ORIGINAL ARTICLE

Retrospective Study

Machine learning to relate PM2.5 and PM10 concentrations to outpatient visits for upper respiratory tract infections in Taiwan: A nationwide analysis

Mei-Juan Chen, Pei-Hsuan Yang, Mi-Tren Hsieh, Chia-Hung Yeh, Chih-Hsiang Huang, Chieh-Ming Yang, Gen-Min Lin

Mei-Juan Chen, Pei-Hsuan Yang, Mi-Tren Hsieh, Chieh-Ming Yang, Gen-Min Lin, Department of Electrical Engineering, National Dong Hwa University, Hualien 974, Taiwan

Chia-Hung Yeh, Department of Electrical Engineering, National Taiwan Normal University, Taipei 106, Taiwan

Chia-Hung Yeh, Chih-Hsiang Huang, Department of Electrical Engineering, National Sun Yat-sen University, Kaohsiung 804, Taiwan

Gen-Min Lin, Department of Medicine, Hualien Armed Forces General Hospital, Hualien 971, Taiwan

Gen-Min Lin, Departments of Medicine, Tri-Service General Hospital, National Defense Medical Center, Taipei 114, Taiwan

ORCID number: Mei-Juan Chen (0000-0003-3382-8296); Pei-Hsuan Yang (0000-0002-6269-5157); Mi-Tren Hsieh (0000-0002-9627-377X); Chia-Hung Yeh (0000-0002-2837-6662); Chih-Hsiang Huang (0000-0002-6490-8072); Chieh-Ming Yang (0000-0002-6614-4635); Gen-Min Lin (0000-0002-5509-1056).

Informed consent statement: Participants were not required to give informed consent to this retrospective study since the analysis of baseline characteristics used anonymized clinical data.

Conflict-of-interest statement: The authors declare that they have no conflicts of interest.

Abstract

AIM
To examine the accuracy of machine learning to relate particulate matter (PM) 2.5 and PM10 concentrations to upper respiratory tract infections (URIs).

METHODS
Daily nationwide and regional outdoor PM2.5 and PM10 concentrations collected over 30 consecutive days obtained from the Taiwan Environment Protection Administration were the inputs for machine learning, using multilayer perceptron (MLP), to relate to the subsequent one-week outpatient visits for URIs. The

URI data were obtained from the Centers for Disease Control datasets in Taiwan between 2009 and 2016. The testing used the middle month dataset of each season (January, April, July and October), and the training used the other months' datasets. The weekly URI cases were classified by tertile as high, moderate, and low volumes.

RESULTS
Both PM concentrations and URI cases peak in winter and spring. In the nationwide data analysis, MLP machine learning can accurately relate the URI volumes of the elderly (89.05% and 88.32%, respectively) and the overall population (81.75% and 83.21%, respectively) with the PM2.5 and PM10 concentrations. In the regional data analyses, greater accuracy is found for PM2.5 than for PM10 for the elderly, particularly in the Central region (78.10% and 74.45%, respectively), whereas greater accuracy is found for PM10 than for PM2.5 for the overall population, particularly in the Northern region (73.19% and 63.04%, respectively).

CONCLUSION
Short-term PM2.5 and PM10 concentrations were accurately related to the subsequent occurrence of URIs by using machine learning. Our findings suggested that the effects of PM2.5 and PM10 on URI may differ by age, and the mechanism needs further evaluation.

Core tip: Particulate matter (PM) 2.5 and PM10 air pollutants can trigger inflammation and predispose the respiratory tract to infections. This study used the multilayer perceptron (MLP) machine learning architecture to relate the daily PM2.5 and PM10 concentrations over 30 consecutive days to the subsequent one-week outpatient visits for upper respiratory tract infections (URIs) in Taiwan between 2008 and 2016. In the nationwide data analysis, PM2.5 and PM10 concentrations can precisely predict the volumes of URI for the elderly (89.05% and 88.32%, respectively) and the overall population (81.75% and 83.21%, respectively). Our findings suggested that machine learning could accurately relate PM2.5 and PM10 concentrations to the outpatient visits for URI, especially for the elderly population.

INTRODUCTION

Particulate matter (PM) 2.5 and PM10, also known as particle pollutions, refer to a mixture of liquid droplets and solid particles in the air with diameters ≤ 2.5 μm and ≤ 10 μm, respectively. In developing countries, PM2.5 accounts for half of PM10 concentrations, whereas in developed countries, PM2.5 is estimated to account for 50%-80% of the PM10 concentrations[1]. Both PM2.5 and PM10 can deposit in the respiratory tract and may trigger inflammatory reactions that increase the plasma interleukin-6 and fibrinogen levels[2]. The inflammation process related to air pollution might decrease innate immunity and predispose robust individuals to acute illnesses, such as upper respiratory tract infections (URIs), and the development of chronic disease, such as lung malignancies[3-5]. Several observational studies have revealed that PM2.5 and PM10 concentrations may be associated with the occurrence of URIs[5,6] and may increase the risk of mortality related to hospitalized pneumonia[7].

Machine learning utilizes computational statistics to explore optimized algorithms, which can learn from and make predictions based on data. Machine learning for potentially hazardous exposure has been successfully applied to predict the occurrence of several clinical diseases, such as myocardial infarction, and the risk of mortality in previous studies[8]. In addition, machine learning, such as artificial neural networks, can provide us with an opportunity for big data training for the prediction of clinical diseases[9]. For example, some models using convolutional neural networks for training with hundreds of thousands of fundus images to predict the presence and the severity of diabetic retinopathy have been well established[10-13]. Carnegie Mellon's Delphi group of the United States Centers of Disease Control has been working to create a machine learning model that accurately tracks the spread of the flu[14].

Since the severity of air pollution varies geographically, the hazardous effect on human health may also differ by region and ethnicity. It is reasonable to create a surveillance system for forecasting the probability of disease occurrence related to regional air pollution. Multilayer perceptron (MLP) artificial intelligence, a type of machine learning similar to the human neural network, is formed by at least three layers of nodes that make use of nonlinear activation for data training[15]. Accordingly, we attempted to establish such an MLP model to relate PM2.5 and PM10 concentrations to the volume of outpatient visits for acute URIs in Taiwan.

MATERIALS AND METHODS

Data collection
The datasets of outpatient visits for URIs were obtained from the website of the Centers for Disease Control (CDC) of Taiwan for the period from January 2009 to December 2016, which is 417 wk in total. Clinical
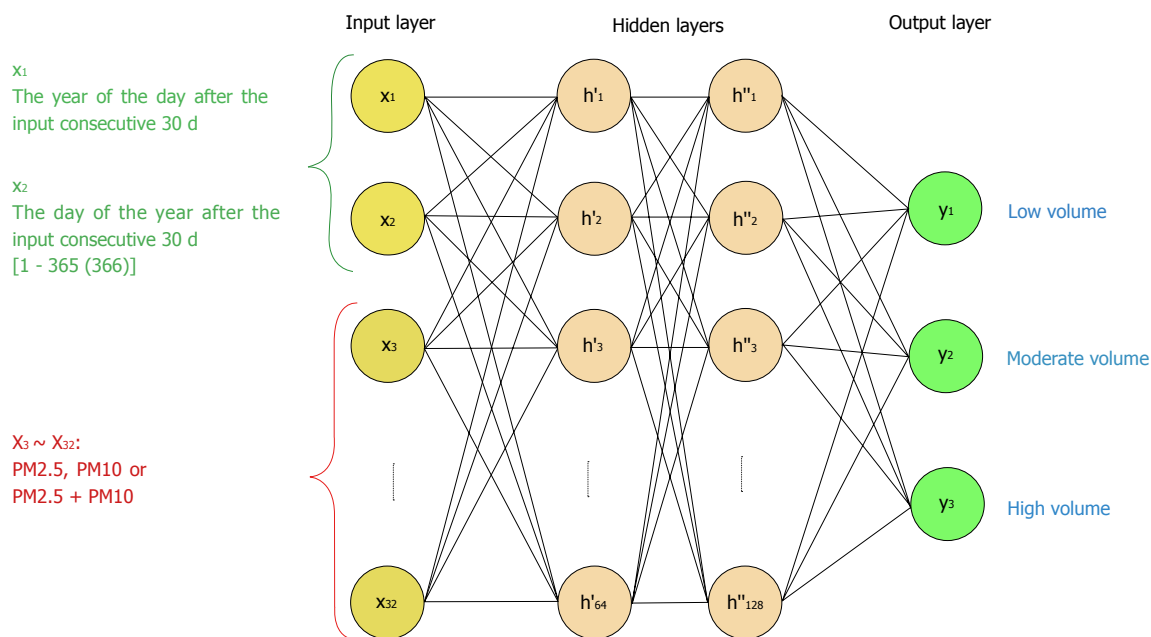
Figure 1 Multilayer perceptron model for the proposed algorithm.

physicians have to diagnose URIs for patients according to clinical symptoms, physical presentations, and objective laboratory data at an outpatient department. The cases of URIs were retrieved from the Taiwan Nationwide Health Insurance records, which is based on the International Classification of Diseases, Ninth Revision, Clinical Modification (ICD-9-CM) codes 465, 487.1, 488.02, 488.01 and 034.0 for acute URIs (Supplemental Table 1). The datasets of PM2.5 and PM10 were obtained from the Taiwan Environment Protection Administration, and the PM concentrations were measured and collected outdoors from 55 ambient air quality monitoring sites spread throughout Taiwan.

MLP model
Figure 1 shows the MLP architecture with a forward and backward propagation learning algorithm to accurately relate PM2.5 and PM10 concentrations to the occurrence of outpatient visits for acute URI, which is classified by tertile as high, moderate, and low volumes. The testing used the middle month datasets of each season in Taiwan, which account for 33% of all datasets [January (winter), April (spring), July (summer), and October (fall)], and the training used the other months' datasets, which account for 67% [December and February (winter), March and May (spring), June and August (summer), and September and November (fall)]. The training and testing procedures were repeated multiple times to determine optimal outcomes.

Statistical analysis
Daily average PM2.5 and PM10 concentrations for 30 consecutive days were used as the inputs of the MLP model to relate to the outputs of subsequent one-

week URI volumes. PM2.5 and PM10 concentrations were normalized before inputting them into the MLP model. Based on the criteria of the Taiwan Environment Protection Administration in Supplemental Table 1, the hazardous to human health cut-off levels of PM2.5 and PM10 were $\geq$ 250.4 $\mu g/m^3$ and $\geq$ 424.0 $\mu g/m^3$, respectively, which were set as 1[16]. In addition, the cut-off levels of PM2.5 and PM10 suggestive of good air quality were $\leq$ 15.4 $\mu g/m^3$ and $\leq$ 54.0 $\mu g/m^3$, respectively, which were set as 0. The PM values within the upper and lower cut-off levels were normalized to between 0 and 1 (Supplemental Table 2). If both PM2.5 and PM10 were treated as the inputs together in the MLP model, the normalization of PM2.5 + PM10 would be the average of the sum of the normalizations of PM2.5 and PM10. One-week volumes of URIs were used as outputs because of the time lag effect[5], since ill patients may seek a medical consultation days after the beginning of the infection, when the symptoms have worsened. Additionally, the accuracy of MLP machine learning for the overall and elderly ($\geq$ 65 years) patients was estimated. The MLP model was tested in a nationwide data analysis and in several regional data analyses of western Taiwan consisting of the Northern (business and economic areas), Central, and Southern regions (industrial areas), and eastern Taiwan, which is represented by the Eastern region (a national park area) (supplemental Figure 1).

RESULTS

Figure 2 shows the average daily concentrations of PM2.5 and PM10 from December 2008 to December 2016 and the average numbers of outpatient visits
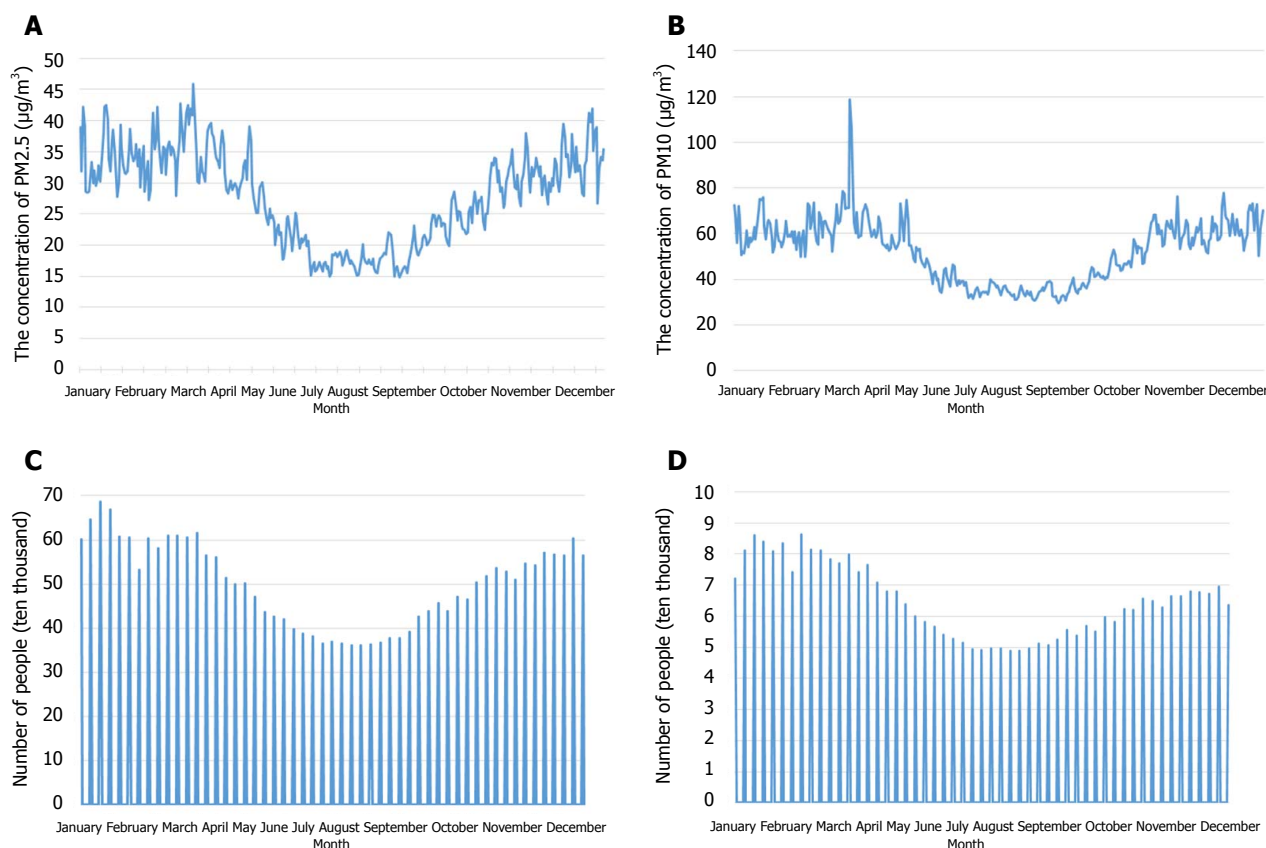
A

B

C

D

Figure 2 The average daily concentrations and weekly numbers of outpatient visits for upper respiratory tract infections in each month. A and B: PM2.5 and PM10, repectively (December 2008 - December 2016); C and D: The overall and the elderly patients, repsectively (January 2009 - December 2016). PM: Particulate matter.

for URIs for the overall population and the elderly population in each month from January 2009 to December 2016. As shown, the concentrations of PM2.5 and PM10 distribute as a diurnal curve and peak in the winter and spring seasons. The PM2.5 concentrations were between 15 μg/m^3 and 46 μg/m^3, and the PM10 concentrations ranged from 30 μg/m^3 to 100+ μg/m^3. The occurence of total and elderly outpatient visits for acute URI are most prevalent in winter and spring as well, which is correlated with the PM2.5 and PM10 levels. With regard to the volume of outpatient visits for URIs, the number of monthly overall URI cases ranges from 35000 to 70000, and the number of monthly elderly URI cases ranges from 4800 to 9000.

Table 1 reveals the averge PM2.5 and PM10 conce-ntrations and the numbers of outpatient visits for URIs in each season by the four regions and in all of Taiwan. In general, the regional patterns of PM concentrations were in line with the nationwide pattern in Taiwan. On average, the Eastern region, followed by the Northern region, had the lowest PM concentrations. Conversely, the Central and Southern regions, consisting of many of the industrial counties of Taiwan, had the highest PM concentrations. Similarly, the regional patterns of outpatient numbers were in line with the nationwide pattern. The Nothern region had the greatest numbers of URIs and the Eastern region had the least numbers

of URIs. The URI prevalence was mostly higher in the winter and spring.

Table 2 demonstrates the nationwide and regional data analysis results of the accuracy of MLP machine learning to relate PM2.5 and PM10 concentrations to the volume of outpatient visits for URIs in the overall population and the elderly population. The na-tionwide data analysis reveals that PM2.5 and PM10 concentrations can correctly relate to the volumes of URI in the elderly population (89.05% and 88.32%, respectively) and the overall population (81.75% and 83.21%, respectively). In the regional analyses, PM2.5 and PM10 concentrations have the greatest accuracy for the elderly population in the Eastern and Northern regions (80.29%/81.75% and 80.43%/76.81%, resp-ectively), which are the two least air polluted areas in Taiwan. By contrast, the accuracy of URI occurrence based on large data mining of PM2.5 and PM10 in all regions of Taiwan is relatively lower at approximately 63% to 73% for the overall population. In addition, PM2.5 has greater accuracy than PM10 for the elderly, particularly in the Central region (78.10% and 74.45%, respectively), whereas PM10 has greater accuracy than PM2.5 for the overall population, particularly in the Northern region (73.19% and 63.04%, respectively). Notably, the MLP performance was better at a nation-wide scale than that at the regional scale. When the

Table 1 Nationwide and regional average particulate matter concentrations between December 2008 and December 2016 and the number of outpatient visits for upper respiratory infections in each season in Taiwan from January 2009 to December 2016

Regions/PM ($\mu g/m3$)		Spring	Summer	Fall	Winter
Taiwan	PM2.5	30.90 ± 12.30	17.97 ± 6.37	28.00 ± 10.20	34.44 ± 13.10
	PM10	58.05 ± 27.47	35.00 ± 7.74	53.96 ± 17.46	62.43 ± 21.31
Northern region	PM2.5	27.37 ± 12.43	18.63 ± 7.43	20.36 ± 10.02	25.95 ± 13.78
	PM10	49.52 ± 33.52	33.30 ± 11.64	38.63 ± 18.51	46.44 ± 24.17
Central region	PM2.5	35.36 ± 16.13	20.12 ± 9.37	32.99 ± 13.90	36.37 ± 15.64
	PM10	60.55 ± 30.03	35.64 ± 11.56	57.15 ± 20.32	60.30 ± 23.70
Southern region	PM2.5	34.99 ± 17.46	17.32 ± 8.71	37.33 ± 15.42	49.14 ± 15.95
	PM10	66.75 ± 32.10	36.51 ± 11.21	70.60 ± 25.93	87.42 ± 24.73
Eastern region	PM2.5	15.37 ± 8.05	10.36 ± 4.94	13.49 ± 7.82	15.62 ± 8.73
	PM10	30.49 ± 15.45	24.04 ± 9.86	31.85 ± 23.11	30.12 ± 14.67
Regions/URI patients (× 10³)					
Taiwan	Overall	523.75 ± 93.61	375.67 ± 35.07	492.46 ± 63.48	607.31 ± 124.83
	Elderly	70.13 ± 11.27	50.62 ± 4.75	60.98 ± 6.91	77.52 ± 17.23
Northern region	Overall	210.90 ± 35.83	148.45 ± 14.46	192.93 ± 26.09	238.55 ± 53.95
	Elderly	25.10 ± 3.99	17.73 ± 1.65	20.93 ± 2.55	27.19 ± 6.71
Central region	Overall	102.77 ± 20.04	72.74 ± 7.30	97.90 ± 12.79	121.29 ± 24.65
	Elderly	13.53 ± 2.23	9.63 ± 0.94	11.85 ± 1.36	14.88 ± 3.27
Southern region	Overall	77.95 ± 14.09	62.25 ± 7.06	78.97 ± 10.53	95.38 ± 18.11
	Elderly	11.78 ± 1.91	9.57 ± 1.17	11.29 ± 1.33	13.74 ± 2.78
Eastern region	Overall	12.91 ± 2.22	8.71 ± 1.12	11.41 ± 1.63	14.35 ± 2.67
	Elderly	2.29 ± 0.34	1.52 ± 0.18	1.86 ± 0.24	2.51 ± 0.54

PM: Particulate matter; URI: Upper respiratory infection.

Table 2 The accuracy of Particulate matter machine learning for PM2.5 and PM10 concentrations to predict the events of outpatient visits for upper respiratory infections by the four regions and in all of Taiwan

Accuracy (%)		Overall population	Elderly population
Taiwan	PM2.5	81.75	89.05
	PM10	83.21	88.32
	PM2.5 + PM10	83.21	89.05
Northern region	PM2.5	63.04	80.43
	PM10	73.19	76.81
	PM2.5 + PM10	65.94	78.99
Central region	PM2.5	69.34	78.10
	PM10	72.26	74.45
	PM2.5 + PM10	69.34	77.37
Southern region	PM2.5	71.01	76.09
	PM10	71.74	74.64
	PM2.5 + PM10	71.74	74.64
Eastern region	PM2.5	67.15	80.29
	PM10	71.53	81.75
	PM2.5 + PM10	71.53	84.67

PM: Particulate matter.

PM2.5 and PM10 concentrations were combined in the MLP model, the accuracy was not improved much for either the elderly or the overall population.

DISCUSSION

In previous studies, the hazardous effect of high levels of PM2.5 and PM10 exposures on the occurrence of URIs had been observed[2,5]. These studies often applied a case-crossover design using the case PM data on the event day of disease occurrence to compare with other control PM data on prospective and retrospective days to see the odds ratio of URIs related to the high PM air pollutions. To our best knowledge, the effect of PM on the respiratory tract may be synergic or cumulative, and using one-day PM concentrations to estimate the URI risk could have bias. Our study used a novel procedure of MLP machine learning, which can process large amounts of successive 30-d PM concentration data from nationwide and regional perspectives to relate to the URI volumes, which would improve the solidity of the relationship.

We found that the concentrations of PM2.5 and PM10 peak in the winter and spring, which could be highly related to several meteorological parameters, such as gravity, outdoor temperature, humidity, wind speed, and rain[17,18]. The URI occurrence may be associated with factors by which the pathogens can grow

rapidly and predispose robust individuals to acute illness. Higher PM levels coinciding with the pathogens' active seasons might contribute to a high prevalence of URIs. In addition, we also showed that the relationship of PM concentrations with acute URI had the best result for the elderly population. This could be explained in part by the fact that the elderly, who had many comorbidities and weaker immunity, were more likely to have acute illness when exposed to multiple air pollutions.

Another important finding was that the MLP models for the PM2.5 and PM10 training data to relate the concentrations of PM2.5 and PM 10 to URI occurrence were more accurate in areas of low air pollution for the elderly. It is reasonable that although PM air pollutants are hazardous to human health, several other toxins, such as sulfide dioxide (SO₂) and carbon monoxide (CO), could also contribute to the development of URIs in areas of heavy air pollution[19]. As a result, the importance of PM on the occurrence of URI might be attenuated if there were many other coexisting air pollutants. In addition, the MLP machine learning more accurately related PM2.5 to the URI volumes for the elderly, whereas the accuracy with PM10 was better for the overall population. Further evaluations are still needed to determine if the effect of PM2.5 and PM10 on the respiratory tract may differ by age, which influences the physical activity and outdoor exposure time. Moreover, there was no additional predictive benefit of putting both PM concentrations together into the MLP models. This was likely because PM2.5 is included in PM10 and they are highly correlated or because of the effect of overfitting in the machine learning process[20].

Notably, the MLP models utilized for the big PM data training had the greatest accuracy at the nationwide scale. This could be because more PM data were involved in the nationwide scale, which facilitated the machine learning and possibly led to the greater accuracy. However, the explanation did not hold true with regard to the regional difference. For instance, the Northern region has the largest population in Taiwan but, paradoxically, yields the lowest accuracy. In contrast, the Eastern region has the smallest population in Taiwan but shows the highest accuracy. Therefore, we speculate that many factors other than sample size, such as the severity of air pollution or the heterogeneity of unrecognized factors, such as migration, may affect the results of the MLP.

Our study has several strengths. First, the PM2.5 and PM10 concentrations as well as the diagnosis of URI were reliable and objectively retrieved from government agencies. Second, a large amount of data from the publicly available website could be easily utilized for ongoing studies, and the results are reproducible. On the other hand, although the MLP is a well-known machine learning method, there are a few limitations in our study. First, we used only PM concentration data in this study, and we may need data for more air pollutants, such as SO₂ and CO, and other

meteorological parameters for further adjustments of the study. Second, details of the baseline characteristics of the patients with URIs, such as sex, body weight, and underlying comorbidities, were lacking, and the results were mainly based on the assumption that all people did not migrate frequently during the study period, which may result in potential bias if the assumptions were inaccurate. Third, we could not provide direct evidence for the cause-effect relationship between PM and acute URIs, which might be due to coincidence merely based on the retrospective nature of the study design.

In conclusion, MLP machine learning could accurately relate short-term PM2.5 and PM10 concentrations to subsequent outpatient visits for URIs. Our findings suggested that the elderly population and areas with less air pollution may have better MLP test results. In addition, the hazardous effect of PM2.5 and PM10 on URIs may differ by age, which is possibly related to daily activity and outdoor exposure time, which needs further evaluation. We also noticed that the performance of the MLP at the nationwide scale was better than that at the regional scale. Whether this finding was because of a larger population sample size or a higher heterogeneity in the nationwide scale is unknown, and this also requires further investigation.

ARTICLE HIGHLIGHTS

Research background
PM2.5 and PM10, also known as particle pollutions, can deposit in the respiratory tract and may trigger inflammatory reactions. Several studies have revealed that PM2.5 and PM10 concentrations may be associated with the occurrence of upper respiratory tract infections (URIs) and increase the mortality related to hospitalized pneumonia. Machine learning utilizes computational statistics to explore optimized algorithms that can learn from and make predictions based on data. Machine learning for potential hazardous exposures has been successfully applied to predict the occurrence of several clinical diseases, such as myocardial infarction, and the related risk of mortality. In addition, machine learning, such as artificial neural networks, can provide us an opportunity for big data training for the prediction of clinical diseases. For example, Carnegie Mellon's Delphi group of the United States Centers of Disease Control has been working to create a machine learning model that accurately tracks the spread of the flu.

Research motivation
Since the severity of air pollution varies geographically, the hazardous effect on human health may also differ by region and ethnicity. It is reasonable to create a surveillance system to forecast the probability of disease occurrence related to regional air pollution. Accordingly, we attempted to establish a model of machine learning to relate PM2.5 and PM10 concentrations to the volume of outpatient visits for acute URIs in Taiwan.

Research objectives
To examine the accuracy of machine learning to relate PM2.5 and PM10 concentrations to URIs.

Research methods
Daily nationwide and regional outdoor PM2.5 and PM10 concentrations collected over 30 consecutive days from the Taiwan Environment Protection Administration were the inputs for the multilayer perceptron (MLP) machine learning to relate to the subsequent one-week outpatient visits for URIs. The URI data were obtained from the Centers for Disease Control datasets in

Taiwan between 2009 and 2016. The testing used the middle month dataset of each season (January, April, July, and October), and the training used the other months' datasets. The weekly URI cases were classified by tertile as high, moderate, and low volumes.

Research results
Both PM concentrations and URI cases peak in the winter and spring. In the nationwide data analysis, MLP machine learning can accurately relate PM2.5 and PM10 concentrations with the URI volumes of the elderly (89.05% and 88.32%, respectively) and the overall population (81.75% and 83.21%, respectively). In the regional data analyses, PM2.5 has greater accuracy than PM10 for the elderly, particularly in the Central region (78.10% and 74.45%, respectively), whereas PM10 has greater accuracy than PM2.5 for the overall population, particularly in the Northern region (73.19% and 63.04%, respectively).

Research conclusions
Machine learning could accurately relate short-term PM2.5 and PM10 concentrations to subsequent URI occurrence. Our findings suggested that the effects of PM2.5 and PM10 on URI may differ by age, and the mechanism needs further evaluation.

Research perspectives
We used MLP machine learning to successfully relate PM concentrations data to the volume of URI cases. Data for more air pollutants and other meteorological parameters can be applied to the current MLP model in future work.

REFERENCES
1 **Grahame TJ**, Klemm R, Schlesinger RB. Public health and components of particulate matter: the changing assessment of black carbon. *J Air Waste Manag Assoc* 2014; **64**: 620-660 [PMID: 25039199 DOI: 10.1080/10962247.2014.912692]
2 **Bind MA**, Baccarelli A, Zanobetti A, Tarantini L, Suh H, Vokonas P, Schwartz J. Air pollution and markers of coagulation, inflammation, and endothelial function: associations and epigene-environment interactions in an elderly cohort. *Epidemiology* 2012; **23**: 332-340 [PMID: 22237295 DOI: 10.1097/EDE.0b013e31824523f0]
3 **Liao Y**, Xu L, Lin X, Hao YT. Temporal Trend in Lung Cancer Burden Attributed to Ambient Fine Particulate Matter in Guangzhou, China. *Biomed Environ Sci* 2017; **30**: 708-717 [PMID: 29122091 DOI: 10.3967/bes2017.096]
4 **Zhou Y**, Li L, Hu L. Correlation Analysis of PM10 and the Incidence of Lung Cancer in Nanchang, China. *Int J Environ Res Public Health* 2017; **14**: pii: E1253 [PMID: 29048397 DOI: 10.3390/ijerph14101253]
5 **Li R**, Jiang N, Liu Q, Huang J, Guo X, Liu F, Gao Z. Impact of Air Pollutants on Outpatient Visits for Acute Respiratory Outcomes. *Int J Environ Res Public Health* 2017; **14**: pii: E47 [PMID: 28067786 DOI: 10.3390/ijerph14010047]
6 **Sinclair AH**, Edgerton ES, Wyzga R, Tolsma D. A two-time-period comparison of the effects of ambient air pollution on outpatient visits for acute respiratory illnesses. *J Air Waste Manag Assoc* 2010; **60**: 163-175 [PMID: 20222529 DOI: 10.3155/1047-3289.60.2.163]
7 **Faustini A**, Stafoggia M, Colais P, Berti G, Bisanti L, Cadum E, Cernigliaro A, Mallone S, Scarnato C, Forastiere F; EpiAir Collaborative Group. Air pollution and multiple acute respiratory outcomes. *Eur Respir J* 2013; **42**: 304-313 [PMID: 23314899 DOI: 10.1183/09031936.00128712]
8 **Ambale-Venkatesh B**, Yang X, Wu CO, Liu K, Hundley WG, McClelland R, Gomes AS, Folsom AR, Shea S, Guallar E, Bluemke DA, Lima JAC. Cardiovascular Event Prediction by *irc Res* 2017; **121**: 1092-1101 [PMID: 28794054 DOI: 10.1161/CIRCRESAHA.117.311312]
9 **Chrysostomou C**, Partaourides H, Seker H. Prediction of Influenza A virus infections in humans using an Artificial Neural Network learning approach. *Conf Proc IEEE Eng Med Biol Soc* 2017; **2017**: 1186-1189 [PMID: 29060087 DOI: 10.1109/EMBC.2017.8037042]
10 **Gulshan V**, Peng L, Coram M, Stumpe MC, Wu D, Narayanaswamy A, Venugopalan S, Widner K, Madams T, Cuadros J, Kim R, Raman R, Nelson PC, Mega JL, Webster DR. Development and Validation of a Deep Learning Algorithm for Detection of Diabetic Retinopathy in Retinal Fundus Photographs. *JAMA* 2016; **316**: 2402-2410 [PMID: 27898976 DOI: 10.1001/jama.2016.17216]
11 **Ting DSW**, Cheung CY, Lim G, Tan GSW, Quang ND, Gan A, Hamzah H, Garcia-Franco R, San Yeo IY, Lee SY, Wong EYM, Sabanayagam C, Baskaran M, Ibrahim F, Tan NC, Finkelstein EA, Lamoureux EL, Wong IY, Bressler NM, Sivaprasad S, Varma R, Jonas JB, He MG, Cheng CY, Cheung GCM, Aung T, Hsu W, Lee ML, Wong TY. Development and Validation of a Deep Learning System for Diabetic Retinopathy and Related Eye Diseases Using Retinal Images From Multiethnic Populations With Diabetes. *JAMA* 2017; **318**: 2211-2223 [PMID: 29234807 DOI: 10.1001/jama.2017.18152]
12 **Gargeya R**, Leng T. Automated Identification of Diabetic Retinopathy Using Deep Learning. *Ophthalmology* 2017; **124**: 962-969 [PMID: 28359545 DOI: 10.1016/j.ophtha.2017.02.008]
13 **Lin GM**, Chen MJ, Lin YY, Lin MH, Yeh CH. An Improvement of Machine Detection for Any Diabetic Retinopathy by Preprocessing Retinal Photographs to Entropy Images in Deep Learning. *APAO Conference. Hong Kong*, Feb 8-11, 2018. Available from: URL: http://2018.apaophth.org/wp-content/uploads/2018/01/Final-Program-2018-1.pdf
14 **Hamilton H**. The CDC uses machine learning and social media to forecast flu outbreaks. Available from: URL: https://www.electronicproducts.com/Programming/Software/The_CDC_uses_machine_learning_and_social_media_to_forecast_flu_outbreaks.aspx
15 **Tang J**, Deng C, Huang GB. Extreme Learning Machine for Multilayer Perceptron. *IEEE Trans Neural Netw Learn Syst* 2016; **27**: 809-821 [PMID: 25966483 DOI: 10.1109/TNNLS.2015.2424995]
16 **Environmental Protection Administration**. AQI value and impact on health. Available from: URL: https://taqm.epa.gov.tw/taqm/en/
17 **Wang J**, Ogawa S. Effects of Meteorological Conditions on PM2.5 Concentrations in Nagasaki, Japan. *Int J Environ Res Public Health* 2015; **12**: 9089-9101 [PMID: 26247953 DOI: 10.3390/ijerph120809089]
18 **Feng J**, Yu H, Mi K, Su X, Li Y, Li Q, Sun J. One year study of PM2.5 in Xinxiang city, North China: Seasonal characteristics, climate impact and source. *Ecotoxicol Environ Saf* 2018; **154**: 75-83 [PMID: 29454989 DOI: 10.1016/j.ecoenv.2018.01.048]
19 **Jaakkola JJ**, Paunio M, Virtanen M, Heinonen OP. Low-level air pollution and upper respiratory infections in children. *Am J Public Health* 1991; **81**: 1060-1063 [PMID: 1854003 DOI: 10.2105/AJPH.81.8.1060]
20 **Srivastava N**, Hinton G, Krizhevsky A, Sutskever I, Salakhutdinov R. Dropout: a simple way to prevent neural networks from overfitting. *J Mach Learn Res* 2014; **15**: 1929-1958

Baishideng®