

RESEARCH ARTICLE

Editorial Process: Submission:08/28/2017 Acceptance:07/16/2019

Cluster Analysis Evaluating PM2.5, Occupation Risk and Mode of Transportation as Surrogates for Air-pollution and the Impact on Lung Cancer Diagnosis and 1-Year Mortality

Sopian Abdul Wahab¹, Astrid Hassan^{1,2}, Mohd Talib Latif^{3,4}, Yasheeny Vadiveel¹, Tamyenthini Jeyabalan¹, Chun Ian Soo¹, Faisal Abdul Hamid¹, Andrea Ban Yu-lin¹, Tidi Hassan^{1*}

Abstract

Objective: Epidemiological studies have reported the close relationship between risk for lung cancers and air pollution in particular, for non-smoking related lung cancers. However, most studies used residential address as proxies which may not estimate accurately an individual's air pollution exposure. Therefore, the aim of this study was to identify risk factors such as occupation and mode of transportation associated with lung cancer diagnosis and death. **Methods:** Subjects with lung cancer (n=514) were evaluated both by chart reviews for clinical data and interviews to determine residential address for ten years, main occupation and main mode of transportation. Annual particulate matter with diameter size less than 2.5 micrometre (PM2.5) concentration were calculated based on particulate matter with diameter size less than 10 micrometre (PM10) data recorded by Malaysian Department of Environment. Logistic regression analysis, cluster analysis and the Cox regression analysis were performed to the studied variables. **Results:** This study concurred with previous studies that lung adenocarcinoma were diagnosed in predominantly younger, female non-smokers compared to the other types of lung cancers. Lung adenocarcinoma subjects had annual PM2.5 that was almost twice higher than squamous cell carcinoma, small cell carcinoma and other histological subtypes (p=0.024). Independent of smoking, the κ -means cluster analysis revealed two clusters in which the high risk cluster involves occupation risk with air pollution of more than four hours per day, main transportation involving motorcycle and trucks and mean annual PM2.5 concentration of more than 30 based on residential address for more than ten years. The increased risk for the high-risk cluster was more than five times for the diagnosis of lung adenocarcinoma (OR=5.69, 95% CI=3.14-7.21, p<0.001). The hazard ratio for the high-risk cluster was 3.89 (95% CI=2.12-4.56, p=0.02) for lung adenocarcinoma mortality at 1 year. **Conclusion:** High-risk cluster including PM2.5, occupation risk and mode of transportation as surrogates for air-pollution exposure was identified and highly associated with lung adenocarcinoma diagnosis and 1-year mortality.

Keywords: PM 2.5- occupation risk- air pollution- Lung cancer

Asian Pac J Cancer Prev, 20 (7), 1959-1965

Introduction

Lung cancer remains the most common cause of death from cancer in Malaysia and worldwide (Liam et al., 2006). Although cigarette smoke is a major cause of lung cancer, air pollution has been classified as group 1 carcinogens in humans by the International Agency for Research on Cancer of the World Health Organization (IARC, 2013). Several epidemiological studies have shown the close relationship between risks for lung cancer especially in non- and or never smokers and different composition of air pollution such as particulate matters

(PM), nitrogen dioxide (NO₂) and sulphur dioxide (SO₂) exposure (Rasschou-Nielsen et al., 2013; Beeson et al., 1998). It has been proposed that the shift of the major histological types of lung cancer from squamous cell carcinomas to adenocarcinomas may relate to outdoor or ambient air pollution (Rasschou-Nielsen et al., 2011).

Like other Asian countries, the high incidence of female non-smokers with adenocarcinoma has also been reported in Malaysia. It seems that risk factors other than smoking may play a more important role in the pathogenesis of female lung cancer (Nakamura et al., 2013; Ha et al., 2014). However, research to conduct

¹Respiratory Unit, Universiti Kebangsaan Malaysia Medical Centre, Faculty of Medicine, University Kebangsaan, Cheras, ²Department of Clinical Oncology, University Technology MARA, Shah Alam, ³School of Environmental and Natural Resource Sciences, Faculty of Science and Technology, ⁴Institute for Environment and Development (Lestari), Universiti Kebangsaan Malaysia, Bangi, Malaysia. *For Correspondence: tidi.maharani.hassan@ppukm.ukm.edu.my

the association between air pollution and lung cancers remains a challenge due to lack of standardized estimation and quantification of ambient air pollution. Most studies used residential address and air quality data as proxies to estimate an individual's air pollution exposure (Nafstad et al., 2003; Chen et al., 2007). To date, there has been no study examining the association between air pollutants and daily activities influencing exposure with lung cancers especially in non-smoking subjects with lung adenocarcinoma in Malaysia or Southeast Asia.

We hypothesized that daily human activity including occupation risk and mode of transportation impact on the total exposure of air pollution which may influence the development of lung cancers and especially lung adenocarcinoma. Using ten years air pollution components index data from air quality monitoring stations and a telephone or face-to-face interview, we aimed to identify risk factors relating to occupation risk and mode of transportation for the diagnosis of lung cancers and assess whether these variables affect survival in subjects with lung cancers. This study also aimed to identify clusters of patients with high risk factors associated with lung cancer diagnosis and death.

Materials and Methods

Our study population included all lung cancer diagnoses recorded in the Department of Pathology, Department of Medicine, Department of Medical Oncology and electronic medical records in Universiti Kebangsaan Malaysia Medical Centre (UKMMC) from 1st January 2010 to 31st July 2016. Clinical demographics, histopathological and radiological data and treatment received were collected retrospectively from chart reviews and discharge registers in addition to the electronic medical records. The final diagnoses were recorded using the International Statistical Classification of Diseases and Related Problems, 10th revision (ICD-10). Face-to-face interview or telephone interview were performed prospectively to subjects or the next-of-kin to confirm residential address for the past ten years and determine the main occupation, average daily ambient air exposure and transportation used by the subjects recruited using a pre-specified questionnaire. The outcomes of lung cancer death were obtained from the hospital's electronic database and confirmed using a face-to-face or telephone interview. Ethical approval was obtained from UKMMC Ethics Committee (ref: FF-2016-440).

Our institution is located in Klang Valley which is the mainstream economic region in Malaysia with extensive physical development of infrastructure, industrialisation and urbanisation. However, as one of the biggest university hospital in Malaysia, the subjects recruited were also from areas outside the Klang Valley. The annual particulate matter with diameter size less than 2.5 micrometre (PM_{2.5}) concentration were derived as predicted concentration, measured as around seventy percent of particulate matter with diameter size less than 10 micrometre (PM₁₀) (Amil et al., 2016). Data for PM₁₀ were obtained from the Air Quality Division, the Department of Environment, the Ministry of Natural

Resources and Environment of Malaysia. Concentrations of PM₁₀ recorded in the database were collected from forty air-monitoring stations that obtain hourly measurements for 24 hours. The residential addresses of the subjects were geocoded, and matched with the surface extinction coefficient data using regression models to estimate annual PM_{2.5} concentrations from the nearest air-monitoring stations.

Statistical analysis

Quantitative data with a normal distribution were expressed using the mean and standard deviation (SD). Qualitative data were described using relative frequencies. Comparisons of categorical variables and continuous variables were performed using the chi-square test and Mann-Whitney test respectively. κ -means cluster analyses were used to identify clusters of subjects. A three-way ANOVA test was used to calculate differences between lung cancer histopathological subtype groups. Unconditional logistic regression was used to estimate odds ratio (OR) and their 95% confidence intervals (CIs) for the association between each studied variables with the diagnosis of lung cancers. A posteriori, clusters' validation was performed to verify the association of the clusters at 52 week-mortality by Kaplan-Meier approach with a log-rank test. The Cox proportional hazards model was used to analyse the association of overall survival rate at 52 weeks with multiple covariates. The Statistical Package of Social Science (SPSS) version 23 and Prism Graph Pad version 7 were used for statistical analysis.

Results

A total of 514 subjects were diagnosed with lung cancer from 1 January 2010 to 31 July 2016 and recruited into this study. Seventy percent (n=360) of the subjects had confirmed adenocarcinoma histology subtype. Eighty-seven percent of subjects diagnosed with both adenocarcinoma and non-adenocarcinoma subtypes were diagnosed at advanced stage (ie stage 3 or 4). Table 1 shows the clinical characteristics and baseline demographics of these subjects. Forty one percent of subjects diagnosed with lung adenocarcinoma were female compared to eighteen percent for non-adenocarcinoma lung cancers (p=0.04). The mean (SD) age for lung adenocarcinoma was significantly less than the non-adenocarcinoma subtypes (59.1 (6.4) vs 67.6 (13.9), p=0.03). The non-adenocarcinoma subtypes had a higher rate of smoking history (of more than 30 pack years) compared to lung adenocarcinoma (75% vs 47%, p=0.02).

High occupational risk defined as exposure to more than four hours to ambient air was more prevalent in lung adenocarcinoma compared to non-adenocarcinoma subtypes (39% vs 12%, p=0.02). Although there were no significant differences for the different kind of transportation that would confer differential risk to air pollution, truck and/or lorry drivers were more commonly seen amongst subjects with lung adenocarcinomas. The mean (SD) PM_{2.5} concentration based on the subjects' residential address was higher compared to the other subtypes (42.2 (19.5) vs 24.2 (8.9), p=0.01).

Table 1. Baseline Demographic and Clinical Characteristics of Lung Adenocarcinoma and Non-adenocarcinoma Lung Cancers (n=154)

Clinical characteristics	Adenocarcinoma (N=360)	Others (N=154)	P-value
Sex, male, n (%)	211 (59)	126 (81)	0.04
Age, years (SD)	59.1(6.4)	67.6 (13.9)	0.03
Smoking, n (%)	168 (47)	116 (75)	0.02
Race			
Malay, n (%)	132 (37)	69 (45)	0.24
Chinese, n (%)	163 (45)	59 (38)	0.07
Indian, n (%)	62 (17)	26 (17)	0.36
Others, n (%)	3 (0.8)	0 (0)	0.71
Level of education			
Primary	138 (38)	61 (40)	0.45
Secondary	171 (48)	69 (45)	0.32
Tertiary	51 (14)	24 (16)	0.62
Occupation* (AP exposure)			
No exposure, n (%)	96 (27)	86 (57)	0.18
Exposure 1-4 hours, n (%)	125 (35)	49 (32)	0.22
Exposure >4 hours, n (%)	139 (39)	19 (12)	0.02
Biomass exposure, n (%)	2 (0.5)	1 (0.6)	0.32
Transportation			
Car	156 (43)	71 (46)	0.22
Public transport	78 (22)	29 (19)	0.34
Motorcycle	112 (31)	54 (35)	0.12
Truck/Lorry	14 (4)	0 (0)	0.03
PM2.5#, concentration (SD)	42.2 (19.5)	24.2 (8.9)	0.01
Lung cancer stages			
Stage 1/2	47 (13)	23 (15)	0.72
Stage 3/4	313 (87)	131 (87)	0.54

Abbreviations, SD; standard deviation; AP, air pollution; PM, particulate matter; *Average outdoor exposure per day; #Based on residential address for ten years

Figure 1 shows the annual concentrations of PM2.5 of the different histological subtypes of lung cancer. Lung adenocarcinoma subjects had annual PM2.5 that was higher than squamous cell carcinoma, small cell carcinoma and other histological subtypes (p=0.024).

κ -means cluster analyses independent of gender, ethnicity and smoking status was able to identify two cluster of patients with 192 subjects composing cluster I (CL I), 241 in cluster 2 (CL II) and 151 not clustered (Table 2). CL I which is categorized as high risk involves occupation risk with air pollution of more than four hours per day, main transportation involving motorcycle and trucks and mean annual PM2.5 concentration of more than 30 based on residential address for more than ten

years. CL II which is categorized as low risk comprises of occupation risk with air pollution of less than four hours per day, main transportation involving car or public transportation and mean annual PM2.5 concentration of less than thirty based on residential address for more than ten years.

Table 3 shows the independent association between the studied variables including the clusters proposed and lung adenocarcinoma diagnosis. Established associations with lung adenocarcinoma include female gender (OR=1.91, 95% CI=1.69-2.56, p=0.03), age more than 55 (OR=2.14, 95% CI=1.67-2.76, p=0.01) and the Chinese race (OR=1.81, 95% CI=1.43-2.45, p=0.02). Smoking pack years of more than thirty showed

Table 2. κ -means Cluster Analysis Identified Two Clusters of Subjects for the Diagnosis of Lung Adenocarcinoma Independent of Gender, Ethnicity and Smoking Status¹

CL I: High risk (N=122)	CL II: Low risk (N=241)
Occupation risk > 4 hours of AP*	Occupation risk < 4 hours of AP*
Transportation (Motorcycle, truck)	Transportation (Car, Public transport)
Annual PM2.5# >30	Annual PM2.5# <30

Abbreviations, CL, cluster analysis; AP, air pollution; PM, particulate matter; *Average outdoor exposure per day; #Based on residential address for ten years; ¹151 patients not clustered

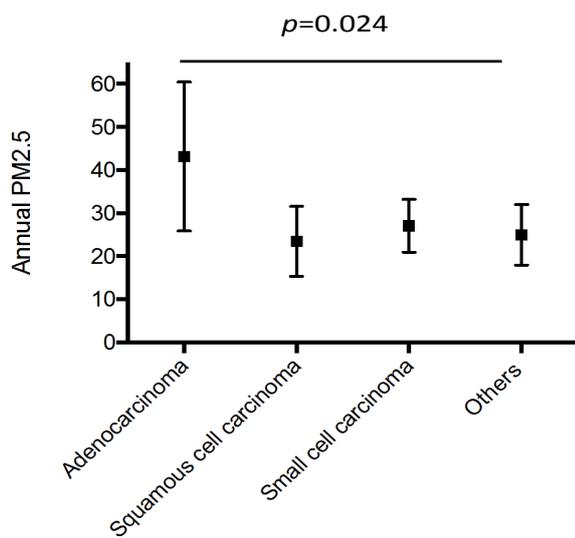


Figure 1. The Annual PM2.5 Concentration Based on 10-year Residential Address in Different Types of Lung Cancers

OR=2.61, 95% CI=1.34-3.92 but was not significant (p=0.06). Occupation risk with ambient air exposure of one to four hours, and more than four hours revealed OR=1.67, 95% CI=1.29-1.99, p=0.04 and OR=2.46, 95% CI=1.28-5.42, p=0.01 respectively. Motorcycle and truck as main transportation showed increased likelihood for lung adenocarcinoma as well (OR=2.42, 95% CI=1.87-3.12, p=0.04 and OR=1.89, 95% CI=1.65-3.22, p=0.03 respectively). Annual PM2.5 of more than 30 showed increased risk with OR=2.12, 95% CI=1.65-3.22, p=0.02 for lung adenocarcinoma diagnosis. With CL I, the increased risk was more than five times likely for the diagnosis of lung adenocarcinoma (OR=5.69, 95% CI=3.14-7.21, p<0.001).

Table 3. Odds Ratio Using Unconditional Logistic Regression Exploring the Independent Association of the Studied Variables with Lung Adenocarcinoma Diagnosis

Variable	Adenocarcinoma	
	OR (95% CI)	p-value
Sex, female	1.91 (1.69-2.56)	0.03
Age > 55	2.14 (1.67-2.76)	0.01
Pack years >30	2.61 (1.34-3.92)	0.06
Race		
Malay	1.12 (0.91-1.95)	0.42
Chinese	1.82 (1.43-2.45)	0.02
Indian	0.97 (0.89-1.20)	0.74
Occupation risk* (AP)		
No exposure	1.19 (0.82-1.82)	0.35
1-4 hours	1.67 (1.29-1.99)	0.04
> 4 hours	2.46 (1.28-5.42)	0.01
Transportation		
Car	1.21 (0.89-1.34)	0.33
Motorcycle	2.42 (1.87-3.12)	0.04
Public	1.68 (1.32-2.91)	0.12
Truck	1.89 (1.24-2.12)	0.03
Annual PM2.5# > 30	2.12 (1.65-3.22)	0.02
Cluster analysis		
High risk	5.69 (3.14-7.21)	<0.001
Low risk	1.23 (0.99-1.76)	0.14

Abbreviations, OR, odds ratio; CI, confidence interval; AP, air pollution; PM, particulate matter; *Average outdoor exposure per day; #Based on residential address for ten years

In Table 4, the independent association between the studied variables including the clusters proposed and lung cancer and specifically lung adenocarcinoma death

Table 4. Adjusted Hazard Ratios (HR) of Clinical Characteristics for Overall Survival at 52-weeks for Subjects Diagnosed with Lung Cancers

Variable	Unadjusted HR (95% CI)	p-value	Adjusted HR (95% CI)	p-value
All (n=514)				
CL: High vs low risk	2.42 (1.34-3.12)	<0.001	2.14 (1.12-2.77)	0.02
Gender (F vs M)	1.34 (1.06-1.82)	0.54	1.12 (0.62-1.52)	0.34
Smoking history	1.44 (1.18-1.98)	0.06	1.25 (1.12-1.75)	0.09
Stage (3/4 vs 1/2)	8.42 (6.72-9.14)	<0.001	7.12 (5.89-8.72)	<0.001
Treatment				
Surgery	1.12 (0.81-1.59)	0.14	0.95 (0.67-1.39)	0.24
Radiation	1.34 (1.21-1.70)	0.45	1.28 (1.12-1.74)	0.67
Chemotherapy	1.72 (1.32-2.42)	0.03	1.65 (1.26-1.98)	0.06
Adenocarcinoma (n=360)				
CL: High vs low risk	4.12 (2.56-5.12)	<0.001	3.89 (2.12-4.56)	0.02
Gender (F vs M)	1.12 (0.89-1.65)	0.24	0.87 (0.58-1.42)	0.32
Smoking history	1.46 (1.08-1.55)	0.09	1.36 (1.03-1.42)	0.23
Stage (3/4 vs 1/2)	6.11 (4.12-7.87)	<0.001	5.28 (3.12-6.99)	0.02
Treatment				
Surgery	0.89 (0.54-1.65)	0.45	0.91 (0.76-1.32)	0.48
Radiation	1.23 (0.89-1.65)	0.32	1.21 (0.87-1.52)	0.64
Chemotherapy	1.52 (1.32-1.89)	0.21	1.42 (1.23-1.74)	0.42

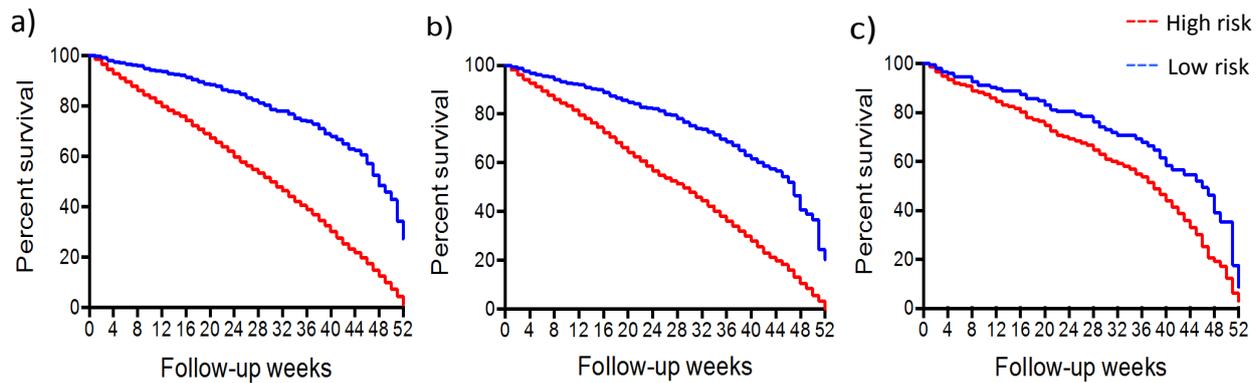


Figure 2. Kaplan Meier Curves for High and Low Risk Clusters for a) All Type of Lung Cancers, b) Lung Adenocarcinoma and c) Non-adenocarcinoma Subtypes

at 1 year (52-weeks) is shown. In the Cox regression analysis, only advanced stage of lung cancers (Stage 3 and 4) and CL I were associated with high risk of death for all lung cancers. Similar observations were found for lung adenocarcinoma death however the unadjusted and adjusted hazard ratios (HR) were much increased compared to lung cancer death from all types of lung cancers. For lung adenocarcinoma death, adjusted HR=3.89, 95% CI=2.12-4.56, $p=0.02$ for CL I vs CL II. Figure 2 shows the Kaplan-Meier curves of the high risk and low risk clusters for all lung cancer, lung adenocarcinoma and non-adenocarcinoma subtypes.

Discussion

This study for the first time used a cluster analysis to include the established PM_{2.5} annual exposure as a surrogate for air pollution, in addition to two most important anthropogenic factors: occupation and mode of transportation to determine the risk of lung cancers and lung adenocarcinoma independent of smoking status. This study also concurs with other previous but dated Malaysian study that lung adenocarcinoma is the most common type of lung cancers, diagnosed at a younger age with a female and/or non-smoker predominance compared to other type of lung cancers. The incidence of lung cancer especially with the histopathological shift towards lung adenocarcinoma in non- and never smokers are increasing, particularly in the Asian female population (Nakamura et al., 2013; Ha et al., 2014). Whilst lung cancer is the number one killer of all cancer death globally, lung cancer in non-smokers, if considered as a separate entity, would be ranked the seventh most deadly cancer, with an estimation of 200,000 lung cancer deaths globally (Sun et al., 2007). Therefore, there is high interest in the field to elucidate and stratify the risk for lung cancers from the perspective of ambient air pollution. The International Agency for Research on Cancer (IARC) agency has recently classified outdoor air pollution as well as PM air pollution as Group 1 carcinogens (IARC, 2013).

Despite robust epidemiological evidence, there are complex challenges in this research field to elucidate the association between air pollution and the risk for lung cancers. One of the major problems is the lack of adequate adjustment for potential confounders.

The most highlightable confounder is the lack of standard measurement to define the accurate amount of exposure to air pollution (Yoo et al., 2015). The estimated exposure of an individual is usually measured from the air pollution index derived from local air quality monitoring stations or residence/occupation in a heavily polluted city as proxies to estimate exposure to air pollution (Hansell et al., 2016). Nevertheless, two persons living in the same household might be exposed to different amount of air pollution depending on their daily activities despite the same residential address in which exposure estimation is usually derived. Indeed, the amount of ambient air exposure depends highly on activities and mode of transportation. For example, in the Klang Valley in which our institution is based, occupations such as stall hawkers and construction workers with daily ambient air exposure more than four hours are intuitively at a higher risk than another person who works in a closed office but lives in the same residential address.

From this study, we observed that subjects diagnosed with lung adenocarcinoma were more likely to be exposed to a higher annual PM_{2.5} concentrations than other type of lung cancers. We examined PM_{2.5} as a surrogate for air pollution exposure based on air-quality monitoring stations and residential address. In most locations, PM_{2.5} constitutes fifty to seventy percent of PM₁₀ and is associated more commonly with health effects (Weichental et al., 2013; Zanobetti et al., 2014). In addition, PM_{2.5} is more likely to be associated with the annual Southeast Asia haze that can remain in the atmosphere for days or weeks and thus be subject to long-range transboundary transport in the air (Amil et al., 2016). Studies have also shown the effects of PM_{2.5} on carcinogenicity in lung cancers (Liu et al., 2015; Yang et al., 2016; Huang et al., 2017) In the logistic regression analysis of this study, the likelihood for lung adenocarcinoma for subjects exposed to outdoor air pollution PM_{2.5} levels was twice as high as compared to other type of lung cancers. Therefore this study confirms previous observations that PM_{2.5} is a strong risk factor for the development of lung adenocarcinoma. This association was also suggested in a large multicenter European study (ESCAPE) which reported a strong association between air particulate matter and lung adenocarcinoma (Raschou-Nielson et al., 2013). The unique aspect of our study is that PM_{2.5} levels were

also clustered with two other important anthropogenic factors as opposed to PM_{2.5} concentrations derived from residential address.

A study sourced from the World Bank estimated that residents of the Greater Kuala Lumpur spend more than 250 million hours a year stuck in traffic (The World Bank, 2015). As roads and expressways are the major arteries for non-public transportation in the Klang Valley, the mode of transportation is likely to influence a person's daily ambient air pollution as well. For example, motorcycles as the main transport mode potentially impact on the total estimated exposure of air pollution for a motorcyclist compared to car drivers. We report here for the first time that motorcyclist has an increased risk of developing lung adenocarcinoma. While this may be confounded by geography in which the Greater Kuala Lumpur placed a higher risk for motorcyclists compared to areas with healthier air pollution indices, this association can be linked to a higher risk of exposure to air pollution as compared to subjects who drove cars or public transportation (Amil et al., 2016). Although associations have been found with nitric oxide (NO_x) and nitrogen oxide (NO₂) in previous studies, it seemed more likely that other highly correlated pollutants from traffic, such as particulate matter with absorbed polycyclic aromatic hydrocarbons and other genotoxic substances, are responsible for a possibly higher risk for lung cancer (Katanoda et al., 2011; Cao et al., 2011). Furthermore, numerous studies showed a higher risk for lung cancer among populations occupationally exposed to diesel engine exhaust (IARC Monograph, 1989; Lipsett et al., 1999). Indeed, one study has reported that motorcycles collectively emit sixteen times more hydrocarbons, three times more carbon monoxide and a 'disproportionately high' amount of other air pollutants (Vasic et al., 2006). Therefore, like truck drivers, the risk for motorcyclists is increased due to the risk of ambient air pollution in addition to the perpetual self-emission contributing to the individual's own ambient air quality. This study also observed that truck drivers have an increased risk of developing lung adenocarcinoma compared to other types of lung cancers. Studies on truck drivers have shown the molecular and epigenetic effects of particulate matter and ambient air exposure compared to controls (Garshick et al., 2008; Sanchez-Guerra et al., 2015). Again, this is likely due to longer daily exposure on the road and in a humid and hot country such as Malaysia, in addition to economic factors, most truck drivers do not use air-condition or filters and are likely to have their vehicle windows open.

We also identified that the high risk cluster of subjects had reduced survival at one-year compared to their counterpart. Although it is inappropriate to conclude from our findings that the biological plausibility of lung cancer mortality is directly associated with high-risk anthropogenic and socioeconomic factors relating to air pollution, this study concurs with previous larger studies that air pollution exposure affects lung cancer survival (Eckel et al., 2016). The carcinogenic relationship between air pollution components and survival is not clearly elucidated. Studies showing changes in expression of genes involved in oxidative stress, DNA damage and repair due

to exposure of air pollution have been reported (Demetriou et al., 2015; Yang et al., 2016). Several epidemiological studies have also shown a positive association between air pollutants to lung cancer risks and mortality (Katanoda et al., 2011; Xu et al., 2013; Eckel et al., 2016). Eckel et al., (2016) examined 352,053 patients with newly diagnosed lung cancer in California and observed reduced survival with higher average of air pollutant components such as nitric oxide and PM_{2.5} and 10. The authors reported that the largest air pollution HRs was for lung adenocarcinoma, the only common histological subtype of lung cancer that developed in a significant number of non-smokers, similar to our study's observations. The authors concluded that air pollution might reduce survival in the susceptible lung cancer subjects by impairing respiratory function.

The present study has several limitations. Firstly, the findings are restricted to a single institution with relatively small numbers as compared to other air pollution studies. Secondly, the clinical characteristics of lung cancers that present to a hospital located in a large urban city such as Kuala Lumpur may render the study less generalizable to other more rural areas with better air quality indices although, as a large teaching hospital in the country, a proportion of the subjects presented were from other regions and states as well (Amil et al., 2016). However, the attempt to use PM_{2.5} annual exposure to high and low risk concentrations dependent on the residential address might negate this limitation. We also acknowledge that the studied variables, though binary stratified, should be measured as continuous variables (age, smoking status, PM_{2.5}, exposure time of more and less than four hours). However, our main goal was to construct an easy-to-predict cluster of factors that can help in daily clinical practice and future research. In addition, occupational exposure was attributed retrospectively to subjects on the basis of their lifetime recall as collected at the interview. Nevertheless, as this study is the first of its kind, this initial observation renders it important as a platform for future studies in which future algorithms can be established to accurately identify higher risk of subjects exposed to ambient air pollution.

Future studies should include other important factors such as the use of personal air monitors and dietary habits that can affect epigenetics as studies have demonstrated strong associations between DNA modifications and consumption of fresh fruits and vegetables, olive oil and antioxidants (Sram et al., 2007; Raaschou-Nielson et al., 2011). In addition, most air pollution studies, both epidemiological and molecular, do not adjust for smoking status whether active or passive. Understandably, due to the complexity of the issue, there are no standard algorithms or biomarkers to evaluate individual's estimate air pollution exposure. This study for the first time identified a high risk cluster of variables that affects diagnosis of lung adenocarcinoma and lung cancer death. However, further validation of these clusters is welcomed in order to endorse or refute their clinical applicability.

Acknowledgements

The authors would like to acknowledge Prof. Dr.

Shamsul Azhar Shah, Department of Community Health for his expert advice on statistics.

Funding Statement

This project does not received any funding from any government and non-government body.

References

- Amil N, Latif MT, Khan MF, Mohammad M (2016). Seasonal variability Of PM2.5 compstion and sources in the Klang valley urban-industrial environment. *Atmos Chem Phys*, **16**, 5357-1.
- Beeson WL, Abbey DE, Knutsen SF (1998). Long-term concentrations of ambient air pollutants and incident lung cancer in California adults: Results from the Ahsmog study. *Environ Health Perspect*, **106**, 813-2.
- Cao J, Yang C, Li J, et al (2011). Association between long-term exposure to outdoor air pollution and mortality in China: A cohort study. *J Hazard Mater*, **186**, 1594-0.
- Chen F, Cole P, Bina W (2007). Time trend and geographic patterns of lung adenocarcinoma in the United States, 1973- 2002. *Cancer Epidemiol Biomarkers Prev*, **16**, 2724-9.
- Demetriou CA, Vineis P (2015). Carcinogenicity of ambient air pollution: Use of biomarkers, lesson learnt and future directions. *Thorax*, **7**, 67-5.
- Eckel SP, Cockburn M, Shu YH, et al (2016). Air pollution affects lung cancer survival. *Thorax*, **71**, 891-8.
- Garshick E, Laden F, Hart JE, et al (2008). Lung cancer and vehicle exhaust in trucking industry workers. *Environ Health Perspect*, **116**, 1327-2.
- Ha SY, Choi SJ, Cho JH, et al (2014). Lung cancer in never-smoker Asian females is driven by oncogenic mutations, most often involving EGFR. *Oncotarget*, **6**, 5465-4.
- Hansell A, Ghosh RE, Blangiardo M, et al (2016). Historic air pollution exposure and long-term mortality risks in England and Wales: Prospective longitudinal cohort study. *Thorax*, **71**, 330-8.
- Huang F, Pan B, Wu J, Chen E, Chen L (2017). Relationship between exposure to Pm2.5 and lung cancer incidence and mortality: A meta-analysis. *Oncotarget*, **8**, 43322-1.
- IARC (2013). Air Pollution and Cancer, International Agency for Research on Cancer (2013).
- IARC (1989). Monographs on the evaluation of carcinogenic risks to humans. Diesel and gasoline engine exhausts and some nitroarenes. International Agency for Research on Cancer. *IARC Monogr Eval Carcinog Risks Hum*, **46**, 1-8.
- Katanoda K, Sobue T, Satoh H, et al (2011). An association between long-term exposure to ambient air pollution and mortality from lung cancer and respiratory diseases in Japan. *J Epidemiol*, **21**, 132-3.
- Liam CK, Pang YK, Leow CH, Poosparajah S, Menon A (2006). Changes in the distribution of lung cancer cell types and patient demography in a developing multiracial Asian country: Experience from a Teaching Hospital. *Lung Cancer*, **53**, 23.
- Lipsett M, Campleman S (1999). Occupational exposure to diesel exhaust and lung cancer: A meta-analysis. *Am J Public Health*, **89**, 1009-7.
- Liu C, Guo H, Cheng X, et al (2015). exposure to airborne PM2.5 suppresses Microrna expression and deregulates target oncogenes that cause Neoplastic transformation in NIH3T3 cells. *Oncotarget*, **6**, 29428-9.
- Malaysia Economic Monitor, The World Bank. June 2015.
- nNafstad P, Håheim LL, Oftedal B, et al (2003). Lung cancer and air pollution: A 27 year follow up of 16,209 norwegian men. *Thorax*, **58**, 1071-6.
- Nakamura H, Saji H (2014). Worldwide trend of increasing primary adenocarcinoma of the lung. *Surg Today*, **44**, 1004-2.
- Raaschou-Nielsen O, Andersen ZJ, Hvidberg M, et al (2011). Lung cancer incidence and long-term exposure to air pollution from traffic. *Environ Health Perspec*, **119**, 860-5.
- Raaschou-Nielsen O, Andersen ZJ, Beelen R, et al (2013). Air pollution and lung cancer incidence in 17 European Cohorts: Prospective analyses from the European study of cohorts for air pollution effects (ESCAPE). *Lancet Oncol*, **14**, 813-2.
- Sram RJ, Beskid O, Rossnerova A, et al (2007). Environmental Exposure to carcinogenic polycyclic Aromatic Hydrocarbons-the interpretation of cytogenetic analysis by fish. *Toxicol Lett*, **172**, 12-0.
- Sanchez-Guerra M, Zheng Y, Osorio-Yanez C, et al (2015). Effects of particulate matter exposure on blood 5-Hydroxymethylation: Results from the Beijing truck driver air pollution study. *Epigenetics*, **10**, 633-2.
- Sun S, Schiller JH, Gazdar AF (2007). Lung cancer in never smokers-A different disease. *Nat Rev Cancer*, **7**, 778.
- Vasic AM, Weilenmann M (2006). Comparison of real-world emissions from two-wheelers and passenger cars. *Environ Sci Technol*, **40**, 149-4.
- Weichenthal SA, Godri-Pollitt K, Villeneuve PJ (2013). Pm2.5, oxidant defence and cardiorespiratory health: A review. *Environ Health*, **12**, 40.
- Xu X, Ha S, Kan H, et al (2013). Health effects of air pollution on length of respiratory cancer survival. *BMC Public Health*, **13**, 800.
- Yang, Chen D, Zhao H, Xiao C (2016). The effects for PM2.5 exposure on non-small-cell lung cancer induced motility and proliferation. *SpringerPlus*, **5**, 2059.
- Yoo EH, Rudra C, Glasgow M, Mu L (2015). Geospatial estimation of individual exposure to air pollutants: Moving from static monitoring to activity-based dynamic exposure assessment. *Ann Am Assoc Geogr*, **105**, 915-6.
- Zanobetti A, Dominici F, Wang Y, Schwartz JD (2014). a National case-crossover analysis of the short-term effect of PM2.5 on hospitalizations and mortality in subjects with diabetes and neurological disorders. *Environ Health*, **3**, 38.



This work is licensed under a Creative Commons Attribution-Non Commercial 4.0 International License.