2023

Effect of Treatment, Stage of Lung Cancer, and Socioeconomic Status on Life Expectancy Within Marginalized Communities

Elsadig Ahmed Elsharif
Walden University

Walden University

College of Health Sciences and Public Policy

This is to certify that the doctoral dissertation by

Elsadig Ahmed Elsharif

has been found to be complete and satisfactory in all respects,
and that any and all revisions required by
the review committee have been made.

Review Committee
Dr. Zin Htway, Committee Chairperson, Public Health Faculty
Dr. Edward Irobi, Committee Member, Public Health Faculty
Dr. Raymond Panas, University Reviewer, Public Health Faculty

Chief Academic Officer and Provost
Sue Subocz, Ph.D.

Walden University
2023

Abstract

Effect of Treatment, Stage of Lung Cancer, and Socioeconomic Status on Life Expectancy Within Marginalized Communities

by

Elsadig Ahmed Elsharif

M.Sc. Medical Microbiology

Georgetown University, Washington, D.C., USA, 2009

M.B.CH.B. (MD)

University of Alexandria, School of Medicine, Egypt, 1985

Dissertation Submitted in Partial Fulfillment

of the Requirements for the Degree of

Doctor of Philosophy

Public Health

Walden University

February 2023

Abstract

Lung cancer is one of the most common cancers in the United States, and it accounts for 25% of cancer deaths. About 70% of cancer cases are diagnosed during late stages, leading to poor outcomes. An estimated 60% of cancer cases involve underserved and disadvantaged communities. However, there are limited studies had addressed effects of treatment, stage of lung cancer, and socioeconomic status on life expectancy within marginalized communities. Research questions examined effect of treatment, stage of lung cancer, and socioeconomic status on life expectancy of lung cancer patients between 2009 and 2019. This study was grounded in the deductive approach theory that facilitates interpretation of causal relationships between variables and concepts. The study was also grounded in the socioecological model, which acknowledges that different contributing factors and determinants exist at different levels of the society and addressing them at all levels will facilitate more effective prevention and control. A quantitative method with a cross-sectional design was used to analyze data from a random sample of 86,998 lung cancer patients. The dataset was obtained from the Surveillance, Epidemiology, and End Results database from the National Cancer Institute. Multiple linear regression was used for descriptive and inferential statistical analyses. Results showed treatment, stage of lung cancer, and socioeconomic status had statistically significant effects on life expectancy of lung cancer patients. Positive social change implications include alleviation of burden of lung cancer by raising awareness, encouraging screening, and advocating to enact new government policies.

Effect of Treatment, Stage of Lung Cancer, and Socioeconomic Status on Life Expectancy Within Marginalized Communities

by

Elsadig Ahmed Elsharif

M.Sc. Medical Microbiology

Georgetown University, Washington, D.C., USA, 2009

M.B.CH.B. (MD)

University of Alexandria, School of Medicine, Egypt, 1985

Dissertation Submitted in Partial Fulfillment

of the Requirements for the Degree of

Doctor of Philosophy

Public Health

Walden University

February 2023

Dedication

I would like to dedicate my doctoral dissertation to the memory of my parents, my mother Asmaa, and my father Elsharif Ahmed, and my family who supported me all the way in my life specially my daughter Lojain who encouraged me to start and continue my dissertation, but she passed away before I complete my study. Also, I dedicate my dissertation to my daughter Lama who encouraged and supported me with technical assistance and advise throughout my dissertation study. She was my cheerleader. Likewise, this dedication is to my wife Hanadi who contributed to my dissertation with support and encouragement as well. Such achievement would not be possible without the unconditional support and encouragement from my family.

Acknowledgments

I would like to acknowledge and thank my Dissertation Committee members, Dr. Zin Htway, chairperson, Dr. Edward Irobi, committtee member, and Dr. Raymond Panas, University Research Reviewer (URR) for their professional and insightful feedback and advice. I am greatful to Dr. Htway for his continuous feedback, support, and encouragement thoughout my dissertation study. Dr. Htway was always available to answer my questions. I admired his positive attitude. I also, appreciate Dr. Irobi's advice and guidence as well as Dr. Panas's insightful and invaliable feedback. Likewise, I would like to acknowledge and thank the National Cancer Institute (NCI) for giving me the opportunity to use their SEER database for my dissertation; as well as I acknowledge those who struggled and died due to lung cancer. If it hadn't been for them and for you all, I wouldn't have such achievement.

Table of Contents

List of Tables iv

List of Figures vi

Chapter 1: Introduction to the Study 1

Background 7

Problem Statement 10

Purpose of Study 14

Research Questions and Hypotheses 15

Theoretical Framework 23

Nature of the Study 25

Definitions 27

Assumptions 30

Scope and Delimitations 32

Limitations 35

Significance 38

Summary 40

Chapter 2: Literature Review 43

Literature Search Strategy 44

Theoretical Foundation 53

Lung Cancer 55

Diagnosis 57

Cancer Staging 59

Contributing Factors to Staging of Lung Cancer....62
Types of Treatment for Lung Cancer....66
Treatment of Choice for NSCLC Based on Stage of Lung Cancer....76
Contributing Factors to Survival/Life Expectancy....82
Summary and Conclusion....101
Chapter 3: Research Method....103
Research Design and Rationale....106
Methodology....108
Data Analysis Plan....110
Threats and Validity....124
Summary....127
Chapter 4: Results....129
Data Collection....136
Study Analyses Results....138
RQ1 138
RQ2 145
RQ3 152
RQ4 159
RQ5 183
RQ6 191
RQ7 202
Summary....213

Chapter 5: Discussion, Conclusions, and Recommendations 216

Interpretation of the Findings 219

RQ1 220

RQ2 221

RQ3 222

RQ4 223

RQ5 227

RQ6 229

RQ7 230

Limitations of the Study 234

Recommendations 235

Implications 237

Conclusion 240

References 242

Appendix A: Walden University IRB Approval Letter 256

Appendix B: NCI Approval for SEER Database Access 257

Appendix C: NIH Certificate of Completion of Training Course on Protecting Human Research Participants 258

List of Tables

Table 1 *Descriptive Statistics of Survival Months, Surgery, Stage of Lung Cancer, Socioeconomic Status, Race, Age, and Geographic Location of Lung Cancer Patients* 139

Table 2 *Descriptive Statistics of Survival Months, Chemotherapy, Stage of Lung Cancer, Socioeconomic Status, Race, Age, and Geographic location of Lung Cancer Patients* 146

Table 3 *Descriptive Statistics of Survival Months, Radiation, Stage of Lung Cancer, Socioeconomic Status, Race, Age, and Geographic Location of Lung Cancer Patients* 153

Table 4 *Descriptive Statistics of Survival Months, Surgery and Chemotherapy, Stage of Lung Cancer, Socioeconomic Status, Race, Age, and Geographic Location of Lung Cancer Patients*..... 160

Table 5 *Descriptive Statistics of Survival Months, Surgery and Radiation, Stage of Lung Cancer, Socioeconomic Status, Race, Age, and Geographic Location of Lung Cancer Patients*..... 167

Table 6 *Descriptive Statistics of Survival Months, Surgery, Radiation and Chemotherapy, Stage of Lung Cancer, Socioeconomic Status, Race, Age, and Geographic Location of Lung Cancer Patients*..... 173

Table 7 *Descriptive Statistics of Survival Months, Radiation and Chemotherapy, Stage of Lung Cancer, Socioeconomic Status, Race, Age, and Geographic Location of Lung Cancer Patients*..... 179

Table 8 *Descriptive Statistics of Survival Months, Stage of Lung Cancer, Surgery, Socioeconomic Status, Race, Age, and Geographic Location of Lung Cancer Patients* 185

Table 9 *Descriptive Statistics of Survival Months, Socioeconomic Status, Surgery, Chemotherapy, Radiotherapy, Surgery and Chemotherapy, Surgery and Radiotherapy, Radiotherapy and Chemotherapy, Surgery, Radiotherapy and Chemotherapy, Race, Age, Geographic Location, and Stage of Lung Cancer* 193

Table 10 *Descriptive Statistics of Survival Months, Race, Age, Geographic Location, Surgery, Stage of Lung Cancer, and Socioeconomic Status of Lung Cancer Patients* 203

List of Figures

Figure 1 *Association Between Surgery and Survival of Lung Cancer Patients*..... 140

Figure 2 *Association Between Chemotherapy and Survival of Lung Cancer Patients*. 147

Figure 3 *Association Between Radiotherapy and Survival of Lung Cancer Patients*.... 154

Figure 4 *Association Between Combination of Surgery/Chemotherapy and Survival of Lung Cancer Patients* 161

Figure 5 *Association Between Combination of Surgery/Radiation and Survival of Lung Cancer Patients*..... 168

Figure 6 *Association Between a Combination of Surgery, Radiation, and Chemotherapy, and Survival of Lung Cancer Patients* 174

Figure 7 *Association Between Combination of Radiation/Chemotherapy and Survival of Lung Cancer Patients* 180

Figure 8 *Association Between Stage of Lung Cancer and Survival of Lung Cancer Patients* 186

Figure 9 *Association Between Socioeconomic Status and Survival of Lung Cancer Patients* 194

Figure 10 *Association Between Race and Survival of Lung Cancer Patients* 204

Figure 11 *Association Between Age and Survival of Lung Cancer Patients* 210

Figure 12 *Association Between Geographic Location and Survival of Lung Cancer Patients* 212

Chapter 1: Introduction to the Study

Lung cancer is one of the most common cancers in the United States and it accounts for 25% of entire cancer deaths (Cancer Treatment Centers of America, 2020). The number of cases of lung cancer in 2020 according to the American Cancer Society was 228,820 and the number of deaths attributed to lung cancer was 135,720 which exceeded the number of deaths due to breast cancer, prostate cancer, and colorectal cancer combined (Cancer Treatment Centers of America, 2020; U.S. Cancer Statistic, n.d.). One male was diagnosed with lung cancer out of every 14 males and one female was diagnosed with lung cancer out of every 17 females (Lung Cancer Foundation of America, 2020). My study was about effect of treatment, stage of lung cancer, and socioeconomic status on life expectancy of marginalized communities. This study underscored the critical role of treatment, stage of lung cancer, and socioeconomic status on outcome (survival status). So, the social problem is lung cancer among marginalized and disadvantaged communities who suffer most due to late diagnosis and late treatment because of lack of access to health care which in turn lead to a short life expectancy.

The aim of this quantitative study was to examine retrospectively cohorts who were diagnosed with lung cancer at different stages, and received different treatment regimens that included surgery, chemotherapy, radiotherapy, or combinations of some of them and evaluate their course of treatments until their last follow-ups or deaths. The purpose of my study was to demonstrate the effect of treatment, stage of lung cancer, and socioeconomic status on life expectancy of marginalized communities. This study highlighted the importance of screening for lung cancer which could facilitate early

diagnosis, potential curative treatment, better prognosis, and a longer life expectancy which this study would contribute to as a potential positive social change (Khorana et al., 2019; Montagne et al., 2021). In this study I had also identified and analyzed the contributing factors for short survival due to lung cancer among marginalized communities.

The major sections in this chapter (chapter 1) included: *Background section* that reflected and highlighted the burden of lung cancer specially among marginalized communities as bout 60% of cancer affects underserved and disadvantaged communities with a higher incidence rate and mortality rate among countries with low to middle income (Montagne et al., 2021). Also, bout 70% of cancer cases were diagnosed at late stages that require multimodality treatment that were mostly palliative with poor prognosis and short survival (Montagne et al., 2021). The background section also included risk factors for lung cancer, a gap in literature and a gap in practice with an overview on my research study.

Problem Statement section, was an introduction to the actual health problem that triggered me to select of the topic of my research study *Effect of Treatment, Stage of Lung Cancer, and Socioeconomic Status on Life Expectancy of Marginalized Communities*, as lung cancer is considered one of the most common cancers in the United States and it accounts for 25% of the entire cancer deaths (Cancer Treatment Centers of America, 2020); which exceeded the number of deaths due to breast cancer, prostate cancer, and colorectal cancer combined (Cancer Treatment Centers of America, 2020). Lung cancer when diagnosed during early stages is mostly curable and 56% of

cases diagnosed with a localized cancer will be at a 5-year survival rate, while about 70% of patients diagnosed with advanced lung cancer, only 15% of them will be at a 5-year-survival rate (Azubuike et al., 2020).

Purpose of Study section, included a discussion about the purpose of this quantitative study as a study of retrospective cohorts who were diagnosed with lung cancer at different stages and received different treatment regimens that included surgery, chemotherapy, radiotherapy, or combinations of some of them, evaluated their socioeconomic status, courses of treatments and responses until their last follow-ups or deaths. As such, the purpose of this study was to examine the effect of treatment, stage of lung cancer, and socioeconomic status on survival/ life expectancy of marginalized communities. In this section some lights were shed on investigating and identifying the contributing factors for short survival due to lung cancer among marginalized communities; as underserved communities' inherent poor lifestyle with other contributing factors such as cigarette smoking, lack of or low education levels, low income, and lack of access to healthcare.

Research Questions and Hypotheses section included a summary of seven research questions and their hypotheses in details that included life expectancy as a dependent variable; the independent variables included surgery, chemotherapy, radiotherapy, a combination of some of them, stage of lung cancer, socioeconomic status, and confounders that included race/ ethnicity, age, and geographic location.

Theoretical Framework section encompassed the framework of the dissertation that was grounded in deductive approach theory (DAT) and the socioecological model.

This section further included explanation of the concept of DAT theory which was based on developing a hypothesis built on an existing theory so as my research study (Research Methodology, n.d.). The DAT concept could be explained further as, the propositions of existing theory could be used as a platform for developing a new hypothesis explaining the so-called deductive approach or DAT (Research Methodology, n.d.). A quantitative method was used in DAT as well as in my study (Research Methodology, n.d.). The logical connections between the framework presented and the nature of my study include, DAT facilitates interpretation of causal relationships between variables and concepts, such as the association between different types of treatment, stage of lung cancer, socioeconomic status, and survival/ life expectancy as in my study. The socioecological model could be applied to my research study as a framework with multifaceted levels of the society where individuals and environment interact with the social system.

Nature of the Study section was an introduction and summary of my research approach as retrospective quantitative analyses of secondary data from the Surveillance, Epidemiology, and End Results (SEER) database which included records of cancer patients in the US managed by the National Cancer Institute (NCI). This method was aligned with my dissertation topic, *Effect of Treatment, Stage of Lung Cancer, and Socioeconomic status* as independent variables on *life expectancy* (months/ years of survival following diagnosis) as a dependent variable. Furthermore, this section included justification for using a quantitative method in this study as I conducted descriptive and inferential statistics analyses to examine effect of treatment, stage of lung cancer, and socioeconomic status on life expectancy of marginalized communities. A cross-sectional

design was used for descriptive statistics and inferential analyses to evaluate the frequency and distribution of treatment, stage of lung cancer, and socioeconomic status on the population sample. The study sample was obtained from SEER database of the NCI for the period from 2009 to 2019 for patients diagnosed with lung cancer during this period and their cause-specific mortality was lung cancer. The statistical analyses were performed by *IBM SPSS Statistics* for Macintosh (Version 27.0).

Definitions section included basic definitions of some words and terms used in the study that are critical to understand. *Assumptions Section* was a summary of the assumptions of my research study which were based on my literature reviews. In this study, I assumed that there was a direct correlation, cause, and effect between type of treatment, stage of lung cancer, socioeconomic status, and survival/ life expectancy of cancer patients. These assumptions were necessary in the context of this study because they helped me to correctly drew a conclusion from the results of my analyses. Assumptions were also considered as requirements to be obtained before I could conduct my analyses.

Scope and Delimitations section was an introduction about a gap in literature which was the scarcity of research about effect of treatment, stage of lung cancer, and socioeconomic status on life expectancy of marginalized communities, which was examined in this research study. Also, in this study I had addressed a gap in practice which was represented by evidence-practice gap across the diagnosis and management process of lung cancer that require clinical practice to align with the recommended evidence-based guidelines to facilitate improvement of the outcome such as a better

prognosis and a longer life expectancy (Rankin et al., 2018). This section also included the external validity and internal validity that were related to this study. The *Limitations Section* expressed the limitations in the cross-sectional design that I used to examine the relationship between independent variables and dependent variable. That was because, cross-sectional design provides inferences about relationship between different variables, but it cannot demonstrate the cause and effect between independent and dependent variables because independent variables (risk factors) and dependent variable (outcome) are measured at the same time (Creswell & Creswell, 2018). Also, this section included limitations of the quantitative method that I also used in my study such as sampling errors that occur when the sample used does not represent the general population which is called sample bias or selection bias. Corrections of these limitations were also discussed in this section.

Significance section included clarification on the importance of this research study about effect of treatment, stage of lung cancer, and socioeconomic status on life expectancy of marginalized communities where many risk factors for lung cancer and barriers to health care are inherited and need to be analyzed and addressed to facilitate better government policies and raise the awareness of the communities. Also, this section included a summary about potential implications for a positive social change that were consistent with my study in addressing the burden of lung cancer such as avoiding risk factors, screening, early diagnosis, and early treatment (World Health Organization [WHO], 2019). Screening can lead to early diagnosis and early treatment which results in

more effective and less expensive treatment with a better outcome (WHO, 2019). And *Summary Section* which included summaries of all the above-mentioned sections.

Background

About 60% of cancer affected underserved and disadvantaged communities with a higher incidence rate and mortality rate among countries with low to middle income (Montagne et al., 2021). About 70% of cancer cases were diagnosed at late stages that required multimodality treatment (Montagne et al., 2021).

Tobacco use was attributed to about 25% of all cancer deaths and tobacco smoke accounted for 87% of deaths due to lung cancer (Baku ski et al., 2019). There are several other risk factors for lung cancer besides cigarette smoking such as secondhand smoke exposure, exposure to radiation, exposure to radon gas that resulted from uranium natural breakdown in soil, rocks, and water which polluted air especially in close living areas, and exposure to other carcinogens in workplace such as asbestos, and arsenic (Mayo Clinic, 2019). E-cigarettes emit formaldehyde into the lungs at a higher level which is also considered a carcinogen (Salamanca et al., 2018). Aging is considered an important risk factor for lung cancer due to accumulation of mutations in somatic cells during the individual's life span (Raniszewska et al., 2021). The International Agency for Research on Cancer (IARC) has recently updated the list of carcinogenic agents to human to include more than 100 agents classified as group one (Monographs Volume 100, parts A-F) (Smith et al., 2016).

Mahase et al. (2018) investigated survival disparities by regional poverty level based on radiotherapy treatment (RT) prescribed for lung cancer patients in these regions.

The purpose of the study was to evaluate the regional poverty level reflected on differences in lung cancer (LC) survival for lung cancer patients received radiotherapy (Mahase et al., 2018). The investigators used retrospective quantitative method for a study sample retrieved from the Surveillance, Epidemiology, and End Results (SEER) database which included patients diagnosed with LC during the period from 2000 to 2009 (Mahase et al., 2018). The multivariate (MVA) study results demonstrated that men had a higher mortality than women; Caucasians had lower mortality compared to African Americans while Asians, Pacific Islanders, and Native Americans had the highest overall survival rates (Mahase et al., 2018). The investigators concluded that RT may offer a positive survival benefit to those who received treatment when accounting for age, gender, race, and socioeconomic status (SES) (Mahase et al., 2018). The study further concluded that an incrementally worse overall survival (OS) rate was associated with increasing regional poverty level even for those who received RT (Mahase et al., 2018).

Burden of lung cancer can be reduced by avoiding risk factors, screening, early diagnosis, and early treatment (WHO, 2019). Lung cancer when diagnosed during early stages is mostly curable and 56% of cases diagnosed with a localized cancer will be at a 5-year survival rate, while about 70% of patients diagnosed with advanced lung cancer, only 15% of them will be at a 5-year-survival rate (Azubuike et al., 2020). Therefore, screening is an important tool that could facilitate early detection and diagnosis of lung cancer and early treatment which could be more effective and less expensive with a better outcome (WHO, 2019).

My research study was an effort to address and fill a gap in literature which was reflected by a scarcity of research about *Effect of Treatment, Stage of Lung Cancer, and Socioeconomic Status on Life Expectancy of Marginalized Communities* despite About 60% of cancer affects underserved and disadvantaged communities (Montagne et al., 2021). Also, my research study was a contribution to address a gap in practice which is represented by evidence-practice gaps across the diagnosis and management process of lung cancer that require clinical practice to align with the recommended evidence-based guidelines to facilitate improvement of the outcome (Rankin et al., 2018). That is because closing the evidence-practice gap will contribute to screening and early diagnosis that will facilitate a possible curable treatment, a better prognosis, and a longer survival/ life expectancy. However, the evidence-practice gap was addressed only in clinical research literature; and despite its public health and epidemiological importance it was not addressed in public health research nor in epidemiology research.

As such, my research study was a contribution and an attempt to raise the awareness of public health and epidemiology scholars and researchers to fill the gap in research and in practice as mentioned above by focusing on underserved communities who inherent poor lifestyle with other contributing factors such as lack of or low education levels, low income, and lack of access to healthcare. In this study I analyzed the underlying causes that made lung cancer stubborn and responsible for 25% of cancer mortality rate (Cancer Treatment Centers of America, 2020). Furthermore, I have recommended the steps that can contribute to early diagnosis and treatment which will improve life expectancy and quality of life among the underserved communities. This

explained the importance of my research study about *Effect of Treatment, Stage of Lung Cancer, and Socioeconomic Status on Life Expectancy of Marginalized Communities* where many risk factors for lung cancer and barriers to health care are inherited by marginalized communities and need to be analyzed and addressed to facilitate better government policies and raise the awareness of the communities (Borrayo et al., 2020). The assessment of survival of cancer patients at the population level is a major contributor to the decoration of healthcare policy for lung cancer (Mar et al., 2020).

Problem Statement

Lung cancer is considered one of the most common cancers in the United States and it accounts for 25% of the entire cancer deaths (Cancer Treatment Centers of America, 2020); which exceeded the number of deaths due to breast cancer, prostate cancer, and colorectal cancer combined (Cancer Treatment Centers of America, 2020). Lung cancer when diagnosed during early stages is mostly curable and 56% of cases diagnosed with a localized cancer will be at a 5-year survival rate, while about 70% of patients diagnosed with advanced lung cancer, only 15% of them will be at a 5-year-survival rate (Azubuike et al., 2020).

Lu et al. (2019) conducted a study to evaluate the changes in incidence, treatment, and survival of lung cancer during the last four decades from 1973 to 2015. SEER database was used by the authors in their study. Joint regression models were used to estimate the changes in incidence, treatment, and survival related to lung cancer. The results based on SEER database, 1,148,341 patients were diagnosed with lung cancer during the period from 1973 to 2015 including 646,662 males and 501,679 females

(whites = 960,808, black = 122,079, other races = 64,010, and unknown = 1,444) (Lu et al., 2019). The average incidence of lung cancer was 59.0/ 100,000 person/ year (Lu et al., 2019). The incidence peak was in 1992 then gradually decreased with a higher incidence rate in males than females and blacks were higher than other racial groups (Lu et al., 2019). The surgical rate for lung cancer was 25%, an increased use of chemotherapy, and a decreased use of radiotherapy (Lu et al., 2019). The 5-year relative survival rate has increased with time but remained low (<21%) (Lu et al., 2019). Chemotherapy combined with radiotherapy were used at a higher rate in late stages of lung cancer than early stages (Lu et al., 2019). The study demonstrated relative decrease in the incidence of lung cancer in the past four decades which was due to advances in the treatment of lung cancer such as hormonal therapy, immunotherapy, and targeted therapy. This conclusion validated my research study about effect of treatment regimen on lung cancer and life expectancy.

Lyu, (2020) investigated the risk factors that contribute to lung cancer survival. The investigator used Kaplan-Meier and Cox proportional hazard models as statistical models to analyze 1,145 patients diagnosed with different types of lung cancer. The dataset was extracted from The Cancer Genome Atlas Project (TCGA) database which is an organization that collects and stores huge numbers of gene data of cancer sequences that contributed to cancer treatment (Lyu, 2020). The dependent variable was survival of lung cancer patients, and the independent variables examined in this study were diagnosis, age, sex, smoking history, stage of lung cancer, fraction genome altered, and mutation count (Lyu, 2020). The results of the study revealed that the stage of lung

cancer is the most influential factor for the survival of lung cancer patients (Lyu, 2020). The impact of the other variables analyzed in this study such as age, sex, and smoking history had significant effect on survival of lung cancer patients only when they interact with time, reflecting their time-variant is associated with survival (Lyu, 2020). The investigator was not able to find any association between two genetic variables and survival (Lyu, 2020). The author argued that the stage of lung cancer is the most influential factor for the survival of lung cancer patients which is in line with my study that was intended to examine the effect of stage of lung cancer on life expectancy.

Rivera et al. (2020) performed a study with the objectives to gather essential available knowledge on disparities for lung cancer screening (LCS) that is characterized by eligibility criteria, access, and implementation to facilitate development of an official statement adopted by the American Thoracic Society to facilitate improving current screening guidelines and allocation of resources for unbiased LCS (Rivera et al., 2020). The authors based their study on available resources that identify disparities in lung cancer outcome among populations with the understanding that LCS can contribute to reduction of mortality due to lung cancer (Rivera et al., 2020). The method used in the study was a multidisciplinary panel that was represented by experts in LCS implementation science, primary healthcare, pulmonary, health behavior, smoking cessation, and epidemiology who all participated jointly on disparity research (Rivera et al., 2020). The multidisciplinary panel investigated available literature on disparity in cancer screening from the historical and emerging evidence perspectives (Rivera et al., 2020).

The results of the study revealed that current LCS guidelines do not consider lung cancer risk, smoking behavior differences between genders, socioeconomic status, race, and ethnicity (Rivera et al., 2020). The results also identified multiple barriers such as cost, and access to screening which contribute to disparity in implementation and dissemination of LCS (Rivera et al., 2020). The statement that resulted from this study identified the impact of LCS eligibility criteria on vulnerable communities who are considered at high risk for lung cancer but do not meet eligibility criteria for screening (Rivera et al., 2020). Furthermore, multiple barriers impact disparity in LCS implementation such as multilevel barriers (Rivera et al., 2020).

The authors recommended strategies that accommodate vulnerable populations for unbiased selection and dissemination of LCS such as addressing racial, ethnic, gender-based differences in smoking behaviors, socioeconomic, risk for lung cancer, access for LCS through health insurance coverage, provide unbiased LCS resources for vulnerable communities, and provide education and resources that contribute to impartial LCS results (Rivera et al., 2020). This study analyzed the contributing factors for the selection bias for lung cancer screening which will lead to a delay in the diagnosis and treatment of lung cancer and will negatively affect the life expectancy. Therefore, this study validates my research study about the effect of socioeconomic status on life expectancy of lung cancer patients in marginalized communities.

My research study was a humble contribution to address and fill a gap in literature which is reflected by a scarcity of research about effect of treatment, stage of lung cancer, and socioeconomic status on life expectancy of marginalized communities despite About

60% of cancer affects underserved and disadvantaged communities (Montagne et al., 2021). Also, my research study was an effort to address a gap in practice which is represented by evidence-practice gaps across the diagnosis and management process of lung cancer that require clinical practice to align with the recommended evidence-based guidelines to facilitate improvement of the outcome (Rankin et al., 2018). That is because closing the evidence-practice gap will contribute to screening and early diagnosis that will facilitate a possible curable treatment, a better prognosis, and a longer survival/ life expectancy. However, the evidence-practice gap was addressed only in clinical research literature; and despite its public health and epidemiological importance it was not addressed in public health or epidemiology research.

Purpose of Study

The purpose of this quantitative study was to investigate retrospectively cohorts who were diagnosed with lung cancer at different stages and received different treatment regimens that included surgery, chemotherapy, radiotherapy, or combinations of some of them, evaluate their socioeconomic status, courses of treatments and responses until their last follow-ups or deaths. As such, the purpose of my study was to examine the effect of treatment, stage of lung cancer, and socioeconomic status on survival/ life expectancy of marginalized communities. This study also had investigated and identified the contributing factors for short survival due to lung cancer among marginalized communities. Underserved communities' inherent poor lifestyle with other contributing factors such as cigarette smoking, lack of or low education levels, low income, and lack

of access to healthcare. These risk factors contribute to lung cancer that is responsible of 25% of cancer mortality (Cancer Treatment Centers of America, 2020).

Lung cancer when diagnosed during early stages is mostly curable and 56% of cases diagnosed with a localized cancer will be at a 5-year survival rate, while about 70% of patients diagnosed with advanced lung cancer, only 15% of them will be at a 5-year-survival rate (Azubuike et al., 2020). This highlights the importance of screening for lung cancer which will facilitate early diagnosis, possible curative treatment, better prognosis, and a longer life expectancy. The assessment of survival of lung cancer patients at the population level is a major contributor to the decoration of healthcare policy for lung cancer (Mar et al., 2020).

Research Questions and Hypotheses

Research Study Variables

Dependent Variable: Survival/ Life expectancy (months/ years of survival following diagnosis).

Independent Variables: Surgery, chemotherapy, radiotherapy, a combination of therapy (surgery and chemotherapy; surgery and radiotherapy; surgery, chemotherapy, and radiotherapy; or chemotherapy and radiotherapy), stage of lung cancer, and socioeconomic status.

Confounders: Race/ ethnicity, age, geographic location.

Research Questions and Hypotheses

RQ1. What is the association between surgery and survival/ life expectancy after controlling for chemotherapy, radiotherapy, a combination of therapy (surgery and

chemotherapy; surgery and radiotherapy; surgery, chemotherapy, and radiotherapy; or chemotherapy and radiotherapy), stage of lung cancer, socioeconomic status, race/ ethnicity, age, and geographic location (urban vs. rural)?

Null hypothesis (H01): There is no association between surgery and survival/ life expectancy after controlling for chemotherapy, radiotherapy, a combination of therapy (surgery and chemotherapy; surgery and radiotherapy; surgery, chemotherapy, and radiotherapy; or chemotherapy and radiotherapy), stage of lung cancer, socioeconomic status, race/ ethnicity, age, and geographic location (urban vs. rural).

Alternative hypothesis (Ha1): There is an association between surgery and survival/ life expectancy after controlling for chemotherapy, radiotherapy, a combination of therapy (surgery and chemotherapy; surgery and radiotherapy; surgery, chemotherapy, and radiotherapy; or chemotherapy and radiotherapy), stage of lung cancer, socioeconomic status, race/ ethnicity, age, and geographic location (urban vs. rural).

Dependent Variable: Survival/ Life expectancy (months/ years of survival following diagnosis).

Independent Variable: Surgery.

Covariates: chemotherapy, radiotherapy, a combination of therapy (surgery and chemotherapy; surgery and radiotherapy; surgery, chemotherapy, and radiotherapy; or chemotherapy and radiotherapy), stage of lung cancer, socioeconomic status, race/ ethnicity, age, and geographic location (urban vs. rural)

RQ2. What is the association between chemotherapy and survival/ life expectancy after controlling for surgery, radiotherapy, a combination of therapy (surgery and

chemotherapy; surgery and radiotherapy; surgery, chemotherapy, and radiotherapy; or chemotherapy and radiotherapy), stage of lung cancer, socioeconomic status, race/ ethnicity, age, and geographic location (urban vs. rural)?

H02: There is no association between chemotherapy and survival/ life expectancy after controlling for surgery, radiotherapy, a combination of therapy (surgery and chemotherapy; surgery and radiotherapy; surgery, chemotherapy, and radiotherapy; or chemotherapy and radiotherapy), stage of lung cancer, socioeconomic status, race/ ethnicity, age, and geographic location (urban vs. rural).

Ha2: There is an association between chemotherapy and survival/ life expectancy after controlling for surgery, radiotherapy, a combination of therapy (surgery and chemotherapy; surgery and radiotherapy; surgery, chemotherapy, and radiotherapy; or chemotherapy and radiotherapy), stage of lung cancer, socioeconomic status, race/ ethnicity, age, and geographic location (urban vs. rural).

Dependent Variable: Survival/ Life expectancy (months/ years of survival following diagnosis).

Independent Variable: Chemotherapy.

Covariates: Surgery, radiotherapy, a combination of therapy (surgery and chemotherapy; surgery and radiotherapy; surgery, chemotherapy, and radiotherapy; or chemotherapy and radiotherapy), stage of lung cancer, socioeconomic status, race/ ethnicity, age, and geographic location (urban vs. rural).

RQ3. What is the association between radiotherapy and survival/ life expectancy after controlling for surgery, chemotherapy, a combination of therapy (surgery and

chemotherapy; surgery and radiotherapy; surgery, chemotherapy, and radiotherapy; or chemotherapy and radiotherapy), stage of lung cancer, socioeconomic status, race/ ethnicity, age, and geographic location (urban vs. rural)?

H03: There is no association between radiotherapy and survival/ life expectancy after controlling for surgery, chemotherapy, a combination of therapy (surgery and chemotherapy; surgery and radiotherapy; surgery, chemotherapy, and radiotherapy; or chemotherapy and radiotherapy), stage of lung cancer, socioeconomic status, race/ ethnicity, age, and geographic location (urban vs. rural).

Ha3: There is an association between radiotherapy and survival/ life expectancy after controlling for surgery, chemotherapy, a combination of therapy (surgery and chemotherapy; surgery and radiotherapy; surgery, chemotherapy, and radiotherapy; or chemotherapy and radiotherapy), stage of lung cancer, socioeconomic status, race/ ethnicity, age, and geographic location (urban vs. rural).

Dependent Variable: Survival/ Life expectancy (months/ years of survival following diagnosis).

Independent Variable: Radiotherapy.

Covariates: Surgery, chemotherapy, a combination of therapy (surgery and chemotherapy; surgery and radiotherapy; surgery, chemotherapy, and radiotherapy; or chemotherapy and radiotherapy), stage of lung cancer, socioeconomic status, race/ ethnicity, age, and geographic location (urban vs. rural).

RQ4. What is the association between a combination of therapy (surgery and chemotherapy; surgery and radiotherapy; surgery, chemotherapy, and radiotherapy; or

chemotherapy and radiotherapy) and survival/ life expectancy after controlling for surgery, chemotherapy, radiotherapy, stage of lung cancer, socioeconomic status, race/ ethnicity, age, and geographic location (urban vs. rural)?

H04: There is no association between a combination of therapy (surgery and chemotherapy; surgery and radiotherapy; surgery, chemotherapy, and radiotherapy; or chemotherapy and radiotherapy) and survival/ life expectancy after controlling for surgery, chemotherapy, radiotherapy, stage of lung cancer, socioeconomic status, race/ ethnicity, age, and geographic location (urban vs. rural).

Ha4: There is an association between a combination of therapy (surgery and chemotherapy; surgery and radiotherapy; surgery, chemotherapy, and radiotherapy; or chemotherapy and radiotherapy) and survival/ life expectancy after controlling for surgery, chemotherapy, radiotherapy, stage of lung cancer, socioeconomic status, race/ ethnicity, age, and geographic location (urban vs. rural).

Dependent Variable: Survival/ Life expectancy (months/ years of survival following diagnosis).

Independent Variable: A combination of therapy (surgery and chemotherapy; surgery and radiotherapy; surgery, chemotherapy, and radiotherapy; and chemotherapy and radiotherapy).

Covariates: Surgery, chemotherapy, radiotherapy, stage of lung cancer, socioeconomic status, race/ ethnicity, age, and geographic location (urban vs. rural).

RQ5. What is the association between stage of lung cancer and survival/ life expectancy after controlling for surgery, chemotherapy, radiotherapy, a combination of

therapy (surgery and chemotherapy; surgery and radiotherapy; surgery, chemotherapy, and radiotherapy; or chemotherapy and radiotherapy), socioeconomic status, race/ ethnicity, age, and geographic location (urban vs. rural)?

H05: There is no association between stage of lung cancer and survival/ life expectancy after controlling for surgery, chemotherapy, radiotherapy, a combination of therapy (surgery and chemotherapy; surgery and radiotherapy; surgery, chemotherapy, and radiotherapy; or chemotherapy and radiotherapy), socioeconomic status, race/ ethnicity, age, and geographic location (urban vs. rural)

Ha5: There is an association between stage of lung cancer and survival/ life expectancy after controlling for surgery, chemotherapy, radiotherapy, a combination of therapy (surgery and chemotherapy; surgery and radiotherapy; surgery, chemotherapy, and radiotherapy; or chemotherapy and radiotherapy), socioeconomic status, race/ ethnicity, age, and geographic location (urban vs. rural).

Dependent Variable: Survival/ Life expectancy (months/ years of survival following diagnosis).

Independent Variable: Stage of lung cancer.

Covariates: Surgery, chemotherapy, radiotherapy, a combination of therapy (surgery and chemotherapy; surgery and radiotherapy; surgery, chemotherapy, and radiotherapy; or chemotherapy and radiotherapy), socioeconomic status, race/ ethnicity, age, and geographic location (urban vs. rural).

RQ6. What is the association between socioeconomic status and survival/ life expectancy after controlling for surgery, chemotherapy, radiotherapy, a combination of

therapy (surgery and chemotherapy; surgery and radiotherapy; surgery, chemotherapy, and radiotherapy; or chemotherapy and radiotherapy), stage of lung cancer, race/ ethnicity, age, and geographic location (urban vs. rural)?

H06: There is no association between socioeconomic status and survival/ life expectancy after controlling for surgery, chemotherapy, radiotherapy, a combination of therapy (surgery and chemotherapy; surgery and radiotherapy; surgery, chemotherapy, and radiotherapy; or chemotherapy and radiotherapy), stage of lung cancer, race/ ethnicity, age, and geographic location (urban vs. rural)?

Ha6: There is an association between socioeconomic status and survival/ life expectancy after controlling for surgery, chemotherapy, radiotherapy, a combination of therapy (surgery and chemotherapy; surgery and radiotherapy; surgery, chemotherapy, and radiotherapy; or chemotherapy and radiotherapy), stage of lung cancer, race/ ethnicity, age, and geographic location (urban vs. rural)?

Dependent Variable: Survival/ Life expectancy (months/ years of survival following diagnosis).

Independent Variable: Socioeconomic status.

Covariates: surgery, chemotherapy, radiotherapy, a combination of therapy (surgery and chemotherapy; surgery and radiotherapy; surgery, chemotherapy, and radiotherapy; or chemotherapy and radiotherapy), stage of lung cancer, race/ ethnicity, age, and geographic location (urban vs. rural)?

.

RQ7. What is the association between race/ ethnicity, age, and geographic location (urban vs. rural) and survival/ life expectancy after controlling for surgery, chemotherapy, radiotherapy, a combination of therapy (surgery and chemotherapy; surgery and radiotherapy; surgery, chemotherapy, and radiotherapy; or chemotherapy and radiotherapy), stage of lung cancer, and socioeconomic status?

H07: There is no association between race/ ethnicity, age, and geographic location (urban vs. rural) and survival/ life expectancy after controlling for surgery, chemotherapy, radiotherapy, a combination of therapy (surgery and chemotherapy; surgery and radiotherapy; surgery, chemotherapy, and radiotherapy; or chemotherapy and radiotherapy), stage of lung cancer, and socioeconomic status.

Ha7: There is an association between race/ ethnicity, age, and geographic location (urban vs. rural) and survival/ life expectancy after controlling for surgery, chemotherapy, radiotherapy, a combination of therapy (surgery and chemotherapy; surgery and radiotherapy; surgery, chemotherapy, and radiotherapy; or chemotherapy and radiotherapy), stage of lung cancer, and socioeconomic status.

Dependent Variable: Survival/ Life expectancy (months/ years of survival following diagnosis).

Independent Variables: Confounders: Race/ ethnicity, age, geographic location (urban vs rural).

Covariates: surgery, chemotherapy, radiotherapy, a combination of therapy (surgery and chemotherapy; surgery and radiotherapy; surgery, chemotherapy, and radiotherapy; or chemotherapy and radiotherapy), stage of lung cancer, and socioeconomic status.

Theoretical Framework

The framework of my dissertation topic could be best expressed by deductive approach theory (DAT) and the socioecological model. The concept of DAT was based on developing a hypothesis built on an existing theory so as my research study (Research Methodology, n.d.). This can be explained further as, the propositions of existing theory can be used as a platform for developing a new hypothesis explaining the so-called deductive approach or DAT (Research Methodology, n.d.). A quantitative method was used in DAT as well as in my study (Research Methodology, n.d.).

The logical connections between the framework presented and the nature of my study include, DAT facilitates interpretation of causal relationships between variables and concepts, such as the association between different types of treatment, stage of lung cancer, socioeconomic status, and survival/ life expectancy, as well as possible quantitative measurements of concepts, and a possible generalizability of research findings to a certain level (Research Methodology, n.d.). Therefore, DAT has served as a conceivable framework for my dissertation. DAT has the capability of testing a known theory or a phenomenon if it is valid in given circumstances (Research Methodology, n.d.). DAT trails the path of logic (Research Methodology, n.d.).; therefore, I considered DAT as a framework for my dissertation topic.

The social ecological model has five stages: (a) *Individual stage*, is the level of the individual knowledge and skills when the individual can gain the knowledge about disease, realize its danger, and what steps should be taken to avoid or control it. This is when the individual behavior starts to change and seeks help.

(b) Interpersonal stage is the stage that addresses the individual's relationship with others. In this phase he or she interacts with his or her circle of influence such as a family and friends. In this stage, the individual shares and gains knowledge about the disease which further helps along the way for seeking screening or treatment. (c) *Organizational stage*, when more sectors of the community become involved such as schools and businesses where health education, counselling, and health insurance can be sought. *(d) Community stage* involves more organizations and affiliates in the community such as hospitals and advocacy groups who come together and participate in raising awareness, funding, and provide more resources which contribute to the community health. And (e) *Public policy stage* when the government tackles prevention by developing agencies, establishing, and enforcing laws that protect the community and provide services to the population at a larger scale than other stages.

The socioecological model could be applied to my research study as a framework with multifaceted levels of the society where individuals and environment interact with the social system. Socioecological model acknowledges that different contributing factors and determinants exist at different levels of the society and addressing them at all levels will facilitate more effective prevention and control. Taking actions at multiple levels at the same time will facilitate prevention of risk factors more effectively for any potential health problem (CDC, 2018). As such, actions to raise the awareness about lung cancer at all levels of the society, as stated in the socioecological model, will facilitate services and funding to improve prevention and control of lung cancer which represent the logical connection between my research study and the socioecological model.

Nature of the Study

My research approach was a retrospective analysis of secondary data from SEER database using a quantitative method. This method was aligned with my dissertation topic, *effect of treatment, stage of lung cancer, and socioeconomic status* as independent variables on *life expectancy* (months/ years of survival following diagnosis) as a dependent variable. To examine the effect of these independent variables on the dependent variable and how that might affect the relationship, I conducted correlational research. Without manipulation of independent variables, correlational research can evaluate the relationship between variables and facilitates explanation of a noticed occurrence between variables (Chiang et al., 2017).

A cross-sectional design was performed to examine the relationship between independent variables (treatment such as surgery, chemotherapy, radiotherapy, combinations of some of them, stages of lung cancer, socioeconomic status) and dependent variable (life expectancy). Cross-sectional design provides inferences about relationship between different variables, but it cannot demonstrate the cause and effect between independent and dependent variables because independent variables (risk factors) and dependent variable (outcome) are measured at the same time; cross-sectional design also reflects the frequency of variables on the sample population (Creswell & Creswell, 2018). To improve the generalizability status of my study sample results on population, I applied Inferential statistics (Creswell & Creswell, 2018).

The statistics test that I used for my analyses was multiple linear regression because it is used when there are more than two measurement variables, one dependent

variable and other independent variables (McDonald, 2014). Such as reflected in my dissertation topic Effect of Treatment, Stage of Lung Cancer, and Socioeconomic Status on Life Expectancy of Marginalized Communities; which had one dependent variable, Life Expectancy of Marginalized Communities, and 9 independent variables, surgery, chemotherapy, radiotherapy, a combination of therapy (surgery and chemotherapy; surgery and radiotherapy; surgery, chemotherapy, and radiotherapy; or chemotherapy and radiotherapy), stage of lung cancer, and socioeconomic status.

A quantitative method was applied in this study because I conducted descriptive and inferential statistics analyses to examine effect of treatment, stage of lung cancer, and socioeconomic status on life expectancy of marginalized communities. A cross-sectional study design was used for inferential analyses; and descriptive statistics was used to evaluate the frequency and distribution of treatment, stage of lung cancer, and socioeconomic status on the population sample. The study sample was obtained from SEER database of the NCI for the period (2009-2019) for patients diagnosed with lung cancer during this period and their cause-specific mortality was lung cancer. The statistical analyses were performed by IBM SPSS Statistics for Macintosh, Version 27.0. The statistical analyses that were conducted on the dataset were descriptive statistics such as frequencies, means, and standard deviations (SD); and inferential statistical tests to examine the study hypotheses. Assessment of burden of disease for lung cancer was obtained from the incidence and that was used for allocation of health resources which both were obtained from SEER database using SEER*Stat (version 8.3.6) during the study period (2009-2019). The source was SEER Research Plus Data 17 Registries

November 2021 Sub [2009-2019], and the population sample size was 86,998 lung cancer patients.

Definitions

Carcinogen: Any substance or agent that can cause cancer (NCI, n.d.).

Carcinogenesis: The process that transforms normal cells into cancer cells (NCI, n.d.).

Comorbidity: The status of someone who has two or more diseases at the same time (NCI, n.d.).

Contraindication: The condition, such as a symptom or medical condition, that prohibits a person from receiving a specific treatment or procedure because it may cause harm (NCI, n.d.).

Covariates: Independents variables that are not part of the study but could influence the outcome (Oxford Languages Dictionary, 2022).

Dependent Variable: From statistical perspective, it is a variable whose values are a product of one or more other variables (CDC, n.d.). It is also called the outcome.

Determinants of Health: The contribution of individual, socioeconomic, and environmental factors to the health status at the individual level or population level (Harris, 2011).

Epigenetics: The study of the process of age and exposure to environmental factors such as chemicals, drugs, and diet that could lead to changes on how genes switch on and off without changing the actual DNA sequence. Such changes may contribute to the individual's risk of disease that may pass from parents to their children (NCI, n.d.).

Health Disparities: Morbidity and mortality differences experienced by sub-population due to various causes (Harris, 2011).

Health Equity: The opportunity that all people will have a chance to live a healthy life regardless of race, ethnicity, gender, sexual orientation, disability, education, job, religion, language, locality, or other factors. The results of lack of health equity will be differences in health outcome such severity of diseases, disabilities, quality of life, and death (NCI, n.d.).

Health Literacy: According to the U.S. Department of Health and Human Services (HHS), in Healthy People 2030, health literacy is defined as personal health literacy and organizational health literacy which were defined separately on the next page (p-29) (NIH, 2021).

Incidence: The number of cases of a disease with an onset during a specific period; often expressed as a rate. Incidence assesses morbidity or other developments within a specific period (Harris, 2011).

Independent Variable: A risk factor, an exposure, or other factors with characteristics that are hypothesized based on observations or measurements to influence an event or outcome (dependent variable) (CDC, n.d.).

Metastasis: Spreading of cancer to other parts of the body outside the original site of the original (primary) cancer. Also, called secondaries (NCI, n.d.).

Morbidity: An assessment of incidence or prevalence of disease in a specific population, locality, or a group of interest (Harris, 2011).

Mortality: An assessment of deaths in a specific population, locality, or a group of interest (Harris, 2011).

Occult stage Lung Cancer: Undetectable lung cancer on images or bronchoscopy but, cancer cells can be found in sputum or bronchial washings (NCI, n.d.).

Organizational Health Literacy: The ability of the organizations to contribute equitably to enable individuals to be health literate as defined in the *Personal Health Literacy* mentioned below (NIH, 2021).

Palliative Treatment/ Therapy: Treatment prescribed to relieve symptoms and pain caused by or related to cancer (NCI, n.d.).

Personal Health Literacy: "The degree to which individuals have the capacity to obtain, process, and understand basic health information and services needed to make appropriate health decisions." (NIH, para 2. 2021).

Prevalence: The number of cases with a disease, infected people, or other attribute present among people during a specific interval of time. It is expressed as a rate most of the time (Harris, 2011).

Risk factor: A personal attributes, an environmental exposure, or inherited characteristics associated with an increase of a specific disease, injury, or other health condition (CDC, n.d.).

Screening: Utilization of technology and procedures to identify individuals who are likely to have a disease, or they developed signs or symptoms of a disease compared to other individuals who are less likely to have the disease (Harris, 2011).

Socioeconomic Status: Classifying people based on their level of education, income, and type of job. Socioeconomic status is classified into high, medium, and low. People who have less access to educational, financial, social, and health resources are classified as people with lower socioeconomic status compared to people with a higher socioeconomic status. Accordingly, people with lower socioeconomic status tend to have poor health with chronic health conditions and disabilities (NCI, n.d.).

Stage of Cancer: The level of spread of cancer from the original site to different parts of the body (American Association of Cancer Research [AACR], 2011).

Surveillance: Systematic monitoring of health status of a population (Harris, 2011).

Survival/ Life expectancy: The average number of years a person of a known age is expected to live based on a statistical projection when current mortality rates did not change (CDC, n.d.).

Assumptions

The assumptions of my research study were based on my literature reviews. In this study, I assumed that there was a direct correlation, cause, and effect between the type of treatment, stage of lung cancer, socioeconomic status, and survival/ life expectancy among marginalized communities. The type of treatment such as surgery, chemotherapy, radiotherapy, or a combination of some of them as examples of the common treatment modalities for lung cancer that were prescribed based on the stage of lung cancer and the health condition of the patient (Khorana et al., 2019; Montagne et al., 2021). For instance, surgery alone is mostly prescribed for patients diagnosed with stage I

when lung cancer is localized and did not metastasize with good prognosis and a longer survival; however, adjuvant (additional) chemotherapy might be prescribed after surgery if the tumor size is large to avoid recurrence (American Cancer Society, 2021). Also, surgery might not be prescribed for stage I lung cancer if the patient's medical condition is deteriorated due to other comorbidities (American Cancer Society, 2021). Such limitations for surgery might not be possible to prove practically from the study sample despite that the patient was diagnosed with stage I lung cancer.

Chemotherapy protocol is a combination of different types of chemotherapy which are selected based on the stage of lung cancer and the responsiveness or resistance of certain types of lung cancer to certain types of chemotherapies such as squamous cell carcinoma, adenocarcinoma, and large cell carcinoma (Mayo Clinic, 2019). Also, the grading system that describes the microscopic cellular changes of cancer cells as well differentiated, moderately differentiated to poorly differentiated will have different levels of response and resistance to different types of chemotherapy and other treatment modalities despite the same stage of lung cancer. Therefore, chemotherapy could be effective or less effective based on responsiveness or resistance of certain types of lung cancer despite that both could be in the of same stage of lung cancer which also might not be possible to identify from the study sample.

Lung cancer can be due to genetic mutations which is diagnosed by Genotyping, an advanced diagnostic method that looks for genetic mutations using PCR-based methods or gene rearrangements using screening immunohistochemistry (Folch et al., 2015). Genotyping besides helping in diagnosis, it facilitates selecting the proper type of

treatment when cancer is correlated with genetic factors (Folch et al., 2015). In such a case, lung cancer due to genetic mutations will not respond to standard therapy protocols such as chemotherapy or radiotherapy but rather it will respond to a different category called targeted therapy drugs with different mechanisms of action that differ from chemotherapy (American Cancer Society, 2021). Patients diagnosed with this type of lung cancer, due to genetic mutations, in my datasets might not be correctly identified because targeted therapy is not one of my independent variables. These assumptions were necessary in the context of my study because they helped me to draw a conclusion from the results of my analyses. Assumptions are also considered as requirements to be obtained before conducting analyses.

Scope and Delimitations

Lung cancer in the United States accounts for 25% of the entire cancer deaths (Cancer Treatment Centers of America, 2020); which exceeded the number of deaths due to breast cancer, prostate cancer, and colorectal cancer combined (Cancer Treatment Centers of America, 2020). Lung cancer when diagnosed during early stages is mostly curable and 56% of cases diagnosed with a localized cancer will be at a 5-year survival rate, while about 70% of patients diagnosed with advanced lung cancer, only 15% of them will be at a 5-year-survival rate (Azubuike et al., 2020). About 60% of cancer affects underserved and disadvantaged communities with a higher incidence rate and mortality rate among countries with low to middle income (Montagne et al., 2021).

My research study was a humble contribution to address and fill a gap in literature which is reflected by a scarcity of research about effect of treatment, stage of lung cancer,

and socioeconomic Status on life expectancy of marginalized communities despite About 60% of cancer affects underserved and disadvantaged communities (Montagne et al., 2021). Also, my research study was an effort to address a gap in practice which is represented by evidence-practice gaps across the diagnosis and management process of lung cancer that require clinical practice to align with the recommended evidence-based guidelines to facilitate improvement of the outcome (Rankin et al., 2018). That is because closing the evidence-practice gap will contribute to screening and early diagnosis that will facilitate a possible curable treatment, a better prognosis, and a longer survival/ life expectancy. However, the evidence-practice gap was addressed only in clinical research literature; and despite its public health and epidemiological importance it was not addressed in public health or epidemiology research.

The internal validity threats that could be related to my study was selection bias which was avoided by random selection of participants. The internal validity threats could also be due to mortality or study attrition which I avoided by selection of a large study sample (n=86,998) to avoid dropouts, incomplete records, or mortality due to other causes (Creswell & Creswell, 2018).

The Study Inclusion Criteria

Patients diagnosed with lung cancer based on pathological laboratory study during the study period (2009-2019).

Lung cancer patients with specified stage of lung cancer based on TNM staging system during the study period.

Lung cancer patients whose specific cause of death was lung cancer during the study period.

The Study Exclusion Criteria

Lung cancer patients who were not diagnosed during the study period (2009-2019).

Lung cancer patients whose diagnoses were not based on pathological laboratory study (missed information).

Lung cancer patients without TNM staging (missed information).

Lung cancer patients whose specific cause of death was not lung cancer or was not identified (missed information). The primary endpoint of the study was considered from the date of the initial diagnosis to the date of cancer specific cause of death or to the date of the last follow up appointment.

The socioecological model could be applied to my research study as a framework with multifaceted levels of the society where individuals and environment interact with the social system. Socioecological model acknowledges that different contributing factors and determinants exist at different levels of the society and addressing them at all levels will facilitate more effective prevention and control. Taking actions at multiple levels at the same time will facilitate prevention of risk factors more effectively for any potential health problem (CDC, 2018). As such, actions to raise the awareness about lung cancer at all levels of the society, as stated in the socioecological model, will facilitate services and funding to better improve prevention and control of lung cancer which represent the logical connection between my research study and the socioecological model.

Threats to external validity in my study might not be experienced immediately as the study sample was obtained from the NCI national database on cancer statistics which represents 34% of cancer statistics in the US (NCI, 2021). On the other hand, with the rapid advances in technology, and advances in lung cancer diagnostic and therapeutic protocols, I highly recommend future research to expand on the same topic theme of my dissertation as threats to external validity will eventually take place as the time passes. To improve the generalizability status of my study sample results on population, I applied inferential statistical tests to examine my study hypotheses (Creswell & Creswell, 2018).

Limitations

The limitation in the study design could be expressed in the cross-sectional design that was performed to examine the relationship between independent variables (treatment such as surgery, chemotherapy, radiotherapy; stage of lung cancer; and socioeconomic status) and dependent variable (life expectancy). That is because, cross-sectional design provides inferences about relationship between different variables, but it cannot demonstrate the cause and effect between independent and dependent variables because independent variables (risk factors) and dependent variable (outcome) are measured at the same time (Creswell & Creswell, 2018). Also, another limitation of cross-sectional design is that the temporal link between the outcome (life expectancy) and the exposure (treatment such as surgery, chemotherapy, radiotherapy; stages of lung cancer; and socioeconomic status) cannot be determined because both (the outcome and the exposure) were examined at the same time (Creswell & Creswell, 2018).

However, the advantage of conducting a cross-sectional design was it allowed me to compare multiple variables at the same time and it saved me time and cost (Creswell & Creswell, 2018). Cross-sectional design was also useful in assessing the burden of disease in a defined population such as lung cancer in marginalized communities in my study and the health needs of a population which both were useful in planning and allocation of health resources (Creswell & Creswell, 2018). As such, these advantages outweighed the limitations which justified using the cross-sectional design in this study.

The limitations of the quantitative method that I used in my study could be sampling errors that might occur when the sample used does not represent the general population which is called sample bias or selection bias. I used a random population sample for my dataset to overcome this limitation; insufficient sample size for statistical analyses could be another quantitative method limitation which would affect the validity of the conclusion. To overcome this limitation, I used *G*Power software* (G*Power 3.1 manual, 2017) for power analysis to minimize bias on sample size.

The Internal Validity Threats

The internal validity threats that could be related to my study was selection bias which was avoided by random selection of the study sample; internal validity threats could be due to mortality or study attrition which was avoided by selection of a large study sample to avoid dropouts, incomplete records, or mortality due to other causes.

Types of Threats to External Validity

Interaction of selection and treatment which rise when characteristics of participants are narrow and cannot be generalized to other people who do not share the

characteristics of participants and the solution to this threat is to conduct additional study on participants with different characteristics (Creswell & Creswell, 2018).

Interaction of setting and treatment is another type of threat to external validity which takes place when characteristics of setting of participants are different from characteristics of setting of other individuals; therefore, they cannot be generalized to the general population; so, to avoid this type of threat, the researcher can conduct additional study on participants with different settings and compare the new results with the initial results (Creswell & Creswell, 2018).

Interaction of history and treatment is a third type of threat to external validity that occur when results of study are time-bound and they cannot be generalized to past or future situations; so, to overcome this threat, the researcher can repeat the study later in time and compare the results with the initial study to see if they are the same or different (Creswell & Creswell, 2018).

Other Validity Threats

Threats to statistical conclusion validity that results from inaccurate inferences drawn from the data due to inadequate statistical power or statistical assumptions were violated (Creswell & Creswell, 2018). Such threats were avoided in this study by using *G*Power software* (G*Power 3.1 manual, 2017) for power analysis and to minimize bias on the size of the sample population.

Threats to construct validity that occur due to inadequate definitions and measures of variables by the researcher (Creswell & Creswell, 2018). In my research study, variables were well defined, and they were accurately measured.

Confounders in this research study were race/ ethnicity, age, and geographical location (urban vs rural) which were avoided by controlling them when statistical analyses were conducted.

Significance

This research study was aimed to be a contribution to address and fill a gap in literature which is reflected by a scarcity of research about effect of treatment, stage of lung cancer, and socioeconomic status on life expectancy of marginalized communities despite About 60% of cancer affects underserved and disadvantaged communities (Montagne et al., 2021). Also, my research study was an effort to address a gap in practice which is represented by evidence-practice gaps across the diagnosis and management process of lung cancer that require clinical practice to align with the recommended evidence-based guidelines to facilitate improvement of the outcome (Melzer et al., 2020; Rankin et al., 2018). That is because closing the evidence-practice gap will contribute to screening and early diagnosis that will facilitate a possible curable treatment, a better prognosis, and a longer survival/ life expectancy. However, the evidence-practice gap was addressed only in clinical research literature; and despite its public health and epidemiological importance it was not addressed in public health or epidemiology research.

So, this study would be one of a few if not the first epidemiological study that addresses the evidence-practice gap. That is because despite my extensive search in literature, I did not come across a single epidemiological research study that addressed

the evidence-practice gap while many clinical research studies were conducted in this regard.

About 60% of cancer affects underserved and disadvantaged communities with a higher incidence rate and mortality rate among countries with low to middle income (Montagne et al., 2021). About 70% of cancer cases were diagnosed at late stages that require multimodality treatment (Montagne et al., 2021). As such, this study had focused on underserved communities who inherent poor lifestyle with other contributing factors such as lack of or low education levels, low income, and lack of access to healthcare. Therefore, the study included identification and analyses of the contributing factors to disparities among lung cancer patients regarding treatment, stage of lung cancer at the time of diagnosis, socioeconomic status, and survival/ life expectancy; together with identification of the evidence-practice gap and suggested pathways to achieve health equity in preventive services, early diagnosis, and early treatment that will eventually lead to a better outcome represented in quality of life and a longer survival.

Positive Social Change

The potential implications for a positive social change that are consistent with my study were addressing the burden of lung cancer specially among marginalized and underserved communities such as avoiding risk factors, screening, early diagnosis, and early treatment (Osarogiagbon, 2018; WHO, 2019). Screening can lead to early diagnosis and early treatment which could be more effective and less expensive with a better outcome (WHO, 2019). This explained the importance of this research study about effect of treatment, stage of lung cancer, and socioeconomic status on life expectancy of

marginalized communities where many risk factors for lung cancer and barriers to health care are inherited and need to be analyzed and addressed to facilitate better government policies and raise awareness of the communities.

Summary

Lung cancer is considered one of the most common cancers in the United States and it accounts for 25% of the entire cancer deaths (Cancer Treatment Centers of America, 2020). About 60% of cancer affects underserved and disadvantaged communities with a higher incidence rate and mortality rate among countries with low to middle income (Montagne et al., 2021). About 70% of cancer cases were diagnosed at late stages that require multimodality treatment (Montagne et al., 2021). So, the social problem is lung cancer among marginalized and disadvantaged communities who suffer most due to late diagnosis and late treatment because of lack of access to health care which in turn lead to a short life expectancy. This study had examined the effect of treatment, stage of lung cancer, and socioeconomic status on life expectancy of marginalized communities. This underscored the critical role of treatment, stage of lung cancer, and socioeconomic status on the outcome (survival status).

The research approach was a retrospective analysis of secondary data from SEER database using a quantitative method. The purpose was to study retrospectively cohorts who were diagnosed with lung cancer at different stages and received different treatment regimens that included surgery, chemotherapy, radiotherapy, or combinations of some of them, evaluate their socioeconomic status, courses of treatments and responses until their last follow-ups or deaths. There are several other risk factors for lung cancer besides

cigarette smoking such as secondhand smoke exposure, exposure to radiation, exposure to radon gas that resulted from uranium natural breakdown in soil, rocks, and water which polluted air especially in close living areas, and exposure to other carcinogens in workplace such as asbestos, and arsenic (Mayo Clinic, 2019). Aging is considered an important risk factor for lung cancer due to accumulation of mutations in somatic cells during the individual's life span (Raniszewska et al., 2021).

This research study was an attempt to address and fill a gap in literature which is reflected by a scarcity of research about effect of treatment, stage of lung cancer, and socioeconomic status on life expectancy of marginalized communities despite about 60% of cancer affects underserved and disadvantaged communities (Montagne et al., 2021). Also, the research study was an effort to address a gap in practice which is represented by evidence-practice gaps across the diagnosis and management process of lung cancer that require clinical practice to align with the recommended evidence-based guidelines to facilitate improvement of the outcome (Melzer et al., 2020; Rankin et al., 2018). The framework of this dissertation was grounded in the deductive approach theory (DAT) and the socioecological model.

A cross-sectional design was performed to examine the relationship between independent variables (treatment such as surgery, chemotherapy, radiotherapy; or a combination of some of them; stages of lung cancer; socioeconomic status) and dependent variable (life expectancy). A multiple linear regression was conducted to examine the association between surgery, chemotherapy, radiotherapy, a combination of therapy (surgery and chemotherapy; surgery and radiotherapy; surgery, chemotherapy,

and radiotherapy; chemotherapy and radiotherapy), stage of lung cancer, socioeconomic status, and life expectancy after controlling for race/ ethnicity, age, and geographic location (urban vs. rural). The statistical analysis was performed by *IBM SPSS Statistics* for Macintosh, Version 27.0. Assessment of burden of disease for lung cancer was obtained from the incidence and that was used for allocation of health resources; both were obtained from SEER database using SEER*Stat (version 8.3.6) during the study period (2009-2019). The potential implications for a positive social change that are consistent with this study will be addressing the burden of lung cancer such as avoiding risk factors, screening, early diagnosis, and early treatment (WHO, 2019). Screening can lead to early diagnosis and early treatment which will be more effective and less expensive with a better outcome (WHO, 2019). Also, the study will raise the awareness of the communities for risk factors and the importance of screening for high-risk groups. The study will advocate for reviewing current government policies toward screening, health education, and health care accessibility.

Chapter 2: Literature Review

Lung cancer is considered one of the most common cancers in the United States and it accounts for 25% of the entire cancer deaths (Cancer Treatment Centers of America, 2020). The number of cases diagnosed with lung cancer in the year 2020 according to the American Cancer Society was 228,820 and the number of deaths attributed to lung cancer was 135,720 which exceeded the number of deaths due to breast cancer, prostate cancer, and colorectal cancer combined (Cancer Treatment Centers of America, 2020; U.S. Cancer Statistic, n.d.). Currently, one male is diagnosed with lung cancer out of every 14 males and one female is diagnosed with lung cancer out of every 17 females (Lung Cancer Foundation of America, 2020).

This was what prompted me to search literature on lung cancer to find why lung cancer is in such a dire situation. In my literature review I came across what is called evidence-practice gaps across the diagnosis and management process of lung cancer that require clinical practice to align with the recommended evidence-based guidelines to facilitate improvement of the outcome (Rankin et al., 2018). The evidence-practice gaps as identified by Rankin et al. (2018) can be summarized in delays of diagnosis and referrals, underutilization of curative and palliative treatments, treatment is influenced by older age and comorbidities, multidisciplinary team is not part of most lung cancer management team, and psychosocial support is not utilized as a part of lung cancer care (McGregor et al., 2017; Rankin et al., 2018). My research study has examined the effect of treatment, stage of lung cancer, and socioeconomic status on life expectancy of marginalized communities. This study had underscored the critical role of treatment,

stage of lung cancer, and socioeconomic status on the outcome (survival status). So, the social problem is lung cancer among marginalized and disadvantaged communities who suffer most due to late diagnosis and late treatment because of lack of access to health care which in turn lead to a short life expectancy.

The aim of this quantitative study was to study retrospectively cohorts who were diagnosed with lung cancer at different stages, and they received different treatment regimens that included surgery, chemotherapy, radiotherapy, or combinations of some of them and evaluate their course of treatments until their last follow-ups or deaths. The purpose of my study was to demonstrate the effect of treatment, stage of lung cancer, and socioeconomic status on life expectancy of marginalized communities. This highlighted the importance of screening for lung cancer which will facilitate early diagnosis, possible curative treatment, better prognosis, and a longer life expectancy. This study had also included identification of the contributing factors for short survival due to lung cancer among marginalized communities.

The major sections in this chapter (chapter 2) included literature Search Strategy, Lung Cancer, Diagnosis, Cancer Staging, Contributing Factors to Stage of Lung Cancer, Types of Treatment for Lung Cancer, Treatment of Choice for Non-small Cell lung cancer (NSCLC) Based on the Stage of Lung Cancer, Contributing Factors to Survival/ Life Expectancy, Theoretical Foundation, Summary, and Conclusion.

Literature Search Strategy

For literature search, two databases were used: The Walden University library database and the Centers for Disease Control and Prevention (CDC) web base. Two

search engines were used, PubMed and Google Scholar, for reviewing scholarly articles relevant to this research study about effect of treatment, stage of lung cancer, and socioeconomic status on life expectancy of marginalized communities using keywords with Microsoft 10. The keywords used for the search were lung cancer, treatment, diagnosis, cancer stage, incidence, prevalence, comorbidity, compliance, survival, marginalized communities, screening, disparities, and contributing factors to survival. The literature review included peer-reviewed articles, seminal, and annals within the last 5 years (2017-2021) and a few older articles that were considered as a foundation for the forehand research study. The literature review included concise summaries and syntheses of the findings to elucidate their relationship with the variables of this study. The seven research questions in this dissertation that were answered by statistical analyses are:

Research Study Variables

Dependent Variable: Survival/ Life expectancy (months/ years of survival following diagnosis).

Independent Variables: Surgery, chemotherapy, radiotherapy, a combination of therapy (surgery and chemotherapy; surgery and radiotherapy; surgery, chemotherapy, and radiotherapy; chemotherapy and radiotherapy), stage of lung cancer, and socioeconomic status.

Confounders: Race/ ethnicity, age, geographic location.

Research Questions and Hypotheses

RQ1. What is the association between surgery and survival/ life expectancy after controlling for chemotherapy, radiotherapy, a combination of therapy (surgery and

chemotherapy; surgery and radiotherapy; surgery, chemotherapy, and radiotherapy; or chemotherapy and radiotherapy), stage of lung cancer, socioeconomic status, race/ ethnicity, age, and geographic location (urban vs. rural)?

Null hypothesis (H01): There is no association between surgery and survival/ life expectancy after controlling for chemotherapy, radiotherapy, a combination of therapy (surgery and chemotherapy; surgery and radiotherapy; surgery, chemotherapy, and radiotherapy; or chemotherapy and radiotherapy), stage of lung cancer, socioeconomic status, race/ ethnicity, age, and geographic location (urban vs. rural).

Alternative hypothesis (Ha1): There is an association between surgery and survival/ life expectancy after controlling for chemotherapy, radiotherapy, a combination of therapy (surgery and chemotherapy; surgery and radiotherapy; surgery, chemotherapy, and radiotherapy; or chemotherapy and radiotherapy), stage of lung cancer, socioeconomic status, race/ ethnicity, age, and geographic location (urban vs. rural).

Dependent Variable: Survival/ Life expectancy (months/ years of survival following diagnosis).

Independent Variable: Surgery.

Covariates: chemotherapy, radiotherapy, a combination of therapy (surgery and chemotherapy; surgery and radiotherapy; surgery, chemotherapy, and radiotherapy; chemotherapy and radiotherapy), stage of lung cancer, socioeconomic status, race/ ethnicity, age, and geographic location (urban vs. rural)

RQ2. What is the association between chemotherapy and survival/ life expectancy after controlling for surgery, radiotherapy, a combination of therapy (surgery and

chemotherapy; surgery and radiotherapy; surgery, chemotherapy, and radiotherapy; or chemotherapy and radiotherapy), stage of lung cancer, socioeconomic status, race/ ethnicity, age, and geographic location (urban vs. rural)?

H02: There is no association between chemotherapy and survival/ life expectancy after controlling for surgery, radiotherapy, a combination of therapy (surgery and chemotherapy; surgery and radiotherapy; surgery, chemotherapy, and radiotherapy; or chemotherapy and radiotherapy), stage of lung cancer, socioeconomic status, race/ ethnicity, age, and geographic location (urban vs. rural).

Ha2: There is an association between chemotherapy and survival/ life expectancy after controlling for surgery, radiotherapy, a combination of therapy (surgery and chemotherapy; surgery and radiotherapy; surgery, chemotherapy, and radiotherapy; or chemotherapy and radiotherapy), stage of lung cancer, socioeconomic status, race/ ethnicity, age, and geographic location (urban vs. rural).

Dependent Variable: Survival/ Life expectancy (months/ years of survival following diagnosis).

Independent Variable: Chemotherapy.

Covariates: Surgery, radiotherapy, a combination of therapy (surgery and chemotherapy; surgery and radiotherapy; surgery, chemotherapy, and radiotherapy; chemotherapy and radiotherapy), stage of lung cancer, socioeconomic status, race/ ethnicity, age, and geographic location (urban vs. rural).

RQ3. What is the association between radiotherapy and survival/ life expectancy after controlling for surgery, chemotherapy, a combination of therapy (surgery and

chemotherapy; surgery and radiotherapy; surgery, chemotherapy, and radiotherapy; or chemotherapy and radiotherapy), stage of lung cancer, socioeconomic status, race/ ethnicity, age, and geographic location (urban vs. rural)?

H03: There is no association between radiotherapy and survival/ life expectancy after controlling for surgery, chemotherapy, a combination of therapy (surgery and chemotherapy; surgery and radiotherapy; surgery, chemotherapy, and radiotherapy; or chemotherapy and radiotherapy), stage of lung cancer, socioeconomic status, race/ ethnicity, age, and geographic location (urban vs. rural).

Ha3: There is an association between radiotherapy and survival/ life expectancy after controlling for surgery, chemotherapy, a combination of therapy (surgery and chemotherapy; surgery and radiotherapy; surgery, chemotherapy, and radiotherapy; or chemotherapy and radiotherapy), stage of lung cancer, socioeconomic status, race/ ethnicity, age, and geographic location (urban vs. rural).

Dependent Variable: Survival/ Life expectancy (months/ years of survival following diagnosis).

Independent Variable: Radiotherapy.

Covariates: Surgery, chemotherapy, a combination of therapy (surgery and chemotherapy; surgery and radiotherapy; surgery, chemotherapy, and radiotherapy; chemotherapy and radiotherapy), stage of lung cancer, socioeconomic status, race/ ethnicity, age, and geographic location (urban vs. rural).

RQ4. What is the association between a combination of therapy (surgery and chemotherapy; surgery and radiotherapy; surgery, chemotherapy, and radiotherapy; or

chemotherapy and radiotherapy) and survival/ life expectancy after controlling for surgery, chemotherapy, radiotherapy, stage of lung cancer, socioeconomic status, race/ ethnicity, age, and geographic location (urban vs. rural)?

H04: There is no association between a combination of therapy (surgery and chemotherapy; surgery and radiotherapy; surgery, chemotherapy, and radiotherapy; or chemotherapy and radiotherapy) and survival/ life expectancy after controlling for surgery, chemotherapy, radiotherapy, stage of lung cancer, socioeconomic status, race/ ethnicity, age, and geographic location (urban vs. rural).

Ha4: There is an association between a combination of therapy (surgery and chemotherapy; surgery and radiotherapy; surgery, chemotherapy, and radiotherapy; or chemotherapy and radiotherapy) and survival/ life expectancy after controlling for surgery, chemotherapy, radiotherapy, stage of lung cancer, socioeconomic status, race/ ethnicity, age, and geographic location (urban vs. rural).

Dependent Variable: Survival/ Life expectancy (months/ years of survival following diagnosis).

Independent Variable: A combination of therapy (surgery and chemotherapy; surgery and radiotherapy; surgery, chemotherapy, and radiotherapy; chemotherapy and radiotherapy).

Covariates: Surgery, chemotherapy, radiotherapy, stage of lung cancer, socioeconomic status, race/ ethnicity, age, and geographic location (urban vs. rural).

RQ5. What is the association between stage of lung cancer and survival/ life expectancy after controlling for surgery, chemotherapy, radiotherapy, a combination of therapy (surgery and chemotherapy; surgery and radiotherapy; surgery, chemotherapy,

and radiotherapy; or chemotherapy and radiotherapy), socioeconomic status, race/ ethnicity, age, and geographic location (urban vs. rural)?

H05: There is no association between stage of lung cancer and survival/ life expectancy after controlling for surgery, chemotherapy, radiotherapy, a combination of therapy (surgery and chemotherapy; surgery and radiotherapy; surgery, chemotherapy, and radiotherapy; or chemotherapy and radiotherapy), socioeconomic status, race/ ethnicity, age, and geographic location (urban vs. rural)

Ha5: There is an association between stage of lung cancer and survival/ life expectancy after controlling for surgery, chemotherapy, radiotherapy, a combination of therapy (surgery and chemotherapy; surgery and radiotherapy; surgery, chemotherapy, and radiotherapy; or chemotherapy and radiotherapy), socioeconomic status, race/ ethnicity, age, and geographic location (urban vs. rural).

Dependent Variable: Survival/ Life expectancy (months/ years of survival following diagnosis).

Independent Variable: Stage of lung cancer.

Covariates: Surgery, chemotherapy, radiotherapy, a combination of therapy (surgery and chemotherapy; surgery and radiotherapy; surgery, chemotherapy, and radiotherapy; chemotherapy and radiotherapy), socioeconomic status, race/ ethnicity, age, and geographic location (urban vs. rural).

RQ6. What is the association between socioeconomic status and survival/ life expectancy after controlling for surgery, chemotherapy, radiotherapy, a combination of therapy (surgery and chemotherapy; surgery and radiotherapy; surgery, chemotherapy,

and radiotherapy; or chemotherapy and radiotherapy), stage of lung cancer, race/ ethnicity, age, and geographic location (urban vs. rural)?

H06: There is no association between socioeconomic status and survival/ life expectancy after controlling for surgery, chemotherapy, radiotherapy, a combination of therapy (surgery and chemotherapy; surgery and radiotherapy; surgery, chemotherapy, and radiotherapy; or chemotherapy and radiotherapy), stage of lung cancer, race/ ethnicity, age, and geographic location (urban vs. rural)?

Ha6: There is an association between socioeconomic status and survival/ life expectancy after controlling for surgery, chemotherapy, radiotherapy, a combination of therapy (surgery and chemotherapy; surgery and radiotherapy; surgery, chemotherapy, and radiotherapy; or chemotherapy and radiotherapy), stage of lung cancer, race/ ethnicity, age, and geographic location (urban vs. rural)?

status.

Dependent Variable: Survival/ Life expectancy (months/ years of survival following diagnosis).

Independent Variable: Socioeconomic status.

Covariates: Surgery, chemotherapy, radiotherapy, a combination of therapy (surgery and chemotherapy; surgery and radiotherapy; surgery, chemotherapy, and radiotherapy; chemotherapy and radiotherapy), stage of lung cancer, race/ ethnicity, age, and geographic location (urban vs. rural).

RQ7. What is the association between race/ ethnicity, age, and geographic location (urban vs. rural) and survival/ life expectancy after controlling for surgery,

chemotherapy, radiotherapy, a combination of therapy (surgery and chemotherapy; surgery and radiotherapy; surgery, chemotherapy, and radiotherapy; or chemotherapy and radiotherapy), stage of lung cancer, and socioeconomic status?

H07: There is no association between race/ ethnicity, age, and geographic location (urban vs. rural) and survival/ life expectancy after controlling for surgery, chemotherapy, radiotherapy, a combination of therapy (surgery and chemotherapy; surgery and radiotherapy; surgery, chemotherapy, and radiotherapy; or chemotherapy and radiotherapy), stage of lung cancer, and socioeconomic status.

Ha7: There is an association between race/ ethnicity, age, and geographic location (urban vs. rural) and survival/ life expectancy after controlling for surgery, chemotherapy, radiotherapy, a combination of therapy (surgery and chemotherapy; surgery and radiotherapy; surgery, chemotherapy, and radiotherapy; or chemotherapy and radiotherapy), stage of lung cancer, and socioeconomic status.

Dependent Variable: Survival/ Life expectancy (months/ years of survival following diagnosis).

Independent Variables: Confounders: Race/ ethnicity, age, geographic location (urban vs rural).

Covariates: surgery, chemotherapy, radiotherapy, a combination of therapy (surgery and chemotherapy; surgery and radiotherapy; surgery, chemotherapy, and radiotherapy; or chemotherapy and radiotherapy), stage of lung cancer, and socioeconomic status.

Theoretical Foundation

The framework of my dissertation topic could be best expressed by deductive approach theory (DAT) and the socioecological model. The concept of DAT was based on developing a hypothesis built on an existing theory so as this research study (Research Methodology, n.d.). This can be explained further as, the propositions of existing theory can be used as a platform for developing a new hypothesis explaining the so-called deductive approach or DAT (Research Methodology, n.d.). A quantitative method was used in DAT as well as in this study (Research Methodology, n.d.).

The logical connections between the framework presented and the nature of this study include, DAT facilitates interpretation of causal relationships between variables and concepts, such as the association between different types of treatment, stage of lung cancer, socioeconomic status, and survival/ life expectancy, as well as possible quantitative measurements of concepts, and a possible generalizability of research findings to a certain level (Research Methodology, n.d.). Therefore, DAT has served as a conceivable framework for my dissertation. DAT has the capability of testing a known theory or a phenomenon if it is valid in given circumstances (Research Methodology, n.d.). DAT trails the path of logic (Research Methodology, n.d.).; therefore, I considered DAT as a framework for my dissertation topic.

The social ecological model has five stages: (a) *Individual stage*, is the level of the individual knowledge and skills when the individual can gain the knowledge about disease, realize its danger, and what steps should be taken to avoid or control it. This is when the individual behavior starts to change and seeks help.

(b) Interpersonal stage is the stage that addresses the individual's relationship with others. In this phase he or she interacts with his or her circle of influence such as a family and friends. In this stage, the individual shares and gains knowledge about the disease which further helps along the way for seeking screening or treatment. (c) *Organizational stage*, when more sectors of the community become involved such as schools and businesses where health education, counselling, and health insurance can be sought. (d) *Community stage* involves more organizations and affiliates in the community such as hospitals and advocacy groups who come together and participate in raising awareness, funding, and provide more resources which contribute to the community health. And (e) *Public policy stage* when the government tackles prevention by developing agencies, establishing, and enforcing laws that protect the community and provide services to the population at a larger scale than other stages.

The socioecological model could be applied to my research study as a framework with multifaceted levels of the society where individuals and environment interact with the social system. Socioecological model acknowledges that different contributing factors and determinants exist at different levels of the society and addressing them at all levels will facilitate more effective prevention and control. Taking actions at multiple levels at the same time will facilitate prevention of risk factors more effectively for any potential health problem (CDC, 2018). As such, actions to raise the awareness about lung cancer at all levels of the society, as stated in the socioecological model, will facilitate services and funding to better improve prevention and control of lung cancer which represent the logical connection between my research study and the socioecological model.

Lung Cancer

Cancer is characterized by uncontrolled cell division and multiplications in any parts of the body which is called primary cancer and can spread to regional or distant parts of the body which is called metastases or secondaries (NCI, 2021). Cancer cells divide and multiply without receiving any signal as normal body cell do and invade other areas of the body while normal cells do not (NCI, 2021). The DNA of cancer cells acquires deletions and duplications of parts of its chromosomes which make cancer cells uncontrolled by the signal that regulates normal cell division and stop dividing by a mechanism called apoptosis or programmed cell death (NCI, 2021). When normal cells transform into cancer cells, they lose their principal function due to DNA changes (NCI, 2021). Lung cancer is a generic name that represents a disease of a diverse characteristics based on histologic and molecular variables that translate into different clinical stages that require different treatment modalities (Montagne et al., 2021).

Lung cancer is a cancer that originates primarily in the lungs (Mayo Clinic, 2019). Lungs are two adjacent organs located in the thoracic cavity (chest) and they are part of the respiratory system (American Lung Association, 2020). The main function of the lungs is gas exchange which is a mechanism that facilitates delivering oxygen from the air we inhale through bloodstream to all body cells and removes carbon dioxide (a waste gas) from body cells through bloodstream and exhale it out of the body (American Lung Association, 2020). Cigarette smoking represents the highest risk factor for lung cancer (Mayo Clinic, 2019). As the number of smoked cigarettes increases and the time spent in smoking increases the risk of lung cancer increases as well (Mayo Clinic, 2019). Non-

smokers can also develop Lung cancer (Mayo Clinic, 2019). Lung cancer in early stages has no signs or symptoms; while late stages of lung cancer are manifested by many signs and symptoms such as persistent cough, hemoptysis (coughing blood), Dyspnea (shortness of breath), hoarseness, chest pain, weight loss, and bone aches (Mayo Clinic, 2019). Cigarette smoke is a carcinogen (a substance that can cause cancer) that causes damage to lung cells beyond repair as the number of smoked cigarettes increases and the time spent in smoking increases; and this will eventually cause lung cells function abnormally and transform into cancer cells (Mayo Clinic, 2019).

Lung cancer has two major types based on the morphology of lung cells; small cell lung cancer (SCLC) which is linked to smoking as a contributing factor, and non-small cell lung cancer (NSCLC) which is further classified into three types of lung cancer, squamous cell carcinoma, adenocarcinoma, and large cell carcinoma based on the cell type within the lung (histological classification) (Mayo Clinic, 2019). NSCLC with its three main types is the most frequently diagnosed primary lung cancer as it accounts for 85% of primary lung cancer and it is diagnosed at late stages which limit treatment to palliative treatment only (Folch et al., 2015; Lu et al., 2019). There are several other risk factors for lung cancer besides cigarette smoking such as secondhand smoke exposure, exposure to radiation, exposure to radon gas that resulted from uranium natural breakdown in soil, rocks, and water which polluted air especially in close living areas, and exposure to other carcinogens in workplace such as asbestos, and arsenic (Mayo Clinic, 2019). E-cigarettes emit formaldehyde into the lungs at a higher level which is also considered a carcinogen (Salamanca et al., 2018). Burden of lung cancer can be

reduced by avoiding risk factors, screening, early diagnosis, and early treatment (WHO, 2019). Screening can lead to early diagnosis and early treatment which results in more effective and less expensive treatment with a better outcome (WHO, 2019). This will explain the importance of my study where many risk factors for lung cancer and barriers to health care are inherited and need to be analyzed and addressed to facilitate better government policies and raise awareness of the communities.

Diagnosis

Minimally invasive procedures are currently used for diagnosis and staging of lung cancer such as endoscopic/ endobronchial ultrasound needle aspiration which is introducing a scope through the trachea and bronchus under ultrasound guide and take a tissue biopsy through a needle linked to the scope (Folch et al., 2015). Transbronchial biopsy and transthoracic image guided core needle biopsy are other minimally invasive procedures that are used for diagnosis and staging of lung cancer in which a needle is passed through skin into the bronchus or into the thorax (chest) under computerized tomography (CT scan) and a tissue biopsy is aspirated through a needle (Folch et al.,2015). These minimally invasive techniques are quick, with lower cost, less complications, and can obtain enough tissue samples for pathologic and molecular diagnoses (Folch et al., 2015).

Genotyping is another advanced diagnostic method that looks for genetic mutations using PCR-based methods or gene rearrangements using screening immunohistochemistry (Folch et al., 2015; Pu et al., 2017). Genotyping besides helping in diagnosis, it facilitates selecting the proper type of treatment when cancer is correlated

with genetic factors (Folch et al., 2015; Roberti et al., 2019). Advanced bronchoscopy techniques such as electromagnetic navigation bronchoscopy is also used for diagnosis and staging as well (Folch et al., 2015). Circulating tumor DNA (ctDNA) in bloodstream, which is released by primary cancer cells, by metastasis, and by minimally residual disease (MRD) can be detected by specialized tests such as digital polymerase chain reaction (dPCR) and next generation sequencing (NGS) (Huang et al., 2019; Li & Liang, 2020). Detection of ctDNA is a valuable diagnostic noninvasive technique that is utilized for screening, diagnosis, prognosis, and evaluation of response to treatment (Huang et al., 2019; Li & Liang, 2020). dPCR and NGS techniques have capabilities of detecting stDNA mutations with high sensitivity and specificity (Li & Liang, 2020). These minimally invasive techniques and noninvasive techniques require advanced technology and multidisciplinary team of professionals which I discussed and recommended in my study to close the evidence-practice gap (Osarogiagbon, 2018). As such, they will eventually attribute to a higher survival and a better quality of life for patients diagnosed with lung cancer (Folch et al., 2015).

About 60% of cancer affects underserved and disadvantaged communities with a higher incidence rate and mortality rate among countries with low to middle income (Montagne et al., 2021). About 70% of cancer cases were diagnosed at late stages that required multimodality treatment (Montagne et al., 2021). As such my research study examined underserved communities who inherent poor lifestyle with other contributing factors such as lack of or low education levels, low income, and lack of access to healthcare. In this study I analyzed the underlying causes that made lung cancer stubborn

and responsible for 25% of cancer mortality rate (Cancer Treatment Centers of America, 2020). Furthermore, in my study I recommended the steps that can contribute to early diagnosis and treatment which will improve life expectancy and quality of life among the underserved communities.

Cancer Staging

Staging of cancer represents the anatomic extent of cancer in the human body which is a useful diagnostic and prognostic tool as well as an indicator for the effect of treatment (Amin et al., 2017). Cancer anatomic classification was first introduced by Pierre Denoix during 1940s-1950s and is expressed as Tumor, Lymph Node, and Metastasis (TNM) (Amin et al., 2017). Then, the American Joint Committee on Cancer (AJCC) published its first edition of cancer staging manual in 1977 and expanded the scope of the anatomic classification (TNM) to include physicians and registrars (Amin et al., 2017). The anatomic classification (TNM) expresses the extent of tumor (T) to involve lymph nodes (N) and metastasis (M) (Amin et al., 2017). Every four to seven years a new edition of AJCC updates cancer staging manual with the latest eighth edition in 2017 (Amin et al., 2017). The second edition of AJCC staging manual (1983) adopted the TNM staging system introduced by the Union for International Cancer Control (UICC) (Amin et al., 2017). The AJCC staging manual in the 1990s became a benchmark for Commission on cancer-accredited hospitals which required AJCC TNM approach for cancer reporting (Amin et al., 2017). The AJCC staging manual is considered the main reference at the population level for determining the initial diagnosis, prognosis, and appropriate treatment (Amin et al., 2017).

The TNM system further specifies the extent of cancer in each category by numbers as follows: For the tumor (T), TX when the tumor cannot be measured, T0 when the tumor cannot be detected, T(is), also called Tumor in situ (T in situ) when the tumor is within the boundaries of the organ or the tissues it originates from without metastasis, T1, T2, T T3, and T4 when the tumor increases in size with 1 is the smallest and 4 is the largest size of the tumor (MD Anderson Cancer Center, 2021).

Regional lymph nodes (N) (small bean-shaped structures) are part of the lymphatic system that distributes the immune system cells throughout the body through a lymphatic fluid; and cancer cells tend to spread into the nearby lymph nodes (MD Anderson Cancer Center, 2021). Lymph node metastasis is expressed in the TNM system as NX when cancer in the adjacent lymph nodes cannot be measured, N0 when cancer cannot be detected in the adjacent lymph nodes, N1, N2, and N3 reflect the number of lymph nodes that are affected by cancer (MD Anderson Cancer Center, 2021).

Distant metastasis (M) is expressed in the TNM system as, MX when metastasis cannot be measured, M0 when no metastasis is detected in any parts of the body, M1 when metastasis is detected in any parts of the body (MD Anderson Cancer Center, 2021).

However, the advances in cancer molecular biology made it possible to better determine the best treatment approach, prognosis, and treatment outcome through biologic factors which led some experts to question the validity of TNM staging system in clinical setting (Amin et al., 2017). Therefore, the eighth AJCC edition of cancer staging manual (2017) considered molecular factors with the anatomic-based

classification (TNM) to bridge the gap between population and individual approach and develop foundation for a conceptual framework for personalized cancer staging (Amin et al., 2017). To merge the anatomic and molecular concepts, the eighth edition of AJCC staging manual added the term *"prognostic stage group"* to the anatomic staging (Amin et al., 2017). The prognostic stage group represents nonanatomic prognostic factors and biomarkers (Amin et al., 2017; Gingras, 2018). Prognostic factors are factors that correlate with cancer prognosis as a predictive measure used for staging (Amin et al., 2017). To maintain the international agreement that was established since 1980s between the AJCC and UICC, the eighth edition of AJCC of cancer staging manual has UICC representatives on expert panels and on the editorial board so a global consensus will be developed on the new cancer staging (Amin et al., 2017). Staging of cancer is a collaborative effort that includes physicians, pathologists, radiologists, and other professionals such as scientists, and researchers who set the rules that can be applied on staging for all types of cancer at different anatomical sites of the human body (Amin et al., 2017).

The stage of lung cancer, as mentioned above, is a useful diagnostic and prognostic tool as well as an indicator for the effect of treatment (Amin et al., 2017). Therefore, I considered it as one of my independent variables that I studied and analyzed its effect on life expectancy of marginalized communities.

Contributing Factors to Staging of Lung Cancer

Cell Type

The origin of lung stem cells (mother cells) remains unclear despite it was widely studied (Raniszewska et al.,2021). That is because the epithelia (the lining layers) of some parts of the lung such as the trachea (windpipe) and bronchioles (tiny air passages inside the lungs) are relatively inactive with low proliferation (Raniszewska et al., 2021). Therefore, the origin of lung cancer stem cells (CSCs) is also unclear; however, the most accepted hypothesis states that they originate from the stem cells of their original normal specific tissues at the anatomical location in the lungs (Raniszewska et al., 2021). As such, squamous cell carcinoma originates from the basal cells of proximal airway such as trachea and bronchi and it expresses stem cell-like behavior; small cell lung cancer (SCLC) originates from bronchiolar exocrine cells and pulmonary neuroendocrine cells and they express the same stem cells behavior of these anatomical sites (Raniszewska et al., 2021).

Grade

The grading system describes the microscopic changes in the appearance of cancer cells and tissues compared to normal body cells and tissues (MD Anderson Cancer Center, 2021). The standard grading scale is from grade1 to grade 4 based on the level of changes from the normal body cells (MD Anderson Cancer Center, 2021).

Grade 1: The appearance of cancer cells and tissue almost looks like the healthy body cells and are called well-differentiated and is considered a low-grade tumor that is less aggressive and with a better prognosis (MD Anderson Cancer Center, 2021).

Grade 2: The cancer cells and tissue in this grade are to some extent look different from the normal cells and are called moderately differentiated. They are intermediate regarding their level of aggressiveness (MD Anderson Cancer Center, 2021).

Grade 3: The cancer cells and tissue are remarkably abnormal, and they are called poorly differentiated and high grade (MD Anderson Cancer Center, 2021).

Grade 4: The cancer cells and tissue look totally abnormal compared to the healthy body cells; they are called undifferentiated; they grow and metastasize faster (MD Anderson Cancer Center, 2021).

The grading system allows physicians to evaluate the extent of aggressiveness and prognosis of cancer which will facilitate developing a treatment plan accordingly (MD Anderson Cancer Center, 2021).

Biomarkers/Tumor Markers

Serum (the clear liquid part separated from coagulated blood) tumor markers' levels are associated with progress/ stage of cancer; therefore, they are widely used as indicators for prognosis and for the evaluation of response to chemotherapy and targeted therapy in advanced stage of NSCLC (Atwater & Massion, 2016; Chu et al., 2018; Ni et al., 2018; Zhang et al., 2020). There are several serum tumor markers with specific and general correlations to certain types of cancers such as carcinoembryonic antigen (CEA), cancer antigen 125 (CA 125), cytokeratin 19 fragment (CYFRA21-1), and squamous-cell carcinoma-related antigen (SCC-Ag) (Bossé & Amos, 2019; Zhang et al., 2020).

Zhang et al. (2020) studied the association between the dynamics of tumor markers and the effect of immunotherapy on advanced cases of NSCLC. The purpose of

this study was to investigate the possibility of using the dynamics of serum tumor markers as predicting factors for the prognosis of advanced NSCLC cases treated with immunotherapy, programmed cell daeath-1/programmed cell death ligand-1 (PD-1/PD-L1) inhibitors (Lozano et al., 2018; Zhang et al., 2020). The method used by the investigators was a longitudinal prospective method on 308 Chinese patients diagnosed with advanced NSCLC and were enrolled in PD-1/PD-L1 inhibitors treatment study at Chinese PLA general hospital in Beijing, China (Zhang et al., 2020). Baseline blood samples were collected before the treatment started and 6 weeks after; CT scan was done for all patients in the study to evaluate the effect of treatment, and serum tumor markers levels were measured with an electrochemical luminescence for SCC-Ag, and with a chemiluminescent microparticle immunoassay for serum CEA, CA 125, and CYFRA21-1 (Zhang et al., 2020).

The result of the study was a decrease of the tumor markers levels by 20% from the baseline after 6 weeks of immunotherapy treatment (immune checkpoint inhibitors [ICIs]) which is considered by the investigators a remarkable improvement (Zhang et al., 2020). The study analysis included the associations between serum tumor markers decline and objective response rate (ORR), progression-free survival (PFS), and overall survival (OS) (Zhang et al., 2020). To balance baseline covariates between different groups, optimization-based method was used (Zhang et al., 2020). The investigators concluded that the dynamic decline of serum tumor markers (CEA, CA 125, CYFRA21-1, and SCC-Ag) from the baseline could predict the efficacy of immunotherapy (PD-1/PD-L1 inhibitors) in advanced stage of NSCLC patients; furthermore, the decline of

associated serum biomarkers levels was associated with a better prognosis (Zhang et al., 2020).

Carcinogens/Carcinogenesis

The International Agency for Research on Cancer (IARC) has recently updated the list of carcinogenic agents to human to include more than 100 agents classified as group 1 (Monographs Volume 100, parts A-F) (Smith et al., 2016). The IARC through an international group of experts has identified ten key characteristics of which one or more are inherent by human carcinogens (Smith et al., 2016). These ten key characteristics are considered the baseline for identifying the mechanism of a carcinogen in humans and they enable the agent/ carcinogen to: Act as an electrophile (bond to human cells) directly or following metabolic activation; genotoxic; alter DNA repair or lead to genomic instability; induce epigenetic changes; induce cellular oxidative stress (imbalance between free radicals formation and ability of cells to clear them); contribute to chronic inflammation; immunosuppressive; alter receptor-mediated effects; lead to immortalization (genetic cell alterations to reproduce indefinitely); and modify cellular proliferation (replication), cellular death, or nutrient supply (Smith et al., 2016).

Smoking Status

Tobacco use is attributed to about 25% of all cancer deaths and tobacco smoke accounts for 87% of deaths due to lung cancer (Bakulski et al., 2019). Tobacco smoking is associated with adenocarcinoma, a type of NSCLC which is the leading cause of cancer death worldwide (Bakulski et al., 2019). Cigarette smoke causes biochemical alterations to the DNA of human cells (DNA methylation) through blood for newborns

from mothers who smoked during pregnancy and for adults due to personal smoking (Bakulski et al., 2019). Cigarette smoke exposure is associated with tissue specific epigenetic alterations or with across all tissues which underscore the value of blood-based methylation biomarkers that will facilitate evaluation of exposure effects in target tissues (Bakulski et al., 2019).

Age

Aging is considered an important risk factor for lung cancer due to accumulation of mutations in somatic cells during the individual's life span (Raniszewska et al., 2021). Although most of mutations do not affect cellular physiology or functions but, some mutated genes can interfere and modify cellular activities such as: Deregulation of cellular metabolism, resistance to cellular death, continuing cellular proliferation, interrupting growth suppressors, activate invasion and metastasis through angiogenesis and mutations (Raniszewska et al., 2021). Types of mutations include somatic point mutations, insertions, deletions, and rearrangements of genes; with NSCLC represents the highest rates of mutations including somatic molecular mutations in an individual with significant variations among populations worldwide regarding types of mutations and frequencies (Raniszewska et al., 2021). Aging is associated with EGFR mutation in lung cancer, and recent molecular advances in NSCLC has led to the development of new targeted therapies (Raniszewska et al., 2021).

Types of Treatment for Lung Cancer

The type of treatment for lung cancer will be assessed and decided based on multiple criteria such as cancer histology, molecular intrinsic factors, stage of lung

cancer, and the condition of the patient (Montagne et al., 2021). As such, treatment could be one type of treatment, a combination of treatment, or multimodality therapy (MMT); as surgical removal of an early localized lung cancer (stage-I) which can be curative, a combination of chemotherapy drugs, or MMT such as chemotherapy and radiotherapy in late stages of lung cancer (stage -IV) which is most likely a palliative treatment that might control further metastasis but it will not cure a lung cancer that was already metastasized at different parts of the body such as bones and brain (Lu et al., 2019; McLouth et al., 2021; Montagne et al., 2021;). Another type of treatment called targeted therapy which is based on diagnostic study of genes (genotyping) that identifies some deletions (Lu et al., 2019; Roberti et al., 2019). A study of trend in cancer treatment during the last four decades (1973-2015) which was done by Lu et al. (2019) has concluded that the rate of surgery was 25%, chemotherapy reflected increase in trend, while radiotherapy showed a decrease in trend (Lu et al., 2019).

Surgery

The innovation of advanced diagnostic methods and techniques for the diagnosis of lung cancer and the innovation of treatment modalities have led to the evolution of lung cancer surgery from invasive thoracic approach to minimally invasive (Montagne et al., 2021). The indication for surgery will be decided based on the stage of lung cancer and the surgical approach will be either resection guided by imaging such as CT scan or ultrasound, or minimally invasive surgical approach, or hybrid approach (Montagne et al., 2021). These innovative surgical approaches have contributed to fast recovery, less complications, and a longer-survival rate (Montagne et al., 2021). Surgery can play a

curative role in early stages of lung cancer when a localized primary tumor can be removed surgically in its entirety; also, surgery can play a palliative role in late stages of lung cancer when metastasis or recurrence of cancer develop (Montagne et al., 2021). NSCLC which accounts for 85% of primary lung cancer (Mayo Clinic, 2019) became more readily early diagnosed due to screening and more responsive to systemic treatment that encompass more advanced medications (Montagne et al., 2021).

Montagne et al. (2021) argued that surgery contributes to diagnosis in early preinvasive stage of lung cancer, to surgical excision of preinvasive tumor, locally, or locally advanced tumor, to palliative surgery that can relieves symptoms such as dyspnea and pain in late stages of lung cancer. The role of surgery is no longer limited to resection of tumor as surgeons became more involved as part of the multidisciplinary team who follow up lung cancer patients from early preinvasive stage before surgery, then surgery, post-operative follow-up, through palliative surgery in case of recurrence or late stages of lung cancer (Montagne et al., 2021). Such evolution of surgical approach as part of multidisciplinary team is a crucial part in my research study as it facilitates early diagnosis and early treatment which contribute to cure from lung cancer that leads to a longer life expectancy and a better quality of life among cancer patients. The involvement of surgeons in the follow-up facilitates detection of residual of tumor after systemic treatment and identify recurrence of disease after remission (Montagne et al., 2021).

Professor Henrik Kehlet, colorectal surgeon was the first to define in 1997 the concept of Enhanced Recovery After Surgery (ERAS) which he constructed with six requirements: Information about the patient before surgery and educating the patient

about what to expect with surgery, minimizing the patient's stress, provide medications to relieve the patient's pain after surgery, exercise for the patient, enteral nutrition, and growth factors (Montagne et al., 2021). The purpose of ERAS is to expedite recovery after surgery by avoiding post-operative complications and to provide the patient with good experience about surgery (Montagne et al., 2021). Now ERAS became popular among many medical specialties. ERAS protocols have been adopted by thoracic surgery to include preoperative programs for educating patients, maintain nutrition, control tobacco and alcohol addiction, management of the respiratory and cardiac associated diseases and physical therapy (Montagne et al., 2021). The results of ERAS are less fasting time before surgery, avoid medications that enhance sedation such as opioid, minimally invasive surgical approach, and facilitates quick mobility of the patient after surgery with less post-operative complications (Montagne et al., 2021).

Chemotherapy

Chemotherapy is a general medical term used to represent drugs that are used to treat cancer by different mechanisms that will lead to either killing of cancer cells or preventing cancer cell division (NCI, n.d.). Chemotherapy can be administered by different routes such as orally, injection, infusion, or on skin based on type and stage of cancer (NCI, n.d.). Chemotherapy can be prescribed as one drug or more than one drug; and can be given alone or in combination with other treatment modalities such as radiation, surgery, or immunotherapy (NCI, n.d.). Despite recent discoveries and implementations of target therapy and immunotherapy, chemotherapy is still considered

the treatment of choice for most cases with advanced NSCLC (Adamowicz et al., 2020; Baxevanos & Mountzios, 2018).

Among many chemotherapeutic agents, platinum-containing drugs are considered the regimens of choice for the treatment of NSCLC (Baxevanos & Mountzios, 2018). There are many novel chemotherapy protocols such as platinum plus pemetrexed combined with bevacizumab for the treatment of advanced non squamous NSCLC (Baxevanos & Mountzios, 2018). For all histological types of NSCLC, a combination of carboplatin and nanoparticle-albumin bound paclitaxel is used (Baxevanos & Mountzios, 2018). Based on the type and stage of lung cancer, chemotherapy can be curative or palliative (Baxevanos & Mountzios, 2018). However, recent clinical data revealed that the efficacy of chemotherapy alone has reached a therapeutic plateau regarding survival rate and the current clinical research recommends chemotherapy to be prescribed with other treatment modalities such as radiotherapy, surgery, or immunotherapy (Baxevanos & Mountzios, 2018).

Baxevanos and Mountzios (2018) explained that despite the application of novel chemotherapeutic agents, there is modest increase of survival. The median overall survival (OS) with the therapeutic plateau ranges between 12 and 14 months according to recent clinical trials (Baxevanos & Mountzios, 2018). Clinical research is promising on increasing the OS with the application of chemotherapy combined with targeted agents and immunotherapy (Baxevanos & Mountzios, 2018).

Radiotherapy

Radiotherapy or radiation therapy is another treatment modality that has been used for the treatment of lung cancer specially NSCLC (American Cancer Society, 2021). Radiotherapy uses high-energy rays to kill cancer cells and depending on the stage of lung cancer and other factors, radiation can be given alone, with surgery, with chemotherapy, or with other types for treatment such as targeted therapy, or immunotherapy (American Cancer Society, 2021). Radiotherapy can be prescribed alone or with chemotherapy when surgery is not possible due to the size or location of the tumor, or the patient is too sick to stand surgery, or the patient does not want surgery (American Cancer Society, 2021). Radiotherapy can be prescribed before surgery to shrink the tumor or after surgery to eradicate the remaining parts of cancer that could not be removed surgically and in both scenarios, radiation can be given alone or with chemotherapy (American Cancer Society, 2021). Radiotherapy can also be prescribed to treat distant metastasis in bones and brain; as well as a palliative treatment in advanced cases of NSCLC to relieve symptoms such as pain or cough (American Cancer Society, 2021).

Types of Radiotherapy Prescribed for NSCLC

External Beam Radiation Therapy (EBRT) is a direct radiation to the lungs, and it is the most used type of radiotherapy for NSCLC and its metastases (American Cancer Society, 2021). EBRT looks like taking an x-ray but with a higher dose of radiation and it is prescribed commonly as five times per week for 5 to 7 weeks; but there are some variations based on the EBRT type and the purpose for prescribing it (American Cancer

Society, 2021). Each EBRT session is given in a few minutes and causes no pain (American Cancer Society, 2021).

EBRT is a new innovative technology that provides more precise treatment as it focuses to the lungs with minimal exposure to the surrounding tissues and it is called Stereotactic body radiation therapy (SBRT) or stereotactic ablative radiotherapy (SABR) which is mostly prescribed for early stage of lung cancer when surgery cannot be done due to patient's health condition or refusal of surgery by the patient (American Cancer Society, 2021). The advantage of SBRT is the duration of treatment which is limited to 1-5 high-dose radiation sessions with beams directed toward the tumor from different directions (American Cancer Society, 2021). The patient will be fitted in a special designed body frame during each SBRT session to restrict lung's movement while the patient is breathing (American Cancer Society, 2021).

There are other techniques used for external body radiation to treat lung cancer specially NSCLC such as three-dimensional conformal radiation therapy (3D-CRT) which utilizes special computers to locate the tumor and direct the radiation beams from several angles preserving the surrounding healthy tissue from the effect of radiation (American Cancer Society, 2021). Other external beam radiation techniques are also available with variable advantages (American Cancer Society, 2021).

Brachytherapy (internal radiation therapy) is the other type of radiotherapy that is used to treat NSCLC which is an internal radiation therapy that facilitates minimizing the tumor size to relieve obstruction of the airway (American Cancer Society, 2021). Brachytherapy is a form of radioactive substance delivered inside or nearby the tumor

through bronchoscopy or during surgery as small pellets which affects the tumor without causing damage to the surrounding tissue and they are either removed after short period or left inside and they weaken after few weeks (American Cancer Society, 2021).

Radiotherapy has some side effects that vary according to the site of the body that was exposed to radiation and they include nausea, vomiting, loss of appetite, weight loss, hair loss, skin irritation, and fatigue (American Cancer Society, 2021). When treatment includes a combination of radiotherapy and chemotherapy, side effects become more intense; however, they usually fade away after completion of treatment (American Cancer Society, 2021). Other side effects when radiation is directed to the chest include, cough, breathing problems, swallowing difficulties, and sore throat; while radiation therapy to the brain might lead to headache, and memory loss (American Cancer Society, 2021).

Radiofrequency ablation (RFA) for NSCLC is used for the treatment of a small size lung cancer located in the outer part of the lung by inserting a needle-like probe through the skin guided by CT scan aimed at the tumor; and a high-energy radio wave (electric current) is passed through the probe which leads to damage of cancer cells by heat generated by the electric current (American Cancer Society, 2021).

Targeted Therapy for NSCLC

As research reveals more knowledge about changes that take place inside the cells of the NSCLC and factors that accelerate their growth, more drugs are invented to target the growth of cancer cells (American Cancer Society, 2021). Targeted drugs have different mechanisms of action that differ from chemotherapy, and they don’t work with

one mechanism of action but rather each group of targeted drugs have a mechanism of action that is different from another group of targeted drugs, and they are used for advanced stages of NSCLC (American Cancer Society, 2021). Some examples of groups of targeted drugs:

Drugs that Target Growth of Blood Vessels of the Tumor. Tumors develop their own blood vessels that supply them with nutrients, a process known as angiogenesis; therefore, these drugs are angiogenesis inhibitors which block the development tumor blood vessels; bevacizumab (Avastin) and ramucirumab (Cyramza) are examples of this group of angiogenesis inhibitors, and they are prescribed for advanced cases of NSCLC with chemotherapy or alone (American Cancer Society, 2021). Some common side effects of these drugs are bleeding, low white blood cell counts which put patients at a risk of infection, high blood pressure, fatigue, and mouth sores (American Cancer Society, 2021).

Drugs that Target Specific Genes of Cancer Cells. Genes targeted by this group of drugs are, KRAS, EGFR, ALK, ROS1, BRAF, MET, and NTRK, which when they exhibit changes from normal, a process called mutation, they promote abnormal cell growth, cell division, and cell spreading (American Cancer Society, 2021; Roberti et al., 2019). Examples of drugs that target specific genes of cancer cells are, sotorasib (Lumakras), osimertinib (Tagrisso), crizotinib (Xalkori), ceritinib (Zykadia), dabrafenib (Tafinlar), capmatinib (Tabrecta), and larotrectinib (Vitrakvi) respectively (American Cancer Society, 2021). Variable side effects result from these targeted drugs such as, nausea, vomiting, diarrhea, dizziness, skin rash, swelling in hands and feet, and changes in vision (American Cancer Society, 2021).

Immunotherapy

Drugs in this category boost the immune system of cancer patients which will facilitate recognition and destruction of cancer cells by the immune system (American Cancer Society, 2021). The immune system recognizes the body's normal cells from others by proteins on the immune cells that function as a checkpoint (American Cancer Society, 2021). For the immune system to initiate an immune response, the checkpoint protein will turn on and it will turn off to avoid attacking the normal body's cells; a function that cancer cells utilize sometimes to avoid recognition and attacks by the immune cells (American Cancer Society, 2021). Immunotherapy drugs are called checkpoint inhibitors because they target immune system checkpoints and therefore, they are prescribed for cases of NSCLC (American Cancer Society, 2021; Davis et al., 2020). The examples of these drugs are nivolumab (Opdivo) and atezolizumab (Tecentriq) (American Cancer Society, 2021). Such mechanism slows down cancer cells growth and it will eventually lead to shrinkage of the tumor which can be removed surgically (American Cancer Society, 2021).

Immunotherapy drugs can be prescribed alone, with chemotherapy, radiotherapy, or with all together (American Cancer Society, 2021; Ko et al., 2018; Lazzari et al., 2018). Side effects of immunotherapy drugs include but not limited to fatigue, skin rash, cough, nausea, joint pain; and other serious side effects which occur less frequent such as infusion reaction/ allergic reaction that is manifested by fever, chills, skin rash, wheezes; and autoimmune reactions that destroy the immune system checkpoints and begin

attacking the body's won cells leading to a life-threatening conditions in the lungs, liver, kidney, and other parts of the body (American Cancer Society, 2021).

Treatment of Choice for NSCLC Based on Stage of Lung Cancer

Treatment option for NSCLC is decided based on the stage of lung cancer among other factors such as the trait of cancer, and general health of the patient (American Cancer Society, 2021).

Treatment of Occult Lung Cancer

This is an early stage when cancer cannot be detected by bronchoscopy or in imaging and cancer cells were only detected in the sputum (American Cancer Society, 2021). Another diagnostic approach for occult lung cancer is bronchial lavage (washings) in which a measured amount of fluid is introduce through a bronchoscope into the broncho-alveoli (inside the lung) and removed to be tested for cancer cells (Wu et al., 2021). Occult lung cancer is not common, but the cancer-specific survival is poor (Wu et al., 2021). However, timely surgical excision and radiotherapy could improve the outcome and survival; but further research is warranted to confirm these findings (Wu et al., 2021).

Treatment of Stage Zero (in situ) NSCLC

In this stage, cancer is called cancer in situ and is confined to superficial layers without invasion in the lung deep layers (American Cancer Society, 2021). Surgical removal is the treatment of choice if the there is no contraindications for surgery, and at this stage cancer is curable (American Cancer Society, 2021). No indication for chemotherapy or radiotherapy in this stage unless there is contraindication for surgery

(American Cancer Society, 2021). As an alternative for surgery, when surgery cannot be done, photodynamic therapy (PDT), brachytherapy, or laser can replace surgery and a cure is expected in stage zero with either of these treatment options (American Cancer Society, 2021).

Treatment of Stage I NSCLC

The type of treatment for stage I NSCLC will depend on the size of the tumor that will vary from surgery only for a small size tumor to surgery followed by adjuvant (additional) chemotherapy if the tumor size is big to avoid recurrence (American Cancer Society, 2021). Also, surgery can be followed by radiotherapy that can be applied to the surgical excision margin (American Cancer Society, 2021). Surgery can be for removing the tumor only, or wedge resection, or segmentectomy, or removing the entire lobe of the lung (lobectomy) (American Cancer Society, 2021). In case there is a contraindication for surgery, radiation will be the second treatment option in the form of stereotactic body radiation therapy (SBRT) or other types of radiotherapy (American Cancer Society, 2021). Radiofrequency ablation (RFA) can be an option if the tumor is small and located in a superficial part of the lung (American Cancer Society, 2021).

Treatment of Stage II NSCLC

The treatment of choice for stage II NSCLC will be surgical resection of cancer and any suspicious lymph nodes that might include metastasis followed by either adjuvant chemotherapy or radiotherapy to treat any remaining cancer cells on the surgical margin or in the lymph nodes and to avoid recurrence of cancer (American Cancer Society, 2021). In patients whose cancer cells include mutations in EGFR gene, adjuvant

chemotherapy with targeted drug therapy such as Osimertinib can be a treatment option as well (American Cancer Society, 2021). In case surgery cannot be done for any medical reason or otherwise, radiotherapy will be the second line of treatment (American Cancer Society, 2021).

Treatment of Stage IIIA NSCLC

Treatment protocol for stage IIIA of NSCLC will be decided based on the size and location of the tumor inside the lung, and the general health condition of the patient (American Cancer Society, 2021). A combination of chemotherapy, radiotherapy, and surgery is often required but the sequence may vary based on the size of tumor, location, and the general condition of the patient (American Cancer Society, 2021). Though, a single treatment modality such as surgery may also be considered if the size of the tumor is not too large or too small and in an accessible location inside the lungs besides, the health condition of the patient should be favorable for surgery (American Cancer Society, 2021). Adjuvant targeted therapy such as Osimertinib can be considered in patients whose cancer cells have mutations in the EGFR gene (American Cancer Society, 2021). Lung cancer patients with a compromised health condition that will not allow for surgery can be given a combination of chemotherapy and radiotherapy (chemoradiation) (American Cancer Society, 2021). When chemotherapy, radiotherapy, and surgery did not give good results, immunotherapy such as pembrolizumab (Keytruda) or cemiplimab (Libtayo) can be considered (American Cancer Society, 2021). Immunotherapy can also be considered as a first line of treatment for NSCLC if the treatment team decided so; therefore, the team should include medical oncologist, radiation oncologist, and thoracic

surgeon to discuss the best treatment option (American Cancer Society, 2021). This addresses the importance of multidisciplinary team as part of the management team for lung cancer patients which is mostly not provided and its absence is considered part of the evidence-practice gap (Rankin et al., 2020). As such, I am considering this in my research study as part of suggested solutions to improve treatment, outcome, and life expectancy.

Treatment of Stage IIIB NSCLC

This stage of lung cancer involves metastases to the adjacent lymph nodes within the chest and the neck as well as metastases to other structures inside the chest; therefore, surgery cannot remove cancer in this stage without additional chemotherapy and radiotherapy (American Cancer Society, 2021). When cancer in this stage responds to chemoradiation, immunotherapy such as pembrolizumab (Keytruda) or cemiplimab (Libtayo) can be prescribed for a year afterward to keep cancer under control (American Cancer Society, 2021). Also, immunotherapy can be considered as a first line of treatment (American Cancer Society, 2021). In this late stage of lung cancer, clinical trials can be considered for some patients who have poor response to conventional therapy (American Cancer Society, 2021).

Treatment of Stage IVA and IVB NSCLC

Lung cancer at this stage has widely spread outside the lung to one or more areas of the body which makes treatment limited to palliative care to relieve symptoms such as pain, and shortness of breath to improve the quality of life; therefore, the chance for cure is very slim depending on the extent of metastases (American Cancer Society, 2021).

Treatment options are decided based on areas affected with metastases, mutations (changes) in specific genes of cancer cells, and the general health of the patient (American Cancer Society, 2021). Treatment can include one or a combination of chemotherapy, radiotherapy, surgery, immunotherapy, targeted therapy, and photodynamic therapy (PDT) (American Cancer Society, 2021). As such, cancer patients at this stage of cancer ought to understand the goal of treatment before they decide to start their treatment.

In stage IVA of NSCLC, cancer has metastasis to one site such as the brain; besides the primary cancer in the lung which may require either surgery, radiotherapy, or both to the brain based on the patient's overall health condition (American Cancer Society, 2021). Treatment of the primary cancer in the lung may include surgery, chemotherapy, radiotherapy, or a combination of two or more depending on the size and site of the tumor in the lung and the extent of metastasis to lymph nodes (American Cancer Society, 2021).

In stage IVB of NSCLC, cancer has metastases almost all over the body and the goal for treatment will be limited to palliative care (American Cancer Society, 2021). However, before initiating any treatment protocol cancer cells should be tested for gene mutations, such as EGFR, ROS1, ALK, KRAS, MET, BRAF, RET, and NTRK (American Cancer Society, 2021). Targeted therapy should be the treatment of choice if any of these genes was found mutated in the cancer cells and a specific gene inhibitor drug should be used such as EGFR inhibitor used for EGFR gene mutation, KRAS inhibitor as sotorasib (Lumakras) used for KRAS gene mutation (American Cancer

Society, 2021). Immunotherapy such as pembrolizumab (Keytruda) or other immunotherapy drugs could be a treatment of choice if cancer cells were found to have a higher level of PD-L1protein; and it can be used alone or combined with chemotherapy depending on the type of immunotherapy drug and the patient's tolerance (American Cancer Society, 2021).

Palliative Procedures for NSCLC

Palliative care includes palliative treatment which I already mentioned and palliative procedures that aimed to relieve cancer patients from complications of cancer such as pleural effusion (buildup of fluids around the lungs) which limits lungs expansion during inspiration and causes shortness of breath (American Cancer Society, 2021). A procedure called thoracocentesis is applied by inserting a needle inside the lower part of the chest under ultrasound guidance to drain the fluid around the lungs (American Cancer Society, 2021). Other procedures are available for the same purpose such as catheter placement which is a thin flexible tube inserted into the lower part of the chest through a cut in the skin and connected to a special botte outside the body to facilitate draining the fluid with the aid of gravity (American Cancer Society, 2021). Lung cancer can also lead to a fluid buildup around the heart called pericardial effusion which limits the heart expansion to accommodate incoming blood from different body organs and leads to a compromise in the function of the heart (American Cancer Society, 2021). A procedure called pericardiocentesis is used to drain the fluid around the heart using a needle inserted inside the chest under echocardiogram guidance to enter the pericardium (a sac around the heart where the fluid is collected) and drain the fluid (American Cancer Society,

2021). These are just a few examples of many complications caused by lung cancer and the procedures that are used to relieve symptoms and improve the quality of life for cancer patients (American Cancer Society, 2021).

Contributing Factors to Survival/Life Expectancy

Screening

Lung cancer when diagnosed during early stages is mostly curable and 56% of cases diagnosed with a localized cancer will be at a 5-year survival rate, while about 70% of patients diagnosed with advanced lung cancer, only 15% of them will be at a 5-year-survival rate (Azubuike et al., 2020). Therefore, screening is an important tool that will facilitate early detection and diagnosis of lung cancer. Preventive services task force lung cancer screening guidelines was established by the United States in 2013 for using low dose computed tomography (LDCT) (Azubuike et al., 2020). LDCT of the chest is considered a primary method for lung cancer screening and a secondary prevention method for early detection of lung cancer to ameliorate the outcome (Azubuike et al., 2020; Raghavan et al., 2020). The US Preventive Services Task Force (USPTF) considered the benefits of using LDCT as a screening measure for lung cancer outweighed the potential risk of low-dose radiation exposure (Azubuike et al., 2020).

Hence, LDCT for lung cancer screening is the standard on care in the US (Azubuike et al., 2020). The lung cancer screening model is endorsed by the American College of Radiologists and it includes: Providing educational materials for patients at the physicians' offices about the importance of lung cancer screening in the form of flyers, pamphlets, and posters; raise the awareness among healthcare providers, their staff, and

the communities about the importance of lung cancer screening and the current guidelines; train medical staff on how to identify patients who are at a high risk for lung cancer and how they can flag such patients by undergoing scripted phone calls and from electronic medical records (EMR) (Azubuike et al., 2020).

A quality improvement project study was performed with the objective to increase providers compliance with the 2013 US Preventive Services Task Force (USPTF) lung cancer screening guidelines which considered LDCT as a screening measure for lung cancer (Azubuike et al., 2020). The study was conducted in Genesee County, Michigan at mid-Michigan Family Medicine Clinic. The justification of conducting this study in Genesee County, Michigan was based on statistics background that was done in 2019 which revealed that Genesee County had 21% adults' cigarettes smokers compared to 20% statewide and 14% nationwide cigarette smokers (Azubuike et al., 2020). This was reflected further in lung cancer incidence rate between 2012 and 2016 as 74.4 cases per 100,000 individuals in Genesee County compared to 64.0 cases in Michigan and 59.2 cases in the US per 100,000 individuals (Azubuike et al., 2020). The mid-Michigan Family Medicine Clinic uses the Epic EMR system and accepts all patients with and without health insurance including low-income patients (Azubuike et al., 2020). The age group of participants were between 55- and 80-year-old with a history of a 30-pack a year smoking and currently smoking or had quit smoking in the last 15 years (Azubuike et al., 2020). The pre-post design was used to evaluate the number of LDCT scans ordered during pre-intervention period compared with post-intervention period (Azubuike et al., 2020).

The goal of the study was to identify cases in their early stages of lung cancer when treatment is curable by implementing public health methods to screen high-risk individuals for lung cancer (Azubuike et al., 2020). The study was completed in 3 months during 2020 and the results revealed that the mortality rate was much lower in lung cancer cases who underwent LDCT screening compared with those cases of lung cancer who did not (Azubuike et al., 2020). The quality improvement project study analysis with Fisher exact test concluded there were statistically significant increase in preintervention LDCT number of scans ordered (n=0) compared to postintervention (n=8) periods (p=.0043) (Azubuike et al., 2020). The limitations of this study were a small sample size and a short period of study; therefore, the researchers recommend future studies to include a larger population and to be conducted for a longer period (Azubuike et al., 2020). Furthermore, the researchers projected that a systematic implementation of their protocol could contribute to more compliance of providers in ordering LDCT (Azubuike et al., 2020).

Stage of Lung Cancer

Assessment of survival of cancer patients at the population level is a major contributor to the decoration of healthcare policy for lung cancer (Mar et al., 2020). TNM stage cancer classification system is a standardized approach in lung cancer survival analyses that will evaluate the impact of new diagnostic and therapeutic advances on prognosis of lung cancer (Mar et al., 2020).

A Mar et al. (2020) evaluated the survival of lung cancer patients who were treated in Basque Health Service using TNM stage in 4-year periods (2003–2006, 2007–

2010 and 2011–2014) and compare the results with survival of lung cancer patients in an equivalent sample of the general population. The researchers used in this study a retrospective observational design on a cohort of 11,635 patients retrieved from Euskadi hospital cancer registry (Mar et al., 2020). The variables included in the data were: TNM stage, age, sex, history, diagnosis date, vital status, and date of death (Mar et al., 2020). The models used in the study were relative survival and Cox and parametric regression models to evaluate changes in survival (Mar et al., 2020). The results of the study revealed that lung cancer 5-year survival probability decreased with increase in stage, for lung cancer patients with stage I from 50%-65% and for lung cancer patients with stage IV from 2-3% (Mar et al., 2020). The researchers also noticed an improvement in survival among patients diagnosed during the periods from 2003-2006 and 2011-2014 as the risk of death in 2003-2006 was 1.66, 1.51, 1.21, and 1.10 for stages I, II, III, and IV respectively (hazard zero); an increase from 11.0% to 17.8% in 5-year relative survival during 2011-2014; and the years of life lost during 2003-2006 and 2011-2014 decreased significantly (6.16 in stage I and 16.21 in stage IV) (Mar et al., 2020).

The researchers concluded that the increase in survival from lung cancer in the Basque Country was statistically significant from 2003 to 2014 among all stages, however, the survival in average was 20% low compared to the general population (Mar et al., 2020). The study also found the years of life lost varies between 8.85 and 17.87 years based on the stage at diagnosis compared to the general population (Mar et al., 2020). The researchers recommended more research on strategies that can improve the survival of lung cancer patients (Mar et al., 2020).

Treatment

Costa et al. (2019) performed a study with the objective to evaluate the possibility of using cancer registry database as a platform to monitor the effect of treatment (Costa et al., 2019). The study sample was 115 patients, and the study was performed during the period from November 1st, 2015, to July 31st, 2016 (Costa et al., 2019). The method used in the study was observational inception cohort where cases of NSCLC were registered to start treatment with immunotherapy drug called nivolumab were monitored retrospectively to assess the characteristics of cancer and the effect of prior treatments (Costa et al., 2019). Nivolumab effect was monitored prospectively, and the outcomes were classified according to the patient's medical record (Costa et al., 2019).

The principal outcome measure used to evaluate treatment effectiveness was, overall survival (OS); secondary outcomes measures were progression free survival (PFS) to assess the effectiveness, and occurrence of adverse drug reaction (ADRs) to assess safety (SPSS version 24 was used to analyze data) (Costa et al., 2019). The study results revealed that 115 patients received nivolumab for NSCLC with most cases diagnosed with non-squamous type (n=107). The median OS was 11.4 months (Cl 95%: 11.1-11.7), with one year survival of 44%. Median PFS was 5.4 months (Cl 95%: 2.8-7.9). Treatment was discontinued in 82 patients due to cancer progression most of the time. 38 patients with ADRs and discontinuation of nivolumab from 21 patients with ADRs (Costa et al., 2019). The authors concluded that cancer registry is a powerful data source for monitoring cancer treatment but that requires education, training, incentives to encourage collecting and reporting data on regular basis (Costa et al., 2019).

Access to Healthcare

The US Preventive Services Task Force (USPSTF) joined with other groups of experts have recommended a group of clinical preventive services for people at average risk for some diseases that include lung cancer (Carey et al., 2020). These services are less likely used by rural residents, people of lower socioeconomic status, and some racial/ ethnic minorities (Carey et al., 2020). The NIH conducted a workshop on June 19-20, 2019, on pathways to prevention and achieving health equity in preventive services.

The workshop evaluated the disparities in using the recommended clinical preventive services that resulted in late diagnosis, early and higher mortality that disproportionately affect racial/ ethnic minorities, rural residents, and poor communities (Annangi et al., 2019; Carey et al., 2020). For example, death due to cancer among black Americans at 183.6 compared to 160.7 among white Americans per 100,000 population (Carey et al., 2020). People with health insurance and live in low socioeconomic communities are 30% less likely to receive recommended colonoscopy for colon cancer screening (Carey et al., 2020). This represents a gap between knowledge evidence that experts recommend and the reality of dysfunctional society that requires research to reduce such disparities in preventive services. The gap between knowledge and evidence base urges the need for research to fill this gap (Carey et al., 2020). Therefore, the NIH workshop on June 19-20, 2019, identified research gaps and recommended the need for enhancing research and methods to reduce disparities with collaborations of clinicians, health systems, and advocacy communities to address the gaps in our knowledge (Carey et al., 2020).

The NIH workshop recommendations were categorized around 5 key question topics, barriers to preventive services attributed to providers, barriers to populations adversely affected by disparities, patient-provider interventions, health information technology intervention, health system interventions (Carey et al., 2020). The NIH workshop panel also pointed out three crossing themes: Community engagement and systems approaches, integration of services and new delivery models, and innovative methods (Carey et al., 2020). These themes are fundamental to be considered to facilitate reaching the goal of creating evidence for decision making by providers, health systems, and the public health community (Carey et al., 2020).

The NIH workshop concluded that simple interventions were mostly unsuccessful in the US to correct disparities in preventable health conditions though some progress has been made (Carey et al., 2020). Therefore, the workshop proceedings and systematic evidence review endorse the recognition that any future improvements in disparities require engagement of stakeholders such as payers, public health system, administrators, community-based organizations, and the public to work collaboratively together (Carey et al., 2020). In this regard, my research study has addressed some of the NIH workshop topics such as evaluation of disparities in using the recommended clinical preventive services that resulted in late diagnosis, early and higher mortality that disproportionately affect racial/ ethnic minorities, rural residents, and poor communities that were related to lung cancer patients in my study (Carey et al., 2020). Also, my study analyzed the contributing factors to disparities, evidence-practice gap, and suggested pathways to

achieve health equity in preventive services, early diagnosis, and early treatment that will eventually lead to a better outcome represented in quality of life and a longer survival.

Geographical Location

VoPham et al. (2022) studied the relationship between the incidence of hepatocellular carcinoma [HCC] (primary liver cancer) in the United States and the emissions of dioxins and dioxin-like compounds. The purpose of this study was to investigate the association between ambient dioxin air emissions from industrial sources at the county level and the risk of HCC in the US (VoPham et al., 2022). A dataset of 90,359 cases of HCC diagnosed during the period (2000-2016) was obtained from SEER database. SEER used geographic information system (GIS) to access county of residence at diagnosis which is linked to a nationwide spatial database of historical dioxin-emitting facilities from1987 to 2007 (VoPham et al., 2022). The statistics test used was Poisson regression with robust variance estimation to calculate incidence rate ratio (IRRs) and confidence intervals (95% CI) to evaluate the association between HCC incidence rate and dioxin emissions at the county level; adjusting for age at diagnosis, sex, race/ ethnicity, year of diagnosis, lifestyle factors, health conditions, and socioeconomic status (VoPham et al., 2022).

The study results revealed in analysis by facility type positive associations between dioxin emissions at the county level from coal-fired power plants and the risk for HCC (adjusted IRR 1.09, 95% CI 1.01-1.17) as well as positive associations between industrial boilers, sewage sludge incinerators and the risk for HCC; but not consistent across both exposure metrics; however, the study results found no association between

dioxin emissions and the risk for HCC based on the number of dioxin-emitting facilities within the county or average annual emissions at the county level (VoPham et al., 2022).

Kelly et al. (2016) conducted a systematic review with an objective to examine the association between differences in travel time or travel distance to healthcare services and the health outcomes to patients. Studies that met the inclusion criteria and reviewed by researchers were 108 studies; the results included 77% of the reviewed studies revealed evidence of association between distance and health outcome whereby patients who lived farther away from healthcare facilities had worse health outcome such as lower survival rates, longer length of hospitalization, and no show for follow-up appointments compared to those who lived nearby healthcare facilities (Kelly et al., 2016). Six of the reviewed studies showed reverse outcome (distance bias effect) whereby people who lived farther away from healthcare facilities had a better health outcome, and the remaining 19 of the reviewed studies showed no relationship between the distance to healthcare facilities and health outcome (Kelly et al., 2016).

The researchers found a large variation in the studied data on geographical locations of patients and healthcare facilities and the methods used to calculate the travel distance and time were not consistent among all the reviewed studies (Kelly et al., 2016). However, the researchers concluded that there was a relationship between travel distance to healthcare facilities and health outcomes and they recommended consideration of this issue within healthcare services discussions (Kelly et al., 2016).

Comorbidities

Comorbidities that are highly prevalent in lung cancer patients are chronic obstructive pulmonary disease (COPD), cardiovascular disease, hypertension, diabetes mellitus (DM), and other types of cancers; while the most common pulmonary comorbidities with lung cancer are bronchial asthma, COPD, and tuberculosis (TB) (Dima et al., 2018).

A study aimed to investigate the effects of comorbidity on lung cancer diagnosis and survival in lung cancer patients in Taiwan was done by Dima et al. (2018) using a nationwide population-based study design. The investigators used a cohort of 101,776 lung cancer patients among them 44,770 with comorbidity and 57,006 without comorbidity during the period from 1995 to 2010 (Dima et al., 2018). The data were obtained from the National Health Insurance Research Database (NHIRD) that were retrieved from the National Health Insurance (NHI) program. The Kaplan-Meier was used in the analyses to compare overall survival between lung cancer patients with comorbidity (chronic bronchitis and hypertension in this study) and without comorbidity (Dima et al., 2018).

The study results revealed that lung cancer patients with comorbidity had a higher overall survival compared to those without comorbidity (Dima et al., 2018). By the end of the study in 2010 the investigators found that lung cancer patients with comorbidity had 14.9% clinical visits while their counterpart without comorbidity had 9.31% clinical visits reflecting that cancer patients with comorbidity had more physicians' visits and that was associated with early diagnosis of lung cancer (Dima et al., 2018). Patients with pre-

existing medical conditions, comorbidity, tend to seek medical care and have more frequent physician's visits which facilitate early diagnosis of lung cancer and eventually they start treatment earlier while patients without comorbidity have less clinic visits and early stages of lung cancer have no symptoms; therefore they were diagnosed in late stages of lung cancer when treatment options are limited due to advanced stages of lung cancer (Dima et al., 2018).

This justifies why lung cancer patients with comorbidity have a higher overall survival compared to those without comorbidity. However, every comorbidity has unique effects on lung cancer and treatment options specially in older patients when adverse effects of comorbidity might have a great challenge in managing the case (Dima et al., 2018). The authors recommended further research with new insights about timely diagnosis, treatment, and long-term surveillance of cancer patients with morbidity (Dima et al., 2018).

The role of comorbidity in management and prognosis on NSCLC was also studied by Nilsson et al. (2017). The source of data was a research database from Lung Cancer Data Base Sweden (LcBaSe) which was a record linkage obtained from the National Lung Cancer Register (NLCR) that includes records of 95% of patients diagnosed with lung cancer in Sweden from 2002 (Nilsson et al., 2017). Charlson Comorbidity Index (CCI) was used to assess comorbidity on a study sample of 19,587 patients diagnosed with NSCLC during the period from 2002 to 2011 (Nilsson et al., 2017).

The researchers used logistic regression and time to event analysis to evaluate the association between comorbidity, treatment, and prognosis (Nilsson et al., 2017). In adjusted analyses, the study results revealed that patients diagnosed with stage IA-IIB lung cancer (NSCLC) with advanced comorbidity were unlikely to qualify for surgery (OR: 0.45; 95% Cl% Cl: 0.36-0.57); while patients in late stages (stage IIIB-IV) with advanced comorbidity were treated with chemotherapy of lower intensity (OR: 0.76; 95% Cl: 0.65-0.89) (Nilsson et al., 2017). Although lung cancer-specific mortality was widely not affected by comorbidity burden in patients with early stages of lung cancer, severe comorbidity in adjusted analyses was associated with increased all-cause mortality (Nilsson et al., 2017).

The researchers concluded that comorbidity contributes to poor prognosis in NSCLC patients, and published survival statistics of lung cancer was not considering comorbidity that was reflected in bad prognosis than it was in fact the case (Nilsson et al., 2017). Therefore, the authors concluded that optimal treatment of comorbidities before and after NSCLC treatment will contribute to a better outcome and prognosis (Nilsson et al., 2017).

Caballero-Vázquez et al. (2021) investigated the risk factors for short-term lung cancer survival on a study sample of 521 patients diagnosed with NSCLC at Virgen de las Nieves Hospital, Granada, Spain during the period between January 1st, 2011, and December 31st, 2016. The purpose of the study was to identify risk factors and assess their effect on prognosis for survival of less or more than one year depending on epidemiological and clinical variables at the time of diagnosis (Caballero-Vázquez et al.,

2021). The data were obtained from electronic medical records of patients and from Granada provincial cancer registry**.** The survival data that reflected the exact date of death were obtained from the regional healthcare system electronic records in Andalusia, Spain (Caballero-Vázquez et al., 2021).

The method used by the investigators was two stepwise binary logistic regression models on a retrospective study of the 521 patients (Caballero-Vázquez et al., 2021). The first model included epidemiological variables such as age, sex, history of smoking, history of lung cancer, and clinical variables such as dyspnea (shortness of breath), cough, dysphonia (change in the pitch or quality of voice), and chest pain as explanatory variables; with the independent risk factors: Age over 70 years, cancer location, significant dyspnea, and dysphonia (Caballero-Vázquez et al., 2021). The second model included therapeutic variables such surgery, chemotherapy, radiotherapy, palliative care, or combination as regressors (Caballero-Vázquez et al., 2021). The purpose of both models was to identify which variables contribute to less than one year survival among these patients (Caballero-Vázquez et al., 2021).

The study results concluded that epidemiological variables such as age, smoking history, history of lung cancer, and clinical variables such as dyspnea, dysphonia, cancer location, and chest pain were suggested as predictors for survival in lung cancer patients at the time of diagnosis; while the therapeutic variables such as surgery, chemotherapy, radiotherapy, or combination is significant for identifying less than one year survival (Caballero-Vázquez et al., 2021). That is because surgery is the treatment of choice for early stages of lung cancer (stage I and II) with a possible curative effect; while other

treatments such as chemotherapy and radiotherapy are inferior to surgery because they are prescribed for late stages of lung cancer (stage III and IV) with palliative effects rather than curative effects (Caballero-Vázquez et al., 2021). This study, based on the history, helps to predict the prognosis of survival at the time of diagnosis as well as it will help to suggest the best therapeutic modality at the time of diagnosis (Caballero-Vázquez et al., 2021). The authors recommended further research to improve survival by developing models with more risk factors (Caballero-Vázquez et al., 2021).

Age

The human body inherits multiple changes at different levels such as cellular and molecular levels as the chronological age increases which lead to a decline in the physiological function that can be expressed in changes in cellular divisions and mutations which eventually make people at a higher risk of cancer at older ages (Gingras, 2020). Cancer researchers were able to document in several studies that most of the age-related-mutagens are carcinogens (Gingras, 2020). As such, age can be considered as a contributing factor for lung cancer and eventually for survival of lung cancer patients. Life expectancy continues to increase globally and with that, the prevalence of age-related diseases is expected to increase as well including lung cancer; therefore, collaborative efforts with stakeholders are required to come with new strategies to facilitate prevention and early diagnosis of such diseases **(**Bürkle et al., 2015).

Botta et al. (2019) underwent a quantitative retrospective study on cancer patients with the objectives to estimate life expectancy (LE) of cancer patients at the time of diagnosis and LE changes from the time of diagnosis through their entire lives to be able

to identify the impact of cancer on their lives. The data used for the study were collected by the network of population-based Italian registries which included 722,737 Italian cancer patients diagnosed during the period from 1985 to 2011and followed until December 31, 2013, for their vital status (Botta et al., 2019). The study included all types of cancers (International Classification of Diseases, Tenth Revision ICD-10 C00-C43, C45-C96) and LE of general population was obtained from the National Institutes of Statistics (Botta et al., 2019). Cohorts studied as centered at mid-point of the age class at diagnosis (ages 17, 22…, 45, 52…, 80 years) (Botta et al., 2019).

The study concluded that the longer the length of time since diagnosis, the more impact of other factors, plus the impact of cancer itself, on cancer survivors' duration and quality of life (Botta et al., 2019). The biggest difference in LE related to sex and age matched general population was noticeable immediately after cancer diagnosis regarding each age class and type of cancer due to the short lethal duration of most aggressive cancers (Botta et al., 2019). Younger cancer patients have a better prognosis than older cancer patients and less mortality risks for non-cancer related causes (Botta et al., 2019). LE has tendency to increase during the first 3-5 years after diagnosis and prognosis improved for each additional year (Botta et al., 2019). The limitations of the study: The study sample represents only 10% of the Italian population, and variability of LE across Italy cannot be excluded even though cancer registries were available in all regions of the country (Botta et al., 2019).

Race/Ethnicity

Racial and ethnic minority groups, as defined by the Centers for Disease Control and Prevention (CDC), include people of color with a wide spectrum of backgrounds and experiences (CDC, 2022). Structural and interpersonal racism compromise mental and physical health of millions of people and challenge them to achieve their optimum health (CDC, 2022). As such, racial inequalities have a negative impact on the U.S. national health (Annangi et al., 2019; CDC, 2022). An increasing number of recent research studies concluded that centuries of racism in the U.S. has a profound negative impact on racial and ethnic minority groups (CDC, 2022). The impact of COVID-19 pandemic disproportionately on black Americans, Hispanics, American Indians, and other racial and ethnic groups has demonstrated the social and racial injustice and inequity those communities have experienced; and the impact of COVID-19 pandemic was only a reminder on how bad the health disparities is (CDC, 2022). COVID-19 data reflected that these communities have higher rates of COVID-19 related hospitalization and death compared with non-Hispanic White populations (CDC, 2022).

Such disparities persist despite controlling for socioeconomic factors and other demographic factors (CDC, 2022; Lake et al., 2020). Mistrust of healthcare system among racial and ethnic minorities was attributed to the history and current experiences of racism; and to avoid further deterioration of health disparities, healthcare providers should engage with these communities to articulate special strategies to curb further widening of mistrust, health inequities, and to deliver evidence-based information (CDC, 2022).

Socioeconomic Status

Mahase et al. (2018) investigated survival disparities by regional poverty level based on radiotherapy treatment (RT) prescribed for lung cancer patients in these regions. The purpose of the study was to evaluate the regional poverty level reflected on differences in lung cancer (LC) survival for lung cancer patients received radiotherapy (Mahase et al., 2018). The investigators used retrospective quantitative method for a study sample retrieved from the Surveillance, Epidemiology, and End Results (SEER) database which included patients diagnosed with LC during the period from 2000 to 2009 (Mahase et al., 2018). The study sample was divided into socioeconomic status (SES) quintiles (five equal groups of populations divided based on distribution of values of the same variable) with quantile 1 represented the highest SES cohort and quantile 5 represented the lowest SES cohort (Mahase et al., 2018). The Kaplan-Meier method with the log-rank test was used to compare overall survival (OS) from diagnosis between demographic and clinical factor levels (Mahase et al., 2018). Multivariate (MVA) Cox proportional hazards regression was used to examine the association of quintile and mortality, adjusting for demographic and clinical factors (Mahase et al., 2018).

The univariate (UVA) results revealed a higher mortality among lung cancer patients who received RT compared to those who did not receive RT (HR: 1.091; CL:1.081-1.102) while the MVA showed a protective effect (HR:0.882; CL: 0.873-0.891) (Mahase et al., 2018). The MVA demonstrated that men had higher mortality than women; Caucasians had lower mortality compared to African Americans while Asians, Pacific Islanders, and Native Americans had the highest overall survival rates (Mahase et

al., 2018). Quintiles 2, 3, 4, and 5 demonstrated higher mortality rates compared to quintile 1 (Mahase et al., 2018). The investigators concluded that RT may offer a positive survival benefit to those who received treatment when accounting for age, gender, race, and SES (Mahase et al., 2018). The study further concluded that an incrementally worse OS rate was associated with increasing regional poverty level even for those who received RT (Mahase et al., 2018).

Health Literacy

The U.S. Department of Health and Human Services (HHS) has expanded the definition of health literacy used in Healthy People 2010 and Healthy People 2020 summarized as: "The degree to which individuals have the capacity to obtain, process, and understand basic health information and services needed to make appropriate health decisions." In the HHS new Healthy People 2030 health literacy is defined as:

Personal Health Literacy: Encompasses the same definition used in Healthy people 2010 and Healthy People 2020 (NIH, 2021).

Organizational Health Literacy: The ability of the organizations to contribute equitably to enable individuals to be health literate as defined in the *Personal Health Literacy* mentioned above (NIH, 2021).

The new HHS definition of health literacy expands the responsibility to include wider sectors of the community, the organizations, instead of limiting the responsibility to individuals which acknowledges the importance of health literacy from the public health perspectives (NIH, 2021). Health literacy is a complex issue that expands beyond

the individual's capability to comprehend and apply as this includes family, community, and systems (NIH, 2021).

Lawrie et al. (2020) conducted a scoping review to evaluate cancer and health literacy among homeless people in the U.S. and Canada. The literature search included 1,124 articles among them 33 articles were found eligible by the authors (Lawrie et al., 2020). The scoping review was conducted during the period from January to March 2016 and it included published articles from year 2000 and after (Lawrie et al., 2020). The method used by the reviewers was Arksey and O'Malley framework and the data sources were Medline, PsycINFO, CINAHL, Scopus, Web of Science, Social Science Abstracts, PubMed, Social Work Abstracts, Embase and Sociological Abstracts (Lawrie et al., 2020).

The review revealed that cancer rate among homeless people in Canada was four times compared to the general population with the most common incident was bronchus and lung cancer; the hospital usage among homeless was three to four times more compared to the general population in both countries; while in the U.S. death rate due to lung cancer was double among homeless compared to the general population (Lawrie et al., 2020). The reviewers found the highest cancer incidence among homeless men and women in the U.S. was bronchus and lung cancer with 88% attributed to tobacco smoking (Lawrie et al., 2020). This expresses disparities in cancer burden among homeless people due to diminished or lack of screening that results in late diagnosis, poor outcome, and short survival or life expectancy (Lawrie et al., 2020).

The review results suggested that low health literacy among homeless people may have contributed significantly to cancer screening access and health care (Lawrie et al., 2020). Furthermore, the review results concluded that lack of health literacy understanding among heath care organizations and providers may play a role to cancer care barriers that consequently affected immediate access to cancer care by population (Lawrie et al., 2020). The limitations of this scoping review were exclusion of the reviewers to the search term *health literacy* to avoid limiting their search yield, and barriers related health literacy could be due to other factors such as geographical location, proximity to health care facilities, and English language for some minorities who speaks other languages (Lawrie et al., 2020).

Summary and Conclusion

This dissertation has contributed to the existing pool of published research studies, especially where limited research studies were done about the effect of lung cancer on survival of marginalizes communities even though about 60% of cancer affects underserved and disadvantage communities (Montagne et al., 2021). There were several areas discussed, analyzed, and promoted in this study that included: The importance of screening for prevention and early diagnosis of lung cancer which will facilitate early treatment and eventually improve the outcome and survival. Screening should be promoted by health care providers and utilize simple less expensive test such blood screening for tumor markers and LDCT; Improving health literacy through community involvement and improving patient-physician communications; lung cancer management should be handled by a team of professionals not a single health care provider as it

requires a multidisciplinary team that may include a pulmonologist or internist, oncologist, surgeon, psychiatrist or psychologist, social worker, and a nurse (Hung et al., 2020; Rankin et al., 2020), as the evaluation and treatment approach should be orchestrated among different specialties to select the best treatment plan (McCann et al., 2021); the evidence-practice gap should be addressed by training health care providers to learn and practice the advanced recommended methods of screening and investigations as well as applying the new recommended treatment protocols (Rankin et al., 2020); promote health education among communities about the risk factors and carcinogens that contribute to lung cancer. Encourage more new research about the epidemiology of lung cancer among marginalized communities and explore new approaches to improve prevention and control of lung cancer among marginalized communities.

The incidence and prevalence of lung cancer as well as the mortality rate cannot be improved without addressing *the effect of treatment, stage of lung cancer, and socioeconomic status on life expectancy of marginalized communities* since about 60% of cancer affects underserved and disadvantaged communities (Montagne et al., 2021).

Chapter 3: Research Method

The purpose of my study was to demonstrate the effect of treatment, stage of lung cancer, and socioeconomic status on life expectancy of marginalized communities. This highlighted the importance of screening for lung cancer which will facilitate early diagnosis, possible curative treatment, better prognosis, and a longer life expectancy. This study had also included identification of the contributing factors for short survival due to lung cancer among marginalized communities. This research was conducted as a quantitative retrospective study on cohorts who were diagnosed with lung cancer at different stages, and they received different treatment regimens that included surgery, chemotherapy, radiotherapy, or combinations of some of them and evaluated their course of treatments until their last follow-ups or deaths.

Surveillance, Epidemiology, and End Results (SEER) was the database source for this quantitative retrospective analysis. SEER is a national program developed by the National Cancer Institute (NCI) to serve as a source of information about cancer statistics in the United States for the purpose of reducing the burden of cancer among the US population (NCI, n.d.). The NCI is one of 27 institutes and centers that comprise the National Institutes of Health (NIH). The Surveillance Research Program (SRP) of the NCI's Division of Cancer Control and Population Sciences (DCCPS) provides support to SEER (NCI, n.d.). The social problem that was investigated in this research study was lung cancer among marginalized and disadvantaged communities who suffer most due to late diagnosis and late treatment because of lack of access to health care which in turn

lead to a short life expectancy. This study has underscored the critical role of treatment, stage of lung cancer, and socioeconomic status on the outcome (survival status).

The major sections in this chapter (chapter 3) included *Introduction section* which outlined the purpose of the study, the source of the dataset, and introduction to the major sections in this chapter. *Research Design and Rationale section* explained a cross-sectional study design that was drawn to examine the relationship between independent variables (treatment such as surgery, chemotherapy, radiotherapy, and a combination of some of them; stages of lung cancer, socioeconomic status) and dependent variable (life expectancy); *Methodology section* summarized a quantitative method that was conducted in the study to facilitate descriptive and inferential statistics analyses to further examine independent and dependent variables*; Data Analysis Plan section* revealed the statistical analyses that were conducted on the dataset which included descriptive statistics such as frequencies, means, and standard deviations (SD); and inferential statistical tests to examine the study hypotheses.

The statistics test used for the analyses was multiple linear regression because it is used when there are more than two measurement variables, one dependent variable and other independent variables (McDonald, 2014). Such as reflected in the dissertation topic Effect of Treatment, Stage of Lung Cancer, and Socioeconomic Status on Life Expectancy of Marginalized Communities; which has 1 dependent variable, Life Expectancy of Marginalized Communities, and 9 independent variables, surgery, chemotherapy, radiotherapy, a combination of therapy (surgery and chemotherapy; surgery and radiotherapy; surgery, chemotherapy, and radiotherapy; chemotherapy and

radiotherapy), stage of lung cancer, and socioeconomic status. Multiple linear regression is useful for prediction of values of dependent variable as well as for suggestions about which independent variables have more effect on the dependent variable (McDonald, 2014). The purpose of multiple linear regression was to find an equation that best predicts the dependent variable as linear function of the independent variables (McDonald, 2014). In this study, a multiple linear regression was applied to understand the linear functional relationships between the dependent variable, Life Expectancy, and the independent variables, surgery, chemotherapy, radiotherapy, a combination of therapy (surgery and chemotherapy; surgery and radiotherapy; surgery, chemotherapy, and radiotherapy; chemotherapy and radiotherapy), stage of lung cancer, and socioeconomic status.

Threats and Validity section included a discussion about validity in quantitative research which can be established by achieving useful inferences from instruments' scores; validity was explained in three traditional forms, *content validity* which is confirming the intended content to be measured was measured; p*redictive or concurrent validity* that inquire about whether the sores of the instrument predict a criterion measure; or whether the scores' results correlate with other results; and *construct validity* is concerned about whether the hypothetical constructs or concepts were measured by the instrument (Creswell & Creswell, 2018). This section also included Threats to Validity with its different types such as internal validity threats, external validity threats, and other types of validity threats. And *Summary section* that summarized the main sections and ideas in this chapter.

Research Design and Rationale

Data for this study involved patients diagnosed with lung cancer at different stages and received different types of treatments during the period (2009-2019). The source of dataset was the Surveillance, Epidemiology, and End Results (SEER) which is a national program developed by the National Cancer Institute (NCI) to serve as a source of information about cancer statistics in the United States for the purpose of reducing the burden of cancer among the US population (NCI, n.d.). The study sample was obtained from SEER database of the NCI for the period (2009-2019) for patients diagnosed with lung cancer during this period and their cause-specific mortality was lung cancer. The statistical analyses were performed by IBM SPSS Statistics for Macintosh, Version 27.0. The statistical analyses that were conducted on the dataset were descriptive statistics such as frequencies, means, and standard deviations (SD); and inferential statistical tests to examine the study hypotheses. Assessment of burden of disease for lung cancer was obtained from the incidence and that was used for allocation of health resources which both were obtained from SEER database using SEER*Stat (version 8.3.6) during the study period (2009-2019). The source was SEER Research Plus Data 17 Registries November 2021 Sub [2009-2019], and the population sample size was 86,998 lung cancer patients.

My research approach was a retrospective analysis of secondary data from SEER database using a quantitative method. This method was aligned with my dissertation topic, *Effect of Treatment, Stage of Lung Cancer, and Socioeconomic status* as independent variables on *life expectancy* (months/ years of survival following diagnosis)

as a dependent variable. To examine the effect of these independent variables on the dependent variable and how that might affect the relationship, I conducted correlational research. Without manipulation of independent variables, correlational research can evaluate the relationship between variables and facilitates explanation of a noticed occurrence between variables (Chiang et al., 2017).

A cross-sectional design was performed to examine the relationship between independent variables (treatment such as surgery, chemotherapy, radiotherapy, combinations of some of them, stages of lung cancer, socioeconomic status) and dependent variable (life expectancy). Cross-sectional design provides inferences about relationship between different variables, but it cannot demonstrate the cause and effect between independent and dependent variables because independent variables (risk factors) and dependent variable (outcome) are measured at the same time; cross-sectional design also reflects the frequency of variables on the sample population (Creswell & Creswell, 2018). To improve the generalizability status of my study sample results on population, I applied Inferential statistics (Creswell & Creswell, 2018).

The statistics test that I used for my analyses was multiple linear regression because it is used when there are more than two measurement variables, one dependent variable and other independent variables (McDonald, 2014). Such as reflected in my dissertation topic *Effect of Treatment, Stage of Lung Cancer, and Socioeconomic Status on Life Expectancy of Marginalized Communities*; which had one dependent variable, *Life Expectancy,* and 9 independent variables, *surgery, chemotherapy, radiotherapy, a combination of therapy (surgery and chemotherapy; surgery and radiotherapy; surgery,*

chemotherapy, and radiotherapy; or chemotherapy and radiotherapy), stage of lung cancer, and socioeconomic status.

The purpose of my study was to demonstrate the impact of different types of treatment that were prescribed based on the stage of lung cancer, the impact of stage of lung cancer at diagnosis, together with the socioeconomic status on life expectancy of marginalized communities. Lung cancer has no symptoms in its early stages (stage I and II) therefore, patients usually diagnosed at late stages (stage III and IV) when curative treatment is mostly not possible. These two factors, late stage of lung cancer and palliative treatment, with low socioeconomic status, contribute to short survival/ life expectancy of marginalized communities who are mostly have no access to health care such as health insurance or financial affordability. Based on the study results, I recommend affordable screening for lung cancer at the communities' levels specially among high-risk groups such as smokers, individuals exposed to carcinogens at their work or at their homes (as asbestos), and individuals with a family history of lung cancer. Despite about 60% of cancer affects underserved and disadvantage communities (Montagne et al., 2021), very few research studies have addressed this issue; therefore, this study would contribute to close this gap in literature.

Methodology

Exploration of patterns and relationships between groups or variables and transforming them into numbers will require statistical methods that can be expressed as descriptive statistics which express patterns of behavior; and can be inferential statistics which express probabilistic arguments that generalize the samples' results to populations

(Rudestam & Newton, 2015). Quantitative method uses statistical methods for data analyses to compare sources of variance of phenomena which facilitate acceptance or rejection of the hypothesis that assumes relationship between the phenomena and make inferences from observations (Rudestam & Newton, 2015). Therefore, I used a quantitative method in this study because descriptive and inferential statistics analyses were conducted to examine the *Effect of Treatment, Stage of Lung Cancer, and Socioeconomic status on life expectancy of marginalized communities.* A cross-sectional study design was used for inferential analyses and descriptive statistics was used to evaluate the frequency and distribution of treatment, stage of lung cancer, and socioeconomic status on the population sample.

My Walden IRB approval number was 08-08-22-0647642; and the study sample was obtained from SEER database of the NCI for the period (2009-2019) for patients diagnosed with lung cancer during this period and their cause-specific mortality was lung cancer. The statistical analyses were performed by IBM SPSS Statistics for Macintosh, Version 27.0. The study sample data were expressed in inferential statistics as multiple linear regression, Confidence Interval (CI), and Pearson Correlation. Assessment of burden of disease for lung cancer was obtained from the incidence and that was used for allocation of health resources which both obtained from SEER database using SEER*Stat (version 8.3.6) during the study period (2009-2019). The source was SEER Research Plus Data 17 Registries November 2021 Sub [2009-2019], and the population sample size was 86,998 lung cancer patients.

The Study Inclusion Criteria. Included patients diagnosed with lung cancer based on pathological laboratory study during the study period (2009-2019); lung cancer patients with specified stage of lung cancer based on TNM staging system during the study period; and lung cancer patients whose specific cause of death was lung cancer during the study period.

The Study Exclusion Criteria. Included lung cancer patients who were not diagnosed during the study period (2009-2019); lung cancer patients whose diagnoses were not based on pathological laboratory study (missed information); lung cancer patients without TNM staging (missed information); and lung cancer patients whose specific cause of death was not lung cancer or was not identified (missed information). The primary endpoint of the study was considered from the date of the initial diagnosis to the date of cancer specific cause of death or to the date of the last follow up appointment.

Data Analysis Plan

The statistical analyses that were conducted on the dataset were descriptive statistics such as frequencies, means, and standard deviations (SD); and inferential statistical tests were used to examine the study hypotheses. The statistics test used for analyses was multiple linear regression because it is used when there are more than two measurement variables, one dependent The statistical analyses that were conducted on the dataset were descriptive statistics such as frequencies, means, and standard deviations (SD); and inferential statistical tests were used to examine the study hypotheses. The statistics test used for analyses was multiple linear regression because it is used when there are more than two measurement variables, one dependent variable and other

independent variables (McDonald, 2014). Such as reflected in my dissertation topic The Effect of Treatment, Stage of Lung Cancer, and Socioeconomic Status on Life Expectancy of Marginalized Communities; which has one dependent variable, Life Expectancy of Marginalized Communities, and 9 independent variables, surgery, chemotherapy, radiotherapy, a combination of therapy (surgery and chemotherapy; surgery and radiotherapy; surgery, chemotherapy, and radiotherapy; chemotherapy and radiotherapy), stage of lung cancer, and socioeconomic status. Multiple linear regression is useful for prediction of values of dependent variable as well as for suggestions about which independent variables have more effect on the dependent variable (McDonald, 2014). The purpose of multiple linear regression is to find an equation that best predicts the dependent variable as linear function of the independent variables (McDonald, 2014). A multiple linear regression was used in this study to understand the functional, linear, relationships between the dependent variable, Life Expectancy of Marginalized Communities, and independent variables, surgery, chemotherapy, radiotherapy, a combination of therapy (surgery and chemotherapy; surgery and radiotherapy; surgery, chemotherapy, and radiotherapy; chemotherapy and radiotherapy), stage of lung cancer, and socioeconomic status.

Research Study Variables

Dependent Variable: Survival/ Life expectancy (months/ years of survival following diagnosis) was obtained from SEER database. The original SEER data Survival/ Life expectancy (months/ years of survival following diagnosis) was transformed to a Scale variable of 60 months of survival following diagnosis.

Independent Variables: Surgery: Dichotomous, code: 0= no surgery; 1= there was surgery.

Chemotherapy: Dichotomous, code: 0= no chemotherapy; 1= there was chemotherapy.

Radiotherapy: Dichotomous, code: 0= no radiotherapy; 1= there was radiotherapy.

A combination of therapy: Ordinal (code: 1=surgery and chemotherapy; 2=surgery and radiotherapy; 3=surgery, chemotherapy, and radiotherapy; 4=chemotherapy and radiotherapy).

Stage of lung cancer: Ordinal, code: 0, 1, 2, 3, 4.

Socioeconomic status: Ordinal, code: 1= below median income, 2= median income, 3= above median income.

Confounders: Race/ ethnicity: Dichotomous, code: 0=Non-black/ non-Hispanic

1= Black/ Hispanic.

Age: Ordinal, code: 0= (40-49), 1= (50-59), 2= (60-69), 3= (70-79), 4= (80-89).

Geographic Location: Dichotomous, code: 0= urban, 1= rural.

Research questions (RQs) and hypotheses (Ha's)

RQ1. What is the association between surgery and survival/ life expectancy after controlling for chemotherapy, radiotherapy, a combination of therapy (surgery and chemotherapy; surgery and radiotherapy; surgery, chemotherapy, and radiotherapy; or chemotherapy and radiotherapy), stage of lung cancer, socioeconomic status, race/ ethnicity, age, and geographic location (urban vs. rural)?

Null hypothesis (H01): There is no association between surgery and survival/ life expectancy after controlling for chemotherapy, radiotherapy, a combination of therapy

(surgery and chemotherapy; surgery and radiotherapy; surgery, chemotherapy, and radiotherapy; or chemotherapy and radiotherapy), stage of lung cancer, socioeconomic status, race/ ethnicity, age, and geographic location (urban vs. rural).

Alternative hypothesis (Ha1): There is an association between surgery and survival/ life expectancy after controlling for chemotherapy, radiotherapy, a combination of therapy (surgery and chemotherapy; surgery and radiotherapy; surgery, chemotherapy, and radiotherapy; or chemotherapy and radiotherapy), stage of lung cancer, socioeconomic status, race/ ethnicity, age, and geographic location (urban vs. rural).

Dependent Variable: Survival/ Life expectancy (months/ years of survival following diagnosis) was obtained from SEER database. The original SEER data Survival/ Life expectancy (months/ years of survival following diagnosis) was transformed to a Scale variable of 60 months of survival following diagnosis.

Independent Variable: Surgery: Dichotomous, code: 0= no surgery; 1= there was surgery.

Covariates: chemotherapy, radiotherapy, a combination of therapy (surgery and chemotherapy; surgery and radiotherapy; surgery, chemotherapy, and radiotherapy; or chemotherapy and radiotherapy), stage of lung cancer, socioeconomic status, race/ ethnicity, age, and geographic location (urban vs. rural).

Test statistic: A *multilinear regression* was performed to examine the association between surgery and survival/ life expectancy after controlling for chemotherapy, radiotherapy, a combination of therapy (surgery and chemotherapy; surgery and radiotherapy; surgery, chemotherapy, and radiotherapy; or chemotherapy and

radiotherapy), stage of lung cancer, socioeconomic status, race/ ethnicity, age, and geographic location (urban vs. rural).

Alpha: 0.05

Power: 0.80

Effect Size: 0.15

Number of Predictors: 9

Calculated Minimum Sample Size: 114

Software: G*Power 3.1.9.7

RQ2. What is the association between chemotherapy and survival/ life expectancy after controlling for surgery, radiotherapy, a combination of therapy (surgery and chemotherapy; surgery and radiotherapy; surgery, chemotherapy, and radiotherapy; or chemotherapy and radiotherapy), stage of lung cancer, socioeconomic status, race/ ethnicity, age, and geographic location (urban vs. rural)?

H02: There is no association between chemotherapy and survival/ life expectancy after controlling for surgery, radiotherapy, a combination of therapy (surgery and chemotherapy; surgery and radiotherapy; surgery, chemotherapy, and radiotherapy; or chemotherapy and radiotherapy), stage of lung cancer, socioeconomic status, race/ ethnicity, age, and geographic location (urban vs. rural).

Ha2: There is an association between chemotherapy and survival/ life expectancy after controlling for surgery, radiotherapy, a combination of therapy (surgery and chemotherapy; surgery and radiotherapy; surgery, chemotherapy, and radiotherapy; or

chemotherapy and radiotherapy), stage of lung cancer, socioeconomic status, race/ ethnicity, age, and geographic location (urban vs. rural).

Dependent Variable: Survival/ Life expectancy (months/ years of survival following diagnosis) was obtained from SEER database. The original SEER data Survival/ Life expectancy (months/ years of survival following diagnosis) was transformed to a Scale variable of 60 months of survival following diagnosis.

Independent Variable: Chemotherapy: Dichotomous, code: 0= no chemotherapy; 1= there was chemotherapy.

Covariates: Surgery, radiotherapy, a combination of therapy (surgery and chemotherapy; surgery and radiotherapy; surgery, chemotherapy, and radiotherapy; or chemotherapy and radiotherapy), stage of lung cancer, socioeconomic status, race/ ethnicity, age, and geographic location (urban vs. rural).

Test statistic: A *multilinear regression* was performed to examine the association between chemotherapy and survival/ life expectancy after controlling for surgery, radiotherapy, a combination of therapy (surgery and chemotherapy; surgery and radiotherapy; surgery, chemotherapy, and radiotherapy; or chemotherapy and radiotherapy), stage of lung cancer, socioeconomic status, race/ ethnicity, age, and geographic location (urban vs. rural).

Alpha: 0.05

Power: 0.80

Effect Size: 0.15

Number of Predictors: 9

Calculated Minimum Sample Size: 114

Software: G*Power 3.1.9.7

RQ3. What is the association between radiotherapy and survival/ life expectancy after controlling for surgery, chemotherapy, a combination of therapy (surgery and chemotherapy; surgery and radiotherapy; surgery, chemotherapy, and radiotherapy; or chemotherapy and radiotherapy), stage of lung cancer, socioeconomic status, race/ ethnicity, age, and geographic location (urban vs. rural)?

H03: There is no association between radiotherapy and survival/ life expectancy after controlling for surgery, chemotherapy, a combination of therapy (surgery and chemotherapy; surgery and radiotherapy; surgery, chemotherapy, and radiotherapy; or chemotherapy and radiotherapy), stage of lung cancer, socioeconomic status, race/ ethnicity, age, and geographic location (urban vs. rural).

Ha3: There is an association between radiotherapy and survival/ life expectancy after controlling for surgery, chemotherapy, a combination of therapy (surgery and chemotherapy; surgery and radiotherapy; surgery, chemotherapy, and radiotherapy; or chemotherapy and radiotherapy), stage of lung cancer, socioeconomic status, race/ ethnicity, age, and geographic location (urban vs. rural).

Dependent Variable: Survival/ Life expectancy (months/ years of survival following diagnosis) was extracted from the SEER database. The original SEER data Survival/ Life expectancy (months/ years of survival following diagnosis) was transformed to a Scale variable of 60 months of survival following diagnosis.

Independent Variable: Radiotherapy: Dichotomous, code: 0= no radiotherapy; 1= there was radiotherapy.

Covariates: Surgery, chemotherapy, a combination of therapy (surgery and chemotherapy; surgery and radiotherapy; surgery, chemotherapy, and radiotherapy; or chemotherapy and radiotherapy), stage of lung cancer, socioeconomic status, race/ ethnicity, age, and geographic location (urban vs. rural).

Test statistic: A *multilinear regression* was performed to examine the association between radiotherapy and survival/ life expectancy after controlling for surgery, chemotherapy, a combination of therapy (surgery and chemotherapy; surgery and radiotherapy; surgery, chemotherapy, and radiotherapy; or chemotherapy and radiotherapy), stage of lung cancer, socioeconomic status, race/ ethnicity, age, and geographic location (urban vs. rural).

Alpha: 0.05

Power: 0.80

Effect Size: 0.15

Number of Predictors: 9

Calculated Minimum Sample Size: 114

Software: G*Power 3.1.9.7

RQ4. What is the association between a combination of therapy (surgery and chemotherapy; surgery and radiotherapy; surgery, chemotherapy, and radiotherapy; or chemotherapy and radiotherapy) and survival/ life expectancy after controlling for

surgery, chemotherapy, radiotherapy, stage of lung cancer, socioeconomic status, race/ ethnicity, age, and geographic location (urban vs. rural)?

H04: There is no association between a combination of therapy (surgery and chemotherapy; surgery and radiotherapy; surgery, chemotherapy, and radiotherapy; or chemotherapy and radiotherapy) and survival/ life expectancy after controlling for surgery, chemotherapy, radiotherapy, stage of lung cancer, socioeconomic status, race/ ethnicity, age, and geographic location (urban vs. rural).

Ha4: There is an association between a combination of therapy (surgery and chemotherapy; surgery and radiotherapy; surgery, chemotherapy, and radiotherapy; or chemotherapy and radiotherapy) and survival/ life expectancy after controlling for surgery, chemotherapy, radiotherapy, stage of lung cancer, socioeconomic status, race/ ethnicity, age, and geographic location (urban vs. rural).

Dependent Variable: Survival/ Life expectancy (months/ years of survival following diagnosis) was extracted from the SEER database. The original SEER data Survival/ Life expectancy (months/ years of survival following diagnosis) was transformed to a Scale variable of 60 months of survival following diagnosis.

Independent Variable: A combination of therapy: Ordinal (code: 1=surgery and chemotherapy; 2=surgery and radiotherapy; 3= surgery, chemotherapy, and radiotherapy; 4=chemotherapy and radiotherapy).

Covariates: Surgery, chemotherapy, radiotherapy, stage of lung cancer, socioeconomic status, race/ ethnicity, age, and geographic location (urban vs. rural).

Test statistic: A *multilinear regression* was performed to examine the association between a combination of therapy (surgery and chemotherapy; surgery and radiotherapy; surgery, chemotherapy, and radiotherapy; or chemotherapy and radiotherapy) and survival/ life expectancy after controlling for surgery, chemotherapy, radiotherapy, stage of lung cancer, socioeconomic status, race/ ethnicity, age, and geographic location (urban vs. rural).

Alpha: 0.05

Power: 0.80

Effect Size: 0.15

Number of Predictors: 9

Calculated Minimum Sample Size: 114

Software: G*Power 3.1.9.7

RQ5. What is the association between stage of lung cancer and survival/ life expectancy after controlling for surgery, chemotherapy, radiotherapy, a combination of therapy (surgery and chemotherapy; surgery and radiotherapy; surgery, chemotherapy, and radiotherapy; or chemotherapy and radiotherapy), socioeconomic status, race/ ethnicity, age, and geographic location (urban vs. rural)?

H05: There is no association between stage of lung cancer and survival/ life expectancy after controlling for surgery, chemotherapy, radiotherapy, a combination of therapy (surgery and chemotherapy; surgery and radiotherapy; surgery, chemotherapy, and radiotherapy; or chemotherapy and radiotherapy), socioeconomic status, race/ ethnicity, age, and geographic location (urban vs. rural).

Ha5: There is an association between stage of lung cancer and survival/ life expectancy after controlling for surgery, chemotherapy, radiotherapy, a combination of therapy (surgery and chemotherapy; surgery and radiotherapy; surgery, chemotherapy, and radiotherapy; or chemotherapy and radiotherapy), socioeconomic status, race/ ethnicity, age, and geographic location (urban vs. rural).

Dependent Variable: Survival/ Life expectancy (months/ years of survival following diagnosis) was extracted from the SEER database. The original SEER data Survival/ Life expectancy (months/ years of survival following diagnosis) was transformed to a Scale variable of 60 months of survival following diagnosis.

Independent Variable: Stage of lung cancer. Numerical/ continuous (stages: 1, 2, 3, 4).

Covariates: Surgery, chemotherapy, radiotherapy, a combination of therapy (surgery and chemotherapy; surgery and radiotherapy; surgery, chemotherapy, and radiotherapy; or chemotherapy and radiotherapy), socioeconomic status, race/ ethnicity, age, and geographic location (urban vs. rural).

Test statistic: A *multilinear regression* was performed to examine the association between stage of lung cancer and survival/ life expectancy after controlling for surgery, chemotherapy, radiotherapy, a combination of therapy (surgery and chemotherapy; surgery and radiotherapy; surgery, chemotherapy, and radiotherapy; or chemotherapy and radiotherapy), socioeconomic status, race/ ethnicity, age, and geographic location (urban vs. rural).

Alpha: 0.05

Power: 0.80

Effect Size: 0.15

Number of Predictors: 9

Calculated Minimum Sample Size: 114

Software: G*Power 3.1.9.7

RQ6. What is the association between socioeconomic status and survival/ life expectancy after controlling for surgery, chemotherapy, radiotherapy, a combination of therapy (surgery and chemotherapy; surgery and radiotherapy; surgery, chemotherapy, and radiotherapy; or chemotherapy and radiotherapy), stage of lung cancer, race/ ethnicity, age, and geographic location (urban vs. rural)?

H06: There is no association between socioeconomic status and survival/ life expectancy after controlling for surgery, chemotherapy, radiotherapy, a combination of therapy (surgery and chemotherapy; surgery and radiotherapy; surgery, chemotherapy, and radiotherapy; or chemotherapy and radiotherapy), stage of lung cancer, race/ ethnicity, age, and geographic location (urban vs. rural).

Ha6: There is an association between socioeconomic status and survival/ life expectancy after controlling for surgery, chemotherapy, radiotherapy, a combination of therapy (surgery and chemotherapy; surgery and radiotherapy; surgery, chemotherapy, and radiotherapy; or chemotherapy and radiotherapy), stage of lung cancer, race/ ethnicity, age, and geographic location (urban vs. rural).

Dependent Variable: Survival/ Life expectancy (months/ years of survival following diagnosis) was extracted from the SEER database. The original SEER data Survival/ Life

expectancy (months/ years of survival following diagnosis) was transformed to a Scale variable of 60 months of survival following diagnosis.

Independent Variable: Socioeconomic status. Ordinal coding: 0=low income, 1=average income, 2=high income.

Covariates: Surgery, chemotherapy, radiotherapy, a combination of therapy (surgery and chemotherapy; surgery and radiotherapy; surgery, chemotherapy, and radiotherapy; or chemotherapy and radiotherapy), race/ ethnicity, age, and geographic location (urban vs. rural).

Test statistic: A *multilinear regression* was performed to examine the association between socioeconomic status and survival/ life expectancy after controlling for surgery, chemotherapy, radiotherapy, a combination of therapy (surgery and chemotherapy; surgery and radiotherapy; surgery, chemotherapy, and radiotherapy; or chemotherapy and radiotherapy), race/ ethnicity, age, and geographic location (urban vs. rural).

Alpha: 0.05

Power: 0.80

Effect Size: 0.15

Number of Predictors: 9

Calculated Minimum Sample Size: 114

Software: G*Power 3.1.9.7

RQ7. What is the association between race/ ethnicity, age, and geographic location (urban vs. rural) and survival/ life expectancy after controlling for surgery, chemotherapy, radiotherapy, a combination of therapy (surgery and chemotherapy;

surgery and radiotherapy; surgery, chemotherapy, and radiotherapy; or chemotherapy and radiotherapy), stage of lung cancer, and socioeconomic status?

H07: There is no association between race/ ethnicity, age, and geographic location (urban vs. rural) and survival/ life expectancy after controlling for surgery, chemotherapy, radiotherapy, a combination of therapy (surgery and chemotherapy; surgery and radiotherapy; surgery, chemotherapy, and radiotherapy; or chemotherapy and radiotherapy), stage of lung cancer, and socioeconomic status.

Ha7: There is an association between race/ ethnicity, age, and geographic location (urban vs. rural) and survival/ life expectancy after controlling for surgery, chemotherapy, radiotherapy, a combination of therapy (surgery and chemotherapy; surgery and radiotherapy; surgery, chemotherapy, and radiotherapy; or chemotherapy and radiotherapy), stage of lung cancer, and socioeconomic status.

Dependent Variable: Survival/ Life expectancy (months/ years of survival following diagnosis) was extracted from the SEER database. The original SEER data Survival/ Life expectancy (months/ years of survival following diagnosis) was transformed to a Scale variable of 60 months of survival following diagnosis.

Independent Variables: Confounders: Race/ ethnicity: Dichotomous, code: 0=Non-black/ non-Hispanic; 1= Black/ Hispanic.

Age: Ordinal, code: 0= (40-49), 1= (50-59), 2= (60-69), 3= (70-79), 4= (80-89).

Geographic Location: Dichotomous, code: 0= urban, 1= rural.

Covariates: surgery, chemotherapy, radiotherapy, a combination of therapy (surgery and chemotherapy; surgery and radiotherapy; surgery, chemotherapy, and radiotherapy; or chemotherapy and radiotherapy), stage of lung cancer, and socioeconomic status.

Test statistic: A *multilinear regression* was performed to examine the association between race/ ethnicity, age, and geographic location (urban vs. rural) and survival/ life expectancy after controlling for surgery, chemotherapy, radiotherapy, a combination of therapy (surgery and chemotherapy; surgery and radiotherapy; surgery, chemotherapy, and radiotherapy; or chemotherapy and radiotherapy), stage of lung cancer, and socioeconomic status.

Alpha: 0.05

Power: 0.80

Effect Size: 0.15

Number of Predictors: 9

Calculated Minimum Sample Size: 114

Software: G*Power 3.1.9.7

Threats and Validity

Validity

Validity in quantitative research could be established by achieving useful inferences from instruments' scores; validity is explained in three traditional forms (Creswell & Creswell, 2018):

*Content Validity***.** Is confirming the intended content to be measured was measured.

*Predictive or Concurrent Validity***.** Inquire about whether the scores of the instrument predict a criterion measure; or whether the scores' results correlate with other results.

Construct Validity. Is concerned about whether the hypothetical constructs or concepts were measured by the instrument. Recent studies considered construct validity is more objective in validity that other forms of validity as it is concerned with whether the scores have achieved the intended goal and led to a positive outcome when applied in real life practice (Creswell & Creswell, 2018).

Threats to Validity

Threats to validity are two main types: a) *Internal validity threats;* b) *External validity threats;* and c*) Other validity threats* such as threats to statistical conclusion validity and threats to construct validity (Creswell & Creswell, 2018).

Internal validity threats*.* Are related to experiences with the sample population or participants that threaten accurate inferences to be drawn from the data about the sample population (Creswell & Creswell, 2018). Internal validity threats can be related to the participants, such as history, maturation, regression, and mortality, or they can be related to the researcher's manipulation of the test, such as diffusion, and compensations, or they can be related to the procedures applied (Creswell & Creswell, 2018). The internal validity threats that could be related to my study were selection bias which were avoided by random selection of participants. Internal validity threats could be due to mortality or study attrition which were avoided by selection of a large study sample to avoid dropouts, incomplete records, or mortality due to other causes.

External validity threats. Happen when the researcher draws incorrect inferences from the sample data that could not be generalized to the general population, or at different settings, or applied to past or future situations (Creswell & Creswell, 2018).

Types of Threats to External Validity

Interaction of selection and treatment which rise when characteristics of participants are narrow and cannot be generalized to other people who do not share the characteristics of participants and the solution to this threat is to conduct additional study on participants with different characteristics (Creswell & Creswell, 2018).

Interaction of setting and treatment is another type of threat to external validity which take place when characteristics of setting of participants are different from characteristics of setting of other individuals; therefore, they cannot be generalized to the general population; so, to avoid this type of threat, the researcher can conduct additional study on participants with different settings and compare the new results with the initial results (Creswell & Creswell, 2018).

Interaction of history and treatment is a third type of threat to external validity that occur when results of study are time-bound and they cannot be generalized to past or future situations; so, to overcome this threat, the researcher can repeat the study later in time and compare the results with the initial study to see if they are the same or different (Creswell & Creswell, 2018).

Threats to external validity in my study might not be experienced immediately as my study sample was extracted from SEER database from the NCI national database on cancer statistics; and the NCI database (SEER) represents 34% of cancer statistic in the

US (NCI, n.d.) which was a random sample that had no narrow characteristics of participants (no interaction of selection) nor different characteristics of setting of the participants from the general population (no interaction of setting). However, with the rapid changes and advances in the diagnostic and therapeutic protocols the interaction of history will take place, when results of the study are time-bound, they cannot be generalized to future situations. Therefore, I highly recommend future research to continue in the same topic theme of my dissertation as threats to external validity will eventually take place as the time passes.

Other validity threats. Threats to statistical conclusion validity that results from inaccurate inferences drawn from the data due to inadequate statistical power or statistical assumptions were violated (Creswell & Creswell, 2018). These threats were avoided in my study by using G*Power calculator (G*Power 3.1 manual, 2017) for power analysis and to minimize bias on the size of the sample population.

Threats to construct validity that occur due to inadequate definitions and measures of variables by the researcher (Creswell & Creswell, 2018). In this research study variables were well defined and were accurately measured to avoid threats to construct validity.

Summary

Data for this study involved patients diagnosed with lung cancer at different stages and received different types of treatments during the period (2009-2019). The source of dataset was from SEER which is a national program developed by the National Cancer Institute (NCI) to serve as a source of information about cancer statistics in the

United States. My research approach was retrospective analyses of secondary data from SEER database using a quantitative method. The statistical analyses that were conducted on the dataset were descriptive statistics such as frequencies, means, and standard deviations (SD); and inferential statistical tests to examine the study hypotheses. A multiple linear regression was performed in the study to examine the linear functional relationships between the dependent variable, *Life Expectancy of Marginalized Communities,* and independent variables, *surgery, chemotherapy, radiotherapy, a combination of therapy (surgery and chemotherapy; surgery and radiotherapy; surgery, chemotherapy, and radiotherapy; or chemotherapy and radiotherapy), stage of lung cancer, and socioeconomic status.*

A cross-sectional design was conducted to examine the relationship between independent variables, treatment (such as surgery, chemotherapy, radiotherapy, combination of therapies), stages of lung cancer, socioeconomic status, and dependent variable (life expectancy). Cross-sectional design provides inferences about relationship between different variables, and it serves well in observational studies such as in this study. Cross-sectional design facilitates analyses of sample population in a specific point in time which also served well in this study.

Chapter 4: Results

The purpose of my study was to demonstrate the effect of treatment, stage of lung cancer, and socioeconomic status on survival/ life expectancy of marginalized communities. This study highlighted the importance of screening for lung cancer which facilitates early diagnosis, possible curative treatment, better prognosis, and a longer life expectancy. My study also identified and analyzed the contributing factors to short survival due to lung cancer among marginalized communities. This was conducted as a quantitative retrospective study on cohorts who were diagnosed with lung cancer at different stages, and they received different treatment regimens that included surgery, chemotherapy, radiotherapy, or combinations of some of them and evaluated their course of treatments until their last follow-ups or deaths.

This chapter included the following sections: An *Introduction* to chapter 4; *Data Collection section* which explained the source of dataset, the period considered for the study, the number of subjects in the study, steps taken to clean the raw data, and the study inclusion and exclusion criteria; *Results* of the SPSS output analyses of the research questions' (RQs) that included statistical assumptions, descriptive statistics results, and inferential analyses results; and a *Summary* of chapter 4.

Research Questions

RQ1. What is the association between surgery and survival/ life expectancy after controlling for chemotherapy, radiotherapy, a combination of therapy (surgery and chemotherapy; surgery and radiotherapy; surgery, chemotherapy, and radiotherapy; or

chemotherapy and radiotherapy), stage of lung cancer, socioeconomic status, race/ ethnicity, age, and geographic location (urban vs. rural)?

Null hypothesis (H01): There is no association between surgery and survival/ life expectancy after controlling for chemotherapy, radiotherapy, a combination of therapy (surgery and chemotherapy; surgery and radiotherapy; surgery, chemotherapy, and radiotherapy; or chemotherapy and radiotherapy), stage of lung cancer, socioeconomic status, race/ ethnicity, age, and geographic location (urban vs. rural).

Alternative hypothesis (Ha1): There is an association between surgery and survival/ life expectancy after controlling for chemotherapy, radiotherapy, a combination of therapy (surgery and chemotherapy; surgery and radiotherapy; surgery, chemotherapy, and radiotherapy; or chemotherapy and radiotherapy), stage of lung cancer, socioeconomic status, race/ ethnicity, age, and geographic location (urban vs. rural).
Test statistic: A *multiple linear regression* was performed to examine the association between surgery and survival/ life expectancy after controlling for chemotherapy, radiotherapy, a combination of therapy (surgery and chemotherapy; surgery and radiotherapy; surgery, chemotherapy, and radiotherapy; or chemotherapy and radiotherapy), stage of lung cancer, socioeconomic status, race/ ethnicity, age, and geographic location (urban vs. rural).

RQ2. What is the association between chemotherapy and survival/ life expectancy after controlling for surgery, radiotherapy, a combination of therapy (surgery and chemotherapy; surgery and radiotherapy; surgery, chemotherapy, and radiotherapy; or

chemotherapy and radiotherapy), stage of lung cancer, socioeconomic status, race/ ethnicity, age, and geographic location (urban vs. rural)?

H02: There is no association between chemotherapy and survival/ life expectancy after controlling for surgery, radiotherapy, a combination of therapy (surgery and chemotherapy; surgery and radiotherapy; surgery, chemotherapy, and radiotherapy; or chemotherapy and radiotherapy), stage of lung cancer, socioeconomic status, race/ ethnicity, age, and geographic location (urban vs. rural).

Ha2: There is an association between chemotherapy and survival/ life expectancy after controlling for surgery, radiotherapy, a combination of therapy (surgery and chemotherapy; surgery and radiotherapy; surgery, chemotherapy, and radiotherapy; or chemotherapy and radiotherapy), stage of lung cancer, socioeconomic status, race/ ethnicity, age, and geographic location (urban vs. rural).

Test statistic: A *multiple linear regression* was performed to examine the association between chemotherapy and survival/ life expectancy after controlling for surgery, radiotherapy, a combination of therapy (surgery and chemotherapy; surgery and radiotherapy; surgery, chemotherapy, and radiotherapy; or chemotherapy and radiotherapy), stage of lung cancer, socioeconomic status, race/ ethnicity, age, and geographic location (urban vs. rural).

RQ3. What is the association between radiotherapy and survival/ life expectancy after controlling for surgery, chemotherapy, a combination of therapy (surgery and chemotherapy; surgery and radiotherapy; surgery, chemotherapy, and radiotherapy; or

chemotherapy and radiotherapy), stage of lung cancer, socioeconomic status, race/ ethnicity, age, and geographic location (urban vs. rural)?

H03: There is no association between radiotherapy and survival/ life expectancy after controlling for surgery, chemotherapy, a combination of therapy (surgery and chemotherapy; surgery and radiotherapy; surgery, chemotherapy, and radiotherapy; or chemotherapy and radiotherapy), stage of lung cancer, socioeconomic status, race/ ethnicity, age, and geographic location (urban vs. rural).

Ha3: There is an association between radiotherapy and survival/ life expectancy after controlling for surgery, chemotherapy, a combination of therapy (surgery and chemotherapy; surgery and radiotherapy; surgery, chemotherapy, and radiotherapy; or chemotherapy and radiotherapy), stage of lung cancer, socioeconomic status, race/ ethnicity, age, and geographic location (urban vs. rural).

Test statistic: A *multiple linear regression* was performed to examine the association between radiotherapy and survival/ life expectancy after controlling for surgery, chemotherapy, a combination of therapy (surgery and chemotherapy; surgery and radiotherapy; surgery, chemotherapy, and radiotherapy; or chemotherapy and radiotherapy), stage of lung cancer, socioeconomic status, race/ ethnicity, age, and geographic location (urban vs. rural).

RQ4. What is the association between a combination of therapy (surgery and chemotherapy; surgery and radiotherapy; surgery, chemotherapy, and radiotherapy; or chemotherapy and radiotherapy) and survival/ life expectancy after controlling for

surgery, chemotherapy, radiotherapy, stage of lung cancer, socioeconomic status, race/ ethnicity, age, and geographic location (urban vs. rural)?

H04: There is no association between a combination of therapy (surgery and chemotherapy; surgery and radiotherapy; surgery, chemotherapy, and radiotherapy; or chemotherapy and radiotherapy) and survival/ life expectancy after controlling for surgery, chemotherapy, radiotherapy, stage of lung cancer, socioeconomic status, race/ ethnicity, age, and geographic location (urban vs. rural).

Ha4: There is an association between a combination of therapy (surgery and chemotherapy; surgery and radiotherapy; surgery, chemotherapy, and radiotherapy; or chemotherapy and radiotherapy) and survival/ life expectancy after controlling for surgery, chemotherapy, radiotherapy, stage of lung cancer, socioeconomic status, race/ ethnicity, age, and geographic location (urban vs. rural).

Test statistic: A *multiple linear regression* was performed to examine the association between a combination of therapy (surgery and chemotherapy; surgery and radiotherapy; surgery, chemotherapy, and radiotherapy; or chemotherapy and radiotherapy) and survival/ life expectancy after controlling for surgery, chemotherapy, radiotherapy, stage of lung cancer, socioeconomic status, race/ ethnicity, age, and geographic location (urban vs. rural).

RQ5. What is the association between stage of lung cancer and survival/ life expectancy after controlling for surgery, chemotherapy, radiotherapy, a combination of therapy (surgery and chemotherapy; surgery and radiotherapy; surgery, chemotherapy,

and radiotherapy; or chemotherapy and radiotherapy), socioeconomic status, race/ ethnicity, age, and geographic location (urban vs. rural)?

H05: There is no association between stage of lung cancer and survival/ life expectancy after controlling for surgery, chemotherapy, radiotherapy, a combination of therapy (surgery and chemotherapy; surgery and radiotherapy; surgery, chemotherapy, and radiotherapy; or chemotherapy and radiotherapy), socioeconomic status, race/ ethnicity, age, and geographic location (urban vs. rural)

Ha5: There is an association between stage of lung cancer and survival/ life expectancy after controlling for surgery, chemotherapy, radiotherapy, a combination of therapy (surgery and chemotherapy; surgery and radiotherapy; surgery, chemotherapy, and radiotherapy; or chemotherapy and radiotherapy), socioeconomic status, race/ ethnicity, age, and geographic location (urban vs. rural).

Test statistic: A *multiple linear regression* was performed to examine the association between stage of lung cancer and survival/ life expectancy after controlling for surgery, chemotherapy, radiotherapy, a combination of therapy (surgery and chemotherapy; surgery and radiotherapy; surgery, chemotherapy, and radiotherapy; or chemotherapy and radiotherapy), socioeconomic status, race/ ethnicity, age, and geographic location (urban vs. rural).

RQ6. What is the association between socioeconomic status and survival/ life expectancy after controlling for surgery, chemotherapy, radiotherapy, a combination of therapy (surgery and chemotherapy; surgery and radiotherapy; surgery, chemotherapy,

and radiotherapy; or chemotherapy and radiotherapy), stage of lung cancer, race/ ethnicity, age, and geographic location (urban vs. rural)?

H06: There is no association between socioeconomic status and survival/ life expectancy after controlling for surgery, chemotherapy, radiotherapy, a combination of therapy (surgery and chemotherapy; surgery and radiotherapy; surgery, chemotherapy, and radiotherapy; or chemotherapy and radiotherapy), stage of lung cancer, race/ ethnicity, age, and geographic location (urban vs. rural)?

Ha6: There is an association between socioeconomic status and survival/ life expectancy after controlling for surgery, chemotherapy, radiotherapy, a combination of therapy (surgery and chemotherapy; surgery and radiotherapy; surgery, chemotherapy, and radiotherapy; or chemotherapy and radiotherapy), stage of lung cancer, race/ ethnicity, age, and geographic location (urban vs. rural)?

status.

Test statistic: A *multiple linear regression* was performed to examine the association between socioeconomic status and survival/ life expectancy after controlling for surgery, chemotherapy, radiotherapy, a combination of therapy (surgery and chemotherapy; surgery and radiotherapy; surgery, chemotherapy, and radiotherapy; or chemotherapy and radiotherapy), stage of lung cancer, race/ ethnicity, age, and geographic location (urban vs. rural).

RQ7. What is the association between race/ ethnicity, age, and geographic location (urban vs. rural) and survival/ life expectancy after controlling for surgery, chemotherapy, radiotherapy, a combination of therapy (surgery and chemotherapy;

surgery and radiotherapy; surgery, chemotherapy, and radiotherapy; or chemotherapy and radiotherapy), stage of lung cancer, and socioeconomic status?

H07: There is no association between race/ ethnicity, age, and geographic location (urban vs. rural) and survival/ life expectancy after controlling for surgery, chemotherapy, radiotherapy, a combination of therapy (surgery and chemotherapy; surgery and radiotherapy; surgery, chemotherapy, and radiotherapy; or chemotherapy and radiotherapy), stage of lung cancer, and socioeconomic status.

Ha7: There is an association between race/ ethnicity, age, and geographic location (urban vs. rural) and survival/ life expectancy after controlling for surgery, chemotherapy, radiotherapy, a combination of therapy (surgery and chemotherapy; surgery and radiotherapy; surgery, chemotherapy, and radiotherapy; or chemotherapy and radiotherapy), stage of lung cancer, and socioeconomic status.

Test statistic: A *multiple linear regression* was performed to examine the association between race/ ethnicity, age, and geographic location (urban vs. rural) and survival/ life expectancy after controlling for surgery, chemotherapy, radiotherapy, a combination of therapy (surgery and chemotherapy; surgery and radiotherapy; surgery, chemotherapy, and radiotherapy; or chemotherapy and radiotherapy), stage of lung cancer, and socioeconomic status.

Data Collection

The study sample was obtained from the Surveillance, Epidemiology, and End Results (SEER) which is a national program developed by the National Cancer Institute (NCI) to serve as a source of information about cancer statistics in the United States for

the purpose of reducing the burden of cancer among the US population (NCI, n.d.). The dataset was obtained for the period (2009-2019) for patients diagnosed with lung cancer during this period and their cause-specific mortality was lung cancer. No discrepancies were noticed in data retrieval compared to data collection plan mentioned in chapter 3.

The statistical analyses were performed by IBM SPSS Statistics for Macintosh, Version 27.0. The study sample dataset was expressed in inferential statistics as multiple linear regression, Confidence Interval (CI), and Pearson Correlation. Assessment of burden of disease for lung cancer was obtained from the incidence and that was used for allocation of health resources which both obtained from SEER database using SEER*Stat (version 8.3.6) during the study period (2009-2019). The study sample size was 86,998 lung cancer patients. First, I retrieved raw data from SEER database that included extended periods of time from year 2000 to 2019 with multiple types of cancers which I cleaned after excluding other types of cancer, coded my variables, and limited the study period to (2009-2019). The following were my study inclusion and exclusion criteria for my data collection:

The Study Inclusion Criteria. Included patients diagnosed with lung cancer based on pathological laboratory study during the study period (2009-2019); lung cancer patients with specified stage of lung cancer based on TNM staging system during the study period; and lung cancer patients whose specific cause of death was lung cancer during the study period.

***The Study Exclusion Criteria*.** Included lung cancer patients who were not diagnosed during the study period (2009-2019); lung cancer patients whose diagnoses

were not based on pathological laboratory study (missed information); lung cancer patients without TNM staging (missed information); and lung cancer patients whose specific cause of death was not lung cancer or was not identified (missed information). The primary endpoint of my study was considered from the date of the initial diagnosis to the date of cancer specific cause of death or to the date of the last follow up appointment.

Study Analyses Results

A quantitative method was used in this study to conduct descriptive and inferential statistics analyses to examine the *Effect of Treatment, Stage of Lung Cancer, and Socioeconomic status on life expectancy of marginalized communities.* A cross-sectional study design was also used for inferential analyses and descriptive statistics to evaluate the frequency and distribution of treatment, stage of lung cancer, and socioeconomic status on the population sample.

RQ1

What is the association between surgery and survival/ life expectancy after controlling for chemotherapy, radiotherapy, a combination of therapy (surgery and chemotherapy; surgery and radiotherapy; surgery, chemotherapy, and radiotherapy; or chemotherapy and radiotherapy), stage of lung cancer, socioeconomic status, race/ ethnicity, age, and geographic location (urban vs. rural)?

Descriptive Statistics Results. The study sample (n= 86,998), survival months, mean was 7.56, and standard deviation was 6.664. Surgery, mean was 0.23, and standard deviation was 0.424. Stage of lung cancer, mean 2.84, and standard deviation was 1.265. Socioeconomic status, (median household income), mean 7.40, and standard deviation

was 2.589. Race (White and Non-White), mean 1.26, and standard deviation was 0.439. Age, mean 14.62, and standard deviation was 2.061. Geographic location (Urban vs. Rural), mean 1.16, and standard deviation was 0.362. (Table 1, Figure 1).

Table 1

Descriptive Statistics of Survival Months, Surgery, Stage of Lung Cancer, Socioeconomic Status, Race, Age, and Geographic Location of Lung Cancer Patients

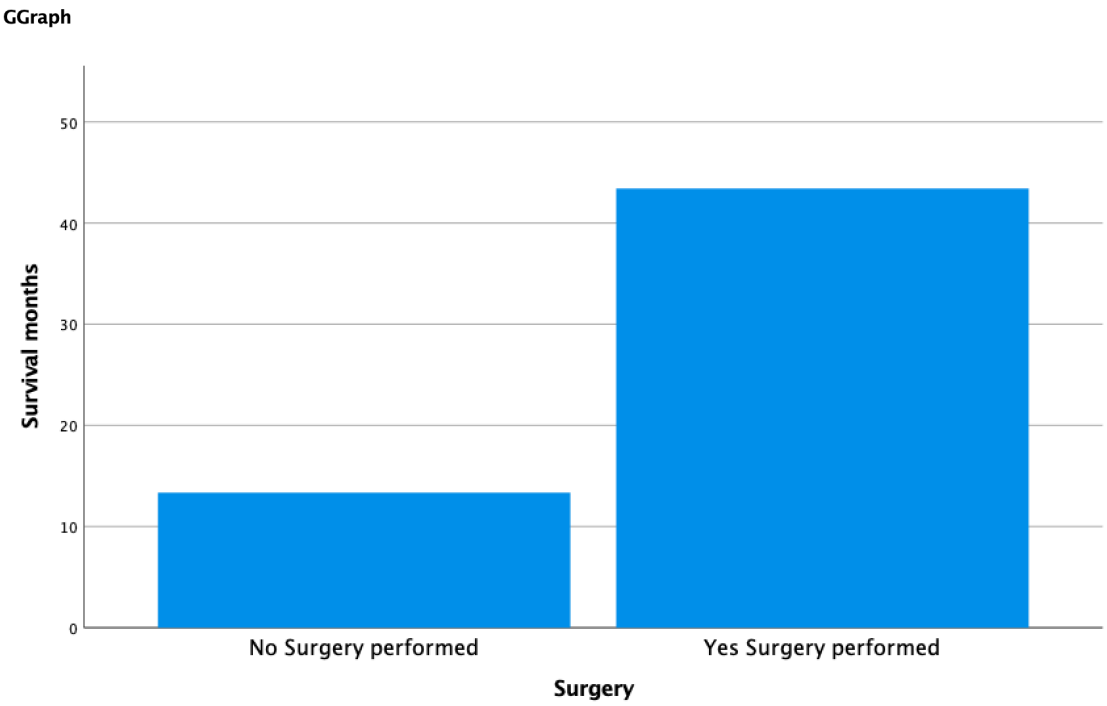
	Mean	**Std. Deviation**	**N**
Survival Months	7.56	6.664	86998
Surgery	0.23	0.424	86998
Stage of Lung Cancer	2.84	1.265	86998
Socioeconomic Status: Median household income	7.40	2.589	86998
Race: White and Non-White	1.26	0.439	86998
Age	14.62	2.061	86998
Geographic Location: Urban vs. Rural	1.16	0.362	86998

Note. Dependent variable: Survival Months.

Data Source: SEER Research Plus Data 17 Registries Nov 2021 Sub [2009-2019].

Figure 1

Association Between Surgery and Survival of Lung Cancer Patients



Note. Data Source: SEER Research Plus Data 17 Registries Nov 2021 Sub [2009-2019].

Statistical Assumptions

I used multiple linear regression for my dataset analyses. To obtain valid results, the following eight assumptions were checked and satisfied to qualify my variables to be analyzed using multiple linear regression (Laerd Statistics, 2018):

Assumption-1: The dependent variable (survival/ life expectancy) should be measured on a continuous scale (interval or ratio).

Assumption-2: The Independent variables should be two or more and they can be continuous or categorical (surgery, chemotherapy, radiotherapy, a combination of therapy

[surgery and chemotherapy; surgery and radiotherapy; surgery, chemotherapy, and radiotherapy]; or chemotherapy and radiotherapy),

Assumption-3: The variables should reflect independence of observations (independence of residuals). In other words, variables should not be related to each other or appear as clustered. This can be verified by Durbin-Watson statistic test using SPSS statistics.

Assumption-4: There should be a linear relationship between the dependent variable and each of the independent variables; as well as between the dependent variable and all the independent variables together. This can be verified by different means such as scatterplots and partial regression plots using SPSS statistics.

Assumption-5: The data should reflect homoscedasticity or homogeneity of variance. This is an assumption of similar variances in different groups along the line were compared. This assumption can be tested using SPSS statistics.

Assumption-6: The data should not reflect multicollinearity that takes place when two or more independent variables are drastically correlated with each other. Such multicollinearity creates misunderstanding of identifying the independent variable that contributes to the variance mentioned in the dependent variable. This also creates a technical problem when calculating a multiple regression model. Multicollinearity can also be tested by using SPSS statistics.

Assumption-7: The dataset should not include significant outliers, high leverage points, or highly influential points. Such observations could have different effects on the regression line which in turn negatively affect the equation of regression used to predict the dependent variable value based on the independent variables. As such, the SPSS output

will reflect reduced predictive accuracy of the results as well as of the statistical significance. These observations can be checked while using SPSS statistics for multiple linear regression. Outliers and leverage points can be detected by using *casewise diagnostics and studentized deleted residuals* when using SPSS statistics. The influential points can be checked while using SPSS statistics by testing a measure of influence known as *Cook's Distance.*

Assumption-8: Residuals (errors) should be approximately normally distributed. This can be verified by a superimposed normal curve histogram and a normal P-P Plot or by a normal Q-Q Plot of the studentized residuals. Both methods can be tested by using SPSS statistics.

Inferential Analyses Results. To investigate What is the association between Surgery and Survival/ life expectancy of lung cancer patients, a simple linear regression was conducted. The predictor was Surgery, and the outcome was Survival/ life expectancy. The predictor variable was found to be statistically significant [B = 4.055, 95% C.I. (3.954, 4.156), $p < .001$], indicating that for lung cancer patients receiving surgery, Survival/ life expectancy increases by approximately 4 months compared to those who did not do Surgery. The model explained approximately 7% of the variability [R-squared = 0.067]. A sensitivity analysis using G*power 3.1.9.7 software [n = 86,998; alpha error prob = 0.05; power = 0.80; predictors = 1] calculated [Effect size f^2 = 9.022085e-05] a small effect. The confidence interval associated with the regression analysis does not contain 0, which means the null hypothesis, there is no association between Surgery and Survival/ life expectancy of lung cancer patients, can be rejected.

To investigate What is the association between Surgery and Survival/ life expectancy of lung cancer patients after controlling for Stage of lung cancer, Socioeconomic status, Race/ ethnicity, Age, and Geographical location (urban vs. rural)?, a multiple linear regression analysis was conducted to evaluate the prediction of Survival/ life expectancy of lung cancer patients from Surgery, Stage of lung cancer, Socioeconomic status, Race/ ethnicity, Age, and Geographical location (urban vs. rural). The results of the multiple linear regression analysis revealed Geographical location (urban vs. rural) not to be a statistically significant predictor to the model ($p > .05$). However, the results of the multiple linear regression analysis revealed a statistically significant association between Surgery, Stage of lung cancer, Socioeconomic status, Race/ ethnicity, and Age.

Controlling for Stage of lung cancer, Socioeconomic status, Race/ ethnicity, and Age, the regression coefficient [B = 1.484, 95% C.I. (1.366, 1.602) $p < .001$] associated with Surgery suggests that for lung cancer patients receiving surgery, Survival/ life expectancy increases by approximately 1.5 months compared to those who did not do Surgery.

Controlling for Surgery, Socioeconomic status, Race/ ethnicity, and Age, the regression coefficient [B = -1.462, 95% C.I. (-1.501, -1.423) $p < .001$] associated with Stage of lung cancer, suggests that with each increase of lung cancer stage (Stage 1, Stage 2, Stage 3, and Stage 4), Survival/ life expectancy decreases by approximately 1.5 months.

Controlling for Surgery, Stage of lung cancer, Race/ ethnicity, and Age, the regression coefficient [B = 0.107, 95% C.I. (0.088, 0.127) $p < .001$] associated with Socioeconomic status, suggests that for median household income level increase, Survival/ life expectancy increases by approximately 0.1 month or 3 days. Controlling for Surgery, Stage of lung cancer, Socioeconomic status, and Age, the regression coefficient [B = 0.097, 95% C.I. (0.001, 0.193) $p = .048$] associated with Race/ ethnicity suggests that Whites compared to non-Whites, have Survival/ life expectancy increases by approximately 0.1 months or 3 days.

Controlling for Surgery, Stage of lung cancer, Socioeconomic status, and Race/ ethnicity, the regression coefficient [B = -0.324, 95% C.I. (-0.344, -0.303) $p < .001$] associated with Age, suggests that with each additional five years of age, Survival/ life expectancy decreases by approximately 0.3 month or 9 days.

Controlling for Surgery, Stage of lung cancer, Socioeconomic status, Race/ ethnicity, and Age, the regression coefficient [B = 1.484, 95% C.I. (1.366, 1.602 $p < .05$] associated with [IV2] suggests that with each additional [IV2], Survival/ life expectancy increases by approximately 1.5 months. The R-squared value of [0.126] associated with this regression model suggests that the association between Surgery, Stage of lung cancer, Socioeconomic status, Race/ ethnicity, and Age account for [12.6%] of the variation in Survival/ life expectancy of lung cancer patients, which means that [87.4%] of the variation in Survival/ life expectancy of lung cancer patients cannot be explained by Surgery, Stage of lung cancer, Socioeconomic status, Race/ ethnicity, and Age alone. A sensitivity analysis using G*power 3.1.9.7 software [n = 86,998; alpha error prob =

0.05; power = 0.80; predictors = 6] calculated [Effect size $f^2 = 0.0001566159$] a small effect. The confidence intervals associated with the regression analysis for the statistically significant predictors do not contain 0, which means the null hypothesis, there is no association between Surgery, Stage of lung cancer, Socioeconomic status, Race/ ethnicity, Age and Survival/ life expectancy of lung cancer patients, can be rejected.

RQ2

What is the association between chemotherapy and survival/ life expectancy after controlling for surgery, radiotherapy, a combination of therapy (surgery and chemotherapy; surgery and radiotherapy; surgery, chemotherapy, and radiotherapy; or chemotherapy and radiotherapy), stage of lung cancer, socioeconomic status, race/ ethnicity, age, and geographic location (urban vs. rural)?

Descriptive Statistics Results. The study sample (n= 86,998), survival months, mean was 7.56, and standard deviation was 6.664. Chemotherapy, mean was 0.40, and standard deviation was 0.489. Stage of lung cancer, mean 2.84, and standard deviation was 1.265. Socioeconomic status, (median household income), mean 7.40, and standard deviation was 2.589. Race (White and Non-White), mean 1.26, and standard deviation was 0.439. Age, mean 14.62, and standard deviation was 2.061. Geographic location (Urban vs. Rural), mean 1.16, and standard deviation was 0.362 (Table 2, Figure 2).

Table 2

Descriptive Statistics of Survival Months, Chemotherapy, Stage of Lung Cancer, Socioeconomic Status, Race, Age, and Geographic location of Lung Cancer Patients

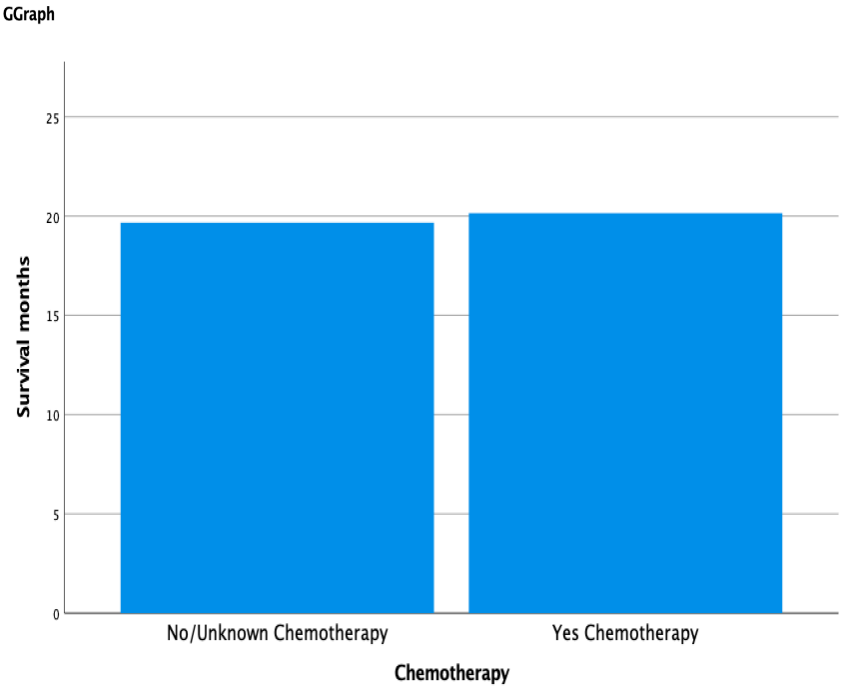
	Mean	**Std. Deviation**	**N**
Survival Months	7.56	6.664	86998
Chemotherapy	0.40	0.489	86998
Stage of Lung Cancer	2.84	1.265	86998
Socioeconomic Status: Median household income	7.40	2.589	86998
Race: White and Non-White	1.26	0.439	86998
Age	14.62	2.061	86998
Geographic Location: Urban vs. Rural	1.16	0.362	86998

Note. Dependent variable: Survival Months

Data Source: SEER Research Plus Data 17 Registries Nov 2021 Sub [2009-2019]

Figure 2

Association Between Chemotherapy and Survival of Lung Cancer Patients



Note. Data Source: SEER Research Plus Data 17 Registries Nov 2021 Sub [2009-2019]

Statistical Assumptions

I used multiple linear regression for my dataset analyses. To obtain valid results, the following eight assumptions were checked and satisfied to qualify my variables to be analyzed using multiple linear regression (Laerd Statistics, 2018):

Assumption-1: The dependent variable (survival/ life expectancy) should be measured on a continuous scale (interval or ratio).

Assumption-2: The Independent variables should be two or more and they can be continuous or categorical (surgery, chemotherapy, radiotherapy, a combination of therapy

[surgery and chemotherapy; surgery and radiotherapy; surgery, chemotherapy, and radiotherapy]; or chemotherapy and radiotherapy),

Assumption-3: The variables should reflect independence of observations (independence of residuals). In other words, variables should not be related to each other or appear as clustered. This can be verified by Durbin-Watson statistic test using SPSS statistics.

Assumption-4: There should be a linear relationship between the dependent variable and each of the independent variables; as well as between the dependent variable and all the independent variables together. This can be verified by different means such as scatterplots and partial regression plots using SPSS statistics.

Assumption-5: The data should reflect homoscedasticity or homogeneity of variance. This is an assumption of similar variances in different groups along the line were compared. This assumption can be tested using SPSS statistics.

Assumption-6: The data should not reflect multicollinearity that takes place when two or more independent variables are drastically correlated with each other. Such multicollinearity creates misunderstanding of identifying the independent variable that contributes to the variance mentioned in the dependent variable. This also creates a technical problem when calculating a multiple regression model. Multicollinearity can also be tested by using SPSS statistics.

Assumption-7: The dataset should not include significant outliers, high leverage points, or highly influential points. Such observations could have different effects on the regression line which in turn negatively affect the equation of regression used to predict the dependent variable value based on the independent variables. As such, the SPSS output

will reflect reduced predictive accuracy of the results as well as of the statistical significance. These observations can be checked while using SPSS statistics for multiple linear regression. Outliers and leverage points can be detected by using *casewise diagnostics and studentized deleted residuals* when using SPSS statistics. The influential points can be checked while using SPSS statistics by testing a measure of influence known as *Cook's Distance.*

Assumption-8: Residuals (errors) should be approximately normally distributed. This can be verified by a superimposed normal curve histogram and a normal P-P Plot or by a normal Q-Q Plot of the studentized residuals. Both methods can be tested by using SPSS statistics.

Inferential Analyses Results. To investigate What is the association between chemotherapy and Survival/ life expectancy of lung cancer patients*,* a simple linear regression was conducted. The predictor was chemotherapy*,* and the outcome was Survival/ life expectancy. The predictor variable was found to be statistically significant [B = 1.672, 95% C.I. (1.583, 1.762), $p < .001$], indicating that for lung cancer patients receiving chemotherapy, *Survival/ life expectancy* increases by approximately 1.7 months compared to those who did not receive chemotherapy. The model explained approximately 2% of the variability [R-squared = 0.015]. A sensitivity analysis using G*power 3.1.9.7 software [n = 86,998; alpha error prob = 0.05; power = 0.80; predictors = 1] calculated [Effect size f^2 = 9.022085e-05] a small effect. The confidence interval associated with the regression analysis for the statistically significant predictors does not

contain 0, which means the null hypothesis, there is no association between chemotherapy and Survival/ life expectancy of lung cancer patients, can be rejected.

To investigate What is the association between Chemotherapy and Survival/ life expectancy of lung cancer patients after controlling for Stage of lung cancer, Socioeconomic status, Race/ ethnicity, Age, and Geographical location (urban vs. rural)?, a multiple linear regression analysis was conducted to evaluate the prediction of Survival/ life expectancy of lung cancer patients from Chemotherapy, Stage of lung cancer, Socioeconomic status, Race/ ethnicity, Age, and Geographical location (urban vs. rural). The results of the multiple linear regression analysis revealed *Geographical location* (urban vs. rural) not to be a statistically significant predictor to the model ($p > .05$). However, the results of the multiple linear regression analysis revealed a statistically significant association between Chemotherapy, Stage of lung cancer, Socioeconomic status, Race/ ethnicity, and Age.

Controlling for Stage of lung cancer, Socioeconomic status, Race/ ethnicity, and Age, the regression coefficient [B = 3.377, 95% C.I. (3.287, 3.467) $p < .001$] associated with Chemotherapy suggests that for lung cancer patients receiving Chemotherapy, Survival/ life expectancy increases by approximately 3.4 months.

Controlling for Chemotherapy, Socioeconomic status, Race/ ethnicity, and Age, the regression coefficient [B = -2.161, 95% C.I. (-2.195, -2.127) $p < .001$] associated with Stage of lung cancer, suggests that with each increase of lung cancer stage (Stage 1, Stage 2, Stage 3, and Stage 4), Survival/ life expectancy decreases by approximately 2.2 months.

Controlling for Chemotherapy, Stage of lung cancer, Race/ ethnicity, and Age, the regression coefficient [B = 0.112, 95% C.I. (0.093, 0.131) p < .001] associated with Socioeconomic status, suggests that for median household income level increase, Survival/ life expectancy increases by approximately 0.1 month or 3 days.

Controlling for Chemotherapy, Stage of lung cancer, Socioeconomic status, and Age, the regression coefficient [B = 0.139, 95% C.I. (0.046, 0.233) p = .003] associated with Race/ ethnicity suggests that Whites compared to non-Whites, have Survival/ life expectancy increases by approximately 0.1 month or 3 days.

Controlling for Chemotherapy, Stage of lung cancer, Socioeconomic status, and Race/ ethnicity, the regression coefficient [B = -0.199, 95% C.I. (-0.219, -0.178) p < .001] associated with Age, suggests that with each additional five years of age, *Survival/ life expectancy* decreases by approximately 0.2 month or 6 days.

Controlling for Chemotherapy, Stage of lung cancer, Socioeconomic status, Race/ ethnicity, and Age, the regression coefficient [B = 3.377, 95% C.I. (3.287, 3.467) p < .05] associated with [IV2] suggests that with each additional [IV2], Survival/ life expectancy increases by approximately 0.2 month or 6 days. The R-squared value of [0.172] associated with this regression model suggests that the association between Chemotherapy, Stage of lung cancer, Socioeconomic status, Race/ ethnicity, and Age account for [17.2%] of the variation in *Survival/ life expectancy of lung cancer patients*, which means that [82.8%] of the variation in *Survival/ life expectancy of lung cancer patients* cannot be explained by *Chemotherapy, Stage of lung cancer, Socioeconomic status, Race/ ethnicity,* and *Age* alone. A sensitivity analysis using G*power 3.1.9.7

software [n = 86,998; alpha error prob = .05; power = 0.80; predictors = 6] calculated [Effect size f^2 = 0.0001566159] a small effect. The confidence interval associated with the regression analysis for the statistically significant predictors does not contain 0, which means the null hypothesis, there is no association between Chemotherapy, Stage of lung cancer, Socioeconomic status, Race/ ethnicity, Age and Survival/ life expectancy of lung cancer patients, can be rejected.

RQ3

What is the association between radiotherapy and survival/ life expectancy after controlling for surgery, chemotherapy, a combination of therapy (surgery and chemotherapy; surgery and radiotherapy; surgery, chemotherapy, and radiotherapy; or chemotherapy and radiotherapy), stage of lung cancer, socioeconomic status, race/ ethnicity, age, and geographic location (urban vs. rural)?

Descriptive Statistics Results. The study sample (n= 86,998), survival months, mean was 7.56, and standard deviation was 6.664. Radiotherapy*,* mean was 0.39, and standard deviation was 0.487. Stage of lung cancer, mean 2.84, and standard deviation was 1.265. Socioeconomic status, (median household income), mean 7.40, and standard deviation was 2.589. Race (White and Non-White), mean 1.26, and standard deviation was 0.439. *Age,* mean 14.62, and standard deviation was 2.061. Geographic location (Urban vs. Rural), mean 1.16, and standard deviation was 0.362 (Table 3, Figure 3).

Table 3

Descriptive Statistics of Survival Months, Radiation, Stage of Lung Cancer, Socioeconomic Status, Race, Age, and Geographic Location of Lung Cancer Patients

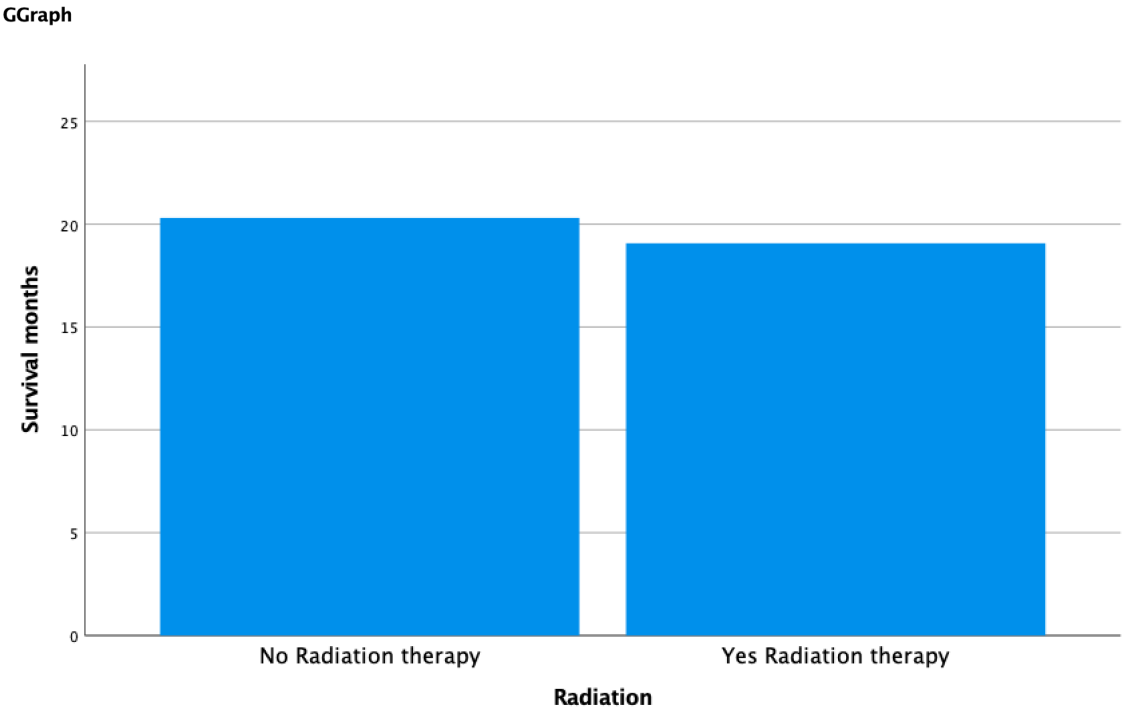
	Mean	**Std. Deviation**	**N**
Survival Months	7.56	6.664	86998
Radiation	0.39	0.487	86998
Stage of Lung Cancer	2.84	1.265	86998
Socioeconomic Status: Median household income	7.40	2.589	86998
Race: White and Non-White	1.26	0.439	86998
Age	14.62	2.061	86998
Geographic Location: Urban vs. Rural	1.16	0.362	86998

Note. Dependent variable: Survival Months.

Data Source: SEER Research Plus Data 17 Registries Nov 2021 Sub [2009-2019].

Figure 3

Association Between Radiotherapy and Survival of Lung Cancer Patients



Note. Data Source: SEER Research Plus Data 17 Registries Nov 2021 Sub [2009-2019].

Statistical Assumptions

I used multiple linear regression for my dataset analyses. To obtain valid results, the following eight assumptions were checked and satisfied to qualify my variables to be analyzed using multiple linear regression (Laerd Statistics, 2018):

Assumption-1: The dependent variable (survival/ life expectancy) should be measured on a continuous scale (interval or ratio).

Assumption-2: The Independent variables should be two or more and they can be continuous or categorical (surgery, chemotherapy, radiotherapy, a combination of therapy

[surgery and chemotherapy; surgery and radiotherapy; surgery, chemotherapy, and radiotherapy]; or chemotherapy and radiotherapy),

Assumption-3: The variables should reflect independence of observations (independence of residuals). In other words, variables should not be related to each other or appear as clustered. This can be verified by Durbin-Watson statistic test using SPSS statistics.

Assumption-4: There should be a linear relationship between the dependent variable and each of the independent variables; as well as between the dependent variable and all the independent variables together. This can be verified by different means such as scatterplots and partial regression plots using SPSS statistics.

Assumption-5: The data should reflect homoscedasticity or homogeneity of variance. This is an assumption of similar variances in different groups along the line were compared. This assumption can be tested using SPSS statistics.

Assumption-6: The data should not reflect multicollinearity that takes place when two or more independent variables are drastically correlated with each other. Such multicollinearity creates misunderstanding of identifying the independent variable that contributes to the variance mentioned in the dependent variable. This also creates a technical problem when calculating a multiple regression model. Multicollinearity can also be tested by using SPSS statistics.

Assumption-7: The dataset should not include significant outliers, high leverage points, or highly influential points. Such observations could have different effects on the regression line which in turn negatively affect the equation of regression used to predict the dependent variable value based on the independent variables. As such, the SPSS output

will reflect reduced predictive accuracy of the results as well as of the statistical significance. These observations can be checked while using SPSS statistics for multiple linear regression. Outliers and leverage points can be detected by using *casewise diagnostics and studentized deleted residuals* when using SPSS statistics. The influential points can be checked while using SPSS statistics by testing a measure of influence known as *Cook's Distance.*

Assumption-8: Residuals (errors) should be approximately normally distributed. This can be verified by a superimposed normal curve histogram and a normal P-P Plot or by a normal Q-Q Plot of the studentized residuals. Both methods can be tested by using SPSS statistics.

Inferential Analyses Results. To investigate What is the association between Radiotherapy, and Survival/ life expectancy of lung cancer patients*,* a simple linear regression was conducted. The predictor was *Radiotherapy,* and the outcome was Survival/ life expectancy. The predictor variable was found to be statistically significant [B = 1.529, 95% C.I. (1.439, 1.620), $p < .001$], indicating that for lung cancer patients receiving Radiotherapy, *Survival/ life expectancy* increases by approximately 1.5 months compared to those who did not receive Radiotherapy. The model explained approximately 1% of the variability [R-squared = 0.013]. A sensitivity analysis using G*power 3.1.9.7 software [n = 86,998; alpha error prob = 0.05; power = 0.80; predictors = 1] calculated [Effect size f^2 = 9.022085e-05] a small effect. The confidence interval associated with the regression analysis for the statistically significant predictors does not

contain 0, which means the null hypothesis, there is no association between Radiotherapy and Survival/ life expectancy of lung cancer patients, can be rejected.

To investigate What is the association between Radiotherapy and Survival/ life expectancy of lung cancer patients after controlling for Stage of lung cancer, Socioeconomic status, Race/ ethnicity, Age, and Geographical location (urban vs. rural)?, a multiple linear regression analysis was conducted to evaluate the prediction of Survival/ life expectancy of lung cancer patients from Radiotherapy, Stage of lung cancer, Socioeconomic status, Race/ ethnicity, Age, and Geographical location (urban vs. rural). The results of the multiple linear regression analysis revealed Geographical location (urban vs. rural) not to be a statistically significant predictor to the model ($p > .05$). However, the results of the multiple linear regression analysis revealed a statistically significant association between Radiotherapy, Stage of lung cancer, Socioeconomic status, Race/ ethnicity, and Age.

Controlling for Stage of lung cancer, Socioeconomic status, Race/ ethnicity, and Age, the regression coefficient [$B = 1.638$, 95% C.I. (1.554, 1.723) $p < .001$] associated with Radiotherapy suggests that for lung cancer patients receiving Radiotherapy, Survival/ life expectancy increases by approximately 1.6 months compared to those who did not receive Radiotherapy.

Controlling for Radiotherapy, Socioeconomic status, Race/ ethnicity, and Age, the regression coefficient [$B = -1.748$, 95% C.I. (-1.781, -1.716) $p < .001$] associated with Stage of lung cancer, suggests that with each increase of lung cancer stage (Stage 1,

Stage 2, Stage 3, and Stage 4), Survival/ life expectancy decreases by approximately 1.8 months.

Controlling for Radiotherapy, Stage of lung cancer, Race/ ethnicity, and Age, the regression coefficient [B = 0.139, 95% C.I. (0.120, 0.159) p < .001] associated with Socioeconomic status*,* suggests that for median household income level increase, Survival/ life expectancy increases by approximately 0.1 month or 3 days compared to those whose median household income level do not increase.

Controlling for Radiotherapy, Stage of lung cancer, Socioeconomic status, and Age, the regression coefficient [B = 0.152, 95% C.I. (0.056, 0.247) p = .002] associated with *Race/ ethnicity,* suggests that Whites compared to non-Whites, have Survival/ life expectancy increases by approximately 0.2 month or 6 days.

Controlling for Radiotherapy, Stage of lung cancer, Socioeconomic status, and Race/ ethnicity, the regression coefficient [B = -0.352, 95% C.I. (-0.373, -0.332) p < .001] associated with Age, suggests that with each additional five years of age, Survival/ life expectancy decreases by approximately 0.4 month or 12 days.

Controlling for Radiotherapy, Stage of lung cancer, Socioeconomic status, Race/ ethnicity, and Age, the regression coefficient [B = 1.638, 95% C.I. (1.554, 1.723) p < .05] associated with [IV2] suggests that with each additional [IV2], Survival/ life expectancy increases by approximately 0.1 month or 3 days. The R-squared value of [0.134] associated with this regression model suggests that the association between Radiotherapy, Stage of lung cancer, Socioeconomic status, Race/ ethnicity, and Age accounts for [13.4%] of the variation in Survival/ life expectancy of lung cancer patients, which means

that [86.6%] of the variation in Survival/ life expectancy of lung cancer patients cannot be explained by Radiotherapy, Stage of lung cancer, Socioeconomic status, Race/ ethnicity, and Age alone. A sensitivity analysis using G*power 3.1.9.7 software [n = 86,998; alpha error prob = 0.05; power = 0.80; predictors = 6] calculated [Effect size f^2 = 0.0001566159] a small effect. The confidence interval associated with the regression analysis for the statistically significant predictors does not contain 0, which means the null hypothesis, there is no association between Radiotherapy, Stage of lung cancer, Socioeconomic status, Race/ ethnicity, Age and Survival/ life expectancy of lung cancer patients, can be rejected.

RQ4

What is the association between a combination of therapy (surgery and chemotherapy; surgery and radiotherapy; surgery, chemotherapy, and radiotherapy; or chemotherapy and radiotherapy) and survival/ life expectancy after controlling for surgery, chemotherapy, radiotherapy, stage of lung cancer, socioeconomic status, race/ ethnicity, age, and geographic location (urban vs. rural)?

Descriptive Statistics Results. The study sample (n= 86,998), survival months, mean was 7.56, and standard deviation was 6.664. Combination of Surgery and Chemotherapy, mean was 0.04, and standard deviation was 0.199. *Stage of lung cancer,* mean 2.84, and standard deviation was 1.265. Socioeconomic status, (median household income), mean 7.40, and standard deviation was 2.589. Race (White and Non-White), mean 1.26, and standard deviation was 0.439. Age, mean 14.62, and standard deviation

was 2.061. Geographic location (Urban vs. Rural), mean 1.16, and standard deviation was 0.362 (Table 4, Figure 4).

Table 4

Descriptive Statistics of Survival Months, Surgery and Chemotherapy, Stage of Lung Cancer, Socioeconomic Status, Race, Age, and Geographic Location of Lung Cancer Patients

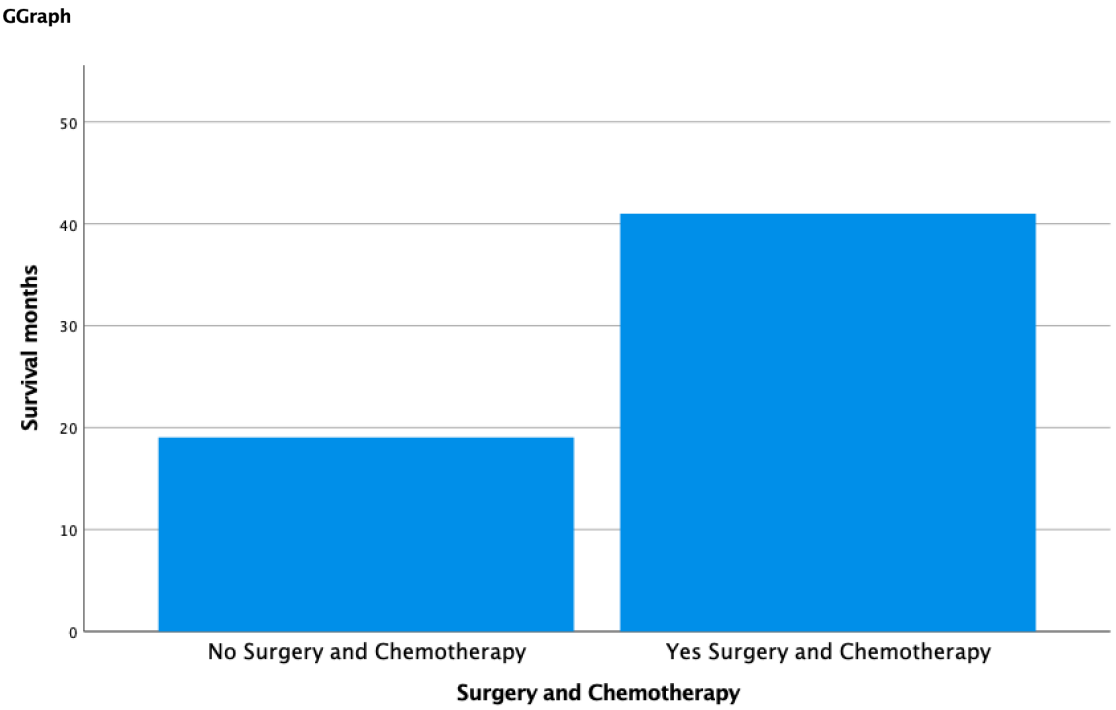
	Mean	**Std. Deviation**	**N**
Survival Months	7.56	6.664	86998
Surgery and Chemotherapy	0.04	0.199	86998
Stage of Lung Cancer	2.84	1.265	86998
Socioeconomic Status: Median household income	7.40	2.589	86998
Race: White and Non-White	1.26	0.439	86998
Age	14.62	2.061	86998
Geographic Location: Urban vs. Rural	1.16	0.362	86998

Note. Dependent variable: Survival Months.

Data Source: SEER Research Plus Data 17 Registries Nov 2021 Sub [2009-2019].

Figure 4

Association Between Combination of Surgery/Chemotherapy and Survival of Lung Cancer Patients



Note. Data Source: SEER Research Plus Data 17 Registries Nov 2021 Sub [2009-2019].

Statistical Assumptions

I used multiple linear regression for my dataset analyses. To obtain valid results, the following eight assumptions were checked and satisfied to qualify my variables to be analyzed using multiple linear regression (Laerd Statistics, 2018):

Assumption-1: The dependent variable (survival/ life expectancy) should be measured on a continuous scale (interval or ratio).

Assumption-2: The Independent variables should be two or more and they can be continuous or categorical (surgery, chemotherapy, radiotherapy, a combination of therapy [surgery and chemotherapy; surgery and radiotherapy; surgery, chemotherapy, and radiotherapy]; or chemotherapy and radiotherapy),

Assumption-3: The variables should reflect independence of observations (independence of residuals). In other words, variables should not be related to each other or appear as clustered. This can be verified by Durbin-Watson statistic test using SPSS statistics.

Assumption-4: There should be a linear relationship between the dependent variable and each of the independent variables; as well as between the dependent variable and all the independent variables together. This can be verified by different means such as scatterplots and partial regression plots using SPSS statistics.

Assumption-5: The data should reflect homoscedasticity or homogeneity of variance. This is an assumption of similar variances in different groups along the line were compared. This assumption can be tested using SPSS statistics.

Assumption-6: The data should not reflect multicollinearity that takes place when two or more independent variables are drastically correlated with each other. Such multicollinearity creates misunderstanding of identifying the independent variable that contributes to the variance mentioned in the dependent variable. This also creates a technical problem when calculating a multiple regression model. Multicollinearity can also be tested by using SPSS statistics.

Assumption-7: The dataset should not include significant outliers, high leverage points, or highly influential points. Such observations could have different effects on the regression

line which in turn negatively affect the equation of regression used to predict the dependent variable value based on the independent variables. As such, the SPSS output will reflect reduced predictive accuracy of the results as well as of the statistical significance. These observations can be checked while using SPSS statistics for multiple linear regression. Outliers and leverage points can be detected by using *casewise diagnostics and studentized deleted residuals* when using SPSS statistics. The influential points can be checked while using SPSS statistics by testing a measure of influence known as *Cook's Distance*.

Assumption-8: Residuals (errors) should be approximately normally distributed. This can be verified by a superimposed normal curve histogram and a normal P-P Plot or by a normal Q-Q Plot of the studentized residuals. Both methods can be tested by using SPSS statistics.

Inferential Analyses Results. To investigate What is the association between a combination of Surgery and Chemotherapy and Survival/ life expectancy of lung cancer patients, a simple linear regression was conducted. The predictor was a combination of Surgery and Chemotherapy, and the outcome was Survival/ life expectancy. The predictor variable was found to be statistically significant [B = 3.340, 95% C.I. (3.119, 3.562), $p < .001$], indicating that for lung cancer patients receiving a combination of Surgery and Chemotherapy, Survival/ life expectancy increases by approximately 3.3 months compared to those who did not receive a combination of Surgery and Chemotherapy. The model explained approximately 1% of the variability [R-squared = 0.010]. A sensitivity analysis using G*power 3.1.9.7 software [n = 86,998; alpha error

prob = 0.05; power = 0.80; predictors = 1] calculated [Effect size f^2 = 9.022085e-05] a small effect. The confidence interval associated with the regression analysis for the statistically significant predictors does not contain 0, which means the null hypothesis, there is no association between a combination of Surgery and Chemotherapy and Survival/ life expectancy of lung cancer patients, can be rejected.

To investigate What is the association between a combination of Surgery and Chemotherapy and Survival/ life expectancy of lung cancer patients after controlling for Stage of lung cancer, Socioeconomic status, Race/ ethnicity, Age, and Geographical location (urban vs. rural)?, a multiple linear regression analysis was conducted to evaluate the prediction of Survival/ life expectancy of lung cancer patients from a combination of Surgery and Chemotherapy, Stage of lung cancer, Socioeconomic status, Race/ ethnicity, Age, and Geographical location (urban vs. rural). The results of the multiple linear regression analysis revealed Geographical location (urban vs. rural) not to be a statistically significant predictor to the model ($p > .05$). However, the results of the multiple linear regression analysis revealed a statistically significant association between a combination of Surgery and Chemotherapy, Stage of lung cancer, Socioeconomic status, Race/ ethnicity, and Age.

Controlling for Stage of lung cancer, Socioeconomic status, Race/ ethnicity, and Age, the regression coefficient [B = 2.093, 95% C.I. (1.883, 2.303) $p < .001$] associated with a combination of Surgery and Chemotherapy suggests that for lung cancer patients receiving a combination of Surgery and Chemotherapy, Survival/ life expectancy

increases by approximately 2.1 months compared to those who did not receive a combination of Surgery and Chemotherapy.

Controlling for a combination of Surgery and Chemotherapy, Socioeconomic status, Race/ ethnicity, and Age, the regression coefficient [B = -1.701, 95% C.I. (-1.734, -1.668) p < .001] associated with Stage of lung cancer, suggests that with each increase of lung cancer stage (Stage 1, Stage 2, Stage 3, and Stage 4), Survival/ life expectancy decreases by approximately 1.7 months.

Controlling for a combination of Surgery and Chemotherapy, Stage of lung cancer, Race/ ethnicity, and Age, the regression coefficient [B = 0.119, 95% C.I. (0.100, 0.139) *p* < .001] associated with *Socioeconomic status,* suggests that for median household income level increase, Survival/ life expectancy increases by approximately 0.1 month or 3 days.

Controlling for a combination of Surgery and Chemotherapy, Stage of lung cancer, Socioeconomic status, and Age, the regression coefficient [B = 0.092, 95% C.I. (0.004, 0.188) p = .060] associated with Race/ ethnicity, suggests that Whites compared to non-Whites, not to be a statistically significant predictor to the model (p > .05).

Controlling for a combination of Surgery and Chemotherapy, Stage of lung cancer, Socioeconomic status, and Race/ ethnicity, the regression coefficient [B = -0.353, 95% C.I. (-0.374, -0.333) p < .001] associated with Age, suggests that with each additional five years of age, Survival/ life expectancy decreases by approximately 0.4 month or 12 days.

Controlling for a combination of Surgery and Chemotherapy, Stage of lung cancer, Socioeconomic status, Race/ ethnicity, and Age, the regression coefficient [B = 2.093, 95% C.I. (1.883, 2.303) $p < .05$] associated with [IV2] suggests that with each additional [IV2], Survival/ life expectancy increases by approximately 0.1 months or 3 days. The R-squared value of [0.124] associated with this regression model suggests that the association between a combination of Surgery and Chemotherapy, Stage of lung cancer, Socioeconomic status, Race/ ethnicity, and Age account for [12.4%] of the variation in Survival/ life expectancy of lung cancer patients, which means that [87.6%] of the variation in Survival/ life expectancy of lung cancer patients cannot be explained by a combination of Surgery and Chemotherapy, Stage of lung cancer, Socioeconomic status, Race/ ethnicity, and Age alone. A sensitivity analysis using G*power 3.1.9.7 software [n = 86,998; alpha error prob = 0.05; power = 0.80; predictors = 6] calculated [Effect size $f^2 = 0.0001566159$] a small effect. The confidence interval associated with the regression analysis does not contain 0, which means the null hypothesis, there is no association between a combination of Surgery and Chemotherapy, Stage of lung cancer, Socioeconomic status, Race/ ethnicity, Age and Survival/ life expectancy of lung cancer patients, can be rejected.

Descriptive Statistics Results. The study sample (n= 86,998), survival months, mean was 7.56, and standard deviation was 6.664. Combination of Surgery and Radiotherapy, mean was 0.00, and standard deviation was 0.069. Stage of lung cancer, mean 2.84, and standard deviation was 1.265. Socioeconomic status, (median household income), mean 7.40, and standard deviation was 2.589. Race (White and Non-White),

mean 1.26, and standard deviation was 0.439. *Age,* mean 14.62, and standard deviation was 2.061. Geographic location (Urban vs. Rural), mean 1.16, and standard deviation was 0.362 (Table 5, Figure 5).

Table 5

Descriptive Statistics of Survival Months, Surgery and Radiation, Stage of Lung Cancer, Socioeconomic Status, Race, Age, and Geographic Location of Lung Cancer Patients

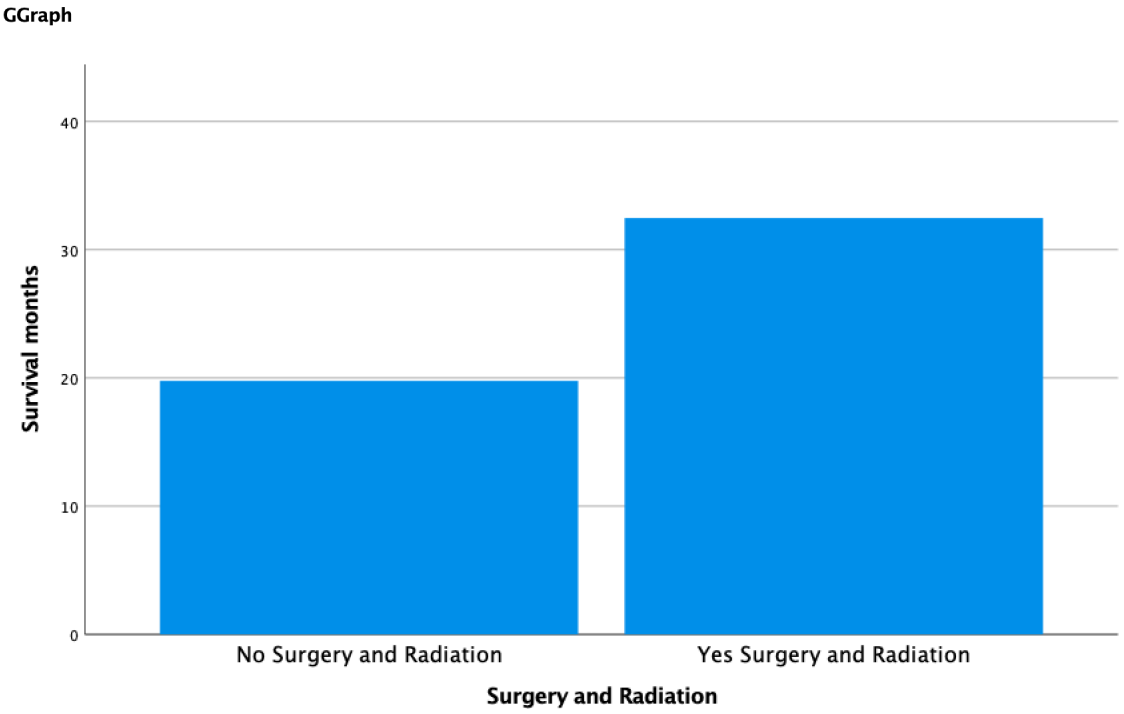
	Mean	**Std. Deviation**	**N**
Survival Months	7.56	6.664	86998
Surgery and Radiation	0.00	0.069	86998
Stage of Lung Cancer	2.84	1.265	86998
Socioeconomic Status: Median household income	7.40	2.589	86998
Race: White and Non-White	1.26	0.439	86998
Age	14.62	2.061	86998
Geographic Location: Urban vs. Rural	1.16	0.362	86998

Note. Dependent variable: Survival Months.

Data Source: SEER Research Plus Data 17 Registries Nov 2021 Sub [2009-2019].

Figure 5

Association Between Combination of Surgery/Radiation and Survival of Lung Cancer Patients



Note. Data Source: SEER Research Plus Data 17 Registries Nov 2021 Sub [2009-2019].

Inferential Analyses Results. To investigate What is the association between a combination of Surgery and Radiotherapy and Survival/ life expectancy of lung cancer patients, a simple linear regression was conducted. The predictor was a combination of Surgery and Radiotherapy, and the outcome was Survival/ life expectancy. The predictor variable was found to be statistically significant [B = 2.046, 95% C.I. (1.404, 2.689), $p < .001$], indicating that for lung cancer patients receiving a combination of Surgery and Radiotherapy, Survival/ life expectancy increases by approximately 2.1 months. The

model explained approximately 0% of the variability [R-squared = 0.000]. A sensitivity analysis using G*power 3.1.9.7 software [n = 86,998; alpha error prob = 0.05; power = 0.80; predictors = 1] calculated [Effect size f^2 = 9.022085e-05] a small effect. The confidence interval associated with the regression analysis does not contain 0, which means the null hypothesis, there is no association between a combination of Surgery and Radiotherapy and Survival/ life expectancy of lung cancer patients, can be rejected.

To investigate What is the association between a combination of Surgery and Radiotherapy and Survival/ life expectancy of lung cancer patients after controlling for Stage of lung cancer, Socioeconomic status, Race/ ethnicity, Age, and Geographical location (urban vs. rural)?, a multiple linear regression analysis was conducted to evaluate the prediction of Survival/ life expectancy of lung cancer patients from a combination of Surgery and Radiotherapy, Stage of lung cancer, Socioeconomic status, Race/ ethnicity, Age, and Geographical location (urban vs. rural). The results of the multiple linear regression analysis revealed Geographical location (urban vs. rural) not to be a statistically significant predictor to the model ($p > .05$). However, the results of the multiple linear regression analysis revealed a statistically significant association between a combination of Surgery and Radiotherapy, Stage of lung cancer, Socioeconomic status, Race/ ethnicity, and Age.

Controlling for Stage of lung cancer, Socioeconomic status, Race/ ethnicity, and Age, the regression coefficient [B = 2.046, 95% C.I. (1.404, 2.689) $p < .001$] associated with a combination of Surgery and Radiotherapy suggests that for lung cancer patients receiving a combination of Surgery and Radiotherapy Survival/ life expectancy increases

by approximately 2.1 months compared to those who did not receive a combination of Surgery and Chemotherapy.

Controlling a combination of Surgery and Radiotherapy, Socioeconomic status, Race/ ethnicity, and Age, the regression coefficient [B = -1.729, 95% C.I. (-1.762, -1.696) p < .001] associated with Stage of lung cancer, suggests that with each increase of lung cancer stage (Stage 1, Stage 2, Stage 3, and Stage 4), Survival/ life expectancy decreases by approximately 1.7 months.

Controlling for a combination of Surgery and Radiotherapy, Stage of lung cancer, Race/ ethnicity, and Age, the regression coefficient [B = 0.124, 95% C.I. (0.105, 0.144) p < .001] associated with Socioeconomic status, suggests that for median household income level increase, Survival/ life expectancy increases by approximately 0.1 month or 3 days.

Controlling for a combination of Surgery and Radiotherapy, Stage of lung cancer, Socioeconomic status, and Age, the regression coefficient [B = 0.090, 95% C.I. (-0.007, 0.186) p = .068] associated with *Race/ ethnicity* suggests that Whites compared to non-Whites not to be a statistically significant predictor to the model (p > .05).

Controlling for a combination of Surgery and Radiotherapy, Stage of lung cancer, Socioeconomic status, and Race/ ethnicity, the regression coefficient [B = -0.369, 95% C.I. (-0.389, -0.349) p < .001] associated with *Age*, suggests that with each additional five years of age, *Survival/ life expectancy* decreases by approximately 0.4 month or 12 days.

Controlling for a combination of Surgery and Radiotherapy, Stage of lung cancer, Socioeconomic status, Race/ ethnicity, and Age, the regression coefficient [B = 1.262,

95% C.I. (0.658, 1.865) p < .05] associated with [IV2] suggests that with each additional [IV2], Survival/ life expectancy increases by approximately 0.1 month or 3 days. The R-squared value of [0.120] associated with this regression model suggests that the association between a combination of Surgery and Radiotherapy, Stage of lung cancer, Socioeconomic status, Race/ ethnicity, and Age account for [12%] of the variation in Survival/ life expectancy of lung cancer patients, which means that [88%] of the variation in Survival/ life expectancy of lung cancer patients cannot be explained by a combination of Surgery and Radiotherapy, Stage of lung cancer, Socioeconomic status, Race/ ethnicity, and Age alone. A sensitivity analysis using G*power 3.1.9.7 software [n = 86,998; alpha error prob = 0.05; power = 0.80; predictors = 6] calculated [Effect size f^2 = 0.0001566159] a small effect. The confidence interval associated with the regression analysis does not contain 0, which means the null hypothesis, there is no association between a combination of Surgery and Radiotherapy, Stage of lung cancer, Socioeconomic status, Race/ ethnicity, Age and Survival/ life expectancy of lung cancer patients, can be rejected.

To investigate What is the association between a combination of Surgery and Radiotherapy and Survival/ life expectancy of lung cancer patients, a simple linear regression was conducted. The predictor was a combination of Surgery and Radiotherapy, and the outcome was Survival/ life expectancy. The predictor variable was found to be statistically significant [B = 2.046, 95% C.I. (1.404, 2.689), p < .001], indicating that for lung cancer patients receiving a combination of Surgery and Radiotherapy, Survival/ life expectancy increases by approximately 2.1 months. The

model explained approximately 0% of the variability [R-squared = 0.000]. A sensitivity analysis using G*power 3.1.9.7 software [n = 86,998; alpha error prob = 0.05; power = 0.80; predictors = 1] calculated [Effect size f^2 = 9.022085e-05] a small effect. The confidence interval associated with the regression analysis for the statistically significant predictors does not contain 0, which means the null hypothesis, there is no association between a combination of Surgery and Radiotherapy and Survival/ life expectancy of lung cancer patients, can be rejected.

Descriptive Statistics Results. The study sample (n= 86,998), survival months, mean was 7.56, and standard deviation was 6.664. Combination of Surgery, Chemotherapy, and Radiotherapy, mean was 0.02, and standard deviation was 0.134. Stage of lung cancer, mean 2.84, and standard deviation was 1.265. Socioeconomic status, (median household income), mean 7.40, and standard deviation was 2.589. Race (White and Non-White), mean 1.26, and standard deviation was 0.439. *Age,* mean 14.62, and standard deviation was 2.061. Geographic location (Urban vs. Rural), mean 1.16, and standard deviation was 0.362 (Table 6, Figure 6).

Table 6

Descriptive Statistics of Survival Months, Surgery, Radiation and Chemotherapy, Stage of Lung Cancer, Socioeconomic Status, Race, Age, and Geographic Location of Lung Cancer Patients

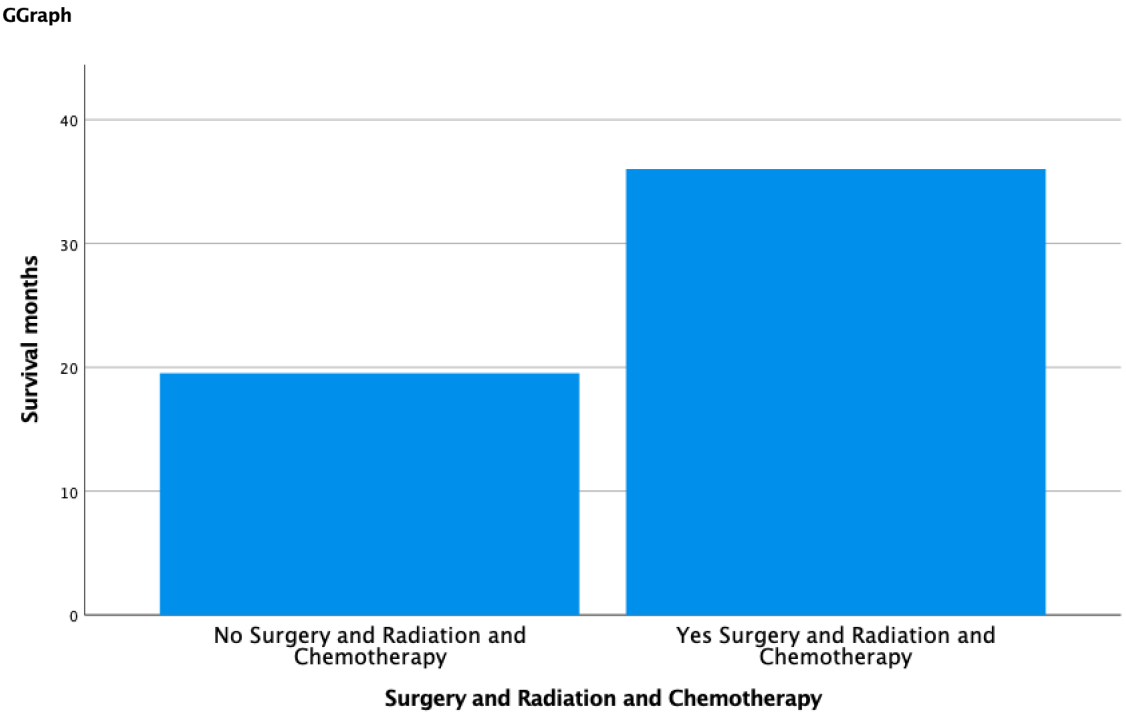
	Mean	**Std. Deviation**	**N**
Survival Months	7.56	6.664	86998
Surgery and Radiation and Chemotherapy	0.02	0.134	86998
Stage of Lung Cancer	2.84	1.265	86998
Socioeconomic Status: Median household income	7.40	2.589	86998
Race: White and Non-White	1.26	0.439	86998
Age	14.62	2.061	86998
Geographic Location: Urban vs. Rural	1.16	0.362	86998

Note. Dependent variable: Survival Months.

Data Source: SEER Research Plus Data 17 Registries Nov 2021 Sub [2009-2019].

Figure 6

Association Between a Combination of Surgery, Radiation, and Chemotherapy, and Survival of Lung Cancer Patients



Note. Data Source: SEER Research Plus Data 17 Registries Nov 2021 Sub [2009-2019].

Inferential Analyses Results. To investigate What is the association between a combination of Surgery, Chemotherapy, and Radiotherapy and Survival/ life expectancy of lung cancer patients, a simple linear regression was conducted. The predictor was a combination of Surgery, Chemotherapy, and Radiotherapy, and the outcome was Survival/ life expectancy. The predictor variable was found to be statistically significant [B = 3.447, 95% C.I. (3.118, 3.776), $p < .001$], indicating that for lung cancer patients receiving a combination of Surgery, Chemotherapy, and Radiotherapy, Survival/ life

expectancy increases by approximately 0.01 month or 3 days compared to those who did not receive a combination of Surgery, Chemotherapy, and Radiotherapy. The model explained approximately 0% of the variability [R-squared = 0.005]. A sensitivity analysis using G*power 3.1.9.7 software [n = 86,998; alpha error prob = 0.05; power = 0.80; predictors = 1] calculated [Effect size f^2 = 9.022085e-05] a small effect. The confidence interval associated with the regression analysis for the statistically significant predictors does not contain 0, which means the null hypothesis, there is no association between a combination of Surgery, Chemotherapy, and Radiotherapy, and Survival/ life expectancy of lung cancer patients, can be rejected.

To investigate What is the association between a combination of Surgery, Chemotherapy, and Radiotherapy and Survival/ life expectancy of lung cancer patients after controlling for Stage of lung cancer, Socioeconomic status, Race/ ethnicity, Age, and Geographical location (urban vs. rural)?, a multiple linear regression analysis was conducted to evaluate the prediction of Survival/ life expectancy of lung cancer patients from a combination of Surgery, Chemotherapy, and Radiotherapy, Stage of lung cancer, Socioeconomic status, Race/ ethnicity, Age, and Geographical location (urban vs. rural). The results of the multiple linear regression analysis revealed Geographical location (urban vs. rural) not to be a statistically significant predictor to the model ($p > .05$). However, the results of the multiple linear regression analysis revealed a statistically significant association between a combination of Surgery, Chemotherapy, and Radiotherapy, Stage of lung cancer, Socioeconomic status, Race/ ethnicity, and Age.

Controlling for Stage of lung cancer, Socioeconomic status, Race/ ethnicity, and Age, the regression coefficient [B = 3.180, 95% C.I. (2.871, 3.489) $p < .001$] associated with a combination of Surgery, Chemotherapy, and Radiotherapy suggests that for lung cancer patients receiving a combination of Surgery, Chemotherapy, and Radiotherapy Survival/ life expectancy increases by approximately 3.2 months.

Controlling a combination of Surgery, Chemotherapy, and Radiotherapy, Socioeconomic status, Race/ ethnicity, and Age, the regression coefficient [B = -1.731, 95% C.I. (-1.764, -1.698) $p < .001$] associated with Stage of lung cancer, suggests that with each increase of lung cancer stage (Stage 1, Stage 2, Stage 3, and Stage 4), Survival/ life expectancy decreases by approximately 1.7 months.

Controlling for a combination of Surgery, Chemotherapy, and Radiotherapy, Stage of lung cancer, Race/ ethnicity, and Age, the regression coefficient [B = 0.124, 95% C.I. (0.104, 0.143) $p < .001$] associated with Socioeconomic status, suggests that for median household income level increase, Survival/ life expectancy increases by approximately 0.1 month or 3 days.

Controlling for a combination of Surgery, Chemotherapy, and Radiotherapy, Stage of lung cancer, Socioeconomic status, and Age, the regression coefficient [B = 0.094, 95% C.I. (-0.002, 0.190) $p = .054$] associated with Race/ ethnicity, suggests that Whites compared to non-Whites not to be a statistically significant predictor to the model ($p > .05$).

Controlling for a combination of Surgery, Chemotherapy, and Radiotherapy, Stage of lung cancer, Socioeconomic status, and Race/ ethnicity, the regression

coefficient [B = -0.356, 95% C.I. (-0.376, -0.336) $p < .001$] associated with Age, suggests that with each additional five years of age, Survival/ life expectancy decreases by approximately 0.4 month or 12 days.

Controlling for a combination of Surgery, Chemotherapy, and Radiotherapy, Stage of lung cancer, Socioeconomic status, Race/ ethnicity, and Age, the regression coefficient [B = 3.180, 95% C.I. (2.871, 3.489) $p < .05$] associated with [IV2] suggests that with each additional [IV2], Survival/ life expectancy increases by approximately 0.1 month or 3 days. The R-squared value of [0.124] associated with this regression model suggests that the association between a combination of Surgery, Chemotherapy, and Radiotherapy, Stage of lung cancer, Socioeconomic status, Race/ ethnicity, and Age account for [12.4%] of the variation in Survival/ life expectancy of lung cancer patients, which means that [87.6%] of the variation in Survival/ life expectancy of lung cancer patients cannot be explained by a combination of Surgery, Chemotherapy, and Radiotherapy, Stage of lung cancer, Socioeconomic status, Race/ ethnicity, and Age alone. A sensitivity analysis using G*power 3.1.9.7 software [n = 86,998; alpha error prob = 0.05; power = 0.80; predictors = 6] calculated [Effect size $f^2 = 0.0001566159$] a small effect. The confidence interval associated with the regression analysis for the statistically significant predictors does not contain 0, which means the null hypothesis, there is no association between a combination of Surgery, Chemotherapy, and Radiotherapy, Stage of lung cancer, Socioeconomic status, Race/ ethnicity, Age and Survival/ life expectancy of lung cancer patients, can be rejected.

Descriptive Statistics Results. The study sample (n= 86,998), survival months, mean was 7.56, and standard deviation was 6.664. Combination of Chemotherapy and Radiotherapy, mean was 0.20, and standard deviation was 0.401. Stage of lung cancer, mean 2.84, and standard deviation was 1.265. Socioeconomic status, (median household income), mean 7.40, and standard deviation was 2.589. Race (White and Non-White), mean 1.26, and standard deviation was 0.439. Age, mean 14.62, and standard deviation was 2.061. Geographic location (Urban vs. Rural), mean 1.16, and standard deviation was 0.362 (Table 7, Figure 7).

Table 7

Descriptive Statistics of Survival Months, Radiation and Chemotherapy, Stage of Lung Cancer, Socioeconomic Status, Race, Age, and Geographic Location of Lung Cancer Patients

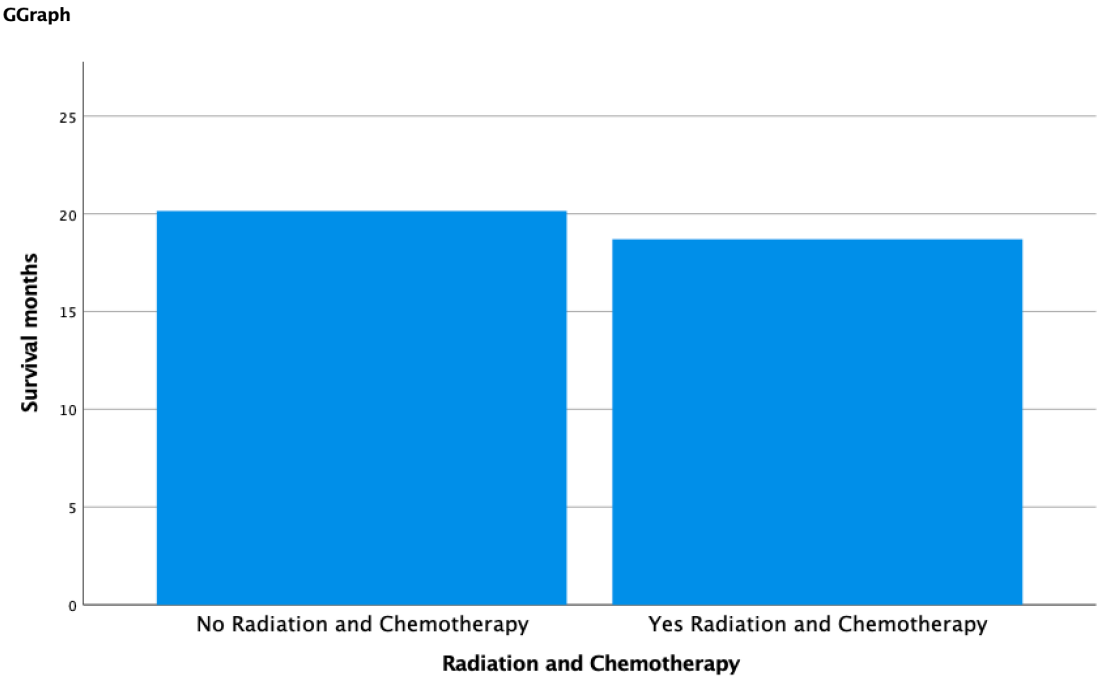
	Mean	**Std. Deviation**	**N**
Survival Months	7.56	6.664	86998
Radiation and Chemotherapy	0.20	0.401	86998
Stage of Lung Cancer	2.84	1.265	86998
Socioeconomic Status: Median household income	7.40	2.589	86998
Race: White and Non-White	1.26	0.439	86998
Age	14.62	2.061	86998
Geographic Location: Urban vs. Rural	1.16	0.362	86998

Note. Dependent variable: Survival Months.

Data Source: SEER Research Plus Data 17 Registries Nov 2021 Sub [2009-2019].

Figure 7

Association Between Combination of Radiation/Chemotherapy and Survival of Lung Cancer Patients



Note. Data Source: SEER Research Plus Data 17 Registries Nov 2021 Sub [2009-2019].

Inferential Analyses Results. To investigate What is the association between a combination of Chemotherapy and Radiotherapy and Survival/ life expectancy of lung cancer patients, a simple linear regression was conducted. The predictor was a combination of Chemotherapy and Radiotherapy, and the outcome was Survival/ life expectancy. The predictor variable was found to be statistically significant [B = 1.387, 95% C.I. (1.277, 1.497), $p < .001$], indicating that for lung cancer patients receiving a combination of Chemotherapy and Radiotherapy, Survival/ life expectancy increases by approximately 0.01 month. The model explained approximately 0% of the variability [R-

squared = 0.007]. A sensitivity analysis using G*power 3.1.9.7 software [n = 86,998; alpha error prob = 0.05; power = 0.80; predictors = 1] calculated [Effect size f^2 = 9.022085e-05] a small effect. The confidence interval associated with the regression analysis for the statistically significant predictors does not contain 0, which means the null hypothesis, there is no association between a combination of Chemotherapy and Radiotherapy, and Survival/ life expectancy of lung cancer patients, can be rejected.

To investigate What is the association between a combination of Chemotherapy and Radiotherapy and Survival/ life expectancy of lung cancer patients after controlling for Stage of lung cancer, Socioeconomic status, Race/ ethnicity, Age, and Geographical location (urban vs. rural)?, a multiple linear regression analysis was conducted to evaluate the prediction of Survival/ life expectancy of lung cancer patients from a combination of Chemotherapy and Radiotherapy, Stage of lung cancer, Socioeconomic status, Race/ ethnicity, Age, and Geographical location (urban vs. rural). The results of the multiple linear regression analysis revealed Geographical location (urban vs. rural) not to be a statistically significant predictor to the model ($p > .05$). However, the results of the multiple linear regression analysis revealed a statistically significant association between a combination of Chemotherapy and Radiotherapy, Stage of lung cancer, Socioeconomic status, Race/ ethnicity, and Age.

Controlling for Stage of lung cancer, Socioeconomic status, Race/ ethnicity, and Age, the regression coefficient [B = 2.414, 95% C.I. (2.307, 2.520) $p < .001$] associated with a combination of Chemotherapy and Radiotherapy suggests that for lung cancer

patients receiving a combination of Chemotherapy and Radiotherapy Survival/ life expectancy increases by approximately 2.4 months.

Controlling a combination of Chemotherapy and Radiotherapy, Socioeconomic status, Race/ ethnicity, and Age, the regression coefficient [B = -1.884, 95% C.I. (-1.917, -1.851) $p < .001$] associated with Stage of lung cancer, suggests that with each increase of lung cancer stage (Stage 1, Stage 2, Stage 3, and Stage 4), Survival/ life expectancy decreases by approximately 1.9 months.

Controlling for a combination of Chemotherapy and Radiotherapy, Stage of lung cancer, Race/ ethnicity, and Age, the regression coefficient [B = 0.132, 95% C.I. (0.113, 0.152) $p < .001$] associated with Socioeconomic status, suggests that for median household income level increase, Survival/ life expectancy increases by approximately 0.1 month or 3 days.

Controlling for a combination of Chemotherapy and Radiotherapy, Stage of lung cancer, Socioeconomic status, and Age, the regression coefficient [B = 0.138, 95% C.I. (0.043, 0.234) $p < .05$] associated with Race/ ethnicity, suggests that Whites compared to non-Whites, have Race/ ethnicity statistically significant predictor to the model ($p < .05$), associated with Race/ ethnicity, suggests Survival/ life expectancy increases by approximately 0.12 month or 3.6 days.

Controlling for a combination of Chemotherapy and Radiotherapy, Stage of lung cancer, Socioeconomic status, and Race/ ethnicity, the regression coefficient [B = -0.299, 95% C.I. (-0.320, -0.279) $p < .001$] associated with Age, suggests that with each

additional five years of age, Survival/ life expectancy decreases by approximately 0.3 month or 9 days.

Controlling for a combination of Chemotherapy and Radiotherapy, Stage of lung cancer, Socioeconomic status, Race/ ethnicity, Age, the regression coefficient [B = 2.414, 95% C.I. (2.307, 2.520) $p < .05$] associated with [IV2] suggests that with each additional [IV2], Survival/ life expectancy increases by approximately 0.1 month or 3 days. The R2 value of [0.140] associated with this regression model suggests that the association between a combination of Chemotherapy and Radiotherapy, Stage of lung cancer, Socioeconomic status, Race/ ethnicity, and Age account for [14%] of the variation in Survival/ life expectancy of lung cancer patients, which means that [86%] of the variation in Survival/ life expectancy of lung cancer patients cannot be explained by a combination of Chemotherapy and Radiotherapy, Stage of lung cancer, Socioeconomic status, Race/ ethnicity, and Age alone. A sensitivity analysis using G*power 3.1.9.7 software [n = 86,998; alpha error prob = 0.05; power = 0.80; predictors = 6] calculated [Effect size f^2 = 0.0001566159] a small effect. The confidence interval associated with the regression analysis for the statistically significant predictors does not contain 0, which means the null hypothesis, there is no association between a combination of Chemotherapy and Radiotherapy, Stage of lung cancer, Socioeconomic status, Race/ ethnicity, Age and Survival/ life expectancy of lung cancer patients, can be rejected.

RQ5

What is the association between stage of lung cancer and survival/ life expectancy after controlling for surgery, chemotherapy, radiotherapy, a combination of therapy

(surgery and chemotherapy; surgery and radiotherapy; surgery, chemotherapy, and radiotherapy; or chemotherapy and radiotherapy), socioeconomic status, race/ ethnicity, age, and geographic location (urban vs. rural)?

Descriptive Statistics Results. The study sample (n= 86,998), survival months, mean was 7.56, and standard deviation was 6.664. Stage of lung cancer, mean 2.84, and standard deviation was 1.265. Surgery, mean 0.23, and standard deviation was 0.424. Socioeconomic status, (median household income), mean 7.40, and standard deviation was 2.589. Race (White and Non-White), mean 1.26, and standard deviation was 0.439. Age, mean 14.62, and standard deviation was 2.061. Geographic location (Urban vs. Rural), mean 1.16, and standard deviation was 0.362 (Table 8, Figure 8).

Table 8

Descriptive Statistics of Survival Months, Stage of Lung Cancer, Surgery, Socioeconomic Status, Race, Age, and Geographic Location of Lung Cancer Patients

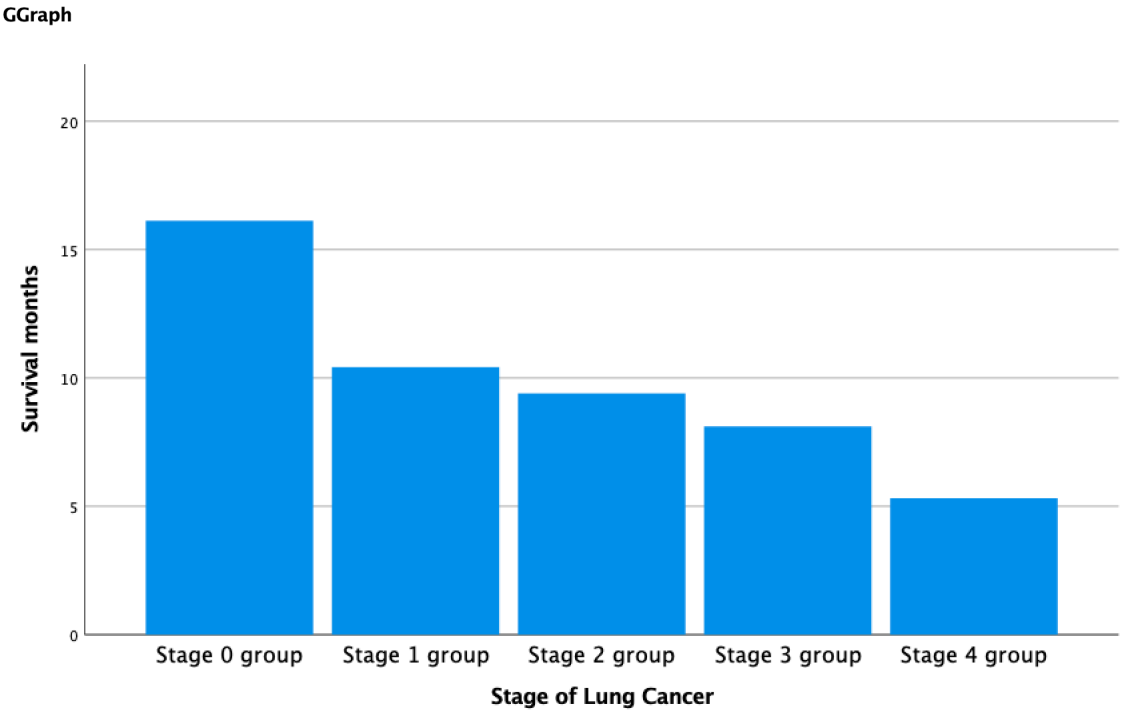
	Mean	**Std. Deviation**	**N**
Survival Months	7.56	6.664	86998
Stage of Lung Cancer	2.84	1.265	86998
Surgery	0.23	0.424	86998
Socioeconomic Status: Median household income	7.40	2.589	86998
Race: White and Non-White	1.26	0.439	86998
Age	14.62	2.061	86998
Geographic Location: Urban vs. Rural	1.16	0.362	86998

Note. Dependent variable: Survival Months.

Data Source: SEER Research Plus Data 17 Registries Nov 2021 Sub [2009-2019].

Figure 8

Association Between Stage of Lung Cancer and Survival of Lung Cancer Patients



Note. Source: SEER Research Plus Data 17 Registries Nov 2021 Sub [2009-2019].

Statistical Assumptions

I used multiple linear regression for my dataset analyses. To obtain valid results, the following eight assumptions were checked and satisfied to qualify my variables to be analyzed using multiple linear regression (Laerd Statistics, 2018):

Assumption-1: The dependent variable (survival/ life expectancy) should be measured on a continuous scale (interval or ratio).

Assumption-2: The Independent variables should be two or more and they can be continuous or categorical (surgery, chemotherapy, radiotherapy, a combination of therapy

[surgery and chemotherapy; surgery and radiotherapy; surgery, chemotherapy, and radiotherapy]; or chemotherapy and radiotherapy),

Assumption-3: The variables should reflect independence of observations (independence of residuals). In other words, variables should not be related to each other or appear as clustered. This can be verified by Durbin-Watson statistic test using SPSS statistics.

Assumption-4: There should be a linear relationship between the dependent variable and each of the independent variables; as well as between the dependent variable and all the independent variables together. This can be verified by different means such as scatterplots and partial regression plots using SPSS statistics.

Assumption-5: The data should reflect homoscedasticity or homogeneity of variance. This is an assumption of similar variances in different groups along the line were compared. This assumption can be tested using SPSS statistics.

Assumption-6: The data should not reflect multicollinearity that takes place when two or more independent variables are drastically correlated with each other. Such multicollinearity creates misunderstanding of identifying the independent variable that contributes to the variance mentioned in the dependent variable. This also creates a technical problem when calculating a multiple regression model. Multicollinearity can also be tested by using SPSS statistics.

Assumption-7: The dataset should not include significant outliers, high leverage points, or highly influential points. Such observations could have different effects on the regression line which in turn negatively affect the equation of regression used to predict the dependent variable value based on the independent variables. As such, the SPSS output

will reflect reduced predictive accuracy of the results as well as of the statistical significance. These observations can be checked while using SPSS statistics for multiple linear regression. Outliers and leverage points can be detected by using *casewise diagnostics and studentized deleted residuals* when using SPSS statistics. The influential points can be checked while using SPSS statistics by testing a measure of influence known as *Cook's Distance*.

Assumption-8: Residuals (errors) should be approximately normally distributed. This can be verified by a superimposed normal curve histogram and a normal P-P Plot or by a normal Q-Q Plot of the studentized residuals. Both methods can be tested by using SPSS statistics.

Inferential Analyses Results. To investigate What is the association between Stage of Lung Cancer and Survival/ life expectancy of lung cancer patients, a simple linear regression was conducted. The predictor was Stage of Lung Cancer, and the outcome was Survival/ life expectancy. The predictor variable was found to be statistically significant [B = -1.709, 95% C.I. (-1.742, -1.676), $p < .001$], indicating that for lung cancer patients with each increase of Stage of Lung Cancer (stage 1, stage 2, stage 3, stage 4), Survival/ life expectancy decreases by approximately 1.7 months. The model explained approximately 11% of the variability [R-squared = 0.105]. A sensitivity analysis using G*power 3.1.9.7 software [n = 86,998; alpha error prob = 0.05; power = 0.80; predictors = 1] calculated [Effect size f^2 = 9.022085e-05] a small effect. The confidence interval associated with the regression analysis for the statistically significant predictors does not contain 0, which means the null hypothesis, there is no association

between Stages of Lung Cancer and Survival/ life expectancy of lung cancer patients, can be rejected.

To investigate What is the association between Stage of Lung Cancer and Survival/ life expectancy of lung cancer patients after controlling for Surgery, Socioeconomic status, Race/ ethnicity, Age, and Geographical location (urban vs. rural)? a multiple linear regression analysis was conducted to evaluate the prediction of Survival/ life expectancy of lung cancer patients from Stage of Lung Cancer. The results of the multiple linear regression analysis revealed Geographical location (urban vs. rural) not to be a statistically significant predictor to the model ($p > .05$). However, the results of the multiple linear regression analysis revealed a statistically significant association between Surgery, Stage of lung cancer, Socioeconomic status, Race/ ethnicity, and Age.

Controlling for Surgery, Socioeconomic status, Race/ ethnicity, and Age, the regression coefficient [$B = - 1.376$, 95% C.I. ($- 1.415$, $- 1.337$) $p < .001$] associated with Stage of lung cancer, suggests that with each increase of lung cancer stage (Stage 1, Stage 2, Stage 3, and Stage 4), Survival/ life expectancy decreases by approximately 1.4 months.

Controlling for Surgery, Stage of Lung Cancer, Race/ ethnicity, and Age, the regression coefficient [$B = 0.107$, 95% C.I. (0.088, 0.127 $p < .001$] associated with Socioeconomic status, suggests that for median household income level increase, Survival/ life expectancy increases by approximately 0.1 month or 3 days.

Controlling for Stage of lung cancer, Socioeconomic status, Race/ Ethnicity, and Age, the regression coefficient [$B = 1.854$, 95% C.I. (1.738, 1.971) $p < .001$] associated

with Surgery, suggests that for lung cancer patients receiving Surgery, Survival/ life expectancy increases by approximately 1.9 months compared to those who did not receive Surgery.

Controlling for Surgery, Stage of lung cancer, Socioeconomic status, and Age, the regression coefficient [B = 0.097, 95% C.I. (0.001, 0.193) p=.048] associated with Race/ ethnicity, suggests that Whites compared to non-Whites, not to be a statistically significant predictor to the model ($p > .05$).

Controlling for Surgery, Stage of lung cancer, Socioeconomic status, and Race/ ethnicity, the regression coefficient [B = -0.324, 95% C.I. (-0.344-0.303) $p < .001$] associated with Age, suggests that with each additional five years of age, Survival/ life expectancy decreases by approximately 0.3 month or 9 days.

Controlling for Surgery, Stage of lung cancer, Socioeconomic status, Race/ ethnicity, and Age, the regression coefficient [B = -1.462, 95% C.I. (-1.501, -1.423) $p < .001$] associated with [IV2] suggests that with each additional [IV2], associated with Stage of lung cancer, suggests that with each increase of lung cancer stage (Stage 1, Stage 2, Stage 3, and Stage 4), Survival/ life expectancy decreases by approximately 0.1 month or 3 days. The R-squared value of [0.126] associated with this regression model suggests that the association between Surgery, Stage of lung cancer, Socioeconomic status, Race/ ethnicity, and Age account for [12.6%] of the variation in Survival/ life expectancy of lung cancer patients, which means that [87.4%] of the variation in Survival/ life expectancy of lung cancer patients cannot be explained by Surgery, Stage of lung cancer, Socioeconomic status, Race/ ethnicity, and Age alone. A sensitivity

analysis using G*power 3.1.9.7 software [n = 86,998; alpha error prob = 0.05; power = 0.80; predictors = 6] calculated [Effect size f^2 = 0.0001566159] a small effect. The confidence interval associated with the regression analysis for the statistically significant predictors does not contain 0, which means the null hypothesis, there is no association between Stage of lung cancer and Survival/ life expectancy of lung cancer patients, can be rejected.

RQ6

What is the association between socioeconomic status and survival/ life expectancy after controlling for surgery, chemotherapy, radiotherapy, a combination of therapy (surgery and chemotherapy; surgery and radiotherapy; surgery, chemotherapy, and radiotherapy; or chemotherapy and radiotherapy), stage of lung cancer, race/ ethnicity, age, and geographic location (urban vs. rural)?

Descriptive Statistics Results. The study sample (n= 86,998), survival months, mean was 7.56, and standard deviation was 6.664. Socioeconomic Status (median household income), mean 7.40, and standard deviation was 2.589. Surgery, mean 0.23, and standard deviation was 0.424. Chemotherapy, mean 0.40, and standard deviation was 0.489. Radiation, mean 0.39, and standard deviation was 0.487. Combination of Surgery and Chemotherapy, mean 0.04, and standard deviation was 0.199. Combination of Surgery and Radiotherapy, mean 0.00, standard deviation was 0.069. Combination of Surgery, Chemotherapy, and Radiotherapy, mean 0.02, standard deviation was 0.134. Combination of Chemotherapy and Radiotherapy, mean 0.20, standard deviation was 0.401. Stage of Lung Cancer, mean 2.84. standard deviation was 1.265. Race, (White

and non-White), mean 1.26, standard deviation was 0.439. Age, mean 14.62, standard deviation 2.061. Geographic location (Urban vs. Rural), mean 1.16, and standard deviation was 0. 362 (Table 9, Figure 9).

Table 9

Descriptive Statistics of Survival Months, Socioeconomic Status, Surgery, Chemotherapy, Radiotherapy, Surgery and Chemotherapy, Surgery and Radiotherapy, Radiotherapy and Chemotherapy, Surgery, Radiotherapy and Chemotherapy, Race, Age, Geographic Location, and Stage of Lung Cancer

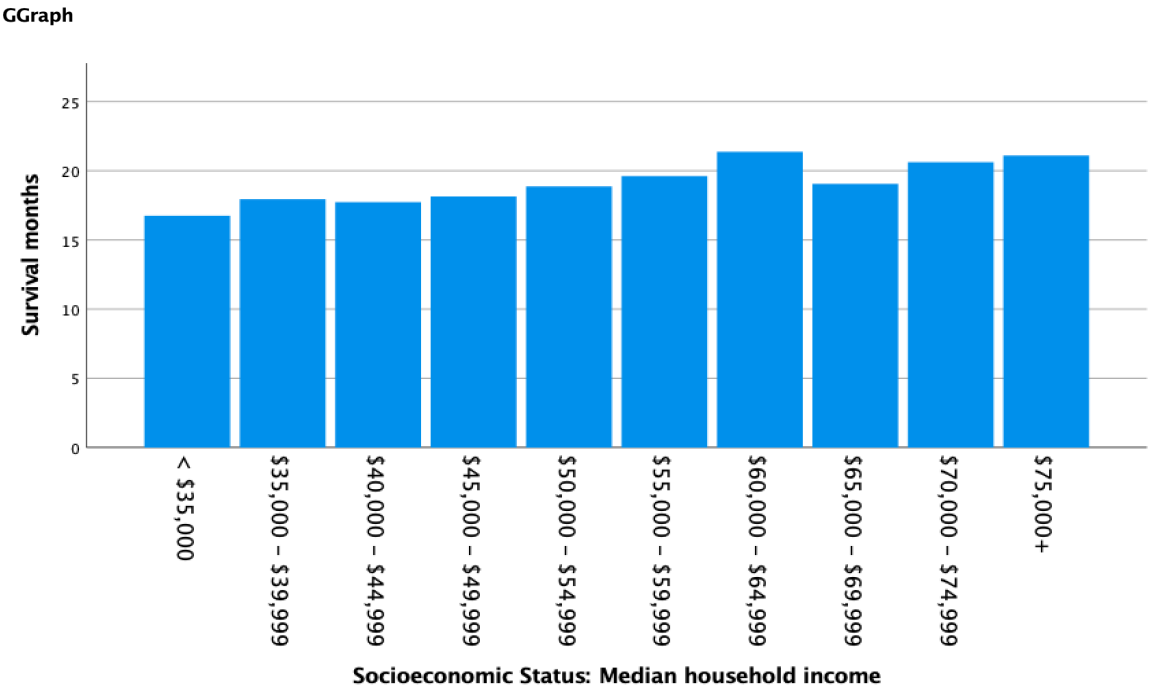
	Mean	**Std. Deviation**	**N**
Survival Months	7.56	6.664	86998
Socioeconomic Status: Median household income	7.40	2.589	86998
Surgery	0.23	0.424	86998
Chemotherapy	0.40	0.489	86998
Radiation	0.39	0.487	86998
Surgery and Chemotherapy	0.04	0.199	86998
Surgery and Radiation	0.00	0.069	86998
Radiation and Chemotherapy	0.20	0.401	86998
Surgery and Radiation and Chemotherapy	0.02	0.134	86998
Race: White and Non-White	1.26	0.439	86998
Age	14.62	2.061	86998
Geographic Location: Urban vs. Rural	1.16	0.362	86998
Stage of Lung Cancer	2.84	1.265	86998

Note. Dependent variable: Survival Months.

Data Source: SEER Research Plus Data 17 Registries Nov 2021 Sub [2009-2019].

Figure 9

Association Between Socioeconomic Status and Survival of Lung Cancer Patients



Note. Data Source: SEER Research Plus Data 17 Registries Nov 2021 Sub [2009-2019].

Statistical Assumptions

I used multiple linear regression for my dataset analyses. To obtain valid results, the following eight assumptions were checked and satisfied to qualify my variables to be analyzed using multiple linear regression (Laerd Statistics, 2018):

Assumption-1: The dependent variable (survival/ life expectancy) should be measured on a continuous scale (interval or ratio).

Assumption-2: The Independent variables should be two or more and they can be continuous or categorical (surgery, chemotherapy, radiotherapy, a combination of therapy

[surgery and chemotherapy; surgery and radiotherapy; surgery, chemotherapy, and radiotherapy]; or chemotherapy and radiotherapy),

Assumption-3: The variables should reflect independence of observations (independence of residuals). In other words, variables should not be related to each other or appear as clustered. This can be verified by Durbin-Watson statistic test using SPSS statistics.

Assumption-4: There should be a linear relationship between the dependent variable and each of the independent variables; as well as between the dependent variable and all the independent variables together. This can be verified by different means such as scatterplots and partial regression plots using SPSS statistics.

Assumption-5: The data should reflect homoscedasticity or homogeneity of variance. This is an assumption of similar variances in different groups along the line were compared. This assumption can be tested using SPSS statistics.

Assumption-6: The data should not reflect multicollinearity that takes place when two or more independent variables are drastically correlated with each other. Such multicollinearity creates misunderstanding of identifying the independent variable that contributes to the variance mentioned in the dependent variable. This also creates a technical problem when calculating a multiple regression model. Multicollinearity can also be tested by using SPSS statistics.

Assumption-7: The dataset should not include significant outliers, high leverage points, or highly influential points. Such observations could have different effects on the regression line which in turn negatively affect the equation of regression used to predict the dependent variable value based on the independent variables. As such, the SPSS output

will reflect reduced predictive accuracy of the results as well as of the statistical significance. These observations can be checked while using SPSS statistics for multiple linear regression. Outliers and leverage points can be detected by using *casewise diagnostics and studentized deleted residuals* when using SPSS statistics. The influential points can be checked while using SPSS statistics by testing a measure of influence known as *Cook's Distance.*

Assumption-8: Residuals (errors) should be approximately normally distributed. This can be verified by a superimposed normal curve histogram and a normal P-P Plot or by a normal Q-Q Plot of the studentized residuals. Both methods can be tested by using SPSS statistics.

Inferential Analyses Results. To investigate Is there an association between Socioeconomic Status and Survival/ life expectancy of lung cancer patients, a simple linear regression was conducted. The predictor was Socioeconomic Status, and the outcome was Survival/ life expectancy. The predictor variable was found to be statistically significant [B = 0.135, 95% C.I. (0.118, 0.152), $p < .001$], indicating that for lung cancer patients with Socioeconomic status, suggests that for median household income level increase, Survival/ life expectancy increases by approximately 0.1 months or 3 days.

The model explained approximately 0.3% of the variability [R-squared = 0.003]. A sensitivity analysis using G*power 3.1.9.7 software [n = 86,998; alpha error prob = 0.05; power = 0.80; predictors = 1] calculated [Effect size f^2 = 9.022085e-05] a small effect. The confidence interval associated with the regression analysis for the statistically

significant predictors does not contain 0, which means the null hypothesis, there is no association between Socioeconomic Status and Survival/ life expectancy of lung cancer patients, can be rejected.

To investigate What is the association between Socioeconomic Status and Survival/ life expectancy of lung cancer patients after controlling for Surgery, Chemotherapy, Radiotherapy, Combination of therapy (Surgery and Chemotherapy, Surgery and Radiotherapy, Surgery, chemotherapy, and Radiotherapy, Chemotherapy and Radiotherapy), Race/ ethnicity, Age, and Geographical location (urban vs. rural)? a multiple linear regression analysis was conducted to evaluate the prediction of Survival/ life expectancy of lung cancer patients from Socioeconomic Status. The results of the multiple linear regression analysis revealed Geographical location (urban vs. rural) not to be a statistically significant predictor to the model ($p > .05$). However, the results of the multiple linear regression analysis revealed a statistically significant association between surgery, chemotherapy, radiotherapy, a combination of therapy (surgery and chemotherapy; surgery and radiotherapy; surgery, chemotherapy, and radiotherapy; or chemotherapy and radiotherapy), Stage of lung cancer, Socioeconomic status, Race/ ethnicity, and Age.

Controlling for Surgery, Chemotherapy, Radiotherapy, a Combination of Therapy (Surgery and Chemotherapy; Surgery and Radiotherapy; Surgery, Chemotherapy, and Radiotherapy; or Chemotherapy and Radiotherapy), Stage of lung cancer, Race/ ethnicity, and Age, the regression coefficient [$B = 0.093$, 95% C.I. (0.074, 0.112) $p < .001$] associated with Socioeconomic status, suggests that for median household income

level increase, Survival/ life expectancy increases by approximately 0.1 months or 3 days.

Controlling for Chemotherapy, Radiotherapy, a Combination of Therapy (Surgery and Chemotherapy; Surgery and Radiotherapy; Surgery, Chemotherapy, and Radiotherapy; or Chemotherapy and Radiotherapy), Stage of Lung Cancer, Socioeconomic status, Race/ ethnicity, and Age, the regression coefficient [B = 3.441, 95% C.I. (3.287, 3.596) $p < .001$] associated with Surgery suggests that for lung cancer patients receiving Surgery, Survival/ life expectancy increases by approximately 3.4 months compared to those who did not receive Surgery.

Controlling for Surgery, Radiotherapy, a Combination of Therapy (Surgery and Chemotherapy; Surgery and Radiotherapy; Surgery, Chemotherapy, and Radiotherapy; or Chemotherapy and Radiotherapy), Stage of lung cancer, Socioeconomic Status, Race/ ethnicity, and Age, the regression coefficient [B = 4.135, 95% C.I. (4.000, 4.270) $p < .001$] associated with Chemotherapy suggests that for lung cancer patients receiving Chemotherapy, Survival/ life expectancy increases by approximately 4.1 months compared to those who did not receive Chemotherapy.

Controlling for Surgery, Chemotherapy, a Combination of Therapy (Surgery and Chemotherapy; Surgery and Radiotherapy; Surgery, Chemotherapy, and Radiotherapy; or Chemotherapy and Radiotherapy), Stage of lung cancer, Socioeconomic status, and Age, the regression coefficient [B = 2.567, 95% C.I. (2.434, 2.700) $p < .001$] associated with Radiotherapy suggests that for lung cancer patients receiving Radiotherapy, Survival/ life

expectancy increases by approximately 2.6 months. compared to those who did not receive Radiotherapy.

Controlling for Surgery, Chemotherapy, Radiotherapy, a Combination of Therapy (Surgery and Radiotherapy; Surgery, Chemotherapy, and Radiotherapy; or Chemotherapy and Radiotherapy), Stage of lung cancer, Socioeconomic status, Race/ ethnicity, and Age, the regression coefficient [B = -2.452, 95% C.I. (-2.710, -2.195) p < .001] associated with Surgery and Chemotherapy, suggests that for lung cancer patients receiving Surgery and Chemotherapy, Survival/ life expectancy decreases by approximately 2.5 months compared to those who did not receive Surgery and Chemotherapy.

Controlling for Surgery, Chemotherapy, Radiotherapy, a Combination of Therapy (Surgery and Chemotherapy; Surgery, Chemotherapy, and Radiotherapy; or Chemotherapy and Radiotherapy), Stage of lung cancer, Socioeconomic status, Race/ ethnicity, and Age, the regression coefficient [B = -1.798, 95% C.I. (-2.402, -1.194) p < .001] associated with Surgery and Radiotherapy, suggests that for lung cancer patients receiving Surgery and Radiotherapy Survival/ life expectancy decreases by approximately 1.8 months. compared to those who did not receive Surgery and Radiotherapy.

Controlling for Surgery, Chemotherapy, Radiotherapy, a Combination of Therapy (Surgery and Chemotherapy, ,Surgery and Radiotherapy; or Chemotherapy and Radiotherapy), Stage of lung cancer, Socioeconomic status, Race/ ethnicity, and Age, the regression coefficient [B = -3.931, 95% C.I. (-4.318, -3.543) p < .001] associated with

Surgery, Chemotherapy, and Radiotherapy, suggests that for lung cancer patients receiving Surgery, Chemotherapy, and Radiotherapy Survival/ life expectancy decreases by approximately 3.9 months compared to those who did not receive a combination of Surgery, Chemotherapy, and Radiotherapy.

Controlling for Surgery, Chemotherapy, Radiotherapy, a Combination of Therapy (Surgery and Chemotherapy, ,Surgery and Radiotherapy; Surgery, Chemotherapy, and Radiotherapy), Stage of lung cancer, Socioeconomic status, Race/ ethnicity, and Age, the regression coefficient [B = -1.954, 95% C.I. (-2.143, -1.764) $p < .001$] associated with Chemotherapy and Radiotherapy suggests that for lung cancer patients receiving Chemotherapy and Radiotherapy Survival/ life expectancy decreases by approximately 2 months. compared to those who did not receive Chemotherapy and Radiotherapy.

Controlling for Surgery, Chemotherapy, Radiotherapy, a Combination of Therapy (Surgery and Chemotherapy, ,Surgery and Radiotherapy; Surgery, Chemotherapy, and Radiotherapy, or Chemotherapy and Radiotherapy), Socioeconomic status, Race/ ethnicity, and Age, the regression coefficient [B = -1.575, 95% C.I. (-1.617, -1.532) $p < .001$] associated with Stage of lung cancer, suggests that with each increase of lung cancer stage (Stage 1, Stage 2, Stage 3, and Stage 4), Survival/ life expectancy decreases by approximately 1.6 months.

Controlling for Surgery, Chemotherapy, Radiotherapy, a Combination of Therapy (Surgery and Chemotherapy, ,Surgery and Radiotherapy; Surgery, Chemotherapy, and Radiotherapy, or Chemotherapy and Radiotherapy), Stage of lung cancer, Socioeconomic status,, and Age, the regression coefficient [B = 0.215, 95% C.I. (0.123, 0.307) $p < .001$]

associated with Race/ Ethnicity, suggests that Whites compared to non-Whites statistically significant predictor to the model ($p < .05$), associated with Race/ ethnicity, suggests Survival/ life expectancy increases by approximately 0.2 month or 6 days.

Controlling for Surgery, Chemotherapy, Radiotherapy, a Combination of Therapy (Surgery and Chemotherapy, ,Surgery and Radiotherapy; Surgery, Chemotherapy, and Radiotherapy, or Chemotherapy and Radiotherapy), Stage of lung cancer, Socioeconomic status, and Race/ Ethnicity, the regression coefficient [B = -0.119, 95% C.I. (-0.140, -0.099) $p < .001$] associated with Age suggests that for lung cancer patients associated with Age, with each additional five years of age, Survival/ life expectancy decreases by approximately 0.1 month or 3 days. The R-squared value of [0.194] associated with this regression model suggests that the association between socioeconomic status, surgery, chemotherapy, radiotherapy, a combination of therapy (surgery and chemotherapy; surgery and radiotherapy; surgery, chemotherapy, and radiotherapy; or chemotherapy and radiotherapy), stage of lung cancer, race/ ethnicity, and age, account for [19.4%] of the variation in Survival/ life expectancy of lung cancer patients, which means that [80.6%] of the variation in Survival/ life expectancy of lung cancer patients cannot be explained by socioeconomic status, surgery, chemotherapy, radiotherapy, a combination of therapy (surgery and chemotherapy; surgery and radiotherapy; surgery, chemotherapy, and radiotherapy; or chemotherapy and radiotherapy), stage of lung cancer, race/ ethnicity, and age, alone. A sensitivity analysis using G*power 3.1.9.7 software [n = 86,998; alpha error prob = .05; power = 0.80; predictors = 11] calculated [Effect size f^2 = 0.0001931495] a small effect. The confidence interval associated with the regression

analysis does not contain 0, which means the null hypothesis, there is no association between Socioeconomic Status and Survival/ life expectancy of lung cancer patients, can be rejected.

RQ7

What is the association between race/ ethnicity, age, and geographic location (urban vs. rural) and survival/ life expectancy after controlling for surgery, chemotherapy, radiotherapy, a combination of therapy (surgery and chemotherapy; surgery and radiotherapy; surgery, chemotherapy, and radiotherapy; or chemotherapy and radiotherapy), stage of lung cancer, and socioeconomic status?

Descriptive Statistics Results. The study sample (n= 86,998), survival months, mean was 7.56, and standard deviation was 6.664. Race (White and non-White), mean 1.26, and standard deviation was 0.439. Age, mean 14.62, standard deviation was 2.061. Geographical location (urban vs. rural), mean 1.16, standard deviation was 0.362. Surgery, mean 0.23, and standard deviation was 0.424. Stage of Lung Cancer, mean 2.84, standard deviation was 1.265. Socioeconomic status (median household income), mean 7.40, and standard deviation was 2.589 (Table 10, Figure 10).

Table 10

Descriptive Statistics of Survival Months, Race, Age, Geographic Location, Surgery, Stage of Lung Cancer, and Socioeconomic Status of Lung Cancer Patients

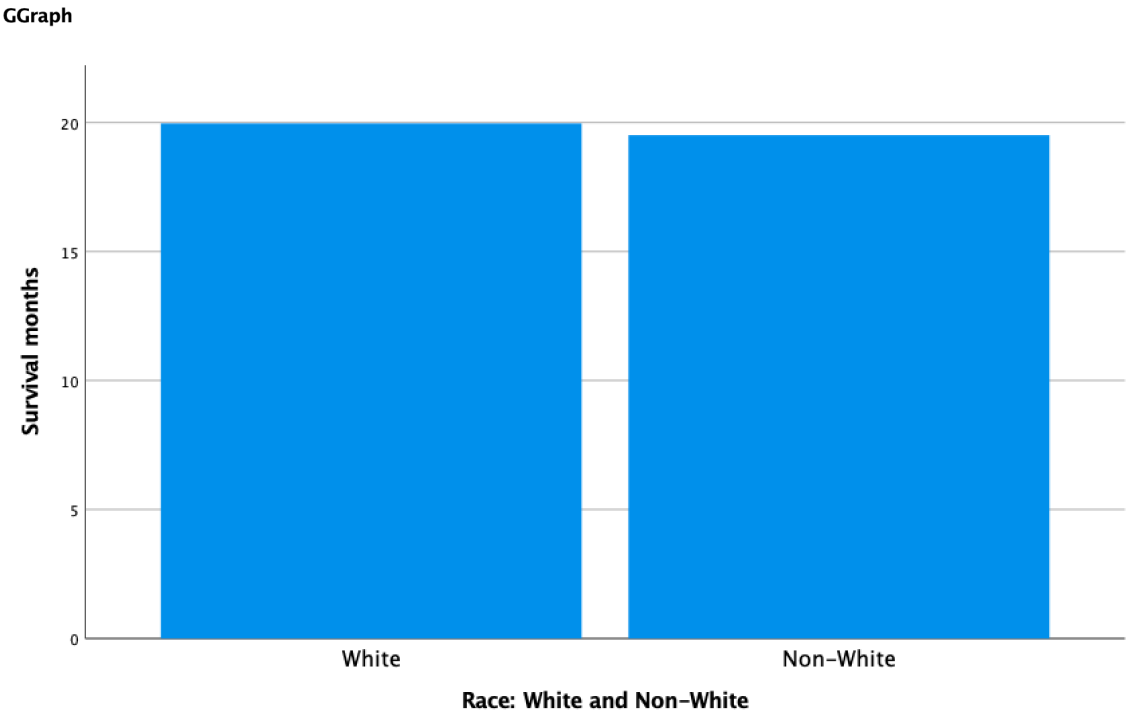
	Mean	**Std. Deviation**	**N**
Survival Months	7.56	6.664	86998
Race: White and Non-White	1.26	0.439	86998
Age	14.62	2.061	86998
Geographic Location: Urban vs. Rural	1.16	0.362	86998
Surgery	0.23	0.424	86998
Stage of Lung Cancer	2.84	1.265	86998
Socioeconomic Status: Median household income	7.40	2.589	86998

Note. Dependent variable: Survival Months.

Data Source: SEER Research Plus Data 17 Registries Nov 2021 Sub [2009-2019].

Figure 10

Association Between Race and Survival of Lung Cancer Patients



Note. Data Source: SEER Research Plus Data 17 Registries Nov 2021 Sub [2009-2019].

Statistical Assumptions

I used multiple linear regression for my dataset analyses. To obtain valid results, the following eight assumptions were checked and satisfied to qualify my variables to be analyzed using multiple linear regression (Laerd Statistics, 2018):

Assumption-1: The dependent variable (survival/ life expectancy) should be measured on a continuous scale (interval or ratio).

Assumption-2: The Independent variables should be two or more and they can be continuous or categorical (surgery, chemotherapy, radiotherapy, a combination of therapy

[surgery and chemotherapy; surgery and radiotherapy; surgery, chemotherapy, and radiotherapy]; or chemotherapy and radiotherapy),

Assumption-3: The variables should reflect independence of observations (independence of residuals). In other words, variables should not be related to each other or appear as clustered. This can be verified by Durbin-Watson statistic test using SPSS statistics.

Assumption-4: There should be a linear relationship between the dependent variable and each of the independent variables; as well as between the dependent variable and all the independent variables together. This can be verified by different means such as scatterplots and partial regression plots using SPSS statistics.

Assumption-5: The data should reflect homoscedasticity or homogeneity of variance. This is an assumption of similar variances in different groups along the line were compared. This assumption can be tested using SPSS statistics.

Assumption-6: The data should not reflect multicollinearity that takes place when two or more independent variables are drastically correlated with each other. Such multicollinearity creates misunderstanding of identifying the independent variable that contributes to the variance mentioned in the dependent variable. This also creates a technical problem when calculating a multiple regression model. Multicollinearity can also be tested by using SPSS statistics.

Assumption-7: The dataset should not include significant outliers, high leverage points, or highly influential points. Such observations could have different effects on the regression line which in turn negatively affect the equation of regression used to predict the dependent variable value based on the independent variables. As such, the SPSS output

will reflect reduced predictive accuracy of the results as well as of the statistical significance. These observations can be checked while using SPSS statistics for multiple linear regression. Outliers and leverage points can be detected by using *casewise diagnostics and studentized deleted residuals* when using SPSS statistics. The influential points can be checked while using SPSS statistics by testing a measure of influence known as *Cook's Distance.*

Assumption-8: Residuals (errors) should be approximately normally distributed. This can be verified by a superimposed normal curve histogram and a normal P-P Plot or by a normal Q-Q Plot of the studentized residuals. Both methods can be tested by using SPSS statistics.

Inferential Analyses Results. To investigate What is the association between Race/ Ethnicity and Survival/ life expectancy of lung cancer patients, a simple linear regression was conducted. The predictor was Race/ Ethnicity, and the outcome was Survival/ life expectancy. The predictor variable was found to be statistically significant [B = -0.218, 95% C.I. (-0.320, -0.116), $p < .001$], indicating that for lung cancer patients Whites compared to non-Whites suggests having Survival/ life expectancy decreases by approximately 0.2 month or 6 days. The model explained approximately 1% of the variability [R-squared = 0.010]. A sensitivity analysis using G*power 3.1.9.7 software [n = 86,998; alpha error prob = 0.05; power = 0.80; predictors = 1] calculated [Effect size f^2 = 9.022085e-05] a small effect. The confidence interval associated with the regression analysis for the statistically significant predictors does not contain 0, which means the

null hypothesis, there is no association between Race/ Ethnicity and Survival/ life expectancy of lung cancer patients, can be rejected.

To investigate What is the association between Race/ Ethnicity and Survival/ life expectancy of lung cancer patients after controlling for Surgery, Stage of lung cancer, Socioeconomic status, Age, and Geographical location (urban vs. rural)? a multiple linear regression analysis was conducted to evaluate the prediction of Survival/ life expectancy of lung cancer patients from Race/ Ethnicity, Surgery, Stage of lung cancer, Socioeconomic status, Age, and Geographical location (urban vs. rural). The results of the multiple linear regression analysis revealed Geographical location (urban vs. rural) not to be a statistically significant predictor to the model ($p > .05$). However, the results of the multiple linear regression analysis revealed a statistically significant association between Race/ Ethnicity, Surgery, Stage of lung cancer, Socioeconomic status, and Age.

Controlling for Surgery, Stage of lung cancer, Socioeconomic status, and Age, the regression coefficient [$B = -0.067$, 95% C.I. (-0.166, 0.031) $p > .05$] associated with Race/ Ethnicity, suggests not to be a statistically significant predictor to the model ($p > .05$).

Controlling for Surgery, Socioeconomic status, Race/ ethnicity, and Age, the regression coefficient [$B = -1.462$, 95% C.I. (-1.501, -1.423) $p < .001$] associated with Stage of lung cancer, suggests that with each increase of lung cancer stage (Stage 1, Stage 2, Stage 3, and Stage 4), Survival/ life expectancy decreases by approximately 1.5 months.

Controlling for Surgery, Stage of lung cancer, Race/ ethnicity, and Age, the regression coefficient [B = 0.107, 95% C.I. (0.088, 0.127) p < .001] associated with Socioeconomic status, suggests that for median household income level increase, Survival/ life expectancy increases by approximately 0.1 month or 3 days.

Controlling for Stage of lung cancer, Socioeconomic status, Race/ ethnicity, and Age, the regression coefficient [B = 1.484, 95% C.I. (1.366, 1.602) p < .001] associated with Surgery, suggests that for lung cancer patients receiving Surgery Survival/ life expectancy increases by approximately 1.5 months. compared to those who did not receive Surgery.

Controlling for Surgery, Stage of lung cancer, Socioeconomic status, and Race/ ethnicity, the regression coefficient [B = -0.324, 95% C.I. (-0.344, -0.303) p < .001] associated with Age, suggests that with each additional five years of age, Survival/ life expectancy decreases by approximately 0.3 months or 9 days.

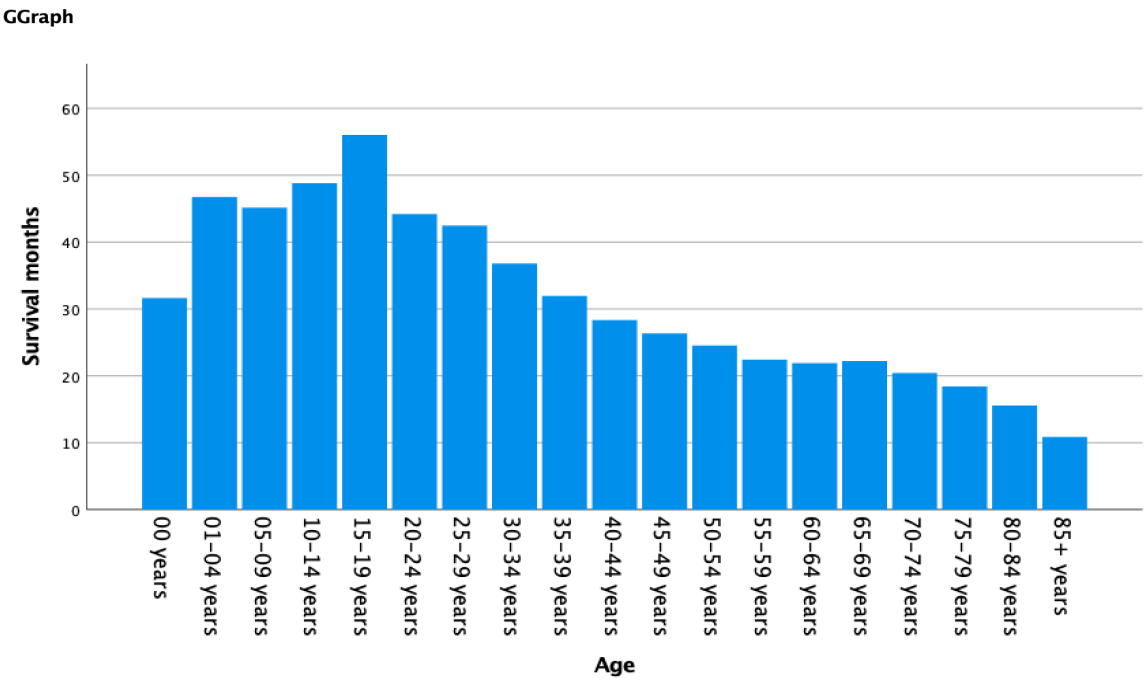
Controlling for Surgery, Stage of lung cancer, Socioeconomic status, Race/ ethnicity, and Age, the regression coefficient [B = 0.097, 95% C.I. (0.001, 0.193) p < .05] associated with [IV2] suggests that with each additional [IV2], Survival/ life expectancy increases by approximately 0.1 month or 3 days. The R-squared value of [0.126] associated with this regression model suggests that the association between Surgery, Stage of lung cancer, Socioeconomic status, Race/ ethnicity, and Age account for [12.6%] of the variation in Survival/ life expectancy of lung cancer patients, which means that [87.4%] of the variation in Survival/ life expectancy of lung cancer patients cannot be explained by Surgery, Stage of lung cancer, Socioeconomic status, Race/ ethnicity,

and Age alone. A sensitivity analysis using G*power 3.1.9.7 software [n = 86,998; alpha error prob = 0.05; power = 0.80; predictors = 6] calculated [Effect size f^2 = 0.0001566159] a small effect. The confidence interval associated with the regression analysis for the statistically significant predictors does not contain 0, which means the null hypothesis, there is no association between race/ ethnicity, age, and Survival/ life expectancy of lung cancer patients, can be rejected.

To investigate What is the association between Age and Survival/ life expectancy of lung cancer patients after controlling for Chemotherapy, Stage of lung cancer, Socioeconomic status, Race/ ethnicity, and Geographical location (urban vs. rural)? a multiple linear regression analysis was conducted to evaluate the prediction of Survival/ life expectancy of lung cancer patients from Age, Stage of lung cancer, Socioeconomic status, Race/ ethnicity, and Geographical location (urban vs. rural). The results of the multiple linear regression analysis revealed Geographical location (urban vs. rural) not to be a statistically significant predictor to the model ($p > .05$). However, the results of the multiple linear regression analysis revealed a statistically significant association between Chemotherapy, Stage of lung cancer, Socioeconomic status, Race/ ethnicity, and Age (Figure 11).

Figure 11

Association Between Age and Survival of Lung Cancer Patients



Note. Data Source: SEER Research Plus Data 17 Registries Nov 2021 Sub [2009-2019].

Controlling for Stage of lung cancer, Socioeconomic status, Race/ ethnicity, and Age, the regression coefficient [B = 1.456, 95% C.I. (1.365, 1.548) p < .001] associated with Chemotherapy, suggests that for lung cancer patients receiving Chemotherapy, Survival/ life expectancy increases by approximately 1.5 months. compared to those who did not receive Chemotherapy.

Controlling for Chemotherapy, Socioeconomic status, Race/ ethnicity, and Age, the regression coefficient [B = -2.161, 95% C.I. (-2.195, -2.127) p < .001] associated with Stage of lung cancer, suggests that with each increase of lung cancer stage (Stage 1,

Stage 2, Stage 3, and Stage 4), Survival/ life expectancy decreases by approximately 2.2 months.

Controlling for Chemotherapy, Stage of lung cancer, Race/ ethnicity, and Age, the regression coefficient [B = 0.112, 95% C.I. (0.093, 0.131) p < .001] associated with Socioeconomic status, suggests that for median household income level increase, Survival/ life expectancy increases by approximately 0.1 month or 3 days.

Controlling for Chemotherapy, Stage of lung cancer, Socioeconomic status, and Age, the regression coefficient [B = -0.218, 95% C.I. (-0.320, -0.116) p < .001] associated with Race/ ethnicity, suggests that Whites compared to non-Whites, have Survival/ life expectancy decreases by approximately 0.2 month or 6 days.

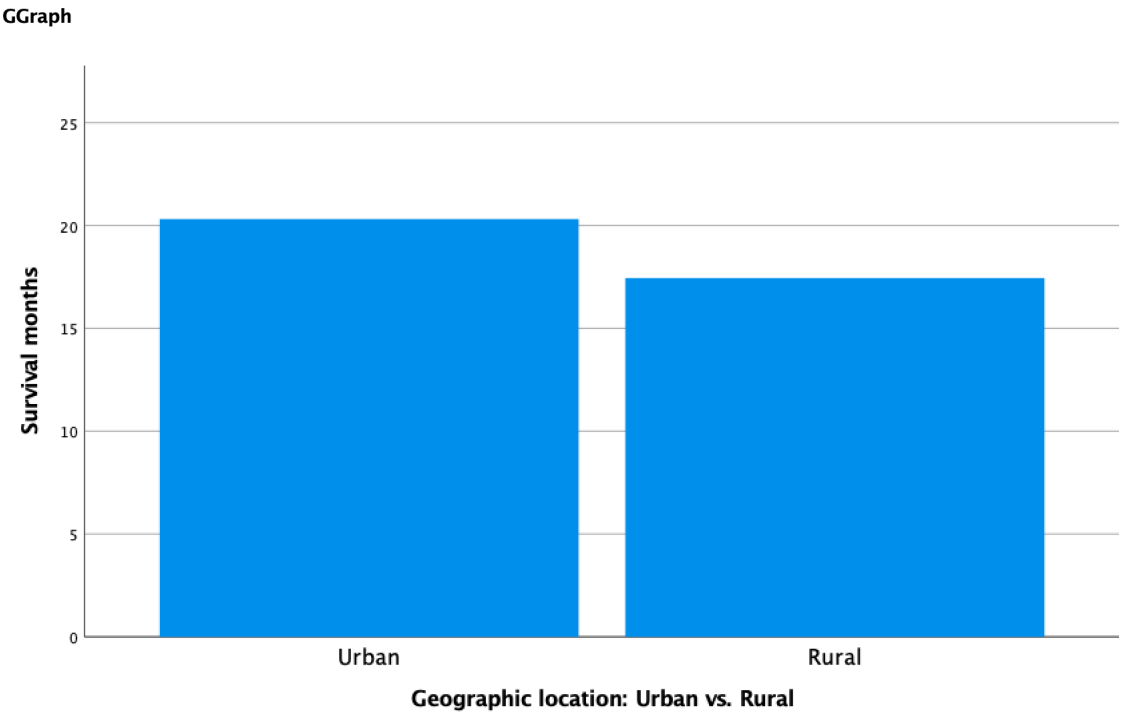
Controlling for Chemotherapy, Stage of lung cancer, Socioeconomic status, and Race/ ethnicity, the regression coefficient [B = -0.314, 95% C.I. (-0.336, -0.293) p < .001] associated with Age, suggests that with each additional five years of age, Survival/ life expectancy decreases by approximately 0.3 months or 9 days.

Controlling for Chemotherapy, Stage of lung cancer, Socioeconomic status, Race/ ethnicity, and Age, the regression coefficient [B = -0.199, 95% C.I. (-0.219, -0.178 p < .001] associated with [IV2] suggests that with each additional [IV2], Survival/ life expectancy decreases by approximately 0.2 month or 6 days. The R-squared value of [0.172] associated with this regression model suggests that the association between Chemotherapy, Stage of lung cancer, Socioeconomic status, Race/ ethnicity, and Age account for [17.2%] of the variation in Survival/ life expectancy of lung cancer patients, which means that [82.8%] of the variation in Survival/ life expectancy of lung cancer

patients cannot be explained by Chemotherapy, Stage of lung cancer, Socioeconomic status, Race/ ethnicity, and Age alone. A sensitivity analysis using G*power 3.1.9.7 software [n = 86,998; alpha error prob = 0.05; power = 0.80; predictors = 6] calculated [Effect size f^2 = 0.0001566159] a small effect. The confidence interval associated with the regression analysis for the statistically significant predictors does not contain 0, which means the null hypothesis, there is no association between race/ ethnicity, age, and Survival/ life expectancy of lung cancer patients, can be rejected (Figure 12).

Figure 12

Association Between Geographic Location and Survival of Lung Cancer Patients



Note. Data Source: SEER Research Plus Data 17 Registries Nov 2021 Sub [2009-2019].

Summary

The purpose of this study was to demonstrate the impact of different types of treatment that were prescribed to patients based on the stage of lung cancer, the impact of stage of lung cancer at diagnosis, together with the socioeconomic status on life expectancy of marginalized communities. The SPSS output analyses results demonstrated that there were statistically significant association between different types of treatments such as: Surgery; Chemotherapy; Radiotherapy; combination of therapy such as: Surgery and Chemotherapy; Surgery and Radiotherapy; Surgery, Chemotherapy, and Radiotherapy; Chemotherapy and Radiotherapy; and Survival/ Life Expectancy.

Statistical Results

The predictor was Surgery, and the outcome was Survival/ life expectancy. The predictor variable was found to be statistically significant [B = 4.055, 95% C.I. (3.954, 4.156), $p < .001$], indicating that for lung cancer patients receiving surgery, Survival/ life expectancy increases by approximately 4 months. The predictor was chemotherapy, and the outcome was Survival/ life expectancy. The predictor variable was found to be statistically significant [B = 1.672, 95% C.I. (1.583, 1.762), $p < .001$], indicating that for lung cancer patients receiving chemotherapy, Survival/ life expectancy increases by approximately 1.7 months. The predictor was Radiotherapy, and the outcome was Survival/ life expectancy. The predictor variable was found to be statistically significant [B = 1.529, 95% C.I. (1.439, 1.620), $p < .001$], indicating that for lung cancer patients receiving Radiotherapy, Survival/ life expectancy increases by approximately 1.5 months. The predictor was a combination of Surgery and Chemotherapy, and the outcome was

Survival/ life expectancy. The predictor variable was found to be statistically significant [B = 3.340, 95% C.I. (3.119, 3.562), $p < .001$], indicating that for lung cancer patients receiving a combination of Surgery and Chemotherapy, Survival/ life expectancy increases by approximately 3.3 months.

The predictor was a combination of Surgery and Radiotherapy, and the outcome was Survival/ life expectancy. The predictor variable was found to be statistically significant [B = 2.046, 95% C.I. (1.404, 2.689), $p < .001$], indicating that for lung cancer patients receiving a combination of Surgery and Radiotherapy, Survival/ life expectancy increases by approximately 2.1 months. The predictor was a combination of Surgery, Chemotherapy, and Radiotherapy, and the outcome was Survival/ life expectancy. The predictor variable was found to be statistically significant [B = 3.447, 95% C.I. (3.118, 3.776), $p < .001$], indicating that for lung cancer patients receiving a combination of Surgery, Chemotherapy, and Radiotherapy, Survival/ life expectancy increases by approximately 0.01 month. The predictor was a combination of Chemotherapy and Radiotherapy, and the outcome was Survival/ life expectancy. The predictor variable was found to be statistically significant [B = 1.387, 95% C.I. (1.277, 1.497), $p < .001$], Indicating that for lung cancer patients receiving a combination of Chemotherapy and Radiotherapy, Survival/ life expectancy increases by approximately 0.01 month.

Also, the SPSS output analyses results demonstrated that there were statistically significant association between Stage of Lung Cancer, Socioeconomic Status and Survival/ Life Expectancy.

Statistical Results

The predictor was Stage of Lung Cancer, and the outcome was Survival/ life expectancy. The predictor variable was found to be statistically significant [B = -1.709, 95% C.I. (-1.742, -1.676), $p < .001$], indicating that for lung cancer patients with each increase of Stage of Lung Cancer (stage 1, stage 2, stage 3, stage 4), Survival/ life expectancy decreases by approximately 1.7 months. The predictor was Socioeconomic Status, and the outcome was Survival/ life expectancy. The predictor variable was found to be statistically significant [B = 0.135, 95% C.I. (0.118, 0.152), $p < .001$], indicating that for lung cancer patients with Socioeconomic status, suggests that for median household income level increase, Survival/ life expectancy increases by approximately 0.1 months or 3 days.

Chapter 5: Discussion, Conclusions, and Recommendations

The purpose of my study was to examine the effect of treatment, stage of lung cancer, and socioeconomic status on survival/ life expectancy of marginalized communities. My study also investigated and identified the contributing factors for short survival due to lung cancer among marginalized communities. Underserved communities' inherent poor lifestyle with contributing factors such as cigarette smoking, lack or low education levels, low income, and lack of access to healthcare. These risk factors contribute to lung cancer that is responsible of 25% of cancer mortality (Cancer Treatment Centers of America, 2020). This highlighted the importance of screening for lung cancer which facilitates early diagnosis, possible curative treatment, better prognosis, and a longer life expectancy. This was demonstrated in my study analyses and results that reflected survival and life expectancy improved in early stages of lung cancer when treatment prescribed lead to greater response than when prescribed for lately diagnosed patients with lung cancer.

A quantitative retrospective study was conducted on cohorts who were diagnosed with lung cancer at different stages during the period (2009-2019), and they received different treatment regimens that included surgery, chemotherapy, radiotherapy, or combinations of some of them and they were evaluated during their course of treatments until their last follow-ups or deaths. A cross-sectional design was performed to examine the relationship between independent variables (treatment such as surgery, chemotherapy, radiotherapy, and a combination of them, stages of lung cancer, and socioeconomic status) and dependent variable (survival/ life expectancy).

Patients are usually diagnosed at late stages of lung cancer (stage III & IV) when curative treatment is mostly not possible. These factors contribute to short life expectancy of marginalized community. Despite about 60% of cancer affects underserved and disadvantage communities (Montagne et al., 2021), very few research studies have addressed this issue; therefore, my study has contributed to close this gap in literature.

The SPSS output analyses results in my study demonstrated that there were statistically significant association between different types of treatments such as: Surgery; Chemotherapy; Radiotherapy; Combination of Therapy such as: Surgery and Chemotherapy; Surgery and Radiotherapy; Surgery, Chemotherapy, and Radiotherapy; Chemotherapy and Radiotherapy; and Survival/ Life Expectancy. Also, the SPSS output analyses results demonstrated that there were statistically significant association between Stage of Lung Cancer, Socioeconomic Status and Survival/ Life Expectancy.

The predictor was Surgery, and the outcome was Survival/ life expectancy. The predictor variable was found to be statistically significant [B = 4.055, 95% C.I. (3.954, 4.156), $p < .001$], indicating that for lung cancer patients receiving surgery, Survival/ life expectancy increases by approximately 4 months. The predictor was chemotherapy, and the outcome was Survival/ life expectancy. The predictor variable was found to be statistically significant [B = 1.672, 95% C.I. (1.583, 1.762), $p < .001$], indicating that for lung cancer patients receiving chemotherapy, Survival/ life expectancy increases by approximately 1.7 months. The predictor was Radiotherapy, and the outcome was Survival/ life expectancy. The predictor variable was found to be statistically significant

[B = 1.529, 95% C.I. (1.439, 1.620), p < .001], indicating that for lung cancer patients receiving Radiotherapy, Survival/ life expectancy increases by approximately 1.5 months.

The predictor was a combination of Surgery and Chemotherapy, and the outcome was Survival/ life expectancy. The predictor variable was found to be statistically significant [B = 3.340, 95% C.I. (3.119, 3.562), p < .001], indicating that for lung cancer patients receiving a combination of Surgery and Chemotherapy, Survival/ life expectancy increases by approximately 3.3 months. The predictor was a combination of Surgery and Radiotherapy, and the outcome was Survival/ life expectancy. The predictor variable was found to be statistically significant [B = 2.046, 95% C.I. (1.404, 2.689), p < .001], indicating that for lung cancer patients receiving a combination of Surgery and Radiotherapy, Survival/ life expectancy increases by approximately 2.1 months. The predictor was a combination of Surgery, Chemotherapy, and Radiotherapy, and the outcome was Survival/ life expectancy. The predictor variable was found to be statistically significant [B = 3.447, 95% C.I. (3.118, 3.776), p < .001], indicating that for lung cancer patients receiving a combination of Surgery, Chemotherapy, and Radiotherapy, Survival/ life expectancy increases by approximately 0.01 month. The predictor was a combination of Chemotherapy and Radiotherapy, and the outcome was Survival/ life expectancy. The predictor variable was found to be statistically significant [B = 1.387, 95% C.I. (1.277, 1.497), p < .001], indicating that for lung cancer patients receiving a combination of Chemotherapy and Radiotherapy, Survival/ life expectancy increases by approximately 0.01 month.

The predictor was Stage of Lung Cancer, and the outcome was Survival/ life expectancy. The predictor variable was found to be statistically significant [B = -1.709, 95% C.I. (-1.742, -1.676), $p < .001$], indicating that for lung cancer patients with each increase of Stage of Lung Cancer (stage 1, stage 2, stage 3, stage 4), Survival/ life expectancy decreases by approximately 1.7 months. The predictor was Socioeconomic Status, and the outcome was Survival/ life expectancy. The predictor variable was found to be statistically significant [B = 0.135, 95% C.I. (0.118, 0.152), $p < .001$], indicating that for lung cancer patients with Socioeconomic status, suggests that for median household income level increase, Survival/ life expectancy increases by approximately 0.1 month or 3 days.

This chapter (5) included besides the Introduction section, Interpretation of the Findings section, Limitations of the Study section, Recommendations section, Implications section, and Conclusion section.

Interpretation of the Findings

The analyses and results of my research study have underscored the critical role of treatment, stage of lung cancer, and socioeconomic status on survival/ life expectancy. Furthermore, my study results addressed lung cancer as a social problem specially among marginalized and disadvantaged communities who suffer most due to late diagnosis and late treatment because of lack of access to health care which in turn lead to a short life expectancy.

RQ1

The predictor was Surgery, and the outcome was Survival/ life expectancy. The predictor variable was found to be statistically significant [B = 4.055, 95% C.I. (3.954, 4.156), $p < .001$], indicating that for lung cancer patients receiving surgery, Survival/ life expectancy increases by approximately 4 months. The model explained approximately 7% of the variability [R-squared = .067]. A sensitivity analysis using G*power 3.1.9.7 software [n = 86,998; alpha error prob = .05; power = .80; predictors = 1] calculated [Effect size f^2 = 9.022085e-05] a small effect. The confidence interval associated with the regression analysis does not contain 0, which means the null hypothesis, there is no association between Surgery and Survival/ life expectancy of lung cancer patients, can be rejected.

The innovation of advanced diagnostic methods and techniques for the diagnosis of lung cancer and the innovation of treatment modalities have led to the evolution of lung cancer surgery from invasive thoracic approach to minimally invasive (Montagne et al., 2021). The indication for surgery will be decided based on the stage of lung cancer and the surgical approach will be either resection guided by imaging such as CT scan or ultrasound, or minimally invasive surgical approach, or hybrid approach (Montagne et al., 2021).

Caballero-Vázquez et al. (2021) has reached the same conclusion when they investigated the risk factors for short-term lung cancer survival on a study sample of 521 patients diagnosed with Non-Small Cell Lung Cancer (NSCLC) at Virgen de las Nieves Hospital, Granada, Spain during the period between January 1st, 2011, and December

31st, 2016. The purpose of the study was to identify risk factors and assess their effect on prognosis for survival of less or more than one year depending on epidemiological and clinical variables at the time of diagnosis (Caballero-Vázquez et al., 2021). Caballero-Vázquez et al. (2021) concluded that epidemiological variables such as age, smoking history, history of lung cancer, and clinical variables such as dyspnea, dysphonia, cancer location, and chest pain were suggested as predictors for survival in lung cancer patients at the time of diagnosis; while the therapeutic variables such as surgery, chemotherapy, radiotherapy, or combination is significant for identifying less than one year survival (Caballero-Vázquez et al., 2021). That is because surgery is the treatment of choice for early stages of lung cancer (stage I &II) with a possible curative effect.

Montagne et al. (2021) argued that surgery contributes to diagnosis in early preinvasive stage of lung cancer, to surgical excision of preinvasive tumor, locally, or locally advanced tumor, to palliative surgery that can relieves symptoms such as dyspnea and pain in late stages of lung cancer.

RQ2

The predictor was chemotherapy, and the outcome was Survival/ life expectancy. The predictor variable was found to be statistically significant [B = 1.672, 95% C.I. (1.583, 1.762), $p < .001$], indicating that for lung cancer patients receiving chemotherapy, Survival/ life expectancy increases by approximately 1.7 months. The model explained approximately 2% of the variability [R-squared = 0.015]. A sensitivity analysis using G*power 3.1.9.7 software [n = 86,998; alpha error prob = 0.05; power = 0.80; predictors = 1] calculated [Effect size f^2 = 9.022085e-05] a small effect. The confidence interval

associated with the regression analysis does not contain 0, which means the null hypothesis, there is no association between chemotherapy and Survival/ life expectancy of lung cancer patients, can be rejected.

Baxevanos & Mountzios, (2018) argued that despite the application of novel chemotherapeutic agents, there is modest increase of survival. The median overall survival (OS) with the therapeutic plateau ranges between 12 and 14 months according to recent clinical trials (Baxevanos & Mountzios, 2018). Despite recent discoveries and implementations of targeted therapy and immunotherapy, chemotherapy is still considered the treatment of choice for most cases with advanced NSCLC (Adamowicz et al., 2020; Baxevanos & Mountzios, 2018).

RQ3

The predictor was Radiotherapy, and the outcome was Survival/ life expectancy. The predictor variable was found to be statistically significant [B = 1.529, 95% C.I. (1.439, 1.620), $p < .001$], indicating that for lung cancer patients receiving Radiotherapy, Survival/ life expectancy increases by approximately 1.5 months. The model explained approximately 1% of the variability [R-squared = 0.013]. A sensitivity analysis using G*power 3.1.9.7 software [n = 86,998; alpha error prob = 0.05; power = 0.80; predictors = 1] calculated [Effect size f^2 = 9.022085e-05] a small effect. The confidence interval associated with the regression analysis does not contain 0, which means the null hypothesis, there is no association between Radiotherapy and Survival/ life expectancy of lung cancer patients, can be rejected.

There are different types of radiotherapy such as external beam radiation therapy (EBRT) which is a direct radiation to the lungs, and it is the most used type of radiotherapy for Non-Small Cell Lung Cancer (NSCLC) and its metastases (American Cancer Society, 2021); Brachytherapy (Internal Radiation Therapy) is another type of radiotherapy that is used to treat NSCLC. Brachytherapy is an internal radiation therapy that facilitates minimizing the tumor size to relieve obstruction of the airway (American Cancer Society, 2021); and Radiofrequency Ablation which is used for the treatment of a small size lung cancer located in the outer part of the lung by inserting a needle-like probe through the skin guided by CT scan aimed at the tumor; and a high-energy radio wave (electric current) is passed through the probe which leads to damage of cancer cells by heat generated by the electric current (American Cancer Society, 2021). Radiotherapy can also be prescribed to treat distant metastasis in bones and brain; as well as a palliative treatment in advanced cases of NSCLC to relieve symptoms such as pain or cough (American Cancer Society, 2021).

RQ4

The predictor was a combination of Surgery and Chemotherapy, and the outcome was Survival/ life expectancy. The predictor variable was found to be statistically significant [B = 3.340, 95% C.I. (3.119, 3.562), $p < .001$], indicating that for lung cancer patients receiving a combination of Surgery and Chemotherapy, Survival/ life expectancy increases by approximately 3.3 months. The model explained approximately 1% of the variability [R-squared = 0.010]. A sensitivity analysis using G*power 3.1.9.7 software [n = 86,998; alpha error prob = 0.05; power = 0.80; predictors = 1] calculated [Effect size f^2

= 9.022085e-05] a small effect. The confidence interval associated with the regression analysis does not contain 0, which means the null hypothesis, there is no association between a combination of Surgery and Chemotherapy and Survival/ life expectancy of lung cancer patients, can be rejected.

The predictor was a combination of Surgery and Radiotherapy, and the outcome was Survival/ life expectancy. The predictor variable was found to be statistically significant [B = 2.046, 95% C.I. (1.404, 2.689), $p < .001$], indicating that for lung cancer patients receiving a combination of Surgery and Radiotherapy, Survival/ life expectancy increases by approximately 2.1 months. The model explained approximately 0% of the variability [R-squared = 0.000]. A sensitivity analysis using G*power 3.1.9.7 software [n = 86,998; alpha error prob = 0.05; power = 0.80; predictors = 1] calculated [Effect size f^2 = 9.022085e-05] a small effect. The confidence interval associated with the regression analysis does not contain 0, which means the null hypothesis, there is no association between a combination of Surgery and Radiotherapy and Survival/ life expectancy of lung cancer patients, can be rejected.

Radiotherapy can be prescribed before surgery to shrink the tumor or after surgery to eradicate the remaining parts of cancer that could not be removed surgically and in both scenarios, radiation can be given alone or with chemotherapy (American Cancer Society, 2021). Radiotherapy can also be prescribed to treat distant metastasis in bones and brain; as well as a palliative treatment in advanced cases of NSCLC to relieve symptoms such as pain or cough (American Cancer Society, 2021).

The predictor was a combination of Surgery, Chemotherapy, and Radiotherapy, and the outcome was Survival/ life expectancy. The predictor variable was found to be statistically significant [B = 3.447, 95% C.I. (3.118, 3.776), p < .001], indicating that for lung cancer patients receiving a combination of Surgery, Chemotherapy, and Radiotherapy, Survival/ life expectancy increases by approximately 0.01 month. The model explained approximately 0% of the variability [R-squared = 0.005]. A sensitivity analysis using G*power 3.1.9.7 software [n = 86,998; alpha error prob = 0.05; power = 0.80; predictors = 1] calculated [Effect size f^2 = 9.022085e-05] a small effect. The confidence interval associated with the regression analysis does not contain 0, which means the null hypothesis, there is no association between a combination of Surgery, Chemotherapy, and Radiotherapy, and Survival/ life expectancy of lung cancer patients, can be rejected.

Lu et al. (2019) conducted a study to evaluate the changes in incidence, treatment, and survival of lung cancer during the last four decades from 1973 to 2015. SEER database was used by the authors in their study (Lu et al., 2019). Joint regression models were used to estimate the changes in incidence, treatment, and survival related to lung cancer (Lu et al., 2019). The results based on SEER database, 1,148,341 patients were diagnosed with lung cancer during the period from 1973 to 2015 including 646,662 males and 501,679 females (whites = 960,808, black = 122,079, other races = 64,010, and unknown = 1,444) (Lu et al., 2019). The average incidence of lung cancer was 59.0/ 100,000 person/ year (Lu et al., 2019).

The incidence peak was in 1992 then gradually decreased with a higher incidence rate in males than females and blacks were higher than other racial groups (Lu et al., 2019). The surgical rate for lung cancer was 25%, an increased use of chemotherapy, and a decreased use of radiotherapy (Lu et al., 2019). The 5-year relative survival rate has increased with time but remained low (<21%) (Lu et al., 2019). Chemotherapy combined with radiotherapy were used at a higher rate in late stages of lung cancer than early stages (Lu et al., 2019). The study demonstrated relative decrease in the incidence of lung cancer in the past four decades which was due to advances in the treatment of lung cancer such as hormonal therapy, immunotherapy, and targeted therapy. The results and conclusion of this study validates my research study results about the effect of treatment regimen on lung cancer patients and survival/ life expectancy.

The association between a combination of Chemotherapy and Radiotherapy, Stage of lung cancer, Socioeconomic status, Race/ ethnicity, and Age account for [14%] of the variation in Survival/ life expectancy of lung cancer patients, which means that [86%] of the variation in Survival/ life expectancy of lung cancer patients cannot be explained by a combination of Chemotherapy and Radiotherapy, Stage of lung cancer, Socioeconomic status, Race/ ethnicity, and Age alone. A sensitivity analysis using G*power 3.1.9.7 software [n = 86,998; alpha error prob = 0.05; power = 0.80; predictors = 6] calculated [Effect size $f^2 = 0.0001566159$] a small effect. The confidence interval associated with the regression analysis does not contain 0, which means the null hypothesis, there is no association between a combination of Chemotherapy and

Radiotherapy, Stage of lung cancer, Socioeconomic status, Race/ ethnicity, Age and Survival/ life expectancy of lung cancer patients, can be rejected.

Clinical research is promising on increasing the overall survival (OS) with the application of chemotherapy combined with targeted agents and immunotherapy (Baxevanos & Mountzios, 2018). Chemotherapy combined with radiotherapy were used at a higher rate in late stages of lung cancer than early stages (Lu et al., 2019).

RQ5

The predictor was Stage of Lung Cancer, and the outcome was Survival/ life expectancy. The predictor variable was found to be statistically significant [B = -1.709, 95% C.I. (-1.742, -1.676), $p < .001$], indicating that for lung cancer patients with each increase of Stage of Lung Cancer (stage 1, stage 2, stage 3, stage 4), Survival/ life expectancy decreases by approximately 1.7 months. The model explained approximately 11% of the variability [R-squared = 0.105]. A sensitivity analysis using G*power 3.1.9.7 software [n = 86,998; alpha error prob = 0.05; power = 0.80; predictors = 1] calculated [Effect size f^2 = 9.022085e-05] a small effect. The confidence interval associated with the regression analysis does not contain 0, which means the null hypothesis, there is no association between Stages of Lung Cancer and Survival/ life expectancy of lung cancer patients, can be rejected.

Staging of cancer represents the anatomic extent of cancer in the human body which is a useful diagnostic and prognostic tool as well as an indicator for the effect of treatment (Amin et al., 2017). The grading system describes the microscopic changes in the appearance of cancer cells and tissues compared to normal body cells and tissues

(MD Anderson Cancer Center, 2021). The standard grading scale is from grade1 to grade 4 based on the level of changes from the normal body cells (MD Anderson Cancer Center, 2021). The grading system allows physicians to evaluate the extent of aggressiveness and prognosis of cancer which will facilitate developing a treatment plan accordingly (MD Anderson Cancer Center, 2021). Lung cancer when diagnosed during early stages is mostly curable and 56% of cases diagnosed with a localized cancer will be at a 5-year survival rate, while about 70% of patients diagnosed with advanced lung cancer, only 15% of them will be at a 5-year-survival rate (Azubuike et al., 2020).

A study was done by Mar et al. (2020) with the objective to evaluate the survival of lung cancer patients who were treated in Basque Health Service using TNM stage in 4-year periods (2003–2006, 2007–2010 and 2011–2014) and compare the results with survival of lung cancer patients in an equivalent sample of the general population. The researchers used in this study a retrospective observational design on a cohort of 11,635 patients retrieved from Euskadi hospital cancer registry (Mar et al., 2020). The variables included in the data were: TNM stage, age, sex, history, diagnosis date, vital status, and date of death (Mar et al., 2020). The results of the study revealed that lung cancer 5-year survival probability decreased with increase in stage, for lung cancer patients with stage I from 50%-65% and for lung cancer patients with stage IV from 2-3% (Mar et al., 2020). The researchers recommended more research on strategies that can improve the survival of lung cancer patients (Mar et al., 2020).

Lyu, (2020) investigated the risk factors that contribute to lung cancer survival. The investigator used Kaplan-Meier and Cox proportional hazard models as statistical

models to analyze 1,145 patients diagnosed with different types of lung cancer (Lyu, 2020). The dataset was extracted from The Cancer Genome Atlas Project (TCGA) database which is an organization that collects and stores huge numbers of gene data of cancer sequences that contributed to cancer treatment (Lyu, 2020). The dependent variable was survival of lung cancer patients, and the independent variables examined in this study were diagnosis, age, sex, smoking history, stage of lung cancer, fraction genome altered, and mutation count (Lyu, 2020). The results of the study revealed that the stage of lung cancer is the most influential factor for the survival of lung cancer patients (Lyu, 2020) which I was able to confirm in my study results of the effect of stage of lung cancer on survival/ life expectancy.

The impact of the other variables analyzed in this study such as age, sex, and smoking history had significant effect on survival of lung cancer patients only when they interact with time, reflecting their time-variant is associated with survival (Lyu, 2020).

RQ6

The predictor was Socioeconomic Status, and the outcome was Survival/ life expectancy. The predictor variable was found to be statistically significant [B = 0.135, 95% C.I. (0.118, 0.152), $p < .001$], indicating that for lung cancer patients with Socioeconomic status, suggests that for median household income level increase, Survival/ life expectancy increases by approximately 0.1 months or 3 days. The model explained approximately 0.3% of the variability [R-squared = 0.003]. A sensitivity analysis using G*power 3.1.9.7 software [n = 86,998; alpha error prob = 0.05; power = 0.80; predictors = 1] calculated [Effect size $f^2 = 9.022085e-05$] a small effect. The

confidence interval associated with the regression analysis does not contain 0, which means the null hypothesis, there is no association between Socioeconomic Status and Survival/ life expectancy of lung cancer patients, can be rejected.

Mahase et al. (2018) investigated survival disparities by regional poverty level based on radiotherapy treatment (RT) prescribed for lung cancer patients in these regions. The purpose of the study was to evaluate the regional poverty level reflected on differences in lung cancer (LC) survival for lung cancer patients received radiotherapy (Mahase et al., 2018). The investigators used retrospective quantitative method for a study sample retrieved from the Surveillance, Epidemiology, and End Results (SEER) database which included patients diagnosed with LC during the period from 2000 to 2009 (Mahase et al., 2018). The study sample was divided into socioeconomic status (SES) quintiles (five equal groups of populations divided based on distribution of values of the same variable) with quantile 1 represented the highest SES cohort and quantile 5 represented the lowest SES cohort (Mahase et al., 2018). The investigators concluded that RT may offer a positive survival benefit to those who received treatment when accounting for age, gender, race, and SES (Mahase et al., 2018). The study further concluded that an incrementally worse OS rate was associated with increasing regional poverty level even for those who received RT (Mahase et al., 2018). As such, my study results confirmed the conclusion of this study.

RQ7

The predictor was Race/ Ethnicity, and the outcome was Survival/ life expectancy. The predictor variable was found to be statistically significant [B = -0.218,

95% C.I. (-0.320, -0.116), $p < .001$], indicating that for lung cancer patients' non-Whites compared to Whites suggests having Survival/ life expectancy decreases by approximately 0.2 month or 6 days. The model explained approximately 1% of the variability [R-squared = 0.010]. A sensitivity analysis using G*power 3.1.9.7 software [n = 86,998; alpha error prob = 0.05; power = 0.80; predictors = 1] calculated [Effect size f^2 = 9.022085e-05] a small effect. The confidence interval associated with the regression analysis does not contain 0, which means the null hypothesis, there is no association between Race/ Ethnicity and Survival/ life expectancy of lung cancer patients, can be rejected.

Racial and ethnic minority groups, as defined by the Centers for Disease Control and Prevention (CDC), include people of color with a wide spectrum of backgrounds and experiences (CDC, 2022). Structural and interpersonal racism compromise mental and physical health of millions of people and challenge them to achieve their optimum health (CDC, 2022). As such, racial inequalities have a negative impact on the U.S. national health (Annangi et al., 2019; CDC, 2022). An increasing number of recent research studies concluded that centuries of racism in the U.S. has a profound negative impact on racial and ethnic minority groups (CDC, 2022). The impact of COVID-19 pandemic disproportionately on black Americans, Hispanics, American Indians, and other racial and ethnic groups has demonstrated the social and racial injustice and inequity those communities have experienced; and the impact of COVID-19 pandemic was only a reminder on how bad the health disparities is (CDC, 2022). COVID-19 data reflected that these communities have higher rates of COVID-19 related hospitalization and death

compared with non-Hispanic White populations (CDC, 2022). Such disparities persist despite controlling for socioeconomic factors and other demographic factors (CDC, 2022; Lake et al., 2020).

To investigate Is there an association between Age and Survival/ life expectancy of lung cancer patients after controlling for Chemotherapy, Stage of lung cancer, Socioeconomic status, Race/ ethnicity, and Geographical location (urban vs. rural)? a multiple linear regression analysis was conducted to evaluate the prediction of Survival/ life expectancy of lung cancer patients from Age, Stage of lung cancer, Socioeconomic status, Race/ ethnicity, and Geographical location (urban vs. rural). The results of the multiple linear regression analysis revealed Geographical location (urban vs. rural) not to be a statistically significant predictor to the model ($p > .05$). However, the results of the multiple linear regression analysis revealed a statistically significant association between Chemotherapy, Stage of lung cancer, Socioeconomic status, Race/ ethnicity, and Age.

Aging is considered an important risk factor for lung cancer due to accumulation of mutations in somatic cells during the individual's life span (Raniszewska et al., 2021). Although most of mutations do not affect cellular physiology or functions but, some mutated genes can interfere and modify cellular activities such as: Deregulation of cellular metabolism, resistance to cellular death, continuing cellular proliferation, interrupting growth suppressors, activate invasion and metastasis through angiogenesis and mutations (Raniszewska et al., 2021).

The logical connections between deductive approach theory (DAT) and the nature of my study include, DAT facilitates interpretation of causal relationships between

variables and concepts, (such as the association between different types of treatment, stage of lung cancer, socioeconomic status, and survival/ life expectancy), as well as possible quantitative measurements of concepts, and a possible generalizability of research findings to a certain level (Research Methodology, n.d.). Therefore, DAT serves as a conceivable framework for my dissertation. DAT has the capability of testing a known theory or a phenomenon if it is valid in given circumstances (Research Methodology, n.d.). DAT trails the path of logic (Research Methodology, n.d.).; therefore, I considered DAT as a framework for my dissertation topic.

The socioecological model can also be applied to my research study as a framework with multifaceted levels of the society where individuals and environment interact with the social system. Socioecological model acknowledges that different contributing factors and determinants exist at different levels of the society and addressing them at all levels (Individual stage, Interpersonal stage, Organizational stage, Community stage, and Public Policy stage) will facilitate more effective prevention and control. Taking actions at multiple levels at the same time will facilitate prevention of risk factors more effectively for any potential health problem (CDC, 2018). As such, actions to raise the awareness about lung cancer at all levels of the society, as stated in the socioecological model, will facilitate services and funding to better improve prevention and control of lung cancer which represent the logical connection between my research study and the socioecological model.

Limitations of the Study

The limitations in the design of my study can be expressed in the cross-sectional design that I used to examine the relationship between independent variables (treatment such as surgery, chemotherapy, radiotherapy, combination of therapy, stages of lung cancer, socioeconomic status) and dependent variable (survival/ life expectancy). That is because, cross-sectional design provides inferences about relationship between different variables, but it cannot demonstrate the cause and effect between independent and dependent variables because independent variables (risk factors) and dependent variable (outcome) are measured at the same time (Creswell & Creswell, 2018). Also, another limitation of cross-sectional design is that the temporal link between the outcome (life expectancy) and the exposure (treatment such as surgery, chemotherapy, radiotherapy, combination of therapy, stages of lung cancer, socioeconomic status) cannot be determined because both (the outcome and the exposure) are examined at the same time (Creswell & Creswell, 2018). The advantage of performing a cross-sectional design is that it allowed me to compare multiple variables at the same time and it saved me time and cost (Creswell & Creswell, 2018). Cross-sectional design is also useful in assessing the burden of disease in a defined population such as the disadvantaged communities in my study and the health needs of a population which both are useful in planning and allocation of health resources (Creswell & Creswell, 2018). As such, these advantages outweighed the limitations of the cross-sectional design which encouraged me to use in my study design.

The limitations of the quantitative method that I conducted in my study are Sampling errors that occur when the sample used does not represent the general population which is called sample bias or selection bias. I used a random population sample for my dataset to overcome this limitation; insufficient sample size for statistical analyses was another limitation which affects the validity of the conclusion. To overcome this limitation, I used G*power 3.1.9.7 software for power analysis to minimize bias on sample size.

The internal validity threats that could be related to my study were selection bias which I avoided by random selection of participants. Also, internal validity threats could be due to mortality or study attrition which I avoided by selection of a large study sample (n=86,998) to avoid dropouts, incomplete records, or mortality due to other causes.

Recommendations

With this comprehensive study about Effect of Treatment, Stage of Lung Cancer, and socioeconomic status on Life Expectancy of Marginalized Communities, I recommend to the government leaders and elected officials to adopt, review, and enact policies that address and urge the need for screening high risk people for lung cancer especially among underserved and marginalized communities through health education and making screening more affordable. This will encourage people to undergo routine lung cancer screening and in turn it will facilitate early diagnosis, potential curable and less expensive treatment which will improve quality of life and increase life expectancy. Also, this will lower health care expenditure especially 60% of cancer affects underserved and disadvantage communities (Montagne et al., 2021).

For hospital administrations, Health insurance companies, and health care providers, my recommendations, based on this study, will be to work on closing the evidence-practice gaps across the diagnosis and management process of lung cancer that require clinical practice to align with the recommended evidence-based guidelines to facilitate improvement of the outcome (Rankin et al., 2018). The evidence-practice gaps as identified by Rankin et al. (2018) can be summarized in delays of diagnosis and referrals, underutilization of curative and palliative treatments, treatment is influenced by older age and comorbidities, multidisciplinary team is not part of most lung cancer management team, and psychosocial support is not utilized as a part of lung cancer care (McGregor et al., 2017; Rankin et al., 2018).

Screening should be promoted by health care providers and utilize simple less expensive test such blood screening for tumor markers and LDCT; Improving health literacy through community involvement and improving patient-physician communications; lung cancer management should be handled by a team of professionals not a single health care provider as it requires a multidisciplinary team that may include a pulmonologist or internist, oncologist, surgeon, psychiatrist or psychologist, social worker, and a nurse (Hung et al., 2020; Rankin et al., 2020), as the evaluation and treatment approach should be orchestrated among different specialties to select the best treatment plan (McCann et al., 2021); the evidence-practice gap should be addressed by training health care providers to learn and practice the advanced recommended methods of screening and investigations as well as applying the new recommended treatment protocols (Rankin et al., 2020); promote health education among communities about the

risk factors and carcinogens that contribute to lung cancer. Encourage more new research about the epidemiology of lung cancer among marginalized communities and explore new approaches to improve prevention and control of lung cancer among marginalized communities.

The incidence and prevalence of lung cancer as well as the mortality rate cannot be improved without addressing the effect of treatment, stage of lung cancer, and socioeconomic status on life expectancy of marginalized communities since about 60% of cancer affects underserved and disadvantaged communities (Montagne et al., 2021).

The advancement in diagnostic technologies as well as treatment modalities such as immunotherapy and targeted therapy have contributed so much to the improvement of control and treatment of lung cancer which in turn have contributed to improvement of quality of life and survival/ life expectancy. Therefore, I recommend further research studies on the effect of advanced diagnostic technologies and new treatment modalities such as immunotherapy and targeted therapy on survival/ life expectancy of lung cancer patients with special emphasis on marginalized and underserved communities.

Implications

The potential implications for positive social change that are consistent with my study will be addressing the burden of lung cancer specially among marginalized and underserved communities such as avoiding risk factors, screening, early diagnosis, and early treatment (Osarogiagbon, 2018; WHO, 2019). Therefore, screening is an important tool that will facilitate early detection and diagnosis of lung cancer. Screening can contribute to early treatment which could be more effective and less expensive treatment

with a better outcome (WHO, 2019). This will explain the importance of my research study about The Effect of Treatment, Stage of Lung Cancer, and Socioeconomic Status on Life Expectancy of Marginalized Communities where many risk factors for lung cancer and barriers to health care are inherited and need to be analyzed and addressed to facilitate better government policies and raise awareness of the communities. The assessment of survival of cancer patients at the population level is a major contributor to the decoration of healthcare policy for lung cancer (Mar et al., 2020).

Rivera et al. (2020) performed a study with the objectives to gather essential available knowledge on disparities for lung cancer screening (LCS) that is characterized by eligibility criteria, access, and implementation to facilitate development of an official statement adopted by the American Thoracic Society to facilitate improving current screening guidelines and allocation of resources for unbiased LCS (Rivera et al., 2020). The authors based their study on available resources that identify disparities in lung cancer outcome among populations with the understanding that LCS can contribute to reduction of mortality due to lung cancer (Rivera et al., 2020). The method used in the study was a multidisciplinary panel that was represented by experts in LCS implementation science, primary healthcare, pulmonary, health behavior, smoking cessation, and epidemiology who all participated jointly on disparity research (Rivera et al., 2020). The multidisciplinary panel investigated available literature on disparity in cancer screening from the historical and emerging evidence perspectives (Rivera et al., 2020).

The results of the study revealed that current LCS guidelines do not consider lung cancer risk, smoking behavior differences between genders, socioeconomic status, race, and ethnicity (Rivera et al., 2020). The results also identified multiple barriers such as cost, and access to screening which contribute to disparity in implementation and dissemination of LCS (Rivera et al., 2020). The statement that resulted from this study identified the impact of LCS eligibility criteria on vulnerable communities who are considered at high risk for lung cancer but do not meet eligibility criteria for screening (Rivera et al., 2020). Furthermore, multiple barriers impact disparity in LCS implementation such as multilevel barriers (Rivera et al., 2020). The authors recommended strategies that accommodate vulnerable populations for unbiased selection and dissemination of LCS such as addressing racial, ethnic, gender-based differences in smoking behaviors, socioeconomic, risk for lung cancer, access for LCS through health insurance coverage, provide unbiased LCS resources for vulnerable communities, and provide education and resources that contribute to impartial LCS results (Rivera et al., 2020). This study analyzed the contributing factors for the selection bias for lung cancer screening which will lead to a delay in the diagnosis and treatment of lung cancer and will negatively affect the life expectancy. Therefore, this study validates my research study about the effect of multiple contributing factors such as treatment, stage of lung cancer, socioeconomic status, race/ ethnicity, education, access to health care on life expectancy of lung cancer patients in marginalized communities.

Conclusion

My research study has contributed to fill a gap in literature which is reflected by a scarcity of research about the Effect of Treatment, Stage of Lung Cancer, and Socioeconomic Status on Life Expectancy of Marginalized Communities despite About 60% of cancer affects underserved and disadvantaged communities (Montagne et al., 2021). The study of these contributing factors (Treatment, Stage of Lung Cancer, and Socioeconomic) on survival/ life expectancy that included analyses and results based on a national database (SEER) and confirming that they are statistically significant would likely be the first dissertation study that included all of them together in one study. As such, this reflected how each predictor contributed to survival/ life expectancy and their impact all together combined on survival/ life expectancy.

Also, my research study has contributed to address a gap in practice which is represented by evidence-practice gaps across the diagnosis and management process of lung cancer that require clinical practice to align with the recommended evidence-based guidelines to facilitate improvement of the outcome (Rankin et al., 2018). That is because closing the evidence-practice gap will contribute to screening and early diagnosis that will facilitate a possible curable treatment, a better prognosis, and a longer survival/ life expectancy. However, the evidence-practice gap was addressed only in clinical research literature; and despite its public health and epidemiological importance it was not addressed in public health or epidemiology research. Therefore, addressing the evidence-practice gaps across the diagnosis and management process of lung cancer should be one of the initiatives of public health due to its implications on alleviating the burden of lung

cancer. The evidence-practice gaps as identified by Rankin et al. (2018) can be summarized in delays of diagnosis and referrals, underutilization of curative and palliative treatments, treatment is influenced by older age and comorbidities, multidisciplinary team is not part of most lung cancer management team, and psychosocial support is not utilized as a part of lung cancer care (McGregor et al., 2017; Rankin et al., 2018).

The social change that could be achieved by this extensive study and analyses in this dissertation will be enacting policies that address the core of these multifactorial health and social problems as well as review of the polices that did not deliver to resolve these problems. Collaborative efforts from all sectors of the community such as individuals, schools, private and government stakeholders should be involved in addressing these health and social problems. As that would eventually improve the survival/ life expectancy of marginalized and underserved communities.

References

Adamowicz, K., Janiszewska, J., & Lichodziejewska-Niemierko, M. (2020). Prognostic value of patient knowledge of cancer on quality of life in advanced lung cancer during chemotherapy. *Journal of Cancer Education*, *35*(1), 93–99. https://doi-org.ezp.waldenulibrary.org/10.1007/s13187-018-1444-3

American Association of Cancer Research. (2011). *Glossary of cancer research terms.* http://www.aacr.org/home/survivors--advocates/glossary-of-cancer-terms-k-z.aspx#R

American Cancer Society. (2021). Treating non-small cell lung cancer. https://www.cancer.org/cancer/acs-medical-content-and-news-staff.html

American Lung Association. (2020). State of lung cancer, state data. https://www.lung.org/our-initiatives/research/monitoring-trends-in-lungdisease/state-of-lung-cancer/states/

Amin, M. B., Greene, F. L., Edge, S. B., Compton, C. C., Gershenwald, J. E., Brookland, R. K., Meyer, L., Gress, D. M., Byrd, D. R., & Winchester, D. P. (2017). The eighth edition of the American Joint Committee on Cancer (AJCC) Cancer Staging Manual: Continuing to build a bridge from a population-based to a more "personalized" approach to cancer staging. *Cancer Journal for Clinicians, 67*(2), 93-99. https://doi.org/10.3322/caac.21388

Annangi, S., Nutalapati, S., Foreman, M. G., Pillai, R., & Flenaugh, E. L. (2019). Potential racial disparities using current lung cancer screening. *Journal of Racial*

and Ethnic Health Disparities, 6(1), 22–26. https://doi.org/10.1007/s40615-018-0492-z

Atwater, T. & Massion, P. P. (2016). Biomarkers of risk to develop lung cancer in the new screening era. *Annual Translational Medicine, 4*(8), 1–6. https://doi.org/10.21037/atm.2016.03.46

Azubuike, U.C., Cooper, D., & Aplin-Snider, C. (2020). Using United States preventive services task force guidelines to improve a family medicine clinic's lung cancer screening rate: A quality improvement project. *The Journal for Nurse Practitioners, 16*(10), e169-e172. https://doi.org/10.1016/j.nurpra.2020.07.001

Bakulski, K. M., Dou, J., Lin, N., & London, S. J. (2019). DNA methylation signature of smoking in lung cancer is enriched for exposure signatures in newborn and adult blood. *Scientific Reports, 9*(4576), 1–13. https://doi.org/10.1038/s41598-019-40963-2

Baxevanos, P., & Mountzios, G. (2018). Novel chemotherapy regimens for advanced lung cancer: Have we reached a plateau? *Annals of Translational Medicine, 6*(8), 139. https://doi.org/10.21037/atm.2018.04.04

Borrayo, E. A., Scott, K. L., Drennen, A., Bendriss, T. M., Kilbourn, K. M., & Valverde, P. (2020). Treatment challenges and support needs of underserved Hispanic patients diagnosed with lung cancer and head-and-neck cancer. *Journal of Psychosocial Oncology, 38*(4), 449–462. https://doi.org/10.1080/07347332.2019.1705453

Bossé, Y., & Amos, C. (2019). A decade of genome-wide association studies

(GWAS) results in lung cancer. *Cancer Epidemiology, Biomarkers, Prevention, 27*(4), 363–379. https://doi.org/10.1158/1055-9965.EPI-16-0794

Botta, L., Dal Maso, L., Guzzinati, S., Panato, C., Gatta, G., Trama, A., Rugge, M., Tagliabue, G., Casella, C., Caruso, B., Michiara, M., Ferretti, S., Sensi, F., Tumino, R., Toffolutti, F., Russo, A. G., Caiazzo, A. L., Mangone, L., Mazzucco, W., Iacovacci, S., … AIRTUM Working Group (2019). Changes in life expectancy for cancer patients over time since diagnosis. *Journal of Advanced Research, 20*, 153–159. https://doi.org/10.1016/j.jare.2019.07.002

Bürkle, A., Moreno-Villanueva, M., Bernhard, J., Blasco, M., Zondag, G., Hoeijmakers, J. H., Toussaint, O., Grubeck-Loebenstein, B., Mocchegiani, E., Collino, S., Gonos, E. S., Sikora, E., Gradinaru, D., Dollé, M., Salmon, M., Kristensen, P., Griffiths, H. R., Libert, C., Grune, T., Breusing, N., … Aspinall, R. (2015). Mark-age biomarkers of ageing. *Mechanisms of Ageing and Development, 151*, 2–12. https://doi.org/10.1016/j.mad.2015.03.006

Caballero-Vázquez, A., Romero-Béjar, J. L., Albendín-García, L., Suleiman-Martos, N., Gómez-Urquiza, J. L., Cañadas, G. R., & Cañadas-De la Fuente, G. A. (2021). Risk factors for short-term lung cancer survival. *Journal of Clinical Medicine, 10*(3), 519. https://doi.org/10.3390/jcm10030519

Cancer Treatment Centers of America. (2020). What you should know about lung cancer. https://www.cancercenter.com/cancer-types/lung-cancer/about

Carey, T., Bekemeier, B., Campos-Outcalt, D., Koch-Weser, S., Millon-Underwood S., & Teutsch, S. (2020). Achieving health equity in preventive services. National

Institutes of Health pathways to prevention workshop. *Annals of Internal Medicine, 172*(4). Doi 10.7326/M19-3171

Centers for Disease Control and Prevention. (2022). Health equity considerations and racial and ethnic minority group. https://www.cdc.gov/coronavirus/2019-ncov/community/health-equity/race-ethnicity.html

Centers for Disease Control and Prevention. (n.d.). Principles of epidemiology in public health practice, (3rd ed.) an introduction to applied epidemiology and biostatistics. https://www.cdc.gov/csels/dsepd/ss1978/glossary.html

Chiang, I. C., Jhangiani, R. S., & Price, P. C. (2017). *Research Methods in Psychology.* (2nd.ed). *Correlational Research, 7*, 130-141. BCcampus Open Education. https://opentextbc.ca/researchmethods/chapter/correlational-research/

Chu, G. C. W., Lazare, K., & Sullivan, F. (2018). Serum and blood-based biomarkers for lung cancer screening: A systematic review. *BioMed Central, 18*(181), 1–6. https://doi.org/10.1186/s12885-018-4024-3

Costa, F. A., Ramos, C., Murteira, R., Almodovar, T., Passos-Coelho, J. L., Carvalho, M. I., Costa, L., Brito, M. J., Ramos, S., Ferreira, M., & Miranda, A. C. (2019). The cancer registry as an ally in monitoring treatment effectiveness. *Pulmonology*, *25*(1), 3–8. https://doi.org/10.1016/j.pulmoe.2018.05.007

Creswell, J. W., & Creswell, J. D. (2018). *Research design: Qualitative, quantitative, and mixed methods* (5th ed.). *Review of the Literature 2*, 23-46. Sage

Davis, A. P., Boyer, M., Lee, J. H., & Kao, S. C. (2020). COVID-19: The use of immunotherapy in metastatic lung cancer. *Immunotherapy, 12*(8), 545-545–548. Doi: http://dx.doi.org.ezp.waldenulibrary.org/10.2217/imt-2020-0096

Dima, S., Chen, K-H., Wang, K-J., Wang, K-M., & Teng, N-C. (2018). Effect of comorbidity on lung cancer diagnosis timing and mortality: A nationwide population-based cohort study in Taiwan. Hindawi. *BioMed Research International*, (1252897), 9. https://doi.org/10.1155/2018/1252897

Folch, E., Costa, D.B.., Wright, J., VanderLaan, P. A. (2015). Lung cancer diagnosis and staging in the minimally invasive age with increasing demands for tissue analysis. *Translational Lung Cancer Research. 4*(4):392-403. Doi 10.3978/j.issn.2218-6751.2015.08.02

Gingras, M. (2020). Age at diagnosis and lung cancer presentation. (28259651). [PhD thesis, Walden University]. ProQuest

Gingras, M. (2018). Scholar-practitioner final project. PUBH-8540: Epidemiology topic seminar A00563966 biomarkers screening for detection of lung and bronchus cancer: a systematic review Abstract.

Harris, J. (2011). Glossary of public health terms. National Conference of State Legislatures, Public Health Accreditation Board. https://idph.iowa.gov/Portals/1/Files/LPHS/LBOH%2010_glossary.pdf

Huang, J. Y., Larose, T. L., Luu, H. N., Wang, R., Fanidi, A., Alcala, K…& Yuan, J.-M. (2019). Circulating markers of cellular immune activation in prediagnostic blood sample and lung cancer risk in the lung cancer cohort consortium (LC3).

International Journal of Cancer, 00(00–00), 1–12. https://doi.org/10.1002/ijc.32555

Hung, H. Y., Tseng, Y. H., Chao, H. S., Chiu, C. H., Hsu, W. H., Hsu, H. S., Wu, Y. C., Chou, T. Y., Chen, C. K., Lan, K. L., Chen, Y. W., Wu, Y. H., & Chen, Y. M. (2020). Multidisciplinary team discussion results in survival benefit for patients with stage III non-small-cell lung cancer. *PloS One, 15(*10), e0236503. https://doi-org.ezp.waldenulibrary.org/10.1371/journal.pone.0236503

Kelley, D. E., Kent, E. E., Litzelman, K., Mollica, M. A., & Rowland, J. H. (2019). Dyadic associations between perceived social support and cancer patient and caregiver health: An actor-partner interdependence modeling approach. *Psycho - Oncology, 28(7)*, 1453-1460. Doi: http://dx.doi.org.ezp.waldenulibrary.org/10.1002/pon.5096

Kelly, C., Hulme, C., Faragher, T., & Clarke, G. (2016). Are differences in travel time or distance to healthcare for adults in global north countries associated with an impact on health outcomes? A systematic review. *BMJ 6* (11). 6: e013059.doi:10.1136/bmjopen-2016-013059

Khorana, A.A., Tullio, K., Elson, P., Pennell, N. A., Grobmyer, S. R., Kalady, M. F., Raymond, D., Abraham, J., Klein, E. A., Walsh, R. M., Monteleone, E.E., Wei, W., Hobbs, B., & Bolwell, B. J. (2019). Time to initial cancer treatment in the United States and association with survival over time: An observational study. *PLOS ONE 14(4)*: e0215108. https://doi.org/10.1371/journal.pone.0213209

Kim, M. L., Matheson, L., Garrard, B., Francis, M., Broad, A., Malone, J., . . . Cheng-Hon Yap. (2019). Use of clinical quality indicators to improve lung cancer care in a regional/rural network of health services. *Australian Journal of Rural Health, 27*(2), 183-187. Doi: http://dx.doi.org.ezp.waldenulibrary.org/10.1111/ajr.12493

Ko, E.C., Raben, D., & Formenti, S.C. (2018). The integration of radiotherapy with immunotherapy for the treatment of non–small cell lung cancer. *Clinical Cancer Research 24*(23). DOI: 10.1158/1078-0432.CCR-17-3620

Lake, M., Shusted, C. S., Hee-Soon J., McIntire, R. K., Zeigler-Johnson, C., Evans, N. R., . . . Barta, J. A. (2020). Black patients referred to a lung cancer screening program experience lower rates of screening and longer time to follow-up. *BMC Cancer, 20,* 1-12. Doi: http://dx.doi.org.ezp.waldenulibrary.org/10.1186/s12885-020-06923-0

Lawrie, K., Charow, R., Giuliani, M., & Papadakos, J. (2020). Homelessness, cancer, and health literacy: A scoping review. *Journal of Health Care for the Poor and Underserved, 31*(1), 81-104. Doi: http://dx.doi.org.ezp.waldenulibrary.org/10.1353/hpu.2020.0010

Lazzari, C., Karachaliou, N., Bulotta, A., Viganó, M., Mirabile, A., Brioschi, E., Santarpia, M., Gianni, L., Rosell, R., & Gregorc, V. (2018). Combination of immunotherapy with chemotherapy and radiotherapy in lung cancer: is this the beginning of the end for cancer? *Therapeutic advances in medical oncology, 10*, 1758835918762094. https://doi.org/10.1177/1758835918762094.

Li, J. J. N., Karim, K., Sung, M., Le, L. W., Lau, S. C. M., Sacher, A., & Leighl, N. B. (2020). Tobacco exposure and immunotherapy response in PD-L1 positive lung cancer patients. *Lung Cancer (Amsterdam, Netherlands), 150,* 159–163. https://doi-org.ezp.waldenulibrary.org/10.1016/j.lungcan.2020.10.023

Li, R.Y., & Liang, Z.Y. (2020). Circulating tumor DNA in lung cancer: real-time monitoring of disease evolution and treatment response. *Chinese Medical Journal, 133*(20), 2476–2485. https://doi-org.ezp.waldenulibrary.org/10.1097/CM9.0000000000001097

Liu, W., Liu, A., Chan, J., Boldt, R. G., Munoz-Schuffenegger, P., & Louie, A. V. (2019). What is the optimal radiotherapy utilization rate for lung cancer? a systematic review. *Translational lung cancer research, 8*(Suppl 2), S163–S171. https://doi.org/10.21037/tlcr.2019.08.12

Lozano, M., Echeveste, J. I., Abengozar, M., Meijias, L. D., Idoate, M. A., Calvo, A…& Andrea, C. E. (2018). Cytology smears in the era of molecular biomarkers in non-small cell lung cancer – Doing more with less. *Molecular testing on cytology smears, 142*, 291–298. https://doi.org/ 10.5858/arpa.2017-0208-RA.

Lung Cancer Foundation of America. (2020). Lung cancer facts. https://lcfamerica.org/lung-cancer-info/lung-cancer-facts/#1543338163380-b2df265a-237f

Lu, T., Yang, X., Huang, Y., Zhao, M., Li, M., Ma, K., Yin, J., Zhan, C., Wang, Q. (2019). Trends in the incidence, treatment, and survival of patients with lung

cancer in the last four decades. 11; 943-53. *Dove Medical Press Limited*. DOI https://doi.org/10.2147/CMAR.S187317

Lyu, R. (2020). Survival analysis of lung cancer patients from TCGA Cohort. *Advances in Lung Cancer, 9,* 1-15. Doi: 10.4236/alc.2020.91001.

Mahase, S., Christos, P., Wang, X., Potters, L., Wernicke, A. G., & Parashar, B. (2018). Survival disparities in the radiotherapeutic management of lung cancer by regional poverty level. *Cureus, 10*(11), e3575. https://doi.org/10.7759/cureus.3575

Mar, J., Arrospide, A., Iruretagoiena, M. L., Clèries, R., Paredes, A., Elejoste, I., . . . Ibarrondo, O. (2020). Changes in lung cancer survival by TNM stage in the Basque country from 2003 to 2014 according to period of diagnosis. *Cancer Epidemiology, 65* Doi: http://dx.doi.org.ezp.waldenulibrary.org/10.1016/j.canep.2020.101668

Mayo Clinic (2019). Lung cancer. https://www.mayoclinic.org/diseases-conditions/lung-cancer/symptomscauses/syc-20374620

McCann, B., Muhr, R., O'Rourke, N., Milroy, R., Kollmeier, J., Misch, D., Horst, J. vd., Morrison, D., Bauer, T., Massalski, O., & Blum, T.G. (2021). ADVANCE-1: An adapted collaborative benchmarking approach in centre-based lung cancer care, *Lung Cancer* (151) 44-52. ISSN 0169-5002, https://doi.org/10.1016/j.lungcan.2020.11.019. (https://www.sciencedirect.com/science/article/pii/S0169500220306942)

McDonald, J.H. (2014). *Handbook of biological statistics (3rd ed.).* multiple regression. Sparky House Publishing.

McGregor, D., Rankin, N., Butow, P., York, S., White, K., Phillips, J., Stone, E., Barnes, D., Jones, R., & Shaw, T. (2017). Closing evidence-practice gaps in lung cancer: Results from multi-methods priority setting in the clinical context. *Asia-Pacific Journal of Clinical Oncology, 13*(1), 28-36. https://doi.org/10.1111/ajco.12499

McLouth, L. E., Weyman, K., Golden, S. L., Cheavens, J. S., Peterman, A., Bursac, V., . . . Weaver, K. E. (2021). Developing pathways, a hope-enhancing intervention for metastatic lung cancer patients receiving cancer treatment. *Psycho - Oncology, 30*(6), 863-873. Doi: http://dx.doi.org.ezp.waldenulibrary.org/10.1002/pon.5650

Melzer, A. C., Golden, S. E., Ono, S. S., Datta, S., Triplette, M., & Slatore, C.G. (2020). We just never have enough time. clinician views of lung cancer screening processes and implementation. *Annals of the American Thoracic Society.17*(10). Doi 10.1513/AnnalsATS.202003-262OC

Montagne, F., Guisier, F., Venissac, N., & Baste, J.-M. (2021). The Role of surgery in lung cancer treatment: Present indications and future perspectives. *Cancers, 13*(15), 3711. doi:10.3390/cancers13153711

National Cancer Institute. (2021). About cancer. National Institute of Health. https://www.cancer.gov/about-cancer/understanding/what-is-cancer#definition

National Cancer Institute. (n.d.). NCI Dictionary of Cancer Terms. https://www.cancer.gov/publications/dictionaries/cancer terms/search/carcinogen/?searchMode=Begins

National Cancer Institute. (n.d.). SEER is an authoritative source for cancer statistics in the United States. https://seer.cancer.gov/

National Institute of Health. (2021). Health literacy. https://www.nih.gov/institutes-nih/nih-office-director/office-communications-public-liaison/clear-communication/health-literacy

Nilsson, J., Berglund, A., Bergstrom, S., Bergqvist, M., & Lambe, M. (2017). The role of comorbidity in the management and prognosis in non-small cell lung cancer: A population-based study. *Acta Oncologica, 56:*7, 949-956, Doi*:* https://doi.org/10.1080/0284186X.2017.1324213

Ni, M., Liu, X., Wu, J., Zhang, D., Tian, J., Wang, T., & Zhang, X. (2018). Identification of candidate biomarkers correlated with the pathogenesis and prognosis of non-small cell lung cancer via integrated bioinformatics analysis. *Frontiers in Genetics, 9*(469), 1–14. https://doi.org/10.3389/fgene.2018.00469

Osarogiagbon, R. (2018). Making the evidentiary case for universal multidisciplinary thoracic oncologic care. *Clinical Lung Cancer 19(4).* Doi 10.1016/j.cllc.2018.05.006

Pu, H. Y., Xu, R., Zhang, M. Y., Yuan, L. J., Hu, J. Y., Huang, G. L., & Wang, H. Y. (2017). Identification of microRNA-615-3p as a novel tumor suppressor in non-small cell lung cancer. *Oncology, 13,* 2403–2410. https://doi.org/10.3892/ol.2017.5684

Raghavan, D., Wheeler, M., Doege, D., Doty, J. D., 2nd, Levy, H., Dungan, K. A., Davis, L. M., Robinson, J. M., Kim, E. S., Mileham, K. F., Oliver, J., & Carrizosa, D.

(2020). Initial results from mobile low-dose computerized tomographic lung cancer screening Unit: Improved outcomes for underserved populations. *The Oncologist, 25*(5), e777–e781. https://doi-org.ezp.waldenulibrary.org/10.1634/theoncologist.2019-0802

Raniszewska, A., Kwiecień, I., Rutkowska, E., Rzepecki, P., & Domagała-Kulawik, J. (2021). Lung cancer stem cells-origin, diagnostic techniques, and perspective for therapies. *cancers, 13*(12), 2996. https://doi.org/10.3390/cancers13122996

Rankin, N., Fradgley, E., & Barnes, D. (2020). Implementation of lung cancer multidisciplinary teams: A review of evidence-practice gaps. *Translational Lung Cancer Research. 9(4).* Doi 10.21037/tlcr.2019.11.32

Rankin, N. M., McGregor, D., Stone, E., Butow, P. N., Young, J. M., White, K., & Shaw, T. (2018). Evidence-practice gaps in lung cancer: A scoping review. *European Journal of Cancer Care. 27(2).* Doi 10.1111/ecc.12588

Research Methodology. (n.d.). Deductive approach (deductive reasoning). https://research-methodology.net/research-methodology/research-approach/deductive-approach-2

Rivera, M. P., Katki, H. A., Tanner, N. T., Triplette, M., Sakoda, L. C., Wiener, R. S., Cardarelli, R., Carter-Harris, L., Crothers, K., Fathi, J. T., Ford, M. E., Smith, R., Winn, R. A., Wisnivesky, J. P., Henderson, L. M., & Aldrich, M. C. (2020). Addressing disparities in lung cancer screening eligibility and healthcare access. An official American thoracic society statement. *American journal of respiratory*

and critical care medicine, 202(7), e95–e112. https://doi.org/10.1164/rccm.202008-3053ST

Roberti, A., Valdes, A. F., Torrecillas, R. Fraga, M. F., & Fernandez, A. F. (2019). Epigenetics in cancer therapy and nanomedicine. *Clinical Epigenetics, 11*(81), 1–18. https://doi.org/10.1186/s13148-019-0675-4

Rudestam, K. E., & Newton, R. R. (2015). *Surviving your dissertation.* (4th ed.), 27-30. ISBN- 13: 978-1452260976-ISBN-10: 1452260974. SAGE

Salamanca, J. C., Meehan-Atrash, J., Vreeke, S., Escobedo, J. O., Peyton, D. H., & Strongin, R. M. (2018). E-cigarettes can emit formaldehyde at high levels under conditions that have been reported to be non-averse to users. Scientific Reports, 8(7559), 1–6. https://doi.org/10.1038/s41598-018-25907-6

Smith, M. T., Guyton, K. Z., Gibbons, C. F., Fritz, J. M., Portier, C. J., Rusyn, I., DeMarini, D. M., Caldwell, J. C., Kavlock, R. J., Lambert, P. F., Hecht, S. S., Bucher, J. R., Stewart, B. W., Baan, R. A., Cogliano, V. J., & Straif, K. (2016). Key characteristics of carcinogens as a basis for organizing Data on mechanisms of carcinogenesis. *Environmental health perspectives, 124*(6), 713–721. https://doi.org/10.1289/ehp.1509912

U. S. Cancer Statistic. (n.d.). Rate of cancer deaths in the United States. https://gis.cdc.gov/Cancer/USCS/DataViz.html

VoPham, T., Bertrand, K. A., Fisher, J. A., Ward, M. H., Laden, F., & Jones, R. R. (2022). Emissions of dioxins and dioxin-like compounds and incidence of

hepatocellular carcinoma in the United States. *Environmental research*, 204(Pt D), 112386. https://doi.org/10.1016/j.envres.2021.112386

Wang, C., Liang, H., Lin, C., Li, F., Xie, G., Qiao, S., & Zhang, X. (2019). Molecular subtyping and prognostic assessment based on tumor mutation burden in patients with lung adenocarcinomas. *International Journal of Molecular Sciences, 20*(4251), 1–13. https://doi.org/10.3390/ijms20174251

World Health Organization. (2019). Cancer: key facts. http://www.who.int/en/news-room/fact-sheets/detail/cancer

World Health Organization. (2019). Cancer. https://www.who.int/cancer/en/

World Health Organization. (2011). Biomarker and Human Biomonitoring. https://www.who.int/health-topics/children-environmental-health

Wu, L. L., Li, C. W., Lin, W. K., Qiu, L. H., & Xie, D. (2021). Incidence and survival analyses for occult lung cancer between 2004 and 2015: a population-based study. *BMC cancer*, *21*(1), 1009. https://doi.org/10.1186/s12885-021-08741-4

Zhang, Z., Yuan, F., Chen, R., Li, Y., Ma, J., Yan, X., Wang, L., Zhang, F., Tao, H., Guo, D., Huang, Z., Zhang, S., Li, X., Zhi, X., Ge, X., Hu, Y., & Wang, J. (2020). Dynamics of serum tumor markers can serve as a prognostic biomarker for Chinese advanced non-small cell lung cancer patients treated with immune checkpoint inhibitors. *Frontiers in Immunology, 11,* 1173. https://doi.org.ezp.waldenulibrary.org/10.3389/fimmu.2020.01173

Appendix A: Walden University IRB Approval Letter

IRB Materials Approved - Elsadig Elsharif

IRB <irb@mail.waldenu.edu>
Mon 8/8/2022 4:23 PM
To: Elsadig Elsharif <elsadig.elsharif@waldenu.edu>
Cc: Zin M. Htway <zin.htway@mail.waldenu.edu>;IRB <irb@mail.waldenu.edu>

Dear Elsadig Elsharif,

This email is to notify you that the Institutional Review Board (IRB) confirms that your doctoral capstone entitled, "The Effect of Treatment, Stage of Lung Cancer, and Socioeconomic Status on Life Expectancy of Marginalized Communities," meets Walden University's ethical standards. Since this project will serve as a Walden doctoral capstone, the Walden IRB will oversee your capstone data analysis and results reporting. Your IRB approval number is 08-08-22-0647642, which expires when your student status ends.

This confirmation is contingent upon your adherence to the exact procedures described in the final version of the documents that have been submitted to IRB@mail.waldenu.edu as of this date. This includes maintaining your current status with the university and the oversight relationship is only valid while you are an actively enrolled student at Walden University. If you need to take a leave of absence or are otherwise unable to remain actively enrolled, this is suspended.

If you need to make any changes to the project staff or procedures, you must obtain IRB approval by submitting the IRB Request for Change in Procedures Form. You will receive confirmation with a status update of the request within 10 business days of submitting the change request form and are not permitted to implement changes prior to receiving approval. Please note that Walden University does not accept responsibility or liability for research activities conducted without the IRB's approval, and the University will not accept or grant credit for student work that fails to comply with the policies and procedures related to ethical standards in research.

When you submitted your IRB materials, you made a commitment to communicate both discrete adverse events and general problems to the IRB within 1 week of their occurrence/realization. Failure to do so may result in invalidation of data, loss of academic credit, and/or loss of legal protections otherwise available to the researcher.

Both the Adverse Event Reporting form and Request for Change in Procedures form can be obtained on the Tools and Guides page of the Walden website: https://academicguides.waldenu.edu/research-center/research-ethics/tools-guides

Doctoral researchers are required to fulfill all of the Student Handbook's Doctoral Student Responsibilities Regarding Research Data regarding raw data retention and dataset confidentiality, as well as logging of all recruitment, data collection, and data management steps. If, in the future, you require copies of the originally submitted IRB materials, you may request them from Institutional Review Board.

Both students and faculty are invited to provide feedback on this IRB experience at the link below:

http://www.surveymonkey.com/s.aspx?sm=qHBJzkJMUx43pZegKlmdiQ_3d_3d

Sincerely,
Libby Munson
Research Ethics Support Specialist
Research Ethics, Compliance, and Partnerships
Walden University
100 Washington Avenue South, Suite 1210
Minneapolis, MN 55401
Email: irb@mail.waldenu.edu
Phone: (612) 312-1283
Fax: (612) 338-5092

Information about the Walden University Institutional Review Board, including instructions for application, may be found at this link: http://academicguides.waldenu.edu/researchcenter/orec

Appendix B: NCI Approval for SEER Database Access

Your SEER Data Access Request has been approved - SAR0040208

NCI at Your Service <nci@servicenowservices.com>
Sat 7/30/2022 6:05 PM
To: Elsadig Elsharif <elsadig.elsharif@waldenu.edu>

NIH NATIONAL CANCER INSTITUTE
Surveillance, Epidemiology, and End Results Program

SEER Incidence Data: Data Access Request Approved

This is an automated message from NCI's Surveillance, Epidemiology, and End Results Program (SEER) Data Access System. Do not reply back to this message.

Your request to upgrade your existing access to SEER Data has been approved for SEER Research Database:

Name:	Elsadig Ahmed Elsharif
Email address:	elsadig.elsharif@waldenu.edu

*****HOW TO ACCESS THE DATA*****

You will receive your SEER*Stat account information to access the data in separate emails. After you receive the account information, you can log into the SEER*Stat software to access the SEER data.

To download SEER*Stat, visit http://seer.cancer.gov/seerstat/download. If you have not installed SEER*Stat before, you will be required to complete the registration form and agree to the Terms of Use.

Appendix C: NIH Certificate of Completion of Training Course on Protecting Human Research Participants

Certificate of Completion

The National Institutes of Health (NIH) Office of Extramural Research certifies that **Elsadig Elsharif** successfully completed the NIH Web-based training course "Protecting Human Research Participants".

Date of completion: 01/01/2018.

Certification Number: 2584514.