Exploring Lung Adenocarcinoma Survival Rate Factors

Introduction:

According to the National Cancer Institute, approximately 40.5% of men and women will be diagnosed with cancer at some point during their life (2018). Of these, lung and bronchus cancers account for nearly 12% of all cancer cases. Lung adenocarcinoma, a subtype of non-small cell lung cancer, remains a significant global health concern with its high mortality rates.

Previous studies have emphasized the heterogeneous nature of lung adenocarcinoma, highlighting its diverse genetic and molecular profiles (Barbar, 2022). The identification of key genetic mutations, such as EGFR and KRAS, has paved the way for targeted therapies, but challenges persist in predicting patient responses and understanding the underlying factors contributing to the disease's aggressiveness (Yang, 2018). Recent literature also underscores the need for comprehensive analyses integrating clinical, genomic, and environmental factors to enhance our understanding of lung adenocarcinoma progression and treatment outcomes.

It is hypothesized that mutations in specific genes associated with lung adenocarcinomas can be used as predictors of survival rate. In addition to this, it is hypothesized that the type of mutation that occurs in these oncogenes will also be useful predictors in lung adenocarcinoma survival rate.

Methods:

Clinical datasets for adenocarcinoma will be collected from the Genomic Commons Data Portal. Data will include case demographics, prevalent oncogenes, and common mutations found in patients with lung adenocarcinoma. This study includes over 20,000 cases that fall within the “Lung” cohort within the GDC data portal.

Statistical tests, such as chi-square tests and logistic regression, will be used to investigate associations between genetic mutations and clinical outcomes. Survival analyses using Kaplan-Meier curves and Cox proportional hazards models will assess the impact of molecular and clinical variables on overall survival.

Results:

The clinical data of 2642 patients diagnosed with a disease includes various statistics like age at diagnosis and death, smoking habits, and disease progression. The mean age at diagnosis is 65.76 years (SD = 10.51), with the majority being diagnosed around 67.05 years, indicating a senior demographic is most affected. The patients' ages at death average 69.18 years (SD = 10.04), with the median significantly higher at 70.48 years, suggesting a slightly younger cohort at diagnosis compared to the time of death. Smoking habits reveal an average of 4.32 cigarettes per day (SD = 6.51) with a median much lower at 2.47, indicating a skewed distribution where

most patients smoke less, but a few smoke substantially more. Pack years smoked average at 46.18 (SD = 29.61), showing a wide variance in smoking duration among patients.

	vars	n	mean	sd	median	trimmed	mad	min	max
age_at_diagnosis	1	2642	65.7613275538457	10.5135937559697	67.0465753424658	66.4483327069374	9.86436739726028	12.0164383561644	90.0602739726027
age_at_death	2	981	69.1822238978956	10.0406873035354	70.4767123287671	69.797553442108	10.0085654794521	40.3808219178082	90.0602739726027
days_to_death	3	1027	808.048685491724	778.946406660962	562	681.702308626975	567.8358	-1	5287
cigarettes_per_day	4	1849	4.32351763633804	6.51292871439834	2.46575342465753	2.74992128606181	1.54352876712329	0	60
pack_years_smoked	5	1823	46.1794349972573	29.6073331957069	40	43.1049074708705	22.239	0	240

Figure 1. Descriptive statistics for continuous clinical data.

The patient demographic largely consists of 67.3% non-Hispanic or Latino individuals, with a gender distribution of 65.3% males compared to 34.7% females. The racial breakdown is predominantly white at 72.6%, followed by 11.5% black or African American, and 5.8% Asian. Vital status data indicates 59.9% of patients are alive, with 40.1% deceased, and pathologic stages data show a notable number of patients in Stage IA (21.4%) and Stage IB (27.5%), suggesting a considerable portion of the cohort is diagnosed early. Regarding primary diagnosis, the data includes 35.7% cases of Adenocarcinoma, and 41% cases of Squamous cell carcinoma. The predominant treatment types are Pharmaceutical Therapy (42.3%) and Radiation Therapy(43.7%). The smoker status shows a significant division with 69.3% smokers versus 30.7% non-smokers.

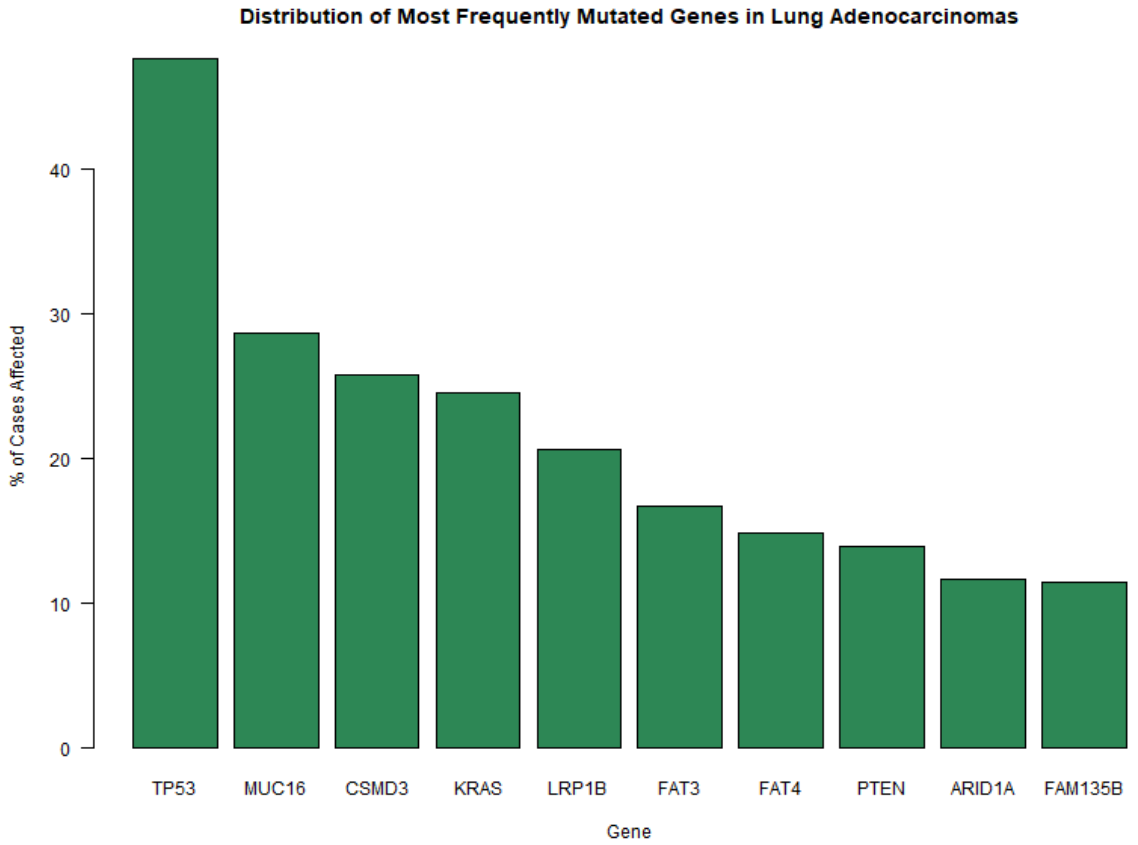

Figure 2. Barplot displaying the most frequently mutated genes found in lung adenocarcinomas

The most frequently mutated gene found in lung adenocarcinomas is tumor protein p53 (TP53). 48.64% of cases within the cohort had at least one mutation within this gene. After this, the most notable genes were MUC16 (28.64%), CSMD3 (25.80%), KRAS (24.57%), and LRP1B (20.60%). It is worth noting that each case can have more than one mutation. All of these genes are protein coding.

References

Barbar, J., Armach, M., Hodroj, M. H., Assi, S., El Nakib, C., Chamseddine, N., & Assi, H. I. (2022). Emerging genetic biomarkers in lung adenocarcinoma. SAGE open medicine, 10, 20503121221132352. https://doi.org/10.1177/20503121221132352

Myers D., Wallen J. (2023). In: StatPearls [Internet]. Lung Adenocarcinoma. Treasure Island (FL): StatPearls Publishing; 2024 Jan-. Available from: https://www.ncbi.nlm.nih.gov/books/NBK519578/

National Cancer Institute. (2018). Cancer of Any Site - Cancer Stat Facts. SEER. https://seer.cancer.gov/statfacts/html/all.html

Genomics Commons Data Portal. Portal.gdc.cancer.gov. Retrieved February 5, 2024, from https://portal.gdc.cancer.gov/projects/TCGA-LUAD

Yang P. (2009). Epidemiology of lung cancer prognosis: quantity and quality of life. Methods in molecular biology (Clifton, N.J.), 471, 469–486. https://doi.org/10.1007/978-1-59745-416-2_24

Yang, S., Yu, X., Fan, Y., Shi, X., & Jin, Y. (2018). Clinicopathologic characteristics and survival outcome in patients with advanced lung adenocarcinoma and KRAS mutation. Journal of Cancer, 9(16), 2930–2937. https://doi.org/10.7150/jca.24425

https://www.ncbi.nlm.nih.gov/books/NBK519578/#:~:text=For%20Stage%201%20disease%2C%205,than%205%25%20survive%20five%20years.

https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6096359/

https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2941142/