

Lung Cancer in Never-Smoking Women

by

Vaishnavi Brahmamdam

BS, Biology, Indiana University, 2019

Submitted to the Graduate Faculty of the
Department of Human Genetics
Graduate School of Public Health in partial fulfillment
of the requirements for the degree of
Master of Public Health

University of Pittsburgh

2022

UNIVERSITY OF PITTSBURGH
GRADUATE SCHOOL OF PUBLIC HEALTH

This essay is submitted

by

Vaishnavi Brahmamdam

on

April 25, 2022

and approved by

Essay Advisor: Brenda Diergaarde, PhD, Associate Professor, Department of Human Genetics,
Graduate School of Public Health, University of Pittsburgh

Essay Reader: Laura Stabile, PhD, Research Associate Professor, Department of Pharmacology
and Chemical Biology, School of Medicine, University of Pittsburgh

Copyright © by Vaishnavi Brahmamdam

2022

Lung Cancer in Never-Smoking Women

Vaishnavi Brahmamdam, MPH

University of Pittsburgh, 2022

Abstract

Lung cancer is the leading cause of cancer death worldwide and the most common cause of cancer mortality in women in the United States. Approximately 80 percent of all lung cancer cases are due to personal cigarette smoking habits, the other 20 percent are attributed to other risk factors. The incidence of lung cancer in never-smokers (LCINS), particularly in women, is rising and of major public health concern. Potential risk factors impacting the development of LCINS and the sex disparity in the observed incidence are not fully understood and requires further investigation.

I conducted a literature review of existing studies assessing the impact of potential risk factors on lung cancer development in never-smoking women. Using the main themes and gaps identified in the literature review, I subsequently developed a questionnaire that aims to improve our understanding of the role environmental and clinical factors play in lung cancer development in never-smoking women. Main themes identified include moderate to strong associations between lung cancer risk in never-smoking women and secondhand smoke, radon, occupational exposures, medical history of chronic respiratory conditions, family history of lung cancer, and air pollution. The greatest gap identified by many of the evaluated studies is the role of reproductive and hormonal history in lung cancer development in this population.

Using the identified main themes and gap, existing questionnaire databases and epidemiological surveys, and articles about questionnaire design and development, I created an

84-question patient questionnaire. The questionnaire was categorized into 12 sections: demographics, cigarette smoking, other smoking habits, secondhand smoke, alcohol, BMI, medical history, family history, reproductive and hormonal history, radon, occupational history, and residential history.

The public health significance of this project is that the questionnaire findings may expand our knowledge of the risk factors involved in lung cancer development in never-smoking women and explain the observed sex disparity in lung cancer incidence in never-smokers. The results may also inform policy reform to address occupational and residential exposures and air pollution regulation. Ultimately, the questionnaire findings may allow public health professionals and practitioners to improve screening methods and treatment options for never-smoking women.

Table of Contents

Preface.....	x
1.0 Introduction.....	1
2.0 Background	3
2.1 Incidence of Lung Cancer	3
2.2 Risk Factors for Lung Cancer Development	5
2.2.1 Smoking.....	5
2.2.2 Secondhand Smoke	6
2.2.3 Residential and Occupational Exposures	8
2.2.4 History of Chronic Health Conditions	9
2.2.5 Family History	11
2.2.6 Diet.....	12
2.3 Lung Cancer and Sex Differences.....	13
2.3.1 Histology	13
2.3.2 Age at Diagnosis	14
2.3.3 Tumor Characteristics.....	15
2.3.4 Reproductive and Hormonal Factors.....	15
2.3.5 Survival	18
2.4 Goal and Specific Aims of this Project	18
3.0 Materials and Methods.....	20
3.1 Literature Review	20
3.2 Development of Patient Questionnaire	21

3.3 Assessment of Questionnaire Readability	24
4.0 Results	26
4.1 Literature Review Table	26
4.2 Lung Cancer Patient Questionnaire	43
4.3 Questionnaire Readability	58
5.0 Discussion.....	59
6.0 Conclusion	65
Bibliography	66

List of Tables

Table 1: Literature Review Summary Table.....	27
--	-----------

List of Figures

Figure 1: Incidence of Cancer Diagnoses in the United States by Sex from 1975 to 2018.	4
Figure 2: Proportion of Never-Smokers Among Non-Small Cell Lung Cancer Patients from 1990-2013	5
Figure 3: Microsoft Word Document Statistics for the Questionnaire.....	58

Preface

I would like to thank the following individuals for their endless support, guidance and insight on the development of this essay. To Dr. Brenda Diergaarde, thank you for sharing with me your desire to start this project and bringing me on board. You have pushed me to learn and grow in countless ways on the journey of epidemiological questionnaire development. Thank you for your kind words of encouragement, understanding and patience. To Dr. Laura Stabile, thank you for your support and expertise in the creation of this essay. To my program advisor, Dr. Andrea Durst, thank you for your continued mentorship in the MPH Human Genetics program.

Lastly, to my family and friends, thank you for your love, support and encouragement every step of the way on my academic journey. I am grateful for each and every one of you and this would not be possible without all of you. A special thank you to my parents, Jyotsna and Ranga, and sister, Sree, for supporting my professional and personal ambitions in life. I owe it all to you.

1.0 Introduction

Lung cancer is the leading cause of cancer death in both sexes in the United States and worldwide. Almost half of these deaths occur in women, making lung cancer the most common cause of cancer death in women with a higher annual mortality rate than breast, uterine, and ovarian cancer (R. L. Siegel, Miller, Fuchs, & Jemal, 2022).

The incidence of lung cancer in never-smokers is rising, and the majority of the cases occur in women (Dubin & Griffin, 2020). Although tobacco use is the driving factor for approximately 80 percent of all lung cancer cases, the rising incidence in never-smokers is a major public health concern (R. L. Siegel et al., 2015). Never-smokers with lung cancer are 2.5 times more likely to be women than men (Jemal et al., 2018). A recent study of a genomic analysis of lung cancer in never-smokers found three novel molecular lung cancer subtypes in this population and that mutations in these tumors were caused by natural processes in the body (T. Zhang et al., 2021). These findings suggest that lung cancer in never-smokers, specifically in women, might be molecularly and evolutionarily distinct from lung cancer in never-smoking men. Environmental and lifestyle factors, including reproductive and hormonal factors, may be responsible for the observed difference in lung cancer incidence (North & Christiani, 2013).

Current literature shows that sex differences do exist between males and females regarding lung cancer etiology. Women are more likely to be diagnosed with lung cancer at a younger age than men (Jemal et al., 2018). In addition, women are more likely to be diagnosed with adenocarcinoma histology than men and lung tumors from women harbor different somatic alterations than those from men, including an increased frequency of mutations in the epidermal growth factor receptor (EGFR) gene and of translocations of the anaplastic lymphoma kinase

(ALK) oncogene (Yang et al., 2005). The latter suggests that indeed there might be sex-related differences in etiology. One of the clearest differences between males and females is in levels of sex hormones, and differences in lung cancer carcinogenesis and outcomes may be due to different exposures to estrogen and progesterone (Chakraborty, Ganti, Marr, & Batra, 2010). A better understanding of the factors involved in risk and outcome of lung cancer among never smoking women could help improve prevention, early detection and treatment strategies.

Risk factors for lung cancer in women have not been extensively investigated and conflicting data exists. In this study, I reviewed existing studies, primarily epidemiological and human studies, rather than animal studies, and developed a questionnaire to be used to assess possible risk factors of lung cancer that may exist among this specific population. This questionnaire will be utilized in a study at the University of Pittsburgh Medical Center (UPMC) Hillman Cancer Center on lung cancer among never-smoking women. This study will also address the sex disparity that exists between never-smoking men and women and the rising incidence of lung cancer. Since this is a growing public health concern, improved knowledge of risk factors can allow for public health professionals and practitioners to develop better screening methods and treatment options.

2.0 Background

2.1 Incidence of Lung Cancer

Lung cancer is estimated to be the second most commonly diagnosed cancer in both females and males in the United States (American Cancer Society, 2021). In 2021, the estimated number of new cases of lung and bronchus cancer were 116,660 in females and 119,100 in males, for a total 235,760 new cases (R. L. Siegel, Miller, Fuchs, & Jemal, 2021). Lung cancer is also the leading cause of cancer-related mortality in both men and women, with an estimated 131,880 deaths due to lung and bronchus cancer in 2021 (R. L. Siegel et al., 2021). Figure 1 below depicts the incidence trend of lung cancer, and other common cancers, from 1975 to 2018 in both sexes. As seen in Figure 1, there is a steep decline in lung cancer incidence in males during this time period, while the incidence in females has increased since 1975 and remained fairly constant since 2005 (R. L. Siegel et al., 2022). Although smoking is the leading cause for lung cancer development, the incidence of LCINS is increasing (Pelosof et al., 2017).

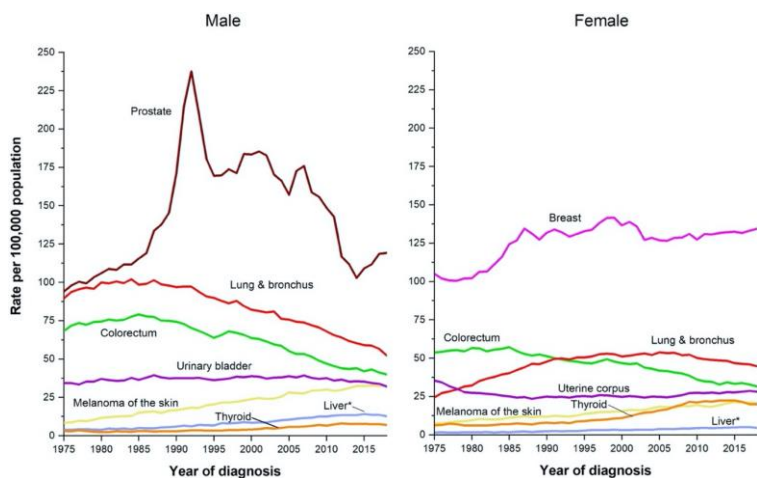
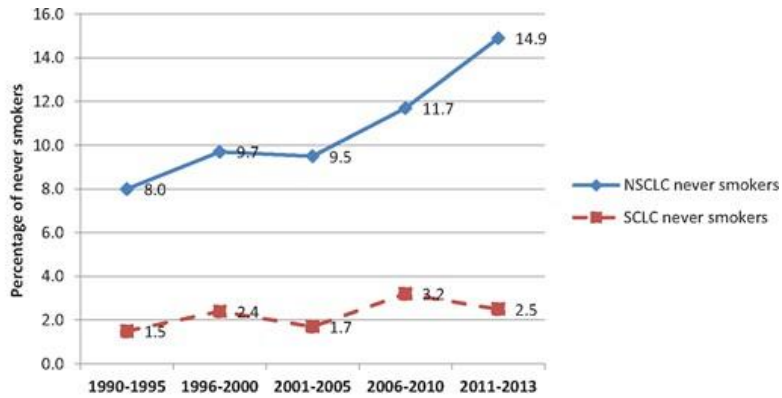


Figure 1: Incidence of Cancer Diagnoses in the United States by Sex from 1975 to 2018.

Figure retrieved from (R. L. Siegel et al., 2022), *Cancer Statistics, 2022*, American Cancer Society Journals, *A Cancer Journal for Clinicians*, Volume 71, Issue 1, page 7-33. Copyright © American Cancer Society.

Although the overall incidence rate of lung cancer in the United States is decreasing due to improved smoking behaviors and advancements in screening (See Figure 1), the incidence rate of lung cancer in never-smokers, who are predominantly female, is steadily increasing (Pelosof et al., 2017). Worldwide statistics show that 53 percent of females with lung cancer and 15 percent of males with lung cancer were never-smokers (Parkin, Bray, Ferlay, & Pisani, 2005). In the United States, the proportion of never-smokers with lung cancer in all non-small cell lung cancer (NSCLC) patients has increased from approximately 8.0 percent in the years 1990-1994 to 14.9 percent in 2011-2014, as depicted in Figure 2 below (Pelosof et al., 2017). Figure 2 also shows the proportion of never-smokers among small cell lung cancer (SCLC) patients. Based on the low incidence of never-smokers with SCLC, the remainder of this essay will focus on never-smokers with NSCLC. The most common NSCLC cell type in never-smokers is adenocarcinoma (D. A. Siegel, Fedewa, Henley, Pollack, & Jemal, 2021). Since never-smokers are 2.5 times more likely to be female than male, this suggests that factors aside from smoking may be contributing to this difference (Jemal et al., 2018).



No. of patients						
All NSCLC	677	1726	2391	3182	2617	
Never smoker NSCLC	54	168	228	372	391	P<0.001
All SCLC	136	297	356	404	317	
Never smoker SCLC	2	7	6	13	8	P=0.36

Figure 2: Proportion of Never-Smokers Among Non-Small Cell Lung Cancer Patients from 1990-2013

Figure retrieved from (Pelosof et al., 2017), Proportion of Never-Smoker Non-Small Cell Lung Cancer Patients at Three Diverse Institutions, Journal of the National Cancer Institute, Volume 109, Issue 7, by permission of Oxford University Press.

2.2 Risk Factors for Lung Cancer Development

2.2.1 Smoking

There are several potential risk factors for lung cancer development, but cigarette smoking is the number one risk factor worldwide. Cigarettes contain chemical carcinogens and with repeated use and increased smoking duration, can lead to an increased risk of developing lung cancer. Cigarette smoking and lung cancer development were originally connected as a causal relationship through several case-control and cohort studies. A landmark study conducted in England in 1947 showed that lung cancer risk was related to the number of cigarettes one smoked.

They found lung cancer risk was 25 times higher in those that smoked cigarettes compared to those who did not smoke cigarettes (Doll & Hill, 1950). A recent meta-analysis of 99 cohort studies published between 1999 and 2016 with 7 million participants and over 50,000 lung cancer cases examined the sex-specific association between cigarette smoking and lung cancer risk. This meta-analysis found that men and women who smoke cigarettes have similar increased risk of lung cancer development, with a relative risk (RR) of 6.99 (95% CI 5.09-9.59) in women and 7.33 (95% CI 4.90-10.96) in men (O'Keeffe et al., 2018). With approximately 82 percent of lung cancer cases directly attributed to smoking, there still remains about 18 percent of cases that may be attributed to other risk factors (Islami et al., 2018).

2.2.2 Secondhand Smoke

Secondhand smoke is both the smoke exhaled by cigarette smokers and side stream smoke that comes from the burning end of a cigarette. Secondhand smoke exposure often occurs at workplaces, public places like bars and restaurants, and at home. Exposure at your home and workplace may be the most crucial because that is where people spend the most amount of time (Office on & Health, 2006). Secondhand smoke contains more than 7,000 toxic chemicals, and a significant amount of them are carcinogenic. Secondhand smoke is classified as a Group A carcinogen by the Environmental Protection Agency, meaning there is enough evidence that suggests secondhand smoke causes cancer in humans (U.S. Environmental Protection Agency, 1992). In 2014, The Surgeon General Report revealed that nonsmokers who have been exposed to a significant amount of secondhand smoke at home or work have a 20 to 30 percent increased risk of developing lung cancer (U.S. Department of Health and Human Services, 2014).

The first epidemiological studies to assess the risk of lung cancer in nonsmokers were published in 1981. All three studies evaluated the relationship between lung cancer risk in women and secondhand smoke exposure at home from their husband. The results of these studies, along with many that have been done since, have consistently delivered the same finding: secondhand smoke has a causal relationship with increasing lung cancer risk, particularly in those that are married to a cigarette smoker (Garfinkel, 1981; Hirayama, 1981; Trichopoulos, Kalandidi, Sparros, & MacMahon, 1981). In addition, a meta-analysis of 40 epidemiologic studies which mostly assessed secondhand smoke exposure in adult life, identified a higher risk of LCINS with exposure to secondhand smoke compared to other types of cancers, with an odds ratio (OR) of 1.245 (95 % CI 1.026-1.511). One particularly strong association was identified in female never-smokers with secondhand smoke exposure compared to female never-smokers without secondhand smoke exposure, with an OR of 1.253 (95% CI 1.142-1.374) (Kim, Ko, Kwon, & Lee, 2018).

Secondhand smoke exposure during adult life and at work confer increased risk of lung cancer development in never-smokers. The Prostate, Lung, Colorectal and Ovarian study evaluated the association between secondhand smoke exposure in 49,569 never-smoking participants and its impact on lung cancer incidence and mortality. Compared to participants with no or little exposure at home during their adult life, participants with secondhand smoke exposure most of their adult life had a higher risk of both lung cancer diagnosis and mortality, with a hazard ratio (HR) of 1.809 (95% CI 1.161-2.819, p=0.009) and 1.923 (95% CI 1.035-3.575, p=0.038), respectively. Compared to participants with no secondhand smoke exposure at work, participants with secondhand smoke exposure for most of their work time also had a higher risk of lung cancer development, with a HR of 2.038 (95% CI 1.313-3.164, p=.002) (Abdel-Rahman, 2020). This

suggests that secondhand smoke plays a significant role in lung cancer development in never-smokers.

2.2.3 Residential and Occupational Exposures

There are several types of exposures one may have at home or work that can contribute to an increased risk of lung cancer development. Aside from secondhand smoke, other exposures include radon, asbestos, arsenic, silica, combustion of biomass fuels (such as wood or crops) and coal, and cooking fumes in the household (Cheng, Weber, Steinberg, & Yu, 2021). Although all of these can contribute to an increased risk of lung cancer, exposure to radon and asbestos are the most common. A case-control study of 445 lung cancer cases, 425 population controls, and 523 hospital controls assessed the relationship between lung cancer risk in never-smokers and a variety of risk factors, of which one was workplace exposures (Darren R. Brenner et al., 2010). The study found that a previous occupational exposure, such as exposure to asbestos, solvents, paints, welding equipment, pesticides, grain elevator dust, wood dust and smoke, or non-tobacco related exhaust, among never-smokers increases their lung cancer risk, with an OR of 2.1 (95% CI 1.3-3.3).

Exposure to radon, an odorless and radioactive gas, is not only common among miners, but also homeowners and renters as it can enter buildings through cracks or gaps and build up indoors. Radon is established as a known environmental cause of lung cancer due to its carcinogenic properties. When radon decays, two products, polonium-218 and polonium-214, release particles which can cause DNA base mutations and chromosomal strand breaks. These DNA mutations and chromosomal strand breaks may occur in the cells lining the respiratory airways and, if not timely repaired, can lead to lung cancer development. A meta-analysis of 22 case-control studies of

residential radon exposure and lung cancer risk identified a combined OR of lung cancer to be 1.29 (95% CI 1.10-1.51). In addition, it showed that with every 100 Bq/m³ increase in residential radon exposure, there was an associated 7 percent increase in lung cancer risk (Z. L. Zhang et al., 2012). A pooled case-control study of 523 never-smoking lung cancer cases and 892 controls which investigated the relationship between residential radon exposure and lung cancer risk revealed an OR of 1.73 (95% CI 1.27-2.35) for those exposed to radon concentrations of more than 200 Bq/m³ versus those exposed to 100 Bq/m³ (Lorenzo-González et al., 2019). This suggests a linear dose-response relationship between residential radon exposure and lung cancer risk among never-smokers.

Asbestos is a group of heat and corrosion-resistant, naturally occurring fibers that is commonly used for insulation purposes, in floor tiles, buildings, and pipes. Inhaling asbestos fibers can cause tissue scarring in the lungs, resulting in lung dysfunction, disability or death. A case-control study found a strong association between high (OR 3.66 and 95%CI 1.61-8.29) and medium (OR 1.25 and 95%CI, 0.47-3.31) asbestos exposure and lung cancer risk, suggesting that asbestos exposure and lung cancer risk have a linear relationship. Moreover, asbestos exposure combined with cigarette smoking further increases lung cancer risk (Yano, Wang, Wang, Qiu, & Wang, 2010).

2.2.4 History of Chronic Health Conditions

Increased lung cancer risk may also be associated with personal history of various chronic health conditions such as asthma, chronic obstructive pulmonary disease (COPD), emphysema, pneumonia and pulmonary tuberculosis. A prospective cohort study of 64,170 participants found that a history of asthma was associated with a 25 percent increase in lung cancer risk, with a HR

of 1.25 (95% CI 1.00-1.57). This positive association existed among all races and ethnicities (Kantor, Hsu, Du, & Signorello, 2019). A large cohort study with 338,548 participants found a 2.6 times higher incidence of lung cancer in never-smoking patients with COPD compared to never-smoking participants without COPD. This risk was comparable for ever-smokers as well, suggesting that patients with COPD are at a higher risk of developing lung cancer, regardless of their smoking status (Park et al., 2020).

In addition to asthma and COPD, emphysema has been linked to lung cancer. A prospective cohort study of 1,166 current and former smokers found that the incidence of lung cancer was three times more likely, with a RR of 3.33 (95% CI 1.41-7.85), in individuals with emphysema (de Torres et al., 2007). Pulmonary tuberculosis, a bacterial infection of the lungs, also confers greater risk of lung cancer. A population-based prospective cohort study of 4,480 participants with tuberculosis and 712,392 without tuberculosis found that incidence of lung cancer was 11-fold higher in participants with tuberculosis compared to those without (Yu et al., 2011). This study also found that the risk of lung cancer in those with tuberculosis was higher in men than women. A meta-analysis conducted using 17 studies including 24,607 lung cancer cases and 81,829 controls from the Lung Cancer Consortium examined the impact of lung diseases on lung cancer risk stratified by smoking status. Findings revealed that among never-smokers, those with previous diagnoses of pneumonia, tuberculosis, or emphysema had elevated lung cancer risk, with a RR of 1.45 (95% CI 1.12-1.63), 1.50 (95% CI 1.03-2.19), and 2.21 (95% CI 1.00-4.90), respectively (D. R. Brenner et al., 2012). This study did not observe a difference in lung cancer risk and previous lung disease by sex. These studies suggest that lung cancer risk is increased in individuals with various co-existing chronic health conditions.

2.2.5 Family History

Although most lung cancer cases are sporadic and due to environmental risk factors as discussed above, having a family history of lung cancer may be a predisposing factor to lung cancer development. A case-control study comprised of female, Singaporean Chinese participants (374 lung cancer cases and 785 controls) investigated the association between family history in first degree relatives and lung cancer risk among never and ever-smokers. Findings showed an increased lung cancer risk among those with family history of lung cancer versus those with no family history (OR = 2.08, 95% CI 1.25-3.247). Though no significant association was found between ever-smokers with family history of lung cancer and lung cancer, there was a significant association found between never-smokers with family history of lung cancer and lung cancer risk (OR = 2.78, 95% CI 1.57-4.90) (Yin, Chan, Seow, Yau, & Seow, 2021). A retrospective cohort study including 230 never-smokers with NSCLC examined the association between family history of lung cancer with NSCLC and tumor mutations. This study found that patients with somatic EGFR mutations were more likely to have a family history of lung cancer, but this association was not as strong in patients whose tumors harbored ALK translocations or KRAS mutations. This suggests that family members of lung cancer patients with known EGFR tumor mutations may be good candidates for earlier cancer screenings (Gaughan, Cryer, Yeap, Jackman, & Costa, 2013).

Genome-wide association studies (GWAS) have identified possible germline mutations and susceptibility loci that may explain lung cancer heritability in families. A GWAS conducted to characterize heritability of previously identified common lung cancer susceptibility regions and possible new loci by genotyping 14,803 lung cancer cases and 12,262 controls, found 18 statistically significant ($p \leq 5 \times 10^{-8}$) loci associated with lung cancer, of which eight were new loci (McKay et al., 2017). Another GWAS comprised of never-smoking lung cancer patients of

European descent found genetic susceptibility to lung cancer in never-smokers associated to three statistically significant single nucleotide polymorphisms (SNPs): rs31490, rs380286, and rs4975616, all in the Chromosome 5 CLPTM1L-TERT region (Hung et al., 2019). Additionally, a study conducted to predict lung cancer risk using common SNPs from lung cancer susceptibility regions identified in previous GWAS found that identifying SNPs in three specific susceptibility regions, 5p15.33, 6p21.33, and 15q25.1, resulted in a small improvement in lung cancer prediction (Weissfeld et al., 2015). It is important to note that familial history cannot be immediately attributed to heritability of genes, but rather families may have similar habits and lifestyle factors contributing to their lung cancer risk (Yin et al., 2021).

2.2.6 Diet

Consumption of fruit and vegetables rich in carotenoids and antioxidants has been hypothesized to decrease lung cancer risk and play a protective role in development (Vieira et al., 2016). A review of the 2007 Continuous Update Project, an international research initiative investigating the impact of diet, nutrition, and physical activity on cancer prevention, revealed that approximately 8.9 percent of lung cancer cases were attributed to low fruit and vegetable intake (Islami et al., 2018; World Cancer Research Fund International). Meta-analysis of 11 prospective studies reporting on the association of cruciferous vegetable intake and lung cancer risk identified a 19 percent decrease in lung cancer risk with intake of 100 grams/day. In the same meta-analysis, 14 prospective studies reporting on the relationship between lung cancer risk and fruit intake showed an 18 percent reduced risk of lung cancer by increasing fruit intake by 200-300 grams/day (Vieira et al., 2016). Furthermore, a multicenter case-control study conducted in Northern China including 1,086 never-smoking cases and 2,172 never-smoking controls examined the effects of

various risk factors on lung cancer in never-smokers, one of which was fruit intake. Findings revealed consumption of fruits 3-5 and 6-7 days per week conferred protection from lung cancer development, with an OR of 0.79 (95% CI 0.65-0.95) and .34 (95% CI 0.27- 0.42), respectively. Additionally, the odds of developing lung cancer was 1.59 times more likely in never-smokers who consumed high-degree alcohol (high-degree denoting liquor and low-degree denoting wine and beer) compared to those that did not (Liang et al., 2019).

2.3 Lung Cancer and Sex Differences

2.3.1 Histology

The histological distribution of lung cancer between males and females differs. There are two major histological forms of lung cancer: NSCLC and SCLC. Although adenocarcinoma, a subtype of NSCLC associated with smoking or former smoking, is the most common histological subtype of lung cancer overall, never-smoking women are more likely to develop adenocarcinoma than never-smoking men. Of the NSCLC subtypes, males are more likely to develop squamous cell carcinoma and females are more likely to develop adenocarcinoma. A population-based study of 2,875 women and 17,686 men found that SCLC and adenocarcinoma were more predominant in women than men. 26.6 percent of women had SCLC versus 19.9 percent in men and 21.6 percent of women had adenocarcinoma versus 9.6 percent in men (Radzikowska, Głaz, & Roszkowski, 2002). A study comprised of 583 lung cancer cases from 1967-1976 and 278 cases from 1991-1999 conducted in Malaysia found that while squamous cell carcinoma was the most frequent cell type from 1967-1976, adenocarcinoma predominated from 1991-1999 among female and male

smokers. However, and rather strikingly, adenocarcinoma predominated more in female never-smokers than male never-smokers from both 1967-1976 and 1991-1999 (Liam, Pang, Leow, Poosparajah, & Menon, 2006). Likewise, a more recent study involving cancer registry and medical record data abstraction of 129,309 lung cancer patients in seven states in the United States aimed to quantify the proportion of never-smokers with lung cancer by sex. This study revealed that squamous cell carcinomas occurred in 6 percent and 8 percent of never-smoking women and men, respectively, while adenocarcinomas occurred in 19.6 percent and 11.8 percent never-smoking women and men, respectively (D. A. Siegel et al., 2021). This may suggest that adenocarcinoma development is different in never-smokers versus smokers, and even differs by sex.

2.3.2 Age at Diagnosis

A systematic analysis of lung cancer incidence in 40 countries showed that women were more likely to be diagnosed with lung cancer at a younger age, 30-49 years, than men, and sex differences in smoking habits may not be the reason for this (Fidler-Benaoudia, Torre, Bray, Ferlay, & Jemal, 2020). A case series of 975 lung cancer patients found that never-smokers not only presented with an advanced stage of disease at diagnosis, but also were diagnosed five years earlier than their current smoker counterparts and ten years earlier than their former smoker counterparts (Toh et al., 2006). However, this conflicts with a retrospective cohort study which found that never-smoking women are actually diagnosed at an older age than ever-smokers. This may be due to either symptom presentation occurring later in this population or delayed screening because lung cancer is more common smokers (Dias, Linhas, Campainha, Conde, & Barroso, 2017). These studies suggest that because never-smoking women not only get diagnosed at a

younger age but also have more advanced stage of lung cancer at time of diagnosis, lung cancer may develop differently in never-smoking women.

2.3.3 Tumor Characteristics

The somatic mutation profile in lung tumors differs in never-smokers compared to ever-smokers and also by sex. A retrospective cohort study comprised of 558 lung cancer patients, including 22.4 percent never-smokers of which 74 percent were women, assessed the prevalence of EGFR mutations and ALK translocations by sex, histologic type, and smoking status. Considering that majority of the never-smokers in this study were women, never-smokers had more EGFR mutations than ever-smokers (36 % vs 8 %; $p < 0.001$). In addition, ALK translocations were also more prevalent in the never-smokers than ever-smokers (26% vs 4%; $p < 0.001$). EGFR exon 19 deletions were the most common in never-smokers, while exon 21 substitutions were more common in ever-smokers (Dias et al., 2017). Another study evaluated the prevalence of EGFR somatic mutations in 219 NSCLC patients. In this study population, 12 percent had an EGFR mutation and 5 percent had a novel, likely germ-line mutation. Of the 12 percent with EGFR mutations, 54 percent and 12 percent were in women and men, respectively (Yang et al., 2005). These studies suggest that tumor characteristics are distinct in never-smoking women versus never-smoking men with lung cancer.

2.3.4 Reproductive and Hormonal Factors

Since lung cancer in never-smokers predominates in women, reproductive and hormonal factors may play a contributing role. Stabile et al. compared estrogen receptor alpha ($ER\alpha$) and

beta ($ER\beta$) and progesterone receptor (PR) in lung tumors and assessed the impact of these levels on patient outcomes. Compared to active smokers with lung cancer, $ER\alpha$ expression was significantly higher in never-smokers and ex-smokers with lung cancer, and PR expression levels were also significantly higher in never-smokers compared to ever-smokers with lung cancer (Stabile et al., 2011). Similarly, an epidemiological cohort study of 140 women with adenocarcinoma – of which 63 were never-smokers and 77 were former or current smokers, investigated genetic and hormonal specificities in never-smoking women. Findings revealed a higher frequency of $ER\alpha$ and $ER\beta$ expression in lung tumors of never-smokers than former and current smokers, suggesting hormonal pathways may play a role in lung cancer development in female never-smokers (Mazières et al., 2013). Expression of aromatase, the enzyme that converts testosterone to estradiol, in never-smoking women with NSCLC has also been assessed. A study comprised of 35 never-smoking women with NSCLC found that women with lower levels of aromatase levels in lung tumors was associated with a greater 5-year survival advantage than women with higher tumor aromatase levels (Mah et al., 2007). Many studies have assessed the association between age at menarche, oral contraceptive use, hormone therapies, pregnancy and menopause and lung cancer risk, however conflicting findings exist.

A prospective cohort study comprised of 89,835 women aged 40-59 aimed to assess the impact of several different reproductive and hormonal factors on lung cancer risk. Over an average of 16.4 years of follow-up, there were 750 incident lung cancer cases and 89,062 non-cases. Findings showed that the average age of menarche was similar in both cases (12.7 years) and non-cases (12.8 years) and was not associated with lung cancer risk (Kabat, Miller, & Rohan, 2007). Use of oral contraceptives, duration of oral contraceptive use, and ever use of hormonal replacement therapy (HRT) were also not associated with lung cancer risk. However, findings

showed an elevated risk in lung cancer in women who used HRT for ten years or longer. Particularly in never-smoking women, there were some significant associations found between parity and age at first birth and lung cancer risk. Compared to ever-smokers, there was a stronger positive association between having five or more pregnancies in never-smokers and lung cancer risk (HR 2.01, 95% CI .99-4.11; $p < .005$). Compared to ever-smokers, there was an inverse association between age at first birth in never-smokers and lung cancer risk. Compared to ever-smokers, never-smoking women who had their first live birth between ages 23-25 had a HR of 0.73 (0.44-1.23) while women who had their first live birth at age 30 or older had a HR of 0.30 (0.10-0.84), suggesting that older age at first live birth may play a protective role in lung cancer development (Kabat et al., 2007).

A prospective case-cohort study was completed using data from 185,017 women aged 50-71 years involved in the NIH-AARP Diet and Health Study to investigate reproductive and hormonal factors and lung cancer risk. Of the 3,512 that developed lung cancer, 276 were never-smokers. Results showed a significant association between lung cancer in never-smokers and age at menarche (RR = 0.55, 95% CI 0.30-1.00 for age 15 vs <11), suggesting that lung cancer risk is twice as likely in individuals that had their first menstrual period before age 11. Somewhat conflicting with the Kabat et al., study, findings from this study did not find any significant association between lung cancer risk in never-smokers and age of first live birth, oral contraceptive use, and age at natural menopause (Brinton et al., 2011). Considering the conflicting and contradictory findings from the previously mentioned studies, more hormonal and reproductive data is needed from never-smoking women to understand the role of reproductive and hormonal history in their lung cancer development.

2.3.5 Survival

Not only do survival differences exist between ever and never-smokers with lung cancer, but studies also suggest that survival differences exist by sex as well, and this may be due to differences in age at diagnosis, histology type, tumor differences and reproductive and hormonal factors. Several studies found significantly longer survival times in never-smokers compared to ever-smokers (Cardona et al., 2019; Cronemberger et al., 2020; Toh et al., 2006; Viñolas et al., 2017). Although survival and quality of life is not the focus of this project, it should be noted that these differences exist, and previously mentioned differences such as age of diagnosis and advanced lung cancer stage at diagnosis may negatively impact survival and prognosis in women more than men.

2.4 Goal and Specific Aims of this Project

The overall goal of this project is to improve our understanding of risk factors for lung cancer development in never-smoking women. We hypothesize that lung cancer in never-smoking women represents a distinct cancer entity. To address this hypothesis:

- 1) I reviewed the existing literature on never-smoking women with lung cancer to understand the current knowledge of this disease in this unique population and identify relevant themes and gaps in literature.
- 2) Utilized the information retrieved from the literature review and subsequently developed a questionnaire to collect detailed information on a diverse set of medical, environmental and lifestyle variables from female never-smoking lung cancer patients.

This questionnaire will be used in a study at the UPMC Hillman Cancer Center about lung cancer among never-smoking women.

3.0 Materials and Methods

Following a literature review of factors associated with lung cancer development in never-smoking women, a patient questionnaire was developed using themes found in the literature review. After questionnaire development, its literacy and readability level were assessed.

3.1 Literature Review

The first step of conducting the literature review search was to specifically define and state our research goal. As stated previously, our goal is to improve our understanding of risk factors for lung cancer development in never-smoking women since lung cancer incidence is rising in this population but is not in men. I performed the literature search using PubMed during January and February of 2022. The complete search period for the literature review, which was used for the background section of this essay, comprised of 1981-2021, but when constructing the literature review table for the purpose of questionnaire development I focused only on studies published in the years of 2006-2021. This latter time period was used because more recently conducted studies and reviews are expected to have the most clinical relevance. Therefore, the inclusion criteria for the literature review table used to define relevant articles was any study or review that 1) examined the impact of potential risk factors on lung cancer development, diagnosis, and survival in never-smokers, specifically never-smoking-women 2) was complemented by free, full text 3) published in the English language, and 4) published since January 2006. Search keywords included combinations and synonyms of the following terms: lung cancer, never-smoker, women,

secondhand smoke, radon, and risk factors. Another method of finding relevant articles was searching in the reference list of identified articles. Evaluation of articles involved reading through the Abstract section of each study to determine if it fulfilled the four inclusion criteria and if the language of the study directly related to my research question, I downloaded the article to EndNote. Many types of study designs were evaluated and included in the table such as cohort studies, case series, medical record abstractions, and case-control studies. Studies conducted in countries worldwide were included in the literature review. All studies and reviews evaluated as pertinent to our research goal were included in a literature review table and categorized by region of study: North America, South America, Europe and Asia. In the table, I indicated author (year), study design, study population, main findings and themes, and limitations and gaps that the authors themselves identified for each study.

3.2 Development of Patient Questionnaire

We referred to three questionnaire development and design research articles to guide us through formatting, layout, logic and common challenges faced when developing this questionnaire. Primary challenges with questionnaire development include: conciseness, recall bias, open or closed-ended questions, wording, question sensitivity, defining unfamiliar terms or measures, and leading questions. (Egholm et al., 2020; Nieuwenhuijsen, 2005; Stehr-Green Paul, 2014). We determined that this questionnaire will best function as a take-home questionnaire for participants, so they will have more time to complete it as opposed to the short amount of time given to fill out a form in a waiting room. It is still important for the questionnaire to not be too lengthy for participants to complete, so any questions, words, or tables lacking function were

removed. Although, we ask about reproductive and hormonal use history in length and detail because existing research shows that this may be a strong factor influencing lung cancer in this population and literature shows conflicting findings. Another way for participants to save time filling out the questionnaire is by including easy-to-follow skip patterns, so they do not need to answer questions that don't pertain to them, and these were included in our questionnaire (Stehr-Green Paul, 2014). Recall bias is more likely to occur with people with the disease because they are more likely to report that they have a certain exposure than a person without the disease (Nieuwenhuijsen, 2005). In this study, all participants will have lung cancer, so minimizing the occurrence of recall bias was a priority in this process. For example, we experienced this issue when deciding how to ask participants about their known occupational exposures to agents such as silicon dust, diesel exhaust, radiation, and coal dust. This was resolved by creating a pros and cons list of the two ways we could ask about these exposures, followed by a discussion.

In addition, we considered the wording of certain questions, particularly for the reproductive and hormonal use history section as these questions may be taboo or sensitive in their culture, trigger personal trauma, and cause reluctance in answering (Beall & Leslie, 2014). To ensure we are receiving accurate information and provide clarity to participants, we defined the following terms: secondhand smoke, one standard drink, first-degree relatives, radon, and elevated radon level. In the case participants do not recognize a therapy by hormone name but rather the brand name, brand names of current and common hormone therapies were searched and provided as examples. This ensures that participants will fill out the questionnaire correctly. Including a comment box at the end gives participants a space to write about anything else they may think of or feel is relevant, which may provide context to a participant's experience and a more holistic understanding of risk factors involved in their lung cancer development.

Developing the patient questionnaire was a highly iterative process. It required continuous refining, rewriting, and rethinking of the structure, function, and sensitivity of the questions we want to ask our target population. The first step was to discuss the aim of the study and how we want to translate that into a questionnaire. The aim was to gather all relevant information of a patient that may be involved in the development of lung cancer in never-smoking women, but this goal must be balanced with time to complete the questionnaire and addressment of the previously mentioned challenges. After conducting the literature review on LCINS, we determined this relevant information to fall into the following twelve categories: demographics, cigarette smoking, other smoking habits, secondhand smoke, alcohol, BMI, medical history, family history, reproductive and hormonal history, radon, occupational history, and residential history. Initially, we did not have all of these categories, but over time and with review of literature we determined these to be the most pertinent.

To assist in question formulation, we used existing patient questionnaires and databases. These include the NIH Common Data Elements (CDE) Repository, National Adult Tobacco Survey Questionnaire (2013-2014), and The Behavioral Risk Factor Surveillance System (Center for Disease Control and Prevention, 2020; National Library of Medicine, 2015; Office of Smoking and Health and National Center for Chronic Disease Prevention and Health Promotion, 2015). Some questions were modeled off questions included in the above questionnaires and databases.

Throughout development, it was important to acknowledge that this would be a non-linear process that would require us to question the reason for asking each question, if it was worth asking, and if we were missing anything. Writing down the justification for asking each question was important so that we could identify how it would be useful in answering our research

questions. These reasonings were documented and thoroughly discussed during weekly meetings to determine inclusion or removal of a question.

3.3 Assessment of Questionnaire Readability

Another important aspect of creating the questionnaire was to consider the health and functional literacy level of our target population. Assessing the literacy and readability of the questionnaire and adjusting accordingly will ensure that participants understand the questions and fill out the questionnaire correctly. The National Assessment of Adult Literacy (NAAL) of 2003 which included more than 19,000 individuals showed that the literacy level of adults age 16 years and older in the United States is lower than international peers. Fifty-four percent of adults had intermediate health literacy, 23 percent had basic health literacy, and 14 percent had below basic health literacy (Cutilli & Bennett, 2009). To assess the questionnaire's readability, I checked the Flesh-Kincaid Grade Level and Flesch Reading Ease through Microsoft Word's Document Statistics. Flesch-Kincaid tests are readability tests used to assess the difficulty of reading an English passage by judging word length and sentence length. The Flesch-Kincaid Grade Level test provides a score in the form of a U.S. grade level. The score is also interpreted as the number of years of education a person may need to understand the material. The Flesch Reading Ease test provides a score out of 100, with higher scores corresponding to lower grade levels. A test score can fall into one of eight ranges: 100-90, 90-80, 80-70, 70-60, 60-50, 50-30, 30-10, and 10-0. These ranges correspond to grade levels from fifth grade to professional degree. Each range also corresponds to a brief description of the material's reading difficulty (Readable, 2021). We did not use other existing readability formulas or health literacy tools, such as Simple Measurement of

Gobbledygook (SMOG) or The CDC Clear Communication Index, because these are better suited to assess public-facing, public health-related materials and are not functionally practical to use considering the format of a questionnaire (e.g., use of tables, bullets, headings and titles) (Baur & Prue, 2014; Wang, Miller, Schmitt, & Wen, 2013)

4.0 Results

4.1 Literature Review Table

In Table 1, the results from the literature review for the patient questionnaire are shown. The literature review comprises the period September 2006 to December 2021. I selected 37 studies from countries in North America, South America, Europe and Asia as relevant to our understanding of potential risk factors for lung cancer development in never-smoking women. Listed are the authors (year of publication), study design, study population, main findings and themes, and limitations and gaps identified by the authors for each study.

Table 1: Literature Review Summary Table

Authors (Year)	Study Design	Study Population	Main Findings and Themes	Limitations and Gaps
North America				
(D. A. Siegel et al., 2021)	Cancer registry and medical record abstraction	129,309 lung cancer patients USA – 7 States	12.9% of the study population were never-smokers. The proportion of never-smoking lung cancer patients was higher in women than men (15.7% vs 9.6%) across all ages, race and ethnicities, and histology types. The proportion of patients with adenocarcinoma specifically was higher in never-smokers.	Generalizability may be limited because the study population was not from a national sample. Authors also suggest further research on never-smokers will help fill gaps on lung cancer risk factors such as occupational exposures, genetic factors, radon, and air pollution.
(Pelosof et al., 2017)	Retrospective study using cancer registries	10,593 NSCLC and 1,510 SCLC patients	The proportion of never-smoker NSCLC patients increased from 8.0% in the years 1990 to 1995 to 14.9% in 2011 to 2013 ($p < .001$). There was no statistically significant increase in the proportion of never-smoker SCLC patients during the same time periods. Lung cancer incidence in the self-reported never-smoking population is increasing. Never-smokers with NSCLC were more likely to be female than male (17.5% vs. 6.9%). The number of never-smoking women with NSCLC increased from 10.2% to 22.1% ($p < 0.001$) from 1990 to 2013. In men the increase was 6.6% to 8.9% ($p < 0.006$).	A limitation of this study is that incidence of never-smokers with lung cancer cases was not measured, rather proportion of never-smokers with lung cancer was over time. Since the study suggests lung cancer is rising among the never-smoking population, it is necessary to identify environmental carcinogens to aid in prevention and treatment.

(Jemal et al., 2018)	Cancer registry and national population-based surveys abstraction	392,108 lung cancer cases from 1995-2014	Study found higher incidence of lung cancer in young women than young men (30-54 years of age), with more of a burden on white and Hispanic women. These differences were not fully explained by sex differences in smoking habits when comparing smoking prevalence between female and male counterparts in this study.	There is a need to research sex-specific risks of lung cancer to understand why there is a higher incidence in young women.
(Abdel-Rahman, 2020)	Prostate, Lung, Colorectal, and Ovary (PLCO) Trial results and supplementary questionnaire abstraction	49,569 never-smoking participants	<p>Compared to those with no secondhand smoke (SHS) exposure at work, participants with exposure at work are at higher risk of lung cancer diagnosis.</p> <p>Compared to those with no or some SHS exposure at home, participants with exposure for most of their adult life are at higher risk of lung cancer diagnosis and mortality and at higher risk for other co-morbidities.</p> <p>Participants with and without SHS exposure during childhood had the same risk of lung cancer diagnosis and mortality.</p>	PLCO trial was designed to assess impact of screening on cancer risk and death, not the impact of environmental tobacco smoke. Though there was a large study population, the number of events (cases and deaths) were small. SHS data was self-reported and may incur bias. There is a need to assess clinical and biological characteristics of never-smokers with lung cancer and strong SHS exposure. This study did not do a sex comparison.

(Myers et al., 2021)	Case series comparison of ever and never-smokers with lung cancer and quantification of their PM _{2.5} exposure	1,005 lung cancer patients	<p>Never-smoking lung cancer patients were more likely to be female, younger, of Asian background, have higher PM_{2.5} exposure, and less likely to have COPD or family history of lung cancer, compared to ever-smokers.</p> <p>Cumulative PM_{2.5} exposure should be included in lung cancer risk assessment.</p>	This study was not designed so they could calculate the true incidence and risk of lung cancer development because it doesn't include non-lung cancer case from the beginning. Along with traditional risk factors, air pollution should be evaluated in future studies and clinical practice as a strong determinant of lung cancer occurrence.
(Al-Zoughool et al., 2013)	Population-based case-control study	44 lung cancer cases, 436 controls – all never-smokers	<p>No association found between environmental tobacco smoke (ETS) exposure from multiple sources inside and outside the home, and lung cancer.</p> <p>On average, controls had more years of education compared to cases. Over 50% of the cases had adenocarcinoma.</p>	The main limitation is that the number of cases was low, impacting the precision of the ORs. Participants may also have not recalled all exposures outside the home.
(Brinton et al., 2011)	Prospective case-cohort study	185,017 women age 50-71 years	Of the 3,512 that developed lung cancer, 276 were never-smokers. Results showed a significant association between lung cancer in never-smokers and age at menarche (RR = 0.55, 95% CI 0.30-1.00 for age 15 vs <11). They did not find any significant association between lung cancer in never-smokers and age of first live birth, oral contraceptive use, age at natural menopause.	Due to conflicting results in existing studies, there should be further investigation into hormonal and reproductive factors contributing to lung cancer in never-smoking women, such as age at menarche, age at menopause,

				hormone use, and parity.
(Baik, Strauss, Speizer, & Feskanich, 2010)	Prospective case-cohort study – questionnaire	107,171 post-menopausal women – 45% never smoking	There is a decreased risk of lung cancer in never-smoking parous women with increased parity (greater number of children). 64% of never-smokers in the study had adenocarcinoma histologic type. Current smokers were also younger at age of menopause than never-smoker counterparts.	Lack of significant findings among never-smokers may be due to small number of cases in this group. More research is necessary as the role of hormones and reproductive factors is still not fully understood.
(Kabat et al., 2007)	Prospective cohort study	89,835 women age 40-59	There was a stronger positive association of parity and lung cancer risk in never-smokers compared to smokers. There was an inverse association between age at first birth and lung cancer risk in never-smokers. There were no other significant associations found between lung cancer risk in never-smokers and age at menarche, oral contraceptive use, hormone replacement therapy use, and duration of each use.	Limitations include lack of information of age of menopause onset for menopausal women in the study and length of menstrual periods and cycle. Since smoking is the most significant contributor to lung cancer, results of the study should be interpreted carefully. Future studies should ask extensively about reproductive factors.
(Clague et al., 2011)	Prospective cohort study	60,592 post-menopausal females, of which 727 with lung cancer diagnosis	No association was found between hormonal therapy use and lung cancer risk in post-menopausal women with 11 years of follow-up.	Limitations of the study include possible confounding due to smoking.

(Darren R. Brenner et al., 2010)	Case-control study	445 cases (35% never-smokers) age 20-84, 425 population controls, and 523 hospital controls	Results show increased lung cancer risk in never-smokers with these risk factors: first-degree relatives with cancer diagnosis before age 50, personal medical history of emphysema among, and occupational exposures (asbestos, solvents, paints, welding equipment, pesticides, grain elevator dust, wood dust and smoke, and non-tobacco related exhaust).	imitations of this study include strong dependence on self-reported exposures and medical history. Future studies should focus on these environmental factors/exposures and genetic mutations in this population. Lung cancer risk associated with family history, occupational exposures, and personal medical history was not compared between sexes.
(Gowda et al., 2019)	Prospective cohort study	65,419 post-menopausal women, 265 lung cancer cases	There were no significant associations found between PM _{2.5} and NO ₂ exposures and lung cancer risk among never-smoking women. There was an increased lung risk associated with individuals residing less than 50 meters from a primary highway compared to those residing >200 meters from a primary highway.	Limitations include lack of information on PM _{2.5} and NO ₂ exposure levels prior to participant baseline and secondhand smoke exposure. Results of this study do not eliminate the associations found in other studies and future studies should focus in detail on traffic-related air pollution near residences.
(Turner et al., 2011)	Prospective cohort study	188,699 never-smoking participants	Over 26-year follow-up period, 1,100 participants developed lung cancer. There was a strong association between lung cancer mortality and PM _{2.5}	The PM _{2.5} level was not assigned at an individual or household level, so this may be a limitation.

			air pollution in both sexes, but was stronger in individuals with normal BMI and personal medical history of chronic lung disease ($p < 0.05$).	Authors mention that air pollution over an extended period of time are more relevant to determining lung cancer risk.
(D. R. Brenner et al., 2012)	Meta-analysis – 17 studies	24,607 lung cancer cases, 81,829 controls	Results showed increased lung cancer risk for never-smokers with history of emphysema, pneumonia, and tuberculosis (RR= 2.21, RR= 1.45, RR=1.50). No risk difference found between sexes.	Limitations include confounding due to occupational or SHS exposures. These diseases would be useful in determining who is at higher risk of lung cancer and may impact the way screening evaluations are done.
(Smith et al., 2012)	Prospective cohort study	271, 238 men and 177,494 women, age 50-71 years	Over the follow-up study period, 6,093 men and 3,344 women were diagnosed with lung cancer, with 166 and 249 in never-smokers, respectively. There was no association found between BMI and lung cancer risk in never-smokers.	Limitations include lack of follow-up about smoking status and other health conditions. At baseline, age of smoking initiation, cigarette content, and number of pack years, and BMI over time is recommended to be asked in future studies.
(Grundy et al., 2017)	Cancer registry analysis	390 never-smoking lung cancer	Population attributable risk of lung cancer was highest in never-smokers, 24.8%. Home radon remediation is crucial for lung cancer prevention.	Main limitation of the study is that radon level data was only collected from homeowners not renters because landlords are not required to remediate high radon levels in Alberta.

				Residential radon is a main non-smoking factor for lung cancer incidence.
(Mah et al., 2007)	Cohort study	422 patients with NSCLC, of which 35 were never-smoking women	Never-smoking women with NSCLC with lower aromatase expression levels had 5-year survival rate of 92% compared to 49% for women with higher levels.	Main limitation of this finding is that the population is small and future studies should confirm this result with a larger population size. Aromatase expression levels is a better survival predictor in women at older ages even though estrogen levels diminish as women age - suggesting growth factor receptors may be involved in this observation.
(Stabile et al., 2011)	Prospective cohort study	183 lung cancer patients, of which 13 were never-smokers	Compared to active smokers, ER α expression was significantly higher in never-smokers and ex-smokers with lung cancer. PR expression levels were also significantly higher in never-smokers compared to ever-smokers.	Future studies should use larger cohort size to confirm predictive findings. Examining these markers in both men and women may help determine who may respond to therapies. It is important to understand and determine the role of sex hormone-related proteins in lung cancer development.

South America				
(Cardona et al., 2019)	Retrospective cohort study	20 patients – 10 smokers, 10 nonsmokers	Median overall survival was significantly longer in never-smokers versus smokers (29.1 months vs 17.3 months). Results suggest never/ever-smokers with SCLC have a better prognosis than smokers with SCLC. Results also show that never/ever-smokers with SCLC have more EGFR, SMAD4, and MET mutations than smoker counterparts.	Limitations include a very small sample size and the retrospective study design. Future studies should include patients from diverse ethnicities and centers. Study also did not look at tumor mutation differences by sex.
(Cronemberger et al., 2020)	Observational retrospective cohort study	370 patients with locally advanced or metastatic non-small cell lung cancer, 54% of patients received genetic testing	Never-smokers had a significantly longer median overall survival time than smoker counterparts (14.6 vs 9.1 months). Never smoker patients were also at a higher risk of receiving a positive result from molecular testing compared to smoker counterparts (51.5% vs 9.1%).	Since a small percentage of participants got tested, were not able to make strong estimates of prevalence of specific mutation for subgroups of populations (like ethnicity). Testing methods in Brazil have also changed since this study was done.
Europe				
(Viñolas et al., 2017)	Prospective, multi-center study	2,035 women	Findings revealed that never-smoking women compared to current or former smoker counterparts were older, had lower level of education, used oral contraceptives more, and over one-third were exposed to SHS of which 82% experienced exposure at home. A higher proportion of EGFR mutations were found in never-smokers compared to current or former smokers	Findings about never-smoking women, like oral contraceptive use and age at diagnosis, conflict with other studies.

			(46.9% vs 19.0%). Median over survival was significantly longer in never-smoking women with EGFR mutation and stage 4 NSCLC compared to those with wildtype.	
(Dias et al., 2017)	Retrospective cohort study	558 patients – 22.4% never-smokers	Never-smokers were more likely to be females, older, and have adenocarcinoma histology. There was a greater prevalence of ALK translocations and EGFR mutations in tumors of never-smokers compared to ever-smokers. Deletions in exon 19 in EGFR was the most common in never-smokers and exon 21 was more frequent in ever-smokers.	As limitations, the number of never-smokers was low and other potential risk factors were not assessed. The higher incidence of lung cancer in female than male never-smokers may be due to other factors such as hormones, radon or cooking oil fumes, all things that this study did not investigate and future studies should.
(Pirie, Peto, Green, Reeves, & Beral, 2016)	Cohort study	1.2 million women – 51% reported as never-smokers	Never-smokers were more likely to have menopause at a later age, had higher fruit and vegetable consumption, less hormone use for menopause, and less likely to have SHS exposure at home as a child or adult compared to ever-smokers. The RR of lung cancer and adenocarcinoma in never-smokers showed a statistically significant association, risk was greater in those with asthma requiring treatment and increased height (165 cm or taller). In never-smoking post-menopausal women currently using hormone therapy, results did not show an increased risk of lung cancer. Self-	One limitation was that the study only assessed SHS exposure from a parent or partner at home, not any other source inside or outside the home. Exposure duration was also not assessed.

			reported non-white ethnicity compared to white was associated with increased lung cancer risk.	
(Fritz & Olsson, 2018)	Epidemiological study and questionnaire	1,159 women	Women diagnosed with lung cancer at a young age (age <40) were majority never-smokers. Never-smokers were also less likely to eat red meat. SHS exposure was found in all women except for two. Among the 70 women identified as diagnosed at a young age, only one woman reported family history as their only risk factor – so environmental exposures play a strong role. Findings suggest about 75% of the never-smokers were exposed to SHS by a parent, partner, or family member regularly at home.	Not all participants may be knowledgeable about all of their environmental exposures. Many women also did not report or knew their family histories.
(Mazières et al., 2013)	Epidemiological cohort study	140 women with adenocarcinoma – of which 63 never-smokers and 77 former or current smokers	The average age of disease presence was higher in never-smokers compared to smoker counterparts (68 versus 58.7 years). Approximately half of the never-smokers had an EGFR mutation (exon 19 and 21) while only 10.7% in former or current smokers. There was a higher frequency of estrogen receptors (alpha and beta) in female never-smokers than smokers – suggesting hormonal pathways are significant in lung cancer in female never-smokers.	Limitations were not discussed in the study article.
(Lorenzo-González et al., 2019)	Pooled case-control study	1415 participants, 523 cases and 892 controls – all never-smokers	OR of lung cancer risk was 1.73 (95% CI 1.27-2.35) for those exposed to residential radon concentrations of more than 200 Bq/m ³ versus those exposed to 100	Limitations include lack of adjustment of results for occupation and history of COPD and no

			Bq/m ³ . Results also show a significant positive association between residential radon exposure and adenocarcinoma. The OR for people exposed to 200 Bq/m ³ or higher versus those exposed to less than 100 Bq/m ³ was 1.44 for females and 2.43 for men.	information of radon exposure at work and only at home. Residential radon exposure should be considered when determining lung cancer risk for never-smokers.
Asia				
(Huang, Qu, & Du, 2019)	Cohort study and retrospective review	8,688 lung cancer patients – 31.85% women and 68.15% men, 92.25% of women were and 14.81% men were never-smokers.	The proportion of never-smokers with lung cancer increased over time (1990-2017), with the most significant increase in those with adenocarcinoma.	The study suggests more attention must be paid to the changes in proportions of lung cancer, specifically focusing on incidence in never and ever-smokers. The authors question whether the increased incidence is due to an actual increased incidence or because of smoking reduction in the overall population.
(Toh et al., 2006)	Hospital-based case series reviews	975 NSCLC patients diagnosed between 1999-2002 – 32.4% were never-smokers	Never-smokers presented with more advanced disease than current or former smoker counterparts. The proportion of females with adenocarcinoma was highest among never-smokers (p =.009). The 4-year survival rates for female smokers and never-smokers was 12.4% and 17.6%, respectively. Never-smokers seemed to	One limitation they identified was that they were not able to identify environmental tobacco smoke (ETS) exposure accurately. Something that could have strengthened their findings was knowledge

			be diagnosed at an earlier age than their current and former smoker counterparts (by 5-10 years)	when patients started smoking, pack-years smoked, and family history of cancer.
(Kurahashi et al., 2008)	Prospective cohort study	28,414 life-long non-smoking women	109/28,414 women were diagnosed with lung cancer, of which 83.7% had adenocarcinoma. The proportion of women in menopause was higher in women with never-smoking husbands when compared to women with husbands who currently smoke. There is a dose-response identified – a husband’s number of cigarettes smoked per day and pack-years was significantly associated with lung adenocarcinoma risk in never-smoking women. There is a 30% excess risk in never-smoking women exposed to passive smoking by their husband. SHS exposure at work also increased cancer risk in never-smoking women.	Limitations include only information being collected at baseline of the study, misclassification of smoking status and true relationship to the woman (husband or relative) could have occurred. The study did not have information on how long husband and wives lived together or spent time together in the same room.
(Yin et al., 2021)	Case-control study	374 lung cancer cases, 785 controls – all female, Singaporean Chinese participants	Family history of lung cancer was associated with higher risk of lung cancer among never-smokers. Further, lung cancer risk among never-smokers who also had low fruit consumption was higher if they had family history of lung cancer, specifically a first-degree relative. This study also found that adenocarcinoma type was most common among never-smokers.	Significant association between family history and lung cancer in never-smoking women cannot be immediately attributed to heritability of genes, but rather families may have similar habits and lifestyle factors contributing to lung cancer risk. Since family history was self-

				reported, there is a limitation of recall bias. The number of family members diagnosed with lung cancer was not collected, which could have strengthened the findings.
(Liang et al., 2019)	Case-control study	1,086 cases and 2,172 controls – 63.54% females and 36.46% males	Findings suggest that family history of lung cancer may increase lung cancer risk by 1.92 fold in never-smokers and OR was the same for both sexes. Never-smokers with a higher education level (college or more) were less likely to develop lung cancer compared to participants with a lower education level. Farmers or workers that are never-smokers were more likely to develop lung cancer, compared to never-smokers with other jobs. Results also revealed a 2.23 increased risk of lung cancer for never-smokers with SHS exposure at work and 2.33 fold increased risk for never-smokers with SHS exposure at home. Fruit intake of 3-5 times a week and 6-7 times a week was deemed a protective factor of lung cancer in never-smokers. Low level of alcohol consumption in never-smokers was not associated with lung cancer.	Smoking was self-reported – so never-smokers may be have been misclassified in the study.
(Cheng et al., 2021)	Review		This article reviews significant risk factors of lung cancer in never-smokers: radon, secondhand smoke, air pollution, occupation-	Exposure assessments must be developed to ask about all potential risk factors from a

			specific exposures, family history, and genetic susceptibility.	large sample population to better understand non-smoking related contributions to lung cancer in never-smokers.
(Weiss et al., 2007)	Prospective cohort study	71,392 women, never-smokers and no prior history of cancer at enrollment	There were 180 incident lung cancer cases. Risk of lung cancer female lifetime nonsmokers is increased for those with later age of menarche, shorter length of reproductive years, and irregular periods Decreased risk of lung cancer in female never-smokers was associated with those with more offspring and later age of menopause.	Since the number of cases was modest and few women had exposures, there was not enough power to make risk associations with exogenous hormones.
(Ha et al., 2015)	Lung tumor resection – genomic analysis	198 lung cancer patients – all female never-smokers	Study results showed female Asian never-smokers with lung cancer most often had EGFR mutation in lung tissue. 79% of female never-smokers had a well-known driver gene mutations.	Limitations and gaps were not identified in the study.
(Park et al., 2020)	Prospective cohort study	338,548 participants, age 40-84 years and no prior history of lung cancer	1,834 participants developed lung cancer. Results show a 2.6 times higher incidence of lung cancer in never-smoking patients with COPD compared to never-smoking participants without COPD. This risk was similar for ever-smokers, suggesting that patients with COPD are at a higher risk of developing lung cancer, regardless of smoking status.	Limitations include lack of information on occupational exposures and emphysema history which may confound risk due to COPD.
(Song, Sung, & Ha, 2008)	Prospective cohort study	170,481 post-menopausal women age 40-64 years	Results show an inverse association between BMI and lung cancer incidence, but the association did not hold true for never-smoking women.	Weight measurements are a better way of determining adiposity, but that was not

				measured. There are a lot of conflicting results from studies about BMI and lung cancer risk.
(Kim et al., 2018)	Meta-analysis	40 studies, of which 12 about lung cancer risk and never-smokers	Findings show that secondhand smoke increases lung cancer risk significantly in never-smokers, and especially in women (OR =1.235).	Author identifies selection and recall bias as possible limitations because only observational studies were included. Because only never-smokers were included in this meta-analysis, findings may be overestimated.

Main themes identified from the literature review:

- Lung cancer in never-smoking women is a distinct disease entity
- Several studies found moderate to strong associations between lung cancer risk in never-smoking women and secondhand smoke, radon, residential and occupational exposures to secondhand smoke and other carcinogens, air pollution, personal history of chronic respiratory conditions, and personal and family history of lung cancer.
- Several studies associate air pollution and lung cancer risk in never-smokers
- Several conflicting studies about association between reproductive and hormonal factors and lung cancer risk in never-smoking women

- Role of reproductive and hormonal factors in lung cancer in never-smoking women is still not fully understood
- Studies assessing lung cancer risk and reproductive and hormonal history ask about age at menarche, age at menopause, oral contraceptive use, hormone therapy use, and duration of hormone therapy use
- Adenocarcinoma subtype is most prevalent in never-smoking women
- EGFR tumor mutations are most prevalent in never-smoking women
- Never-smoking women more likely to be of Asian descent
- Longer overall survival time for never-smokers with lung cancer than ever-smokers

Gaps and future directions identified by the authors:

- Conflicting results exist on the association between lung cancer risk in never-smoking women and the following potential risk factors: reproductive and hormonal history, diet, and BMI
- Future studies should ask never-smokers with lung cancer detailed questions about the number of family members diagnosed with cancer and at what age
- Future studies should ask never-smokers with lung cancer how long they have lived with the family member that smoked cigarettes at home as a child and during adult life
- Future studies should ask never-smokers with lung cancer detailed questions about environmental exposures
- Future studies should ask never-smoking women with lung cancer detailed questions about reproductive and hormonal history

- Future studies should ask never-smoking women age at smoking initiation, cigarette content, and number of pack years to confirm their true smoking status
- Several studies reporting results based on self-reported data experience recall bias

4.2 Lung Cancer Patient Questionnaire

Utilizing the themes and gaps identified in the literature review search, the following questionnaire was developed. Questions were organized into twelve categories: demographics, cigarette smoking, other smoking habits, secondhand smoke, alcohol, BMI, medical history, family history, reproductive and hormonal history, radon, occupational history, and residential history. Residential and radon history is of particular interest to western Pennsylvania because this area has high air pollution and many homes with high radon. Allegheny County ranked within the top 2 percent of all U.S. counties for cancer risk from all Hazardous Air Pollutants (HAPS) with diesel exhaust particulates (DPM) representing the single strongest driver of risk, exceeding the next pollutant by 5 times (Michanowicz et al., 2013). Though I identified diet as a possible risk factor for lung cancer in never-smoking women and included it in the background section, we did not include questions about diet in this questionnaire for two reasons: 1) the studies that assessed diet and lung cancer risk in never-smokers are prospective studies, and our project is not prospective in design, as we only include lung cancer patients as participants and 2) we are limited by considering the length of the questionnaire. The questionnaire takes approximately 30-40 minutes to complete. Participants will take the questionnaire home or receive an electronic version by email to complete at their own pace.

- White
- Other
- Don't Know
- Prefer not to answer

5. Do you consider yourself to be Hispanic or Latino?

- Yes
- No
- Unknown

6. What is the highest level of education you have completed or the highest degree you have received?

- 8th grade or less
- Some high school, but did not graduate
- High school or GED
- Some college or 2-year degree
- 4-year college graduate
- More than 4-year college graduate

7. What is your current marital status?

- Married or living as married
- Divorced
- Widowed
- Separated
- Never married
- Prefer not to answer

B. Cigarette smoking:

8. In your entire life, have you smoked 100 or more cigarettes?

- Yes
- No (skip to question 13)
- Don't know

9. At what age did you start smoking cigarettes regularly?

- _____

10. How many cigarettes do/did you usually smoke in a day?

- _____ cigarettes/day

11. Since you started smoking cigarettes, have you ever quit for one year or longer?
- Yes, **if so how many times did you quit for 1 year or longer:** _____
 - No
 - Don't Know

12. What is your current smoking status?
- Currently smoking
 - Quit smoking, if so at what age did you last quit: _____

C. Other smoking habits

13. Have you ever used any of the following tobacco products one or more times during your lifetime? (If not applicable, skip to question 14)

Tobacco Product	Ever Used (yes/no)	Age you started using regularly?	Are you currently using?	How often?	If not currently using, at what age did you stop?
Regular pipe				/day /week /month	
E-cigarette or vape				/day /week /month	
Cigar, cigarillo, or little filtered cigar				/day /week /month	
Water pipe or hookah				/day /week	

				/month	
Smokeless tobacco (e.g. moist snuff, dip, spit, chew tobacco, snus, or dissolvable)				/day /week /month	
Other, please specify: _____				/day /week /month	

D. Secondhand smoke

Secondhand smoke is the smoke you inhale when you are *not* smoking yourself, but rather the smoke you inhale when cigarette smokers exhale and the smoke that comes from a cigarette's burning end.

14. During your childhood, did you live with a regular cigarette smoker who smoked inside your home?

- Yes
- No (skip to question 16)
- Don't know

15. For how many years were you exposed to secondhand smoke as a child?

- 1-5 years
- 6-10 years
- 11-15 years
- 15 + years

16. Are you exposed to secondhand smoke on a regular basis in your current household?

- Yes
- No (skip question 18)
- Don't know

17. For how many years have you been exposed to secondhand smoke in your current household?

- 1-5 years
- 6-10 years
- 11-15 years
- 15 + years

18. Have you been exposed to secondhand smoke on a regular basis at any previous or current job?

- Yes
- No (skip to question 20)
- Don't know

19. For how many years have you been exposed to secondhand smoke at any previous or current job?

- 1-5 years
- 6-10 years
- 11-15 years
- 15 + years

E. Alcohol

20. Before you were diagnosed with lung cancer, did you drink alcoholic beverages?

- Yes
- No (skip to question 23)

21. Before you were diagnosed with lung cancer, how often do you drink alcoholic beverages?

- Monthly or less
- 2-4 times a month
- 2-3 times a week
- 4 or more times a week

22. Before you were diagnosed, how many standard drinks containing alcohol do you have on a typical day when you are drinking? (1 standard drink = 12 oz of beer, 5 oz of wine, 8 oz of malt liquor, or 1.5 oz or "shot" of 80-proof liquor.)

- None
- 1 or 2
- 3 or 4
- 5 or 6
- 7 to 9
- 10 or more

23. Do you currently drink alcohol beverages?

- Yes
- No (skip to question 26)

24. Currently, how often do you drink alcoholic beverages?

- Monthly or less
- 2-4 times a month

- 2-3 times a week
- 4 or more times a week

25. Currently, how many standard drinks containing alcohol do you have on a typical day when you are drinking?

- None
- 1 or 2
- 3 or 4
- 5 or 6
- 7 to 9
- 10 or more

F. BMI

26. What is your current height?

- ___ ft ___ in

27. What is your current weight?

- _____lbs.

28. During your 20s, what was your weight?

- _____lbs.

29. During your 30s, what was your weight?

- _____lbs.

30. During your 40s, what was your weight?

- _____lbs.

31. During your 50s, what was your weight?

- _____lbs.

32. During your 60s, what was your weight?

- _____lbs.

G. Medical history

33. At what age were you diagnosed with lung cancer?

- _____

34. What type of lung cancer have you been diagnosed with?
- Non-small cell lung cancer (NSCLC)
 - Small-cell lung cancer (SCLC)
35. During the year before your lung cancer diagnosis, did you have any of the following symptoms? Check all that apply.
- Coughing
 - Chest pain
 - Shortness of breath
 - Wheezing
 - Coughing up blood
 - Feeling tired constantly
 - Weight loss
 - Pneumonia
 - Headache
 - Other, please specify _____
36. Have you ever been diagnosed with any of the following conditions? Check all that apply.
- Asthma
 - Chronic obstructive pulmonary disease (COPD)
 - Chronic bronchitis
 - Emphysema
 - Pneumococcal pneumonia
 - Tuberculosis
 - None
 - Don't Know
37. Are you currently receiving treatment for lung cancer?
- Yes, **if so what treatment:** _____
 - No
38. What type of treatment have you received for lung cancer? Check all that apply.
- Surgery
 - Radiation therapy
 - Chemotherapy
 - Targeted therapy
 - Immunotherapy
 - Other, please specify _____
 - None
39. Have you been diagnosed with head and neck cancer?
- Yes, **if so at what age:** _____

No

40. Have you been diagnosed with any other cancers?

- Yes
- No (skip to question 42)
- Don't Know

41. What other cancer(s) have you been diagnosed with and at what age?

- _____, _____ years
- _____, _____ years
- _____, _____ years

H. Family history

42. Has anyone in your family that is related to you by blood, ever been diagnosed with lung cancer? (Consider first-degree relatives only: parents, brothers or sisters, half siblings, or children).

- Yes
- No (skip to question 44)
- Don't know

43. Which family member(s) was diagnosed with lung cancer and at what age? Check all that apply.

Family member	Lung cancer (Y/N)	Age at diagnosis
Mother		
Father		
Brother		
Sister		
Child		

44. Have any of your first-degree relatives been diagnosed with head and neck cancer?

- Yes
- No (skip to question 46)
- Don't know

45. Which family member(s) was diagnosed with head and neck cancer and at what age? Check all that apply.

Family member	Lung cancer (Y/N)	Age at diagnosis
---------------	-------------------	------------------

Mother		
Father		
Brother		
Sister		
Child		

I. Reproductive and Hormonal history

46. Have you ever had a menstrual period?

- Yes
- No (skip to question 48)
- Don't know
- Prefer not to answer (skip to question 48)

47. How old were you when your menstrual periods began?

- Younger than 10 years
- 10-11 years
- 12-13 years
- 14-15 years
- 16 years or older

48. Have you used any of the following forms of contraception and for how long?
Check all that apply.

Form of contraception	Have used (Y/N)	<6 months	6 months – 1 year	1-5 years	5+ years
Pill					
IUD					
Injection					
Implant					
Patch					
Vaginal					
Other, please specify:					

- None
- Prefer not to answer

49. Have you ever been pregnant?

- Yes
- No (skip to question 56)

- Don't know
- Prefer not to answer (skip to question 56)

50. At what age did you first get pregnant?

51. Including live births, stillbirths, miscarriages, abortions, and tubal and ectopic pregnancies, how many times have you been pregnant?

52. How many live births have you experienced?

53. At what age did you have your first child?

54. Did you breastfeed any of your children?

- Yes
- No (skip to question 56)

55. How long did you breastfeed your child? (If you have multiple children, please provide an average number of months).

_____ months

56. Have you tried to get pregnant but have not been able to do so?

- Yes
- No (skip to question 59)
- Prefer not to answer

57. Have you ever used any type of assisted reproductive technology to have children?

- Yes
- No (skip to question 59)
- Prefer not to answer

58. What type of assisted reproductive technology have you used? If you have, please specify how many times you used each. Check all that apply.

- In vitro fertilization (IVF), _____ times
- Intrauterine insemination (IUI), _____ times
- Gamete intrafallopian transfer (GIFT), _____ times
- Zygote intrafallopian transfer (ZIFT), _____ times
- Frozen embryo transfer, _____ times
- Other, specify _____
- Don't know
- None

59. Are you still having menstrual periods?
- Yes (skip to question 68)
 - No
 - Never had a menstrual period (skip to question 68)
60. How old were you when you had your last menstrual period?
- Younger than 40 years
 - 40-44 years
 - 45-49 years
 - 50-54 years
 - 55 years or older
61. If you no longer have your menstrual period, what was the cause or reason for it?
- Natural menopause
 - Surgical (complete hysterectomy-ovaries and uterus removed)
 - Surgical (partial hysterectomy-only uterus removed)
 - Medications or chemotherapy
 - Don't know
 - Other, please specify _____
62. Have you ever used any type of hormonal therapy for menopause?
- Yes
 - No (skip to question 68)
63. Which of the following hormonal therapies do you or have you used for menopause? Check all that apply.
- Estrogen only (e.g. Premarin®, Estrace®, Cenestin)
 - Progesterone only (e.g. Provera®, Prometrium®, Aygestin®)
 - Estrogen and Progesterone (e.g. Activella®, Climara Pro®, Angeliq®)
 - Testosterone
 - Other, please specify _____
 - None
 - Don't know
64. At what age did you start this hormonal therapy for menopause?
- _____
65. Are you still using hormonal therapy for menopause?
- Yes
 - No
66. How long have or did you use hormonal therapy for menopause?
- < 6 months
 - 6 months to 1 year
 - 1 to 5 years
 - 5 to 10 years

10+ years

67. If you have stopped using hormonal therapy for menopause, at what age did you stop?

68. Which of the following hormones do you or have you used for reasons other than menopause? Check all that apply.

- Estrogen only (e.g. Estrace®, Alora®, Climera®, Delestrogen®)
- Progesterone only (e.g. Provera®, Prometrium®, Aygestin®)
- Estrogen and Progesterone (e.g. Activella®, Climara Pro®, Angeliq®)
- Testosterone (e.g. Delatestryl®, Depo®, Aveed®, Androgel®)
- Other, please specify _____
- None (skip to section J)
- Don't know

69. At what age did you start taking these hormones for reasons other than menopause?

70. Are you still using these hormones for reasons other than menopause?

- Yes
- No

71. How long have or did you use these hormones for reasons other than menopause?

- < 6 months
- 6 months to 1 year
- 1 to 5 years
- 5 to 10 years
- 10+ years

72. If you have stopped using these hormones, at what age did you stop?

J. Radon

73. Radon is a colorless and odorless radioactive gas that exists in soil, but sometimes can enter your home through cracks or gaps and build up. From your knowledge, has your current or previous home had elevated radon levels (Elevated radon level = 4 pCi/L or greater)

- Yes
- No
- Don't know

74. From your knowledge, has your current or a previous home been tested for its radon levels?

- Yes
- No
- Don't know

75. Has your current or previous home received radon mitigation or remediation?

- Yes
- No
- Don't know

K. Occupational history

76. Have you worked for 12 months or more in any of the following occupations or industries?

- | | |
|------------------------------------|--|
| Asbestos work | Foundry or steel milling |
| Chemicals or plastic manufacturing | Sandblasting |
| Coal mining | Tile work |
| Welding | Insulation installment |
| Painting | Motor vehicle manufacturing and repair |
| Radiologic technology | |

- Yes
- No

77. What is the total number of years you worked in this occupation or industry?

Occupation or Industry	Total number of years worked
Asbestos work	
Welding	
Chemicals or plastic manufacturing	
Foundry or steel milling	
Sandblasting	
Coal mining	
Tile work	
Insulation installment	
Painting	
Motor vehicle manufacturing and repair	
Radiologic technologist	

78. What is your current occupation?

□ _____

L. Residential history (try to be as specific as possible)

79. Where did you live most of your pre-teen life? (age 0-12)

□ _____
Town/City, State, Country

80. While you lived there, were there nearby steel plants or coal mines or did a lot of pollution occur? Please share anything you remember or were aware of.

□ _____

81. Where did you live most of your teen life? (age 13 to 18)

□ _____
Town/ City, State, Country

82. While you lived there, were there nearby steel plants or coal mines or did a lot of pollution occur? Please share anything you remember or were aware of.

□ _____

83. Where have you lived most of your adult life? (age 18+)

□ _____
Town/City, State, Country

84. While you lived there, were there nearby steel plants or coal mines or did a lot of pollution occur? Please share anything you remember or were aware of.

□ _____

Comments Box: (Is there anything else you would like to share?)

Thank you for participating!

4.3 Questionnaire Readability

Microsoft Word Document statistics are shown below in Figure 2. It indicated a Flesch-Kincaid Grade Level of 5.8, meaning the developed questionnaire reads at a fifth grade level. It also showed a Flesch Reading Ease score of 70.4 which, based on the Flesch Reading Ease score breakdown, is defined as fairly easy for the average adult to read and understand.

Counts	
Words	2,053
Characters	10,910
Paragraphs	419
Sentences	122
Averages	
Sentences per Paragraph	1.1
Words per Sentence	9.8
Characters per Word	4.6
Readability	
Flesch Reading Ease	70.4
Flesch-Kincaid Grade Level	5.8
Passive Sentences	10.6%

Figure 3: Microsoft Word Document Statistics for the Questionnaire

5.0 Discussion

Lung cancer incidence in never-smoking women is increasing and the literature review analysis provided an understanding of various risk factors playing a role in this observation. The literature review identified areas of potential risk factors to focus on in the patient questionnaire. The primary risk factors I identified were: secondhand smoke, radon, air pollution, reproductive and hormonal factors, occupational history, family history of cancer, medical history, and BMI. We developed and categorized our questions based on the main themes and gaps in literature the authors identified from the papers in Table 1.

Prior to Section A, we ask participants to provide their personal information including their name, address, and contact information. It is important to include this in a questionnaire to ensure we can easily identify participants in the study records and potentially update their answers in the future (Stehr-Green Paul, 2014). Section A asked participants general demographic questions. It is essential to ask these questions so we can characterize the participant population and characteristics like race or educational level can be used to see if they affect lung cancer risk in never-smoking women (Stehr-Green Paul, 2014). One particular case series study conducted in Canada of 1,005 lung cancer patients found that never-smokers with lung cancer were more likely to be of Asian background than non-Asian background, with an OR of 6.4 (95% CI 2.76-5.82, $p < 0.001$) (Myers et al., 2021). A few studies have also assessed the impact of education level on risk of lung cancer development in never-smokers, so we felt it would be useful to include a question about education level in the questionnaire (Al-Zoughool et al., 2013; Liang et al., 2019; Viñolas et al., 2017).

Section B asks questions about the participants' cigarette smoking habits. These questions establish the participants' current cigarette smoking status and describes their history of use. Although we anticipate the participants to be never-smokers, they all may not be, so it is important to confirm their current cigarette smoking status. To assess their status, we ask the question, "In your entire life, have you smoked 100 or more cigarettes?" and participants who answer no are confirmed to be never-smokers. Smith et al. suggested that it would be useful for future studies about risk factors of lung cancer in never-smokers to ask participants their age at smoking initiation and cigarette content and identify participants' number of pack years. For these reasons, we included questions asking the age they started smoking cigarettes regularly, how many they do or did smoke in a day, and if they have ever quit for one or more years. We can utilize the answers to these questions to calculate pack years. Section C asks questions about other smoking habits. We ask about other smoking habits because often the focus is only on cigarette smoking because that is how never-smokers are defined. Other smoking habits, like cigar or e-cigarette use, do affect lung cancer risk. This section only includes one table for participants to fill out, as this was the best way to shorten the questionnaire and ask about their habits in a concise manner (Nieuwenhuijsen, 2005).

Section D asks participants about known secondhand smoke exposure in various settings. Several studies have found strong associations between lung cancer risk in never-smokers and secondhand smoke exposure in childhood, current household, and at a current or previous job (Abdel-Rahman, 2020; Kim et al., 2018), and conflicting data exists, therefore exposures in each of these settings were questioned.

Section E asks questions about alcohol use during the year prior to their lung cancer diagnosis and current alcohol use. One case-control study found that low-alcohol use is not

associated with lung cancer risk in never-smokers, but the odds of developing lung cancer was 1.59 times more likely in never-smokers who consumed high-degree liquor compared to those that did not. Asking about previous and current alcohol use will allow us to better evaluate alcohol as a potential risk factor. Section F asked about current height and weight during their 20s, 30s, 40s, 50s, and 60s, so we could determine BMI over time. Studies have been conducted to assess the association between BMI and lung cancer risk in never-smokers, but conflicting data exists (Smith et al., 2012; Song et al., 2008; Turner et al., 2011). Due to the conflicting studies and authors of these studies identifying BMI and lung cancer risk as a gap in literature, we included questions about it.

Section G and Section H ask questions about medical history and family history, respectively. One main theme identified from the literature review table was that several studies found significant associations between lung cancer risk in never-smoking women and personal history of chronic respiratory conditions (D. R. Brenner et al., 2012; Darren R. Brenner et al., 2010; Liang et al., 2019; Park et al., 2020; Turner et al., 2011). For this reason, we included a question in Section G about diagnoses of several chronic respiratory conditions. Another main theme identified from the studies in Table 1 was that there may be a significant association between lung cancer risk in never-smoking women and family history of cancer, especially lung cancer (Darren R. Brenner et al., 2010; Coté et al., 2012; Gaughan et al., 2013; Liang et al., 2019; Yin et al., 2021). Yin et al. stated that asking participants how many members were diagnosed with cancer and at what age for each would have strengthened their findings, so we included questions to address those gaps.

Section I asks participants about their reproductive and hormonal history. One of the main, and likely most significant themes, identified in several studies of the literature review table was

that the role of reproductive and hormone history in lung cancer risk in never-smoking women is not fully understood and conflicting results exist. Due to these reasons, this was the longest and most detailed section of the questionnaire because improving our understand of this risk factor may fill this gap in literature. Studies that assess the role of reproductive and hormone history and lung cancer in never-smoking women asked about age at menarche, age at menopause, oral contraceptive use, hormone therapy use, and duration of hormone therapy use, so these were all questions we included in Section I (Baik et al., 2010; Brinton et al., 2011; Chakraborty et al., 2010; Clague et al., 2011; Dias et al., 2017; Kabat et al., 2007; Mazières et al., 2013; Weiss et al., 2007; Yao, Gu, Zhu, Yuan, & Song, 2013). This section also required extensive thought and revision regarding sensitivity, structure and design of questions. We strategically used tables, skip patterns, fill-in-the-blank questions, and provided brand names of common hormone therapies to ensure participants are able to fill out the questionnaire accurately.

Section J asks three questions about radon exposure in the home, an established risk factor for lung cancer development (Grundy et al., 2017; Lorenzo-González et al., 2019). These authors suggested future studies ask subjects about known residential elevated radon levels and residential radon remediation. Radon may not be a term that people are familiar with, so to provide further clarity we defined the term (Nieuwenhuijsen, 2005). Section K asks questions about participants' occupational history. Authors of studies in Table 1 found an increased risk of lung cancer in never-smokers associated with certain occupational exposures such as, asbestos, solvents, paints, welding equipment, pesticides, grain elevator dust, wood dust and smoke, and non-tobacco related exhaust (Darren R. Brenner et al., 2010; Liang et al., 2019; D. A. Siegel et al., 2021). Using the specific exposures assessed in these studies, we compiled a list of common occupations that have these exposures and created a table for ease of use by participants.

Section L asks participants about their residential history so we can get a better picture of the environment they have lived in. Many authors of the studies in the literature review table suggested that epidemiological studies should prioritize retrieving information about air pollution, as it may be a factor for lung cancer development in never-smokers (Gowda et al., 2019; Myers et al., 2021). Although questions in this section are structured as leading questions, a way of phrasing that generally should be avoided because it may influence participants' response, we made an exception because we felt it was the best way to ask about nearby coal mines, steel plants, and air pollution. In this section we felt it was appropriate to utilize open-ended questions to give participants the space to share their experience, memory, and knowledge about their surroundings at various periods in their life (Nieuwenhuijsen, 2005). This was also a reason for including the comments box.

I assessed the readability of the patient questionnaire using the Flesch-Kincaid readability tool. Results indicated a Flesch-Kincaid Grade Level of fifth grade with a score of 5.8 and a Flesch Reading Ease score of 70.4, meaning the questionnaire is fairly easy for the average adult to read. According to a study done by the Barbara Bush Foundation for Family Literacy, over half of the United States population between the ages of 16 and 74 read at a sixth-grade level or below (Barbara Bush Foundation for Family Literacy, 2019). The Flesch-Kincaid Grade Level score falls under the sixth-grade threshold, so we can be assured that the average adult will have the functional literacy skills to complete the questionnaire. Based only on the Flesch-Kincaid readability scores, this questionnaire is written at an appropriate level for our target population, adult women. However, one readability tool may not be enough or be the most holistic approach to assess readability of the questionnaire.

This project has some limitations. The literature search did not involve a quality assessment, so the studies in Table 1 used to develop the questionnaire may not all be considered high-quality studies. In addition, the time period used and focus on English-only papers may have resulted in missing particular risk factors that could have been assessed in the questionnaire.

The questionnaire was made to be self-administered and this design has several benefits. Participants can complete the questionnaire at their own pace, in the format they prefer, paper or online. Because it is self-administered, this removes the chance of interviewer bias, which can affect the way questions are asked and answered. It also increases likelihood of participants to answer questions, particularly ones that are socially undesirable. Although, a potential limitation of this design is that people cannot ask someone a question if they do not understand a question in the questionnaire. This is also a fairly lengthy questionnaire with a total of 84 questions, so this may reduce response rates. However, because it is designed to be completed at home and we are transparent about the time it takes to complete, we do not expect results to be affected by this issue significantly. Lastly, because this questionnaire includes mostly retrospective survey questions, there is concern for recall bias as many of the questions requires participants to recall details from as early as their childhood.

6.0 Conclusion

The incidence of lung cancer in never-smoking women is rising, and current literature shows that various risk factors may be involved, but these factors need to be studied further as conflicting data exist (Cheng et al., 2021). The results from this questionnaire will fill this gap in literature by bettering our understanding of what factors may be driving lung cancer incidence to increase in this population and addressing the existing sex disparity in incidence. The public health significance of this project is that not only will the questionnaire results expand our knowledge of the risk factors involved, but may also inform policy reform to address occupational and residential exposures and air pollution. In addition, the questionnaire findings may allow public health professionals and practitioners to develop better screening methods and treatment options for never-smoking women. Future directions for this project will involve contacting current lung cancer patients to get their input and feedback on the current draft of the questionnaire. Once the questionnaire receives final approval, it will be used in a study comprised of female never-smoking lung cancer patients at UPMC Hillman Cancer Center.

Bibliography

- Abdel-Rahman, O. (2020). Incidence and Mortality of Lung Cancer Among Never Smokers in Relationship to Secondhand Smoking: Findings From the PLCO Trial. *Clin Lung Cancer, 21*(5), 415-420.e412. doi:10.1016/j.clcc.2020.04.009
- Al-Zoughool, M., Pintos, J., Richardson, L., Parent, M., Ghadirian, P., Krewski, D., & Siemiatycki, J. (2013). Exposure to environmental tobacco smoke (ETS) and risk of lung cancer in Montreal: a case-control study. *Environ Health, 12*, 112. doi:10.1186/1476-069x-12-112
- American Cancer Society. (2021). *Cancer Facts & Figures 2021*. Retrieved from Atlanta: <https://www.cancer.org/content/dam/cancer-org/research/cancer-facts-and-statistics/annual-cancer-facts-and-figures/2021/cancer-facts-and-figures-2021.pdf>
- Baik, C. S., Strauss, G. M., Speizer, F. E., & Feskanich, D. (2010). Reproductive factors, hormone use, and risk for lung cancer in postmenopausal women, the Nurses' Health Study. *Cancer Epidemiol Biomarkers Prev, 19*(10), 2525-2533. doi:10.1158/1055-9965.Epi-10-0450
- Barbara Bush Foundation for Family Literacy. (2019). 130 million adults in the U.S. lack proficiency in literacy Retrieved from <https://map.barbarabush.org/>
- Baur, C., & Prue, C. (2014). The CDC Clear Communication Index is a new evidence-based tool to prepare and review health information. *Health Promot Pract, 15*(5), 629-637. doi:10.1177/1524839914538969
- Beall, C. M., & Leslie, P. W. (2014). Collecting women's reproductive histories. *American journal of human biology : the official journal of the Human Biology Council, 26*(5), 577-589. doi:10.1002/ajhb.22543
- Brenner, D. R., Boffetta, P., Duell, E. J., Bickeböllner, H., Rosenberger, A., McCormack, V., . . . Hung, R. J. (2012). Previous lung diseases and lung cancer risk: a pooled analysis from the International Lung Cancer Consortium. *Am J Epidemiol, 176*(7), 573-585. doi:10.1093/aje/kws151
- Brenner, D. R., Hung, R. J., Tsao, M.-S., Shepherd, F. A., Johnston, M. R., Narod, S., . . . McLaughlin, J. R. (2010). Lung cancer risk in never-smokers: a population-based case-control study of epidemiologic risk factors. *BMC Cancer, 10*, 285-285. doi:10.1186/1471-2407-10-285
- Brinton, L. A., Gierach, G. L., Andaya, A., Park, Y., Schatzkin, A., Hollenbeck, A. R., & Spitz, M. R. (2011). Reproductive and hormonal factors and lung cancer risk in the NIH-AARP Diet and Health Study cohort. *Cancer epidemiology, biomarkers & prevention : a publication of the American Association for Cancer Research, cosponsored by the American Society of Preventive Oncology, 20*(5), 900-911. doi:10.1158/1055-9965.EPI-10-1325
- Cardona, A. F., Rojas, L., Zatarain-Barrón, Z. L., Ruiz-Patiño, A., Ricaurte, L., Corrales, L., . . . Arrieta, O. (2019). Multigene Mutation Profiling and Clinical Characteristics of Small-Cell Lung Cancer in Never-Smokers vs. Heavy Smokers (Geno1.3-CLICaP). *Front Oncol, 9*, 254. doi:10.3389/fonc.2019.00254

- Center for Disease Control and Prevention. (2020). *Behavioral Risk Factor Surveillance System Questionnaire*. Retrieved from <https://www.cdc.gov/brfss/questionnaires/pdf-ques/2020-BRFSS-Questionnaire-508.pdf>
- Chakraborty, S., Ganti, A. K., Marr, A., & Batra, S. K. (2010). Lung cancer in women: role of estrogens. *Expert review of respiratory medicine*, 4(4), 509-518. doi:10.1586/ers.10.50
- Cheng, E. S., Weber, M., Steinberg, J., & Yu, X. Q. (2021). Lung cancer risk in never-smokers: An overview of environmental and genetic factors. *Chinese journal of cancer research = Chung-kuo yen cheng yen chiu*, 33(5), 548-562. doi:10.21147/j.issn.1000-9604.2021.05.02
- Clague, J., Reynolds, P., Sullivan-Halley, J., Ma, H., Lacey, J. V., Jr., Henderson, K. D., . . . Bernstein, L. (2011). Menopausal hormone therapy does not influence lung cancer risk: results from the California Teachers Study. *Cancer Epidemiol Biomarkers Prev*, 20(3), 560-564. doi:10.1158/1055-9965.Epi-10-1182
- Coté, M. L., Liu, M., Bonassi, S., Neri, M., Schwartz, A. G., Christiani, D. C., . . . Hung, R. J. (2012). Increased risk of lung cancer in individuals with a family history of the disease: a pooled analysis from the International Lung Cancer Consortium. *Eur J Cancer*, 48(13), 1957-1968. doi:10.1016/j.ejca.2012.01.038
- Cronemberger, E., Baldotto, C., Marinho, F., De Marchi, P., Araújo, L. H., Franke, F., . . . Castro, G., Jr. (2020). Real-World Molecular Testing and Treatment Patterns in Brazilian Patients with Newly Diagnosed Locally Advanced or Metastatic NSCLC. *Clinics (Sao Paulo)*, 75, e1777. doi:10.6061/clinics/2020/e1777
- Cutilli, C. C., & Bennett, I. M. (2009). Understanding the health literacy of America: results of the National Assessment of Adult Literacy. *Orthopedic nursing*, 28(1), 27-34. doi:10.1097/01.NOR.0000345852.22122.d6
- de Torres, J. P., Bastarrika, G., Wisnivesky, J. P., Alcaide, A. B., Campo, A., Seijo, L. M., . . . Zulueta, J. J. (2007). Assessing the relationship between lung cancer risk and emphysema detected on low-dose CT of the chest. *Chest*, 132(6), 1932-1938. doi:10.1378/chest.07-1490
- Dias, M., Linhas, R., Campainha, S., Conde, S., & Barroso, A. (2017). Lung cancer in never-smokers - what are the differences? *Acta Oncol*, 56(7), 931-935. doi:10.1080/0284186x.2017.1287944
- Doll, R., & Hill, A. B. (1950). Smoking and carcinoma of the lung; preliminary report. *Br Med J*, 2(4682), 739-748. doi:10.1136/bmj.2.4682.739
- Dubin, S., & Griffin, D. (2020). Lung Cancer in Non-Smokers. *Missouri medicine*, 117(4), 375-379.
- Egholm, C. L., Packness, A., Stokholm, J., Rasmussen, K., Ellervik, C., Simonsen, E., . . . Jepsen, R. (2020). Questionnaire development for the Lolland-Falster Health Study, Denmark: an iterative and incremental process. *BMC Medical Research Methodology*, 20(1), 52. doi:10.1186/s12874-020-00931-1
- Fidler-Benaoudia, M. M., Torre, L. A., Bray, F., Ferlay, J., & Jemal, A. (2020). Lung cancer incidence in young women vs. young men: A systematic analysis in 40 countries. *Int J Cancer*, 147(3), 811-819. doi:10.1002/ijc.32809
- Fritz, I., & Olsson, H. (2018). Lung cancer in young women in southern Sweden: A descriptive study. *Clin Respir J*, 12(4), 1565-1571. doi:10.1111/crj.12712
- Garfinkel, L. (1981). Time Trends in Lung Cancer Mortality Among Nonsmokers and a Note on Passive Smoking. *JNCI: Journal of the National Cancer Institute*, 66(6), 1061-1066. doi:10.1093/jnci/66.6.1061

- Gaughan, E. M., Cryer, S. K., Yeap, B. Y., Jackman, D. M., & Costa, D. B. (2013). Family history of lung cancer in never smokers with non-small-cell lung cancer and its association with tumors harboring EGFR mutations. *Lung Cancer*, 79(3), 193-197. doi:10.1016/j.lungcan.2012.12.002
- Gowda, S. N., DeRoos, A. J., Hunt, R. P., Gasset, A. J., Mirabelli, M. C., Bird, C. E., . . . Bhatti, P. (2019). Ambient air pollution and lung cancer risk among never-smokers in the Women's Health Initiative. *Environ Epidemiol*, 3(6), e076. doi:10.1097/ee9.0000000000000076
- Grundy, A., Brand, K., Khandwala, F., Poirier, A., Tamminen, S., Friedenreich, C. M., & Brenner, D. R. (2017). Lung cancer incidence attributable to residential radon exposure in Alberta in 2012. *CMAJ Open*, 5(2), E529-e534. doi:10.9778/cmajo.20160053
- Ha, S. Y., Choi, S.-J., Cho, J. H., Choi, H. J., Lee, J., Jung, K., . . . Kim, J. (2015). Lung cancer in never-smoker Asian females is driven by oncogenic mutations, most often involving EGFR. *Oncotarget*, 6(7), 5465-5474. doi:10.18632/oncotarget.2925
- Hirayama, T. (1981). Non-smoking wives of heavy smokers have a higher risk of lung cancer: a study from Japan. *Br Med J (Clin Res Ed)*, 282(6259), 183-185. doi:10.1136/bmj.282.6259.183
- Huang, C., Qu, X., & Du, J. (2019). Proportion of lung adenocarcinoma in female never-smokers has increased dramatically over the past 28 years. *Journal of Thoracic Disease*, 11(7), 2685-2688. Retrieved from <https://jtd.amegroups.com/article/view/30211>
- Hung, R. J., Spitz, M. R., Houlston, R. S., Schwartz, A. G., Field, J. K., Ying, J., . . . Gorlova, O. Y. (2019). Lung Cancer Risk in Never-Smokers of European Descent is Associated With Genetic Variation in the 5(p)15.33 TERT-CLPTM1L1 Region. *J Thorac Oncol*, 14(8), 1360-1369. doi:10.1016/j.jtho.2019.04.008
- Islami, F., Goding Sauer, A., Miller, K. D., Siegel, R. L., Fedewa, S. A., Jacobs, E. J., . . . Jemal, A. (2018). Proportion and number of cancer cases and deaths attributable to potentially modifiable risk factors in the United States. *CA Cancer J Clin*, 68(1), 31-54. doi:10.3322/caac.21440
- Jemal, A., Miller, K. D., Ma, J., Siegel, R. L., Fedewa, S. A., Islami, F., . . . Thun, M. J. (2018). Higher Lung Cancer Incidence in Young Women Than Young Men in the United States. *New England Journal of Medicine*, 378(21), 1999-2009. doi:10.1056/NEJMoa1715907
- Kabat, G. C., Miller, A. B., & Rohan, T. E. (2007). Reproductive and hormonal factors and risk of lung cancer in women: a prospective cohort study. *Int J Cancer*, 120(10), 2214-2220. doi:10.1002/ijc.22543
- Kantor, E. D., Hsu, M., Du, M., & Signorello, L. B. (2019). Allergies and Asthma in Relation to Cancer Risk. *Cancer Epidemiol Biomarkers Prev*, 28(8), 1395-1403. doi:10.1158/1055-9965.Epi-18-1330
- Kim, A. S., Ko, H.-J., Kwon, J.-H., & Lee, J.-M. (2018). Exposure to Secondhand Smoke and Risk of Cancer in Never Smokers: A Meta-Analysis of Epidemiologic Studies. *International journal of environmental research and public health*, 15(9), 1981. doi:10.3390/ijerph15091981
- Kurahashi, N., Inoue, M., Liu, Y., Iwasaki, M., Sasazuki, S., Sobue, T., & Tsugane, S. (2008). Passive smoking and lung cancer in Japanese non-smoking women: a prospective study. *Int J Cancer*, 122(3), 653-657. doi:10.1002/ijc.23116
- Liam, C. K., Pang, Y. K., Leow, C. H., Poosparajah, S., & Menon, A. (2006). Changes in the distribution of lung cancer cell types and patient demography in a developing multiracial

- Asian country: experience of a university teaching hospital. *Lung Cancer*, 53(1), 23-30. doi:10.1016/j.lungcan.2006.03.009
- Liang, D., Wang, J., Li, D., Shi, J., Jing, J., Shan, B., & He, Y. (2019). Lung Cancer in Never-Smokers: A Multicenter Case-Control Study in North China. *Frontiers in Oncology*, 9. doi:10.3389/fonc.2019.01354
- Lorenzo-González, M., Ruano-Ravina, A., Torres-Durán, M., Kelsey, K. T., Provencio, M., Parente-Lamelas, I., . . . Barros-Dios, J. M. (2019). Lung cancer and residential radon in never-smokers: A pooling study in the Northwest of Spain. *Environ Res*, 172, 713-718. doi:10.1016/j.envres.2019.03.011
- Mah, V., Seligson, D. B., Li, A., Márquez, D. C., Wistuba, II, Elshimali, Y., . . . Goodglick, L. (2007). Aromatase expression predicts survival in women with early-stage non small cell lung cancer. *Cancer Res*, 67(21), 10484-10490. doi:10.1158/0008-5472.Can-07-2607
- Mazières, J., Rouquette, I., Lepage, B., Milia, J., Brouchet, L., Guibert, N., . . . Fouret, P. (2013). Specificities of lung adenocarcinoma in women who have never smoked. *J Thorac Oncol*, 8(7), 923-929. doi:10.1097/JTO.0b013e3182904dfb
- McKay, J. D., Hung, R. J., Han, Y., Zong, X., Carreras-Torres, R., Christiani, D. C., . . . SpiroMeta, C. (2017). Large-scale association analysis identifies new lung cancer susceptibility loci and heterogeneity in genetic susceptibility across histological subtypes. *Nature genetics*, 49(7), 1126-1132. doi:10.1038/ng.3892
- Michanowicz, D., Malone, S., Ferrar, S., Kelso, M., Clougherty, J., Kriesky, J., & Fabisiak, J. (2013). *PRETA Air: Hazardous Air Pollutants. Pittsburgh Regional Environmental Threats Analysis (PRETA) Report* Retrieved from Pittsburgh, Pennsylvania
- Myers, R., Brauer, M., Dummer, T., Atkar-Khattra, S., Yee, J., Melosky, B., . . . Lam, S. (2021). High-Ambient Air Pollution Exposure Among Never Smokers Versus Ever Smokers With Lung Cancer. *J Thorac Oncol*, 16(11), 1850-1858. doi:10.1016/j.jtho.2021.06.015
- National Library of Medicine. (2015). NIH Common Data Elements Repository, from National Institute of Health. Retrieved from <https://cde.nlm.nih.gov/home>
- Nieuwenhuijsen, M. J. (2005). Design of exposure questionnaires for epidemiological studies. *Occupational and environmental medicine*, 62(4), 272-214. doi:10.1136/oem.2004.015206
- North, C. M., & Christiani, D. C. (2013). Women and lung cancer: what is new? *Semin Thorac Cardiovasc Surg*, 25(2), 87-94. doi:10.1053/j.semtcvs.2013.05.002
- O'Keeffe, L. M., Taylor, G., Huxley, R. R., Mitchell, P., Woodward, M., & Peters, S. A. E. (2018). Smoking as a risk factor for lung cancer in women and men: a systematic review and meta-analysis. *BMJ open*, 8(10), e021611-e021611. doi:10.1136/bmjopen-2018-021611
- Office of Smoking and Health and National Center for Chronic Disease Prevention and Health Promotion. (2015). *National Adult Tobacco Survey Questionnaire*. Retrieved from https://www.cdc.gov/tobacco/data_statistics/surveys/nats/pdfs/2013-2014-questionnaire-tag508.pdf
- Office on, S., & Health. (2006). Publications and Reports of the Surgeon General. In *The Health Consequences of Involuntary Exposure to Tobacco Smoke: A Report of the Surgeon General*. Atlanta (GA): Centers for Disease Control and Prevention (US).
- Park, H. Y., Kang, D., Shin, S. H., Yoo, K. H., Rhee, C. K., Suh, G. Y., . . . Kwon, O. J. (2020). Chronic obstructive pulmonary disease and lung cancer incidence in never smokers: a cohort study. *Thorax*, 75(6), 506-509. doi:10.1136/thoraxjnl-2019-213732

- Parkin, D. M., Bray, F., Ferlay, J., & Pisani, P. (2005). Global Cancer Statistics, 2002. *CA: A Cancer Journal for Clinicians*, 55(2), 74-108. doi:<https://doi.org/10.3322/canjclin.55.2.74>
- Pelosof, L., Ahn, C., Gao, A., Horn, L., Madrigales, A., Cox, J., . . . Schiller, J. (2017). Proportion of Never-Smoker Non-Small Cell Lung Cancer Patients at Three Diverse Institutions. *Journal of the National Cancer Institute*, 109(7), djw295. doi:10.1093/jnci/djw295
- Pirie, K., Peto, R., Green, J., Reeves, G. K., & Beral, V. (2016). Lung cancer in never smokers in the UK Million Women Study. *Int J Cancer*, 139(2), 347-354. doi:10.1002/ijc.30084
- Radzikowska, E., Głaz, P., & Roszkowski, K. (2002). Lung cancer in women: age, smoking, histology, performance status, stage, initial treatment and survival. Population-based study of 20 561 cases. *Ann Oncol*, 13(7), 1087-1093. doi:10.1093/annonc/mdf187
- Readable. (2021). Flesch Reading Ease and the Flesch Kincaid Grade Level. Retrieved from <https://readable.com/readability/flesch-reading-ease-flesch-kincaid-grade-level/>
- Siegel, D. A., Fedewa, S. A., Henley, S. J., Pollack, L. A., & Jemal, A. (2021). Proportion of Never Smokers Among Men and Women With Lung Cancer in 7 US States. *JAMA Oncology*, 7(2), 302-304. doi:10.1001/jamaoncol.2020.6362
- Siegel, R. L., Jacobs, E. J., Newton, C. C., Feskanich, D., Freedman, N. D., Prentice, R. L., & Jemal, A. (2015). Deaths Due to Cigarette Smoking for 12 Smoking-Related Cancers in the United States. *JAMA Internal Medicine*, 175(9), 1574-1576. doi:10.1001/jamainternmed.2015.2398
- Siegel, R. L., Miller, K. D., Fuchs, H. E., & Jemal, A. (2021). Cancer Statistics, 2021. *CA: A Cancer Journal for Clinicians*, 71(1), 7-33. doi:<https://doi.org/10.3322/caac.21654>
- Siegel, R. L., Miller, K. D., Fuchs, H. E., & Jemal, A. (2022). Cancer statistics, 2022. *CA: A Cancer Journal for Clinicians*, 72(1), 7-33. doi:<https://doi.org/10.3322/caac.21708>
- Smith, L., Brinton, L. A., Spitz, M. R., Lam, T. K., Park, Y., Hollenbeck, A. R., . . . Gierach, G. L. (2012). Body mass index and risk of lung cancer among never, former, and current smokers. *Journal of the National Cancer Institute*, 104(10), 778-789. doi:10.1093/jnci/djs179
- Song, Y. M., Sung, J., & Ha, M. (2008). Obesity and risk of cancer in postmenopausal Korean women. *Journal of clinical oncology : official journal of the American Society of Clinical Oncology*, 26(20), 3395-3402. doi:10.1200/jco.2007.15.7867
- Stabile, L. P., Dacic, S., Land, S. R., Lenzner, D. E., Dhir, R., Acquafondata, M., . . . Siegfried, J. M. (2011). Combined Analysis of Estrogen Receptor β -1 and Progesterone Receptor Expression Identifies Lung Cancer Patients with Poor Outcome. *Clinical Cancer Research*, 17(1), 154-164. doi:10.1158/1078-0432.Ccr-10-0992
- Stehr-Green Paul, S.-G. J. K., Nelson Amy,. (2014). Developing a Questionnaire. In *Methods in Field Epidemiology* (Vol. 2).
- Toh, C. K., Gao, F., Lim, W. T., Leong, S. S., Fong, K. W., Yap, S. P., . . . Tan, E. H. (2006). Never-smokers with lung cancer: epidemiologic evidence of a distinct disease entity. *Journal of clinical oncology : official journal of the American Society of Clinical Oncology*, 24(15), 2245-2251. doi:10.1200/jco.2005.04.8033
- Trichopoulos, D., Kalandidi, A., Sparros, L., & MacMahon, B. (1981). Lung cancer and passive smoking. *Int J Cancer*, 27(1), 1-4. doi:10.1002/ijc.2910270102
- Turner, M. C., Krewski, D., Pope, C. A., 3rd, Chen, Y., Gapstur, S. M., & Thun, M. J. (2011). Long-term ambient fine particulate matter air pollution and lung cancer in a large cohort of never-smokers. *Am J Respir Crit Care Med*, 184(12), 1374-1381. doi:10.1164/rccm.201106-1011OC

- U.S. Department of Health and Human Services. (2014). *The Health Consequences of Smoking- 50 Years of Progress: A Report of the Surgeon General*. Retrieved from Atlanta, GA: <https://www.hhs.gov/sites/default/files/consequences-smoking-exec-summary.pdf>
- U.S. Environmental Protection Agency. (1992). Respiratory Health Effects of Passive Smoking (Also Known as Exposure to Secondhand Smoke or Environmental Tobacco Smoke ETS).
- Vieira, A. R., Abar, L., Vingeliene, S., Chan, D. S., Aune, D., Navarro-Rosenblatt, D., . . . Norat, T. (2016). Fruits, vegetables and lung cancer risk: a systematic review and meta-analysis. *Ann Oncol*, 27(1), 81-96. doi:10.1093/annonc/mdv381
- Viñolas, N., Garrido, P., Isla, D., Provencio, M., Majem, M., Artal, A., . . . Felip, E. (2017). Lung Cancer in Never-Smoking Women: A Sub-Analysis of the Spanish Female-Specific Database WORLD07. *Cancer Invest*, 35(5), 358-365. doi:10.1080/07357907.2017.1295461
- Wang, L. W., Miller, M. J., Schmitt, M. R., & Wen, F. K. (2013). Assessing readability formula differences with written health information materials: application, results, and recommendations. *Res Social Adm Pharm*, 9(5), 503-516. doi:10.1016/j.sapharm.2012.05.009
- Weiss, J., Hou, L., Shu, X.-O., Ji, B.-T., Yang, G., Li, H., . . . Zheng, W. (2007). Hormonal risk factors for lung cancer in female lifetime non-smokers from Shanghai, China. *Cancer Research*, 67(9_Supplement), 2612-2612.
- Weissfeld, J. L., Lin, Y., Lin, H. M., Kurland, B. F., Wilson, D. O., Fuhrman, C. R., . . . Diergaard, B. (2015). Lung Cancer Risk Prediction Using Common SNPs Located in GWAS-Identified Susceptibility Regions. *J Thorac Oncol*, 10(11), 1538-1545. doi:10.1097/jto.0000000000000666
- World Cancer Research Fund International. About the Global Cancer Update Programme. Retrieved from <https://www.wcrf.org/diet-activity-and-cancer/global-cancer-update-programme/about-the-global-cancer-update-programme/>
- Yang, S. H., Mechanic, L. E., Yang, P., Landi, M. T., Bowman, E. D., Wampfler, J., . . . Jen, J. (2005). Mutations in the tyrosine kinase domain of the epidermal growth factor receptor in non-small cell lung cancer. *Clin Cancer Res*, 11(6), 2106-2110. doi:10.1158/1078-0432.Ccr-04-1853
- Yano, E., Wang, X., Wang, M., Qiu, H., & Wang, Z. (2010). Lung cancer mortality from exposure to chrysotile asbestos and smoking: a case-control study within a cohort in China. *Occup Environ Med*, 67(12), 867-871. doi:10.1136/oem.2009.051615
- Yao, Y., Gu, X., Zhu, J., Yuan, D., & Song, Y. (2013). Hormone replacement therapy in females can decrease the risk of lung cancer: a meta-analysis. *PloS one*, 8(8), e71236-e71236. doi:10.1371/journal.pone.0071236
- Yin, X., Chan, C. P. Y., Seow, A., Yau, W.-P., & Seow, W. J. (2021). Association between family history and lung cancer risk among Chinese women in Singapore. *Scientific Reports*, 11(1), 21862. doi:10.1038/s41598-021-00929-9
- Yu, Y. H., Liao, C. C., Hsu, W. H., Chen, H. J., Liao, W. C., Muo, C. H., . . . Chen, C. Y. (2011). Increased lung cancer risk among patients with pulmonary tuberculosis: a population cohort study. *J Thorac Oncol*, 6(1), 32-37. doi:10.1097/JTO.0b013e3181fb4fcc
- Zhang, T., Joubert, P., Ansari-Pour, N., Zhao, W., Hoang, P. H., Lokanga, R., . . . Landi, M. T. (2021). Genomic and evolutionary classification of lung cancer in never smokers. *Nature genetics*, 53(9), 1348-1359. doi:10.1038/s41588-021-00920-0

Zhang, Z. L., Sun, J., Dong, J. Y., Tian, H. L., Xue, L., Qin, L. Q., & Tong, J. (2012). Residential radon and lung cancer risk: an updated meta- analysis of case-control studies. *Asian Pac J Cancer Prev*, 13(6), 2459-2465. doi:10.7314/apjcp.2012.13.6.2459