Specific Aims

This proposal focuses on enhancing the accuracy of mutation timing in non-small cell lung cancer (NSCLC) through advanced computational modeling using a Hidden Markov Model (HMM). Our project leverages the extensive genomic data from the PCAWG dataset to develop a robust model that will be specifically applied to understand NSCLC progression in African Americans.

Aim 1: Develop a Hidden Markov Model using the PCAWG dataset to accurately time mutations in NSCLC *Hypothesis: A specifically designed HMM can effectively use the PCAWG dataset to model mutation timing in NSCLC.*

- Construct an initial HMM framework specifically tailored to capture the progression dynamics of NSCLC based on the PCAWG dataset
- Refine the model to incorporate single nucleotide variants (SNVs) from the PCAWG dataset as primary emissions for mutation timing
- Optimize the model's parameters, including transition probabilities and emission distributions, to enhance prediction accuracy
- Prepare the model framework for application on a broader dataset by ensuring scalability and adaptability

Aim 2: Apply the refined HMM to the Black Lung Cancer Dataset to study mutation timing in NSCLC among African Americans *Hypothesis: Applying the HMM to the Black Lung Cancer Dataset will reveal unique mutational timing and patterns specific to the African American population with NSCLC.*

- Adapt the HMM to accommodate genetic and mutational diversity specific to the African American NSCLC population
- Analyze the mutation timings across different cities to identify common and unique mutational patterns
- Correlate mutational timelines with available clinical outcomes to assess prognostic implications of mutation patterns

Aim 3: Validate the HMM's performance in predicting accurate mutation timings in NSCLC *Hypothesis: The HMM will provide accurate and clinically relevant predictions of mutation timings, enhancing our understanding of NSCLC progression in African Americans.*

- Compare the HMM predictions with known clinical and pathological data to validate accuracy
- Perform statistical validation using cross-validation methods to evaluate the model's robustness and reliability
- Publish the validation results and methodology in a peer-reviewed computational biology journal

This computational approach aims to significantly advance our understanding of NSCLC progression by providing a precise temporal framework for mutation development, which could ultimately influence treatment strategies and outcomes for African American populations.

Research Strategy

Significance

Lung Cancer and Environmental Exposures Lung cancer remains the leading cause of cancer-related mortality worldwide, with non-small cell lung cancer (NSCLC) accounting for 85% of lung cancer cases in the US.[1] Akin to tobacco smoking, exposure to the complex mixture of air pollution, particularly fine particulate matter ($PM_{2.5}$) and nitric oxide (NO), poses a major risk factor for developing lung cancer. In heavily polluted cities like Los Angeles, exposure to these pollutants significantly increases the risk of developing lung cancer.[2,3] In 2014, the Nurse's Health Study found that living within 200 meters of a highway and a $10\ \mu g/m^3$ increase in $PM_{2.5}$ levels were associated with an increased risk of lung cancer (HR = 1.57; 95% CI: 1.26, 1.77).[4] Furthermore, a 2019 meta-analysis estimated that a $10\ \mu g/m^3$ increase in $PM_{2.5}$ exposure in Europe and North America increased lung cancer risk by 25%.[5]

Despite the clear evidence linking air pollution exposure to elevated lung cancer risk, the precise molecular mechanisms by which these complex pollutant mixtures initiate and promote NSCLC remain poorly understood, representing a critical knowledge gap. This study will investigate lung cancer in Blacks, an understudied group that exhibits a high prevalence of aggressive, early-onset tumors that are often driven by distinct molecular profiles like EGFR mutations.[6] Elucidating the environmental drivers and biological pathways of lung carcinogenesis in this subgroup could reveal novel diagnostic approaches.

Addressing Lung Cancer Inequities in Blacks Although African Americans/Blacks (Blacks) have lower smoking rates compared to non-Hispanic Whites, they experience significantly higher lung cancer incidence and mortality rates, especially among men.[7–9] This disparity is striking, as Blacks tend to initiate smoking later in life and consume fewer cigarettes compared to their White counterparts.[8,10] Black women, despite smoking fewer cigarettes, have the same or higher incidence of lung cancer as White women.

Current lung cancer screening guidelines based on pack-years and age[11] fail to adequately identify Blacks at risk. Blacks are diagnosed with lung cancer at a significantly younger age than Whites, often before reaching the screening threshold of 30 pack-years or age 55.[12] The molecular drivers underlying these aggressive, early-onset lung cancers in the Black population remain unclear. However, disparities in environmental exposures, particularly air pollution, are suspected to play a role.[6] Evidence shows that Blacks are consistently exposed to significantly higher levels of $PM_{2.5}$ and NO compared to non-Hispanic Whites. This study will utilize a multi-regional cohort of non-smokers, former smokers, and smokers, to identify the molecular connections between air pollutants and lung cancer in Blacks.

Furthermore, existing studies do not account for how social determinants of health in Blacks may modulate susceptibility to cancers.[9] Addressing this gap is crucial for accurately assessing risk and developing prevention strategies in diverse populations.

Characterization of Environmental Exposure Outdoor air pollution, including $PM_{2.5}$, is classified as a Group 1 carcinogen by the International Agency for Research on Cancer (IARC).[13] Past studies demonstrate a clear link between residing near major roadways and an elevated risk of developing lung cancer.[13] Exhaust from combustion engines releases a mixture of carcinogenic compounds into the atmosphere near major roadways. These pollutants include polycyclic aromatic hydrocarbons (PAHs), nitrogen oxides, and toxic heavy metals such as arsenic, nickel, and lead.[14] Previous studies have attempted to map air pollution levels using census tract data. However, these methods only detect a limited subset of pollutants, failing to capture the full complexity of environmental pollutants. Moreover, existing research does not account for how rising global temperatures associated with climate change may alter the chemical composition and carcinogenic potency of air pollution over time. Another major shortcoming is the lack of integration of social determinants of health, such as obesity, diabetes, and chronic inflammatory conditions, which may exacerbate susceptibility to cancer.

Potential for Transformative Impact This study will employ advanced geospatial methods to precisely quantify individual exposures to air pollutants in Black communities in LA, Chicago, New Orleans, Charlestown SC, Richmond VA, and Rochester NY. Crucially, it will integrate this environmental exposure data with social determinants of health and biological factors that modulate disease susceptibility in these communities. Black populations in LA have historically faced disproportionately higher exposure to air pollution due to factors like redlining, the placing of industrial facilities near their neighborhoods, and a lack of green spaces. Despite having some of the lowest rates of smoking in the US, LA suffers from some of the worst highway-generated air pollution. By precisely characterizing these elevated exposures and combining them with data on obesity, diabetes, chronic inflammation, and other risk factors prevalent in Black communities, the goal is to develop a comprehensive model elucidating how environmental drivers interact synergistically with social and biological parameters to initiate and promote aggressive, early-onset NSCLC in this population.

This multidisciplinary approach, combining external exposure assessments with internal susceptibility factors, will provide novel mechanistic insights into the environmental carcinogenesis pathways driving the excess lung cancer burden observed in Black communities. By integrating precise air pollution exposure data with epidemiological cohorts and molecular tumor profiling from Black NSCLC patients, this study will generate a comprehensive model of how environmental insults precipitate lung carcinogenesis in the context of social and biological vulnerabilities in this underserved population. Insights from this

innovative approach have the potential to transform our understanding of air pollution's role in NSCLC etiology in Blacks and identify new opportunities for targeted prevention, early detection, and treatment strategies. This is particularly important as while the rates of most lung cancers are declining, the incidence of NSCLC in non-smoking women of color is rapidly rising in LA and other cities.

Innovation

While lung cancer in smokers is well-studied, far less is known about environmental contributors in Black non-smokers. Investigating subtype-specific mechanisms could reveal novel vulnerabilities for targeted therapies. This proposal introduces several innovative elements to advance our understanding of air pollution's role in lung cancer. While previous studies have relied on census tract pollution maps that fail to capture individual exposures and pollutant complexity, this study pioneers advanced geospatial monitoring and modeling to precisely quantify personal exposures to $PM_{2.5}$, PAHs, NO, and metals — critical for defining exposure-response relationships. Existing studies have not accounted for how social determinants like obesity, diabetes, and inflammation modulate environmental cancer risk. This proposal uniquely integrates high-resolution exposure data with comprehensive individual health parameters to model the complex interplay between external exposures and internal susceptibility. By combining exposure science with epidemiology and molecular biology, this study will provide a comprehensive model that elucidates the environmental drivers and biological mechanisms underlying environmentally induced NSCLC initiation and progression.

Under the mentorship of Dr. Paul Spellman, I will leverage his scientific expertise to identify causal mutational signatures linked to specific carcinogens. I will employ the cutting-edge techniques developed by Dr. Spellman to meticulously pinpoint the molecular alterations induced by environmental exposures, thereby providing a direct link between the components of air pollution and the pathogenesis of lung cancer. Through the implementation of the proposed research, I aim to elucidate the precise mutational signatures associated with exposure to $PM_{2.5}$, PAHs, NO, and toxic metals found within the complex mixture of air pollution. By rigorously identifying these causal relationships, I will establish a direct and unambiguous connection between environmental exposures and the development of lung cancer.

In summary, this proposal leverages advanced exposure monitoring integrated with epidemiological and molecular approaches to generate an unprecedented multidisciplinary model of environmental lung carcinogenesis. This multidisciplinary approach has the potential to transform our understanding of the role of air pollution in NSCLC and identify new avenues for prevention and early detection, especially in high-risk never-smoker populations.

Approach

Aim 1: Develop a Hidden Markov Model using the PCAWG dataset to accurately time mutations in NSCLC

Rationale Lorem ipsum dolor sit amet, consectetuer adipiscing elit. Ut purus elit, vestibulum ut, placerat ac, adipiscing vitae, felis. Curabitur dictum gravida mauris. Nam arcu libero, nonummy eget, consectetuer id, vulputate a, magna. Donec vehicula augue eu neque. Pellentesque habitant morbi tristique senectus et netus et malesuada fames ac turpis egestas. Mauris ut leo. Cras viverra metus rhoncus sem. Nulla et lectus vestibulum urna fringilla ultrices. Phasellus eu tellus sit amet tortor gravida placerat. Integer sapien est, iaculis in, pretium quis, viverra ac, nunc. Praesent eget sem vel leo ultrices bibendum. Aenean faucibus. Morbi dolor nulla, malesuada eu, pulvinar at, mollis ac, nulla. Curabitur auctor semper nulla. Donec varius orci eget risus. Duis nibh mi, congue eu, accumsan eleifend, sagittis quis, diam. Duis eget orci sit amet orci dignissim rutrum.

1.1. Construct an initial HMM framework specifically tailored to capture the progression dynamics of NSCLC based on the PCAWG dataset Nam dui ligula, fringilla a, euismod sodales, sollicitudin vel, wisi. Morbi auctor lorem non justo. Nam lacus libero, pretium at, lobortis vitae, ultricies et, tellus. Donec aliquet, tortor sed accumsan bibendum, erat ligula aliquet magna, vitae ornare odio metus a mi. Morbi ac orci et nisl hendrerit mollis. Suspendisse ut massa. Cras nec ante. Pellentesque a nulla. Cum sociis natoque penatibus et magnis dis parturient montes, nascetur ridiculus mus. Aliquam tincidunt urna. Nulla ullamcorper vestibulum turpis. Pellentesque cursus luctus mauris.

Figure 1: An example figure with caption to explain.

1.2. Refine the model to incorporate single nucleotide variants (SNVs) from the PCAWG dataset as primary emissions for mutation timing Nulla malesuada porttitor diam. Donec felis erat, congue non, volutpat at, tincidunt tristique, libero. Vivamus viverra fermentum felis. Donec nonummy pellentesque ante. Phasellus adipiscing semper elit. Proin fermentum massa ac quam. Sed diam turpis, molestie vitae, placerat a, molestie nec, leo. Maecenas lacinia. Nam ipsum ligula, eleifend at, accumsan nec, suscipit a, ipsum. Morbi blandit ligula feugiat magna. Nunc eleifend consequat lorem. Sed lacinia nulla vitae enim. Pellentesque tincidunt purus vel magna. Integer non enim.

Praesent euismod nunc eu purus. Donec bibendum quam in tellus. Nullam cursus pulvinar lectus. Donec et mi. Nam vulputate metus eu enim. Vestibulum pellentesque felis eu massa.

1.3. Optimize the model's parameters, including transition probabilities and emission distributions, to enhance prediction accuracy Quisque ullamcorper placerat ipsum. Cras nibh. Morbi vel justo vitae lacus tincidunt ultrices. Lorem ipsum dolor sit amet, consectetuer adipiscing elit. In hac habitasse platea dictumst. Integer tempus convallis augue. Etiam facilisis. Nunc elementum fermentum wisi. Aenean placerat. Ut imperdiet, enim sed gravida sollicitudin, felis odio placerat quam, ac pulvinar elit purus eget enim. Nunc vitae tortor. Proin tempus nibh sit amet nisl. Vivamus quis tortor vitae risus porta vehicula.

Challenges & Alternative Approaches Suspendisse vel felis. Ut lorem lorem, interdum eu, tincidunt sit amet, laoreet vitae, arcu. Aenean faucibus pede eu ante. Praesent enim elit, rutrum at, molestie non, nonummy vel, nisl. Ut lectus eros, malesuada sit amet, fermentum eu, sodales cursus, magna. Donec eu purus. Quisque vehicula, urna sed ultricies auctor, pede lorem egestas dui, et convallis elit erat sed nulla. Donec luctus. Curabitur et nunc. Aliquam dolor odio, commodo pretium, ultricies non, pharetra in, velit. Integer arcu est, nonummy in, fermentum faucibus, egestas vel, odio.

Aim 2: Apply the refined HMM to the Black Lung Cancer Dataset to study mutation timing in NSCLC among African Americans

Rationale Lorem ipsum dolor sit amet, consectetuer adipiscing elit. Ut purus elit, vestibulum ut, placerat ac, adipiscing vitae, felis. Curabitur dictum gravida mauris. Nam arcu libero, nonummy eget, consectetuer id, vulputate a, magna. Donec vehicula augue eu neque. Pellentesque habitant morbi tristique senectus et netus et malesuada fames ac turpis egestas. Mauris ut leo. Cras viverra metus rhoncus sem. Nulla et lectus vestibulum urna fringilla ultrices. Phasellus eu tellus sit amet tortor gravida placerat. Integer sapien est, iaculis in, pretium quis, viverra ac, nunc. Praesent eget sem vel leo ultrices bibendum. Aenean faucibus. Morbi dolor nulla, malesuada eu, pulvinar at, mollis ac, nulla. Curabitur auctor semper nulla. Donec varius orci eget risus. Duis nibh mi, congue eu, accumsan eleifend, sagittis quis, diam. Duis eget orci sit amet orci dignissim rutrum.

2.1. Adapt the HMM to accommodate genetic and mutational diversity specific to the African American NSCLC population Nam dui ligula, fringilla a, euismod sodales, sollicitudin vel, wisi. Morbi auctor lorem non justo. Nam lacus libero, pretium at, lobortis vitae, ultricies et, tellus. Donec aliquet, tortor sed accumsan bibendum, erat ligula aliquet magna, vitae ornare odio metus a mi. Morbi ac orci et nisl hendrerit mollis. Suspendisse ut massa. Cras nec ante. Pellentesque a nulla. Cum sociis natoque penatibus et magnis dis parturient montes, nascetur ridiculus mus. Aliquam tincidunt urna. Nulla ullamcorper vestibulum turpis. Pellentesque cursus luctus mauris.

2.2. Analyze the mutation timings across different cities to identify common and unique mutational patterns Nulla malesuada porttitor diam. Donec felis erat, congue non, volutpat at, tincidunt tristique, libero. Vivamus viverra fermentum felis. Donec nonummy pellentesque ante. Phasellus adipiscing semper elit. Proin fermentum massa ac quam. Sed diam turpis, molestie vitae, placerat a, molestie nec, leo. Maecenas lacinia. Nam ipsum ligula, eleifend at, accumsan nec, suscipit a, ipsum. Morbi blandit ligula feugiat magna. Nunc eleifend consequat lorem. Sed lacinia nulla vitae enim. Pellentesque tincidunt purus vel magna. Integer non enim. Praesent euismod nunc eu purus. Donec bibendum quam in tellus. Nullam cursus pulvinar lectus. Donec et mi. Nam vulputate metus eu enim. Vestibulum pellentesque felis eu massa.

2.3. Correlate mutational timelines with available clinical outcomes to assess prognostic implications of mutation patterns Quisque ullamcorper placerat ipsum. Cras nibh. Morbi vel justo vitae lacus tincidunt ultrices. Lorem ipsum dolor sit amet, consectetuer adipiscing elit. In hac habitasse platea dictumst. Integer tempus convallis augue. Etiam facilisis. Nunc elementum fermentum wisi. Aenean placerat. Ut imperdiet, enim sed gravida sollicitudin, felis odio placerat quam, ac pulvinar elit purus eget enim. Nunc vitae tortor. Proin tempus nibh sit amet nisl. Vivamus quis tortor vitae risus porta vehicula.

Challenges & Alternative Approaches Fusce mauris. Vestibulum luctus nibh at lectus. Sed bibendum, nulla a faucibus semper, leo velit ultricies tellus, ac venenatis arcu wisi vel nisl. Vestibulum diam. Aliquam pellentesque, augue quis sagittis posuere, turpis lacus congue quam, in hendrerit risus eros eget felis. Maecenas eget erat in sapien mattis porttitor. Vestibulum porttitor. Nulla facilisi. Sed a turpis eu lacus commodo facilisis. Morbi fringilla, wisi in dignissim interdum, justo lectus sagittis dui, et vehicula libero dui cursus dui. Mauris tempor ligula sed lacus. Duis cursus enim ut augue. Cras ac magna. Cras nulla. Nulla egestas. Curabitur a leo. Quisque egestas wisi eget nunc. Nam feugiat lacus vel est. Curabitur consectetuer.

Aim 3: Validate the HMM's performance in predicting accurate mutation timings in NSCLC

Rationale Lorem ipsum dolor sit amet, consectetuer adipiscing elit. Ut purus elit, vestibulum ut, placerat ac, adipiscing vitae, felis. Curabitur dictum gravida mauris. Nam arcu libero, nonummy eget, consectetuer id, vulputate a, magna. Donec vehicula augue eu neque. Pellentesque habitant morbi tristique senectus et netus et malesuada fames ac turpis egestas. Mauris ut leo. Cras viverra metus rhoncus sem. Nulla et lectus vestibulum urna fringilla ultrices. Phasellus eu tellus sit amet tortor gravida placerat. Integer sapien est, iaculis in, pretium quis, viverra ac, nunc. Praesent eget sem vel leo ultrices bibendum. Aenean faucibus. Morbi dolor nulla, malesuada eu, pulvinar at, mollis ac, nulla. Curabitur auctor semper nulla. Donec varius orci eget risus. Duis nibh mi, congue eu, accumsan eleifend, sagittis quis, diam. Duis eget orci sit amet orci dignissim rutrum.

3.1. Compare the HMM predictions with known clinical and pathological data to validate accuracy Nam dui ligula, fringilla a, euismod sodales, sollicitudin vel, wisi. Morbi auctor lorem non justo. Nam lacus libero, pretium at, lobortis vitae, ultricies et, tellus. Donec aliquet, tortor sed accumsan bibendum, erat ligula aliquet magna, vitae ornare odio metus a mi. Morbi ac orci et nisl hendrerit mollis. Suspendisse ut massa. Cras nec ante. Pellentesque a nulla. Cum sociis natoque penatibus et magnis dis parturient montes, nascetur ridiculus mus. Aliquam tincidunt urna. Nulla ullamcorper vestibulum turpis. Pellentesque cursus luctus mauris.

3.2. Perform statistical validation using cross-validation methods to evaluate the model's robustness and reliability Nulla malesuada porttitor diam. Donec felis erat, congue non, volutpat at, tincidunt tristique, libero. Vivamus viverra fermentum felis. Donec nonummy pellentesque ante. Phasellus adipiscing semper elit. Proin fermentum massa ac quam. Sed diam turpis, molestie vitae, placerat a, molestie nec, leo. Maecenas lacinia. Nam ipsum ligula, eleifend at, accumsan nec, suscipit a, ipsum. Morbi blandit ligula feugiat magna. Nunc eleifend consequat lorem. Sed lacinia nulla vitae enim. Pellentesque tincidunt purus vel magna. Integer non enim. Praesent euismod nunc eu purus. Donec bibendum quam in tellus. Nullam cursus pulvinar lectus. Donec et mi. Nam vulputate metus eu enim. Vestibulum pellentesque felis eu massa.

3.3. Publish the validation results and methodology in a peer-reviewed computational biology journal Quisque ullamcorper placerat ipsum. Cras nibh. Morbi vel justo vitae lacus tincidunt ultrices. Lorem ipsum dolor sit amet, consectetuer adipiscing elit. In hac habitasse platea dictumst. Integer tempus convallis augue. Etiam facilisis. Nunc elementum fermentum wisi. Aenean placerat. Ut imperdiet, enim sed gravida sollicitudin, felis odio placerat quam, ac pulvinar elit purus eget enim. Nunc vitae tortor. Proin tempus nibh sit amet nisl. Vivamus quis tortor vitae risus porta vehicula.

Challenges & Alternative Approaches Fusce mauris. Vestibulum luctus nibh at lectus. Sed bibendum, nulla a faucibus semper, leo velit ultricies tellus, ac venenatis arcu wisi vel nisl. Vestibulum diam. Aliquam pellentesque, augue quis sagittis posuere, turpis lacus congue quam, in hendrerit risus eros eget felis. Maecenas eget erat in sapien mattis porttitor. Vestibulum porttitor. Nulla facilisi. Sed a turpis eu lacus commodo facilisis. Morbi fringilla, wisi in dignissim interdum, justo lectus sagittis dui, et vehicula libero dui cursus dui. Mauris tempor ligula sed lacus. Duis cursus enim ut augue. Cras ac magna. Cras nulla. Nulla egestas. Curabitur a leo. Quisque egestas wisi eget nunc. Nam feugiat lacus vel est. Curabitur consectetuer.

Glossary

References

[1] Julian R. Molina, Ping Yang, Stephen D. Cassivi, Steven E. Schild, and Alex A. Adjei. “Non–Small Cell Lung Cancer: Epidemiology, Risk Factors, Treatment, and Survivorship”. In: *Mayo Clinic proceedings. Mayo Clinic* 83.5 (May 2008). https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2718421/, pp. 584–594. ISSN: 0025-6196.

[2] Christine D. Berg, Joan H. Schiller, Paolo Boffetta, Jing Cai, Casey Connolly, Anna Kerpel-Fronius, Andrea Borondy Kitts, David C.L. Lam, Anant Mohan, Renelle Myers, Tejas Suri, Martin C. Tammemagi, Dawei Yang, and Stephen Lam. “Air Pollution and Lung Cancer: A Review by International Association for the Study of Lung Cancer Early Detection and Screening Committee”. In: *Journal of Thoracic Oncology* 18.10 (Oct. 2023), pp. 1277–1289. ISSN: 15560864. DOI: 10.1016/j.jtho.2023.05.024.

[3] Yanqian Huang, Meng Zhu, Mengmeng Ji, Jingyi Fan, Junxing Xie, Xiaoxia Wei, Xiangxiang Jiang, Jing Xu, Liang Chen, Rong Yin, Yuzhuo Wang, Juncheng Dai, Guangfu Jin, Lin Xu, Zhibin Hu, Hongxia Ma, and Hongbing Shen. “Air Pollution, Genetic Factors, and the Risk of Lung Cancer: A Prospective Study in the UK Biobank”. In: *American Journal of Respiratory and Critical Care Medicine* 204.7 (Oct. 2021), pp. 817–825. ISSN: 1073-449X. DOI: 10.1164/rccm.202011-4063OC.

[4] Robin C. Puett, Jaime E. Hart, Jeff D. Yanosky, Donna Spiegelman, Molin Wang, Jared A. Fisher, Biling Hong, and Francine Laden. “Particulate Matter Air Pollution Exposure, Distance to Road, and Incident Lung Cancer in the Nurses’ Health Study Cohort”. In: *Environmental Health Perspectives* 122.9 (Sept. 2014), pp. 926–932. ISSN: 1552-9924. DOI: 10.1289/ehp.1307490.

[5] Hung-Ling Huang, Yung-Hsin Chuang, Tzu-Hsuan Lin, Changqing Lin, Yen-Hsu Chen, Jen-Yu Hung, and Ta-Chien Chan. “Ambient Cumulative PM2.5 Exposure and the Risk of Lung Cancer Incidence and Mortality: A Retrospective Cohort Study”. In: *International Journal of Environmental Research and Public Health* 18.23 (Nov. 2021), p. 12400. ISSN: 1661-7827. DOI: 10.3390/ijerph182312400.

[6] Anita Marcinkiewicz, Aleksandra Ochotnicka, Karolina Borowska-Waniak, Kinga Skorupińska, Dominik Michalik, and Maja Borowska. “THE IMPACT OF AIR POLLUTION ON THE OCCURRENCE OF LUNG CANCER: A LITERATURE REVIEW”. In: *Archiv Euromedica* 13.5 (Oct. 2023). ISSN: 2199885X. DOI: 10.35630/2023/13/5.507.

[7] Nadine Belony, Bing Ren, Phuc Pham, Matthew Gregory, Pablo E. Puente, Nazarius S. Lamango, Ite A. Offringa, and Yong Huang. “Abstract 4038: Study of Therapeutic Effects of Polyisoprenylated Cysteinyl Amide Inhibitors on Lung Cancer Cells of Black Patients Using 3D-Printed Alveolar Model”. In: *Cancer Research* 82.12_Supplement (June 2022), p. 4038. ISSN: 0008-5472. DOI: 10.1158/1538-7445.AM2022-4038.

[8] Jose Thomas Thaiparambil, Zheng Yin, and Randa El-Zein. “Abstract 1948: Novel Insights into Lung Cancer & Chronic Obstructive Pulmonary Disease Racial Disparities”. In: *Cancer Research* 83.7_Supplement (Apr. 2023), p. 1948. ISSN: 0008-5472. DOI: 10.1158/1538-7445.AM2023-1948.

[9] Celina I. Valencia, Francine C. Gachupin, Yamilé Molina, and Ken Batai. “Interrogating Patterns of Cancer Disparities by Expanding the Social Determinants of Health Framework to Include Biological Pathways of Social Experiences”. In: *International Journal of Environmental Research and Public Health* 19.4 (Feb. 2022), p. 2455. ISSN: 1661-7827. DOI: 10.3390/ijerph19042455.

[10] Karyn Hede. “Drilling Down to the Causes of Racial Disparities in Lung Cancer”. In: *JNCI: Journal of the National Cancer Institute* 102.18 (Sept. 2010), pp. 1385–1387. ISSN: 0027-8874. DOI: 10.1093/jnci/djq371.

[11] Rebecca Landy, Li C. Cheung, Corey D. Young, Anil K. Chaturvedi, and Hormuzd A. Katki. “Absolute Lung Cancer Risk Increases among Individuals with $>$15 Quit-Years: Analyses to Inform the Update of the American Cancer Society Lung Cancer Screening Guidelines”. In: *Cancer* 130.2 (Jan. 2024), pp. 201–215. ISSN: 0008-543X. DOI: 10.1002/cncr.34758.

[12] Rafael Meza, Jihyoun Jeon, Iakovos Toumazis, Kevin Ten Haaf, Pianpian Cao, Mehrad Bastani, Summer S. Han, Erik F. Blom, Daniel E. Jonas, Eric J. Feuer, Sylvia K. Plevritis, Harry J. de Koning, and Chung Yin Kong. “Evaluation of the Benefits and Harms of Lung Cancer Screening With Low-Dose Computed Tomography: Modeling Study for the US Preventive Services Task Force”. In: *JAMA* 325.10 (Mar. 2021), pp. 988–997. ISSN: 1538-3598. DOI: 10.1001/jama.2021.1077.

[13] Shilpa N. Gowda, Anneclaire J. DeRoos, Rebecca P. Hunt, Amanda J. Gassett, Maria C. Mirabelli, Chloe E. Bird, Helene G. Margolis, Dorothy Lane, Matthew R. Bonner, Garnet Anderson, Eric A. Whitsel, Joel D. Kaufman, and Parveen Bhatti. “Ambient Air Pollution and Lung Cancer Risk among Never-Smokers in the Women’s Health Initiative”. In: *Environmental Epidemiology* 3.6 (Oct. 2019), e076. ISSN: 2474-7882. DOI: 10.1097/EE9.0000000000000076.

[14] Xian-Jun Yu, Min-Jun Yang, Bo Zhou, Gui-Zhen Wang, Yun-Chao Huang, Li-Chuan Wu, Xin Cheng, Zhe-Sheng Wen, Jin-Yan Huang, Yun-Dong Zhang, Xiao-Hong Gao, Gao-Feng Li, Shui-Wang He, Zhao-Hui Gu, Liang Ma, Chun-Ming Pan, Ping Wang, Hao-Bin Chen, Zhi-Peng Hong, Xiao-Lu Wang, Wen-Jing Mao, Xiao-Long Jin, Hui Kang, Shu-Ting Chen, Yong-Qiang Zhu, Wen-Yi Gu, Zi Liu, Hui Dong, Lin-Wei Tian, Sai-Juan Chen, Yi Cao, Sheng-Yue Wang, and Guang-Biao Zhou. “Characterization of Somatic Mutations in Air Pollution-Related Lung Cancer”. In: *EBioMedicine* 2.6 (June 2015), pp. 583–590. ISSN: 2352-3964. DOI: 10.1016/j.ebiom.2015.04.003.