

## Original Article

# Decision Tree of Occupational Lung Cancer Using Classification and Regression Analysis

Tae-Woo KIM<sup>1</sup>, Dong-Hee KOH<sup>1</sup> and Chung-Yill PARK<sup>2</sup><sup>1</sup>Occupational Safety and Health Research Institute, Korea Occupational Safety and Health Agency, Incheon  
<sup>2</sup>Department of Preventive Medicine, College of Medicine, The Catholic University of Korea, Seoul, Korea

**Objectives:** Determining the work-relatedness of lung cancer developed through occupational exposures is very difficult. Aims of the present study are to develop a decision tree of occupational lung cancer.

**Methods:** 153 cases of lung cancer surveyed by the Occupational Safety and Health Research Institute (OSHRI) from 1992-2007 were included. The target variable was whether the case was approved as work-related lung cancer, and independent variables were age, sex, pack-years of smoking, histological type, type of industry, latency, working period and exposure material in the workplace. The Classification and Regression Test (CART) model was used in searching for predictors of occupational lung cancer.

**Results:** In the CART model, the best predictor was exposure to known lung carcinogens. The second best predictor was 8.6 years or higher latency and the third best predictor was smoking history of less than 11.25 pack-years. The CART model must be used sparingly in deciding the work-relatedness of lung cancer because it is not absolute.

**Conclusion:** We found that exposure to lung carcinogens, latency and smoking history were predictive factors of approval for occupational lung cancer. Further studies for work-relatedness of occupational disease are needed.

**Key Words:** Occupational lung cancer, Decision tree, Latency, Smoking, CART

## Introduction

Worldwide, lung cancer is the leading cause of death by cancer [1,2] and most common cancer among occupational related cancers [3-5]. Approximately 90% of men and 60% of women developing lung cancer are smokers [6]. According to many studies, work-related lung cancer occupies 5-20% of total lung cancer cases, a wide variance [7-9]. However, because work-related lung cancer is differentiated from other lung cancers only by occupational exposure, not by histological morphology, pathologic form, clinical manifestation or outcome, distinction

between occupational and non-occupational lung cancer is very difficult [10]. Therefore, doctors specializing in occupational and environmental medicine are typically consulted when compensating for occupational lung cancer [10].

Different approval rules for occupational diseases exist in each nation. In Korea, approval and compensation of occupational diseases are prescribed by Article 34 of the Industrial Accident Compensation Insurance Act. The work-relatedness of disease is decided by medical experts considering exposure experience and duration. The decision standards of work-relatedness for lung cancer in this Article included: worker exposed to tar and worked more than 10 years; worker exposed to chrome or chrome compounds and worked more than 2 years and worker exposed to asbestos. However, these standards were only applicable to industrial workers exposed to carcinogens such as crystalline silica [11], nickel [12], cadmium [13], arsenic [14], working in industries such as aluminum production [15], iron and steel founding [16], and rubber industry [17]. There-

**Received:** April 29, 2010, **Accepted:** September 16, 2010

**Correspondence to:** Chung-Yill PARK

Department of Preventive Medicine, College of Medicine, The Catholic University of Korea

62, Yeouido-dong, Yeongdeungpo-gu, Seoul 150-713, Korea

**Tel:** +82-2-3779-1405, **Fax:** +82-2-782-6017

**E-mail:** cypark@catholic.ac.kr

fore, because the determination of work-relatedness involves vague standards, many specialists' decisions may be subjective.

In Korea Workers' Compensation and Welfare Service (COMWEL), which decides work-relatedness of diseases, clinical and occupational doctors formed the Occupational Disease Award Commission to determine work-relatedness. If a work-relatedness decision was very difficult, COMWEL sent the case to Occupational Safety and Health Research Institute (OSHRI) to perform an epidemiologic investigation. The OSHRI survey team, comprised of occupational and environmental medicine specialists and hygienists, surveyed the workplace environment for probable exposure materials via material safety data sheets or analysis of bulk samples, job history, general characteristics of the worker and family history. If possible, this team measured exposure concentration, but in many cases, measurement was not performed. They also extensively reviewed literature related with each case. The Occupational Disease Specialist Commission formed by university professors of occupational medicine and researchers of OSHRI discussed the work-relatedness of each case. After this discussion, researchers sent brief survey reports including the commission's recommendation for work-relatedness to COMWEL. Finally, COMWEL decided the work-relatedness of each case taking into consideration OSHRI's recommendation.

In Korea, because only workers approved as having an occupational disease are compensated by COMWEL, the decision of work-relatedness is very important to workers and COMWEL. In spite of this importance, comprehensive studies on this decision-making process are very sparse. The aim of the present study is to find variables affecting the work-relatedness of lung cancer and to derive a decision tree of occupational lung cancer to help objectively decide work-relatedness.

## Materials and Methods

### Materials

From 1992 to 2007, 160 cases of lung cancer were referred to COMWEL for determination of work-relatedness. Among them, 6 cases postponing their decision and 1 case not having brief report were excluded. Therefore, 153 cases were included in this study.

### Methods

Each of OSHRI's survey reports for lung cancer was reviewed, and 15 variables that probably affected decisions on occupational lung cancer were found; however, only 8 of them were included in this study. These variables were work-relatedness of lung cancer (occupational vs non-occupational lung cancer),

demographic characteristics (sex, age diagnosed with lung cancer), health habit (smoking history), clinical factors (histological classification) and occupational factors (industry classification, exposed or probably exposed materials, latent period and working period). Occupational lung cancer was defined as primary lung cancer related with work. Industries were classified by International Standard Industrial Classification (ISIC Rev. 4) and divided into manufacturing and non-manufacturing industries. One major potential lung carcinogen to humans among several exposed or probable exposed materials was selected as the exposure material. By International Agency for Research on Cancer (IARC) classification, these exposure materials were classified as known (carcinogenic, IARC group 1) or unknown (probably, possibly and not carcinogenic, IARC group 2A, 2B and 4) carcinogen to lung. When cases of exposure to diesel engine exhaust (DEE) were reviewed, it was found that polycyclic aromatic hydrocarbons (PAHs) included in DEE [18] had a greater effect on the decision of occupational lung cancer than only DEE. Therefore, although DEE was IARC group 2A [19], DEE as PAHs was classified as IARC group 1 for this study. Latency was defined as the time from the date the worker was first exposed to exposure materials to diagnosis date. If the exposure period was not accurate, latency was defined as the period from the date first worked to diagnosis date. Working period was the period employed in the workplace that workers were exposed to known or unknown lung carcinogens.

### Statistical analysis

Age diagnosed with lung cancer, smoking history, latency and working period were analyzed by independent t-test. Sex, histological classification and industry classification were analyzed by chi-square test.

For deriving the occupational lung cancer prediction tree, Classification and Regression Tree (CART) analysis was used. CART analysis is a tree-building technique and suited to the generation of clinical decision rules [20,21]. It is a non-parametric technique that can handle highly skewed or multi-modal numerical data, as well as categorical predictors [22]. It can uncover complex interactions between predictors that are most important in determining the outcome variable to be explained [23]. The CART method is a relatively automatic machine learning method, which produces decision trees that are easily to interpret [20].

In CART analysis, the decision of occupational lung cancer was selected as the target variable and 7 variables were included as confounding variables. For validation of the tree, objects were divided randomly into training (122 cases, 80%) and validation data (31 cases, 20%) by data partition. The decision

tree was derived from training data and evaluated by validation data. A multivariate logistic regression model was constructed from training data and tested in the validation data. For comparisons of accuracy of CART and the logistic regression model, area under receiver operating characteristic curves (AUCs) of two models was used in training and validation data.

## Results

### General characteristics and exposure materials

Among the 153 cases were 57 decisions of occupational lung cancer cases (37.3%) and 96 decisions of non-occupational lung cancer (62.7%). Mean age of the two groups was statistically similar (50.8 vs 50.3 years). Males comprised a greater proportion of occupational lung cancer than non-occupational lung cancer, but the difference was not significant ( $p = 0.620$ ). Smoking history in non-occupational lung cancer was approximately 5 pack-years more than occupational but not significant ( $p = 0.089$ ). Latency in occupational lung cancer was approximately 7 years longer than in non-occupational lung cancer ( $p <$

0.001), and the working period of occupational lung cancer was 6 years longer ( $p < 0.001$ ). According to industrial classification, proportion of occupational lung cancer in non-manufacturing industry was significantly higher than manufacturing industry (53.1% vs 29.8%,  $p = 0.003$ ). According to histological classification, adenocarcinoma was most common followed by squamous cell carcinoma and small cell lung carcinoma (Table 1).

In this study, there were 24 kinds of exposure material. Exposure cases to asbestos were most common (45 cases, 29.4%), followed by crystalline silica (21 cases, 13.7%) and dust (18 cases, 11.8%). Known lung carcinogens included asbestos, crystalline silica, hexavalent chromium (Cr(VI)), PAHs, Radon and Cokes oven Emission (COE). There were 92 cases of exposure to known carcinogens. While cases exposed to Cr(VI) had the highest proportion of occupational lung cancer (71.4%), workers exposed to PAHs had the lowest proportion (50%). There was only 1 case each of Radon and COE exposure (Table 2).

### CART model

In the training model, the first prediction factor of decision

**Table 1.** Descriptive characteristics of the survey populations by occupational and non-occupational lung cancer

Variables	Occupational lung cancer (n = 57)	Non-occupational lung cancer (n = 96)	p-value
Age (Mean $\pm$ SD, years)	50.8 $\pm$ 7.7	50.3 $\pm$ 8.7	0.728
Gender (%)			0.620
Male	55 (39.6)	84 (60.4)	
Female	2 (14.3)	12 (85.7)	
Smoking (pack-years)	12.2 $\pm$ 14.7	17.3 $\pm$ 17.1	0.089
Latent period <sup>‡</sup> (years)	22.8 $\pm$ 8.7	15.5 $\pm$ 9.5	< 0.001
Working period <sup>§</sup> (years)	21.0 $\pm$ 7.3	14.7 $\pm$ 9.1	< 0.001
Industry (%)			0.003
Manufacturing	30 (29.8)	73 (70.2)	
Non-manufacturing	26 (53.1)	20 (46.9)	
Histological type (%)			0.754
Adenocarcinoma	22 (37.9)	36 (62.1)	
SCC*	16 (37.2)	27 (62.8)	
SCLC <sup>†</sup>	5 (33.3)	10 (66.7)	
Others	2 (66.7)	1 (33.3)	

\*SCC: squamous cell carcinoma.

<sup>†</sup>SCLC: small cell lung cancer.

<sup>‡</sup>Latent period: from the date that worker was first exposed by exposure materials to diagnosis date.

<sup>§</sup>Working period: the period employed in workplace that workers were exposed to known or unknown lung carcinogen.

for occupational lung cancer was carcinogenesis to lung of exposure materials. Forty-seven cases (62.7%) among 75 cases exposed to known lung carcinogens were determined to be occupational lung cancer. All cases exposed to unknown lung carcinogens were decided as lung cancer developing by non-occupational causes. Among cases exposed to known lung carcin-

ogens, the second predictor was latency of 8.6 years or higher. Forty-seven (71.2%) among 66 cases with a latency period of 8.6 years or higher were decided as occupational lung cancer. All cases with a latency of less than 8.6 years were decided as non-occupational. In cases exposed to known lung carcinogens and having a latency of 8.6 years or higher, smoking less than

**Table 2.** Exposed material during work in the workplace

( ) = %

Exposure material	IARC group	Carcinogen to lung	Occupational lung cancer (n = 57)	Non-occupational lung cancer (n = 96)	Total
Asbestos	1	Known	29 (64.4)	16 (35.6)	45 (29.4)
Crystalline silica			12 (57.1)	9 (42.9)	21 (13.7)
PAHs*			5 (50.0)	5 (50.0)	10 (6.5)
Cr(IV) <sup>†</sup>			8 (80.0)	2 (20.0)	10 (6.6)
Lead Chromate			2 (50.0)	2 (50.0)	4 (2.6)
Radon			0 (0.0)	1 (100)	1 (0.7)
COE <sup>‡</sup>			1 (100)	0 (0.0)	1 (0.7)
VCM <sup>§</sup>		Unknown	0 (0.0)	2 (100)	2 (1.3)
Radiation			0 (0.0)	2 (100)	2 (1.3)
Formaldehyde			0 (0.0)	1 (100)	1 (0.7)
Lead	2A		0 (0.0)	2 (100)	2 (1.3)
MMVF <sup>  </sup>	2B		0 (0.0)	2 (100)	2 (1.3)
Hydrochloric acid			0 (0.0)	1 (100)	1 (0.7)
Carbon black			0 (0.0)	1 (100)	1 (0.7)
Dust	3		0 (0.0)	18 (100)	18 (11.8)
Organic solvents			0 (0.0)	8 (100)	8 (5.2)
Welding fume			0 (0.0)	3 (100)	3 (2.0)
Paints			0 (0.0)	2 (100)	2 (1.3)
Dye			0 (0.0)	1 (100)	1 (0.7)
Harmful gas			0 (0.0)	1 (100)	1 (0.7)
MWF <sup>¶</sup>			0 (0.0)	1 (100)	1 (0.7)
Tar			0 (0.0)	1 (100)	1 (0.7)
Cr(III)**			0 (0.0)	11 (100)	11 (7.1)
Missing			0 (0.0)	4 (100)	4 (2.6)
Total			57 (37.3)	96 (62.7)	153 (100)

\*PAHs: polycyclic aromatic hydrocarbons.

<sup>†</sup>Cr(IV): hexavalent chromium.

<sup>‡</sup>COE: coke oven emission.

<sup>§</sup>VCM: vinyl chloride monomer.

<sup>||</sup>MMVF: man-made vitreous fiber.

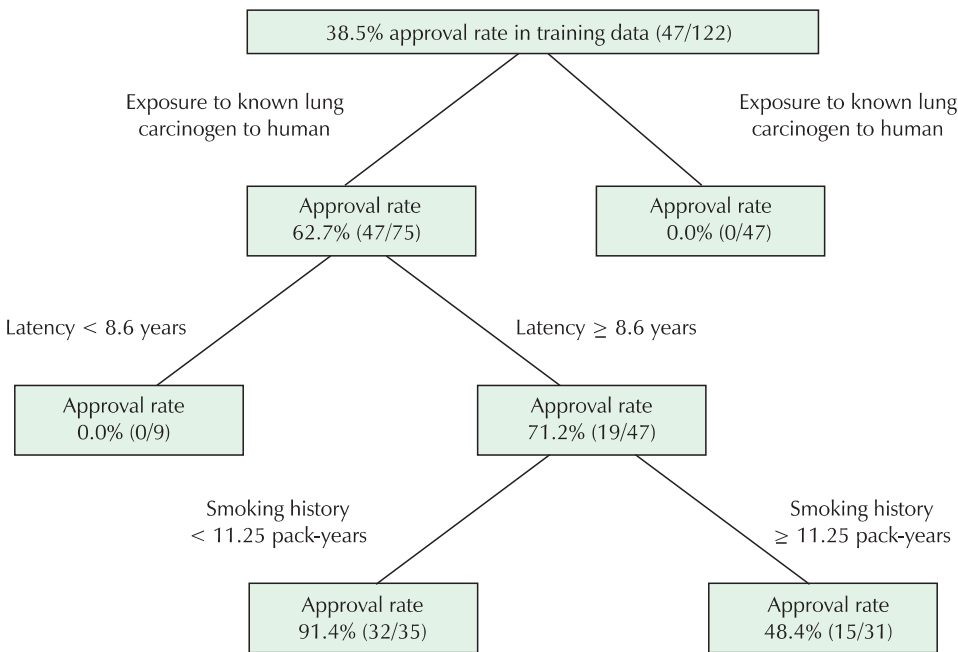
<sup>¶</sup>MWF: metal working fluid.

\*\*Cr(III): trivalent chromium.

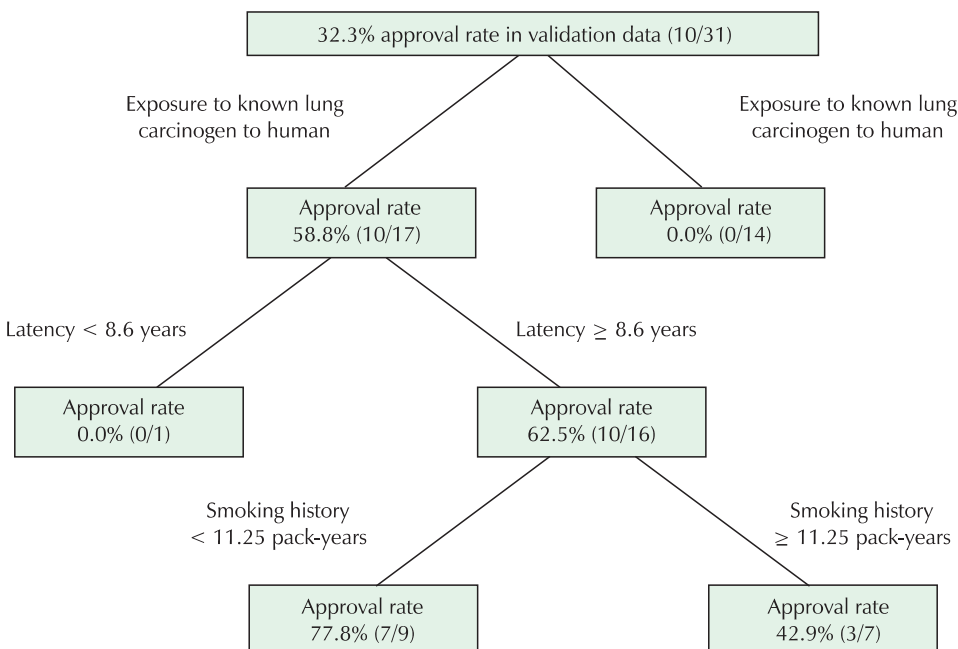
11.25 pack-years provided additional prediction value. Thirty-two (91.4%) of 35 cases were decided as occupational lung cancer. Fifteen (48.4%) of 31 cases having 11.25 pack-years or higher were decided as lung cancer developed by occupational exposure (Fig. 1).

Applying the rule of training data in the validation tree, 17 (58.8%) among 31 cases exposed to known lung carcinogens were decided as occupational lung cancer. Ten (62.5%) of 16

cases exposed to known lung carcinogens and having a latency of 8.6 years or higher were decided as occupational lung cancer. In 9 cases exposed to known lung carcinogens, having 8.6 years or higher latency and smoking less than 11.25 pack-years, 7 cases (77.8%) were decided as occupational lung cancer (Fig. 2). In logistic regression, carcinogenesis to lung of exposure materials, smoking amount and working period were analyzed as significant variables. With an increase of 1 pack-year, work-



**Fig. 1.** Predictors of occupational lung cancer in training data. Each node is based on available data for each predictive variable. Each approval rate for each predictor is marked in box. Each predictor was written beside line.



**Fig. 2.** Decision tree of occupational lung cancer applied in validation data. Each node is based on available data for each predictive variable. Each approval rate for each predictor is marked in box. Each predictor was written beside line.

relatedness for occupational lung cancer decreased 0.94 times. Increasing one year of the working period increased work-relatedness 1.15 times.

Based on the AUCs, the accuracy of the CART model (0.914) was more than that of the logistic regression model (0.824).

## Discussion

When deciding the work-relatedness of lung cancer, 4 steps should be considered: whether or not the worker was exposed to lung carcinogen or worked in an industry known for development of lung cancer, exposure dose of lung carcinogen, sufficient latency period for developing lung cancer and smoking amount [10]. In the CART model used in this study, 3 of the 4 steps were included (not exposure dose) and the accuracy of this model was very high.

Our result that occupational exposure by known lung carcinogen was the first predictor of an occupational lung cancer decision was not surprising. Problems remain when deciding the work-relatedness of lung cancer in workers exposed to probable lung carcinogens such as man-made vitreous fibers [24]. Studies investigating the proportion of occupational cancer have included workers exposed to probable or possible carcinogens, as well as known carcinogen to humans [7-9]. In this model, only workers exposed to known lung carcinogens were determined to have lung cancer by occupational exposure. In Korea, if workers exposed to possible lung carcinogens request COMWEL to determine the work-relatedness of their lung cancer, the result may be negative.

In Korea, the work-relatedness of lung cancer miners is determined by other laws for miners; therefore, we had only 1 survey report for workers exposed to radon. Known carcinogens to humans included in this model were asbestos [25], crystalline silica [11], PAHs [26], Cr(IV) [27] and COE [28]. It is known that these materials have a synergic effect for the development of lung cancer [29]. Because we included only 1 major lung carcinogen, these interactions were ignored. In each survey report, synergic effects of exposure materials actually showed a partially positive effect on the work-relatedness decision, but the effect was unclear. The interaction acted as an additional reason for a decision in workers strongly suspected of having occupational lung cancer.

In the CART model, sufficient latency in the work-relatedness decision was 8.6 years or higher. Latency of a solid tumor is generally 10-12 years [30]. The cut point of latency in our result was shorter because a worker exposed to asbestos for 9 years was decided as having occupational lung cancer. Except

for 1 case exposed to asbestos, latency in them was higher than 10 years. Purposes of CART were accurate prediction and classification. Therefore, latency values of 8.6 year or higher in this model should be estimated to exclude the 9 cases of non-occupational lung cancer. This result was a disadvantage of CART that prediction error should have occurred around the splitting point of interval variables [31].

In smoking studies using CART, high smoking quantity increased risk of lung cancer [32]. A dose-response relationship between smoking and lung cancer has been found, and the risk of lung cancer in the smoker group in the first two decades of smoking showed twice the odds ratio (OR) when compared to the non-smoker group [33]. In another study, lung cancer risk in the smoker group smoking more than 30 years increased than when smoking for less than 30 years [34]. In this study, most workers smoked less than 11.25 pack-years were classified as occupational lung cancer. This result could be interpreted as the carcinogenic effect of smoking was extremely important in the decision process for smokers with low smoking quantity or non-smokers. In cases that were definitely exposed to carcinogen and not exposed to environmental materials, such as asbestos in slate, were decided as occupational lung cancer.

Considering the many causes of lung cancer, the synergic effect between smoking and occupational exposure to carcinogens has been shown [35]. In this model, this effect was found in workers smoking more than 9.75 pack-years. Although many studies supported this interaction [36,37], many studies have shown contradictory results [38,39]. Therefore the interaction between smoking and carcinogens is indefinite [40]. According to Hertz-Picciotto, proof that arsenic and smoking have an synergistic interaction in the development of lung cancer is compelling [39]. In present study, low smoking amount had a favorable effect on the work-related evaluation, while a high smoking amount did not have any clear implication.

Asbestos is carcinogenic in human lungs, but 16 of 45 cases exposed to asbestos were decided as non-occupational lung cancer. In the review of 16 cases, 8 cases had a short working period of less than 8 years or irregular exposure and 4 cases were heavy smokers. In others, we suspected exposure to asbestos, but could not find certain evidence of exposure in bulk samples or work environment evaluation.

The KOSHA team performed exposure assessment (personal and area airborne asbestos fiber concentration) in 11 of 45 cases exposed to asbestos, and 6 cases of the 11 were occupational lung cancer. Nine of the 11 cases were estimated by Time-Weighted Average (TWA) and 2 cases during short time. In cases decided as occupational lung cancer, chrysotile and tremolite were found in airborne samples. In cases estimated

by TWA, airborne personal asbestos concentration was 0.003-0.4 fiber/cc, and area concentration was 0.002 fiber/cc. In cases estimated during 14 minutes, airborne personal and area concentration were 0.4 and 0.13 fiber/cc. In cases decided as non-occupational lung cancer, although chrysotile and tremolite were found, asbestos fiber concentration was very low (0 or 0.005 fiber/cc). In only a short time sample, personal and area asbestos concentration were 0.027 fiber/cc (22 minutes) and 0.159 fiber/cc (21 minutes). Although this case was exposed to high concentrations of asbestos, the KOSHA team determined it was non-occupational lung cancer due to short latency and the nature of his duties (supervising construction work) (Table 3). According to the asbestos exposure assessment, we know that low asbestos fiber concentration affects the decision of work-relatedness. However, we could not find asbestos exposure concentration standards for determining work-relatedness because many cases exposed to asbestos had no estimate of exposure concentration.

Although both latency and working period were statistically significant in t-test for all cases, only latency was included. This result might be interpreted that latency was more important than working period in approval of occupational lung cancer. Also, the latency in our model could already reflect the

effect of the working period.

In the chi-square test for industry classification of all cases, a significant increase in occupational lung cancer was found in the non-manufacturing industry compared with the manufacturing industry. In the CART model, industry classification was not included. When all cases were classified according to the first predictor (occupational exposure by known lung carcinogen), significance of industry classification disappeared. This result revealed the characteristic of CART analysis presenting decision processes for classifying patterns.

Our study has several limitations. First, occupational decision by OSHRI may be inaccurate. For decision accuracy, a commission composed of 2 professors of occupational and environmental medicine, 11 researchers (9 specialists of occupational and environmental medicine and 2 hygienists) and several residents of OSHRI discussed and decided work-relatedness of diseases. Therefore, most of OSHRI's decisions may be relatively reliable. Second, this study was based on only Korean data. Each nation has their own rules and social consensus for occupational decisions. Therefore, our results may only be applicable in Korea. Third, because of the small sample size, the prediction cut point was unstable. Although the accuracy of this model was very high, it could be due to

**Table 3.** Industry, latent period, working period, fiber type, airborne asbestos concentration and work-relatedness in cases measured by personal and area sampling

Case	Industry	Latent period (years)	Working period (years)	Fiber type	Airborne concentration* (fiber/cc)		Work-relatedness
					Personal sample	Area sample	
1	Non-manufacturing	18	18	Chrysotile	< 0.01		Occupational
2	Manufacturing	21	21	Chrysotile	0.4		Occupational
3	Manufacturing	21	21	Chrysotile	0.7 (14 min <sup>†</sup> )	0.13 (14 min <sup>†</sup> )	Occupational
4	Manufacturing	50	50		0	0	Non-occupational
5	Non-manufacturing	3.5	12	Tremolite	0.003	0	Occupational
		8.5		Chrysotile	0.002	0.002	
6	Manufacturing	15	21	Tremolite	0.005	0	Non-occupational
		6					
7	Non-manufacturing	11	11	Chrysotile	0.003		Occupational
8	Non-manufacturing	9.5	9.5	Chrysotile	0	0.003	Non-occupational
9	Manufacturing	5.5	21	Chrysotile	0.027 (22 min <sup>†</sup> )	0.159 (21 min <sup>†</sup> )	Non-occupational
10	Manufacturing	28	28	Chrysotile	0.055		Occupational
11	Non-manufacturing	8	8		0	0	Non-occupational

\*Airborne concentration: estimated Time-Weighted Average (TWA) except for case 3 and 9.

<sup>†</sup>Min: measured minutes.

small objects. Minimizing the instability, analysis was repeatedly performed for an optimal tree. Fourth, exposure dose of the worker was not included in this analysis. In Korea, estimation of hazardous materials exposure - except for dust - in the workplace started after late 1990; therefore, exposure data from before this does not exist. We presumed, qualitatively, past exposure dose according to many studies done in other nations. We estimated recent exposure dose and collected current exposure data from workers' companies. However, because most of them were very low, we did not use them for occupational decisions. Absence of past exposure dose may be the starting point of this study. If exposure estimation were performed, this study would not be needed. Although not completed, we thought that the working period could partially estimate occupational exposure. However, the working period was not included in the decision tree. Fifth, selection bias may have existed because this study included only cases reviewed by COMWEL. Up to 2007, most cases were decided by OSHRI, except for the lung cancer cases among miners that were decided by COMWEL. Therefore, this bias may be low. Finally, we did not carry out cross-validation with other data. Our data was unique occupational decision data in Korea. Therefore, we do not know the accurate validity of these results.

In conclusion, we found that exposure to lung carcinogen, latency and smoking history were prediction factors of approval for occupational lung cancer. This decision tree must be considered as a minimal decision standard of work-relatedness for lung cancer, because doctors that decide work-relatedness must take into account the intricacies of each case. To make accurate decision standards for occupational lung cancer, additional studies for elevating validation have to be performed.

## Acknowledgements

This work was supported by the Korean Occupational Safety and Health Agency (KOSHA), Occupational Safety and Health Research Institute.

## References

- Alberg AJ, Ford JG, Samet JM. American College of Chest Physicians. Epidemiology of lung cancer: ACCP evidence-based clinical practice guidelines. 2nd ed. Chest 2007;132:29S-55S.
- Ferlay J, Autier P, Boniol M, Heanue M, Colombet M, Boyle P. Estimates of the cancer incidence and mortality in Europe in 2006. Ann Oncol 2007;18:581-92.
- Doll R, Peto R. The causes of cancer: quantitative estimates of avoidable risks of cancer in the United States today. J Natl Cancer Inst 1981;66:1191-308.
- Driscoll T, Nelson DI, Steenland K, Leigh J, Concha-Barrientos M, Fingerhut M, Prüss-Ustün A. The global burden of disease due to occupational carcinogens. Am J Ind Med 2005;48:419-31.
- Steenland K, Burnett C, Lalich N, Ward E, Hurrell J. Dying for work: the magnitude of US mortality from selected causes of death associated with occupation. Am J Ind Med 2003;43:461-82.
- Peto R, Lopez AD, Boreham J, Thun M, Heath C Jr. Mortality from smoking in developed countries 1950-2000: indirect estimates from national vital statistics. Oxford (UK): Oxford University Press; 1994.
- Boffetta P, Kogevinas M. Introduction: epidemiologic research and prevention of occupational cancer in Europe. Environ Health Perspect 1999;107:229-31.
- Fingerhut M, Nelson DI, Driscoll T, Concha-Barrientos M, Steenland K, Punnett L, Prüss-Ustün A, Leigh J, Corvalan C, Eijkemans G, Takala J. The contribution of occupational risks to the global burden of disease: summary and next steps. Med Lav 2006;97:313-21.
- Kim JI, Kim JH, Kang D, Kim JW, Kim JE, Ahn JH, Lee CH, Lee HJ, Kang JU, Son JK, Sul JK, Kim YK, Jung KY, Kim JY. Epidemiologic characteristics of occupational lung cancer in the Busan area. Korean J Occup Environ Med 2006;18:53-8.
- Rosenstock L, Cullen MR, Brodtkin CA, Redlich CA. Textbook of clinical occupational and environmental medicine. 2nd ed. Philadelphia (PA): Elsevier Saunders Pub; 2005. p. 727-41.
- Steenland K, Sanderson W. Lung cancer among industrial sand workers exposed to crystalline silica. Am J Epidemiol 2001;153:695-703.
- Grimsrud TK, Berge SR, Martinsen JI, Andersen A. Lung cancer incidence among Norwegian nickel-refinery workers 1953-2000. J Environ Monit 2003;5:190-7.
- Verougstraete V, Lison D, Hotz P. Cadmium, lung and prostate cancer: a systematic review of recent epidemiological data. J Toxicol Environ Health B Crit Rev 2003;6:227-55.
- Järup L, Pershagen G, Wall S. Cumulative arsenic exposure and lung cancer in smelter workers: a dose-response study. Am J Ind Med 1989;15:31-41.
- Björ O, Damber L, Edström C, Nilsson T. Long-term follow-up study of mortality and the incidence of cancer in a cohort of workers at a primary aluminum smelter in Sweden. Scand J Work Environ Health 2008;34:463-70.
- Smith GH, Williams FL, Lloyd OL. Respiratory cancer and air pollution from iron foundries in a Scottish town: an epidemiological and environmental study. Br J Ind Med 1987;44:795-802.
- Kogevinas M, Sala M, Boffetta P, Kazerouni N, Kromhout H, Hoar-Zahm S. Cancer risk in the rubber industry: a review



- of the recent epidemiological evidence. *Occup Environ Med* 1998;55:1-12.
18. Wichmann HE. Diesel exhaust particles. *Inhal Toxicol* 2007;19:241-4.
  19. IARC Monograph on the evaluation of the carcinogenic risks to humans. Volume 46 diesel and gasoline engine exhausts and some nitroarenes. Lyon (France): International Agency for Research on Cancer (IARC); 1989. 458 p.
  20. Lewis RJ. An introduction to classification and regression tree (CART) analysis [Internet]. 2000 [cited 2010 Apr 12]. Available from: <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.95.4103&rep=rep1&type=pdf>.
  21. Timofeev R. Classification and regression trees (cart) theory and applications: Master Thesis. Berlin (Germany): Humboldt University; 2004. p. 7-8.
  22. Breiman L, Friedman J, Olshen RA, Stone CJ. Classification and regression trees. 1st ed. New York: Chapman & Hall; 1998. p. 18-22.
  23. Steinberg D, Colla P. CART--classification and regression trees. 1st ed. San Diego (CA): Salford Systems; 1997. p. 180-1.
  24. IARC Monograph on the evaluation of the carcinogenic risks to humans. Volume 81 man-made vitreous fibres. Lyon (France): International Agency for Research on Cancer; 2002. 418 p.
  25. Berman DW, Crump KS. A meta-analysis of asbestos-related cancer risk that addresses fiber size and mineral type. *Crit Rev Toxicol* 2008;38:49-73.
  26. IARC Monograph on the evaluation of the carcinogenic risks to humans. Volume 92 some non-heterocyclic polycyclic aromatic hydrocarbons and some related exposures. Lyon (France): International Agency for Research on Cancer; 2010. 853 p.
  27. Cole P, Rodu B. Epidemiologic studies of chrome and cancer mortality: a series of meta-analyses. *Regul Toxicol Pharmacol* 2005;43:225-31.
  28. Costantino JP, Redmond CK, Bearden A. Occupationally related cancer risk among coke oven workers: 30 years of follow-up. *J Occup Environ Med* 1995;37:597-604.
  29. Hengstler JG, Bolm-Audorff U, Faldum A, Janssen K, Reifnerath M, Götte W, Jung D, Mayer-Popken O, Fuchs J, Gebhard S, Bienfait HG, Schlink K, Dietrich C, Faust D, Epe B, Oesch F. Occupational exposure to heavy metals: DNA damage induction and DNA repair inhibition prove co-exposures to cadmium, cobalt and lead as more dangerous than hitherto expected. *Carcinogenesis* 2003;24:63-73.
  30. LaDou J. Current occupational and environmental medicine. 4th ed. New York: McGrawHill Pub; 2007. p. 224-61.
  31. Scalable classification and regression tree construction [Internet]. 2007 [cited 2010 Apr 10]. Available from: <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.86.413&rep=rep1&type=pdf>.
  32. Spitz MR, Hong WK, Amos CI, Wu X, Schabath MB, Dong Q, Shete S, Etzel CJ. A risk model for prediction of lung cancer. *J Natl Cancer Inst* 2007;99:715-26.
  33. Cassidy A, Myles JP, van Tongeren M, Page RD, Liloglou T, Duffy SW, Field JK. The LLP risk model: an individual risk prediction model for lung cancer. *Br J Cancer* 2008;98:270-6.
  34. Etzel CJ, Kachroo S, Liu M, D'Amelio A, Dong Q, Cote ML, Wenzlaff AS, Hong WK, Greisinger AJ, Schwartz AG, Spitz MR. Development and validation of a lung cancer risk prediction model for African-Americans. *Cancer Prev Res (Phila)* 2008;1:255-65.
  35. Subramanian J, Govindan R. Lung cancer in never smokers: a review. *J Clin Oncol* 2007;25:561-70.
  36. Brown T. Silica exposure, smoking, silicosis and lung cancer--complex interactions. *Occup Med (Lond)* 2009;59:89-95.
  37. Liddell FDK. The interaction of asbestos and smoking in lung cancer. *Ann Occup Hyg* 2001;45:341-56.
  38. Cassidy A, 't Mannetje A, van Tongeren M, Field JK, Zaridze D, Szeszenia-Dabrowska N, Rudnai P, Lissowska J, Fabianova E, Mates D, Bencko V, Foretova L, Janout V, Fevotte J, Fletcher T, Brennan P, Boffetta P. Occupational exposure to crystalline silica and risk of lung cancer: a multicenter case-control study in Europe. *Epidemiology* 2007;18:36-43.
  39. Hertz-Picciotto I, Smith AH, Holtzman D, Lipsett M, Alexeeff G. Synergism between occupational arsenic exposure and smoking in the induction of lung cancer. *Epidemiology* 1992;3:23-31.
  40. Case BW. Asbestos, smoking, and lung cancer: interaction and attribution. *Occup Environ Med* 2006;63:507-8.