How Spatial Factors Affect Respiratory Health in Coal Resource-based Cities: a Case-control Study in Eastern Yunnan Province, China

Yuxing Zhao & Lan Wang

Abstract:
Urban planning can affect respiratory health. Due to industrial and mining pollution, unique urban and rural spatial structure, etc., the built environment factors of coal resource-based towns have special impact paths on respiratory health. Lung cancer is one of the cancers with the highest incidence and mortality rates worldwide. Qujing City, located in the eastern part of Yunnan Province, China, is dominated by the coal mining industry. Since the survey in the 1970s, it has been the area with the highest incidence of lung cancer in China. Compared with the national average, the standardized mortality rate is 3 times that of men and 5 times that of women. Existing studies are mostly centered on the field of epidemiology, among which there are few studies on spatial influencing factors and the research indicators are single. This study aims to explore the association between environmental factors and the risk of lung cancer in residents in coal resource-based urban and rural areas at different stages of the life course, and propose corresponding intervention measures.

The study adopted a case-control study based on life course theory. Lung cancer patients and healthy controls were interviewed in two local hospitals, and individual-level influencing factors that could not be obtained through external data were obtained through questionnaires. Among them, non-spatial factors such as smoking history were confounding factors and were controlled; spatial influencing factors such as migration history and living environment history were the core content of the study. On this basis, Landsat remote sensing images were identified through supervised classification methods in ENVI5.6, and pollution sources such as mining land and industrial land were extracted. Frequency matching was established according to gender and age (±2 years old), and 51 cases in each case group and control group were obtained. The exposure ratio in the case group and the control group was compared through spatial statistics and the use of non-conditional Logistic regression model.

The study found that the case group and the control group showed obvious urban-rural differences in space. Compared with the control group, the hotspots were mainly concentrated in rural areas, and the case clustering was strongest when the analysis radius was 2km. After controlling for factors such as fuel use and housing type, long-term residence within 2km of a coal mine was still a significant risk factor after adjustment (OR: 2.433; 95%CI: 0.951-6.222). The living environment has a heterogeneous effect on the incidence of lung cancer at different stages of the life course. After adjustment, living in earthen houses in early adulthood (25-44 years old) (OR: 2.891; 95%CI: 1.144-7.311), living in brick and wood houses in middle-aged and elderly (45 years old and above) is a risk factor in univariate analysis (OR: 2.759; 95%CI: 1.159-6.567), while living in concrete houses during the development period (0-24 years old) (OR: 0.108; 95%CI: 0.013-0.894) and early adulthood (OR: 0.292; 95%CI: 0.103-0.823) may have a protective effect. Although cooking in the kitchen and the main room, and the type of stove, were not included in the final model, the univariate analysis still indicated the health risks of poor ventilation and indoor pollution. The above research provides a basis for promoting the implementation of health equity planning at the micro and medium levels.

Keywords: Resource-based cities; Lung cancer; Case-control study; Residence history; Life Course; Health equity

1. Introduction

Urban and rural space is the main carrier of human activities, directly or indirectly affecting people's physical and mental health through complex paths. "Healthy city planning" is based on reducing health risks, optimizing health resources, and promoting healthy behaviors. It regulates and shapes space. It is an important policy tool for achieving health promotion and disease prevention and control, and plays an important role in "preventing disease" at the population level (Wang L, 2023). Respiratory health is an important research area in the field of healthy cities. A large number of scholars have explored the ways and mechanisms to promote respiratory health around four types of urban space elements: land use and facilities, development intensity and form, road traffic system, green space and open space. However, compared with other types of towns, the impact path of built environment factors on respiratory health in coal resource-based towns is special: for example, industrial and mining pollution has a stronger explanatory power for respiratory health results (Wei H, et al., 2022), and lower aggregation levels and more small and medium-sized settlements may also have some impact on the action path (Lu S, et al., 2020).

Qujing City, Yunnan Province, located in the eastern Yunnan region of China, is extremely rich in coal resources and has formed a coal industry chain dominated by coal-fired power, coking, and chemical industry. Various pollutants generated by widely distributed coal mines, coal washing plants, coking plants, etc., bring long-term health risks to local residents. Eastern Yunnan is a high-incidence area of lung cancer in China. A retrospective survey of causes of death in the 1970s showed that the lung cancer mortality rate in Xuanwei was 4.1 times and 7.9 times that of the national male and female lung cancer mortality rates during the same period. At present, compared with the national average, the standardized mortality rate is 3 times that of men and 5 times that of women. Since the 1970s, a large number of studies from an epidemiological perspective have shown that factors such as coal type, frequency and amount of use (Barone-Adesi, et al., 2012), stove type (Hu W, et al., 2014), and years of work in coal mines (Hosgood III HD, et al., 2012) are related to the incidence of lung cancer. However, in existing studies, the number of studies on spatial influencing factors is small and the research indicators are single. This study aims to explore the spatial influencing factors of lung cancer in resource-based towns in eastern Yunnan and propose corresponding spatial intervention measures.

2. Materials and methods

2.1.Study Area

This study takes Fuyuan County and Xuanwei City under the jurisdiction of Qujing City, Yunnan Province as the research objects. Fuyuan County has a total area of 3,348 square kilometers and a total population of 840,000; Xuanwei City has a total area of 6,070 square kilometers and a total population of 1.19 million. The two places are located in the eastern part of Yunnan Province, which is a slope area where the Yunnan Plateau transitions to the Guizhou Plateau and belongs to the Wumeng Mountain System. Most areas have relatively straight peak lines and relatively gentle mountain tops, which are shallowly cut mountains with small terrain slopes and small relative height differences. The local plains between the mountains are called "Bazi", which are usually economic centers with prosperous agriculture and dense population.

Fuyuan and Xuanwei both belong to the coal-bearing belt of the eastern Yunnan coalfield, which is distributed to the east of the Niu Shoushan uplift. The coal-bearing strata are consistent with the tectonic axis of the earth, and the coal types are distributed in strips from southwest to northeast. The coal in the two places is shallowly buried and easy to mine. Since the late Ming Dynasty and early Qing Dynasty, people have mined coal for cooking and heating in the two places.

Through field research, it was found that there are a large number of pollution sources such as coal mines and thermal power plants adjacent to towns, coal is easily available, local residents burn coal indoors, and pollution diffusion conditions are poor, which poses health risks.

2.2. Research Framework

This study adopts the research framework of epidemiology. First, the current situation study in descriptive epidemiology is used to describe the distribution of diseases or health and other factors affecting the incidence of lung cancer, which can reveal the phenomenon and provide clues for etiology research. In this study, the average nearest neighbor, nearest neighbor hierarchical cluster analysis, and multi-distance spatial cluster analysis were used. Subsequently, the case-control study in analytical epidemiology was used to test the above hypothesis and explore the impact of built environment exposure at different stages on the incidence of lung cancer from the perspective of life course. In this study, the logistic regression analysis method was used.

2.3. Data Collection

2.3.1.Health Outcomes

The cases were new lung cancer cases from two local hospitals in Qujing in the past five years, and the controls were patients from departments of the two hospitals that were not related to respiratory diseases (such as orthopedics, etc.). Through face-to-face interviews and filling out questionnaires, individual-level confounding factors and spatial influencing factors that could not be obtained through external data were obtained. 162 questionnaires were obtained from April 7 to August 2024. After excluding samples that had lived outside the research scope for a long time, whose addresses were not accurately located, and whose some indicators were missing, a total of 138 samples were obtained, including 72 cases and 66 controls. A database was established in EpiData3.1, and frequency matching was performed according to gender and age (plus or minus 2 years old), and 51 cases and 51 controls were obtained. Oral informed consent was obtained from all study subjects.

2.3.2.Spatial Factors

Spatial influencing factors focus on the meso-micro scale. The meso-level includes spatial relationships (the distance between the residence and the pollution source), wind environment (wind direction), and topography (relatively flat areas, surface undulation). The micro-level includes residential history (house structure, ventilation conditions), stove type, and cooking location.

The address data of the samples came from the questionnaire. The urban samples were accurate to the road intersection, and the rural samples were accurate to the natural village. The spatial location was performed in ArcGIS Pro. The pollution sources include mining land and industrial land. Using the remote sensing images of the Landsat-8 satellite in December 2019, the results of the third land survey in Fuyuan County were used as samples, and the maximum likelihood method was used for supervised classification. Mining land and industrial land were identified in ENVI5.6. The wind direction is plus or minus 30 degrees from the local dominant wind direction.

Relatively flat areas (RFA) are areas with a relative height of no more than 30 meters within a range of 300m*300m around the grid according to GDEMV3 30M resolution digital elevation data. The relief degree of land surface (RDLS) represents the average altitude, the proportion of flat areas, and the elevation range of a specific area. The calculation formula is as follows:

$$RDLS=ALT/1000+Range(H)*[1-P(A)/A]/BM$$

Where ALT is the average altitude within 2 km of the sample, Range(H) is the difference between the highest and lowest altitudes within 2 km, P(A) is the area of relatively flat area (RFA) within 2 km, A is the total area of 2 km, and BM is the benchmark altitude of mainland China (500 m).

Micro-level data were obtained through questionnaires.

2.3.3. Non-spatial Factors

Non-spatial influencing factors include the respondent’s smoking history (cumulative smoking duration, smoking amount, passive smoking), occupational exposure history (occupation type, employment duration, whether engaged in coal-related industries), life fuel use history (fuel use type, use duration, use amount), personal medical history and family genetic history. The above data were obtained through questionnaires.

2.4.Analytical Methods

2.4.1.Spatial Analysis

1-Average Nearest Neighbor

The average nearest neighbor is the ratio of the “average observed distance” to the “expected average distance”. If the index is less than 1, it is a clustered mode, and if the index is greater than 1, it is a discrete mode.

2-Nearest Neighbor Hierarchical Cluster Analysis

The nearest neighbor hierarchical cluster analysis method is used to explore the hot spots of spatial distribution of point data (Zhan D, Meng B, 2013). This method defines a threshold distance (random nearest neighbor distance) for clustering, and only points closer than the threshold distance will be selected for clustering. The first clustering forms the first-order clustering, followed by subsequent clustering to form second-order and third-order clustering until all clusters converge into a single cluster or clustering fails. This study used Crimestat 4.02 software to compare the hot spots of the current residence distribution of the case group and the control group to explore the differences in distribution and agglomeration between the two.

The study found significant differences in their distribution: the clusters of the case group were located in rural areas; the clusters of the control group were located in Fuyuan County, Housuo Town and Yingshang Town.

3-Multi-distance Spatial Cluster Analysis (Ripley's K function)
Ripley's K function is used to analyze the spatial distribution pattern of points at different distance scales (Zhan D, Meng B, 2013). Its principle is to calculate the number of other points within a certain distance r around the point (K observed value) and compare it with the theoretical value of completely random distribution (K expected value). If the K observed value at a specific distance is greater than the K expected value, the distribution is more clustered than the random distribution at that distance; if the K observed value is less than the K expected value, the dispersion is higher.

2.4.2. Case-control Study Based on Life Course Theory
Case-control study is the most commonly used analytical epidemiological research method, mainly used to explore the causes or risk factors of diseases and test etiological hypotheses. Its basic principle is to use a group of patients who have been diagnosed with a specific disease as the case group, and a group of individuals who do not have the disease but are comparable as the control group. Through inquiries and other methods, the subjects' previous exposure history to various possible risk factors is collected, and the differences in the exposure ratios of each factor between the case group and the control group are compared to see whether they are statistically significant.

Chronic diseases are affected by a variety of factors over a long period of time, and the occurrence of chronic diseases cannot be explained entirely by exposure in adulthood (Zheng Y, He D, 2015). Exposure at different ages may directly or indirectly affect health in adulthood (Bever A M, Song M, 2023). The life course theory holds that in the course of life, various factors independently, cumulatively, and interactively affect health and disease in adulthood (Kuh D, et al., 2003). Exposure to air pollution at a younger age may have more harmful effects because organ systems are still developing (Portengen L, et al., 2023). Lung development begins in early pregnancy and continues to grow until the age of 25, when lung function reaches a physiological stable period (Mahmoud O, et al., 2023). Since the age of 45, the incidence of lung cancer has increased rapidly. According to the life course, it is divided into three time periods, namely the development period (0-24 years old), early adulthood (25-44 years old), and middle-aged and elderly (45 years old and above). The association between risk exposure and lung cancer incidence in the three time periods was calculated separately.

SPSS29.0 was used for data analysis in this study. Cross-tabulation was used for univariate studies, and conditional logistic regression was used for multivariate studies to calculate the odds ratio (OR) and 95% confidence interval (95% CI) between risk factors and lung cancer incidence. Each continuous variable (each risk factor for cumulative exposure duration) was converted into a categorical variable using a dichotomy method, and the critical point was the average value of the control group.

3. Results
3.1.Spatial Statistics
The average nearest neighbor calculation was performed on the case group and the control group in Fuyuan and Xuanwei, and it was found that the degree of clustering of the case group in both places was higher than that of

the control group (table). Through the nearest neighbor hierarchical cluster analysis method, it was found that the clustering areas of the case group were mostly located in rural areas, while the clustering areas of the control group were located in urban areas, and there was a significant difference in the distribution of the two. Through multi-distance spatial cluster analysis, it was found that the case group was clustered within 21.5km, and the degree of clustering was the highest at 2km, while the control group was clustered within 21km, and the degree of clustering was the highest at 5km (Figure 1).

Table 1 Average nearest neighbor of the case group and the control group

Group		Nearest neighbor ratio	Z score	P value
Fuyuan	Case group	0.47	-7.38	0.00**
	Control group	0.82	-2.57	0.01**
Xuanwei	Case group	0.46	-4.53	0.00**
	Control group	1.01	0.09	0.93

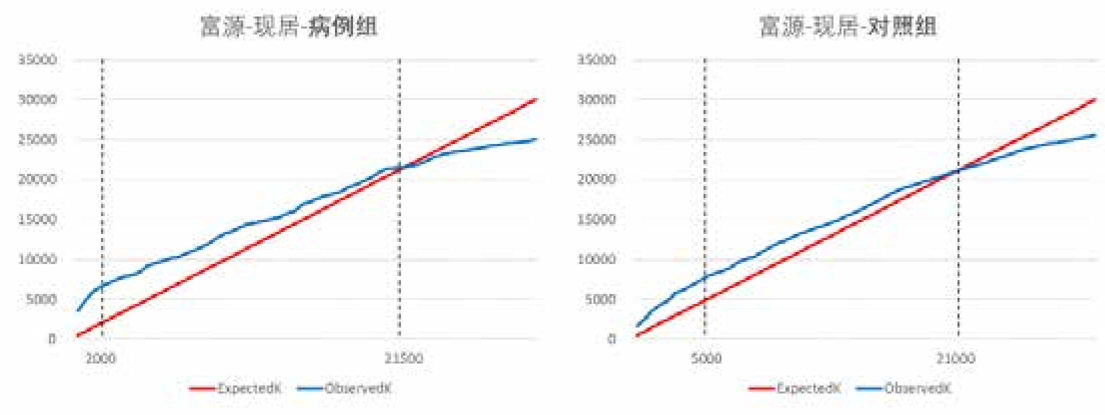

Fig.1 Multi-distance spatial clustering of case group and control group

In summary, the study found that the degree of clustering in the case group was higher than that in the control group, and the spatial distribution and scale of the clusters were significantly different, indicating that the local incidence of lung cancer was affected by spatial factors.

3.2.Case-control Study Results

3.2.1. Univariate Analysis

The odds ratios of spatial influencing factors are shown in Table 2.

In terms of the spatial relationship between residence and pollution sources, living within 2km of a coal mine for a cumulative period of 50 years is associated with a 4-fold risk of lung cancer (95%CI: 1.691-9.256). Living within 1km of a coal mine for a cumulative period of 30 years is also associated with an increased risk of lung cancer (OR: 2.220; 95%CI: 1.004-4.908). From the perspective of different life stages, long-term residence within 1km of a coal mine in middle-aged and elderly people is associated with a 2.4-fold risk of lung cancer (95%CI: 1.088-5.368); living within 1km of a coal mine during development and early adulthood is also a risk factor, but it is not significant in the study. Living within 1km of an industry for 26 years is less associated with lung cancer. However, in terms of stages, long-term residence within 1km of the surrounding industry during the developmental period was negatively correlated with the incidence of lung cancer (OR: 0.475; 95%CI: 0.212-1.066), while long-term residence within 1km of the surrounding industry during the middle-aged and elderly period was positively correlated with the incidence of lung cancer (OR: 2.218; 95%CI: 0.958-5.131), although the two were not statistically significant. The area of mining land within 2km of the residence was weakly associated with the incidence of lung cancer, and it was not significant. Being born in the downwind area closest to the coal mine was positively correlated with the risk of lung cancer (OR: 1.680; 95%CI: 0.739-3.818), but it was not statistically significant. In terms of housing structure, from the perspective of the whole life course, living in a civil house for a total of 28 years was associated with a 2-fold risk of lung cancer, but it was not significant (95%CI: 0.928-4.499); living in civil houses and concrete houses for a long time was weakly associated with the incidence of lung cancer. In terms of life course, during the developmental period, living in concrete houses was negatively correlated with the risk of lung cancer (OR: 0.108; 95%CI: 0.013-0.894); living in earthen houses and brick-wood houses was not significantly associated with the risk of lung cancer. In early adulthood, living in earthen houses was positively correlated with the risk of lung cancer (OR: 3.391; 95%CI: 1.499-7.668), living in concrete houses was negatively correlated with the risk of lung cancer (OR: 0.292; 95%CI: 0.103-0.823), and living in brick-wood houses was weakly associated with the risk of lung cancer. In middle-aged and elderly age, living in brick-wood houses for 3 years was positively correlated with the risk of lung cancer (OR: 2.759; 95%CI: 1.159-6.567), living in concrete houses for 6 years was positively correlated with the risk of lung cancer (OR: 3.853; 95%CI: 1.665-8.915), and living in earthen houses was weakly associated with the risk of lung cancer.

In terms of cooking location, cooking in the main room for 9 years in early adulthood was positively correlated with the risk of lung cancer (OR: 2.220; 95%CI: 1.004-4.908), and cooking in the kitchen in middle-aged and elderly age was positively correlated with the risk of lung cancer (OR: 6.500; 95%CI: 2.722-15.521).

Table 2 Spatial influencing factors and lung cancer risk

Characteristic	Case (n=51)		Control (n=51)		OR(95%CI)
	No	Yes	No	Yes	
Living within 2km of a coal mine for 50 years	**12**	**39**	**28**	**23**	**3.957(1.691-9.256)**
Living within 1km of a coal mine for 30 years	**19**	**32**	**29**	**22**	**2.220 (1.004-4.908)**
0-24 years old for 13 years	17	34	23	28	1.643 (0.737-3.663)
25-44 years old for 9 years	17	34	24	27	1.778(0.798-3.960)
45 years old and above for 9 years	**22**	**29**	**33**	**18**	**2.417(1.088-5.368)**
The mining area within 2km of the residential area exceeds 90,000 m2 for 30 years	24	27	25	26	1.082(0.497-2.353)
Living within 1km of the industrial area for a total of 26 years	28	23	30	21	1.173(0.536-2.571)
0-24 years old for 12 years	35	16	26	25	0.475(0.212-1.066)
25-44 years old for 8 years	29	22	30	21	1.084 (0.494-2.379)
45 years old and above for 6 years	29	22	38	13	2.218(0.958-5.131)
The birthplace is located in the downwind area closest to the coal mine	30	21	36	15	1.680(0.739-3.818)
Living in a civil house for a total of 28 years	20	31	29	22	2.043(0.928-4.499)
0-24 years old for 17 years	9	42	14	37	1.7669(0.685-4.551)
25-44 years old for 6 years	**20**	**31**	**35**	**16**	**3.391(1.499-7.668)**
45 years old and above for 3 years	38	13	39	12	1.112 (0.451-2.743)
Living in a brick-and-wood house for a total of 16 years	25	26	31	20	1.612 (0.735-3.537)
0-24 years old for 5 years	45	6	39	12	0.433(0.149-1.263)
25-44 years old for 6 years	30	21	32	19	1.179(0.532-2.613)
45 years old and above for 3 years	**29**	**22**	**40**	**11**	**2.759 (1.159-6.567)**

Living in concrete house for a total of 12 years	33	18	31	20	0.845(0.379-1.888)
0-24 years old for 1 year	**50**	**1**	**43**	**8**	**0.108(0.013-0.894)**
25-44 years old for 6 years	**45**	**6**	**35**	**16**	**0.292(0.103-0.823)**
45 years old and above for 6 years	**22**	**29**	**38**	**13**	**3.853(1.665-8.915)**
Cooking in the main room for a total of 37 years	21	30	28	23	1.739(0.794-3.810)
0-24 years for 22 years	10	41	12	39	1.262(0.489-3.252)
25-44 years for 9 years	**19**	**32**	**29**	**22**	**2.220(1.004-4.908)**
45 years and above for 5 years	36	15	34	17	0.833(0.361-1.926)
Cooking in the kitchen for a total of 19 years	22	29	28	23	1.605(0.734-3.507)
0-24 years old for 2 years	44	7	42	9	0.742(0.254-2.174)
25-44 years old for 9 years	32	19	26	25	0.618(0.280-1.360)
45 years old and above for 8 years	**12**	**39**	**34**	**17**	**6.500 (2.722-15.521)**
Use fire pit for a total of 18 years	10	41	18	33	2.236(0.910-5.493)
0-24 years old for 16 years	9	42	16	35	2.133(0.840-5.416)
25-44 years old for 2 years	40	11	43	8	1.478(0.540-4.048)
45 years old and above for 1 year	46	5	48	3	1.739(0.393-7.697)
Use floor stove for a total of 21 years	22	29	27	24	1.483(0.679-3.237)
0-24 years old for 7 years	42	9	36	15	0.514(0.201-1.315)
25-44 years old for 10 years	19	32	24	27	1.497(0.679-3.300)
45 years old and above for 4 years	**29**	**22**	**40**	**11**	**2.759(1.159-6.567)**

The odds ratios of non-spatial factors are shown in Table 3.

Using bituminous coal for 11 years in early adulthood and 3 years in middle-aged and elderly age were positively correlated with the risk of lung cancer (OR: 2.401; 95%CI: 1.040-5.546, OR: 3.540; 95%CI: 1.472-8.782). Using a floor stove for 4 years in middle-aged and elderly age was positively correlated with the risk of lung cancer (OR: 2.759; 95%CI: 1.159-6.567).

Table 3 Non-spatial influencing factors and lung cancer risk

Characteristic	Case (n=51)		Control (n=51)		OR (95%CI)
	No	Yes	No	Yes	
Cooked for 19 years	24	27	29	22	1.483(0.679-3.237)
Used bituminous coal for 33 years	15	36	23	28	1.971(0.871-4.461)
0-24 years old for 18 years	12	39	13	38	1.112(0.451-2.743)
25-44 years old for 11 years	**13**	**38**	**23**	**28**	**2.401(1.040-5.546)**
45 years old and above for 3 years	**29**	**22**	**42**	**9**	**3.540(1.472-8.782)**
Using anthracite for a total of 7 years	37	14	36	15	0.908(0.384-2.148)
0-24 years old for 4 years	47	4	43	8	0.457(0.129-1.628)
25-44 years old for 2 years	45	6	43	8	0.717(0.230-2.236)
45 years and over for 2 years	39	12	45	6	2.308(0.792-6.726)
Total firewood usage for 25 years	34	17	30	21	0.714(0.319-1.600)
0-24 years old for 12 years	29	22	26	25	0.789(0.362-1.721)
25-44 years old for 8 years	33	18	29	22	0.719(0.324-1.597)

45 years and over for 5 years	35	16	38	13	1.336(0.563-3.170)
Have smoked before	23	28	25	26	1.171(0.538-2.549)
0-24 years old ever smoked	32	19	28	23	0.732(0.328-1.595)
25-44 years old ever smoked	28	23	26	25	0.854(0.392-1.860)
45 years and over ever smoked	31	20	29	22	0.850(0.386-1.873)
Worked in coal-related jobs	33	18	32	19	0.919(0.410-2.060)
Have a family medical history	32	19	37	14	1.569(0.680-3.624)
Have a personal medical history	46	5	47	4	1.277(0.323-5.057)

3.2.2. Conditional Logistic Regression

Due to the limited sample size of this study, the significant variables in the univariate study were included in the binary logistic regression model using the stepwise method. The input variables were living within 2km of the coal mine for a total of 50 years, living within 1km of the coal mine for a total of 30 years, living within 1km of the coal mine for a total of 9 years for people aged 45 and above, living in earthen houses for 6 years for people aged 25-44, living in brick and wood houses for 6 years for people aged 25-44, living in concrete houses for 1 year for people aged 0-24, living in concrete houses for 6 years for people aged 25-44, living in concrete houses for 6 years for people aged 45 and above, cooking in the main room for 9 years for people aged 25-44, cooking in the kitchen for 8 years for people aged 45 and above, using bituminous coal for 11 years for people aged 25-44, using bituminous coal for 3 years for people aged 45 and above, and using floor stoves for 4 years for people aged 45 and above.

The odds ratio of the independent effect of each influencing factor is shown in Table 4. The study found that long-term residence in earthen houses in early adulthood was positively correlated with the risk of lung cancer (OR: 2.891; 95%CI: 1.144-7.311); the use of bituminous coal for 3 years in middle-aged and elderly people was positively correlated with the risk of lung cancer (OR: 3.612; 95%CI: 1.431-9.731); and living within 2km of a coal mine for a cumulative period of 50 years was positively correlated with the risk of lung cancer (OR: 2.433; 95%CI: 0.951-6.222).

Table 4 Independent effects of various influencing factors on lung cancer risk

Characteristic	Case(n=51)		Control(n=51)		OR(95%CI)	P Value
	No	Yes	No	Yes		
25-44岁居住土木房达6年	20	31	35	16	2.891(1.144-7.311)	0.025
45岁后使用烟煤达3年	29	22	42	9	3.612(1.431-9.731)	0.011
居住在煤矿2km内达50年	12	39	28	23	2.433(0.951-6.222)	0.064

4. Discussion

4.1. Spatial statistical analysis

In spatial statistics, it was found that the degree of agglomeration of the case group was higher than that of the control group, indicating that in addition to the agglomeration of towns and villages, there are other spatially related factors that cause the cases to agglomerate. The study found that when the research range was 2km, the degree of agglomeration of the case group was the highest, and when the research range was 5km, the degree of agglomeration of the control group was the highest, indicating that with a range of 2km, the environmental factors around the study sample were a suitable distance for the impact of lung cancer. The analysis of the agglomeration hotspots of the case group and the control group found that the hotspots of the control group were in Fuyuan County, Housuo Town, and Yingshang Town, which were urbanized areas with concentrated populations; while the hotspots of the case group were mainly distributed in rural areas. There were significant urban-rural differences between the two, and the urban-rural differences are the content that needs to be paid attention to in subsequent studies. In addition to geographical location, urban-rural differences also include different living environment quality, fuel use types, occupations, living habits, etc. The combined effect of multiple factors has created lung cancer hotspots in rural areas.

4.2. Univariate analysis

In the univariate analysis, long-term residence within 2km of a coal mine was positively correlated with the risk of lung cancer; living within 1km of a coal mine in middle-aged and elderly age was positively correlated with the risk of lung cancer, while similar trends were shown in the developmental period and early adulthood, but not significant. The middle-aged and elderly stage of the sample corresponds to the period from 2000 to the present. Some studies have shown that the local incidence of lung cancer has dropped significantly after the 1970s thanks to the improvement of stoves (Lan Q, et al., 2002). Although some studies have shown that early life exposure has a greater impact on lung cancer (Portengen L, et al., 2023), the exposure risk in middle-aged and elderly age should not be ignored.

Studies have shown that long-term residence within 1km of industry is not significantly associated with the risk of lung cancer, but in terms of life course, living around industry in the developmental period tends to be a protective factor, while living around industry in middle-aged and elderly age tends to be a risk factor. This may be because industrial areas have changed from “development centers” in the early years to “environmentally degraded areas” in recent years: in the developmental period of the research sample (corresponding to the 1960s-1990s), industrial areas are often the places where various resources are most concentrated, such as schools for children of workers and miners, employee hospitals, and unit housing, providing better education and health resources, thereby promoting early health; while in the middle-aged and elderly period of the research sample (corresponding to the 2000s to the present), non-industrial areas have generally achieved rapid development, while the aging of infrastructure around industrial areas, the accumulation of industrial pollution, and the deterioration of the ecological environment have increased the risk of exposure to coal dust, heavy metals, etc., which has become a health burden. In addition, people with higher socioeconomic abilities may choose to move out of the most polluted industrial areas and move into new urban areas far away from industry. In subsequent studies, analysis of urban-rural differences and socioeconomic levels should be included. (Subsequent analysis of socioeconomic level, whether living in urban/mining areas, and facilities poi)

In terms of the structure and materials of the house, living in earthen houses in early adulthood and brick-wood houses in middle-aged and elderly age are risk factors for lung cancer; long-term residence in concrete houses is not significantly associated with the risk of lung cancer, but in terms of life course, it is a protective factor in the developmental period and early adulthood, while it is a risk factor in the middle-aged and elderly age. The house structure may affect indoor ventilation, and thus affect the diffusion of pollutants. Tile houses with earthen structures and brick-wood structures are typical local residential forms, usually two-story, with the first floor for the main room (living, cooking) and bedrooms, and the second floor for bedrooms or for storing sundries; there are doors and windows on the front facade, no windows on the back wall, and poor ventilation conditions. Tile houses with earthen structures are usually older, while brick-wood structures are more recent. The form of concrete houses is no different from that of other regions, with complete rooms and through-the-hall wind conditions. Living in concrete houses during development and early adulthood is a protective factor, probably because concrete houses have better ventilation capacity and are easier to be equipped with smoke exhaust systems, and are also a sign of a higher socioeconomic level. Living in earthen houses in early adulthood and brick-wood houses in middle-aged and elderly age are risk factors, which may be due to the poor ventilation of earthen and brick-wood houses, and are also related to poor socioeconomic levels. The study found that living in concrete houses in middle-aged and elderly age is a risk factor, which is contrary to common perception. This may be because concrete houses are more common in urban areas and industrial and mining areas, thus exposing people to environmental pollution for a long time. In addition, due to the good airtightness of the house, if the windows are not opened for ventilation, pollutants may accumulate. (Subsequently supplemented with floor plans, socioeconomic levels, and whether living in urban areas/mining areas)

In terms of cooking location, cooking in the living room in early adulthood is a risk factor; cooking in the kitchen in middle-aged and elderly age is a risk factor; cooking in the kitchen during development and early adulthood tends to be a protective factor, although not significant. This may be because there is a clear partition between the kitchen in development and early adulthood, which can effectively reduce indoor air pollution in the living area; in middle-aged and elderly age, due to the change of family roles, they need to mainly undertake cooking tasks, and pollution is easy to accumulate in the kitchen, resulting in higher pollution exposure. (Further analysis will be conducted based on whether or not the person is responsible for cooking at home)

In terms of stove use, the use of fire pits in each life course showed a tendency to be a risk factor, although it was not significant, which is consistent with the conclusion of existing studies that pollution generated by burning organic matter indoors cannot be discharged. Using a floor stove in middle-aged and elderly age is a risk factor, probably because it is still more polluting than gas stoves. (Gas stoves will be added later)

Among non-spatial influencing factors, the duration of coal use in early adulthood and middle-aged and elderly age is a risk factor, which is consistent with existing studies.

4.3. Multivariate Analysis

In the multivariate analysis, it was found that long-term living within 2km of a coal mine, living in earthen houses in early adulthood, and using bituminous coal in middle-aged and elderly age were risk factors for lung cancer.

Although some variables (such as cooking in the kitchen, using floor stoves, and living in concrete houses for many years) were shown as risk factors or protective factors in the univariate analysis, they were not included in the final model. This may be due to the following reasons: 1) There was collinearity between the variables, or the variables themselves were mediators or proxy indicators of other key exposure factors. For example, "cooking in the living room at the age of 25-44" and "living in earthen houses" both reflected fuel exposure and ventilation conditions under low living conditions. "Brick and wood houses" are the intermediate type of housing structure, and their ventilation and pollution isolation capabilities are between earthen houses and concrete houses. Their protective or risk effects may be masked; 2) The exposure time of some variables is relatively short, such as using floor stoves for 4 years for people over 45 years old, and the exposure period is relatively short; 3) The direction of action of some variables is inconsistent at different stages of life. For example, concrete houses are usually regarded as variables with protective effects, but there are also proxy variables such as other environmental factors such as proximity to industrial and mining areas that are eliminated.

Through multivariate logistic regression model, the study revealed the core influencing factors of meso-scale, micro-scale and confounding factors among many influencing factors.

5. Conclusions

5.1. Researching Conclusions

This study used a case-control study based on life course theory. Through spatial statistics, univariate and multivariate logistic regression analysis, it systematically explored the association between environmental factors and lung cancer risk in residents in coal resource-based urban and rural areas at different life course stages, and the following conclusions were drawn:

1) There are significant urban-rural differences in spatial distribution, and 2km is the sensitive radius of environmental exposure to lung cancer. The case group showed higher spatial agglomeration, and hotspots were

mainly distributed in rural areas; the hotspots of the control group were concentrated in densely populated urban areas. Spatial scale analysis found that when the analysis radius was 2km, the agglomeration of lung cancer cases was the strongest, indicating that environmental exposure within a 2km range had a strong impact on the incidence of lung cancer and should be used as a key spatial unit for future environmental health intervention.

2) Even if the use of fuel factors such as bituminous coal is taken into account, long-term residence around coal mines is still a significant risk factor for lung cancer. Both univariate and multivariate analyses showed that living within 2km of a coal mine for more than 50 years significantly increased the risk of lung cancer. The risk of living within 1km of a coal mine in the middle-aged and elderly (after 45 years old) is higher, indicating that the long-term accumulation of industrial pollution and the exposure effect at an advanced age cannot be ignored.

3) The characteristics of the living environment at different stages of the life course have heterogeneous effects. Living in earthen houses in early adulthood (25–44 years old) is a risk factor for lung cancer, which may be related to poor ventilation conditions and the use of highly polluting fuels; living in brick and wood houses in middle-aged and elderly (≥45 years old) is also a risk factor, suggesting that environmental aging may increase the health burden on the basis of unimproved quality of life. In contrast, living in concrete houses during development and early adulthood showed a certain protective effect, which may be related to better ventilation conditions and higher socioeconomic status; but living in concrete houses in middle-aged and elderly people showed a risk tendency, suggesting that concrete houses may be a proxy indicator of exposure to industrial areas or highly enclosed polluted environments.

4) Although the location of cooking in the kitchen and the main room and the type of stove were not included in the final model, they are still important risk factors. Cooking in the main room (early adulthood) and cooking in the kitchen (middle-aged and elderly) are both potential risk factors for lung cancer, reflecting the problem of no isolation between the living area and the cooking area and the diffusion of indoor pollutants. The use of floor stoves in middle-aged and elderly people is also a risk factor. Although these variables were not retained in the multivariate model, they still have certain public health significance as behavioral indicators of ventilation and combustion pollution.

5.2. Planning inspiration: meso- and micro-level interventions

In terms of spatial intervention, at the meso level, we should identify high-risk areas for lung cancer and make the 2km area around coal mines the key intervention area for public health; distinguish urban and rural differences and strengthen pollution detection around industrial and mining areas. At the micro level, we should promote the renovation of rural houses and the renewal of ventilation facilities, incorporate the concept of “through-hallway wind” into local architectural design, and improve the effect of natural ventilation; promote the supply of clean energy (such as electricity, gas, etc.) in rural areas.

In terms of policy recommendations, industrial and mining planning should conduct health impact assessment (HIA) and no longer be dominated by economic benefits alone; for residents who have lived in coal mines and highly polluted areas for a long time, lung screening should be vigorously promoted; and clean energy and healthy buildings should be promoted.

5.3. Research deficiencies and follow-up work

Socioeconomic attribute variables have not been included. The current model lacks variables such as the respondents’ education level and occupation type, which may misestimate the actual effects of residential location and house structure. The changes in residential history at the urban and rural levels have not been sorted out. Urban and rural differences include multiple influencing factors such as environmental pollution, fuel type, ventilation conditions, and medical accessibility. For example, living in a concrete house for people over 45 years old is a risk factor. It may not be a problem with the house itself, but "the house is located in an urban area or an industrial and mining area." Subsequent research needs to further analyze whether people have moved from rural areas to cities during their life course. The interactive effects of multiple variables have not been taken into account, ignoring the synergistic effects between variables.

References

Lu Shuo, Zhang Wenzhong, Yu Jianhui, et al. Identification of evolutionary stages and development characteristics of resource-based cities[J]. Acta Geographica Sinica, 2020, 75(10):2180-2191.

Wang Jiechen, Lu Min, Yuan Zhenyu, et al. Study on the spatial distribution pattern of ATM outlets in Nanjing based on Ripley's K function[J]. Geographical Science, 2016, 36(12):1843-1849.DOI:10.13249/j.cnki.s gs.2016.12.009.

Wang Lan. Healthy city science and planning evidence-based practice[J]. Journal of Urban Planning, 2023, (06): 27-31. DOI: 10.16361/j.upf.202306005.

Wei Haitao, Liu Yan, Tian Zhihui, Wang Xiaolei, He Xiaohui, Wei Jing. Spatial distribution pattern of lung cancer and the influence of environmental factors in Henan Province[J]. Acta Geographica Sinica, 2022, 77(01): 245-258.

Zhan Dongsheng, Meng Bin. Spatial agglomeration characteristics of Beijing residents' residence and employment based on social attributes[J]. Acta Geographica Sinica, 2013, 68(12): 1607-1618.

Zheng Yuan, He Dian. Development and application of life course epidemiology[J]. Chinese Journal of Disease Control and Prevention, 2015, 19(02): 196-199. DOI: 10.16462/j.cnki.zhjbkz.2015.02.024.Barone-Adesi F, Chapman R S, Silverman D T, et al. Risk of lung cancer associated with domestic use of coal in Xuanwei, China: retrospective cohort study[J]. Bmj, 2012, 345.

Bever A M, Song M. Early-life exposures and adulthood cancer risk: a life course perspective[J]. JNCI: Journal of the National Cancer Institute, 2023, 115(1): 4-7.

Hosgood III H D, Chapman R S, Wei H, et al. Coal mining is associated with lung cancer risk in Xuanwei, China[J]. American journal of industrial medicine, 2012, 55(1): 5-10.

Hu W, Downward G S, Reiss B, et al. Personal and indoor PM2. 5 exposure from burning solid fuels in vented and unvented stoves in a rural region of China with a high incidence of lung cancer[J]. Environmental science & technology, 2014, 48(15): 8456-8464.

Kuh D, Ben-Shlomo Y, Lynch J, et al. Life course epidemiology[J]. Journal of epidemiology and community health, 2003, 57(10): 778.

Lan Q, Chapman R S, Schreinemachers D M, et al. Household stove improvement and risk of lung cancer in Xuanwei, China[J]. Journal of the National Cancer Institute, 2002, 94(11): 826-835.

Mahmoud O, Granell R, Peralta G P, et al. Early-life and health behaviour influences on lung function in early adulthood[J]. European Respiratory Journal, 2023, 61(3).

Portengen L, Downward G, Bassig B A, et al. Methylated polycyclic aromatic hydrocarbons from household coal use across the life course and risk of lung cancer in a large cohort of 42,420 subjects in Xuanwei, China[J]. Environment international, 2023, 173: 107870.