ABO Blood Groups and Risk of Lung Cancer: A Retrospective Cohort Study Using Electronic Health Record

Yi-Horng Lai* and Fen-Fen Huang

Department of Healthcare Administration, Asia Eastern University of Science and Technology, New Taipei City, Taiwan
Email: yhlai1108@gmail.com (Y.L.); fl005@mail.aeust.edu.tw (F.F.H.)
*Corresponding author

Manuscript received May 7, 2024; revised June 4, 2024; accepted July 17, 2024; published November 27, 2024

***Abstract*—Detecting early indications of lung cancer is challenging, resulting in delayed identification and advanced stage diagnoses for most patients. Conventional diagnostic approaches for lung cancer face issues such as low accuracy, missed diagnoses, and misinterpretations. Consequently, there is a growing recognition of the crucial role early detection plays in managing lung cancer. This study aimed to explore the correlation between blood type and lung cancer through the analysis of real-world data. The goal is to improve public awareness of the significant risks associated with lung cancer, promote active engagement in relevant screenings, and underscore the importance of early detection, timely treatment, and overall well-being restoration, particularly for individuals with an elevated risk. Utilizing the TriNetX database, the research performed a real-world analysis of data to investigate the correlation between ABO blood groups and the risk of lung cancers. The lung cancer risk ratio between blood types A and AB is 0.934 (95% CI: 0.876–0.996). Blood type A patients have a 12-year cumulative incidence rate of 1.710%, while blood type B patients have a rate of 1.823%. The risk ratio of lung cancer for blood type B vs. AB is 0.918 (95% CI: 0.861–0.979). The 12-year cumulative incidence rate for blood type B patients is 1.683%, and for blood type AB patients, it is 1.823%. There were no statistically significant differences in the risk ratios for all other groups. The findings of this study suggest that individuals with blood type AB face an elevated risk of developing lung cancer in comparison to those with blood type A. Moreover, individuals with blood type AB exhibit a greater susceptibility to lung cancer when contrasted with individuals having blood type B.**

***Keywords*—ABO blood types, respiratory system cancer, lung cancer, real-world data**

I. Introduction

Over the past few years, there has been a consistent rise in the prevalence of lung cancer, establishing it as one of the most widespread forms of cancer worldwide. According to data from the World Health Organization (WHO) [1], in 2019, among the 55.4 million global deaths, the top ten primary causes contributed to 55% of all fatalities. Within these factors, tracheal, bronchus, and lung cancers held the 6th position, demonstrating a discernible upward trajectory. Moreover, among the leading ten contributors to mortality in middle and high-income nations, there has been a substantial surge in the incidence of lung cancer-related fatalities, marking a notable rise of 411,000 deaths. When considering all other income brackets collectively, the overall fatality count has more than two folded.

Detecting the early signs of lung cancer is challenging, making early identification difficult and leading to diagnoses at advanced stages for most patients. Traditional diagnostic methods for lung cancer encounter issues such as low accuracy, missed diagnoses, and misdiagnoses. Consequently, there is an increasing acknowledgment of the critical importance of early diagnosis in lung cancer.

Humanity's understanding of complex and sophisticated physiological mechanisms within the human body has its roots in a multitude of scientific discoveries. In the year 1900, a pivotal advancement occurred in the field of scientific inquiry, thanks to Karl Landsteiner at the University of Vienna. Landsteiner was the scientist who revealed the ABO blood group system during his research pursuits. This discovery sheds light on the reasons behind the diverse outcomes of blood transfusions, explaining why some are successful while others can lead to fatal consequences.

Over the years, numerous research papers have explored the relationship between blood types and various diseases. For instance, individuals with blood type O who contract cholera may display more pronounced symptoms compared to those with non-O blood types, as discussed by Anstee in 2010 [2]. The likelihood of developing metabolic conditions like hyperlipidemia and diabetes varies among individuals with different blood types. Furthermore, extensive research on the association between blood type and cancer risk has been ongoing since 1950, encompassing various cancer types such as gastric cancer, salivary gland cancer, colorectal cancer, ovarian cancer, uterine corpus cancer, and cervical cancer, as highlighted by Garratty in 2000 [3].

In recent years, the application of big data analytics in the medical field, particularly in the realm of intelligent healthcare, has gained significant attention [4]. Big data analytics entails the exploration of vast raw data to discern trends, patterns, and correlations, enabling informed decision-making [5]. By employing well-established statistical analysis techniques and leveraging modern tools to analyze larger datasets, intelligent healthcare spans various areas, including the prediction, diagnosis, and treatment of conditions such as lung cancer.

Iodice *et al.* gathered data from the tumor registry, encompassing 15,359 cancer patients treated at the European Institute of Oncology between 2000 and 2003, all of whom had specified ABO blood types. Employing a case-control analysis, they juxtaposed the distribution of ABO blood types among patients with different cancer types against those with other cancer forms. They observed a significant reduction in the prevalence of blood type O among individuals with exocrine pancreatic cancer compared to those with other cancer types [6]. Concurrently, findings from The Nurses' Health Study revealed that among 316 participants, individuals with blood groups A, AB, or B exhibited a heightened propensity for pancreatic cancer compared to those with blood type O [7]. In another study involving 820

participants from a Taiwan cohort, where blood types were determined in a laboratory environment, it was identified that individuals with blood types A, B, and AB faced an elevated risk of pancreatic cancer relative to those with blood type O [8].

Based on research conducted by Edgren *et al.*, which examined 688 cases of gastric cancer in the SCANDAT database, there seems to be an increased likelihood of gastric cancer among individuals with blood type A [9]. Similarly, a study involving 524 participants from a Taiwanese cohort, where blood types were determined in the laboratory, discovered that those with blood type A exhibited a greater susceptibility to stomach cancer compared to individuals with blood type O [10].

Urun *et al.* investigated 2,044 individuals diagnosed with lung cancer in Turkey and discovered a correlation between non-O blood type and increased lung cancer risk [11]. Similarly, Eren *et al.* studied 587 lung cancer patients at a public hospital, noting a higher prevalence of blood type B among lung cancer patients compared to non-lung cancer patients [12]. Additionally, Utkan *et al.* analyzed ABO distributions among 1,954 patients in Turkey, revealing a significant association between ABO blood type and lung cancer. Specifically, individuals with blood types other than O were found to have an elevated risk of developing lung cancer [13].

Sun *et al.* carried out a study encompassing 1065 participants from a Taiwan cohort, whose blood types underwent determination via laboratory tests. The results uncovered no noteworthy association between blood type and the susceptibility to lung cancer [10]. Furthermore, a retrospective case series analysis involving 500 lung cancer patients utilized hospital case records from Nanakali Hospital and Rizgary Teaching Hospital. The study's outcomes concluded that there is no statistically significant correlation between ABO blood groups and the overall susceptibility to lung cancer, nor across various histopathological types [14]. Similarly, data from a cohort of 964 male lung cancer patients in Shanghai, China, observed over a 25-year span, indicated no significant link between blood types and the risk of developing lung cancer [15].

The aim of this study was to explore the correlation between blood type and lung cancer using real-world data. The goal is to raise public awareness about the key risks linked to lung cancer, promote active engagement in relevant screenings, and advocate for early detection, timely treatment, and overall health restoration, particularly for individuals at a heightened risk.

II. Materials and Research Method

A. Research Data

Established in 2013, TriNetX is a global health research network that provides de-identified Electronic Health Records (EHR) sourced from over 120 major healthcare institutions, covering a vast patient pool of over 70 billion individuals [16]. By consolidating real-world EHR data from various regions and facilitating the exchange of de-identified information, TriNetX expedites the generation of real-world evidence. This methodology not only streamlines the research process but also contributes to cost savings and time efficiency in the field of health science. The participating healthcare institutions are predominantly situated in the United States and Europe, with some representation from Asia and Australia. Most of these institutions are either academic medical centers or research-oriented non-academic medical organizations.

Via TriNetX, researchers have the capability to investigate a de-identified database in order to pinpoint patients who meet particular criteria. This encompasses details like inpatient and outpatient diagnosis codes, treatment procedures, medication usage, and laboratory data. Notably, it omits information gathered by alternative clinical research initiatives. The data is represented in terms of unique patient counts, ensuring each patient is only tallied once and excluding those with solely medical record numbers or lacking diagnoses or codes. This aids researchers in verifying the adequacy of potential patients for their study.

TriNetX obtains its data from EHR offered by healthcare institutions, encompassing both (1) organized data and (2) unstructured data that undergoes processing through natural language processing techniques. Additionally, contributions are made by (1) cancer registry data and (2) genomic information sourced from third-party genetic testing laboratories.

This research employed a retrospective observational approach, utilizing data sourced from the TriNetX US Collaborative Network, encompassing information from 83 healthcare institutions. The study spanned from January 2010 to March 2023 and employed the International Statistical Classification of Diseases and Related Health Problems, 10th Revision (ICD-10). The investigation specifically concentrated on major cancer ICD-10 diagnosis codes, namely C33 and C34.

As TriNetX is a de-identified database devoid of personally identifiable patient information, the currently published TriNetX articles have received Institutional Review Board (IRB) exemption from review requirements in multiple locations.

B. Data Analysis

Traditional randomized controlled trials frequently utilize precise inclusion and exclusion criteria to protect participants and reduce patient variability, thereby ensuring the consistency of trial outcomes. To minimize bias, experimental methodologies include strategies such as random allocation, control groups, and blind designs. The implementation procedures are carried out meticulously to uphold the quality of results and improve the overall reliability of research findings. The adoption of this research design is widespread in medical research.

However, it may easily result in a discrepancy between real-world situations and the conditions of clinical trials, potentially affecting the applicability of research findings to practical situations. An increasing number of studies endorse leveraging real-world data for big data analysis [17]. The advantage of conducting big data research with real-world data is its intrinsic relevance to real-world situations.

This research employs a generational comparative analysis to explore the association between ABO blood types and lung cancer (ICD-10: C33, C34), utilizing real-world data sourced from TriNetX. The objective is to discern the correlation

between blood types and the incidence of lung cancer.

The research cohort was chosen between January 2010 and March 2023. Employing the TriNetX, a comparative analysis of A, B, O, and AB blood types was performed for a generational control study focusing on primary lung cancer. Following propensity score matching for age, gender, and race, individual generational control studies assessed the occurrences of primary lung cancer cases, cumulative incidence rates at 3, 6, 9, and 12 years, along with the risk ratio.

III. Results

A. Descriptive Statistics

In this study, a total of 117,680,314 cases were selected from the TriNetX US Collaborative Network database. As shown in Table 1 and Fig. 1. There was a total of 1,012,452 individuals with blood type A, including 387,113 males (38%) and 623,768 females (62%). The average age was 49 ± 21.4 years, with 75% being White, 13% Black or African American, and 2% Asian. There was a total of 346,988 individuals with blood type B, including 126,762 males (37%) and 219,760 females (63%). The average age was 47 ± 21.2 years, with 55% being White, 28% Black or African American, and 5% Asian. There was a total of 104,334 individuals with blood type AB, including 39,003 males (37%) and 65,147 females (63%). The average age was 48.3 ± 21.5 years, with 66% being White, 20% Black or African American, and 5% Asian. There was a total of 1,255,002 individuals with blood type O, including 466,521 males (37%) and 786,658 females (63%). The average age was 48.3 ± 21.5 years, with 65% being White, 19% Black or African American, and 2% Asian.

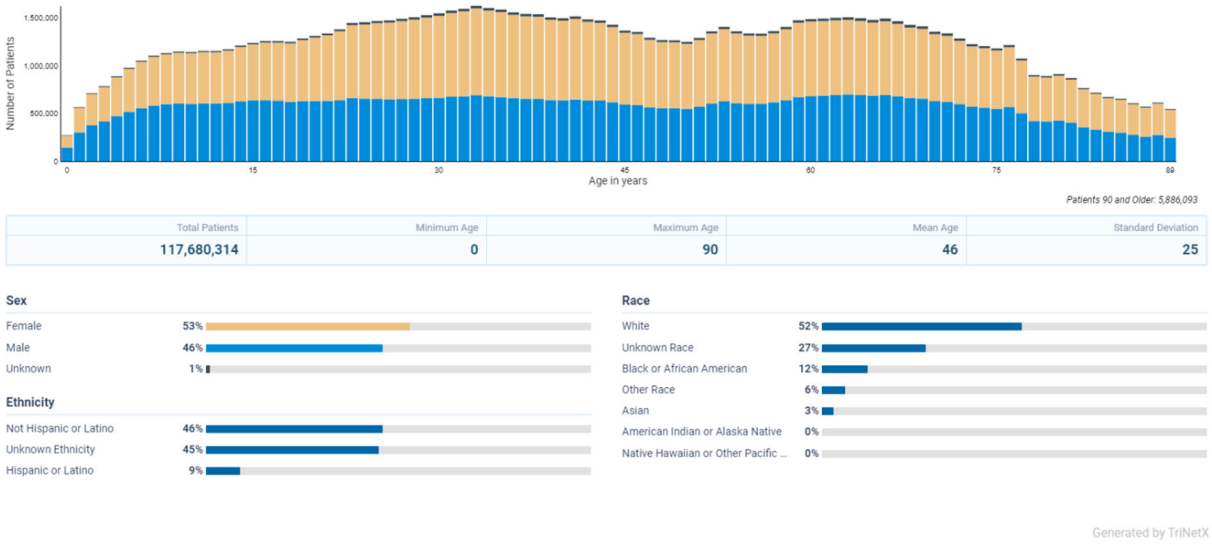

Fig. 1. Demographics of research data in this study.

Table 1. Data characteristics

Blood Type	Type A	Type B	Type AB	Type O
Patients	1,012,452	346,988	104,334	1,255,002
Age at Index	49 ± 21.4	47 ± 21.2	48.3 ± 21.5	47.5 ± 21.4
Sex				
Female	623,768 (62%)	219,760 (63%)	65,147 (63%)	786,658 (63%)
Male	387,113 (38%)	126,762 (37%)	39,003 (37%)	466,521 (37%)
Unknown	1,571 (0%)	466 (0%)	184 (0%)	1,823 (0%)
Race (Patients / % of Cohort)				
White	762,953 (75%)	189,690 (55%)	68,511 (66%)	814,118 (65%)
Black or African American	126,210 (13%)	98,408 (28%)	20,900 (20%)	241,493 (19%)
Unknown Race	61,953 (6%)	23,696 (7%)	6,550 (6%)	91,143 (7%)
Other Race	41,029 (4%)	16,626 (5%)	3,416 (3%)	78,098 (6%)
Asian	15,282 (2%)	16,840 (5%)	4,495 (5%)	22,035 (2%)
American Indian or Alaska Native	3,565 (0%)	1,051 (0%)	282 (0%)	6,006 (1%)
Native Hawaiian or Other Pacific Islander	1,460 (0%)	677 (0%)	180 (0%)	2,109 (0%)
ICD-10 Diagnoses （Patients/ % of Cohort）				
Factors influencing health status and contact with health services (Z00-Z99)	130,572 (62%)	96,327 (63%)	60,353 (62%)	144,630 (62%)

Symptoms, signs and abnormal clinical and laboratory findings, not elsewhere classified (R00-R99)	110,114 (52%)	84,867 (55%)	53,437 (55%)	120,587 (52%)
Diseases of the musculoskeletal system and connective tissue. (M00-M99)	74,064 (35%)	57,967 (39%)	36,669 (38%)	76,284 (34%)
Endocrine, nutritional and metabolic diseases (E00-E89)	74,504 (35%)	57,686 (37%)	37,461 (38%)	78,360 (34%)
Diseases of the circulatory system (I00-I99)	69,925 (33%)	54,487 (35%)	36,367 (37%)	71,391 (31%)

Table 2. Degree of discreteness of blood type groups

Blood Type	A vs B	A vs AB	A vs O	B vs AB	B vs O	AB vs O
Age at Index	0.0936*	0.0330*	0.0709*	0.0602*	0.0224*	0.0377*
Sex						
Female	0.0357*	0.0171*	0.0221*	0.0185*	0.0135*	0.0050
Unknown Gender	0.0055*	0.0052*	0.0026	0.0107*	0.0029	0.0078
Male	0.0353*	0.0176*	0.0219*	0.0177*	0.0133*	0.0043*
Race						
White	0.4442*	0.2137*	0.2305*	0.2260*	0.2091*	0.0167*
Unknown Race	0.0286*	0.0065*	0.1863*	0.0220*	0.0171*	0.0391*
Black or African American	0.4023*	0.2062*	0.0456*	0.1954*	0.2154*	0.0199*
Other Race	0.0360*	0.0414*	0.0984*	0.0772*	0.0628*	0.1390*
Asian	0.1914*	0.1671*	0.0060*	0.0261*	0.1739	0.1493*
American Indian or Alaska Native	0.0086*	0.0147	0.0197*	0.0061	0.0282	0.0341*
Native Hawaiian or Other Pacific Islander	0.0124*	0.0071*	0.0194*	0.0053	0.0064	0.0011*

Table 3. Distribution of characteristics among groups after marching (A vs B, A vs AB, A vs O)

Group	1		2		3	
	A	B	A	AB	A	O
Patients	340,521	340,521	102,737	102,737	996,296	996,296
Age at Index	47.2 ± 21.2	47.2 ± 21.2	48.5 ± 21.5	48.5 ± 21.5	49.2 ± 21.4	49.2 ± 21.4
Gender						
Female	214,974 (63.13%)	214,974 (63.12%)	63,944 (62.24%)	63,942 (62.23%)	611,505 (61.37%)	611,617 (61.38%)
Male	125,117 (36.74%)	125,126 (36.74%)	38,612 (37.58%)	38,614 (37.58%)	383,263 (38.46%)	383,168 (38.45%)
Unknown Gender	430 (0.12%)	454 (0.13%)	181 (0.17%)	181 (0.17%)	1,528 (0.15%)	1,511 (0.15%)
Race						
White	187,950 (55.19%)	187,950 (55.19%)	67,871 (66.06%)	67,871 (66.06%)	756,322 (75.91%)	756,323 (75.91%)
Black or African American	98,334 (28.87%)	98,334 (28.87%)	20,880 (20.32%)	20,880 (20.32%)	126,069 (12.65%)	126,061 (12.65%)
Unknown Race	20,977 (6.16%)	20,984 (6.16%)	5,712 (5.56%)	5,712 (5.56%)	53,626 (5.38%)	53,681 (5.38%)
Other Race	16,344 (4.80%)	16,337 (4.79%)	3,331 (3.24%)	3,331 (3.24%)	40,068 (4.02%)	40,065 (4.02%)
Asian	15,188 (4.46%)	15,188 (4.46%)	4,481 (4.36%)	4,481 (4.36%)	15,188 (1.52%)	15,218 (1.52%)
American Indian or Alaska Native	1,051 (0.30%)	1,051 (0.30%)	282 (0.27%)	282 (0.27%)	3,565 (0.35%)	3,566 (0.35%)
Native Hawaiian or Other Pacific Islander	677 (0.19%)	677 (0.19%)	180 (0.17%)	180 (0.17%)	1,458 (0.14%)	1,382 (0.13%)

The top five ICD-10 diagnoses in the day preceding the initial blood type data acquisition were: (1) Factors influencing health status and contact with health services (Z00-Z99); (2) Symptoms, signs and abnormal clinical and laboratory findings, not elsewhere classified (R00-R99); (3) Diseases of the musculoskeletal system and connective tissue (M00-M99); (4) Endocrine, nutritional and metabolic diseases (E00-E89); (5) Diseases of the circulatory system (I00-I99).

By comparing database systems, Table 2 illustrates the extent of dispersion of fundamental data across various blood types. The standard deviations for age, gender, white, black or African American, and Asian populations exceed 0.1 for both A-type and B-type blood patients, suggesting statistically significant differences between the two groups in terms of age, gender, and ethnicity. Similarly, significant differences are observed between type A and type AB, type A and type O, type B and type AB, type B and type O, as well as type AB and blood type O patients, based on the standard deviations for age, gender, white, black or African American, and Asian populations, all surpassing 0.1.

B. Propensity Score Matching

Propensity score matching the four blood types in pairs,

creating a total of six groups. Subsequently, comparing them through age, gender, and ethnicity matching. Detailed information is provided in Tables 3 and 4.

Following the pairing groups of patients with blood type A vs. B (Group 1), the total cases for each blood type amount to 340,521 individuals. The recorded age falls within the range of 47.2 ± 21.2 years, with no statistically significant distinctions observed between the two blood type patient cohorts. Gender distribution ratios similarly reveal no statistically significant differences. Additionally, there are no significant variations in ethnicity proportions among white, black or African American, and Asian populations.

Upon pairing patients with blood type A vs. AB (Group 2), the total cases for each blood type are 102,737 people. The registered age is 48.5 ± 21.5 years, and there are no statistically significant differences between the two blood type patient groups. Gender distribution ratios exhibit no statistically significant distinctions, and ethnicity proportions among white, black or African American, and Asian populations remain unchanged.

In the case of patients paired with blood type A vs. O (Group 3), the combined cases for each blood type reach 996,296 individuals. The registered age is 49.2 ± 21.4 years, and there are no statistically significant differences between the two blood type patient groups. Gender distribution ratios do not show statistically significant variations, and there are no significant differences in ethnicity proportions among white, black or African American, and Asian populations.

Table 4. Distribution of characteristics among groups after marching (B vs AB, B vs O, AB vs O)

Group	4		5		6	
	B	AB	B	O	AB	O
Patients	102,737	102,737	342,130	342,130	102,737	102,737
Age at Index	48.5 ± 21.5	48.5 ± 21.5	47.2 ± 21.2	47.2 ± 21.2	48.5 ± 21.5	48.5 ± 21.5
Gender						
Female	63,951 (62.24%)	63,942 (62.23%)	216,127 (63.17%)	216,131 (63.17%)	63,942 (62.23%)	63,943 (62.24%)
Male	38,605 (37.57%)	38,614 (37.58%)	125,548 (36.69%)	125,549 (36.69%)	38,614 (37.58%)	38,613 (37.584%)
Unknown Gender	181 (0.17%)	181 (0.17%)	455 (0.13%)	450 (0.13%)	181 (0.17%)	181 (0.17%)
Race						
White	67,871 (66.06%)	67,871 (66.06%)	187,950 (54.93%)	187,949 (54.93%)	67,871 (66.06%)	67,873 (66.06%)
Black or African American	20,880 (20.32%)	20,880 (20.32%)	98,334 (28.74%)	98,334 (28.74%)	20,880 (20.32%)	20,881 (20.32%)
Unknown Race	5,714 (5.56%)	5,712 (5.56%)	20,984 (6.13%)	20,981 (6.13%)	5,712 (5.56%)	5,712 (5.56%)
Other Race	3,331 (3.24%)	3,331 (3.24%)	16,337 (4.77%)	16,337 (4.77%)	3,331 (3.24%)	3,331 (3.24%)
Asian	4,482 (4.36%)	4,481 (4.36%)	16,797 (4.91%)	16,797 (4.91%)	4,481 (4.36%)	4,481 (4.36%)
American Indian or Alaska Native	281 (0.27%)	282 (0.27%)	1,051 (0.30%)	1,051 (0.30%)	282 (0.27%)	282 (0.27%)
Native Hawaiian or Other Pacific Islander	178 (0.17%)	180 (0.175)	677 (0.19%)	681 (0.19%)	180 (0.17%)	177 (0.17%)

Subsequently, pairing patients with blood type B vs. AB (Group 4) results in 102,737 cases for each blood type. The registered age remains at 48.5 ± 21.5 years, with no statistically significant differences between the two blood type patient groups. Gender distribution ratios reveal no statistically significant distinctions, and there are no significant differences in ethnicity proportions among white, black or African American, and Asian populations.

For patients paired with blood types B vs. O (Group 5), the total cases for each blood type equal 342,130 individuals. The registered age is 47.2 ± 21.2 years, and there are no statistically significant differences between the two blood type patient groups. Gender distribution ratios show no statistically significant variations, and there are no significant differences in ethnicity proportions among white, black or African American, and Asian populations.

Finally, after pairing patients with blood type AB vs. O (Group 6), the combined cases for each blood type reach 102,737 people. The registered age is 48.5 ± 21.5 years, and there are no statistically significant differences between the two blood type patient groups. Gender distribution ratios do not exhibit statistically significant distinctions, and there are no significant differences in ethnicity proportions among white, black or African American, and Asian populations.

C. Blood Type and Lung Cancer Risk

The comparative results of lung cancer risk among various blood type pairing groups, stratified by age, gender, and ethnicity, are presented in Table 5.

In comparison with Group 2, the risk ratio of blood type A to blood type AB is 0.934 (95% CI: 0.876 ~ 0.996). The cumulative incidence rate over 12 years for blood type A patients is 1.710%, while for blood type AB patients, it is 1.823%. In comparison with Group 4, the risk ratio of blood type B to blood type AB is 0.918 (95% CI: 0.861 ~ 0.979). The 12-year cumulative incidence rate for blood type B patients is 1.683%, and for blood type AB patients, it is 1.823%. The risk ratios for all other groups did not reach statistically significant differences.

In terms of different gender patients, the comparative results of lung cancer risk among various blood types after age and race pairing are shown in Table 6. There are 364,695 male and 364,695 female patients in the A blood type group. In the B blood type group, there are 120,406 male and 120,406 female patients. The AB blood type group consists of 37,021 male and 37,021 female patients. In the O blood

type group, there are 441,367 male and 441,367 female patients. Across different blood type groups, there is no statistically significant difference in the risk of developing lung cancer between different genders.

Table 5. The risk of lung cancer among different blood types

	1		2		3		4		5		6	
	A	B	A	AB	A	O	B	AB	B	O	AB	O
Patients	341,092	341,092	102,877	102,877	997,562	997,562	102,877	102,877	342,721	342,721	102,877	102,877
Patients with Outcome	5,489	5,478	1,776	1,901	18,022	18,331	1,745	1,901	5,494	5,603	1,901	1,828
Cumulative Incidence												
At the end of the 3rd year	1.382%	1.383%	1.474%	1.577%	1.553%	1.575%	1.474%	1.577%	1.380%	1.394%	1.577%	1.531%
At the end of the 6th year	1.513%	1.515%	1.617%	1.732%	1.700%	1.725%	1.602%	1.732%	1.512%	1.531%	1.732%	1.671%
At the end of the 9th year	1.563%	1.570%	1.681%	1.792%	1.759%	1.790%	1.655%	1.792%	1.567%	1.589%	1.792%	1.731%
At the end of the 12th year	1.590%	1.592%	1.710%	1.823%	1.789%	1.816%	1.683%	1.823%	1.589%	1.589%	1.823%	1.753%
Risk Ratio	1.002 (0.965, 1.04)		0.934 (0.876, 0.996)		0.983 (0.963, 1.003		0.918 (0.861, 0.979)		0.981 (0.945, 1.017)		1.040 (0.976, 1.108)	

Table 6. The risk of lung cancer among different blood types and genders

Blood Type	A		B		AB		O	
Gender	Male	Female	Male	Female	Male	Female	Male	Female
Patients	364,695	364,695	120,406	120,406	37,021	37,021	441,367	441,367
Patients with Outcome	8,428	8,530	2,560	2,609	897	888	9,909	9,860
Cumulative Incidence								
At the end of the 3rd year	1.985%	2.030%	1.837%	1.895%	2.064%	2.096%	1.92%	1.952%
At the end of the 6th year	2.175%	2.211%	2.016%	2.066%	2.296%	2.277%	2.109%	2.124%
At the end of the 9th year	2.245%	2.289%	2.089%	2.138%	2.372%	2.345%	2.188%	2.193%
At the end of the 12th year	2.284%	2.323%	2.117%	2.165%	2.404%	2.382%	2.218%	2.223%
Risk Ratio	0.988 (0.959, 1.018)		0.981 (0.93, 1.036)		1.01 (0.922, 1.107)		1.005 (0.978, 1.033	

IV. Discussion and Conclusion

A. Discussion

This study utilizes the TriNetX system for data collection and analysis to investigate differences in the risk of developing lung cancer among individuals with different blood types. The research data in this study reveal notable differences in age, gender, and racial proportions among patients with different blood types. Therefore, following age, gender, and race matching, risk analyses for lung cancer are conducted within each matched subgroup.

In generational studies involving age, gender, and race pairing, the risk ratio comparing individuals with blood type A to those with blood type AB is 0.934. The cumulative incidence rate over 12 years for individuals with blood type A is 1.710%, while for individuals with blood type B, it is 1.823%. The risk ratio comparing individuals with blood type B to those with blood type AB is 0.918. The cumulative incidence rate over 12 years for individuals with blood type B is 1.683%, while for individuals with blood type AB, it is 1.823%.

The results of this study indicate that individuals with blood type AB have a higher risk of developing lung cancer compared to those with blood type A. Additionally, individuals with blood type AB have a higher risk of lung cancer compared to those with blood type B. The result of this study is similar to Urun's study [11]. Urun's research revealed a correlation between ABO blood types and the incidence of lung cancer, indicating that individuals with a non-O blood type face an elevated risk of developing lung cancer.

B. Conclusion

The arrangement of A, B, and H antigens on cell surfaces has the potential to impact tumor expression. As these antigens are implicated in tumor development, metastasis, and prognosis, they may contribute to processes like cell recognition, signal transduction, and adhesion [18]. The risk of cancer development varies across individuals with distinct ABO blood types. The crucial factor is the manifestation of surface A, B, and H antigens on cells throughout different bodily tissues, which is associated with the ABO genotype. Consequently, this display influences susceptibility to cancer and the probability of tumor cell metastasis [19].

In cross-generational studies based on age, gender, and racial pairing, individuals with blood type A have a lower risk of developing lung cancer compared to those with blood type AB, with a risk ratio of 0.934. Individuals with blood type B have a lower risk of developing lung cancer compared to

those with blood type AB, with a risk ratio of 0.918. There is no statistically significant difference in the risk of developing lung cancer among individuals with other blood types. Therefore, it is recommended that individuals with blood type AB pay attention to the risk of developing lung cancer and undergo regular check-ups at the hospital.

Conflict of Interest

The author declare that he has no competing interests.

Author Contributions

The author of this manuscript is Yi-Horng Lai and Fen-Fen Huang. The author has made substantial contributions to the conception and design, acquisition, analysis and interpretation of data, and he was involved in drafting the manuscript. The author read and approved the final manuscript.

Funding

Publication costs are not funded by any organization.

References

[1] World Health Organization. (2020). The top 10 causes of death. [Online]. Available: https://www.who.int/news-room/fact-sheets/detail/the-top-10-causes-of-death

[2] D. J. Anstee, "The relationship between blood groups and disease," *Blood*, vol. 115, no. 23, pp. 4635–4643, 2010. https://doi.org/10.1182/blood-2010-01-261859

[3] G. Garratty, "Blood groups and disease: A historical perspective," *Transfusion Medicine Reviews*, vol. 14, no. 4, pp. 291–301. 2000. https://doi.org/10.1053/tmrv.2000.16228

[4] Y. Inoue, Y. Takekuma, T. Miyai, H. Kashiwagi, Y. Sato, M. Sugawara, and S. Imai, "Use of Japanese big data from electronic medical records to investigate risk factors and identify their high-risk combinations for linezolid-induced thrombocytopenia," *European Journal of Clinical Pharmacology*, vol. 79, no. 3, pp. 415–425, 2023. https://doi.org/10.1007/s00228-023-03455-x

[5] R. Jiang, S. Han, Y. Yu, and W. Ding, "An access control model for medical big data based on clustering and risk," *Information Sciences*, vol. 621, pp. 691–707, 2023. https://doi.org/10.1016/j.ins.2022.11.102

[6] S. Iodice, P. Maisonneuve, E. Botteri, M. T. Sandri, and A. B. Lowenfels, "ABO blood group and cancer," *European Journal of Cancer*, vol. 46, no. 18, pp. 3345–3350, 2010. https://doi.org/10.1016/j.ejca.2010.08.009

[7] B. M. Wolpin, A. T. Chan, P. Hartge, S. J. Chanock, P. Kraft, D. J. Hunter, E. L. Giovannucci, and C. S. Fuchs, "ABO blood group and the risk of pancreatic cancer," *JNCI: Journal of the National Cancer Institute*, vol. 101, no. 6, pp. 424–431, 2009. https://doi.org/10.1093/jnci/djp020

[8] W. Sun, C. Wen, and J. Lin *et al.*, "ABO blood types and cancer risk-A cohort study of 339,432 subjects in Taiwan," *Cancer Epidemiology*, vol. 39, no. 2, pp. 150–156, 2015. https://doi.org/10.1016/j.canep.2014.12.006

[9] G. Edgren, H. Hjalgrim, K. Rostgaard, R. Norda, A. Wikman, M. Melbye, and O. Nyrén, "Risk of gastric cancer and peptic ulcers in relation to ABO blood type: A cohort study," *American Journal of Epidemiology*, vol. 172, no. 11, pp. 1280–1285, 2010. https://doi.org/10.1093/aje/kwq299

[10] W. Sun, C. Wen, and J. Lin *et al.*, "ABO blood types and cancer risk-A cohort study of 339,432 subjects in Taiwan," *Cancer Epidemiology*, vol. 39, no. 2, pp. 150–156, 2015. https://doi.org/10.1016/j.canep.2014.12.006

[11] Y. Urun, G. Utkan, and A. K. Cangir *et al.*, "Association of ABO blood group and risk of lung cancer in a multicenter study in Turkey," *Asian Pacific Journal of Cancer Prevention*, vol. 14, no. 5, pp. 2801–2803, 2013. https://doi.org/10.7314/APJCP.2013.14.5.2801

[12] C. Eren, A. G. Akbıyık, and İ. N. Ökten, "Analysis of the relationship between blood groups and lung cancer risk depending on histology," Düzce Üniversitesi Sağlık Bilimleri Enstitüsü Dergisi, vol. 11, no. 3, pp. 386–390, 2021. https://doi.org/10.33631/duzcesbed.867713

[13] G. Utkan, Y. Urun, A. K. Cangir *et al.*, "ABO blood group and the risk of lung cancer: Multicenter, case-control, observational study," *Journal of Clinical Oncology*, vol. 30, no. 15, pp. 1596–1596, 2012. https://doi.org/10.1200/jco.2012.30.15_suppl.1596

[14] Z. S. Hussein, "Relationship between the ABO blood group and lung cancer susceptibility," *Medical Journal of Babylon*, vol. 18, no. 2, pp. 80–82, 2021. https://www.iasj.net/iasj/article/207516

[15] J.Y.Huang, R. Wang, Y. T. Gao, and J. M. Yuan, "ABO blood type and the risk of cancer—Findings from the Shanghai Cohort Study," *Plos One*, vol. 12, no. 9, e0184295, 2017. https://doi.org/10.1371/journal.pone.0184295

[16] TriNetX. (2023). LLC, About TRINETX, LLC. [Online]. Available: https://trinetx.com/

[17] Y. Tanaka, S. Tanaka, T. Takahashi, and N. Kato, "Clinical features of polymyalgia rheumatica patients in Japan: Analysis of real-world data from 2015 to 2020," *Modern Rheumatology*, vol. 34, no. 1, pp. 201–207, 2024. https://doi.org/10.1093/mr/road026

[18] P. A. Edwards, "Heterogeneous expression of cell-surface antigens in normal epithelia and their tumours, revealed by monoclonal antibodies," *British Journal of Cancer*, vol. 51, no. 2, pp. 149–160, 1985. https://doi.org/10.1038%2Fbjc.1985.24

[19] M. Franchini, E. J. Favaloro, G. Targher, and G. Lippi, "ABO blood group, hypercoagulability, and cardiovascular and cancer risk," *Critical Reviews in Clinical Laboratory Sciences*, vol. 49, no. 4, pp. 137–149, 2012. https://doi.org/10.3109/10408363.2012.708647