degree of

Doctor of Philosophy

Major: Business Administration

The University of Memphis

August 2019

Acknowledgements

I am thankful to God the Almighty, for giving me strength, health, and intellect to conduct this dissertation. I am very grateful to the chair of my dissertation committee Dr. Joonhyung Lee for always going above and beyond in his great mentorship to teach me how to learn. I am also indebted to my outstanding doctoral committee members including Professor William Smith, Professor Albert Okunade, and Dr. Jamin Speer for your sincere advice, guidance and motivation. You made me a better researcher and I will be forever grateful. I also extend my gratitude to Ms. Janis Lamar for her kindness. I would like to acknowledge the support and encouragement of all my classmates, other faculty and university staff who contribute greatly to make my PhD journey memorable and academically worthwhile. Most Importantly, I would like to thank my beloved parents, Jean Mathieu and Mariam, for their unconditional love and for teaching me to always believe in my dreams. Thank you for teaching me the importance of education and hard work. I extend my thanks to my wonderful siblings Kuma and Nina and to all my friends specially Dr. Thierry Nianogo. Finally, I was very fortunate to have a nice husband Demba and my darling daughter Djene in my life. Thank you for putting up with me all those years and being my refuge.

Abstract

Sidketa Ida Fofana, Ph.D. The University of Memphis. August, 2019. Essays in Applied Economics. Major Professor Joonhyung Lee, Ph.D.

My dissertation consists of two essays in applied Microeconomics. I study the main determinants of Diabetes and breast cancer. My first essay is on diabetes, it assesses racial differences and the impact of health care coverage in self-Care management and quality of care for $9,805$ diabetes patients in Texas. Using a multiple logistic regression model, I find that Hispanics with diabetes in Texas are still struggling to improve their self-management and gain access to quality care compared to Black and White non-Hispanics. For instance, 41.4% of Hispanics fail to perform daily foot care compared to 34.2% of White non-Hispanic and 25% of Black non-Hispanics. Furthermore, Hispanics are less likely to have a provider checking their AIC (OR: 0.54, 95%, CI, .45-.63) and Black (OR: 0.87, CI 0.67-1.12) compared to Whites. My results also indicate that having health care coverage and taking a diabetes self-management class significantly improves self-management and considerably reduces the race disparity On my second essay, I take advantage of this 20-year cohort study of cancer survival data in Texas to study the main factors that can explain why some breast cancer patients live longer than others. Using a survival analysis which consists of performing a log-rank test, a survival time regression and a Cox proportional hazards regression, and dividing the data in groups based on the survival time then running a multinomial logistic regression, my results suggest that stage at diagnostic is the most important drivers of breast cancer survival, in fact, compared to stage1 survivors, survivors with stage IV are more likely to die with a hazard ratio of (14.02). I also find that being diagnosed with advanced grade will lead to short survival time. Furthermore, there are some racial disparities in survival time. Finally, I find that most of the disparities in terms of stage, grade, age, race and income occur in the first five years of survival. Those two essays lead to some policy recommendations such as facilitating access to

quality of care for minorities in case of diabetes and promoting early breast cancer screening and diagnostic in vulnerable communities.

Contents

List of Tables

List of Figures

Chapter 1

Introduction

My dissertation is composed of two essays in applied Microeconomics more precisely in health economics. According to Fuchs (2010), health economics is a behavioral science and it also provides major insights for health policy issues and health services research. In my research, I explore those two routes of the field. I investigate positive issues using real life data to analyze behaviors and relations within variables that affects conditions such as diabetes or breast cancer. I then make useful recommendations for policy makers to improve the general well being of economic agents, and increase their longevity.

My dissertation is centered on certain chronic diseases with high incidence on the United State population. In fact, diabetes is a chronic disease that can lead to serious complications if not properly taken care of. An estimate of 9.4% of the United State (US) population (30.3 million people) had the disease in 2015, according to the National Diabetes Statistics Report. Additionally, an estimate of 33.9% of the US adults aged 18 years or older (84.1 million people) had pre-diabetes in 2015. Moreover, from a CDC report, breast cancer is the second most common cancer after skin cancer among women in the United States across all races and ethnicities. In fact, in 2014, $236,968$ women and $2,141$ men in the United States

were diagnosed with breast cancer. Additionally, $41,211$ women and 465 men in the United States died from it (CDC). In the U.S., one in eight women will be diagnosed with invasive breast cancer during her lifetime according to National Cancer Institute. Those two conditions are burdens to tax payers, they are public health issue and cancer mortality is associated with huge productivity loss. As an illustration, after adjusting for age group and sex, average medical expenditures among people diagnosed with diabetes were about 2.3 times higher than expenditures for people without diabetes (American Diabetes Association). In 2017, the total direct and indirect estimated cost of those diagnosed with diabetes in the US was about $327 billion including $237 billion in direct medical costs and $90 billion in reduced productivity (American Diabetes Association, 2018). I am using sophisticate statistical tool such as logistic regression, Cox proportional hazards regression, and multinomial logistic regression to analyses the data. In the case of diabetes, the data was collected from the Texas Behavioral Risk Factor Surveillance Survey (BRFSS) $2008-2013$. It is a phone survey that collects state data about U.S. residents regarding their health-related risk behaviors, chronic health conditions, and use of preventive service. In my study, I have a large sample size of $9,805$ diabetes patients. For the socio-economic study of breast cancer survival, data for breast cancer cases was provided by the Texas Cancer Registry (TCR), Cancer Epidemiology and Surveillance Branch, Texas Department of State Health Services. It is a retrospective cohort study in Texas and I include only the diagnosis of breast cancer patients recorded by the SEER Program in 1995 and 1996, and we follow them up to 2015 with a large sample size of $21,107$ survivors of breast cancer. My study in diabetes concludes that self-care management and quality of care are heavily impacted by the race differential in Texas. In fact, 41.4% of Hispanics fail to perform daily foot care compare to 34.2% of White non-Hispanic and 25% of Black non-Hispanic. Only 50.98% of Hispanics attended class compare to 57.89% of white

and 63.55% of Black non-Hispanics. Furthermore, Hispanics are less likely to have a provider checking their AIC (OR: 0.54, 95%, CI, .45-.63) and Black (OR: 0.87,CI 0.67-1.12) compared to Whites. Hispanics were less likely to get flu shot (OR: .62, 95%, CI,.56-.68), and Black (OR: .64, 95%, .56-.74) compared to White non-Hispanic. Hispanics were also less likely to get pneumonia shot (OR: .35, 95%, CI, .32-.38), and Black (OR: .56, 95%, $.48 - .64$) compared to White non-Hispanic. However, having health care coverage and taking a diabetes self-management course are contributing factors to the reduction of racial disparity in self-care management. From my research on breast cancer, I find that stage at diagnostic is the most important drivers of breast cancer survival, in fact, Compared to stage1 survivors, survivors with stage IV are more likely to die with a hazard ratio of (14.02). My findings also suggest that most of the disparities in terms of stage, grade, age, race and income occur in the first five year of survival. I conclude that Policy makers need to promote early screening and diagnostic in vulnerable communities.

Chapter 2

Racial differences and impact of health care coverage in self-care management and quality of care on diabetes in Texas

2.1 Introduction

Diabetes is a chronic disease that can lead to serious complications if not properly taken care of. An estimate of 9.4% of the United State (US) population (30.3 million people) had the disease in 2015, according to the National Diabetes Statistics Report. Additionally, an estimate of 33.9% of the US adults aged 18 years or older (84.1 million people) had prediabetes in 2015. Prediabetes is a condition that can lead to type 2 diabetes if there is no significant change toward a healthier lifestyle. In 2014, 7.2 million hospital discharges and 14.2 million emergency department visits were reported with a diabetes diagnosis among US adults aged 18 years or older. After adjusting for age group and sex, average medical expenditures

among people diagnosed with diabetes were about 2.3 times higher than expenditures for people without diabetes (American Diabetes Association); in 2012, the total direct and indirect estimated cost of those diagnosed with diabetes in the US was about $245 billion, including $176 billion in direct medical costs and $69 billion in reduced productivity and total direct and indirect cost rise to $327 billion in 2017 including $237 billion in direct medical costs and $90 billion in reduced productivity(American Diabetes Association, 2018), and this constitutes a burden to tax payers. Complications due to diabetes, such as cardiovascular diseases and kidney failure are very costly. Also, in 2015, diabetes was the seventh leading cause of death in Texas and in the US, according to Center of Disease Control and Prevention (CDC). Therefore, self-management and quality of care for diabetes patients are extremely crucial to avoid further health complications and reduce morbidity, mortality, and medical cost significantly. The southern states, specifically, have the highest incidence and prevalence of the disease. In my study, I focus on Texas, where diabetes is a growing epidemic that affects 11.2% of the population, according to the CDC, and diabetes in Texas is taking a drastic physical and financial toll on the state. In fact, in 2017, the total diabetes cost was about 25.60 billion in Texas. It is a public health concerns that include the region of the Rio Grande Valley where 26% of the population suffers from diabetes and the incident of prediabetes is still growing. The literature suggests that there might be some racial disparities in self-management and quality of care in diabetes patients in the past (Sloan, Padrotn, and Platt (2009); Nwasuruba, Osuagwu, Bae, Singh, and Egede (2007); Adams, Zhang et al. (2005)). My contribution to the literature on diabetes is to investigate if there is a racial difference in self-management and quality of care on diabetes patients from 2008-2013 in Texas. In addition, this paper seeks to understand how having health care coverage and taking a diabetes class will impact self-management and quality of care in Texas. The methodology I use to

analyze the data is the multiple logistic regression models. The data was collected from the Texas Behavioral Risk Factor Surveillance Survey (BRFSS) 2008 -2013.

2.2 Background

Diabetes is a chronic disease that is associated with abnormally high levels of sugar or glucose in the blood due to the bodiesŠ inability to produce or respond to the hormone insulin. Type 1, type 2, and gestational diabetes are the most common types of diabetes. Type 1 diabetes is often diagnosed in children and young adults, although it can appear at any age. People with this type will need to take insulin every day to stay alive, and it accounts for less than 5% of all diagnostic diabetes in the US (CDC). Improving glycemic control in patients with type 1 diabetes substantially reduces their risk of microvascular complications and cardiovascular disease, according to Lind, Svensson et al. (2014). Type 2 diabetes is mostly diagnosed in middle-age and older people, yet it can occur at any age. In type 2, the body does not produce or use insulin well, and between $90 - 95\%$ of all diagnosed with diabetes have type 2. Gestational diabetes can develop during pregnancy in some women. In most cases, this type goes away after women give birth, yet according to Kim, Newton and Knopp (2005), it is a risk factor for type 2 diabetes later in life Finally, in cost benefit analyze we can observe that it is much cheaper to take the diabetes self-care management class and have access to the quality of care than pay for the complications. The table below 2.1 is showing those costs. Therefore, it is extremely urgent to find solutions of avoiding costly complications due to diabetes for taxes payers, patients and insurance companies.

Table 2.1: Cost analysis for people living with diabetes in the USA

	Average Cost
Individual Diabetes Self-Management, Initial Visit per hour	$20 – $182.00
Individual Diabetes Self -Management, Follow Up Per hour	$20 – $91.00
Cost for flu shot	$15 – $40 with no insurance
Cost for pneumonia shot	$125.99 – $229.99(ex: Pneumovax 23)
Group Diabetes Management Two-Hour Class Sessions	$20 – $140.00
Gym membership per month	$10 – $20.00
Doctor visit	$100 – $400.00
People with diagnosed diabetes incur average medical expenditures	$16,750.00, which about$9,600is attributed to diabetes, in 2017
The costs of hospitalizations involving stroke	average cost is between 20,396 and 23,256
A foot or leg amputation costs in initial hospital costs,	Between $30,000 and $60,000

2.3 Literature Review

Referring to the concept of health capital and the demand for health, Grossman (1972) constructs a model of the demand for the commodity "good health." To him, individual demands good health because there is positive utility associated with being healthy, and having good health determines the amount of time available for home and income production. In the case of diabetes patients, good health can be viewed as avoiding further health complications due to the diabetes. Thus, those patients will get positive utility of being healthy. They would like to be more productive at home and at work and avoid losing time being hospitalized or sick. Additionally, the state will save more money if all diabetes patients have good health. Therefore, I must analyze and control the risk factors that can lead to

diabetes complications. So, according to the National Health and Nutrition Examination Survey (NHANES) and National Center for Health Statistics Centers for Disease Control and Prevention, the main risk factors for complications among adults aged 18 years or older with diagnosed diabetes in the USA in 2011-2014 are, obesity, physical inactivity, smoking, high blood pressure, and high blood glucose. Obesity is the key risk factor for complications in diagnosed diabetes with around 87% of people with the disease in the US being obese or overweight (National Diabetes statistics Report, 2017). Obese people have added pressure on their body's ability to use insulin to accurately control blood sugar levels and are, consequently, more likely to develop diabetes or have more complications from it. Therefore, good self-management for diabetes patients will require healthier nutrition and exercise to reduce the bodyŠs weight. Additionally, salty, fatty, and sweet foods need to be avoided to prevent future complications, such as cardiovascular disease, kidney damage, amputations, loss of sight, nerve damage, and depression. For instance, Hu and Malik (2010) assess risk of obesity and type 2 diabetes due to sugar-sweetened beverages (SSB). Hu and Malik (2010) review the literature on the impact of SSB on obesity and diabetes until 2006, and their conclusion is that epidemiological studies clearly indicate that regular consumption of SSBs can lead to weight gain and substantially increase risk of developing chronic diseases including MetSyn, T2DM and CHD. Therefore, self-management for diabetic patients should include controlling their intake of sugar-sweetened beverages, such as soda.

Furthermore, Chou, Shin-Yi, Mi, Grossman, and Saffer (2004) study the factors that may be responsible for the 50% increase in the number of obese adults in the US since the late 1970s. They use the Behavioral Risk Factor Surveillance System for the years 1984Ű1999 and their findings suggest fast-food and full-service restaurants as culprits in undesirable weight outcomes. Also, technological innovations and the realization of economies of scale that led to reductions in the

fast-food restaurant price may have been stimulated in part by efforts to accommodate the increased demand for consumption of food away from home. Therefore, to fight obesity among diabetes patients we need to reeducate them on staying away from unhealthy food and cooking more for themselves. Likewise, smoking is a risk factor for complications with diagnosed diabetes. In fact, smoking increases insulin resistance and is associated with central fat accumulation. As a result, smoking is surely a risk factor for complications for diabetes patients. Moreover, smoking increases the risk of metabolic syndrome and diabetes, and these factors increase risk of cardiovascular disease, which is one of the major complications that leads diagnosed diabetes patients to be hospitalized, according to Chiolero et al. (2008) and Attvall et al. (1993). Moreover, blood glucose control is important for self-management and quality of care for diabetes patients to check the status of the disease. Therefore, the A1C test is recommended for blood glucose control. The A1C test is a blood test that provides information about a personŠs average levels of blood glucose, also called blood sugar, over the past 3 months. The A1C test is sometimes called the hemoglobin A1C, HbA1c, or glycohemoglobin test. The A1C test is the primary test used for diabetes management and diabetes research. Below 5.7% is normal, between 5.7% and 6.4% the person is considered to suffer from prediabetes, and above 6.5% is considered as diabetic (National institute of Diabetes and Digestive and Kidney Disease). Therefore, the American Diabetes Association recommends people who have the disease to use A1C or any test at least 2 times a year to control their blood glucose. Shah, Langenberget al.(2015) studied a cohort of $1,921,260$ individuals in England, of whom $1,887,062$ (98.2%) did not have diabetes and $34,198$ (1.8%) had type 2 diabetes. Their goal was to assess the incidence of cardiovascular diseases and type 2 diabetes. From a follow up of 5.5 years, they found that $6,137$ (17.9%) of type 2 diabetes patients have cardiovascular disease and the most common were peripheral arterial disease, heart

failure, ischaemic stroke, stable angina, and non-fatal myocardial infarction. Furthermore, they find that for individuals aged 40 years old without cardiovascular disease, the overall estimated risk of developing any cardiovascular disease by age 80 years was 30.7% (95% CI $30.3 - 31.0$) for women without diabetes and 44.3% $(43.8 - 44.7)$ for men without diabetes, compared to 58.2% $(54.9 - 61.4)$ for women with type 2 diabetes and 67.4% $(64.4 - 70.4)$ for men with type 2 diabetes. Therefore, diabetes and cardiovascular diseases are highly correlated. This is very worrisome when we are aware that cardiovascular diseases are the first leading cause of death in Texas and in the US.

Therefore, self-management and access to quality of care are vital for diabetes patients to avoid or minimize the risk of cardiovascular disease, heart failure, and stoke. Those complications are very debilitating and very costly for individuals who have it. Also, Kate, Sobel et al (1999) conducted a randomized trial to study if a chronic disease self-management program can improve health status while reducing hospitalization. Their study was a six-month randomized, controlled trial at community-based sites comparing treatment subjects with wait-list control subjects. Participants were 952 patients, 40 years of age or older with a physician-confirmed diagnosis of heart disease, lung disease, stroke, or arthritis. They measured health status, health behaviors and health service utilization, as determined by mailed, self-administered questionnaires. They found that treatment subjects, when compared with control subjects, demonstrated improvements with 6 months of weekly minutes of exercise, frequency of cognitive symptom management, communication with physicians, self-reported health, health distress, fatigue, disability, and social/role activities limitations. They also had fewer hospitalizations and fewer days in the hospital. Yet, no differences were found in pain/physical discomfort, shortness of breath, or psychological well-being. This paper reinforces the importance of self-management and quality of care for diabetic patients to avoid

complications and reduce mortality due to the diseases. Furthermore, this paper supports my idea that perhaps if all diabetes patients have access to health care coverage and classes on how to manage it, their condition may improve.

Texas is unique as far as race and ethnicity are concerned. Indeed, it is very racially and ethnically diverse with more than 40% Hispanics, 12% Black and 40% White non-Hispanics; based on the literature, racial and ethnic disparities persist for access to medical care in significant measure for several disease categories and service types, according to Mayberry et al.(2000). Egede Le and Zheng D (2003) studied racial/ethnic differences in adult vaccination among individuals with diabetes with a main goal of examining whether differences in access to health care, health coverage, and socioeconomic status explained racial differences in influenza and pneumococcal vaccination rates in individuals with diabetes. They found that indeed, racial disparity in vaccination rates for adults with diabetes is independent of access to care, health care coverage, and socioeconomic status. If those conclusions are accurate, then perhaps having health care coverage may not have any significant impact on race disparities in terms of self-management for diabetes patients. State leaders may need to find better approaches that reduce race disparity and have more significant impacts on improving self-management and quality of care for diabetes patients. Karter, Moffet et al.(2002) also studied ethnical Disparities in Diabetic Complications in an Insured Population. Using a longitudinal observational study conducted from January 1, 1995 to December 31, 1998 in California, they had a total sample of $62,432$ diabetic patients which was composed of 64% white, 14% black, 12% Asian and 10% Latino participants. Their study concluded that for end-stage renal disease, there was an elevated incident among ethnic minorities despite uniform medical care coverage. Furthermore, Karter, Moffet et al. provided new evidence that rates of other complications are similar or lower relative to those of whites. According to their studies, genetic origin

or unmeasured environmental factors or a combination of both can explain those ethnical disparities. In my paper, I want to investigate if there is a racial disparity in self-management care for diabetic patients in Texas and if those disparities persist over time.

In addition, Sloan, Padroťn, and Platt (2009) study preferences, beliefs, and self-management of Diabetes, they conclude that individualsŠ beliefs about control over life events and longevity matter to both investments in care and subjective measures of health outcomes. While such beliefs were associated with self-assessed outcomes, they were not able to establish a link between these beliefs and the most widely used single objective measure of diabetes outcomes. Even after controlling for differences in preferences and beliefs, clinical, cognitive, and demographic factors, substantial differences in health outcomes, both subjectively and objectively measured, but not in health investment, by race and ethnicity remained. This implies that blacks and Hispanics may realize lower returns on health investments, at least for diabetes care, than do non-Hispanic whites, even after controlling for socioeconomic factors, cognitive status, and preferences and beliefs. I investigate if there is a racial disparity in Texas on self-management and quality of care for diabetes patients. One of the prime conditions for a diabetes patient to successfully self-manage their condition is to take a diabetes self-management education. In fact, diabetes self-management education offers a valuable opportunity for individuals living with diabetes to gain the knowledge, skills, and motivation to successfully cope with their disease and reduce costly complications. This will then lead to reduced racial disparity if it exists.

The motivation for this paper also comes from Sloan, Padroťn, and Platt (2009) and Nwasuruba, Osuagwu, Bae, Singh, and Egede (2007), who also studied racial differences in diabetes self-management behaviors and quality of care in Texas. Nwasuruba and al (2007) study was conducted among $1,720$ adults with diabetes

from $2002 - 2004$ in Texas, Behavioral Risk Factor Surveillance Survey. They used a multiple logistic regression models, and their paper concluded that Hispanics have poorer access to care and poorer health status compared to whites or blacks. I contribute to the literature in many ways. In fact, I use recent data from $2008 - 2013$. I pool data for the 6-year period, which allows me to have a bigger sample size of $9,805$, and more importantly to better help to answer the questions I have raised. Additionally, I am accessing the importance of race on self-management and quality of care for diabetes patients mainly because the literature has mixed results on these issues. One the one hand, Nwasuruba et al. (2007) found that there was a racial disparity in quality of diabetes care, while Brown et al. (2005) concluded that minority race/ethnicity was not consistently associated with worse processes or outcomes, and not all differences favored whites. The only notable socioeconomic position disparity in their study was in rates of dilated eye examinations. Social disparities in health may be reduced in managed-care settings. I believe that by identifying the sub groups that are struggling to improve self-management or have less access to quality of care, I can better assist them and make resourceful recommendations to state leaders and public health officials.

Furthermore, my study helps understand how having health care coverage and taking diabetes management classes will impact self-management and quality of care in Texas. I expect that having health care coverage and a class will be the solution to improving diabetes self-management and quality of care. Consequently, it can lead to the elimination of racial disparities among diabetes patients in Texas. In fact, Lillie-Blanton, and Hoffman (2005) studied the role of health insurance coverage in reducing racial/ethnic disparities in health care, their paper reviews evidences from studies specifically designed to quantify the contribution of health insurance to racial/ethnic disparities in access. They compared measures of access to health care between whites and a specific racial/ethnic group; measured the

effects of racial/ethnic differences in social, economic, or health system factors that may contribute to disparities in access to health care; and quantified the contribution of racial/ethnic differences in health insurance on disparities in access to care. They found that health insurance accounts for much of the variation in racial/ethnic disparities in access to a usual source of care. This paper is relevant since, health care coverage for patients with diabetes may reduce the racial disparity in Texas. In addition, I am inspired by the health production function of Frank A. Sloan and Chee-Ruey Hsieh (2012), Orem's self-care deficit nursing theory, and the clinical guidance of American diabetes Association (ADA) 2018. I then have a health production function which relates inputs to outputs:

$$Output_{ij}(H) = f_i(self - caremanagment_{ij}, accesstoqualityofcare_{ij}, \tag{2.1}$$

$$self - caremanagment_{ij} * accesstoqualityofcare_{ij}, geneticmakeup_{ij})$$

$Output_{ij}(H)$ In this case of diabetes, refers to good outcome of diabetes patients which is to avoid further complications due to the condition such as cardiovascular disease, amputation, depression. In this case, I use general health reported by the patients and patients BMI. I consider good outcome as someone reporting having excellent and good health or someone having a BMI less than 25 (normal weight). In model (1), i represents individual patient and j represents that individual race/ ethnicity. For race/ ethnicity, I consider Black non-Hispanics, White non-Hispanics, and Hispanics. Keys assumptions for model (1): Genetic makeup does not vary over time and is outside the decision maker's or diabetes patient control. It is considered as exogenous to diabetes patients then.

Since, diabetes patients can control their self-care management and access to quality of care, I consider those variables as endogenous.

Self-care management and access to quality of care are complement.

Based on Orem‘s self-care deficit nursing theory, I define self-care as a human regulatory function that persons must perform for themselves to maintain life. Self-care must be learned, and it must be deliberately performed. As a result, persons through this self-care management process, exercise intellectual and practical skills to manage themselves, to sustain the motivation in an effective manner for essential daily care. The way an individual engages in self-care will vary due to influences from their culture, environment, and outside influences such as race and ethnicity. Engaging in self-care and dependent-care are affected by a personŠs limitations in knowing what to do, when to do it and how to do it. Based on the American Diabetes Association clinical guidance, for diabetes patients self-care management is composed by physical activity (N_{pa}), smoking (N_s), home glucose testing (N_{hgt}), home foot exam (N_{hfe}), and attending diabetes classes (N_{adc}) . I will empirically test how each race/ ethnicity performs on each component. The basic conditioning factors influence both the need and abilities of the person for performing self-care based on Orem‘s nursing theory. Basic conditioning factors are factors that affect all persons such as age, gender, developmental state, health state, health care system, socio-cultural orientation, family system factors, patterns of living, environmental factors and available resources. In my study, I will check those demographic characteristics by race. Another input that is central for diabetes patients is access to quality of care. Based on the ADA guidance, for access to quality of care I consider AIC test by provider (Z_{atp}), foot exam by provider (Z_{fep}), dilated eye exam by provider (Z_{deep}), received the flu shot (Z_{rfs}), and received the pneumonia shot (Z_{rps}). Finally, self-care management and access to quality of care are considered complement. Therefore, any diabetes patient who would like to avoid complications need to perform both. Why f_i is important for individual with diabetes? Based on the index of “quality-adjusted life years” developed by

Zeckhauser and Shepard (1976) and Weinstein and Stason (1977), I build the following utility function: Let‘s consider an individual who has diabetes with a lifetime horizon of T years along which he will enjoy a health status H_t and I consider that the individual value the quality of life. I assume that each year when the individual with diabetes avoid complication due to the disease, he can work and earn income Y_t Let‘s the health profile of an individual be

$$(H_1, t_1, Y_1; H_2, t_2, Y_2; \cdots ; H_n, t_n, Y_n)$$

where health status H_t is maintained for t_t periods, income Y_t is earned at t_t periods and $t_1 + t_2 + t_3 + \cdots \ t_n = T$ I get this utility function:

$$U(H_1, t_1, Y_1; H_2, t_2, Y_2; \cdots ; H_n, t_n, Y_n)$$

I assume that for the individual with diabetes who value life, there will be some medical costs or costs related to good self-care management and quality of care such as health insurance, gym membership. I will denote that cost M I assume that if the individual with diabetes is very sick, he/she will not be able to work,

$$Y_t = 0$$

.

The budget constraint will then be:

$$Y_t(H_t) = qX_t + pM_t$$

Where Y_t is spent between medical care M_t and other consumption goods X_t at time t, and p and q being the prices. This budget constraint will be concave rather

than linear. It is because individual with poor income will not be able to work so income will be zero. Therefore, for the individual to maximize utility, he/she will need to conduct an excellent self- care manage of his/her diabetes and have access to great quality of care. I will then need to analyze all those components by race/ ethnicity to see if there are some disparities. Policy makers should make sure that diabetes patients have access to good health care coverage and take diabetes self-care management mainly for the poorest, and sick. In fact, from the budget constraint I can observe that if you are very sick you cannot work, and if you earn very low income, you may not be able to sustain medical cost. That is why I will analyze the race disparity among diabetes patients who have health care coverage and have taken the diabetes self-care management classes.

2.4 Data Description and Characteristics

I use data from the Texas Behavioral Risk Factor Surveillance Survey (BRFSS) from 2008 to 2013. It is a phone survey that collects state data about U.S. residents regarding their health-related risk behaviors, chronic health conditions, and use of preventive services. Each year, more than $400,000$ adult interviews are collected. It is a well-recognized tool for health researches in the US and it has been used for epidemiology, public health, health economist in their research publications. In this paper, I use data from 2008 to 2013. This approach helps increase the sample size and allows for subgroup investigation. The BRFSS does not distinguish between diabetes types 1 and 2, so I will not be able tell the difference either between type 1 and type 2. Yet, I know that less than 5% of all diagnosis diabetes are type 1 and can then assume most results are for type 2 patients. I exclude gestational diabetes in my study. I am using a multiple logistic regression model to see the racial disparity and impact of health care coverage in diabetes self-management and

quality of care in Texas. In my study, I define race or ethnicity as White non-Hispanics, Black non-Hispanics, and Hispanics. I create four age categories, which are 18-34, 35-49, 50-64 and 65 and plus years, and education is categorized as less than high school graduate, high school graduate, and some college education or more. Income is categorized as less than $\$25,000.00$, less than $\$50,000.00$, and $\$50,000.00$ and more. Marital status is dichotomized as married and not married. Employment is dichotomized as employed and non-employed. Health status is categorized as excellent, very good, or good vs fair or poor. Body mass index (BMI) is categorized as less than 25 (normal or underweight), over 25 (overweight and obese). I report patients who take insulin.

Table 2.2 displays the demographic characteristics of adults with diabetes in Texas by race from 2008 to 2013, 55.52% of the sample are white non-Hispanic, 10.43% are Black non-Hispanic and 34.04% are Hispanic. I can see that White non-Hispanics are much older than Black non-Hispanics and Hispanics in fact 57.7% of White non-Hispanic are over 65 years old compared to 41.05% of black non-Hispanic and only 39.6% of Hispanic being over 65 years old. On the education factor, Hispanics are less educated than Black and White non-Hispanic, 48.52% have no high school diploma compared to 17.06% for black non-Hispanics and only 8.82% for white non-Hispanics. Income disparity is observed, and Hispanics are poorer than black non-Hispanics and White non-Hispanics. Hispanics are more likely to report fair or poor health more than Black non-Hispanics and White non-Hispanics and less likely to have a regular provider. Black non-Hispanics are less likely to be employed or married and more likely to be overweight and obese compared to White non-Hispanic and Hispanic.

Self-management for adults with diabetes in Texas Referring to the American Diabetes Association guidelines of 2017 and past literature such as Nwasuruba et al. (2009), Chandler and Monnat (2015), I use 5 variables to assess self-management for

Table 2.2: Demographic characteristics of adults with diabetes in Texas by race $2008 - 2013$

$N = 9805$	NHW $n = 5444$	NHB $n = 1023$	H $n = 3338$
Age			
$18 - 34$	1.24	2.64	3.05
$35 - 49$	13.97	25.7	28.34
$50 - 64$	27.07	30.6	29.02
65+	57.7	41.05	39.6
Women	56.15	66.86	64.89
Education			
<High School	8.82	17.06	48.52
HS GED	28.48	33.33	24
Some college+	62.7	49.61	27.47
Income			
$<$ 25,000	34.36	59.59	66.58
25,000 to $<$ 50,000	29.59	21.93	19.41
50,000+	36.04	18.48	14
Married	52.97	32.31	53.81
Employed	35.6	31.2	44.31
Health status			
$Exc./VG/G$	57.16	46.41	37.8
Fair/ Poor	42.84	53.59	62.2
Has a regular provider	93.83	85.3	73.68
BMI			
<25(normal)	16.98	12.35	12.77
>25(Obese and Overweight)	83.02	87.65	87.22
Takes insulin			
Yes	25.63	37.6	27.97

Notes: NHW=Non-Hispanic white, NHB=Non-Hispanic Black, H=Hispanic

diabetes patients and 5 variables to assess access of quality of care for diabetes patients. Therefore, to assess self-management I consider: physical activity (Y_{pa}), smoking (Y_s), home glucose testing (Y_{hgt}), home foot exam (Y_{hfe}), and attending diabetes classes (Y_{adc}) I am using a Chi- square test to test if I reject or not the null hypothesis I then have for self-care management:

$$H_0 : p_{Wk} = p_{Bk} = p_{Hk}$$

$$H_a : p_{Wk} \neq p_{Bk} or p_{Bk} \neq p_{Hk} or p_{Wk} \neq p_{Hk}$$

Where k is equal to physical activity, smoking, home glucose testing, home foot exam, or attending diabetes classes, and for instance when K is physical activity, p_{Wk} means proportion of white that are engaged in physical activity.

Table 2.3: Comparison of diabetes self-management by race/ ethnicity in Texas in %

	NHW	NHB	H	P-value
Home glucose testing				0.068
>Once daily (yes)	63.25	69.46	65.62	
< once daily or Never (No)	36.75	30.54	34.37	
Home foot care				0
>Once daily	65.8	75	58.6	
<once daily or Never (No)	34.2	25	41.4	
Attended diabetes class				0
Yes	57.89	63.55	50.98	
No	42.1	36.44	49.02	
Smoking				0
Yes	23.37	36.32	28.88	
No	76.62	63.68	77.12	
Physical activity				0.062
Yes	58.26	59.77	56.17	
No	41.74	40.22	43.83	

Notes: NHW=Non-Hispanic white, NHB=Non-Hispanic Black, H=Hispanic

2.3 shows a comparison of diabetes self-management by race/ ethnicity in Texas from 2008 to 2013. I notice that there is race disparity in self-management. In fact, Hispanics (34.37%) never perform home glucose testing or perform it less than one a day compared to Black non-Hispanics (30.54%) and White (36.75%) non-Hispanic. As far as home foot care is concerned, Hispanics (41.4%) never perform home glucose testing or perform it less than one time a day, compared to Black (25%) and white (34.2%) non-Hispanic. Black non-Hispanics (63.55%) attended more diabetes class compared to 57.89% White non-Hispanics and 50.98% Hispanics. They are more likely to smoke 36.32% compared to White non-Hispanics (23.37%) and

Hispanics (28.88%). For physical activities, Black non-Hispanics (59.77%) are more likely to exercise than White non-Hispanics (58.26%) and Hispanics (56.17%). I can conclude that for diabetes self-care management, Hispanics are doing slightly worse than Black and white non-Hispanic. Access to quality of care for adults with diabetes in Texas 2008-2010

To examine access to quality of care for diabetes patients, I consider: AIC test by provider (Z_{atp}), foot exam by provider (Z_{fep}) , dilated eye exam by provider (Z_{deep}), received the flu shot (Z_{rfs}), and received the pneumonia shot (Z_{rps}). I am using a Chi- square test to test if I reject or not the null hypothesis I then have for access to quality of care:

$$H_0 : p_{Wk} = p_{Bk} = p_{Hk}$$

$$H_a : p_{Wk} \neq p_{Bk} or p_{Bk} \neq p_{Hk} or p_{Wk} \neq p_{Hk}$$

Where k is equal to AIC test by provider, foot exam by provider, dilated eye exam by provider, received the flu shot, or received the pneumonia shot. For instance, p_{wk} when K is received the pneumonia shot, mean proportion of white that received the pneumonia shot.

Table 2.4 shows a comparison for access to quality of care for adults with diabetes in Texas by race. Some race disparity in access to quality of care are observed with all p-values being less than 0.01. In fact, Hispanic (47.67%) never have the A1C test by provider or not having it once a year compared to Black (35.89%) and White (32.84%) non-Hispanic. As far as the foot exam by provider, Hispanics (38.32%) never have it or having it less than once a year. compared to Black (15.53%) and white (27.81%) non-Hispanic. In terms of a dilated eye exam by provider, 33.17% of Hispanics never have it or having it less than once a year

Table 2.4: Access to quality of care for adults with diabetes in Texas by race $2008 - 2013$

	NHW	NHB	H	P-value
AIC test by provider				0
Yes	67.15	64.11	52.33	
NO	32.84	35.89	47.67	
Foot exam by provider				0
Yes	72.18	84.47	61.68	
No	27.81	15.53	38.32	
Dilated eye exam by provider				0
Yes	69.68	75.37	66.83	
No	30.32	24.63	33.17	
flu shot				0
Yes	64.42	53.8	47.04	
No	35.58	46.2	52.96	
pneumonia shot				0
Yes	68.95	44.74	56.18	
No	31.05	55.26	43.81	

Notes: NHW=Non-Hispanic white, NHB=Non-Hispanic Black, H=Hispanic

compared to 24.63% of Black and 30.32% of white non-Hispanics. Less Hispanic (47.04%) get the flu shot compared to White non-Hispanics (64.42%) and Black non-Hispanics (53.8%). Finally, only 56.18% of Hispanics get the pneumonia vaccine compared to 44.74% of Black non-Hispanics and 68.95% of White non-Hispanics. Therefore, as far as access to quality of care for adults with diabetes in Texas is concerned, White non-Hispanics are having better access to quality of care than Black non-Hispanics and Hispanics.

2.5 Econometrics Model and results

I consider several specifications; each of the specifications is run on self-care management and on access to quality of care. The first specification consists of running a logistic regression where the dependent variable Y takes on one or zero, it takes one if the patient took diabetes classes and zero if not, say, and the

independent variable is race; I defineπ as the probability that Y takes on one. This first specification can be written as:

This first specification can be written as:

$$\boldsymbol{l}ogit(\pi_i) = \ln(\frac{\pi_i}{1-\pi_i}) = \boldsymbol{\beta}_0 + \boldsymbol{\beta}_1 Bl_i + \boldsymbol{\beta}_2 HS_i \tag{2.2}$$

Where i is individual, Bl_i is equal to one if individual i is black person and zero if not, and HS_i is equal to one if individual i is Hispanic and zero if not, here white race is the control variable.

Table 2.5: Unadjusted odds Ratio of self-care management, access to quality of care and health output for diabetes patients in Texas.

	BNH Odd ratio,95% CI, n= 1023	H odd ratio, 95% CI n=3338
Physical activity	1.06, (.93-1.22)	.918*,(.84-1)
Home glucose testing	1.32*, (1.02-1.69)	1.11, (0.94-1.3)
Home foot care	1.56*, (1.19-2.03)	.74*, (.63-.86)
Attended diabetes class	1.27*, (.99-1.61)	.76*,(.65-.88)
Smoking	1.87*, (1.52-2.3)	1.33*,(1.14-1.54)
A1C test by provider	0.87,(.67-1.12)	.54*,(.45-.63)
Foot exam by provider	2.1*, (1.52-2.88)	.62*,(.52-.73)
Dilated eye exam by provider	1.33*, (1.02-1.73)	.88,(0.74-1.03)
Flu Shot	0.64*,(.56-.74)	0.62*, (.56-.68)
Pneumonia Shot	o.56*,(.48-.64)	0.35*,(.32-.38)
Health Plan	0.38*,(.31-.47)	0.18*, (.16-.21)
General Health	0.65*,(.57-.74)	0.46*,(.42-.5)
BMI	0.69*,(.56-.84)	0,72*,(.63-.82)

Notes: $\star$ means significance at 5% level of significance,and NHB=Non-Hispanic Black, H=Hispanic

Table 2.5 shows that there is a race disparity in term of self- care management and access to quality of care. In fact, Hispanic are less likely to check their feet at home (Odd Ratio (OR): 0.74, 95%, CI, $0.63 - 0.86$)) compared to white. Black non-Hispanic are likely to smoke (OR: 1.87, 95%, CI, $1.52 - 2.3$) compared to White non-Hispanic and Hispanic (OR: 1.33). Hispanic are less likely to have a provider

check their AIC (OR:0.54, 95%, CI, $.45 - .63$) and Black (OR: 0.87) compared to Whites. Hispanics were less likely to get flu shot (OR: .62, 95%, CI, $.56 - .68$), and Black (OR: .64, 95%, $.56 - .74$) compared to White non-Hispanic. Hispanics were also less likely to get pneumonia shot (OR:.35, 95%, CI, $.32 - .38$), and Black (OR:.56, 95%, $.48 - .64$) compared to White non-Hispanic. Hispanic are less likely to report having excellent, very good and good (OR: 0.46, 95%, CI, $0.42 - 0.5$) compared to Black (OR: 0.65, 95%, CI, $0.57 - 0.74$), and White non-Hispanic.

The second specification consists of running a logistic regression where the dependent variable Y takes on one (with π) as probability) or zero, it takes one if the patient took diabetes classes and zero if not, say, and the independent variable are race, sex, education, marital status, employment status, age, and Income. This second specification can be modeled as:

$$\boldsymbol{logit}(\pi_i) = \boldsymbol{\beta}_0 + \boldsymbol{\beta}_1 Bl_i + \boldsymbol{\beta}_2 HS_i + \boldsymbol{\beta}_1 sex_i + \boldsymbol{\beta}_2 HSG_i \tag{2.3}$$

$$+\boldsymbol{\beta}_1 CG_i + \boldsymbol{\beta}_2 Ma_i + \boldsymbol{\beta}_1 Ep_i + \boldsymbol{\beta}_2 AL49_i + \boldsymbol{\beta}_1 AL64_i + \boldsymbol{\beta}_2 AG65_i + \boldsymbol{\beta}_1 IL_i + \boldsymbol{\beta}_1 IG_i$$

where i is individual i, Bl_i is equal to one if individual i is black person and zero if not, and HS_i is equal to one if individual i is Hispanic and zero if not, here white race is the control variable; sex_i is equal to one if individual i is a male and zero if not; HSG_i is equal to one if individual i is a high school graduate and zero if not; CG_i is equal to one if individual i is a college graduate and zero if not; Ma_i is equal to one if individual i is married and zero if not; Ep_i is equal to one if individual i is employed and zero if not; $AL49_i$ is equal to one if individual i is between 15 and 49 years old and zero if not; $AL64_i$ is equal to one if individual i is between 49 years old and 64 and zero if not; $AL65_i$ is equal to one if individual i is more than 65 years old and zero if not; IL_i is equal to one if individual i earns between $\$25,000.00$ and \$50000 and zero if not; IG_i is equal to one if individual i earns more

than $\$50,000.00$ and zero if not.

Table 2.6: Adjusted OR of diabetes self-care management and access to quality of care in Texas controlling for age, income, education, marital status

	BNH	H	Sex	HG	Col	Mar	Empl
Physical activity	1.22*	1.25*	.79*	1.11	1.54*	1.06	1.17*
Home glucose testing	1.27	0.98	1.44*	0.85	0.82	1.01	0.62*
Home foot care	1.67*	0.95	1.15	1.78*	1.8*	1.1	0.9
Attended diabetes class	1.45*	0.96	1.38*	1.61*	2.01*	1	0.86
Smoking	1.25	0.8*	1.37*	1.09	0.82*	0.76	0.97
A1C test by provider	1.08	0.84	1.17	1.4*	1.87*	1.1	0.85
Foot exam by provider	2.48*	0.9	0.86	1.6*	1.67*	0.94	0.8*
Dilated eye exam by provider	1.7*	1.3*	1.1	1.26	1.53*	1.14	0.88
Flu Shot	0.8*	0.83*	1.08	1	1.16*	1*	0.77
Pneumonia	0.69*	0.55*	1.31*	1.45*	1.56*	0.89*	0.46*

Notes: $\star$ means significance at 5% level of significance, NHB=Non-Hispanic Black, H=Hispanic

Table 2.7: Adjusted OR of diabetes self-care management and access to quality of care in Texas controlling for age, income, education, marital status.

	<50k	>50k	<49	<64	>65
Physical activity	1.13*	1.71*	0.68*	0.62*	0.69*
Home glucose testing	0.93	0.83	1.06	1.1	1.14
Home foot care	0.89	0.87	1.22	1.31	1.37
Attended diabetes class	1.1	1.35*	2.2	2.1	1.63
Smoking	0.83	0.63*	0.87	0.43*	0.13*
A1C test by provider	1.2	1.59*	1.84	2.13	1.8
Foot exam by provider	1.01	1.15	1.55	1.98*	2*
Dilated eye exam by provider	1.15	1.48*	1.44	2.16*	4.05*
Flu Shot	1.17*	1.3*	1.14	1.82*	2.86*
Pneumonia shot	1.17*	1.13	1.44	2.41*	5.21*

Notes: $\star$ means significance at 5% level of significance

2.6 and 2.7 Indicate that there is a significant difference for self-care management and access to quality of care of diabetes patients in Texas based on age, income, education, and marital status. In term of age, I observe that patients

over 65 except for physical activity are more likely to perform self-care management successfully than others. They have better access to quality of care compared to the rest and this may be the fact that in the USA 65-year-old individuals and others have Medicare. This finding suggests that White non-Hispanics who are getting diabetes at older age compare to Black non-Hispanics and Hispanics are doing better in term of self-care management and access to quality of care because they have Medicare. Therefore, having access to great health care may be the solution to reduce racial disparity among diabetes patients in Texas, and it will also reduce considerably the costly complications due to diabetes. Except for physical activities and foot exam by provider, women are doing much better than men. College graduate are doing better than people with only high school diploma and much better than people without high school diploma. Employed people are less likely to successfully perform self-care management and have lower access to quality of care compare to non-employed people. I observe those households who make over $\$50,000.00$ a year are having a better access to quality of care than the rest. Yet, even after controlling for those socio-economic factors, racial disparity still exists mainly for flu shot and pneumonia shot. Another concern of this paper is to examine how having access to health care and taking diabetes self-management classes help to reduce race disparity in term of self-care management and access to quality of care among patients with diabetes in Texas. I then run model (1) we perform specification 1 on diabetes patients who had diabetes self-care management classes and access to quality of care. The results are presented in Table 9.

Table 2.8 displays the demographic characteristics of adults with diabetes who have taken a diabetes self-care management and have access to health care in Texas by race 2008-2013. It is a similar demographic characteristic than the adults with diabetes in Texas.

On table 2.9, we observe an improvement of self-care management across all

Table 2.8: Demographic characteristics of adults with diabetes who have taken a diabetes self-care management and have access to health care in Texas by race 2008-2013

$N = 1592$	NHW $n = 1007$	NHB $n = 191$	H $n = 394$
Age			
18-34	1.09	16.48	2.03
35-49	16.48	24.61	26.14
50-64	28.8	35.6	29.19
65+	53.62	39.27	42.64
Women	58.29	66.49	66.5
Education			
< High School	6.27	8.42	36.25
HS GED	26.86	29.47	26.73
Some college+	66.86	62.11	37.01
Income			
< 25,000	29.68	45.78	59.09
25,000 to < 50,000	29.11	30.12	21.81
50,000+	41.21	24.1	19.09
Married	56.54	34.21	53.94
Employed	37.81	31.75	36.64
Health status			
Exc./ VG/ G	57.97	53.16	36.76
Fair/ Poor	42.03	46.84	63.24
Has a regular provider	100	100	100
BMI			
<25(normal)	17.16	8.6	14.28
>25(Obese and Overweight)	82.83	91.4	85.7
Takes insulin			
Yes	31.71	41.36	34.77

Notes:NHW=Non-Hispanic white, NHB=Non-Hispanic Black, H=Hispanic

,

races among diabetes patients who have taken a diabetes self-care management and have access to health care in Texas from 2008-2013. We will then run the regression to see if the race disparity disappears or remain for self-care management. 2.10 we observe a small improvement of access to quality to care across all races among diabetes patients who have taken a diabetes self-care management and have access to health care in Texas from 2008-2013. We then run the regression to see if the

Table 2.9: Comparison of diabetes patients‘ self-management by race/ ethnicity among diabetes patients who have taken a diabetes self-care management and have access to health care in Texas from 2008-2013

	NHW	NHB	H	P-value
Home glucose testing				0.152
>Once daily (yes)	69.67	73.54	74.55	
<once daily or Never (No)	30.33	26.45	25.45	
Home foot care				0.011
>Once daily	67.07	77.96	67.1	
< once daily or Never (No)	32.93	22.04	32.9	
Attended diabetes class				0
Yes	1	1	1	
No	0	0	0	
Smoking				0.006
Yes	20.2	33.33	29.11	
No	79.8	66.66	70.88	
Physical activity				0.967
Yes	61.06	61.78	61.68	
No	38.94	38.22	38.32	

Notes: NHW=Non-Hispanic White, NHB=Non-Hispanic Black, H=Hispanic

race disparity disappears or remain for access to quality of care. This specification can be written as: Where i is individual i , Bl_i is equal to one if individual i is black person and zero if not, and HS_i is equal to one if individual i is Hispanic and zero if not, here white race is the control variable, where The second specification consists of running a logistic regression where the dependent variable Y takes on one (with π as probability) or zero, it takes one if the patient took diabetes classes and zero if not, say, and the independent variable are race, sex, education, marital status, employment status, age, and Income. This second specification can be modeled as: where i is individual i, Bl_i is equal to one if individual i is black person and zero if not, and HS_i is equal to one if individual i is Hispanic and zero if not, here white race is the control variable; sex_i is equal to one if individuali is a male and zero if not; HSG_i is equal to one if individual i is a high school graduate and zero if not; CG_i is equal to one if individual i is a college graduate and zero if not; Ma_i is equal

Table 2.10: Comparison of diabetes patients access to quality of care by race/ ethnicity among diabetes patients who have taken a diabetes self-care management and have access to health care in Texas from 2008-2013

	NHW	NHB	H	P-value
AIC test by provider				0
Yes	74.58	71.1	63.49	
NO	25.42	28.9	36.51	
Foot exam by provider				0
Yes	0.8	92.35	76.58	
No	19.94	7.65	23.42	
Dilated eye exam by provider				0.125
Yes	74.72	81.58	76.41	
No	25.28	18.42	23.59	
flu shot				0
Yes	74.15	57.75	67	
No	25.85	42.25	33	
pneumonia shot				0
Yes	70.93	56.5	53.5	
No	29.07	43.5	46.49	

Notes: NHW=Non-Hispanic White, NHB=Non-Hispanic Black, H=Hispanic

Table 2.11: Adjusted odds Ratio of self-care management, access to quality of care and health output controlling diabetes patients who have taken a diabetes self-care management and have access to health care in Texas from 2008 to 2013.

	BNH Odd ratio,95% CI $n = 1023$	H odd ratio, 95% CI $n = 3338$	Dummy
Physical activity	1.06,(.93-1.22)	.93,(.85-1.01)	1.18*(1.06-1.32)
Home glucose testing	1.31*,(1.02-1.69)	1.22*,(1.04-1.44)	1.85*(1.59-2.14)
Home foot care	1.56*,(1.19-2.03)	.77*, (.65-.903)	1.34*(1.16-1.56)
Smoking	1.87*, (1.52-2.3)	1.32*,(1.14-1.54)	.86(.71-1.03)
A1C test by provider	0.86,(.67-1.12)	.59*,(.45-.63)	2.13*(1.83-2.49)
Foot exam by provider	2.1*, (1.53-2.92)	.71*,(.59-.84)	2.68*(2.28-3.16)
Dilated eye exam by provider	1.33*,(1.02-1.74)	.96,(0.82-1.14)	1.85*(1.59-2.16)
Flu Shot	0.64*,(.56-.74)	0.64*, (.58-.70)	1.72*(1.53-1.93)
Pneumonia Shot	o.56*,(.48-.64)	0.36*,(.32-.39)	1.22*(1.09-1.38)
General Health	0.65*, (.57-.74)	0.46*, (.42-.5)	1.05(.94-1.18)
BMI	0.69*, (.56-.84)	0,72*, (.63-.82)	1(.86-1.17)

Notes: $\star$ means significance at 5% level of significance, NHB=Non-Hispanic Black, H=Hispanic

to one if individual i is married and zero if not; Ep_i is equal to one if individual i is employed and zero if not; $AL49_i$ is equal to one if individual i is between 15 and 49 years old and zero if not; $AL64_i$ is equal to one if individual i is between 49 years

old and 64 and zero if not; $AL65_i$ is equal to one if individual i is more than 65 years old and zero if not; IL_i is equal to one if individual i earns between $\$25,000.00$ and $\$50,000.00$ and zero if not; IG_i is equal to one if individual i earns more than $\$50,000.00$ and zero if not, where

Table 2.12: Adjusted odds Ratio of self-care management, access to quality of care controlling for diabetes patients who have taken a diabetes self-care management and have access to health care in Texas from 2008 to 2013.

	NHB	H	Dummy
Physical activity	1.22*	1.26*	1.11
Home glucose testing	1.21	1.03	2.01*
Home foot care	1.64*	0.96	1.32*
Smoking	1.25	0.8*	0.85
A1C test by provider	1.04	0.87	2.02*
Foot exam by provider	2.45*	0.96	2.76*
Dilated eye exam by provider	1.7*	1.35*	1.73*
Flu Shot	0.8*	0.85*	1.69*
Pneumonia Shot	0.69*	0.55*	1.26*

Notes: $\star$ means significance at 5% level of significance, and the dummy is 0(if individual i did not take diabetes classes or has not access to health care) or 1(if individual i took diabetes classes and has access to health care) NHB=Non-Hispanic Black, H=Hispanic

From table 2.11 and 2.12 we can conclude that: In term of self-care management, besides smoking, I do observe a reduction of race disparity when I add the dummy to the regression. Furthermore, all the coefficients attached to the dummy are significant and positive, this implies that having taken diabetes self-care management classes and having access to quality of care lead diabetes patients to conduct a better self-care management. For instance, from table 10 individuals who have taken diabetes self-care management classes and have access to health are more likely to conduct home glucose testing compared to others with an odds ratio (OR) of 1.82. And this odds ratio increases from 1.82 to 2.01 when we control for socio-economic factors on table 11. As far access to quality of care is concerned, I still observe some race disparities even though there are some reductions. Furthermore, all the coefficients attached to the dummy are significant and positive

which implies that having taken diabetes classes and having access to quality of care lead to a better access to quality of care for diabetes patients. For instance, from table 10 individuals who have taken diabetes self-care management class and have access to health are more likely (OR: 2.68) compared to others to have their Foot exam by provider. This odds ratio increases from 2.68 to 2.76∗ when we control for some socio-economic factors on table 11. However, the disparity persists on flu shot and pneumonia shot. I recommend that flu shot, and pneumonia shot be made more available to all diabetes patients in Texas, and perhaps made them free for patients who cannot afford those two shots.

2.6 Discussion and Conclusion

My findings suggest that Hispanics with diabetes in Texas are still struggling to improve their self-management and access to quality of care compared to Black and White non-Hispanic. However, having health care coverage and having taken a diabetes self-management class improves the overall self-management and quality of care across all races. In terms of self-management, the racial disparity reduces significantly; this is not the case when it comes to access of quality of care. I recommend that Texas expands good quality insurance coverage and diabetes self-management classes to all diabetes patients as soon as they are diagnosed with the condition, mainly for minorities. This will considerably reduce the racial disparities and the number of diabetes health complications by enhancing diabetes patient outcomes. In an economic sense, this will probably be cost effective for the state and the tax payers in the long-run.

However, this paper has some limitations. In fact, when the BRFSS is conducting a phone survey, there may be a bias. This implies that respondents can report good behaviors which may not reflect their actual behaviors. Additionally, I

would like to be able to incorporate respondentsŠ diet as part of the self-management. I hope that in the future, the BRFSS will incorporate more precise questions on individual diets.

Chapter 3

Socio-economic study of breast cancer survival using survival analysis: a 20 years cohort study in Texas

3.1 Introduction

Based on the Centers of Disease Control and Prevention (CDC) report, breast cancer is the second most common cancer after skin cancer among women in the United States across all races and ethnicities. In fact, in 2014, 236,968 women and 2,141 men in the United States were diagnosed with breast cancer. Additionally, 41,211 women and 465 men in the United States died from it (CDC). In the U.S., one in eight women will be diagnosed with invasive breast cancer during her lifetime according to National Cancer Institute. In this study, I focus on the state of Texas, where the total financial cost of hospitalization due to breast cancer in 2016 was around 252 million. 13,183 Texas Medicaid beneficiaries have received acute care

for breast cancer in the fiscal year 2014 with an average expenditure per beneficiary of $4,637.63$. In 2015, there were $16,136$ new cases of female breast cancer and, in the same year, it claimed the life of $2,849$ patients in Texas. It is a burden to tax payers, a public health issue and cancer mortality is associated with huge productivity loss. In addition the number of new cases is still increasing every year and the population is getting older. I take advantage of this 20 years cohort study of cancer survival data in Texas from the Texas Cancer Registry (TCR), Cancer Epidemiology and Surveillance Branch, Texas. I include only the diagnosis of breast cancer patients recorded in 1995 and 1996, and we follow them up to 2015 with a large sample size of $21,626$ survivors to study the main drivers that can explain why some breast cancer patients live longer than others. I performed a 2-step survival analysis on the entire data sample, the first consisted of running a log-rank test, a survival time regression and then the second is consisting of running Cox proportional hazards regression . Finally, I conducted a multinomial logistic regression analysis, I divided the data based on the survival time into 4 groups where group 1 is composed by survivors who live up to 5 years, group 2 is made by survivors living between 5 and 14 years after diagnostics. Group 3 is only composed of survivors who died after 14 years, and finally, group 4 is formed by only survivors who are still alive after the end of the study in 2015. The incomplete observed responses are censored and dismissed. All those different techniques help better understand the main drivers of breast cancer survival and the disparities that may exist within those variables. Knowing those factors makes it easy to identify the riskiest survivors of breast cancer, and I am able to make effective and resourceful recommendations to help new patients have a much longer life. This is a unique approach and to my best knowledge I am the first one to explore the breast cancer survival this way. The rest of the paper is organized as follows: a background and a literature review, then I have a data description and characteristics, econometrics

model and results, and finally a conclusion.

3.2 Background and Literature review

According to the Centers of Disease Control and Prevention, cancer is a disease in which cells in the body grow out of control. When cancer starts in the breast, it is called breast cancer. It can spread outside the breast through blood vessels and lymph vessels. When breast cancer spreads to other parts of the body, it is said to have metastasized. The most common kinds of breast cancer are invasive ductal carcinoma, and Invasive lobular carcinoma. The main risk factors of breast cancer listed by the CDC, Gull Z, Anwar Z, Sherazi BA (2017), and Kelsey,JL, Horn-Ross, PL (1993) who identify different risk factors which lead to breast cancer, include older age, genetic mutations, early menstrual period, late or no pregnancy, starting menopause after age 55, not being physically active, being overweight or obese after menopause, having dense breast, using a combination of hormonal therapy, taking oral contraception‘s, family history of breast cancer, drinking alcohol, smoking, cigarettes and women from high socioeconomic status . It is, therefore, very difficult to totally prevent breast cancer. Yet, a mammogram is a screening tool that can be used by physicians to find breast cancer early, sometimes up to three years before it manifests the first physical symptoms. A mammogram is an X-ray picture of the breast. A mammography can also be used as a diagnostic tool to confirm the presence of breast cancer. However, from information provided by National Cancer Institute, a mammography is not perfect since errors in diagnostics can happen. In fact, a mammography can make false positive results mostly for young women and women with dense breast size. It can also find false negative results that can delay treatment and lead to a false impression of security for affected women. Finally, the risk of harm from the radiation exposure from a mammography is low, but repeated

x-rays have the potential to cause cancer. Cathy J.Bradley, K Robin Yabroff and al. (2008). Recently, there is a new type of mammogram for breast cancer detection tomosynthesis called 3 D mammogram. It reduces the need for additional screenings, and improves detection of breast cancer mainly for dense breast tissue. In January 2018, Texas required all insurance companies to cover 3 D mammogram for Texas patients. We need to continue investing in research for better screening and diagnostic tools in the future which will lead to early detection and reduce breast cancer mortality. Since breast cancer can be a recurrent or chronic disease, the main goal for survivors is to maximize their survival time. Therefore, it is important to understand why some survivors live longer than others. Besides, cancer mortality is associated with a huge productivity loss, as proven by Cathy J.Bradley, K Robin Yabroff and al (2008) who study productivity costs of cancer mortality in the United States: 2000-2020.They developed models using the human capital approach, which relies on earnings as a measure of productivity, to estimate the value of productivity loss as a result of cancer mortality. A model that included costs of care giving and household work were used. The values of forgone earnings for employed individuals and imputed forgone earnings for informal care giving were then estimated for the years 2000Ű2020. The annual productivity cost from cancer mortality in the base model was approximately $115.8 billion in 2000; the projected value was $147.6 billion for 2020. Including imputed earnings loss due to care giving and household activity increased the base model total productivity cost to $232.4 billion in 2000 and to $308 billion in 2020. Their paper is useful to our analysis, as it shows productivity costs and all related costs due to cancer mortality. Therefore, it is important to understand that the longer cancer survivors live the more it helps minimize the loss of productivity. Furthermore, considering the lifetime earnings of women and men in the USA based on their educational attainment, I can observe that an early death has substantial productivity loss associated with it, see Fig 3.1 a

and Fig 3.2 for more details. For instance, a woman with a bachelorŠs degree has an expected lifetime earnings estimate of $1.32 million (in net income) and $1.43 million (in gross income), then, it is extremely important that policy makers take this into account and have policies that will lead to an extended length of survival for breast cancer survivors, even if those policies imply spending more money on medical care and facilitating the access to quality of care for new diagnose breast cancer patients.

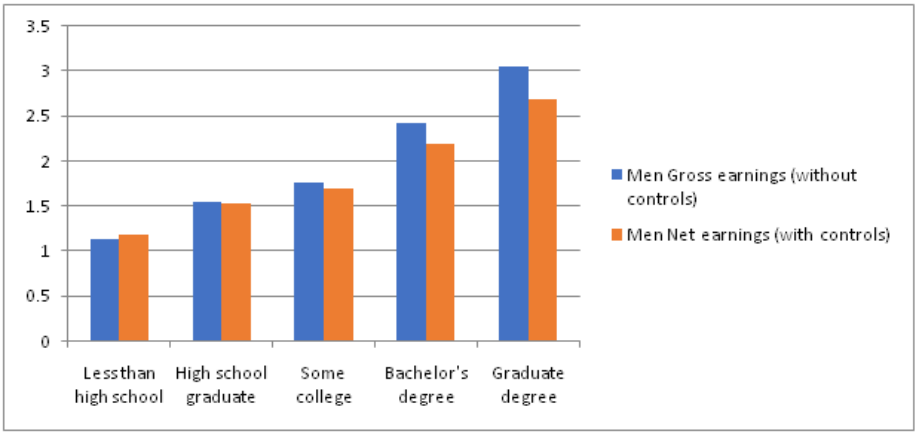

Figure 3.1: Men estimated lifetime earnings by educational attainment (in millions of dollars)

In addition, there are costs associated with breast cancer survivors, according to Donatus U. Ekwueme, Justin G. Trogdon, Olga A. Khavjou, Gery P. Guy Jr. (2016) who study productivity costs associated with breast cancer among survivors aged 18Ű44. They find that, per capita, younger women with breast cancer had annual losses of $2,293 (95% CI=$1,069, $3,518) from missed work and $442 (95% CI=$161, $723) from lack of productivity at home. Total annual breast cancerŰassociated productivity costs for younger women were $344 million (95% CI=$154 million, $535 million). Older women with breast cancer had lower per capita work loss productivity costs of $1,407 (95% CI=$899, $1,915) but higher total work loss productivity costs estimated at $1,072 million (95% CI=$685 million, $1,460 million) than younger women. As a result, the benefit of survival

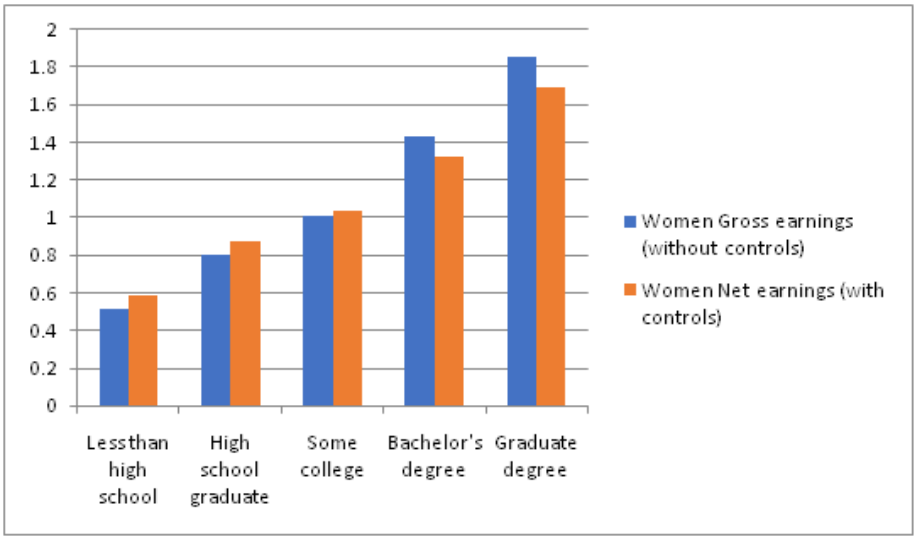

Figure 3.2: Women estimated lifetime earnings by educational attainment (in millions of dollars)

exceeds the cost of death in terms of finance, economic, labor productivity.

We also need to mention that most of those women are someone daughters, mothers, or grandmothers, which implies emotional and social loss as well. In the case of the premature death of a young mother, for instance, the hiring of a caregiver to tend to her children can be challenging and costly. In the literature, personal characteristics such as race, age, lifestyles, physical activities, diet, and treatment, and stage of detection of the cancer, such as mammography, radiotherapy or chemotherapy, can explain the survival time for breast cancer patients. Breast cancer survival and race: Anne McCarthy, Jianing Yang, and Katrina Armstrong (2015) study an increasing disparity in breast cancer mortality from 1979 to 2010 for African American women aged 20 to 49. They find that disease-specific mortality rates declined over time for selected conditions, but mortality rates were persistently higher for black women for breast cancer, cervical cancer, colorectal cancer, ischemic heart disease, and stroke. The mortality rate ratio increased for breast cancer across the studyŠs period. The annual mortality rate ratio for black women compared to white women was 1.36 in 1979 compared to 2.00 in 2010.

Furthermore, Neal A. Chatterjee, Yulei He, and Nancy L. Keating (2013) assess a racial difference in breast cancer mortality by stage at diagnosis, since mammography became available. They calculated adjusted odds of distant (versus local or regional) tumors for $143,249$ white and $13,571$ black women aged 50 to 69 years, diagnosed with breast cancer between 1982 and 2007. In conclusion, in the mammography era, racial disparities remain in stage at diagnosis which implies that early diagnostic is a key to better prognostics in the case of breast cancer.

Finally, J. William Eley, Holly A. Hill, Vivien W Chen et al. (1994) examine the ability of recognized prognostic factors for breast cancer to account for the observed poorer survivorship in blacks compared to their white counterparts. Their multivariable survival models were used to estimate the hazard ratio (relative risk of mortality) for blacks compared to whites, adjusting for various combinations of potential explanatory factors. They find that approximately 75% of the racial difference in survival was explained by the prognostic factors studied. Socio-demographic variables appeared to act largely through racial differences in stage at diagnosis, which may be amenable to change through improved access to and use of screening for black women. Therefore, early screening is vital to improve survival time of breast cancer patients. In fact, Robert S. Kirsner, Fangchaoet al. (2010) study earlier stage at diagnosis and improved survival among Medicare health maintenance organizations (HMO)patients with breast cancer. They used a linkage of two national databases, the Medicare database from the Centers for Medicare and Medicaid Services (CMS) and the National Cancer Institute‘s (NCI) Surveillance, Epidemiology, and End Results (SEER) program database to evaluate differences in demographic data, stage at diagnosis, and survival in patients with breast cancers over the period $1985-2001$. Their results have indicated that Medicare patients enrolled in HMOs were diagnosed at an earlier stage of diagnosis than fee for services (FFS) patients. HMO patients diagnosed with breast cancer

had improved survival rates, and these differences remained even after controlling for potential confounders. Specifically, breast cancer patients enrolled in HMOs had 9% increased probability of survival (hazard ratio [HR] 0.91, 95% confidence interval [CI] $0.88 - 0.93$) than their counterparts enrolled in FFS. These findings persisted even when patients had a cancer diagnosis before their breast cancer. These findings show that improved survivorship among breast cancer patients in HMOs compared to FFS is likely due to a combination of factors, including but not limited to earlier stage at the time of diagnosis, which reinforced the idea of early diagnosis is a key to improving survival time for breast cancer patients. Fiona McKenziea*, Lis Ellison-Loschmanna and Mona Jeffreysb (2011) who investigate reasons for ethnic inequalities in breast cancer survival in New Zealand find that inequalities persisted after adjustment for subtype variables (ER/PR/HER2), while adjusting for access to care variables (extent/size) eliminated the ethnic inequalities in excess mortality. They conclude that ethnic disparities in breast cancer survival in New Zealand can be attributed to deprivation and differential access to health care rather than differences in breast cancer subtypes. I then control for race and ethnicity in my study to see if there is any disparity on survival time due to race, given that Texas is diverse in regard to race and ethnicities.

Age at diagnosis and survival of breast cancer: According to Charles E Phelps in his book Health Economics (1992) any human being stock of health wears out over time and this route is call aging, and as the stock of health diminishes enough, we are losing our capacity to function and finally we are going to die. Therefore, age at diagnosis is an important factor for breast cancer survivorship. Adami, Malker, Holmberg, Persson, and Stone (1986) analyzed the relation between age at diagnosis and relative survival in $57,068$ women in Sweden in whom breast cancer was diagnosed in 1960 to 1978. Their findings suggest that women who were 45 to 49 years old had the best prognosis, with a relative survival exceeding that of the

youngest patients. They also find that the relative survival declined noticeably after the age of 49 mainly in women 50 to 59 and women older than 75. They conclude that the long-term annual mortality rate due to breast cancer approached 1 to 2 percent at the premenopausal ages but exceeded 5 percent throughout the period of observation in the oldest age group. Additionally, Candyce H. Kroenke, Bernard Rosner, Wendy Y. Chen, Ichiro Kawachi, Graham A. Colditz, and Michelle D. Holmes (2004) studied changes in physical and psychosocial function before and after breast cancer by age at diagnosis. From a large sample size of $122,969$ women from the Nurses' Health Study (NHS) and NHS 2, ages 29 to 71 years, who responded to pre- and post-functional status assessments. Among them, $1,082$ women were diagnosed with breast cancer between 1992 and 1997. Functional status was measured using the Medical Outcomes Study Short Form 36 (SF-36). Mean change in health-related quality of life (HRQoL) scores was computed across categories representing the combination of incident breast cancer (yes or no) and age at diagnosis (40, 41 to 64, or 65 years). They conclude that compared to women 40 years without breast cancer, women with breast cancer experienced significant functional declines. Young (age 40) women who developed breast cancer experienced the largest relative declines in HRQoL (as compared to middle-aged and elderly women) in multiple domains, including physical roles(18.8 v 11.5 and 7.5 points, respectively), bodily pain(9.0 v 2.7 and 2.7 points), social functioning (11.3 v 4.3 and 4.4 points), and mental health (3.1 v 0.0 and 0.4 points). Much of the decline in HRQoL among elderly (age 65) women with breast cancer was age related. In my study, I will then see if age at diagnostic still an important factor of survival time for breast cancer patients in Texas, and which age ranges are the most at risk. Furthermore, obesity, smoking and alcohol, and the existence of comorbidity are also influencing negatively the survival time for breast cancer patients. Obesity and breast cancer: In terms of obesity, most recent studies agree that obesity is a

determinant for both getting breast cancer and surviving long from it. In fact, Maliniak, Patel, McCullough, Campbell, Leach, Gasptur and Gaudet (2017) study obesity, physical activity, and breast cancer survival among older breast cancer survivors in the Cancer Prevention Study-II Nutrition Cohort for $5,254$ between 1992 and 2013. They conclude that Higher BMI, pre- or post-diagnosis, was associated with a higher risk of breast cancer-specific mortality in older patients, independent of comorbidity and stage at diagnosis. Likewise, Connor et al. (2016) conducted a study on the relationship between obesity and quality of life (QOL) among Hispanic and non-Hispanic white breast cancer survivors and population-based controls from the Long-term Quality of Life Study, a 12- to 15-year follow-up study of breast cancer cases/survivors and controls from New Mexico ($n = 451$). They use Body Mass Index (BMI) as a measure of obesity, and follow-up interviews were modeled with composite scores for physical and mental health from the SF-36 Quality of Life Survey. Interaction between ethnicity and BMI and change in BMI were evaluated. All models were adjusted for age, ethnicity, Charlson Index, depression, fatigue, and physical activity. They found that baseline obesity ($b = -6.58$, $p = 0.04$) was significantly associated with decreased mental health among survivors, but not among controls. Obesity at baseline and follow-up was significantly associated with decreased physical health among survivors (baseline $b = -10.51$, $p = 0.004$; follow-up $b = -7.16$, $p = 0.02$) and controls (baseline $b = -11.07$, $p = 0.001$; follow-up $b = -5.18$, $p = 0.04$). No significant interactions between ethnicity and BMI were observed. Smoking and breast cancer: Unhealthy habits, such as smoking, are a factor for breast cancer survival length. First, not smoking cigarettes and avoiding exposure to secondhand smoke has multiple health benefits. Secondly, numerous studies have showed that smoking can lead to early death among breast cancer survival. For instance, Yoichiro et al. (2015) studied use a prospective cohort to study smoking and survival after breast cancer diagnosis in

Japanese women. They followed between January 1997 and December 2007, 941 female patients aged 21 years or over at the Miyagi Cancer Center Hospital (MCCH) that were newly diagnosed as having breast cancer. They found that only a limited number of patients with a longer duration of smoking might have a higher risk of all-cause and breast cancer-specific death by stratification of hormone receptor status among premenopausal women. It has been suggested that this potential relationship might be related to the estrogen-like substances in active tobacco smoke, which can exert estrogenic effects. Therefore, smoking should be avoided if breast cancer survivors would like to extend their survival time. Alcohol and breast cancer survival: Evidence for the association between alcohol consumption with breast cancer survival are conflicting in the literature; excessive consumption of alcohol beverage can have a negative impact for any one health mainly in the case of chronic condition such as breast cancer survival. For instance, Yun-Jiu Gou, Ding-XiongXie, Ke-Hu Yang, Ya-Li Liu, Jian-Hua Zhang, Bin Li, Xiao-Dong He (2013) conducted a Meta-analysis of Cohort Studies on alcohol consumption and breast cancer survival. After including 25 cohort studies, they find that only alcohol consumption of less than 20 g/d was associated with increased breast cancer mortality, but not with increased breast cancer recurrence. However, Anne M. Weaver, Susan E. McCann, Jing Nie, Stephen B. Edge, Thomas H. Nochajski, Marcia Russell, Maurizio Trevisan, and Jo L. Freudenheim (2013) examined the link between alcohol intake from all sources, assessed by cognitive lifetime drinking history, and all cause breast cancer mortality among women with breast cancer ($N = 1,097$) who participated in a population-based case-control study. Weaver et al. concluded that there are no associations between drinking status or total volume of alcohol intake and breast cancer or causing mortality. High-intensity alcohol consumption may be associated with decreased survival rates in postmenopausal women with breast cancer. Co morbidity and breast cancer

survival: The existence of co morbidity, such as diabetes and asthma, is a factor of early death due to breast cancer. In fact, Santorelli et al. (2017) study racial differences in the effects of co morbidity on breast cancer-specific survival. They are using retrospective cohort study of $68,090$ women 66+ years, who were diagnosed with stage $I - III$ breast cancer in the United States from 1994 to 2004. Their findings suggest that diabetes without complications was associated with a significantly increased hazard for breast cancer-specific death among white breast cancer patients, and aggressive tumor characteristics explained some of this effect. Also, hyper tension was associated with an earlier stage of breast cancer diagnosis for both black and white women ($p < 0.01$). For black women, hyper tension was also associated with a less aggressive tumor grade ($p = 0.02$) and for white women hyper tension was associated with a less aggressive hormone receptor status ($p < 0.01$). Stage and grade at diagnostic and breast cancer survival: the stage and the grade of the cancer at time of diagnosis play a major role on the length of survivorship and on the treatment the patient is expected to receive. In fact, staging is used to evaluate the size of a tumor, whether it has spread and how far it has spread. Understanding the stage of the cancer helps physicians predict the possible outcome and layout a treatment plan for each patient. The main method used for defining the stage of a cancer is the TNM (tumor, nodes, and meta-stasis) system. The TNM system is often used to categorize cancer into five stages: stage 0, stage 1, stage 2, stage 3 and stage 4. In stage 0 abnormal cells are found in the lining of the breast duct or in the lobules of the breast. It is a non-invasive. Stage 1 usually means that a cancer is relatively small and contained within the breast. Stage 2 usually means the cancer has not started to spread into surrounding tissue, but the tumor is larger than in Stage 1. Sometimes Stage 2 means that cancer cells have spread into lymph nodes close to the tumor. Usually when the cancer is larger, and it may have started to spread into surrounding tissues and there are cancer cells

in the lymph nodes in the area, we call it stage 3. Finally, in stage 4, Cancer has spread to other parts of the body including bones and organs, such as lungs, brain, or liver. Therefore, the stage at time of diagnostic is key for survival time. Some studies have confirmed this. According to the American Cancer Society, the rate of survival from breast cancer by stage after 5 years from diagnosis is the following: close to 100% for stage0, 98% for stage1, 93% for stage 2, 72% for stage 3, and then decrease dramatically for stage 4 since only 23% are likely to survive. We need to diagnose the tumor at the early stage if we aim to improve the survival time for breast cancer patients. I control for stage at diagnostic in this study.

Beside the stage of the tumor, the grade of the tumor is an important factor for both survival and treatment options. The grade of a tumor indicates what the cells look like and gives an idea of how quickly the cancer may grow and spread. Tumors are graded between 1 and 3. Grade 1: the cancer cells look small and uniform like normal cells and are usually slow-growing compared to other grades of breast cancer. Grade 2: the cancer cells are slightly bigger than normal cells, varying in shape and are growing faster than normal cells. Grade 3: the cancer cells look different than normal cells and are usually faster-growing than normal cells. Neighborhood poverty level and breast cancer survival: Access to quality care is crucial for breast cancer prognostics. Since there is some cost associated with taking good care of survivors, cost of healthy food, need to maintain physical activity such as going to the gym, have access to yoga studios. As a result, financial status plays an important role, and to account for it, I control for neighborhood poverty level based on the census tract of diagnosis address considering that the data does not provide the survivors income or education level. In fact, from the literature, the neighborhood poverty level where the survivors live and their length of survival time correlate. Boscoe et al. (2010) studied 3 million tumors diagnosed between 2005 and 2009 from 16 states plus Los Angeles, and they were assigned 1 of 4 groupings

based on the poverty rate of the residential census tract at time of diagnosis. The sex-specific risk ratio of the highest-to-lowest poverty category was measured using Poisson regression, adjusting for age and race for 39 cancer sites. Boscoe et al. (2010) finds that there is a negligible relationship between local poverty rate and cancer incidence overall, but 32 of 39 individual cancer sites show such an association, with 14 sites associated with higher poverty and 18 sites associated with lower poverty. This includes 19 sites with stronger evidence of a relationship as indicated by a monotonic increase or decrease across all 4 poverty categories. The combined categories of HPV-associated and tobacco-associated cancers also are positively associated with poverty. I need to mention that breast cancer incidence rate is higher among women with high standards of living due to their lifestyles such as getting married at an older age or not getting married at all, having baby after 30 years old, not breastfeeding and perhaps this high incidence can be explain by Charles E Phelps hypothesis of "life in a fast Lane". This hypothesis suggests that an increase in people incomes comes at the expense of health hazardous industrial processes which cause a decline in health status. I also look at where the survivors live such as metropolitan, urban, or rural areas since that can impact what types of care the survivors receives. For instance, there are more cancer hospitals in big cities than in rural areas and healthier food stores in cities than in rural areas. All these factors can have a significant impact on breast cancer patients' prognostic.

3.3 Data Description and Characteristics

This is a retrospective cohort study in Texas. Data for breast cancer cases was provided by the Texas Cancer Registry (TCR), Cancer Epidemiology and Surveillance Branch, Texas Department of State Health Services,1100 West 49th Street, Austin, TX 78756, https://www.dshs.texas.gox/tcr/. Since November 2017,

this data has been submitted to the National Program of Cancer Registries (NACR) and North American National Association of Central Cancer Registries (NAACCR). This data contains cases of incidence from 1995 up to 2015. TCR reported a total of $2,083,875$ incidence files of Texas residents with malignant and in-situ cancers from $1995-2015$. TCR is part of the National Cancer InstituteŠs Surveillance Epidemiology and End Results (SEER), which holds high quality information on cancer. I include only the diagnosis of breast cancer patients recorded by the SEER Program in 1995 and 1996, and we follow them up to 2015. The Texas Cancer Registry program recorded $15,968$ cases of breast cancer in white non-Hispanics women and $2,223$ cases of breast cancer in black non-Hispanics women diagnosed, $3,204$ Hispanics and 231 others. In this data, 99.40% are women. I then use the length of survival for each patient which is by months. I perform 2 types of analyses. In Texas, the total population in 1995 was $18,723,991$ where 58.59% are White non-Hispanics, 27.37% are Hispanics, 11.67% are Black non-Hispanics and 2.36% are others. Additionally, 50.52% are female and 49.48% are male. First, I use a survival analysis on length of survival to deeply understand the correlation between survival length and some variables such as race, age, stage, grade, and poverty level of the survivors. In table 1 below I despite the data descriptive and characteristics of breast cancer survivors diagnose in 1995 and 1996 in Texas and since 99% of breast cancer patients are female, I will provide a data descriptive of female in Texas in 1995 by race and age.

From 3.1, I notice that White non-Hispanics are much older (46.76% are 65 and older), than Black non-Hispanics (33.88% are 65 and older), than Hispanics (31.4% are 65 and older), and other are much younger since only 15.56% of them are 65 and older. This age gap alone can explain survival length disparities among racial groups. In fact, a 70 year old woman may not have the same survival time than a 45 year old woman. Beside, from the general population in Texas in 1995, I remark

Table 3.1: Data descriptive and characteristics of Breast cancer survivors diagnose in 1995 and 1996 in Texas

Variables	n	NHW	FwTx	NHB	FBTx	H	FHTx	Other	FOTx
Race									
White non-Hispanic	15573	15573	5572421						
Black non-Hispanic	2134			2134	1121795				
Hispanic	3150					3150	2544590		
Other	225							225	220657
Age									
0 and 39	1626	5.81%	55.89%	13.12%	67.67%	12.98%	73.21%	12.44%	67.21%
40 and 49	3938	16.77%	14.72%	23.29%	13.04%	23.37%	11.21%	38.22%	16.68%
50 and 64	6507	30.66%	14.23%	29.71%	10.24%	32.25%	8.97%	33.78%	10.83%
65 and 74	4859	24.68%	7.86%	18.56%	4.87%	18.67%	4%	12%	3.43%
Over 75	4177	22.08%	7.3%	15.32%	4.18%	12.73%	2.61%	3.56%	1.85%
Gender									
Female	20982	99.41%	100%	99.25%	100%	99.62%	100%	97.78%	
Male	125	0.59%		0.75%		0.38%		2.22%	
Treatment									
Chemo and Radiation	3136	16.33%		14.83%		15.39%		18.31%	
Chemo only	16231	82.67%		84.67%		83.83%		80.75%	
Radiation only	179	1%		0.5%		0.77%		0.94%	
Grade									
Grade I	2545	13.59%		8.06%		7.56%		7.56%	
Grade II	6495	32.74%		24.65%		25.33%		29.78%	
Grade III	62271	27.34%		38.75%		34.95%		34.22%	
Grade IV	420	1.65%		3.09%		2.89%		2.67%	
Unknow grade	5376	24.68%		25.45%		29.27%		25.78%	
Stage									
Stage I	11307	56.82%		44.05%		44.32%		49.78%	
Stage II	5440	24.15%		29.29%		31.4%		26.22%	
Stage III	803	3.44%		4.97%		4.7%		5.78%	
Stage IV	1312	5.35%		9.98%		8.1%		4.44%	
Unknown or un-staged	2245	10.24%		11.72%		11.49%		13.78%	
Poverty level									
$0-5\%$	3832	22%		6%		6.48%		30.67%	
$5\%-10\%$	4811	26.73%		10.36%		11.65%		25.33%	
$10\%-20\%$	7172	36.1%		28.88%		27.43%		29.78%	
$20\%-100\%$	5279	15.17%		54.76%		54.44%		14.22%	
Areas where survivors lived									
Metropolitan areas	17375	80.5%		88.66%		85.87%		96.89%	
Urban areas	3396	17.64%		10.36%		13.33%		3.11%	
Rural areas	335	1.86%		0.98%		0.79%		0%	

Notes: Where FWTx means total White non -Hispanic female in Texas in 1995, FBTx means total Black non-Hispanics female in Texas in 1995, FHTx means total Hispanics female in Texas in 1995 and FOTx means total other female in Texas in 1995.
Source: general population in Texas in 1995 is from the Texas Department of State Health Services
Notes: NHW=Non-Hispanic White, NHB=Non-Hispanic Black, H=Hispanic, Chemo= chemotherapy

that Whites are much older than everyone else. In fact, only 55.89% of White non-Hispanics are less than 39 years old, while 67.67 of Black non-Hispanics, 67.21% of other and 73.21% of Hispanics are in that same age group. Moreover, only 1.85% of others are 75 year old and over while 2.61% of Hispanic, 4.18% of Black

non-Hispanics and 7.3% of White non-Hispanics are in this age group. This disparity in age distribution of female in Texas in 1995 may explain why we have a racial disparity in getting diagnosed with breast cancer when in the literature we know that most people get breast cancer after the age of 40 year old. Finally, from this data 73.86% of all diagnosed in 1995 and 1996 are White non-Hispanics, 10.12% are Black non-Hispanics, 14.94% are Hispanics and only 1.06% are other races and ethnicities. Based on grade, White non-Hispanics are being diagnose with early grade (46.33% are diagnosed with grade 1 and 2) than others (37.34% are diagnosed with grade 1 and 2), than Hispanics (32.89% are diagnosed with grade 1 and 2) and then Black (32.71% are diagnosed with grade 1 and 2). From the literature we also know that being diagnose with early grade and stage yield better prognostic. Furthermore, White non-Hispanics are being diagnose with early stage (56.82% were diagnose with stage 1) compared to Black non-Hispanics (44.05% were diagnose with stage 1), Hispanics (44.32% were diagnose with stage 1) and others (49.78% were diagnose with stage 1). I observe that White non-Hispanics and other are living in affluent neighborhood compared to Black non- Hispanics and Hispanics. In fact, 54.76% of Black non-Hispanics live in a neighborhood where the poverty lane is over 20% compared to 54.44% for Hispanics, and only 15.17% of White non-Hispanics and 14.22% of others. Those income discrepancies may play a role on survival length disparities I perform a log-rank test which is a hypothesis test to compare distributions of two samples. It is a nonparametric test and appropriate to use when the data has right censored like in the case of our study. Figure 3.3 shows the relation between survival probabilities for breast cancer patients with respect to different stages at diagnostic; from it I observe that the stage at diagnostic significantly plays a major role on the survival time. For instance, after 5 years of survival around only 9% of survivors with stage 1 died, however, at the same time around 80% of survival with stage 4 at diagnostics died. It is then crucial to

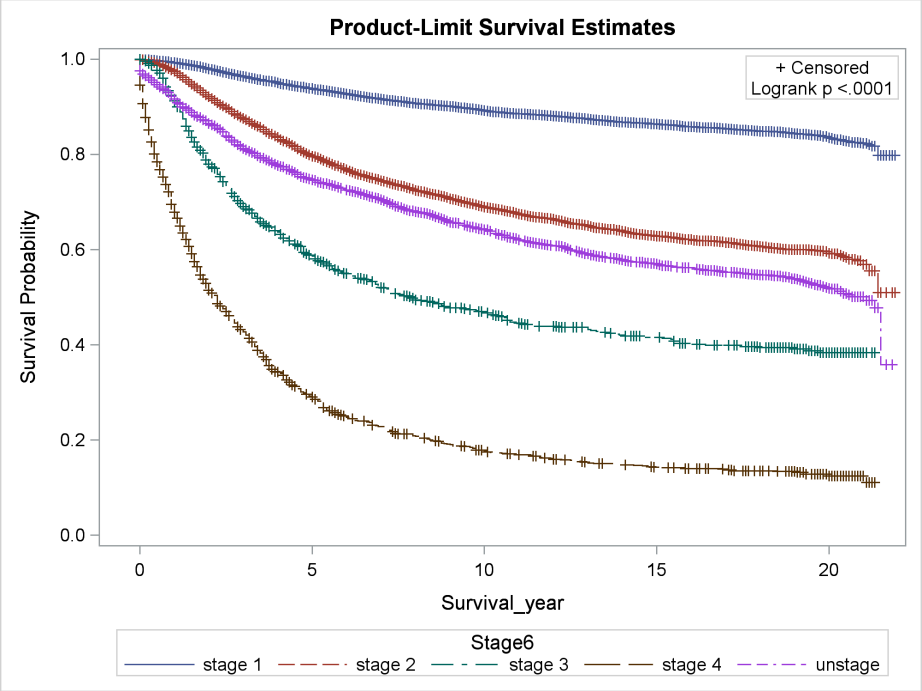

Figure 3.3: Comparing survival by stage

promote and encourage early diagnostic to improve survival time among breast cancer patients

Figure 3.4 displays the difference in the survival probability of breast cancer patients among race and ethnicity groups. I notice that there is a significant disparity of survival probability based on race and ethnicity. For example, after 5 years of survival around 30% of Black-non Hispanics survivors died, around 20% of Hispanics survivors died, around 15% of White-non Hispanics died and only 5% of other survivors died. Therefore, it is important to intensify the awareness of breast cancer within the Black- non Hispanics and Hispanics communities.

Figure 3.5 presents the relation between survival probabilities for breast cancer patients with regard to different grade at diagnostic, I note that the grade at diagnostic significantly influences the survival time. For instance, after 5 years of survival around only 5% of survivors with grade I died, however, at the same time around 35% of survival with grade 4 at diagnostics died. It is crucial to promote and

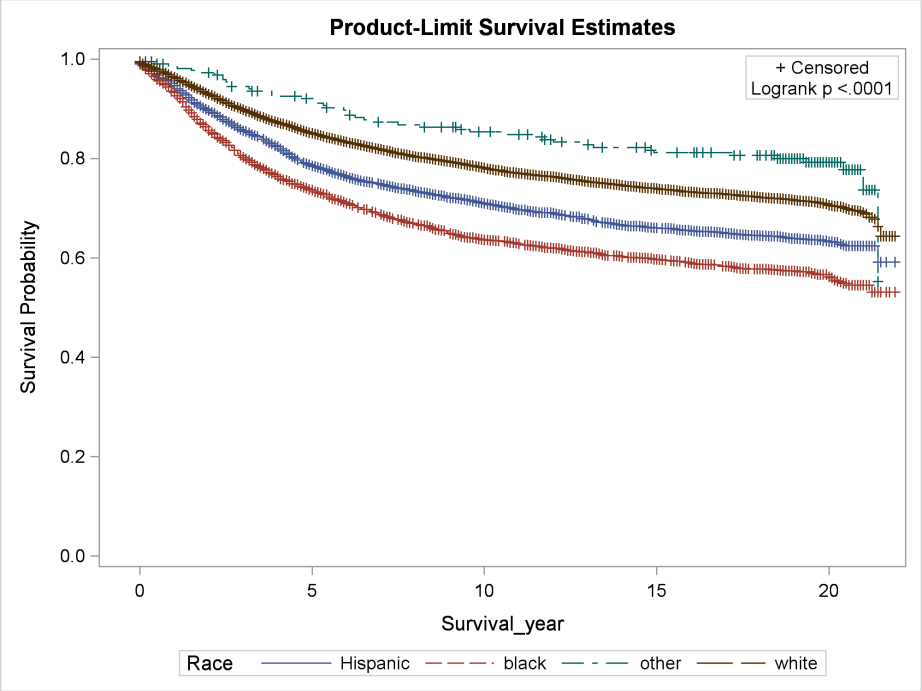

Figure 3.4: Comparing survival by Race and ethnicity

encourage early diagnostic to improve survival time among breast cancer patients.

Figure 3.6 depicts the relation between survival probability from breast cancer and the survivor‘s neighborhood poverty line; I can see that there is a significant disparity on the survival time. As an illustration, after 5 years of survival around only 10% of survivors living a high income neighborhood died, however, at the same time around 25% of survival from poor neighborhood died. We should then better assist and help survivors from under- privilege neighborhood to improve their survival time.

Figure 3.7 portrays the relation between survival probability from breast cancer and type of treatment received by the survivors; I remark that the type of treatment has a slight impact on the survival time. For instance, after 5 years of survival around 25% of survivors who receive chemotherapy died, however, at the same time around 15% of survival who receive radiation only died. This may be due to the fact that the type of treatment has a correlation with the stage and grade of the disease;

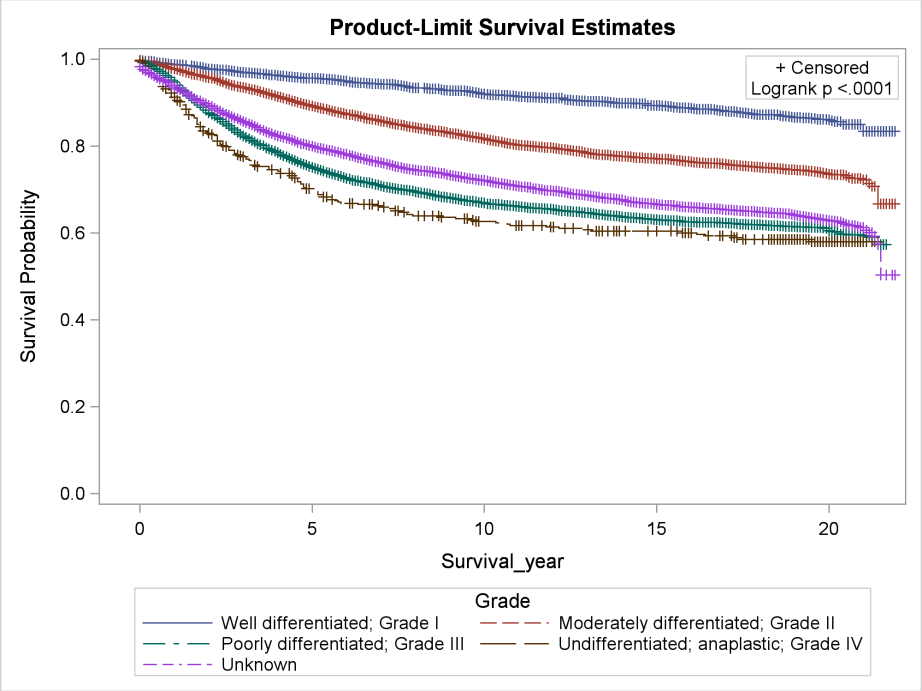

Figure 3.5: Comparing survival by grade

therefore, I cannot just suggest that all patients need to receive radiation only.

3.4 Econometrics model and results

3.4.1 Model 1: Survival Time Regression Model

$$\mathbf{log}(s_i(t)) = \boldsymbol{\beta}_0 + \boldsymbol{\beta}_1 Gender_i + \boldsymbol{\beta}_2 Race_i + \boldsymbol{\beta}_3 Age_i + \boldsymbol{\beta}_4 Treatment_i + \tag{3.1}$$

$$\boldsymbol{\beta}_5 Grade_i + \boldsymbol{\beta}_6 Stage_i + \boldsymbol{\beta}_7 IncomeProxy_i + \boldsymbol{\beta}_8 GeographicAreas_i + \boldsymbol{\epsilon}_i$$

$$H_0 : \beta_k = 0$$

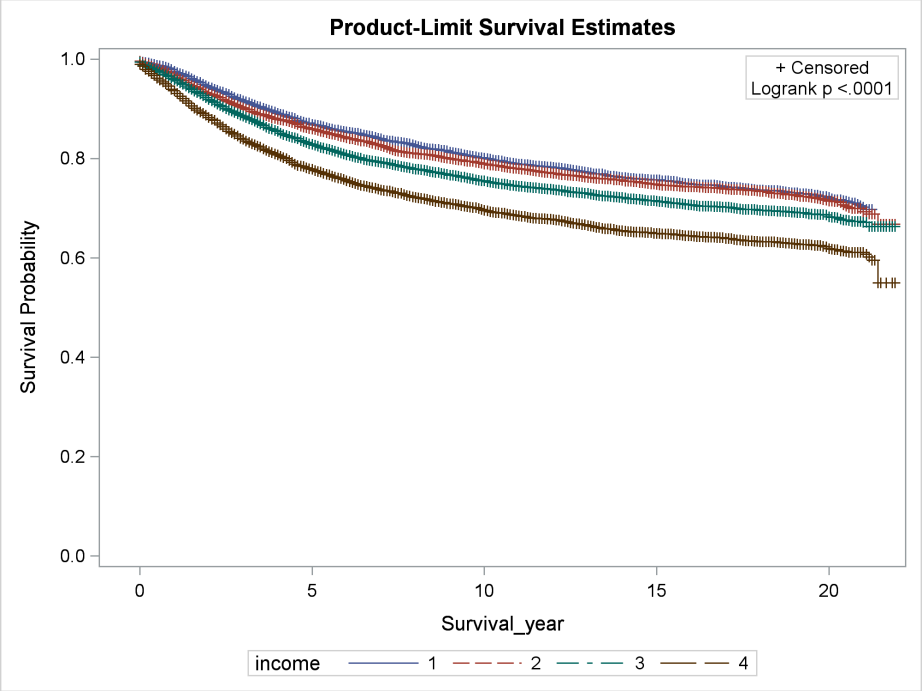

Figure 3.6: Comparing survival by neighborhood poverty line

$$H_a : \beta_k \neq 0$$

Where $\epsilon_i \sim N(0, 1)$, and $k = 0, \cdots K$ is the number of explanatory variables and $s_i(t)$ represent individual survival time

Table 2: Survival Time Regression

From 3.2, based on the Estimate no interaction terms, the most relevant results indicate that Blacks have the lowest survival length ($-0.37^{\star\star}$) compared to White, and others race are living longer ($0.57^{\star\star}$) compared to White. Compared to survivors diagnostic with grade I, survivors with grade II, III, and IV are living less long. Compared to the Stage 1 survivors Stage II, III, IV and un-stage survivors are dying earlier. It is imperative to diagnostic patients with breast cancer in early stage and grade if we aim to increase substantially the length of survival. Compared to survivors living in neighborhood poverty line between $20\% - 100\%$, survivors living in more influent neighborhood live much longer. Moreover, when I Interact

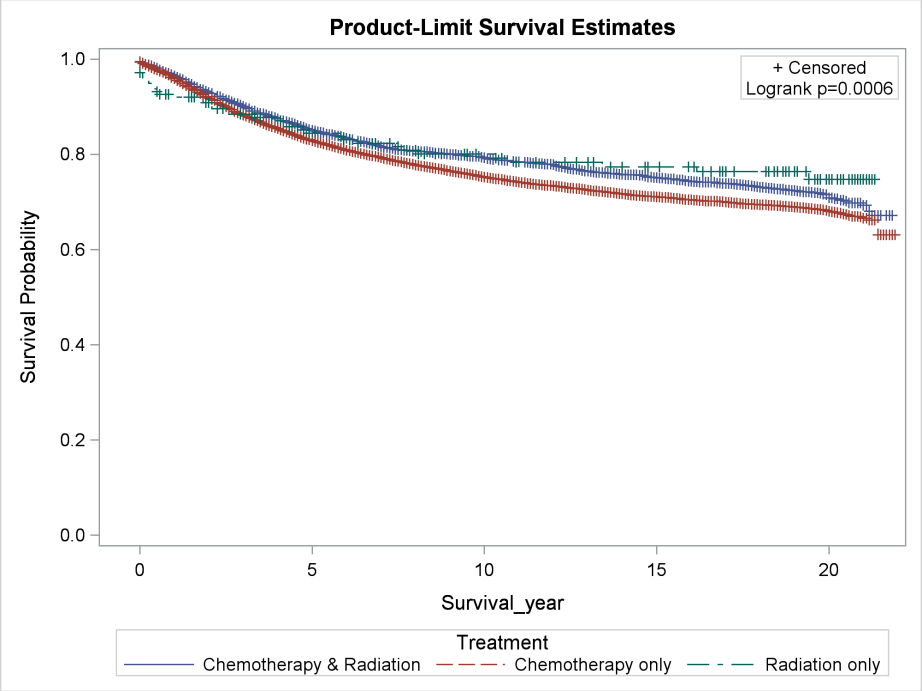

Figure 3.7: Comparing survival by treatment received

stage and age, age and race, and stage and race I find some significant results. For instance, being black and being between 40 to 49 year old is not a good prognostic for a breast cancer survivor in Texas.

3.4.2 Model 2: Cox proportional hazards regression

$$h_i(t) = (\boldsymbol{\delta}_1 Gender_i + \boldsymbol{\delta}_2 Race_i + \boldsymbol{\delta}_3 Age_i + \boldsymbol{\delta}_4 Treatment_i + \boldsymbol{\delta}_5 Grade_i + \boldsymbol{\delta}_6 Stage_i \quad (3.2)$$

$$+\boldsymbol{\delta}_7 IncomeProxy_i + \boldsymbol{\delta}_8 GeographicAreas_i + \boldsymbol{\epsilon}_i) h_0(t)$$

,

$$H_0 : \delta_j = 0$$

Table 3.2: Survival Time Regression

Variables	Estimate without interaction terms	Estimate with Interaction terms
	8.79**	8.8**
female	0.26	-0.3
Hispanic	-0.04	-0.15
black	-0.37**	-0.33**
other	0.57**	0.2
Between 40 and 49	0.17	0.26
Between 50 and 64	0.07	0.22
between 65 and 74	-0.03	0.17
over 75	-0.43	-0.2
Chemo and Radiation	-0.33	-0.33
Chemo only	-0.48**	-0.5**
Grade II	-0.72**	-0.72**
Grade III	-1.23**	-1.23**
Grade IV	-1.35**	-1.34**
Unknown	-0.97**	-0.96**
stage 2	-1.28**	-1.3**
stage 3	-2.14**	-2.1**
stage 4	-3.19**	-2.78**
Un-stage 5	-1.45**	-1.2**
$0\%- < 5\%$	0.31**	0.32**
$5\%- < 10\%$	0.3**	0.31**
$10\%- < 20\%$	0.2**	0.21**
Metropolitan area	0.21	0.21
Urban area	0.2	0.21
Between 40 and 49 * Black		-0.42**
un-stage*Hispanic		0.39**
un-stage*Black		0.37**
Between 40 and 49 * Stage4		-0.47**
Between 50 and 59 * Stage4		-0.53**
Between 65 and 74 * Stage4		-0.74**
over 75 * Stage4		-0.41**
over 75 * Stage5		-0.58**

Notes:⋆⋆ means 5% level of statistical significance (p-values are less than 5%), Chemo= chemotherapy

$$H_a : \delta_j \neq 0$$

Where the subscript i on an explanatory variable denotes the value of that variable for the i th individual, and $j = 1, \cdots, J$ where J is the number of explanatory variables.

Table 3.3: Cox proportional hazards regression model estimate

	Estimate	Hazard Ratio
female	0.24	1.27
Hispanic	0.03	1.04
black	0.29**	1.34
other	-0.45**	0.63
Between 40 and 49	-0.13**	0.88
Between 50 and 64	-0.06	0.943
between 65 and 74	-0.01	0.99
over 75	0.28**	1.32
Chemo and Radiation	0.14	1.15
Chemo only	0.26	1.3
Grade II	0.58**	1.78
Grade III	0.99**	2.71
Grade IV	1.1**	3.02
Unknown grade	0.8**	2.21
stage 2	1.04**	2.82
stage 3	1.72**	5.6
stage 4	2.64**	14.02
stage 5	1.21**	3.37
$0\%- < 5\%$	-0.25**	0.78
$5\%- < 10\%$	-0.24**	0.78
$10\%- < 20\%$	-0.16**	0.85
Metropolitan area	-0.19	0.83
Urban area	-0.17	0.85

Notes: $\star\star$ means 5% level of statistical significance (p-values are less than 5%), Chemo= chemotherapy

From table 3.3, I see that compared to whites, blacks present higher hazard ratio ($1.34^{\star\star}$) which means that blacks has 1.34 times risk of dying compared to whites. Furthermore, compared to survivors age between 0 and 39, survivors 75 and over present higher hazard ratio ($1.32^{\star\star}$). Compared to survivors diagnostics with grade I, survivors with more advanced grade are more likely to died early; For instance grade IV has a hazard ratio ($3.02^{\star\star}$). Moreover, compared to survivors diagnostics

with stage 1 breast cancer, survivors with more advanced stage are more likely to die early; for example, survivors with stage IV have a hazard ratio (14.02**) that means that at any particular time, survivors diagnosed with stage IV have 14.02 time vulnerability of dying compared to those diagnosed with Stage I. We then need to promote early diagnostics to improve survival length. Compared to survivors living in a less influent neighborhood, survivors in a more affluent one live much longer. In fact, survivors living in a $0-5\%$ poverty line have a hazard ratio (0.78**). Therefore, based on those regressions, the most significant factor is stage at diagnostics, then grade, follow by income, then age and finally race. Those findings suggest that stage is a factor of paramount importance. I then would like to better understand characteristics of breast cancer survivors in Texas by stage. I then conducted a Cox proportional hazards regression on survivors that were diagnosed with stage 1 only.

Table 3.4 describes the characteristics of Breast cancer survivors in Texas by stage, it indicates that White non-Hispanics and other are diagnose with early stage cancer than Black-non Hispanics and Hispanics. Therefore, policy makers need to promote early screening in Black non-Hispanics and Hispanics communities. This data description also shows that younger survivors are being diagnosed with slightly later stage cancer. Finally, survivors from affluent neighborhoods are being diagnosed with early stage compared to others. I will then run a Cox proportional hazards regression model estimate among survivors diagnostic with stage 1.

Table 5: Cox proportional hazards regression model estimate among survivors diagnostic with stage 1

Table 3.5 portrays that among survivors diagnostic with stage 1, black non-Hispanics have a high hazard of dying compared to white non Hispanics of 1.45. Compared to grade 1, survivors diagnose with more advanced grade are more likely to die early, as an example the hazard ratio for survivors with stage 1 and grade IV

Table 3.4: Data descriptive and characteristics of Breast cancer survivors in Texas by stage

	n	Stage 1 $n = 11,307$	Stage 2 $n = 5440$	Stage 3 $n = 803$	Stage 4 $n = 1312$	Un-Stage $n = 2245$
Race						
White non-Hispanic	15573	78.33%	69.21%	66.75%	63.54%	71.27%
Black non-Hispanic	2134	8.32%	11.5%	13.2%	16.25%	11.17%
Hispanic	3150	12.36%	18.2%	18.43%	19.45%	16.18%
Other	225	0.99%	1.09%	1.62%	0.76%	1.39%
Age						
0 and 39	1626	6.34%	10%	9.96%	8.92%	7.48%
40 and 49	3938	17%	22.54%	20.55%	17.84%	17.42%
50 and 64	6507	30.94%	31.32%	33.25%	32.93%	26.99%
65 and 74	4859	25.04%	20.46%	17.93%	22%	22%
Over 75	4177	20.69%	15.68%	18.31%	18.6%	26.46%
Treatment						
Chemo and Radiation	3136	16.41%	16.51%	17.34%	15.58%	12.25%
Chemo only	16231	82.76%	82.65%	81.5%	83.69%	86.05%
Radiation only	179	0.83%	0.83%	1.16%	0.72%	1.7%
Grade						
Grade I	2545	17.06%	7.33%	3.99%	3.2%	6.37%
Grade II	6495	34.34%	32.83%	28.27%	20.12%	14.92%
Grade III	62271	25.27%	41.19%	50.56%	37.04%	12.52%
Grade IV	420	1.63%	2.74%	2.74%	2.52%	1.43%
Unknown grade	5376	21.7%	15.9%	14.45%	37.12%	64.77%
Poverty level						
$0 - 5\%$	3832	19.4%	17.77%	15.25%	13.72%	16.54%
$5\% - 10\%$	4811	24.11%	21.89%	22.38%	20.73%	19.84%
$10\% - 20\%$	7172	33.79%	33.93%	35.5%	33.31%	35.09%
$20\% - 100\%$	5279	22.7%	26.41%	26.88%	32.24%	28.53%

is at 3.56. The best prognostic will occur when we can diagnose breast cancer patients at stage 1 and grade 1. Being diagnosed with stage 1 and living in an affluent neighborhood will lead to better prognostic.

Table 3.5: Cox proportional hazards regression model estimate among survivors diagnostic with stage 1

	Estimate	Hazard Ratio
female	0.98	2.67
Hispanic	0.15	1.17
black	0.38**	1.45
other	-0.43	0.65
Between 40 and 49	-0.14	0.87
Between 50 and 64	-0.14	0.87
between 65 and 74	-0.137	0.87
over 75	0.19	1.2
Chemo and Radiation	0.29	1.33
Chemo only	0.41	1.5
Grade II	0.73**	2.06
Grade III	1.17**	3.2
Grade IV	1.27**	3.56
Unknown grade	0.75**	2.12
$0\%- < 5\%$	-0.26**	0.77
$5\%- < 10\%$	-0.28**	0.75
$10\%- < 20\%$	-0.08	0.92
Metropolitan area	-0.39	0.68
Urban area	-0.34	0.72

3.4.3 Multinomial Logistic Regression:

My next step is to investigate when those disparities occur; is it in the first 5 years of survival or is it later. Therefore, I conducted a multinomial logistic regression; I create 4 groups of breast cancer survivors, the first group only comprises of survivors who died between 0 and 5 years after diagnostic of breast cancer, the second group is made of survivors who died between 5 years and 14 years, the third group is only constituted of survivors who died after 14 years, and finally the fourth group is formed by only survivors that are still alive after the end of the study in 2015. The incomplete observed responses are censored and classified in group 5. For the rest of the study, I focus on the 4 first groups only. I am

starting this analysis by showing the data descriptive and characteristics of breast cancer survivors in Texas by groups. Table 6: Data descriptive and characteristics of Breast cancer survivors in Texas by groups

From Table 3.6 I observe that:

In terms of race, Black non-Hispanic have lower survival rate compared to White non-Hispanics, Hispanics and others. In fact, 37% of Black non-Hispanics died within 5 years of diagnostic (group 1) compared to 31.36% of Hispanics, 27% of White non-Hispanics, and only 13% of others in the same period. Furthermore, when I consider group 4, which is for survivors who are still alive after 2015, I see that only 25.3% of Black non-Hispanic are in that group compared to 30% of White non-Hispanics, 32.7% of Hispanics, and 54.35% of others. Therefore, there is a racial disparity in term of survival length of breast cancer in Texas.

In terms of age, breast cancer patients that are 75 and older have a lower survival time compared to survivors that are younger. For instance, 46.07% of survivors that are 75 and over did not live beyond 5 years of diagnosis compared to 26.12% for survivors who are less than 40 years old. When I consider group 4, only 1.5% of people 75 and over survives compared to 55.27% of survivors that are between 40 and 49 years old, and 52.11% of people between 0 and 39 year old did survive. This can be explained by the fact that survivors over 65 have existing comorbidity compared to the younger survivors and the life expectancy of women in the USA is at 81.3 years.

In terms of grade and stage, I can notice that as the grade or stage progress, the survival time reduces. Based on the stage, only 17.13% of all diagnostics with stage I died within 5 years compared to 78.39% of all diagnostics with stage IV. However, those 17.13% represented $1,776$ survivors and the 78.39% represent $1,012$ survivors. Therefore, promoting early diagnostics through better screening strategy is crucial yet not enough, and we are going to investigate why some survivors diagnosed with

stage I live only 5 years and others live beyond 20 years. Additionally, 37.1% of patients who were diagnosed with stage 1are still alive after 2015 compared to only 4.73% patients who were diagnosed with stage IV. Moreover, based on grade, 16.05% of patients diagnosed with grade I died within 5 years compared to 41.8% of patients diagnosed with grade IV.

In terms of poverty level of where the survivor lives, I can observe that there is a positive relation between the lower poverty level and the survival length. Indeed, 22.08% of survivors that live in neighborhood where only 5% of the residents live under the poverty level died within 5 years of diagnosis compared to 35.52% of people who live in the neighborhood that have more than 20% of their residents living under the poverty line. This may be explained by the income of the survivor, and in health economics we know income and access to quality of care are positively related. Furthermore, this can be explained by the fact that in affluent neighborhoods, there are better hospitals and amenities like gyms, organic food stores, and more recreational activities that all contribute to length of survival. In terms of where the survivor lives, there is some slight disparity since patients living in metropolitan areas who are living much longer than patients living in rural areas, which can also be explained by availability of services and hospitals that specialize on cancer treatments.

Multinomial logistic regression analysis

I model the multinomial logistic regression as follows:

Let Y be a categorical response with 4 categories , that is Y takes 4 values, 1 are those in group 1, 2 are those in group 2, 3 is for group 3 and 4 are those in group 4;

Let $\pi_j(x) = P(Y = j|x)$ at a fixed setting x for explanatory variables, with $\sum_{j=1}^{4} \pi_i = 1$ Logit models pair each response category with a baseline category, here

I consider group 4 as the baseline category. So, the model is as follows:

$$log\frac{\pi_j(x)}{\pi_4(x)} = \alpha_j + \beta_j^{'}(x)$$

, $j = 1, 2, 3$ With

$$\beta_j^{'} = (\beta_{1j}, \beta_{2j} \cdots \beta_{kj})$$

and $x^{'} = (race, gender, age, treatment, grade, stage, poverty, city)$

Table 7: Multinomial logistic Regression on Breast cancer Survival time

Table 3.7 describes the multinomial logistic Regression analysis with group 4 as the reference category. The most relevant results indicate that for all other covariates fixed for instance, for stage 3 estimated odds that those whose length of survival is less than 5 years (group 1) instead of those whose length of survival is still alive (group 4) are exp (2.34) = 10.42 times the estimated for odds for stage 1, and this is statistically significant with a p-value < 0.0001; the Wald 95% confidence interval is $(8.28, 13.11)$.

Furthermore, stage 3 estimated odds for those in group 2 are $\exp 1.25 = 3.47$ times the estimated for odds for stage 1 and it is statistically significant. The odd estimate decreased in group 3 to $\exp 0.48 = 1.56$. For stage 4, the estimated odds become $\exp 3.96 = 52.36$ times the estimated odds for stage 1 with a p-value < 0.001, the Wald 95% confidence interval is $(39.75, 68.97)$, and the odd decreases to 5.88 in group 2 and to 2.02 in group 3. So, the conclusion is that early diagnosis is key for longer survival, and the disparities mostly occur in the first 5 years of survival.

For Black non-Hispanics estimated odds that those whose length of survival is less than 5 years (group 1) instead of those whose length of survival is still alive (group 4) are $\exp 0.53 = 1.62$ times the estimated for odds for White non-Hispanics, and this is statistically significant with a p-value < 0.0001; the Wald 95% confidence

interval is $(1.39, 1.89)$. Furthermore, Black non-Hispanic estimated odds for those in group 2 are $\exp 0.34 = 1.38$ times the estimated for odds for White non-Hispanics and it is statistically significant. The odd estimate decreased in group 3 to $\exp 0.15 = 1.2$. Therefore, the most racial gap occurs in the first five years of survival this is very worrisome and implies that the productivity loss due to breast cancer mortality for black non-Hispanics is too enormous.

For survivors 75 and over estimated odds that those whose length of survival is less than 5 years (group 1) instead of those whose length of survival is still alive (group 4) are $\exp 4.86 = 128.74$ times the estimated for odds for survivors less than 39, and this is statistically significant with a p-value < 0.0001; the Wald 95% confidence interval is $(94.62, 175.17)$. Furthermore survivors 75 and over estimated odds for those in group 2 are $\exp 4.75 = 115.4$ times the estimated for odds for survivors less than 39 and it is statistically significant. The odd estimate decreased in group 3 to $\exp(4.42) = 83.34$. Even for survivors between the age of 65 and 74, the odd is still high. We need to assist the old survivors to improve their survival time. Finally, living in affluent neighborhood will lead to better prognostic.

3.5 Conclusion

This study allows us to investigate the main factors that can explain the survival length of breast cancer patients in Texas from a 20-year retrospective cohort data. My findings suggest that stage at diagnostic is the most important drivers of breast cancer survival. In fact, Compared to stage I survivors, survivors with stage IV are more likely to die with a hazard ratio of (14.02). I also find that being diagnosed with advanced grade will lead to short survival time. For instance, compared to survival diagnosed with grade 1, grade IV patients have a hazard ratio of (3.02). Furthermore, I find that there are some racial disparities in survival time. In fact,

compared to white non-Hispanics, black non-Hispanics have the shorter length of survivors with a hazard ratio of (1.34), and being black non-Hispanics and being between 40 to 49 year old is not a good prognostic for a breast cancer survivor in Texas. I observe that being compared to survivors between the age of 0 and 39, survivors 75 and older had a higher hazard of dying with a hazard ratio of 1.32. Living in affluent neighborhood areas improve survival time. Finally, I find that most of the disparities in terms of stage, grade, age, race and income occur in the first five year of survival. Policy makers need to promote early screening and diagnostic in vulnerable communities such as black non-Hispanics and Hispanics. They also need to better assist survival over 75 year old and black non-Hispanics patients between the age of 40 and 49, and the poor to improve survival time, since survival length has huge socio-economic implications. Yet, this paper has some limitations, as I believe this study would have been stronger if we had access to more survivorsŠ characteristics, such as education level, and accurate household income.

Table 3.6: Data descriptive and characteristics of Breast cancer survivors in Texas by groups

		Group 1	Group 2	Group 3	Group 4
Race	n	5631	5353	2817	5932
White non-Hispanic	14616	27%	28%	15%	30%
Black non-Hispanic	2011	37%	25.8%	11.44%	25.3%
Hispanic	2911	31.36%	23.32%	12.6%	32.7%
Other	195	13%	19%	13.33%	54.35%
Age					
Between 0 and 39	1493	26.12%	16.28%	5.49%	52.11%
Between 40 and 49	3550	21.07%	15.63%	8.03%	55.27%
Between 50 and 64	5942	22.53%	21.66%	14.56%	41.25%
between 65 and 74	4645	27.02%	33.89%	24.22%	14.88%
75 and over	4124	46.07%	41.2%	11.23%	1.5%
Gender					
Female	19639	28.49%	27.09%	14.27%	30.14%
Male	115	30.43%	33.04%	14.78%	21.74%
Treatment					
Chemo and Radiation	2945	25.94%	25.16%	16.13%	32.6%
Chemo only	15157	28.38%	27.25%	13.93%	30.43%
Radiation only	169	31.36%	26.03%	15.97%	26.63%
Grade					
Grade I	2362	16.05%	31%	19.94%	33.02%
Grade II	6035	23.18%	29.5%	15.66%	31.66%
Grade III	5897	34.07%	22.88%	11.48%	31.57%
Grade IV	390	41.8%	20.51%	8.97%	28.71%
Unknown grade	5070	33.17%	27.95%	13.65%	25.23%
Stage					
Stage I	10452	17.13%	28.37%	17.4%	37.1%
Stage II	5091	31.07%	28.05%	13.08%	27.79%
Stage III	774	51.81%	24.93%	7.36%	15.89%
Stage IV	1291	78.39%	13.79%	3.1%	4.73%
Unknown or un-staged	2146	39.42%	27.68%	11.09%	21.8%
Poverty level					
$0-5\%$	3546	22.08%	23.32%	13.85%	40.75%
$5\%-10\%$	4484	23.8%	27.27%	15.14%	33.78%
$10\%-20\%$	6738	29.84%	27.8%	14.94%	27.41%
$20\%-100\%$	4975	35.52%	28.78%	12.88%	22.81%
Areas where survivors lived					
Metropolitan areas	16231	27.6%	26.61%	14.2%	31.58%
Urban areas	3205	32.54%	29.8%	14.41%	23.24%
Rural areas	317	34.38%	26.18%	16.4%	23.03%

Notes: s_t is for survivor time

Table 3.7: Multinomial logistic Regression on Breast cancer Survival time

	Group 1 n=5114 -1.99**	Odds Ratio	Group 2 n=4909 -1.7**	Odds Ratio	Group 3 n=2614 -2.43**	Odds Ratio
female	-0.08	0.85	-0.16	0.73	-0.05	0.9
Hispanic	0.07	1.03	-0.07	0.91	-0.06	0.99
black	0.53**	1.62	0.35**	1.38	0.14	1.2
other	-0.65**	0.5	-0.31	0.72	-0.04	1
Between 40 and 49	-0.18	0.83	-0.04	0.96	0.38**	1.46
Between 50 and 64	0.36**	1.43	0.64**	1.9	1.32**	3.74
between 65 and 74	1.85**	6.34	2.2**	9.1	2.9**	18.22
over 75	4.86**	128.74	4.75**	115.4	4.42**	83.34
Chemo and Radiation	-0.02	1.14	0.029	1.32	0.1	1.28
Chemo only	0.17**	1.38	0.22**	1.6	0.05	1.22
Grade II	-0.3**	1.62	-0.04	1.2	-0.06	0.99
Grade III	0.31**	2.99	0.06	1.32	0	1.06
Grade IV	0.71**	4.44	0.1	1.38	-0.02	1.03
Unknown	0.06	2.31	0.11**	1.4	0.12	1.2
stage 2	1.24**	3.44	0.73**	2.07	0.4**	1.5
stage 3	2.34**	10.42	1.25**	3.47	0.48**	1.56
stage 4	3.96**	52.36	1.78**	5.88	0.7**	2.02
Un-stage	1.37**	3.96	0.61**	1.84	0.25**	1.3
$0\%- < 5\%$	-0.27**	0.54	0.26**	0.59	-0.2**	0.73
$5\%- < 10\%$	-0.15**	0.61	-0.02	0.76	0	0.9
$10\%- < 20\%$	0.07	0.76	0.02	0.8	0.07	0.96
Metropolitan area	-0.09	0.79	-0.03	1	-0.06	0.86
Urban area	-0.05	0.82	0.06	1.1	-0.02	0.9

Notes: $\star\star$ means 5% level of statistical significance (p-values are less than 5%), Chemo= chemotherapy

Bibliography

[1] **Autor, Alyce S. Adams, PHD, Fang Zhang, PHD, Connie Mah, MS, Richard W. Grant, MD, MPH, Ken Kleinman, SCD, James B. Meigs, MD, MPH and Dennis Ross-Degnan, SCD,** 2005. "Race Differences in Long-Term Diabetes Management in an HMO"Diabetes care;31(5); 916 – 921

[2] **Autor, Andrew J. Karter, PhD; Assiamira Ferrara, MD, PhD; Jennifer Y. Liu, MPH; Howard H. Moffet. MPH, Lynn M Ackerson. PHD and Joe.V. Selby. MD, MPH** (2002)"Ethnic Disparities in Diabetic Complications in an Insured Population."JAMA

[3] **Autor, Anne Marie McCarthy, ScM, PhD, Jianing Yang, MS, and Katrina Armstrong, MD** (2015)"Increasing Disparities in Breast Cancer Mortality From 1979 to 2010 for US Black Women Aged 20 to 49 Years"American Journal of Public Health,105: S446 – $S448$

[4] **Anne M. Weaver, Susan E. McCann, Jing Nie, Stephen B. Edge, Thomas H. Nochajski, Marcia Russell, Maurizio Trevisan, Jo L. Freudenheim** (2013)"Alcohol intake over the life course and breast cancer survival in Western New York exposures and breast cancer (WEB) study: quantity and intensity of intake"Breast Cancer Res Treat 139:245Ú253

[5] **Autor, Anoop Dinesh Shah, Dr., MRCP, Claudia Langenberg, PhD, Eleni Rapsomaniki, PhD, Spiros Denaxas, PhD,Mar Pujades-Rodriguez, PhD, Chris P Gale, FRCP,John Deanfield, Prof, FRCP, Liam Smeeth, Prof, FRCGP,Adam Timmis, Prof, FRCP, and Harry Hemingway, Prof, FRCP,** 2015. "Type 2 diabetes and incidence of cardiovascular diseases: a cohort study in 1.9 million people"Lancet Diabetes Endocrinol

[6] **Autor, Arnaud Chiolero, David Faeh, Fred Paccaud and Jacques Cornuz,** 2011. "Consequences of smoking for body weight, body fat distribution, and insulin resistance." American Journalof Clinical Nutrition; 87(4):801 – 9

[7] **Autor, Attvall S, Fowelin J, Lager I, Von Schenck H, Smith U** 1993. "Smoking induces insulin resistance–a potential link with the insulin resistance syndrome" J Intern Med; 233(4);327 – 32

[8] **Avonne E. Connor1, Richard N. Baumgartner, Christina M. Pinkston2, Stephanie D. Boone, Kathy B. Baumgartner** 2016. "Obesity, ethnicity, and quality of life among breast cancer survivors and women without breast cancer: the long-term quality of life follow-up study" Cancer Causes Control,27;115 – 124

[9] **Bernd Holleczek a, Silvia Rossi b, Agius Domenic c, KaireInnosd, Pamela Minicozzi e, Silvia Francisci b, Monika Hacklf, Nora Eisemanng, Hermann Brennerh,i,j, and the EUROCARE-5 Working Group1** 2015. "On-going improvement and persistent di?erences in the survival for patients with colon and rectum cancer across Europe 1999Ű2007, Results from the EUROCARE-5 study" European Journal of Cancer 51 ;2158 – 2168.

[10] **Brown AF, Gregg EW, Stevens MR, Karter AJ, Weinberger M, Safford MM, Gary TL, Caputo DA, Waitzfelder B, Kim C, Beckles GL.** 2005. "Race, ethnicity, socioeconomic position, and quality of care for adults with diabetes enrolled in managed care: the Translating Research Into Action for Diabetes (TRIAD) study" Diabetes Care; 28(12);2864 – 70

[11] **Candyce H. Kroenke, Bernard Rosner, Wendy Y. Chen, Ichiro Kawachi, Graham A. Colditz, and Michelle D. Holmes** 2004. "Functional Impact of Breast Cancer by Age at Diagnosis" JOURNAL OF CLINICAL ONCOLOGY, VOLUME 22, No 10,

[12] **Catherine Kim, MD, MPH, Katherine M. Newton, PHD and Robert H. Knopp,MD** 2002. " Gestational Diabetes and the Incidence of Type 2 Diabetes, a systematic review" American Diabetes Association.; 25(10): 1862 – 1868

[13] **Cathy J. Bradley K. Robin Yabroff Bassam Dahman Eric J. Feuer Angela Mariotto Martin L. Brown** 2008"Productivity Costs of Cancer Mortality in the United States: 2000 – 2020" Journal of the National Cancer Institute, Volume 100, Issue 24

[14] **CE Phelps** 1992 "Life in the fast lanes" Harper Collins Publishers, New York, health Economics

[15] **Chiagozie Nwasuruba, Christie Osuagwu, Sejong Bae, Karan P. Singh, Leonard E. Egede** 2015. "ŠRacial differences in diabetes self-management and quality of care in Texas" Journal of Diabetes and its Complications; Volume 23, Issue 2, Pages 112 – 118

[16] **Chou, Shin-Yi, Michael Grossman, and Henry Saffer** 2004. "An Economic Analysis of Adult Obesity: Results from the Behavioral Risk Factor Surveillance System. " Journal of Health Economics, 23(3): 565 – 587

[17] **Derek S. Brown, PhD, Justin G. Trogdon, PhD,Donatus U. Ekwueme, PhD,LindaChamiec-Case, BA, , Gery P. Guy Jr., PhD,**

Florence K. Tangka, PhD, Chunyu Li, MD, PhD, Katrina F. Trivers, PhD,Juan L. Rodriguez, MPH 2016"Health State Utility Impact of Breast Cancer in U.S. Women Aged 18 − 44 YearsŠ. American Journal of Preventive Medicine" American Journal of Preventive Medicine

[18] **DeSantis CE, Lin CC, Mariotto AB, Siegel RL, Stein KD, Kramer JL, Alteri R, Robbins AS, Jemal A** 2014. "cancer treatment and survivorship statistics " CA Cancer J Clin 64(4):252 − 271. doi:10.3322/caac.21235

[19] **Donatus U. Ekwueme, PhD ,Justin G. Trogdon, PhD, Olga A. Khavjou, MA, Gery P. Guy Jr., PhD** 2016. "Productivity Costs Associated with Breast Cancer Among Survivors Aged 18 − 44 Years "American Journal of Preventive Medicine

[20] **Egede LE, and Zheng D,** 2003. "Racial/ethnic differences in adult vaccination among individuals with diabetes." American Journal Public health, 93: 324 − 329

[21] **Fiona McKenziea*, Lis Ellison-Loschmanna and Mona Jeffreysb** 2011. "Investigating reasons for ethnic inequalities in breast cancer survival in New Zealand" Ethnicity and Health, Vol. 16, No. 6

[22] **Francis P Boscoe, PhD, Christopher J Johnson, MPH,Recinda L Sherman, PhD, David G Stinchcomb, MS, Ge Lin, PhD,and Kevin A Henry, PhD,** 2010."The Relationship Between Area Poverty Rate and Site-Specific Cancer Incidence in the United States" Journal of Registry Management; 37(4): 148 − 151

[23] **Frank A. Sloan, Norma A. Padroťn, and Alyssa C. Platt** 2009. "Preferences, Beliefs, and Self-Management of Diabetes" Health Services Research

[24] **Frank B. Hu, MD, PHD and Vasanti S. Malik, MSc ,.** 2010. "Sugar-sweetened beverages and risk of obesity and type 2 diabetes: Epidemiologic evidence"physiol behave;100(1): 47 − 54

[25] **Grossman, M.** 1972. "On the Concept of Health Capital and the Demand for Health." Journal of Political Economy 80 (2): 223 − 55

[26] **Gull Z, Anwar Z, Sherazi BA** 2017 "Study of Treatment and Risk Factors Leading to Breast Cancer" Isra Med J.; 9(4): 226 − 28

[27] **Hans-Oloy Adami,MD, Brigitta Malker, BSc, Lars Holmberg, MD, Ingemar Persson, MD., and Betty Stone, PHD** 1986 "The Relation between Survival and Age at diagnosis in Breast Cancer" N Engl J Med; 315:559 − 563

[28] **Kelsey, JL, Horn-Ross, Pl.** 1993 "Breast cancer: magnitude of the problem and descriptive epidemiology" Epidemiol Rev; 15:7 − 16

[29] **J. William Eley, MD, MPH; Holly A. Hill, MD, PhD; Vivien W. Chen, PhD; et al** 1994 "Racial Differences in Survival From Breast Cancer Results of the National Cancer Institute Black/White Cancer Survival Study" JAMA, 272(12):947 − 954

[30] **Lorig, Kate R. RN, DrPH*; Sobel, David S. MD, MPH; Stewart, Anita L. PhD; Brown, Byron William Jr. PhD; Bandura, Albert PhD; Ritter, Philip PhD; Gonzalez, Virginia M. MPH; Laurent, Diana D. MPH; Holman, Halsted R. MD** 1999. "Evidence Suggesting That a Chronic Disease Self-Management Program Can Improve Health Status While Reducing Hospitalization: A Randomized Trial." Medical care, Vol. 37, No. 1 (Jan., 1999), pp. 5 − 14

[31] **Kevin Griffith, Leigh Evans and Jacob Bor** 2017. "The Affordable Care Act reduced socioeconomic disparities in health care Access" Health Affairs

[32] **Maliniak ML; Patel AV; McCullough ML; Campbell PT; Leach CR; Gapstur SM; Gaudet MM** 2017. "Obesity, physical activity, and breast cancer survival among older breast cancer survivors in the Cancer Prevention Study-II Nutrition Cohort" Breast Cancer Research and Treatment

[33] **Marcus Lind, M.D., Ph.D., Ann-Marie Svensson, Ph.D., Mikhail Kosiborod, M.D., Soffia Gudbjörnsdottir, M.D., Ph.D., Aldina Pivodic, M.Sc., Hans Wedel, Ph.D., Sofia Dahlqvist, Mark Clements, M.D., Ph.D., and Annika Rosengren, M.D., Ph.D** 2014 "Glycemic Control and Excess Mortality in Type 1 Diabetes." New England Journal of Medicine

[34] **Marilyn L. Kwan, Lawrence H. Kushi, Erin Weltzien, Emily K. Tam, Adrienne Castillo, Carol Sweeney, and Bette J. Caan** 2010. "Alcohol Consumption and Breast Cancer Recurrence and Survival Among Women With Early-Stage Breast Cancer: The Life After Cancer Epidemiology Study" J Clin Oncol,

[35] **Marsha Lillie-Blanton and Catherine Hoffman** 2005. "The Role of Health Insurance Coverage in Reducing Racial/Ethnic Disparities in Health Care" Health affairs

[36] **Mayberry. R. M, Mili, F, and Ofili, E.** 2000. "Racial and ethnic differences in access to medical care " Medical Care Research and Review; 57 Suppl 1:108 − 45

[37] **Melissa L. Santorelli, Kim M. Hirshfield, Michael B. Steinberg, Yong Lin, George G. Rhoads, Elisa V. Bandera, Kitaw Demissie** 2017. "Racial differences in the effects of comorbidity on breast cancer-specific survival " Cancer Causes Control 28:809 − 817

[38] **National Diabetes Statistics Report** 2017. "estimates of Diabetes and its burden in the United States"

[39] **Neal A. Chatterjee, MD, Yulei He, PhD, and Nancy L. Keating, MD, MPH** 2013. "Racial Differences in Breast Cancer Stage at Diagnosis in the Mammography Era " American Journal of Public Health, Vol 103, No. 1

[40] **Raeven Faye chandler and Shannon M. Monnat** 2015. "Racial/Ethnic Differences in Use of Health Care Services for Diabetes Management" Health Education and Behavior. 42(6) 783 – 792.

[41] **Robert S. Kirsner, M.D., Ph.D., Fangchao Ma, M.D., Ph.D., Lora E. Fleming, M.D., Ph.D. Daniel G. Federman, M.D. Edward Trapido, Sc.D. Robert Duncan, Ph.D. Panta Rouhani, Ph.D., M.P.H. and James D. Wilkinson, M.D., M.P.H.** 2010. "Earlier Stage at Diagnosis and Improved Survival Among Medicare HMO Patients with Breast Cancer" Journal womenŠs health vol 19 No 9

[42] **Richard Zeckhauser and Donald Shepard** 1976. "Where Now for Saving Lives? " Duke University School of Law

[43] **Shi Q, Fonseca V, Krousel-Wood M, Zhao Y, Nellans FP, Luo Q, Shi L.** 2016. "Will the Affordable Care Act (ACA) Improve Racial/Ethnic Disparity of Eye Examination Among US Working-Age Population with Diabetes?"Diabetes care,37(5):1321 – 8.

[44] **Weinstein MC and Stason WB** 1997. "Foundations of cost-effectiveness analysis for health and medical practices" New England Journal of Medicine, 296(13):716 – 21.

[45] **Wang G, Zhang z, Ayala C, Dunet Do, Fang J, George MG** 2014. "Costs of hospitalization for stroke patients aged 18-64 years in the United States."May-Jun;23(5):861 – 8. doi

[46] **Yoichiro Kakugawa, Masaaki Kawai, Yoshikazu Nishino, Kayoko Fukamachi, Takanori Ishida, Noriaki Ohuchi and Yuko Minami** 2015. "Smoking and survival after breast cancer diagnosis in Japanese women: A prospective cohort study"Cancer Science 106 (2015) 1066 – 1074

[47] **Yun-Jiu Gou, Ding-XiongXie, Ke-Hu Yang, Ya-Li Liu, Jian-Hua Zhang, Bin Li, Xiao-Dong He** 2013. "Alcohol Consumption and Breast Cancer Survival: A Metaanalysis of Cohort Studies."Asian Pacific Journal of Cancer Prevention, Vol 14,

[48] $https://www.niddk.nih.gov/health-information/diabetes/overview/what-is-diabetes$

[49] $https://www.niddk.nih.gov/health-information/diabetes/overview/tests-diagnosis/a1c-test\#1$

[50] $https://www.cdc.gov/nchs/fastats/leading-causes-of-death.htm$

[51] $http://care.diabetesjournals.org/content/diacare/suppl/2017/12/08/41.Supplement_1.DC1/DC_41_S1_Combined.pdf$

[52] $http://www.diabetes.org/advocacy/news-events/cost-of-diabetes.html$

[53] $https://www.ssa.gov/policy/docs/research-summaries/education-earnings-alt.html$

[54] $https://www.dshs.state.tx.us/chs/popdat/downloads.sht$