BMC Public Health

Research article

Identification and classification of high risk groups for Coal Workers' Pneumoconiosis using an artificial neural network based on occupational histories: a retrospective cohort study

Hongbo Liu[1,2], Zhifeng Tang[1], Yongli Yang[3], Dong Weng[1], Gao Sun[2], Zhiwen Duan[1] and Jie Chen*[1]

Address: [1]Division of Pneumoconiosis, School of Public Health, China Medical University, 92 North 2nd Road, Shenyang, 110001, PR China, [2]Department of Health Statistics, School of Public Health, China Medical University, 92 North 2nd Road, Shenyang, 110001, PR China and [3]Division of Occupational Disease, General Hospital of Tiefa Colliery, Diaobingshan, Liaoning Province, 112700, PR China

Email: Hongbo Liu - liuhbabc@163.com; Zhifeng Tang - tzfcmu@163.com; Yongli Yang - yl_yang2005@yahoo.com.cn; Dong Weng - cruise00@gmail.com; Gao Sun - sungao@mail.cmu.edu.cn; Zhiwen Duan - dzhiwen@yahoo.com.cn; Jie Chen* - chenjie@mail.cmu.edu.cn

* Corresponding author

Published: 29 September 2009

BMC Public Health 2009, **9**:366 doi:10.1186/1471-2458-9-366

Received: 24 March 2009
Accepted: 29 September 2009

This article is available from: http://www.biomedcentral.com/1471-2458/9/366

Abstract

Background: Coal workers' pneumoconiosis (CWP) is a preventable, but not fully curable occupational lung disease. More and more coal miners are likely to be at risk of developing CWP owing to an increase in coal production and utilization, especially in developing countries. Coal miners with different occupational categories and durations of dust exposure may be at different levels of risk for CWP. It is necessary to identify and classify different levels of risk for CWP in coal miners with different work histories. In this way, we can recommend different intervals for medical examinations according to different levels of risk for CWP. Our findings may provide a basis for further emending the measures of CWP prevention and control.

Methods: The study was performed using longitudinal retrospective data in the Tiefa Colliery in China. A three-layer artificial neural network with 6 input variables, 15 neurons in the hidden layer, and 1 output neuron was developed in conjunction with coal miners' occupational exposure data. Sensitivity and ROC analyses were adapted to explain the importance of input variables and the performance of the neural network. The occupational characteristics and the probability values predicted were used to categorize coal miners for their levels of risk for CWP.

Results: The sensitivity analysis showed that influence of the duration of dust exposure and occupational category on CWP was 65% and 67%, respectively. The area under the ROC in 3 sets was 0.981, 0.969, and 0.992. There were 7959 coal miners with a probability value < 0.001. The average duration of dust exposure was 15.35 years. The average duration of ex-dust exposure was 0.69 years. Of the coal miners, 79.27% worked in helping and mining. Most of the coal miners were born after 1950 and were first exposed to dust after 1970. One hundred forty-four coal miners had a probability value ≥0.1. The average durations of dust exposure and ex-dust exposure were 25.70 and 16.30 years, respectively. Most of the coal miners were born before 1950 and began to be exposed to dust before 1980. Of the coal miners, 90.28% worked in tunneling.

Conclusion: The duration of dust exposure and occupational category were the two most important factors for CWP. Coal miners at different levels of risk for CWP could be classified by the three-layer neural network analysis based on occupational history.

Background

Coal workers' pneumoconiosis (CWP) is an incurable occupational lung disease caused by inhaling respirable coal mine dust, and especially those who have worked underground for many years, are at risk of developing CWP, even at low exposure levels [1,2]. Consequently, in the early 1950s some countries began to implement laws and regulations and initiate technically feasible control measures to minimize coal miners'dust exposure [3-5]. Recent studies have shown that the number of coal miners with CWP is decreasing in these countries year-by-year [6-8]. However, in developing countries, owing to the growing world economy, the increase in coal production and utilization results in numerous miners exposed to the health hazards of coal mine dust [9,10]. In China, coal is the major energy resource (about 70% of electricity is generated in coal-burning power plants). The estimated number of underground miners at present is > 6 millions [11]. It has been reported that the number of new CWP patients is > 4000 cases per year and CWP accounts for about 48% of the total number of cases of pneumoconiosis in China [12].

At present, a chest X-ray is the gold standard of monitoring and diagnosing CWP. In the US all miners working in an underground coal mine must be offered a chest X-ray every 5 years by the mines [13]. Current regulations in China require that coal miners must be offered a chest X-ray every 2-3 years (GBZ188-2007). Coal miners with different occupational categories and durations of dust exposure may be at different levels of risk for CWP [14,15]. It is necessary to identify and classify different levels of risk for CWP of coal miners with a different work history. In this way, we can recommend different intervals of medical examinations according to different levels of risk for CWP.

An artificial neural network is potentially more successful than a traditional statistical model in predicting clinical outcome or others [16,17]. It has been widely applied to predict, diagnose, and classify disease in many fields [18,19]. In the field of occupational health, many studies using the artificial neural network have been reported [20,21]. In our study, we designed an artificial neural network which used a Bayesian learning algorithm by introducing probabilistic treatment of the Bayesian inference technique. It can overcome some difficult problems, such as local trapping, over-fitting, and overtime in training. Also, it is proposed to have significant advantages over the conventional neural network approach [22].

This study conducted a longitudinal retrospective investigation at the Tiefa Colliery in northeast China. We constructed an artificial neural network model based on occupational histories to predict the risk for CWP in miners. We classified different levels of risk for CWP of coal miners and recommended different intervals of medical examinations according to different levels of risk for CWP. It could provide the basis for further emending the measures of CWP prevention and control and it is important for strengthening the surveillance of occupational hazards in coal mines.

Methods

Study settings

A retrospective cohort study was conducted in the Tiefa Colliery in northeast China. The colliery was founded in the early 1950s. In the Tiefa Colliery, the type of coal mined is Kennel coal, and the mining technique used is longwall. In the 1950s, dust prevention measures were poor. At the start of the 1960s, wet operation and ventilation devices were used to decrease dust concentration. In the late 1960s and later, completely mechanized coal mining equipment was installed in the plants, and the dust concentration decreased noticeably. There were further improvements in decreasing dust concentration between the late 1970s and the early 1980s by using other advanced machinery, which were evidently effective in reducing dust exposure levels. Thus, the longitudinal trend in dust exposure levels in the colliery should decline.

Study population

We investigated all coal miners who had been exposed to dust for at least 5 years. Every investigated miner had detailed records of their occupational history, past physical examination cards, and chest X-rays. Those coal miners having parted from dusty jobs or having deceased were also included in the study if their duration of dust exposure was ≥ 5 years in the mine. The data were collected in December 2007. Most of data in the study were elicited from personnel records in the Manpower Resource Section of the Tiefa Colliery. The records of health status were obtained from the Department of Industry Hygiene and Occupational Disease of this colliery. The use of the mine data in this study was approved by the Manpower Resource Section and the Department of Industry Hygiene and Occupational Disease of the Tiefa Colliery in January 2008.

The duration of dust exposure was calculated for each coal miner by the accumulation of the periods of all jobs with dust exposure. The duration of each job was calculated by taking the difference between the starting date of the exposure and the ending date. The duration of ex-dust exposure started on the first day of parting from dusty jobs, and ended on the date when the study ended (31 December 2007), or when miners were diagnosed with CWP. Occupational categories were divided into four groups according to the way a miner was exposed to dust (tunneling, mining, combining, and helping), and also the composition of the dust. Both tunneling and mining miners worked in the underground areas in direct contact

with the area in which dust was produced. Helping miners performed maintenance, transportation, and washing the plant and cinders; helping miners did not come in direct contact with the working surface of the mining or tunneling. Coal miners were defined as combining if their occupational history contained tunneling and other occupational categories, and their duration of tunneling was > 2 years, but not more than one-half of the time exposed to dust. Table 1 describes all the variables selected for in this study, and provides the interpretation of the values used for coding of global variables. The variables in the global dataset were normalized to [0, 1].

Diagnostic outcome

Chest radiographs of CWP patients and other investigated miners were read and diagnosed independently by five qualified experts who were members of the Tieling Municipal Pneumoconiosis Diagnosis Committee. The diagnosis was based on the Chinese diagnostic standard for pneumoconiosis and corresponding standard films of pneumoconiosis.

Statistical analysis

Development of the artificial neural network

The data were divided at random into 3 subsets in a 3:1:1 ratio as follows: training set (8793 subjects), validation set (2931 subjects), and testing set (2931 subjects). The training set was used for the adjustment of weights during training. As the artificial neural network can be overtrained to recognize specific cases in the training set and result in good performance in the training set but not in the testing sets, the validation set was to decide when to stop training in order to minimize the potential bias.

The three-layer neural network model was built and trained using a Bayesian learning algorithm, which had 6 input nodes and 1 output neuron (0, absence of CWP; 1, presence of CWP). To determine the optimal number of

Table 1: Possible values and characteristics of coal miners with and without CWP

Variable Name	Possible values and number of examples			
	Category	With CWP (n = 236)	Without CWP (n = 14419)	χ^2(P)
Birth (year of birthday)	(7)Before 1930	52	201	1821.977 (<0.0001)
	(6)1930-	129	696	
	(5)1940-	40	871	
	(4)1950-	14	3652	
	(3)1960-	1	6168	
	(2)1970-	0	2736	
	(1)After 1980	0	95	
Age (age in years of first exposure to dust)	(7)< 20	60	6328	60.003 (<0.0001)
	(6)20-	100	5786	
	(5)25-	51	1614	
	(4)30-	16	454	
	(3)35-	8	167	
	(2)40-	1	61	
	(1)≥45	0	9	
Year (year of first exposure to dust)	(7)Before 1950	14	33	1992.257 (<0.0001)
	(6)1950-	149	669	
	(5)1960-	57	759	
	(4)1970-	15	3026	
	(3)1980-	1	6164	
	(2)1990-	0	3764	
	(1)After 2000	0	4	
Job (occupational category)	(4)Tunneling	201	4949	274.335 (<0.0001)
	(3)Combining	9	670	
	(2)Mining	23	2892	
	(1)Helping	3	5908	
Time 1(years)	Duration of dust exposure	23.38 ± 7.07	18.31 ± 7.05	120.561 (<0.001)
Time 2(years)	Duration of ex-dust exposure	6.09 ± 7.96	2.49 ± 5.72	47.613 (<0.001)

The number of ( ) is possible value of different variables in neural network analysis

neurons, we randomly split all data into 5 subsets of equal size. At any number of neurons, it was trained on all but one subset, and tested on the remaining one. We started from 6 neurons, and gradually increased the number of neurons. When there were 15 neurons, the performance of the trained neural network on output sample in the testing set began to deteriorate. Hence, the intermediate layer had 15 neurons in the current neural network model. The number of training epochs was set to 300, the learning rate was 0.05, and the training goal was set at 0.001. Default settings were used for the remaining parameters. The analysis of the neural network was performed using the MATLAB Neural Network Toolbox (2006).

Sensitivity analysis

Neural network models have been long criticized for being black box solutions mainly because of their inability to generate interpretable parameters for each input variable. To mitigate this problem, sensitivity analysis was adapted to explain their inference mechanism [23]. In our study, every input variable to the network varied between the mean ± standard deviation, while all others were fixed at their respective means, and the corresponding change was recorded as a percentage deviation in the output. It could help illustrate the effect of changing a single input variable on the network output.

ROC curve

The receive operating characteristic (ROC) methodology is a computational methodology which has a very important connection to the artificial neural network applied to classification applications [24]. An important feature of the ROC curve is that it readily incorporates prevalence and misclassification cost factors in decision-making. In our study, the performance of the artificial neural network was tested using ROC curve analysis. The area under the ROC (AUROC) measured the accuracy of the output value of the neural network which distinguished coal miners with or without CWP in the future. We calculated sensitivity, specificity, positive and negative predictive values (PPV and NPV), and Youden's index of output. The analysis of the ROC curve was performed using SPSS11.5 (SPSS Institute, Inc., Chicago, IL, USA).

In this study, continuous variables were compared by the Kruskal-Wallis test when appropriate; categorical variables were compared by a χ^2 test.

Results

Basic characteristics

This study included 14,655 coal miners. By the end of 2007, all subjects had worked for an average of 18.39 years, and on average, 2.55 years had elapsed since exposure to dust stopped. There were 236 coal miners with CWP and 14,419 coal miners without CWP. Table 1 shows that there were statistically significant differences in every input variable between coal miners with and without CWP.

Sensitivity analysis

The sensitivity analysis of the 6 variables is outlined in Figure 1. The value shown for each input variable is a measure of its relative importance, with 0 representing a variable that has no effect on the prediction and 1 representing a variable that completely dominates the prediction. The x-axis is the input variables; the y-axis is the percent change on the output variable. The most important factors in this predictive model were occupational category (job) and dust-exposed duration (time 1), and the influence on the output variance among coal miners were 67% for occupational category and 65% for dust-exposed duration.

Accuracy of the ROC curve

When the output value of the neural network was analyzed by the ROC curve in 3 sets, the AUROC was as follows: 0.98 (in the training set), 0.97 (in the validation set), and 0.99 (in the testing set; Figure 2). The sensitivity at the optimum cut-off of 0.018 in the 3 sets was 93.0%, 91.5%, and 97.9%. In the testing set, NPV was 100.0%. Sensitivity, specificity, positive and negative predictive values, and the percentage of cases correctly identified as negative at different cutoff points in the testing set are shown in Table 2. For a decision threshold of 0.015, sensitivity and negative predictive values were 100%, which indicated that this model predicted the absence of CWP with the highest accuracy. This procedure would have correctly identified 87.8% of all coal miners not exposed to high risk for CWP in the future.

Occupational characteristics of coal miners without CWP

We further divided coal miners without CWP into four groups according to the probability values of their outputs: <0.001 (7959 subjects), 0.001- (5024 subjects), 0.015- (1292 subjects), and 0.1- (144 subjects), as shown in Table 3. Coal miners with a probability value < 0.001 were at a low risk for CWP. Their average durations of dust exposure and ex-dust exposure were 15.35 and 0.69 years, respectively. Of the coal miners, 79.27% worked in helping and mining. Most of the coal miners were born after 1950 and were first exposed to dust after 1970. Coal miners with a probability value ≥0.015 were at high risk for CWP. There were 144 coal miners with a probability value ≥0.1. The average durations of dust exposure and ex-dust exposure were 25.70 and 16.30 years, respectively. Most of the coal miners were born before 1950 and began to be exposed to dust before 1980; 90.28% worked in tunneling, but there were 4 coal miners working in helping

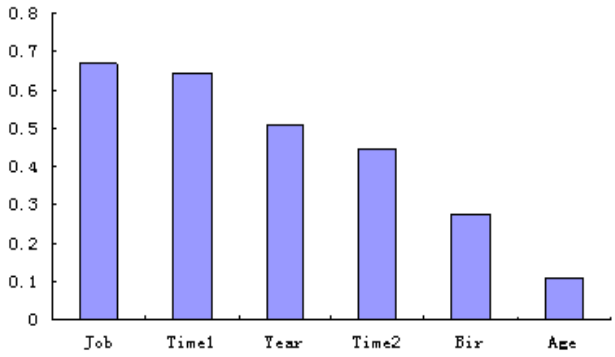

Figure 1
Sensitivity analysis of input variables.

and 10 in mining. The coal miners with a probability value ≥0.1 were at the highest risk for CWP.

Discussion

CWP results from exposure to coal mine dust for a long time, and can be prevented by decreasing the dust concentration in the workplace. In developed countries, effective dust control measures have resulted in a low incidence of CWP in coal mines [25,26]. Moreover, a steady decrease in the number of coal miners exposed to dust in developed countries has played a crucial role. In the United States, the number of employees in mining industries has decreased from 1.7 million in the early 1990s to 0.2 million in the early 2000s [27,28]. But in developing countries, especially China, more and more miners are exposed to high concentrations of dust, work in collieries lacking effective surveillance, and are at a high risk for CWP. The surveillance and monitoring of occupational hazards in coal mines require an enormous budget for a developing country. It is important to identify and classify the high risk groups for CWP for strategic monitoring and managing of coal miners.

In the present study, based on the occupational histories of coal miners, we constructed a three-layer neural network which extends a back-propagation learning algorithm by introducing probabilistic treatment of the Bayesian inference technique for the synaptic weight [29]. Results of sensitivity analysis showed the importance of predictors. As can be seen, the two most important variables influencing the prediction of CWP were duration of dust exposure and occupational category. These results were consistent with some of the earlier studies, which revealed that duration of dust exposure and occupational category are important factors effecting the occurrence of CWP [15,30,31]. Sensitivity analysis showed that the neural network using Bayesian approaches could achieve its predictive purpose.

As demonstrated by ROC curves, the predictive accuracy of BNN in conjunction with occupational exposure data was reasonably high in the training, validation, and testing sets. The AUROC were 0.98, 0.97, and 0.99, respectively. Suarthana [32] fitted a multivariable logistic regression model according to respiratory symptoms, exposure level, and lung functions to predict the probability of an individual worker having pneumoconiosis. The AUROC of the three models were 0.79 (0.74-0.85), 0.79 (0.74-0.85), and 0.81 (0.75-0.86), respectively. Although there were different influencing variables in the two models, the neural network had a superior accuracy for individual classification.

Our study evaluated the influence of the different cut-off points on the accuracy of ROC. When we chose a high cut-off point, the number of coal miners at high risk for CWP was few, but there was a low sensitivity and many coal miners with CWP could not be identified to be at high risk for CWP. The purpose of primary prevention is to prevent healthy coal miners from becoming CWP patients. Thus, we should choose a lower cut-off point to improve the sensitivity, and to reduce the number of missed coal miners at high risk for CWP, who had a relatively low probability of output. A probability value of 0.015 yielded 100% sensitivity in the testing set, which meant that all coal miners with CWP would be captured.

The probability of all coal miners without CWP was predicted by the neural network in this study. Most of the

Table 2: ROC analysis of neural network output at different cutoff points in the testing set

Cutoff point	Sensitivity%	Specificity%	PPV%	NPV%	Youden Index%	Negatives correctly identified%
0.015	100.0	89.3	13.2	100.0	89.3	87.8
0.02	95.7	92.3	16.9	99.9	88.0	90.9
0.04	93.6	96.2	28.4	99.9	89.8	94.7
0.06	91.5	97.4	36.4	99.9	88.9	96.0
0.08	89.4	98.1	42.9	99.8	87.5	96.6
0.1	87.2	98.6	50.6	99.8	85.8	97.2
0.3	70.2	99.8	86.8	99.5	70.0	98.7
0.5	63.8	99.9	90.9	99.4	63.7	98.9
0.6	63.8	100.0	100.0	99.4	63.8	99.0

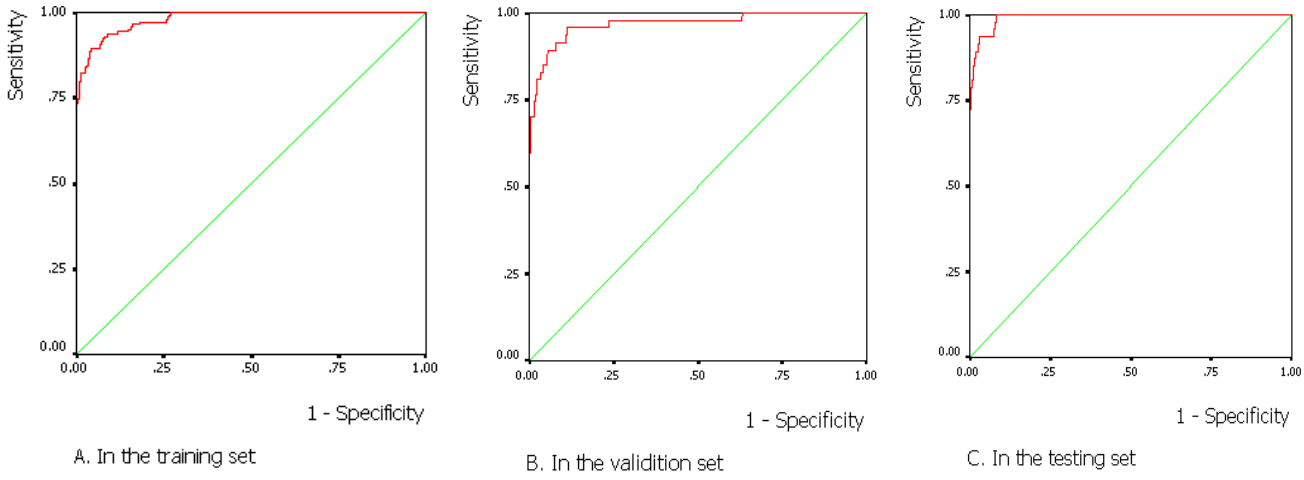

Figure 2
ROC curve of the neural network output in three sets.

coal miners with a probability > 0.015 were tunneling miners (69.01%). Generally speaking, tunneling miners were at the highest risk for CWP [14], and they should have high probability values predicted. But, there were 4 helping miners and 10 mining miners with a probability > 0.1. We further analyzed that their average duration of dust exposure was 32.14 years and 28.57 years, respectively, and their average duration of ex-dust exposure was 13.25 years and 9.86 years, respectively. All of these factors resulted in high probability values, which indicated a high risk of CWP [33,34]. Therefore it is obvious that these occupational histories could be used to predict coal miners at high risk for CWP.

Therefore, according to the probability values predicted by the neural network and the characteristics of occupational histories of coal miners without CWP, we suggest that coal miners with a duration of dust exposure >25 years and with a duration of ex-dust exposure >10 years, or coal miners with a duration of dust exposure >25 years and working in tunneling or combining, would be at high risk for CWP. They should undergo a medical examination every year. Other coal miners with a duration of dust exposure >25 years would be at relatively high risk for CWP and should undergo a medical examination every 2 years. Coal miners with a duration of dust exposure of 15-25 years and with the sum of dust exposure and ex-dust exposure >25 years, or coal miners with a dust exposure of 15-25 years and working in tunneling or combining, would be at moderate risk for CWP, and should undergo a medical examination every 3 years. All other coal miners with a duration of dust exposure <15 years, or with the sum of dust exposure and ex-dust exposure <25 years, would be at low risk for CWP, and interval between medical examinations can be prolonged to 5 years.

Study limitations

The Tiefa Colliery is located in northeast China. Although it is a typical state-run colliery in China, only one colliery was studied, and may not be representative of all collieries. The predictive model based on Tiefa Colliery should not be directly applied to other collieries without necessary adaptations. However, it is feasible that the predictive model in this study is applied to identify the risk of CWP in other collieries after appropriately adjusting and retraining the neural network model.

Dust concentration and cumulative dose exposure are the most important factors to influence the occurrence of CWP [35]. Because of the administration system problems in Tiefa Colliery, we failed to obtain the data of the dust measurements. This is another limitation of this study. As some studies showed that the duration of dust exposure and occupational category were closely related to cumulative dose exposure and dust concentration [14,36], we speculated that occupational histories in this study can be used to reflect dust exposure level.

In addition, many studies on CWP have demonstrated significant influence of cigarette smoking on coal miners's health. Smoking mainly influence on lung function and respiratory symptoms. CWP is a lung disease caused by coal mine dust, which has significant changes of pathology, mainly including lung fibrosis. Although smoking influence on CWP is weakly, it is an important confounder factor. In this study, there is no mention on the

Table 3: Occupational characteristics of coal miners with different probability values of output(%)

Variable Name	Category	The probability value of output among coal miners without CWP			
		<0.001 (n = 7959)	**0.001- (n = 5024)**	**0.015- (n = 1292)**	**0.1- (n = 144)**
Bir	(7)Before 1930	0(0.00)	93(1.85)	97(7.51)	11(7.64)
	(6)1930-	13(0.16)	340(6.77)	269(20.82)	74(51.39)
	(5)1940-	84(1.05)	375(7.46)	356(27.55)	56(38.89)
	(4)1950-	643(8.08)	2436(48.49)	570(44.12)	3(2.08)
	(3)1960-	4439(55.77)	1729(34.41)	0(0.00)	0(0.00)
	(2)1970-	2685(33.73)	51(1.01)	0(0.00)	0(0.00)
	(1)After 1980	95(1.19)	0(0.00)	0(0.00)	0(0.00)
Age	(7)< 20	4011(50.39)	1943(38.67)	320(24.77)	54(37.50)
	(6)20-	3088(38.80)	1985(39.51)	650(50.31)	63(43.75)
	(5)25-	557(7.00)	869(17.30)	169(13.08)	19(13.19)
	(4)30-	215(2.70)	122(2.43)	112(8.67)	5(3.47)
	(3)35-	71(0.89)	65(1.29)	28(2.17)	3(2.08)
	(2)40-	16(0.20)	34(0.68)	11(0.85)	0(0.00)
	(1)≥45	1(0.01)	6(0.12)	2(0.15)	0(0.00)
Year	(7)Before 1950	0(0.00)	13(0.25)	15(1.16)	5(3.47)
	(6)1950-	3(0.04)	270(5.37)	307(23.76)	89(61.80)
	(5)1960-	23(0.29)	341(6.79)	354(27.40)	41(28.47)
	(4)1970-	423(5.31)	2084(41.48)	512(39.63)	7(4.86)
	(3)1980-	3779(47.48)	2282(45.42)	101(7.82)	2(1.39)
	(2)1990-	3727(46.83)	34(0.68)	3(0.23)	0(0.00)
	(1)After 2000	4(0.05)	0(0.00)	0(0.00)	0(0.00)
Job	(4)Tunneling	1468(18.44)	2490(49.56)	861(66.64)	130(90.28)
	(3)Combining	182(2.29)	443(8.82)	45(3.48)	0(0.00)
	(2)Mining	1744(21.91)	943(18.77)	195(15.09)	10(6.94)
	(1)Helping	4565(57.36)	1148(22.85)	191(14.78)	4(2.78)
Time1	Duration of dust exposure	15.35 ± 5.85	20.57 ± 6.41	26.91 ± 8.01	25.70 ± 7.08
Time2	Duration of ex-dust exposure	0.69 ± 2.54	3.46 ± 6.51	8.30 ± 8.52	16.30 ± 9.25

The number of ( ) is possible value of different variables in neural network analysis

smoking habits of the studied population, duration of smoking and number of cigarette smoked. This is another limitation of this study. Because of the retrospective cohort study and the large studied population, it is difficult to collect smoking habit information. In our study, we want to make use of occupational histories to predict the risk for CWP, so we didn't take into account the influence of cigarette smoking.

Conclusion

Duration of dust exposure and occupational category were two most important factors of CWP. Coal miners at different levels of risk of CWP could be identified by the three-layer neural network analysis based on occupational history. Coal miners with duration of dust exposure >25 years and ex-dust exposure >10 years, or coal miners with duration of dust exposure >25 years and working in tunneling or combining would be at high risk for CWP. Coal miners with duration of dust exposure <15 years, or with the sum of dust exposure and ex-dust exposure <25 years would be at low risk for CWP.

Abbreviations

CWP: coal workers' pneumoconiosis; ROC: receive operating characteristic; AUROC: area under ROC; PPV: positive predictive values; NPV: negative predictive values

Competing interests

The authors declare that they have no competing interests.

Authors' contributions

HL conducted the study, performed statistical analysis, and wrote the initial draft and revised the manuscripts after consultation with the other authors. ZT designed and collected the preliminary data, and helped revise the manuscript. YY participated in the design and acquisition of preliminary data. DW collected the preliminary data, and helped revise the manuscript. GS participated in the study design and interpretation of the data. ZD made contributions to the study design and data analysis. JC conceived the study and participated in the interpretation of the data, and the drafting and revision of the paper. All authors have read and approved the submitted version of the manuscript.

Acknowledgements

This work was supported by a grant from Liaoning Provincial Natural Science Foundation (No. 20052219).

References

1. Antao VC, Petsonk EL, Sokolow LZ, Wolfe AL, Pinheiro GA, Hale JM, Attfield MD: **Rapidly progressive coal workers' pneumoconiosis in the United States: geographic clustering and other factors.** *Occup Environ Med* 2005, **62:**670-674.
2. CDC: **Advanced cases of coal workers' pneumoconiosis-two countries, Virginia, 2006.** *MMWR* 2005, **55:**909-913.
3. Parobeck : **Effect of the 2.0 mg/m3 coal mine dust standard on underground environmental dust levels.** *Am Ind Hyg Assoc J* 1975, **36(8):**604-609.
4. Kinnear P: **The politics of coal dust: industrial campaigns for the regulation of dust disease in Australian coal mining, 1939-49.** *Labour Hist* 2001, **80:**65-82.
5. Weeks JL: **The Mine Safety and Health Administration's criterion threshold value policy increases miners' risk of pneumoconiosis.** *Am J Ind Med* 2006, **49(6):**492-498.
6. Scott DF, Grayson RL, Metz EA: **Disease and illness in U.S. mining, 1983-2001.** *J Occup Environ Med* 2004, **46:**1272-1277.
7. Goodwin S, Attfield MD: **Temporal trends in coal workers' pneumoconiosis prevalence: validating the national coal study results.** *Occup Environ Med* 1998, **40:**1065-1071.
8. Marek K, Lebecki K: **Occurrence and prevention of coal miners' pneumoconiosis in Polan.** *Am J Ind Med* 1999, **36:**610-617.
9. Bridbord K, Costello J, Gamble J, Groce D, Hutchison M, Jones W, Merchant J, Ortmeyer C, Reger R, Wagner WL: **Occupational safety and health implications of increased coal utilization.** *Environ Health Perspect* 1979, **33:**285-302.
10. **World Energy Council** [http://www.worldenergy.org/publications/energy_policy_scenarios_to_2050/default.asp]. Accessed March 18, 2009
11. Liu FD, Pan ZQ, Liu SL, Chen L, Ma JZ, Yang ML, Wang NP: **The estimation of the number of underground coal miners and the annual dose to coal miners in China.** *Health Phys* 2007, **93:**127-132.
12. Zhang M, Du XY, Li T, Wang D: **Design for monitor and early warning system based on data from national occupational diseases report database.** *Chinese J Safety Sci Tech* 2009, **5(1):**80-84.
13. *Federal Coal Mine Health and Safety Act of 1969, Pub. L. No. 91-173, S. 2917* [http://www.msha.gov/solicitor/coalact/69act.htm]. (December 30, 1969)
14. Mamuya SH, Bråtveit M, Mashalla Y, Moen BE: **High prevalence of respiratory symptoms among workers in the development section of a manually operated coal mine in a developing country: A cross sectional study.** *BMC Public Health* 2007, **7:**17-24.
15. Attfield MD, Morring K: **An investigation into the relationship between coal workers' pneumoconiosis and dust exposure in U.S. coal miners.** *Am Ind Assoc J* 1992, **53(8):**486-492.
16. Green M, Bjork J, Forberg J, Ekelund U, Edenbrandt L, Ohlsson M: **Comparison between neural networks and multiple logistic regression to predict acute coronary syndrome in the emergency room.** *Artif Intell Med* 2006, **38:**305-318.
17. Abbod MF, Catto JW, Linkens DA, Hamdy FC: **Application of artificial intelligence to the management of urological cancer.** *J Urol* 2007, **178:**1150-1156.
18. Ahmed FE: **Artificial neural networks for diagnosis and survival prediction in colon cancer.** *Mol Cancer* 2005, **6:**4-29.
19. William GB: **Application of artificial neural networks to clinical medicine.** *Lancet* 1995, **346:**1135-1138.
20. Black J, Benkei G, Smith K, Fritschi L: **Artificial neural networks and job-specific modules to assess occupational exposure.** *Ann Occup Hyg* 2004, **48:**595-600.
21. Zurada J, Karwowski W, Marras WS: **A neural network-based system for classification of industrial jobs with respect to risk of low back disorders due to workplace design.** *Appl Ergon* 1997, **28(1):**49-58.
22. Lampinen J, Vehtati A: **Bayesian approach for neural networks-review and case studies.** *Neural Netw* 2001, **14(3):**257-274.
23. Fish KE, Blodgett JG: **A visual method for determining variable importance in an artificial neural network model: an empirical benchmark study.** *J Targ Meas Anal Market* 2003, **11(3):**244-254.
24. Dayhoff JE, Deleo JM: **Artificial neural network-Opening the black box.** *Cancer* 2001, **91:**1616-1635.
25. Attfield MD, Althouse RB: **Surveillance data on US coal miners' pneumoconiosis 1970 to 1986.** *Am J Public Health* 1992, **82:**971-977.
26. CDC: **Pneumoconiosis prevalence among working coal miners examined in federal chest radiograph surveillance programs--United States, 1996-2002.** *MMWR Morb Mortal Wkly Rep* 2003, **52(15):**336-340.
27. National Institute for Occupational Safety and Health: **Work related lung disease surveillance report, 1991. Cincinnati, Ohio: US Dept of Health and Human Services, Centers for Disease Control, NIOSH.** DHHS (NIOSH) publication No. 91-113.
28. Castranova V, Vallyathan V: **Silicosis and coal workers' pneumoconiosis.** *Environ Health Perspect* 2000, **108(Suppl 4):**675-684.
29. Yao XQ, Zhu HQ, She ZS: **A dynamic Bayesian network approach to protein secondary structure prediction.** *BMC Bioinformatics* 2008, **9:**49-61.
30. Heederik D, Attfield M: **Characterization of dust exposure for the study of chronic occupational lung disease: A comparison of different exposure assessment strategies.** *Am J Epidemiol* 2000, **151:**982-990.
31. Takigawa T, Kishimoto T, Nabe M, Nishide T, Wang DH, Seki A, Uchida G, Kira S: **The current state of workers' pneumoconiosis in relationship to dusty working environments in Okayama Prefecture, Japan.** *Acta Med Okayama* 2002, **56(6):**303-308.
32. Suarthana E, Moons KG, Heederik D, Meijer E: **A simple diagnostic model for ruling out pneumoconiosis among construction workers.** *Occup Environ Med* 2007, **64(9):**595-601.
33. Attfield MD, Seixas NS: **Prevalence of pneumoconiosis and its relationship to dust exposure in a cohort of U.S. bituminous coal miners and ex-miners.** *Am J Ind Med* 1995, **27:**137-151.
34. Kenny LC, Hurley F, Warren ND: **Estimation of the risk of contracting pneumoconiosis in the UK coal mining industry.** *Ann Occup Hyg* 2002, **46(suppl 1):**257-260.
35. Onder M, Onder S: **Evaluation of occupational exposures to respirable dust in underground coal mines.** *Ind Health* 2009, **47(1):**43-49.
36. Naidoo RN, Robins TG, Murray J: **Respiratory outcomes among South African coal miners at autopsy.** *Am J Ind Med* 2005, **48(3):**217-224.

Pre-publication history

The pre-publication history for this paper can be accessed here:

http://www.biomedcentral.com/1471-2458/9/366/prepub