

OPEN

An association between air pollution and daily most frequently visits of eighteen outpatient diseases in an industrial city

Tang-Tat Chau¹ & Kuo-Ying Wang^{2*}

Toxic effects of air pollutants were individually identified in various organs of the body. However, the concurrent occurrences and the connection of diseases in multiple organs arise from air pollution has not been concurrently studied before. Here we hypothesize that there exist connected health effects arise from air pollution when diseases in various organs were considered together. We used medical data from hospital outpatient visits for various organs in the body with a disease-air pollution model that represents each of the diseases as a function of the environmental factors. Our results show that elevated air pollution risks (above 40%) concurrently occurred in diseases of spondylosis, cerebrovascular, pneumonia, accidents, chronic obstructive pulmonary disease (COPD), influenza, osteoarthritis (OA), asthma, peptic ulcer disease (PUD), cancer, heart, hypertensive, diabetes, kidney, and rheumatism. Air pollutants that were associated with elevated health risks are particular matters with diameters equal or less than $2.5 \mu\text{m}$ ($PM_{2.5}$), nitrogen dioxide (NO_2), ozone (O_3), particular matters with diameters equal or less than $10 \mu\text{m}$ (PM_{10}), carbon monoxide (CO), and nitrogen oxide (NO). Concurrent occurrences of diseases in various organs indicate that the immune system tries to connectively defend the body from persistent and rising air pollution.

Air pollution is the most important environmental risk factor for health¹. The most severe health risks associated with air pollution are the mortality rates, which have been studied the longest². The 1952 London smog revealed an increase of 3000 mortality rates associated with continuous increases in sulfur dioxide (SO_2) concentrations from 70 ppbv to 340 ppbv in 4 weeks². Based on a sample of 8111 people for a period of 14 to 16 years (a total of 111,076 person-years), Dockery *et al.* found that $PM_{2.5}$ contributes to excess mortality in six U.S. cities³. Hence an estimated latency period of mortality rates⁴ of weeks to years between causes (increased in air pollution levels) and effects (rising mortality rates in the population).

Before people got killed by diseases associated with air pollution, people were struck with various diseases associated with air pollution. The toxicity and pathophysiological changes in the organs resulting from air pollution exposure were best found in experimental studies⁵. Air pollutants induced 1.5 to 2 times heritable deoxyribonucleic acid (DNA) mutations in mice placed 1-km compared with mice placed 30-km downwind to two steel mills⁶. Deposited soots changed the physical outlook of bird specimens⁷. Elevated concentrations of selected persistent organic pollutants of polycyclic aromatic hydrocarbons (PAHs), o,p'-isomers of dichlorodiphenyltrichloroethane (DDT), metabolites, and α -hexachlorocyclohexane (HCH) were measured in the placenta and were associated with the increase in neural tube defects in pregnant women⁸. Elevated levels of insecticide p,p'-dichlorodiphenyldichloroethylene (p,p'-DDE) were measured in maternal serum specimens from early pregnancy and were associated with autism among offspring⁹. Elevated levels of DDT were measured in the maternal serum samples collected during the peak years (1959–1967) of DDT use in the United States compared with other samples of breast cancer collected in later decades, and found utero DDT exposure and risks of breast cancer in young women and a possible association with more aggressive tumors¹⁰.

Traffic-related air pollution has shown to trigger microglial activation and neuronal atrophy¹¹, and affect embryonic and adult neurons through glutamatergic mechanisms in rodent models¹². Structural brain magnetic resonance imaging scans revealed that older women with greater $PM_{2.5}$ exposures had significantly smaller white

¹Department of Family Medicine, Landseed International Hospital, Tao-Yuan, Taiwan. ²Department of Atmospheric Sciences, National Central University, Chung-Li, Taiwan. *email: kuoying@mail.atm.ncu.edu.tw

matter present in frontal and temporal lobes and corpus callosum⁵. Combustion-derived iron-rich magnetite nanoparticles from airborne particulate matter pollution were found in the human brain¹³, showing that air pollution particles were directly transported and deposited in the human brain. Mice exposed to elevated levels of NO_2 (2.5–5.0 mg/m³) were found to show the deterioration of spatial learning and memory, aggravated amyloid β_{42} accumulation, promoted pathological abnormalities, and cognitive defects related to Alzheimer's disease¹⁴. Humans exposed to NO_2 and $PM_{2.5}$ were positively correlated with the incidence of dementia in London, England¹⁵. Toxins accompanying air pollutants into the body exert chronic activation of microglia, leading to direct neuronal damage and neuronal death in the central nervous system¹⁶. Delicate communication exists between the brain and the immune system^{17,18}, and evidence that connects health and emotions through neural activity in the brain¹⁹. These pathophysiological findings reveal that air pollutants affected the central nervous system, leading to the decline of brain function²⁰ and fatal diseases¹⁸.

Pathophysiological evidence convincingly showed pathways of various diseases arose from air pollution²¹. Hence, more diseases associated with air pollution remained to be discovered in the population²². Kampa and Castanas²³ and Jay²⁴ showed that diseases associated with air pollution were found in various parts of human organs: cardiovascular and circulatory system, digestive and excretory system, endocrine system, integumentary and exocrine system, lymphatic and immune system, muscular and skeletal system, nervous system, renal and urinary system, reproductive system, respiratory system, and sensory system. Darnall²⁵ hypothesized a whole body approach to study pains, which are often associated with diseases. FitzGerald *et al.*²⁶ proposed that humans are model organisms for future medicine for treating diseases.

Though air pollution concentrations have gradually declined in the developed country, air pollution levels have continuously increased in the developing countries²⁷. Population in the industrial cities are particularly at risks of air pollution^{2,3,28}. In this work, we hypothesize that hospital outpatient visits for diseases contain health signals from the entire body organs that are related to air pollution. Diseases associated with air pollution in various organs were individually studied before, but the concurrence of diseases from various organs that were associated with air pollution hasn't been studied in an integrated and connected approach and perspective. To prove this hypothesis, in this work we studied 18 diseases that were recorded during the hospital outpatient visits. These 18 diseases represent the mostly visited diseases for the outpatients to the Taiwan Landseed Hospital during the period 2007–2011. Wang and Chau²⁹ used daily outpatient visits to find an association between air pollution and daily outpatients of respiratory diseases. Chau and Wang³⁰ found an association between temperatures and daily outpatients for accidents. Note that increased in temperature was found to deteriorate mental health³¹. In this work, we scaled up the analysis by Wang and Chau²⁹ and Chau and Wang³⁰ to investigate the association of daily hospital visits of 18 diseases with air pollution. These 18 diseases arise from various body organs that were impacted by air pollution^{23,24}. A total of 1.7 million medical visits from 18 diseases during 2007–2011 were combined with half a million air pollution and meteorological data for big data analysis. We show that measurable data of medical visits and environmental measurements reveal diseases that arise from various organs are concurrently and connectively linked to air pollution.

Data and Methods

A disease-air pollution model. A multivariate disease-air pollution model used in this work were used in previous studies in finding an association between hospital outpatient visits for respiratory diseases and environmental factors²⁹, and between hospital outpatient visits for accidents and environmental factors³⁰. In this section, we describe detailed computational steps used in this work with the same multivariate model.

The time series of daily (t) medical records $M_{i,j}(t)$ from each disease i of the 18 diseases for each year j of 2007–2011 and each group of 3-age groups of outpatients (0–15, 16–65, and >65 years old) were calculated with respect to 12 time series of daily environmental parameters (7 air pollution parameters of PM_{10} , $PM_{2.5}$, O_3 , CO , NO_2 , NO , SO_2 ; and 5 meteorological parameters of temperature, rainfall, wind direction, wind speed, and relative humidity):

$$M_{i,j}(t) = \begin{pmatrix} \beta_{i,j,0} + \\ \beta_{i,j,1}[PM_{10}(t)]_j + \\ \beta_{i,j,2}[PM_{2.5}(t)]_j + \\ \beta_{i,j,3}[O_3]_j(t) + \\ \beta_{i,j,4}[CO]_j(t) + \\ \beta_{i,j,5}[NO]_j(t) + \\ \beta_{i,j,6}[NO_2]_j(t) + \\ \beta_{i,j,7}[SO_2]_j(t) + \\ \beta_{i,j,8}[T]_j(t) + \\ \beta_{i,j,9}[P]_j(t) + \\ \beta_{i,j,10}[WD]_j(t) + \\ \beta_{i,j,11}[WS]_j(t) + \\ \beta_{i,j,12}[RH]_j(t) \end{pmatrix} = \beta_{i,j,0} + x_j B_{i,j} \quad (1)$$

where

$$B_{i,j} = \begin{pmatrix} \beta_{i,j,1} \\ \beta_{i,j,2} \\ \beta_{i,j,3} \\ \beta_{i,j,4} \\ \beta_{i,j,5} \\ \beta_{i,j,6} \\ \beta_{i,j,7} \\ \beta_{i,j,8} \\ \beta_{i,j,9} \\ \beta_{i,j,10} \\ \beta_{i,j,11} \\ \beta_{i,j,12} \end{pmatrix}. \quad (2)$$

The transposed of the parameter array x_i is

$$(x_j)^t = \begin{pmatrix} [PM_{10}]_j \\ [PM_{2.5}]_j \\ [O_3]_j \\ [CO]_j \\ [NO]_j \\ [NO_2]_j \\ [T]_j \\ [P]_j \\ [WD]_j \\ [WS]_j \\ [RH]_j \end{pmatrix}. \quad (3)$$

Here i represents each of the 18 diseases, $i = 1, \dots, 18$; j represents each year of 2007, \dots , 2011; M is the outpatient numbers; T is temperature; P is daily accumulation precipitation; WD is wind direction; WS is wind speed, and RH is relative humidity. In the application of Eq. 1, medical data $M_{i,j}$ and parameter arrays x_i are known from real-world data. Hence, we solve Eq. 1 for finding the association coefficient matrix $B_{i,j}$.

Notice that time-series data of daily outpatient number $M_{i,j}$ of a disease, and 12 environmental factors are known from input data. The holidays and weekends were removed from the time-series data to avoid the artificial effect of low outpatients on these public holidays^{29,30}. Except for the removal of medical and environmental data from public holidays, no other processes were applied to the input time-series data. All environmental data were tested for normal distribution using the Chi-square test³². Except for rainfall and wind directions, the environmental data resembles the patterns of normal distribution²⁹. Test of the normal distribution of medical data also exhibit patterns of normal distribution^{29,30} (shown also in Supplementary Materials). All environmental data that were used as model variables were normalized by the maximum values of each variable to ensure that the computation were on the same unitless ground when making the comparison³³.

The environmental factors are PM_{10} , $PM_{2.5}$, O_3 , CO , NO , NO_2 , SO_2 , T , P , WD , WS , and RH . The daily outpatient visits were computed according to the 3 age groups of outpatients (0–15, 16–65, and higher than 65 years old as in Wang and Chau²⁹ and Chau and Wang³⁰). A spectrum of 18 diseases was sorted based on the International Classification of Diseases, 9th Revision (ICD-9) as shown in Table 1. These 18 diseases include the diseases of the respiratory system, such as allergic rhinitis (AR), asthma, pneumonia, and COPD; accidents; mental disorders (anxiety, dissociative, somatoform disorders); the digestive system, such as PUD, and chronic liver diseases and cirrhosis; circulatory system, such as cerebrovascular disease, heart disease, and hypertensive disease; diabetes mellitus; malignant neoplasm (cancer); genitourinary system (nephritis, nephritis, and nephrosis); musculoskeletal system, such as other disorders of soft tissues, OA and allied disorders, and spondylosis and allied disorders; and influenza. Figure 1 shows a schematic diagram of distribution of association coefficients $\beta_{i,j}$ with respect to diseases and environmental factors.

As the time-series patterns of outpatient visits to the hospital were shown to lag the time-series patterns of environmental factors, the association coefficients in Eq. 1 were solved for eight time-lag scenarios. These 8 scenarios are the hospital outpatient visits for diseases lag an environmental factor at 0-day, 1-day, 2-day, and so on, to 7-day.

Numerical solutions of the disease-air pollution model. In order to compute disease air pollution risks and air pollution risks, we need to compute association coefficients shown in Eq. 1. For a disease i , we take partial derivatives of Eq. 1 with respect to each of the 12 environmental variables. Hence,

Disease	ICD-9 Codes
Diseases of the respiratory system	
Allergic Rhinitis	477
Asthma	493
Pneumonia	480-486
Chronic obstructive pulmonary disease	490-493
Accidents	800-949
Mental disorders	
Anxiety, dissociative and somatoform disorders	300
Diseases of the digestive system	
Peptic ulcer	533
Chronic liver disease and cirrhosis	571
Diseases of the circulatory system	
Cerebrovascular disease	430-438
Heart disease	390-392, 393-398, 410-414, 420-429
Hypertensive disease	401-405
Endocrine disorders	
Diabetes mellitus	250
Malignant neoplasm	140-208
Diseases of the genitourinary system	
Nephritis, nephrotic syndrome, and nephrosis	580-589
Diseases of the musculoskeletal system and connective tissue	
Other disorders of soft tissues	729
Osteoarthritis and allied disorders	715
Spondylosis and allied disorders	721
Influenza vaccine injection	

Table 1. List The International Classification of Diseases, 9th Revision (ICD-9), For This Wok.

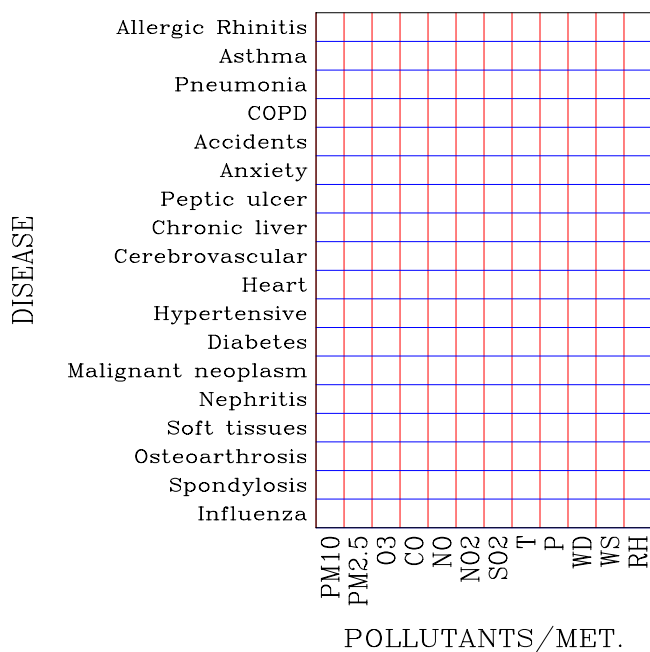


Figure 1. A schematic diagram showing the distribution of correlation coefficient matrix of diseases (y-axis) with respect to parameters of air pollutants and meteorological factors (x-axis). Each empty box represents a correlation coefficient to be determined from the disease-air pollution model shown in Eq. (1).

$$\frac{\partial M_{i,j}}{\partial [PM_{10}]_j} = \left(\begin{array}{l} \beta_{i,j,1} + \\ \beta_{i,j,2} \frac{\partial [PM_{2.5}]_j}{\partial [PM_{10}]_j} + \\ \beta_{i,j,3} \frac{\partial [O_3]_j}{\partial [PM_{10}]_j} + \\ \beta_{i,j,4} \frac{\partial [CO]_j}{\partial [PM_{10}]_j} + \\ \beta_{i,j,5} \frac{\partial [NO]_j}{\partial [PM_{10}]_j} + \\ \beta_{i,j,6} \frac{\partial [NO_2]_j}{\partial [PM_{10}]_j} + \\ \beta_{i,j,7} \frac{\partial [SO_2]_j}{\partial [PM_{10}]_j} + \\ \beta_{i,j,8} \frac{\partial [T]_j}{\partial [PM_{10}]_j} + \\ \beta_{i,j,9} \frac{\partial [P]_j}{\partial [PM_{10}]_j} + \\ \beta_{i,j,10} \frac{\partial [WD]_j}{\partial [PM_{10}]_j} + \\ \beta_{i,j,11} \frac{\partial [WS]_j}{\partial [PM_{10}]_j} + \\ \beta_{i,j,12} \frac{\partial [RH]_j}{\partial [PM_{10}]_j} \end{array} \right) \quad (4)$$

and

$$\frac{\partial M_{i,j}}{\partial [PM_{2.5}]_j} = \left(\begin{array}{l} \beta_{i,j,1} \frac{\partial [PM_{10}]_j}{\partial [PM_{2.5}]_j} + \\ \beta_{i,j,2} + \\ \beta_{i,j,3} \frac{\partial [O_3]_j}{\partial [PM_{2.5}]_j} + \\ \beta_{i,j,4} \frac{\partial [CO]_j}{\partial [PM_{2.5}]_j} + \\ \beta_{i,j,5} \frac{\partial [NO]_j}{\partial [PM_{2.5}]_j} + \\ \beta_{i,j,6} \frac{\partial [NO_2]_j}{\partial [PM_{2.5}]_j} + \\ \beta_{i,j,7} \frac{\partial [SO_2]_j}{\partial [PM_{2.5}]_j} + \\ \beta_{i,j,8} \frac{\partial [T]_j}{\partial [PM_{2.5}]_j} + \\ \beta_{i,j,9} \frac{\partial [P]_j}{\partial [PM_{2.5}]_j} + \\ \beta_{i,j,10} \frac{\partial [WD]_j}{\partial [PM_{2.5}]_j} + \\ \beta_{i,j,11} \frac{\partial [WS]_j}{\partial [PM_{2.5}]_j} + \\ \beta_{i,j,12} \frac{\partial [RH]_j}{\partial [PM_{2.5}]_j} \end{array} \right) \quad (5)$$

Here $\partial M_{i,j}/\partial [PM_{10}]_j$ means variations of outpatients $M_{i,j}$ for disease i with respect to the variations of $[PM_{10}]_j$ in year j ; and $\partial M_{i,j}/\partial [PM_{2.5}]_j$ means variations of outpatients $M_{i,j}$ for disease i with respect to the variations of $[PM_{2.5}]_j$ in year j . For the terms on the right-hand side of Eq. 4, $\partial [PM_{2.5}]_j/\partial [PM_{10}]_j$ means the variations of $[PM_{2.5}]_j$ associated with the changes in $[PM_{10}]_j$, and so on for the rest of other terms. These terms represent changes with respect to $[PM_{10}]_j$, indicating the coupling effects between environmental factors.

Similarly, we take partial directives of $M_{i,j}$ for a disease i with respect to the rest of the other environmental variables. Hence, we write this system of 12 equations with respect to a disease i in j year as

$$A_{i,j} = J_{i,j} B_{i,j}. \tag{6}$$

where $A_{i,j}$ is

$$A_{i,j} = \begin{pmatrix} \frac{\partial M_{i,j}}{\partial [PM_{10}]_j} \\ \frac{\partial M_{i,j}}{\partial [PM_{2.5}]_j} \\ \frac{\partial M_{i,j}}{\partial [O_3]_j} \\ \frac{\partial M_{i,j}}{\partial [CO]_j} \\ \frac{\partial M_{i,j}}{\partial [NO]_j} \\ \frac{\partial M_{i,j}}{\partial [NO_2]_j} \\ \frac{\partial M_{i,j}}{\partial [SO_2]_j} \\ \frac{\partial M_{i,j}}{\partial [T]_j} \\ \frac{\partial M_{i,j}}{\partial [P]_j} \\ \frac{\partial M_{i,j}}{\partial [WD]_j} \\ \frac{\partial M_{i,j}}{\partial [WS]_j} \\ \frac{\partial M_{i,j}}{\partial [RH]_j} \end{pmatrix}. \tag{7}$$

The Jacobian $J_{i,j}$ of this disease-air pollution model for a disease i is written as

$$J_{i,j} = \begin{pmatrix} 1 & \frac{\partial f_{i,j,2}}{\partial x_{i,j,1}} & \frac{\partial f_{i,j,3}}{\partial x_{i,j,1}} & \dots & \frac{\partial f_{i,j,11}}{\partial x_{i,j,1}} & \frac{\partial f_{i,j,12}}{\partial x_{i,j,1}} \\ \frac{\partial f_{i,j,1}}{\partial x_{i,j,2}} & 1 & \frac{\partial f_{i,j,3}}{\partial x_{i,j,2}} & \dots & \frac{\partial f_{i,j,11}}{\partial x_{i,j,2}} & \frac{\partial f_{i,j,12}}{\partial x_{i,j,2}} \\ \frac{\partial f_{i,j,1}}{\partial x_{i,j,2}} & \frac{\partial f_{i,j,2}}{\partial x_{i,j,2}} & 1 & \dots & \frac{\partial f_{i,j,11}}{\partial x_{i,j,2}} & \frac{\partial f_{i,j,12}}{\partial x_{i,j,2}} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ \frac{\partial f_{i,j,1}}{\partial x_{i,j,11}} & \frac{\partial f_{i,j,2}}{\partial x_{i,j,11}} & \frac{\partial f_{i,j,3}}{\partial x_{i,j,11}} & \dots & 1 & \frac{\partial f_{i,j,12}}{\partial x_{i,j,11}} \\ \frac{\partial f_{i,j,1}}{\partial x_{i,j,12}} & \frac{\partial f_{i,j,2}}{\partial x_{i,j,12}} & \frac{\partial f_{i,j,3}}{\partial x_{i,j,12}} & \dots & \frac{\partial f_{i,j,11}}{\partial x_{i,j,12}} & 1 \end{pmatrix}. \tag{8}$$

Here $f_{i,j,1} = [PM_{10}]_j$, $f_{i,j,2} = [PM_{2.5}]_j$, $f_{i,j,3} = [O_3]_j$, $f_{i,j,4} = [CO]_j$, $f_{i,j,5} = [NO]_j$, $f_{i,j,6} = [NO_2]_j$, $f_{i,j,7} = [SO_2]_j$, $f_{i,j,8} = [T]_j$, $f_{i,j,9} = [P]_j$, $f_{i,j,10} = [WD]_j$, $f_{i,j,11} = [WS]_j$, and $f_{i,j,12} = [RH]_j$. Similarly,

$x_{i,j,1} = [PM_{10}]_j$, $x_{i,j,2} = [PM_{2.5}]_j$, $x_{i,j,3} = [O_3]_j$, $x_{i,j,4} = [CO]_j$, $x_{i,j,5} = [NO]_j$, $x_{i,j,6} = [NO_2]_j$, $x_{i,j,7} = [SO_2]_j$, $x_{i,j,8} = [T]_j$, $x_{i,j,9} = [P]_j$, $x_{i,j,10} = [WD]_j$, $x_{i,j,11} = [WS]_j$, and $x_{i,j,12} = [RH]_j$.

We solve Eq. 6 for the unknown association coefficients $\beta_{i,j}$ for the disease-air pollution model Eq. 1 for a disease i in year j . Notice that the diagonal elements of the Jacobian matrix $J_{i,j}$ are one, while the non-diagonal elements represent confounding effects between the environmental factors. If the environmental factors are independent of each other, then the elements from the non-diagonal elements of the Jacobian matrix $J_{i,j}$ will become zero. This reduces $J_{i,j}$ to a simple unitary matrix.

We solved Eq. 6 for $B_{i,j}$ by matrix inversion of the Jacobian matrix $J_{i,j}$, followed by a matrix multiplication,

$$B_{i,j} = J_{i,j}^{-1}A_{i,j}. \quad (9)$$

The elements of the matrix $A_{i,j}$ and the Jacobian matrix $J_{i,j}$ were calculated using the linear regression method³². The effectiveness of the model in the analysis of the relationship (matrix $B_{i,j}$) between air pollution and disease is evaluated against a set of analytical distribution of air pollution with known association coefficients. The results will be shown in the Results section.

Disease air pollution risk R_D . Starr³⁴ measured the fatality risk of activity as total fatalities with respect to total person-hour of exposure to the activity considered. This method leads to a fatality risk as to the statistical probability of fatalities per person-hour of exposure. In this work, we follow Starr³⁴ to compute the risk of increased visits for a disease as the statistical probability of the occurrences of positive association coefficients with respect to all association coefficients.

From Eq. 9, we compute and plot distribution of association coefficients $\beta_{i,j}$ for each of the 18 diseases over 5 years and 3 age groups of patients. For each disease and each age group of outpatients, there are 35 association coefficients $\beta_{i,j}$ (5 years \times 7 air pollution factors) for each age group of outpatients. The statistical significance of the association coefficients are tested using the Student's t test^{29,32}. The detailed distribution of the values of the association coefficients $\beta_{i,j}$ and the P values of the association coefficients for each disease i and each age group of outpatients is shown in the Supplementary Materials.

A positive association coefficient $\beta_{i,j}$ represents an increase in $M_{i,j}$ disease outpatients with respect to an increase in a specific air pollution factor. Hence, the occurrences of total positive and statistically significant association coefficients $\beta_{i,j}$ (P-value < 0.05) with respect to all 35 association coefficients $\beta_{i,j}$ indicates the probability of increases in the outpatient visits for a disease with respect to all pollutants. We use this probability to describe the risk of disease associated with air pollutants, and is computed as followed:

$$R_D = \frac{\text{Total number of occurrences of positive } \beta_{i,j} \text{ for a disease}}{\text{Occurrences of all } \beta_{i,j} \text{ for a disease}} \quad (10)$$

Note that the risk in the context of Eq. 10 essentially means the probability associated with an event of interest—for example, the probability of increased outpatient visits for a disease³⁵. Figure 2 shows a flowchart of air pollution risks computed in this work.

Air pollution health risk R_A . When Eq. 1 is applied through 18 diseases over 5 years, we compute the probability of statistically significant and positive association coefficients $\beta_{i,j}$ (P < 0.05) for each air pollutant with respect to all diseases. Following equation gives air pollution health risk with respect to each air pollutant.

$$R_A = \frac{\text{Total number of occurrences of positive } \beta_{i,j} \text{ from a pollutant}}{\text{Occurrences of all } \beta_{i,j} \text{ from a pollutant}} \quad (11)$$

Daily outpatients, air pollution, and meteorological data. The left-hand side of Eq. 1 requires the input of daily outpatient visits for each of the 18 diseases. The data of daily outpatient visits for the 18 diseases from the Taiwan Landseed Hospital were used in this work^{29,30}. Each medical record contains the residential location of the outpatient. These residential locations were converted to geographical locations (in longitudes and latitudes) for computing the distances between the residential places of the outpatients to the emission sources. The diseases, coded according to the ICD-9 codes (Table 1), during 2007–2011 were used in this study. The hourly air pollution and meteorological data were obtained from the ambient air monitoring station in Chung-Li operated by the Taiwan Environmental Protection Administration (EPA³⁶). The ambient air pollution and meteorological variables were measured at an hourly frequency. The instruments used to make air pollution measurements were calibrated once per day, at the midnight of each day of measurements. The calibration gases were traceable to the United States (US) National Institute of Standards and Technology (NIST). Taiwan EPA strictly follows methods and references set up by the US EPA³⁷ to operate a network of 84 automatic ambient air stations that continuously monitor ambient air pollution in Taiwan. A professional instrument company in Taiwan has operated all 84 automatic ambient air stations. Taiwan EPA has also set up another independent procedure to conduct the audit of all instruments and data quality for each of the ambient air stations by another professional instrument company. This mechanism, an independent operator of the instruments and another independent auditor of the instruments, ensure the data quality, references, and the monitoring methods meet the requirements all the time. All monitoring data are calibrated, online, and openly accessible on the English website (click on Data Service on

the left column of <https://taqm.epa.gov.tw/taqm/en/default.aspx> to obtain hourly monitoring data). Taiwan EPA has been widely used in international studies³⁸.

These hourly data were used to compute daily maximum values for each variable used in this work. The time-series plots of daily air pollution measurements PM_{10} , $PM_{2.5}$, O_3 , CO, NO_2 , NO, and SO_2 ; and temperature data were shown in Wang and Chau²⁹. Tests of normality for air pollutants and meteorological variables were also shown in Wang and Chau²⁹. Tests of normality for data from 18 diseases are shown in the Supplementary Materials.

Results

Verification of the disease-air pollution model. Equation 1 is the core to the air pollution health impact assessment. Hence, we design a set of analytical solutions to test Eq. 1. Following is a list of equations used in this analytical test:

$$\begin{aligned}
 [PM_{10}]_k &= 30 + k \\
 [PM_{2.5}]_k &= 250 + k \times \sin(2k) \\
 [O_3]_k &= 250 + k \times \sin(3k) \\
 [CO]_k &= 250 + k \times \sin(4k) \\
 [NO]_k &= 250 + k \times \sin(5k) \\
 [NO_2]_k &= 250 + k \times \sin(6k) \\
 [SO_2]_k &= 250 + k \times \sin(7k) \\
 [T]_k &= 250 + k \times \sin(8k) \\
 [P]_k &= 250 + k \times \sin(9k) \\
 [WD]_k &= 250 + k \times \sin(10k) \\
 [WS]_k &= 250 + k \times \sin(11k) \\
 [RH]_k &= 250 + k \times \sin(12k)
 \end{aligned} \tag{12}$$

Here k represent Julian days.

By substituting Eq. 12 into Eq. 1, and let association coefficients $\beta_{i,j}$ with respect to each parameter equal to unity, we can write time-series outpatients (OP) for a disease as a function of 12 environmental factors:

$$\begin{aligned}
 OP_{i,k} &= [PM_{10}]_{i,k} + [PM_{2.5}]_{i,k} + [O_3]_{i,k} + [CO]_{i,k} + [NO]_{i,k} \\
 &+ [NO_2]_{i,k} + [SO_2]_{i,k} + [T]_{i,k} + [R]_{i,k} + [WD]_{i,k} + [WS]_{i,k} + [RH]_{i,k}
 \end{aligned} \tag{13}$$

Hence, Eq. 13 shows an analytical distribution of daily outpatient numbers OP with respect to the analytical environmental factors given by Eq. 12. Figure 3 shows time series distribution of outpatients computed according to the analytical solutions, shown in Eq. 13, as each of the 12 analytical functions is included in the outpatient equation for a given year (here shown as 2007).

With the time series OP (from Eq. 13) and 12 parameters (Eq. 12), we compute association coefficients $\beta_{i,j}$, as shown in Eq. 1. Table 2 shows results of calculated associate coefficients, and comparisons with the analytical results. The calculated and analytical results are almost the same, verifying the algorithms we use to compute association coefficients $\beta_{i,j}$.

The association coefficients β shown in Table 2 were calculated from the daily analytical outpatient numbers given by Eq. 12. These β can then be used in Eq. 1 to predict daily time-series outpatient numbers. Figure 4 compares predicted daily outpatient numbers with analytic outpatient numbers. The comparisons show a good agreement between the predicted and analytical outpatient numbers. These results verify the algorithms used in computing the association coefficients from given input daily outpatients numbers and air pollution and meteorological data.

Air pollution source burden per outpatient. The environmental burden of pollution sources in the studied area can be appreciated by examining the spatial distribution between the emission sources and outpatients for diseases. Figure 5(a) shows the spatial distribution of outpatients for the diseases of the respiratory system during the period 2007–2011, while Fig. 5(b) shows a distribution of outpatients with emergency visits. These two figures show a total of 239,610 of outpatients for the respiratory diseases during 2007–2011. Figure 5(c) shows the spatial distribution of the registered industrial emission sources registered on the online governmental system Taiwan Emissions Data Set³⁹. There are 26,132 industrial emissions shown in this figure. Figure 6(d) overlaps emission sources with all outpatients (normal and emergency room visits). It is clear that the spatial distribution of outpatients and air pollution sources are very close to each other.

How close are the emission sources to the location of the residency of outpatients? We calculated the number of registered industrial emission sources with respect to the distance to each outpatient. Figure 6(a) shows a log distribution between the averaged number of industrial emission sources and the distance to an outpatient. On average, there is an industrial emission source within 180 meters of an outpatient. The industrial emissions for an outpatient are 0.334 at 100 meters, 36.135 sources at 1000 meters of distance, and 3370.175 emission sources at 10 km of distance. Figure 6(a) exhibits a log-linear relationship between the emission sources and distances to the patient. The following equation can approximate this relationship:

$$y = C_1 \times (x^{r_1}) \tag{14}$$

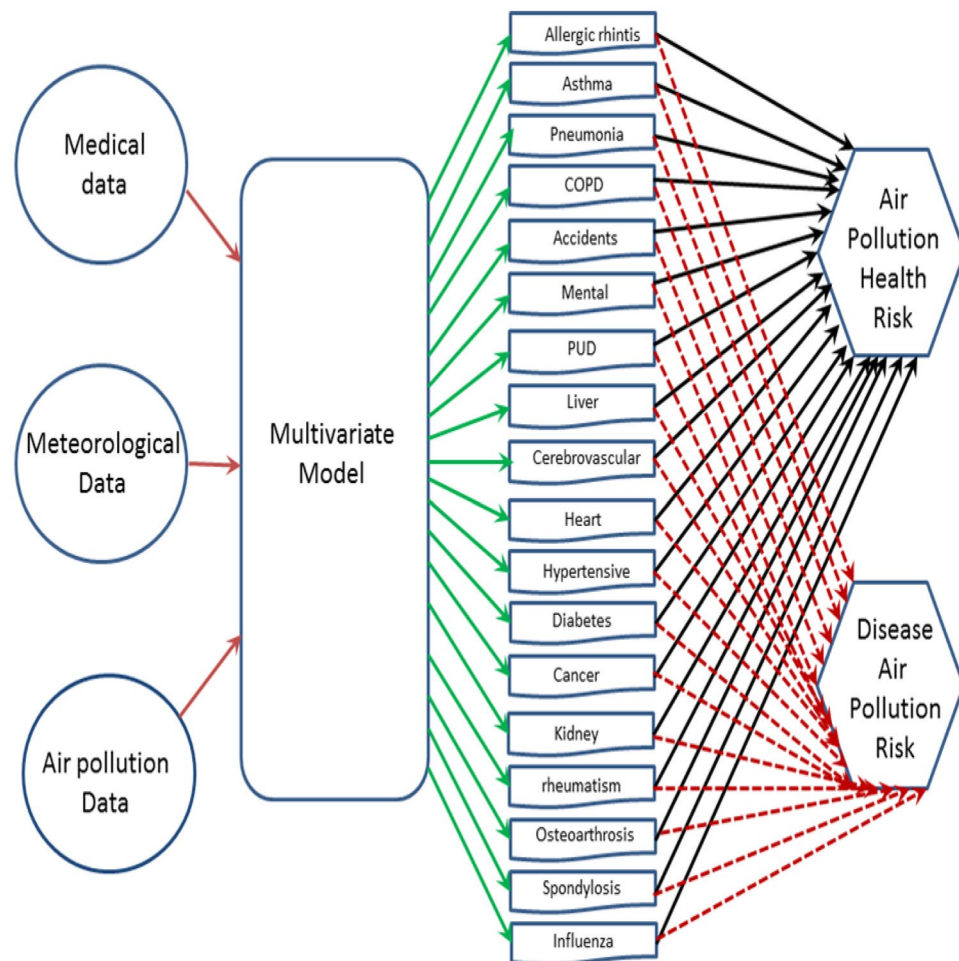


Figure 2. Calculation of the disease air pollution risks (R_D), and the air pollution health risks (R_A) in this work. The inputs are medical data from Taiwan Landseed Hospital, and air pollution and meteorological data from EPA ambient air monitoring station in Chung-Li. The multivariate model is the diseases-air pollution model used in this work.

Here $C_1 = 3.31 \times 10^{-5}$, $r_1 = 2.002$, y is the number of emission source, and x is the distance to an outpatient. Verification of this equation at $x = 1000$ meters, predicts $y = 33.551$ emission source. The real emission sources at 1000 meters are 36.135. Hence the errors from this equation are 7.2%.

On the other hand, how many outpatients are exposed to registered industrial emissions? Figure 6(b) shows a log distribution of averaged outpatient numbers with respect to the distance to an emission stack from 100 meters to 10 kilometers. Figure 6(b) also exhibits a log-linear relationship between the outpatient numbers and distance to the industrial emission source. There are about 3 outpatients at 100 distance from an industrial emission source, and 30,236 outpatients at 10 kilometers from an industrial emission source. The following equation can approximate this relationship:

$$y = C_2 \times (x^{r_2}) \quad (15)$$

Here $C_2 = 0.0003$, $r_2 = 2.002$.

Hence, on average, there are 3 outpatients from an industrial emission source at 100 meters of distance; and one industrial emission source from an outpatient at 150 meters of distance.

Disease distribution in outpatients: an age-based perspective. Table 3 shows a list of total outpatient visits according to disease and year during 2007–2011. A total of 1.78 million outpatient visits during this period. The most substantial 26.46% of outpatients were resulting from the diseases of the circulatory system (7.63% from the cerebrovascular diseases, 7.76% from the heart diseases, and 11.07% from the hypertensive diseases). The second largest outpatients of 19% were from the diseases of the musculoskeletal system and connective tissue. Followed by 13.80% of outpatients from the diseases of the genitourinary system, and 10.32% from the accidents. A 10.12% from the diseases of the respiratory system, and 9.31% from the diseases of endocrine disorders (diabetes mellitus). A 4.95% from the diseases of the digestive system, and 4.68% from the diseases of

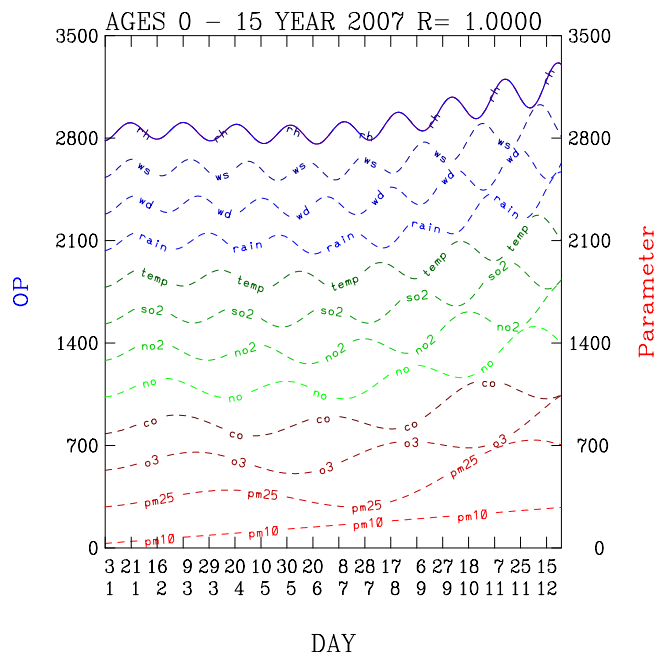


Figure 3. Time-series distribution of 12 environmental factors and total outpatients for a disease. The magnitudes of the OP are increased as the contribution from each of the environmental factors are gradually added.

β	Computed	Analytical Solution
β_{12}	1.00007451	1
β_{11}	1.00012779	1
β_{10}	1.00016189	1
β_9	1.00018036	1
β_8	1.00018334	1
β_7	1.00016534	1
β_6	1.00014186	1
β_5	1.00013232	1
β_4	1.00010610	1
β_3	1.00005054	1
β_2	1.00002706	1
β_1	1.00000393	1
β_0	-0.33797091	0

Table 2. List Computed and Analytical Association Coefficients β .

mental disorders (anxiety, dissociative and somatoform disorders). A 2.95% from the diseases of malignant neoplasm (cancers), and 0.89% from influenza.

These statistics reveal that the diseases of the circulatory systems, connecting tissues in bones, and genitourinary systems account for nearly 60% of all outpatients; and account for 80% of all outpatients when outpatients of accidents and respiratory systems are added.

Figure 7 shows daily outpatients for the 18 diseases from 2007 to 2011 and each age group of outpatients. The youngest group outpatients (0–15 years old) were mostly suffered from the diseases of respiratory systems and accidents. The leading causes of diseases of the respiratory systems were allergic rhinitis and pneumonia than, followed by asthma and the COPD. There were few visits for the young bodies from the diseases of mental disorders, digestive system, circulatory system, diabetes, cancer, genitourinary system, musculoskeletal system.

For the 16–65 years old outpatients, the diseases have spread to the entire body. The leading causes of diseases were resulting from mental disorders (anxiety, dissociative, and somatoform disorders), the circulatory system, and the musculoskeletal system. The leading causes of the diseases of the respiratory system are allergic rhinitis, asthma, and COPD. Pneumonia is not a significant cause of the respiratory disease as the 0–15 years old outpatients.

For the most senior group of the outpatients (higher than 65 years old), the leading causes of diseases were resulting from the circulatory system, diabetes, genitourinary system, and musculoskeletal system. Diseases of

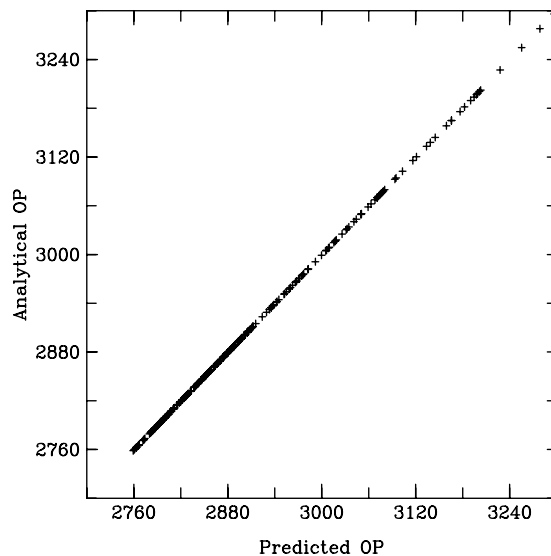


Figure 4. A scattered plot distribution of predicted OP (x-axis) versus analytical OP (y-axis).

the digestive system and respiratory system, accident, and mental disorder are comparatively low. Apparently, for those who can survive into old ages, their mental disorders are less a problem than the physical functionality of a body.

Disease air pollution risks. Figure 8 shows computed the distribution of disease air pollution risks R_D with respect to a spectrum of 18 diseases and 8 time-lags scenarios between the hospital visits for diseases and environmental factors in 3 age groups of outpatients.

For the 0–15 years old patients, higher than 40% of air pollution risks are mostly associated with the diseases of the respiratory system (pneumonia), accidents, and influenza. Immediate air pollution risks are mostly associated with AR (35–40%), pneumonia (35–40%), and influenza (40–45%). One-day delay air pollution risks are associated with pneumonia (45–50%) and influenza (45–50%). Pneumonia and influenza are two diseases that exhibit the most persistent occurrence of risks that are mostly higher than 40%.

We note that the air pollution risks for the following diseases are below 20% for the 0–15 years old patients in all 8 scenarios: mental (no mental problem yet), chronic liver and cirrhosis (liver is still excellent), heart and hypertensive (heart working very good and blood vessels are very clean), diabetes (no abnormality in body sugar production, regulation, and accumulation), and kidney (the genitourinary system is in good shape).

The occurrences of elevated air pollution risks for ages 0–15 outpatients are ranked as pneumonia (6), influenza (3), and accidents (1). The most sensitive diseases for air pollution risks for young outpatients is influenza. A total of 3 diseases out of 18 diseases showing high air pollution risks ($R_D > 40\%$).

For 0–15 years old outpatients, diseases with low health risks ($R_D < 20\%$) are mostly from the mental, digestive system (PUD, liver), circulation system, and cancer, and kidney. Diseases with high health risks ($R_D > 40\%$) are mostly from the respiratory system and influenza. Key air pollutants are $PM_{2.5}$, O_3 , NO, PM_{10} , and NO_2 .

For the 16–65 years old outpatients, there exist more widespread occurrence of air pollution risks to the entire body than the 0–15 years old outpatients. The model shows that diseases of the respiratory system (pneumonia and COPD), accident, one disease of the digestive system (PUD), two diseases of the circulatory system (cerebrovascular and heart), three diseases of the musculoskeletal system, and influenza all exhibit higher the 40% of air pollution risks. The immediate impacts of elevated air pollution risk ($R_D > 40\%$) occur in COPD, accidents, PUD, OA, and spondylosis. The 1-day delay impacts of elevated air pollution risks ($R_D > 40\%$) occur in COPD, accidents, and influenza. For 2-day to 6-day delay impacts of air pollution risks ($R_D > 40\%$), these are diseases mostly in pneumonia, COPD, accidents, PUD, cerebrovascular, heart, rheumatism, OA, spondylosis (2-day delay), and influenza. Up to 7-day delay of air pollution risks ($R_D > 40\%$) are seen in pneumonia, accidents, OA, and spondylosis.

The occurrences of elevated air pollution risk ($R_D > 40\%$) for the 16–65 years old outpatients are ranked as accidents (7), spondylosis (7), COPD (5), OA (4), influenza (5), pneumonia (3), PUD (4), cerebrovascular (2), heart (1), and rheumatism (1). This group of outpatients is most sensitive to air pollution risks for diseases of COPD, accidents, PUD, OA, and spondylosis. For the 16–65 years old outpatients, a total of 10 diseases out of 18 diseases analyzed here showing high air pollution risks ($R_D > 40\%$), exhibiting the highest distribution of disease air pollution risks.

For the outpatient with ages higher than 65 years old, elevated air pollution risks ($R_D > 40\%$) mainly distributed in the diseases of the respiratory system (COPD), diseases of the circulatory system (heart, hypertensive, and cerebrovascular diseases), cancer, and musculoskeletal system (spondylosis disease). Immediate impact (0-day delay) of elevated air pollution risks ($R_D > 40\%$) concentrate on the diseases of the circulatory system (heart and

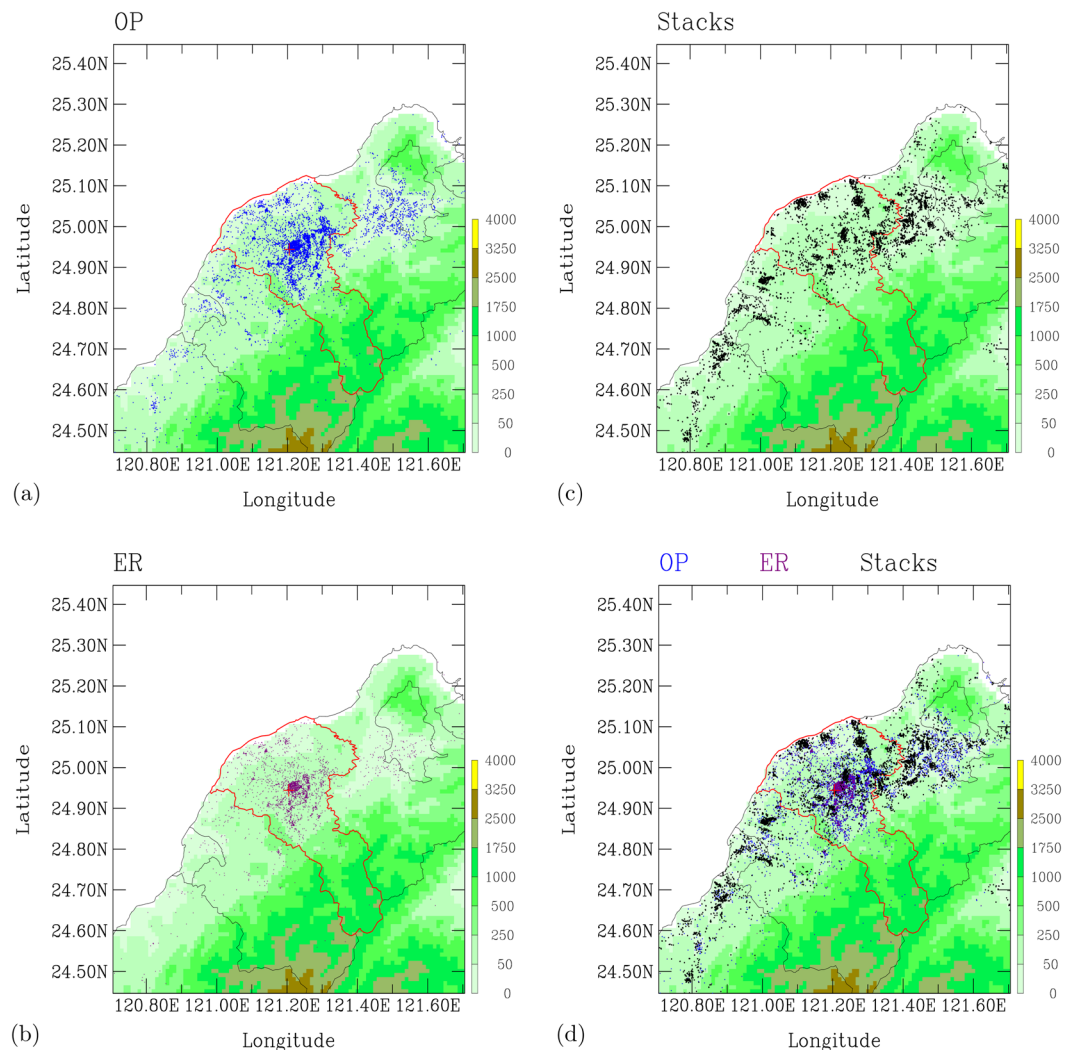


Figure 5. Spatial distribution of outpatients for the respiratory diseases and industrial emission source during 2007–2011. **(a)** Outpatients. **(b)** Outpatients from emergency room visits. **(c)** Registered industrial source emissions. **(d)** Composite of outpatients and industrial emission sources. Shaded colors indicate terrain heights, in units of meters.

hypertensive), and spondylosis disease. The 7-day delayed effects of air pollution risks ($R_D > 40\%$) appear in the diseases of cerebrovascular, cancer, and spondylosis. The most prevalent diseases with elevated air pollution risks ($R_D > 40\%$) are cerebrovascular diseases and spondylosis diseases.

The occurrences of elevated air pollution risks ($R_D > 40\%$) for the elderly outpatients are from diseases of spondylosis (8), cerebrovascular (6), cancer (1), hypertensive (1), heart (1) and COPD (1).

The diseases of the circulatory system (cerebrovascular) and the musculoskeletal system (spondylosis) of the elderly outpatients exhibit the most sensitive and the most persistent of up to 7 days of the air pollution risks. A total of 6 diseases show the occurrence of high air pollution risks (more than 40%) from 18 diseases.

Air pollution health risks. Figure 9 shows air pollution health risks R_A with respect to 7 pollutants and 8 time-lag scenarios in 3 groups of outpatients. For young outpatients (0–15 years old), occurrences of the 20%–40% of health risks at time lags of 1–7 days are mostly associated with $PM_{2.5}$, NO , O_3 , and NO_2 . The immediate impacts (0-day delay) on health risk (20%–40%) are associated with PM_{10} and NO .

For the 16–65 years old outpatients, elevated health risks ($R_A > 40\%$) mostly occur at time lags of 1–7 days and are associated with $PM_{2.5}$, O_3 , NO_2 , and CO . Occurrences of immediate impacts on health risks ($R_A = 20\%$ –40%) are associated with PM_{10} , $PM_{2.5}$, O_3 , CO , NO and NO_2 .

For the elderly outpatients (>65 years old), immediate health risks ($R_A = 20\%$ –40%) are mostly associated with PM_{10} , CO , NO , and NO_2 . The health risks R_A at time lags of 1–7 days are prevalently associated with NO_2 , followed by associations with $PM_{2.5}$, O_3 , CO , NO , PM_{10} , and SO_2 .

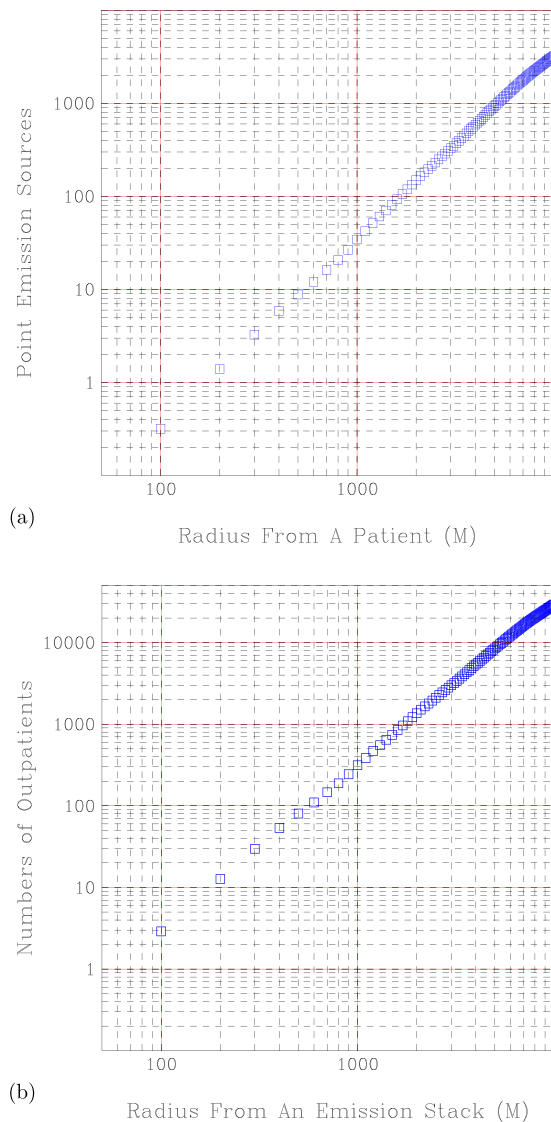


Figure 6. A scattered plot analysis of spatial closeness between emission sources and outpatients during 2007–2011. **(a)** Averaged number of industrial emission sources versus distance from an outpatient. **(b)** Averaged number of Outpatients versus distance from an industrial emission source.

Above results reveal that $PM_{2.5}$ is the most hazardous air pollutant from the analysis of outpatient data. The second most hazardous air pollutant is O_3 , followed by NO_2 and CO. Air pollutants PM_{10} and NO are most active than other air pollutants in producing immediate (0-day) delay health impacts ($R_A = 20\%–40\%$) for all outpatients.

Health risks: a whole body perspective. Our analysis revealed that the health risks of air pollution depends on types of air pollutants, physical ages of outpatients, and the response of the body to changing air pollution concentrations. Figure 10 summarizes the distribution of air pollution health risks (R_A) and disease health risks (R_D) with respect to immediate (0 Day) and delayed (1–7 Days) time responses in 3 age groups of outpatients.

For the 0–15 years old outpatients, PM_{10} and NO contribute to the immediate impact of the health risks, while $PM_{2.5}$, O_3 , NO , and NO_2 contribute to the time lags of 1–7 days of health risks. Diseases of immediate and with 20%–40% health risks are AR, pneumonia, cerebrovascular, cancer, rheumatism, OA, and influenza ($R_D > 40\%$). Diseases with time lags of 1–7 days and with the 20%–40% health risks are AR, asthma, pneumonia (high health risks), COPD, accidents ($R_D > 40\%$), PUD, cerebrovascular, cancer, rheumatism, OA, spondylosis, and influenza ($R_D > 40\%$). As the body response time increases from 0-day to 1–7 days, representative air pollutants for diseases change from PM_{10} and NO to $PM_{2.5}$, O_3 , NO , and NO_2 . Changes of diseases with increases in health risks and increases in body response times are asthma, pneumonia, COPD, accidents, and spondylosis. Hence, $PM_{2.5}$, O_3 , and NO_2 correspond to increases in health risks for asthma, pneumonia, COPD, accidents, and spondylosis.

Disease	2007	2008	2009	2010	2011	Total	Percent
Respiratory system							10.12
Allergic Rhinitis	5118	6803	8413	9041	9021	38396	2.15
Asthma	7779	7475	7910	8988	12458	44610	2.50
Pneumonia	7878	7036	7066	6989	9051	38020	2.13
COPD	11061	11225	12159	12082	13076	59603	3.34
Accidents	37228	36703	34782	37147	38264	184124	10.32
Mental disorders	15437	15078	16770	16776	19490	83551	4.68
Digestive system							4.95
Peptic ulcer	9106	7839	7817	6453	6807	38022	2.13
Chronic liver/cirrhosis	11222	10309	10247	9310	9239	50327	2.82
Circulatory system							26.46
Cerebrovascular disease	22024	25963	28388	29434	30334	136143	7.63
Heart disease	24151	26173	28107	26365	33671	138467	7.76
Hypertensive disease	38511	42048	42591	40240	34089	197479	11.07
Diabetes mellitus	28662	29382	32705	35325	39906	165980	9.31
Malignant neoplasm	9219	10103	9457	10098	13655	52532	2.95
Genitourinary system	44595	50552	50367	49677	50965	246156	13.80
Musculoskeletal system							19.00
Other disorders of soft tis.	11440	11814	19203	22400	28963	93820	5.26
Osteoarthritis/allied dis.	25645	26063	22152	26545	32890	133295	7.47
Spondylosis/allied dis.	16355	21102	23518	24707	26075	111757	6.27
Influenza vaccine injection	164	54	11836	2334	1507	15895	0.89
Yearly Total	317816	338247	365578	364923	397003	1780070	100.00

Table 3. List of Total Outpatient Visits According to Disease and Year.

For the 16–65 years old outpatients, air pollutants with the immediate health risks of 20%–40% are PM_{10} , $PM_{2.5}$, O_3 , NO , and NO_2 . Diseases with immediate and health risks of 20%–40% are asthma, COPD ($R_D > 40\%$), accidents ($R_D > 40\%$), PUD ($R_D > 40\%$), cerebrovascular, heart, hypertensive, rheumatism, OA ($R_D > 40\%$), spondylosis ($R_D > 40\%$), and influenza. The highest health risks ($R_A > 40\%$) are associated with the increases in time lags of 1–7 days and are associated with $PM_{2.5}$, O_3 , CO , and NO_2 . The main changes of air pollutants with increasing health risks are $PM_{2.5}$, O_3 , CO , NO_2 , and SO_2 ($R_A = 20\%$ – 40%). Increases in time responses to time lags of 1–7 days resulting in the increases of health risks of diseases of AR, pneumonia ($R_D > 40\%$), mental, liver, cerebrovascular ($R_D > 40\%$), heart ($R_D > 40\%$), diabetes, cancer, kidney, rheumatism ($R_D > 40\%$), and influenza ($R_D > 40\%$).

For the elderly outpatients, PM_{10} , CO , NO , and NO_2 are associated with the immediate health risk of 20%–40%, while all 7 seven air pollutants are associated with the 1–7 days delay of health risks ($R_A = 20\%$ – 40%). Air pollutants $PM_{2.5}$, O_3 , and SO_2 have become important in the health risks ($R_A = 20\%$ – 40%) as the time lags increase to 1–7 days. The corresponding changes of diseases with increases in health risks are AR, pneumonia, COPD ($R_D > 40\%$), accidents, PUD, cerebrovascular ($R_D > 40\%$), diabetes, cancer ($R_D > 40\%$), kidney, and OA.

Figure 10 shows that the most hazardous air pollutants for health risks of $R_A > 40\%$ are $PM_{2.5}$, O_3 , CO , and NO_2 . Diseases with $R_D > 40\%$ of health risks are pneumonia, COPD, accidents, PUD, heart, hypertensive, cerebrovascular, heart, cancer, rheumatism, OA, spondylosis, and influenza. The time lags of 1–7 days of health risks are accompanied by more diseases and higher risks than the immediate health risks. The 16–65 years old outpatients are most susceptible to elevated health risks for more diseases across the body from air pollution than the other two groups of outpatients.

For the 0–15 years old outpatients, concurrent occurrences of elevated health risk of $R_D > 40\%$ is influenza for immediate health risks; and diseases of pneumonia, accidents, and influenza for time lags of 1–7 days of health risks. For the 16–65 years old outpatients, concurrent occurrences of elevated health risk of $R_D > 40\%$ are diseases of COPD, accidents, PUD, OA, and spondylosis for the immediate health risks; and diseases of pneumonia, COPD, accidents, PUD, cerebrovascular, heart, rheumatism, osteoarthritis, spondylosis, and influenza for time lags of 1–7 days of health risks. For the elderly outpatients, concurrent occurrences of elevated health risk of $R_D > 40\%$ are diseases of heart, hypertensive, and spondylosis for the immediate health risks; and diseases of COPD, cerebrovascular, cancer, and spondylosis for time lags of 1–7 days of health risks. Hence, Fig. 10 validates our hypothesis that air pollution has widespread and concurrent impact on diseases from various organs in the body.

Discussion

Our results of the elevated health risks are consistent with previous findings that diseases of the respiratory and circulatory systems were found to be associated with the air pollution^{23,40–42}. Long-term air pollution exposure was found to be an important risk factor for cancers^{43,44}.

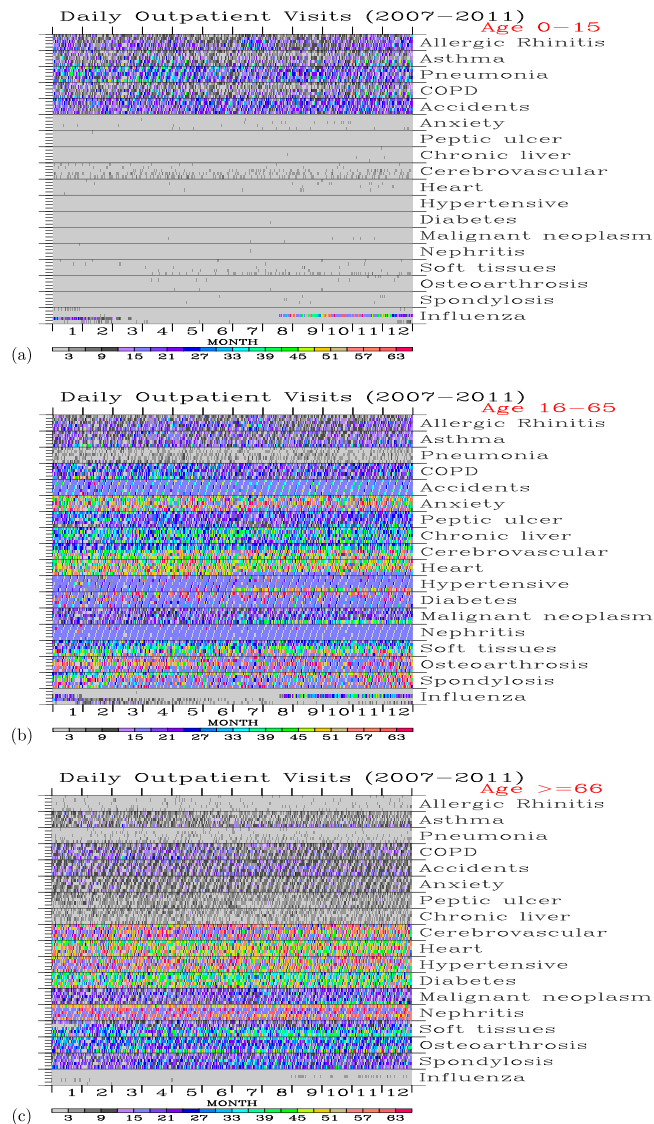


Figure 7. Daily counts of outpatient visits (shaded colors) for the (a) 0–15, (b) 16–65, and (c) above 65 year-old age group of people for each of the ICD-9 diseases during the 2007–2011 period.

For the association of diseases of the digestive system with air pollution, studies showed that air pollution affects the gut microbiota and linked to the inflammatory disease^{45–49}, and confirmed from experiments with mice^{47,50,51}. Inhaled air pollutants are quickly cleared from the lungs and transported by the mucociliary into the intestine⁵². Cytotoxin substance produced by gut bacteria *Campylobacter pylori* were shown to present in patients with PUD than those with chronic gastritis only⁵³. The human gut is populated with as many as 100 trillion microbes⁵⁴. The human-gut microbiota interactions producing host-microbiota metabolism, signaling, and immune-inflammatory processes that physiologically connect the gut, liver, muscle, and brain⁵⁵, affecting health and diseases^{56,57}, and is now an innovative method for fighting against cancer⁵⁸. Hence, elevated health risks of PUD found in this work concurs with evidence found in previous works.

For NO, which is the Molecule of the Year 1992⁵⁹, participates in the widespread signaling of cellular and organ function in the body⁶⁰. NO is a unique signaling molecule in the vascular system⁶¹ that touches nearly all areas of life⁶². Also, Maher *et al.*¹³ found that NO_2 and $PM_{2.5}$ were positively associated with the incidence of dementia in London, England.

For the links between air pollution and brain and accidents, studies showed the pathways for air pollutants to attack the brain central nervous system include (1) by entering the nose and travel through the olfactory bulb to enter the brain; (2) by affecting the lining of the nasal epithelium, causing the release of the inflammatory cytokines that damage the brain; (3) by entering the lung, activating the release of the proinflammatory cytokines that transported by the blood to the brain area to cause adverse brain effects²⁰. As such, exposure to air pollution is a contributor to the decline of cognitive function⁶³. Functional near-infrared spectroscopy (fNIRS) measurements of brain activity of train drivers confirm variations of brain waves during various stages of driving processes⁶⁴. Neuropsychological tests showed that exposure to CO leads to dysfunctions in memory, attention and

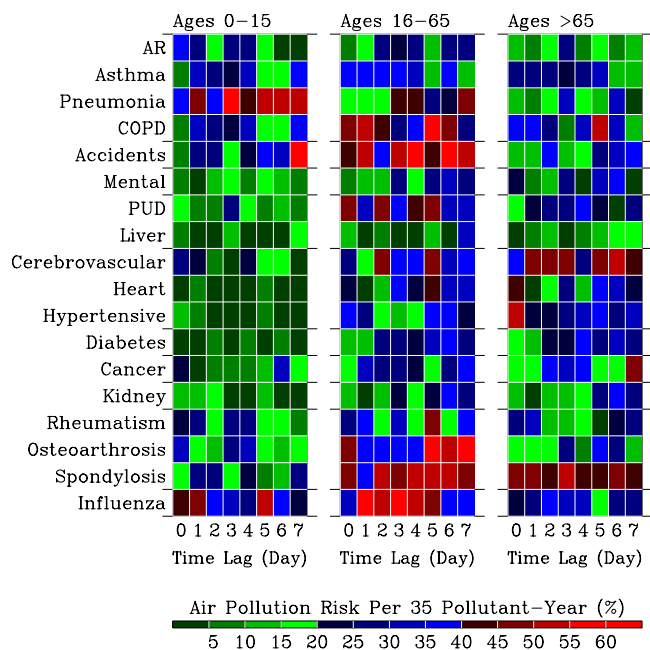


Figure 8. Disease air pollution risks R_D . Distribution of R_D with respect to a spectrum of 18 diseases and 8 time-lags scenarios in 3 age groups of outpatients.

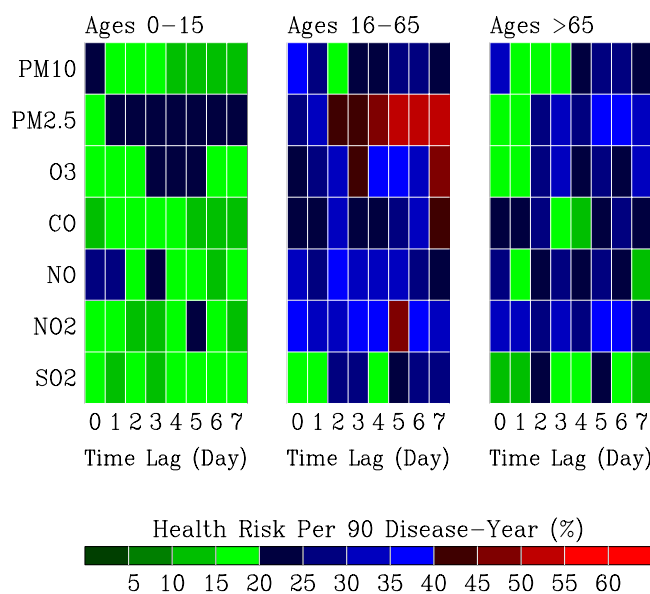


Figure 9. Air pollution health risks R_A . Distribution of R_A with respect to 7 pollutants and 8 time-lag scenarios in 3 groups of outpatients.

concentration, tracking skills, visuomotor skills, visuospatial planning and processing⁶⁵. CO exposure was tested to show a highly significant deficit in careful driving skills⁶⁶. Experiments have confirmed that abnormally low numbers of hypocretin-producing neurons in the brain, resulting in the reduced levels of protein hypocretin⁶⁷, which in turns signal the bone marrow to increase levels of white blood cells that can migrate to arterial walls, leading to the formation of plaque and atherosclerosis⁶⁸.

Immune pathways that link fever and gut immunity are shown in a study demonstrating trafficking of immune cells to enhance immune response and fight infection⁶⁹. T lymphocyte cell receptor called α -4 integrins when combined to the heat-shock proteins (hsps), produced by T cell during the fever condition, was able to pull together hsps activated T cells and stick on the blood-vessel walls to infection sites⁶⁹. This α -4 integrins-related pathway could have important roles in inflammation and gut immunity.

Our analysis shows that the most frequent occurrences of diseases associated with elevated health risks of $R_D > 40\%$ are diseases of spondylosis (in bones), COPD (in the lung), accidents (in the brain), cerebrovascular

Risk Types	Time Lag	Ages 0–15	Ages 16–65	Ages > 65
Air Pollution Health Risks RA	0 Day	PM10 PM2.5 O3 CO NO NO2 SO2	PM10 PM2.5 O3 CO NO NO2 SO2	PM10 PM2.5 O3 CO NO NO2 SO2
	1–7 Days	PM10 PM2.5 O3 CO NO NO2 SO2	PM10 PM2.5 O3 CO NO NO2 SO2	PM10 PM2.5 O3 CO NO NO2 SO2
Disease Air Pollution Risks RD	0 Day	AR Asthma Pneumonia COPD Accidents Mental PUD Liver Cerebrovascular Heart Hypertensive Diabetes Cancer Kidney Rheumatism Osteoarthritis Spondylosis Influenza	AR Asthma Pneumonia COPD Accidents Mental PUD Liver Cerebrovascular Heart Hypertensive Diabetes Cancer Kidney Rheumatism Osteoarthritis Spondylosis Influenza	AR Asthma Pneumonia COPD Accidents Mental PUD Liver Cerebrovascular Heart Hypertensive Diabetes Cancer Kidney Rheumatism Osteoarthritis Spondylosis Influenza
	1–7 Days	AR Asthma Pneumonia COPD Accidents Mental PUD Liver Cerebrovascular Heart Hypertensive Diabetes Cancer Kidney Rheumatism Osteoarthritis Spondylosis Influenza	AR Asthma Pneumonia COPD Accidents Mental PUD Liver Cerebrovascular Heart Hypertensive Diabetes Cancer Kidney Rheumatism Osteoarthritis Spondylosis Influenza	AR Asthma Pneumonia COPD Accidents Mental PUD Liver Cerebrovascular Heart Hypertensive Diabetes Cancer Kidney Rheumatism Osteoarthritis Spondylosis Influenza



Figure 10. Distribution of air pollution health risks (R_A) and disease health risks (R_D) with respect to immediate (0 Day) and delayed (1–7 Days) time responses in 3 age groups of outpatients.

(in blood vessels), and influenza (brain linked to raising body temperature to kill invaded pathogen associated with air). The concurrence of these diseases in various organs of the body indicate the connections that had occurred within the body.

What mechanisms that can cause such a wide spread of impacts on diseases with increasing air pollution levels? Why the concurrent increase in the diseases of the musculoskeletal system, digestive system, and influenza? In a whole body perspective, the widespread distribution of diseases in various part of body organs indicate that the body immune system and the brain has been affected by the increased in concentrations of air pollution. In the context of the immune system⁷⁰, the neurogenic inflammation causes the diseases of the musculoskeletal system triggered by the immune system’s response to the foreign toxins from inhaled air pollutants⁷¹. Air pollution caused oxidative damage and inflammation and resulted in the cytotoxic responses in the body^{72–76}. Air pollution was found to cause inflammation, leading to accelerated bone loss and increased risk of bone fracture and osteoporosis⁷⁷. Exposure to air pollution is associated with oxidative damage to DNA and lipids in humans⁷⁸.

Hence, the inflammation processes are the responses of the immune system to external air pollution. Cytokines, mediators (messengers) involved in cell growth and activation, are produced during the processes of inflammation and immunity responses in the body^{79,80}. Concentrations of pro-inflammatory cytokines such as tumor necrosis factor (TNF)- α (TNF – α), interleukin (IL)-6 (IL-6), chemokine macrophage inflammatory protein 2 (MIP-2), and NO were measured in the cell culture where macrophages of mice were exposed to $PM_{2.5}$ and PM_{10} ⁷³. Pro-inflammatory cytokines TNF – α , IL-1, IL-6, granulocyte macrophage colony stimulating factor (GM-CSF), and chemokines IL-8 are abundant in patients with rheumatoid arthritis (RA)⁷⁹. The cardiovascular diseases, resulting from the plaque accumulation on the vascular wall, are associated with the active IL-1 β produced by vascular wall cells and lesional leukocytes⁸¹. Cytokines IL-1, IL-6, IL-8, IL-17, leukocyte inhibitory factor (LIF), TNF- α , interferon (IFN)- γ (INF- γ), produce cartilage destruction⁸². Hence, chemicals that are able to inhibit the pro-inflammatory cytokines IL-1, TNF- α , interferon (IFN)- γ (INF- γ), IFN- β ; to stimulate the production of anti-inflammatory cytokines IL-4, IL-6, IL-10, IL-13, transforming growth factor (TGF)- β and insulin-like growth factor (IGF)-1 are potentially good for OA therapy⁸². Since IL-1 caused an increase in DNA damage of chondrocytes in OA cartilage⁸³, inhibition of IL-1 is proposed as a strategy for the treatment of RA⁸⁴.

Impacts of inhaled toxins such as smoke were found to be associated with the loss of bone mineral density and risk of hip fracture^{85,86}.

Therefore, the first line of defense against air pollution is skin and mucous membranes^{87–89}. The second line of defense against air pollution is phagocytes, antimicrobial substances, natural killer cells (NK cells), inflammation, and fever. The third line of defense against invading air pollution are B lymphocytes, T lymphocytes, and Antigen-presenting cells (APCs)⁷⁰. If further studies confirm our results shown here, then the degradation of the immune system are severe health risks that were overlooked before. Indeed, studies have shown excessive antibiotic usage in Taiwan⁹⁰, consistent with elevated pollution combined with high density of the population.

In the context of quality of air means the quality of life⁹¹, the closeness of outpatients and industrial emission sources reveal very grim pictures on the quality of health for people living in an industrial city, as shown in this work. This work was deeply inspired by the seminal work of John Snow on the analysis of the 1854 London cholera outbreak^{92,93}. The impacts of air pollutants in biological organisms are profound and heritable, as shown from mice placed 1-km downwind of steel mills produced 1.5- to 2.0-fold of heritable DNA mutations in the pups⁶, and other pathophysiological studies reviewed in this work. The strength of this work is that we can learn more about the impacts of industrial emissions on public health if we study all diseases associated with people as they exposed, breathed, and aged^{22,25,26}.

The limitation of this study is that more physiological evidence is needed to further understand the mechanisms hidden in the mountain of health and environmental data. New milestones in the health risk assessment of air pollution have been achieved in recent years. These milestones were propelled by the advancements in new tools such as computational hardware (e.g. GPU computers), diversified observational data⁷, non-conventional environmental data obtained and powered by the development of open-source smart sensors⁹⁴, experimental methods⁶, artificial intelligence⁹⁵, and big data analysis for public health⁹³. Application of these new tools can better quantify the health impact of air pollution on a finer spatial-temporal resolution than the environmental data currently available by sparse measurements.

Conclusions

In this work, we showed a very polluted industrial city, characterized by between 3 outpatients per 100 meters of an industrial emission source and an industrial emission source per 180 meters to an outpatient. A total of 1.78 million outpatients over a spectrum of 18 diseases had visited the Landsead Hospital for medical services during 5 years of 2007–2011. We calculated that the most hazardous air pollutants are (in health risk order) $PM_{2.5}$, NO_2 , PM_{10} , O_3 , CO, and NO.

In the whole body perspective, the diseases of the respiratory system, the musculoskeletal system, and the circulatory systems exhibit the highest air pollution risks. Diseases of the musculoskeletal system reveal the degradation of the immune system and central nervous systems that have been constantly bombarded by the imported toxic materials as people continuously exposed and breathed in a very polluted atmospheric environment. This impact of air pollution on the body immune system is an important finding of this work.

Accession codes. All data used in this work are openly available at <https://doi.org/10.6084/m9.figshare.7217459>.

Received: 9 May 2019; Accepted: 19 January 2020;

Published online: 11 February 2020

References

1. WHO. AirQ+: key features, WHO Regional Office for Europe (2016).
2. Anderson, H. R. Air pollution and mortality: A history. *Atmos. Environ.* **43**, 142–152 (2009).
3. Dockery, D. W. *et al.* An association between air pollution and mortality in six U.S. cities. *N. Engl. J. Med.* **329**(24), 1753–1759 (1993).
4. Heinsohn, R. J. & Kabel, R. L. *Sources and control of air pollution*. (Prentice Hall, Upper Saddle River, New Jersey, 1999).
5. Chen, J. C. *et al.* Ambient air pollution and neurotoxicity on brain structure: Evidence from Women's Health Initiative Memory Study. *Ann. Neuro.* **78**, 466–476 (2015).
6. Somers, C. M., Yauk, C. L., White, P. A., Parfett, C. L. & Quinn, J. S. Air pollution induces heritable DNA mutations. *P.N.A.S.* **99**(25), 15904–15907, <https://doi.org/10.1073/pnas.252499499> (2002).
7. DuBay, S. G. & Fuldner, C. C. Bird specimens track 135 years of atmospheric black carbon and environmental policy. *P.N.A.S.* **114**(43), 11321–11326, <https://doi.org/10.1073/pnas.1710239114> (2017).
8. Ren, A. *et al.* Association of selected persistent organic pollutants in the placenta with the risk of neural tube defects. *P.N.A.S.* **108**(31), 12770–12775, <https://doi.org/10.1073/pnas.1105209108> (2011).
9. Brown, A. S. *et al.* Association of maternal insecticide levels with autism in offspring from a national birth cohort. *Am. J. Psychiatry.* PMID, 30111184, <https://doi.org/10.1176/appi.api.2018.17101129> (2018).
10. Cohn, B. A. *et al.* DDT exposure in utero and breast cancer. *J. Clin. Endocrinol. Metab.* **100**, 2865–2872 (2015).
11. Woodward, N. C. *et al.* Traffic-related air pollution impact on mouse brain accelerates myelin and neuritic agin changes with specificity for CA1 neurons. *Neurobiol. Aging* **53**, 48–58, <https://doi.org/10.1016/j.neurobiolaging.2017.01.007> (2017).
12. Morgan, T. E. *et al.* Glutamatergic neurons in rodent models respond to nanscale particulate urban air pollutants *in vivo* and *in vitro*. *Environ. Health Perspect.* **119**(7), 1003–1009, <https://doi.org/10.1289/ehp.1002973> (2011).
13. Maher, B. A. *et al.* Magnetite pollution nanoparticles in the human brain. *P.N.A.S.* **113**(39), 10,797–10,801, <https://doi.org/10.1073/pnas.1605941113> (2016).
14. Yan, W., Yun, Y., Ku, T., Li, G. & Sang, N. NO_2 promotes Alzheimer's disease-like progression: cyclooxygenase-2-derived prostaglandin E_2 modulation and monoacylglycerol lipase inhibition-targeted medication. *Sci. Rep.* **6**, 22429, <https://doi.org/10.1038/srep.22429> (2016).
15. Carey, I. M. *et al.* Are noise and air pollution related to the incidence of dementia? A cohort study in London, England. *B.M.J. Open* **8**, e022404, <https://doi.org/10.1136/bmjopen-2018-022404> (2018).
16. Block, M. L. & Calderón-Garcidueñas, L. Air pollution: Mechanisms of neuroinflammation & CNS disease. *Trends. Neurosci.* **32**(9), 506–516, [10.1016/j.tins.2009.05.009](https://doi.org/10.1016/j.tins.2009.05.009) (2009).

17. Sternberg, E. M. & Gold, P. W. The mind-body interaction in disease. *Sci. Am.* **7**, 8–15 (1997).
18. Drew, L. An age-old story of dementia. *Nature* **559**, S2–S3, <https://doi.org/10.1038/d41586-018-05718-5> (2018).
19. Sternberg, E. M. *The Balance Within: The Science Connecting Health and Emotions*. (W.H. Freeman and Company, New York, 2000).
20. Underwood, E. The polluted brain. *Science* **355**(6323), 342–345, <https://doi.org/10.1126/science.355.6323.342> (2017).
21. Brunekreef, B. & Holgate, S. T. Air pollution and health. *Lancet* **360**, 1233–1242 (2002).
22. Pearce, N. Traditional epidemiology, modern epidemiology, and public health. *Am. J. Public Health* **86**, 678–683 (1996).
23. Kampa, M. & Castanas, E. Human health effects of air pollution. *Environ. Pollut.* **151**, 362–367 (2008).
24. Jay, S. Are the adverse human effects of air pollution underestimated in the literature? Implications for science, medicine and public policy. Health Supplemental Files, Paper 1, <https://doi.org/10.5703/1288284316637> (2018).
25. Darnall, B. To treat pain, study people in all their complexity. *Nature* **557**, 7 (2018).
26. FitzGerald, G. *et al.* The future of humans as model organisms. *Science* **361**(6402), 552–553, <https://doi.org/10.1126/science.aau7779> (2018).
27. Stocker, T. F. *et al.* (Eds) Climate change 2013: The physical science basis. Intergovernmental Panel on Climate Change, Working Group I Contribution to the IPCC Fifth Assessment Report (AR5), Cambridge University Press, UK, New York, <https://doi.org/10.1017/CBO9781107415324> (2013).
28. Zhang, X., Chen, X. & Zhang, X. The impact of exposure to air pollution on cognitive performance. *P.N.A.S.* **115**(37), 9193–9197, <https://doi.org/10.1073/pnas.1809474115> (2018).
29. Wang, K.-Y. & Chau, T. T. An association between air pollution and daily outpatient visits for respiratory disease in a heavy industry area. *PLoS One* **8**(10), e75220, <https://doi.org/10.1371/journal.pone.0075220> (2013).
30. Chau, T. T. & Wang, K.-Y. Association between daily hospital outpatient visits for accidents and daily ambient air temperatures in an industrial city. *PLoS One* **11**(1), e0145003, <https://doi.org/10.1371/journal.pone.0145003> (2016).
31. Burke, M. *et al.* Higher temperatures increase suicide rates in the United States and Mexico. *Nature Climate Change* **8**, 723–729 (2018).
32. Press, W. H., Teukolsky, S. A., Vetterling, W. T. & Flannery, B. P. *Numerical recipes in Fortran*. 2nd ed., Vol. 77, (933p. Cambridge University Press, Cambridge, UK, 1992).
33. Sola, J. & Sevilla, J. Importance of input data normalization for the application of neural networks to complex industrial problems. *IEEE Trans. Nucl. Sci.* **44**(3), 1464–1468, <https://doi.org/10.1109/23.589532> (1997).
34. Starr, C. Social benefit versus technological risk. *Science* **165**(3899), 1232–1238 (1969).
35. Kleinbaum, D. G., Kupper, L. L., Nizam, A. & Muller, K. E. *Applied regression analysis and other multivariable methods*. 4th ed., 906p. Duxbury Press, CA 94002, USA, 2008).
36. Fang, S. H. & Chen, H. W. Air quality and pollution control in Taiwan. *Atmos. Environ.* **30**, 735–741 (1996).
37. US EPA. EPA quality manual for environmental programs. CIO 2105-P-01-0 (formerly 5360 A1), May 5, 2000, United States Environmental Protection Agency, Office of Environmental Information, Quality Staff, Washington, D.C. 20460 (2010).
38. Lary, D. J. *et al.* Estimating the global abundance of ground level presence of particulate matter ($PM_{2.5}$). *Geospatial. Health* **8**(3), S611–S630 (2014).
39. Chen, T.-F., Tsai, C.-Y. & Chang, K.-H. Performance evaluation of atmospheric particulate matter modeling for East Asia. *Atmos. Environ.* **77**, 365–375 (2013).
40. Suwa, T. *et al.* Particulate air pollution induces progression of atherosclerosis. *J. Am. Coll. Cardiol.* **39**(6), 935–942 (2002).
41. Calderón-Garcidueñas, L. *et al.* Pediatric respiratory and systemic effects of chronic air pollution exposure: Nose, lung, heart, and brain pathology. *Toxicol. Pathol.* **35**, 154–162 (2007).
42. Laumbach, R., Meng, Q. & Kippen, H. What can individuals do to reduce personal health risks from air pollution? *J. Thorac. Dis.* **7**(1), 96–107 (2015).
43. Gorham, E. D., Garland, C. F. & Garland, F. C. Acid haze air pollution and breast and colon cancer mortality in 20 Canadian cities. *Can. J. Public Health.* **80**(2), 96–100 (1989).
44. Pope, C. A. III *et al.* Lung cancer, cardiopulmonary mortality, and long-term exposure to fine particulate air pollution. *J.A.M.A.* **287**(9), 1132–1141 (2002).
45. Kaplan, G. *et al.* The inflammatory bowel diseases and ambient air pollution: A novel association. *Am. J. Gastroenterol.* **105**, 2412–2419, <https://doi.org/10.1038/ajg.2010.252> (2010).
46. Ananthakrishnan, A. N., McGinley, E. L., Binion, D. G. & Saeian, K. Ambient air pollution correlates with hospitalizations for inflammatory bowel disease: An ecologic analysis. *Inflamm. Bowel Dis.* **17**, 1138–1145 (2011).
47. Mutlu, E. A. *et al.* Particulate matter air pollution causes oxidant-mediated increase in gut permeability in mice. *Part. Fibre. Toxicol.* **8**, 19 (2011).
48. Kaplan, G. G. *et al.* Non-specific abdominal pain and air pollution: A novel association. *PLoS One* **7**(10), e47669, <https://doi.org/10.1371/journal.pone.0047669> (2012).
49. Salim, S. Y., Kaplan, G. G. & Madsen, K. L. Air pollution effects on the gut microbiota: A link between exposure and inflammatory disease. *Gut Microbes* **5**(2), 215–219 (2013).
50. Kish, L. *et al.* Environmental particulate matter induces murine intestinal inflammatory responses and alters the gut microbiome. *PLoS One* **8**(4), e62220, <https://doi.org/10.1371/journal.pone.0062220> (2013).
51. Salim, S. Y. *et al.* Exposure to ingested airborne pollutant particulate matter increases mucosal exposure to bacteria and induces early onset of inflammation in neonatal IL-10-deficient mice. *Inflamm. Bowel Dis.* **20**, 1129–1138 (2014).
52. Möller, W. *et al.* Mucociliary and long-term particle clearance in the airways of health nonsmoker subjects. *J. Appl. Physiol.* **97**, 2200–2206 (2004).
53. Figura, N. *et al.* Cytotoxin production by *Campylobacter pylori* strains isolated from patients with peptic ulcers and from patients with chronic gastritis only. *J. Clin. Microbiol.* **27**(1), 225–226 (1989).
54. Ley, R. E., Peterson, D. A. & Gordon, J. I. Ecological and evolutionary forces shaping microbial diversity in the human intestine. *Cell* **124**, 837–848, <https://doi.org/10.1016/j.cell.2006.02.017> (2006).
55. Nicholson, J. K. *et al.* Host-gut microbiota metabolic interactions. *Science* **336**(6086), 1262–1267 (2012).
56. Kinross, J. M., Drazi, A. W. & Nicholson, J. K. Gut microbiome-host interactions in health and disease. *Genome. Med.* **3**, 14 (2011).
57. Wang, Z. *et al.* Gut flora metabolism of phosphatidylcholine promotes cardiovascular disease. *Nature* **472**, 57–65 (2011).
58. Guglielmi, G. How gut microbes are joining the fight against cancer. *Nature* **557**, 482–484, <https://doi.org/10.1038/d41586-018-05208-8> (2018).
59. Koshland, D. E. The Molecule of the Year. *Science* **258**, 1861 (1992).
60. Luiking, Y. C., Engelen, P. K. J. & Deutz, N. E. P. Regulation of nitric oxide production in health and disease. *Curr. Opin. Clin. Nutr. Metab. Care* **13**(1), 97–104, <https://doi.org/10.1097/MCO.0b013e328332f99d> (2010).
61. Ignarro, L. J. Nitric oxide as a unique signaling molecule in the vascular system: A historical overview. *J. Physiol. Pharmacol.* **53**(4), 503–514 (2002).
62. Brüne, B. Nitric oxide: NO apoptosis or turning it ON? *Cell Death Differ.* **10**, 864–869 (2003).
63. Power, M. C., Adar, S. D., Yanosky, J. D. & Weuve, J. Exposure to air pollution as a potential contributor to cognitive function, cognitive decline, brain imaging, and dementia: a systematic review of epidemiologic research. *Neurotoxicology* **56**, 235–253, <https://doi.org/10.1016/j.neuro.2016.06.004> (2016).

64. Kojima, T., Tsunashima, H., Shiozawa, T., Takada, H. & Sakai, T. Measurement of train driver's brain activity by functional near-infrared spectroscopy (fNIRS). *Opt. Quantum Electron* **37**, 1319–1338 (2006).
65. Amitai, Y., Zlotogorski, Z., Golan-Katzav, V., Wexler, A. & Gross, D. Neuropsychological impairment from acute low-level exposure to carbon monoxide. *Arch. Neurol.* **55**, 845–848 (1998).
66. Wright, G., Randell, P. & Shephard, R. J. Carbon monoxide and driving skills. *Arch. Environ. Health* **27**(6), 349–354, <https://doi.org/10.1080/00039896.1973.10666400> (1973).
67. Liblau, R. S. Put to sleep by immune cells. *Nature* **562**, 46–48 (2018).
68. Tall, A. R. & Jelic, S. How broken sleep promotes cardiovascular disease. *Nature*, <https://doi.org/10.1038/d41586-019-00393-6> (2019).
69. Lin, C. *et al.* Fever promotes T lymphocytes trafficking via a thermal sensory pathway involving heat shock protein 90 and $\alpha 4$ integrins. *Immunity* **50**, 137–151, <https://doi.org/10.1016/j.immuni.2018.11.013> (2018).
70. Parham, P. *The Immune System*. 2ed. (Garland Science, New York & London, 2005).
71. Meggs, W. J. Neurogenic inflammation and sensitivity to environmental chemicals. *Environ. Health Perspect.* **101**, 234–238 (1993).
72. van Eeden, S. F. *et al.* Cytokines involved in the systemic inflammatory response induced by exposure to particulate matter air pollutants (PM_{10}). *Am. J. Respir. Crit. Care Med.* **164**, 826–830 (2001).
73. Jalava, P. I. *et al.* Associations of urban air particulate composition with inflammatory and cytotoxic responses in RAW 246.7 cell line. *Inhal. Toxicol.* **21**(12), 994–1006, <https://doi.org/10.1080/08958370802695710> (2009).
74. Sawyer, K., Mundandhara, S., Ghio, A. J. & Madden, M. C. The effects of ambient particulate matter on human alveolar macrophage oxidative and inflammatory responses. *J. Toxicol. Environ. Health A* **73**(1), 41–57, <https://doi.org/10.1080/15287390903248901> (2009).
75. Bind, M.-A. *et al.* Air pollution and markers of coagulation, inflammation, and endothelial function: Associations and epigenetic-environment interactions in an elderly cohort. *Epidemiology* **23**(2), 332–340 (2012).
76. Hu, R. *et al.* $PM_{2.5}$ exposure elicits oxidative stress responses and mitochondrial apoptosis pathway activation in HaCaT keratinocytes. *Chin. Med. J.* **130**(18), 2205–2214 (2019).
77. Prada, D. *et al.* Association of air particulate pollution with bone loss over time and bone fracture risk: analysis of data from two independent studies. *Lancet Planet Health* **1**, e337–47 (2017).
78. Möller, P. & Loft, S. Oxidative damage to DNA and lipids as biomarkers of exposure to air pollution. *Environ. Health Perspect.* **118**(8), 1126–1136 (2010).
79. Feldmann, M., Brennan, F. M. & Maini, R. N. Role of cytokines in rheumatoid arthritis. *Annu. Rev. Immunol.* **14**, 397–440 (1996).
80. Calderón-Garcidueñas, L. *et al.* Long-term air pollution exposure is associated with neuroinflammation, an altered innate immune response, disruption of the blood-brain barrier, ultrafine particulate deposition, and accumulation of amyloid β -42 and α -synuclein in children and young adults. *Toxicol. Pathol.* **36**, 289–310 (2008).
81. Libby, M. Interleukin-1 beta as a target for atherosclerosis therapy. *J. Am. Coll. Cardiol.* **70**, 2278–2279 (2017).
82. Fernandez-Cuadros, M. E., Perez-Moro, O. S. & Mirón-Canelo, J. A. Could ozone be used as a feasible future treatment in osteoarthritis of the knee? *Divers. Equal. Health Care* **13**(3), 232–239 (2016).
83. Davies, C. M., Guilak, F., Weinberg, J. B. & Fermo, B. Reactive nitrogen and oxygen species in interleukin-1-mediated DNA damage associated with osteoarthritis. *Osteoarthr. Cartil.* **16**, 624–630 (2008).
84. Towns, M. & Bathon, J. Inhibition of interleukin-1 as a strategy for the treatment of rheumatoid arthritis. The Johns Hopkins arthritis center, <http://www.hopkinsarthritis.org/arthritis-info/rheumatoid-arthritis/ill.html>, last updated: 1 November 2011 (2011).
85. Krall, E. A. & Dawson-Hughes, B. Smoking and bone loss among postmenopausal women. *J. Bone. Miner. Res.* **6**(4), <https://doi.org/10.1002/jbmr.5650060404> (1991).
86. Law, M. R. & Hackshaw, A. K. A meta-analysis of cigarette smoking, bone mineral density and risk of hip fracture: recognition of a major effect. *B.M.J.* **315**, 841–846 (1997).
87. Weschler, C. J. *et al.* Transdermal uptake of diethyl phthalate and di(n-butyl) phthalate directly from air: Experimental verification. *Environ. Health Perspect.* **123**, 928–934, <https://doi.org/10.1289/ehp.1409151> (2015).
88. Hüls, A. *et al.* Traffic-related air pollution contributes to development of facial lentigenes: Further epidemiological evidence from Caucasians and Asians. *J. Invest. Dermatol.* **136**, 1053–1056, <https://doi.org/10.1061/j.jid.2015.12.045> (2016).
89. Svoboda, E. When skin's defence against pollution fails. *Nature* **563**, S89–S90, <https://doi.org/10.1038/d41586-018-07431-9> (2018).
90. Chen, C. *et al.* Behavior, attitudes and knowledge about antibiotic usage among residents of Changhua, Taiwan. *J. Microbiol. Immunol. Infect.* **38**, 53–59 (2005).
91. Darcin, M. Association between air quality and quality of life. *Environ. Sci. Pollut. Res. Int.* **21**(3), 1954–1959, <https://doi.org/10.1007/s11356-013-2101-3> (2013).
92. Cameron, D. & Jones, I. G. John Snow, the Broad Street pump and modern epidemiology. *Int. J. Epidemiol.* **12**, 393–396 (1983).
93. Khoury, M. & Ioannidis, J. P. A. Big data meets public health. *Science* **346**(6213), 1054–1055, <https://doi.org/10.1126/science.aaa2709> (2014).
94. Reis, S. *et al.* Integrating modelling and smart sensors for environmental and human health. *Env. Model. Soft.* **74**, 238–246, <https://doi.org/10.1016/envsoft.2015.06.003> (2015).
95. Lary, D. J., Lary, T. & Sattler, B. Using machine learning to estimate global $PM_{2.5}$ for environmental health studies. *Environ. Health Insights* **9**(S1), 41–52, <https://doi.org/10.4137/EHI.S15664> (2015).

Acknowledgements

We thank Taiwan EPA for the ambient air monitoring data that are openly accessible for research at <http://www.epa.gov.tw>; and Tien-Mu Hsiao and Chuan-Fa Hsieh of the Taiwan Landseed Hospital for very inspirational discussions during the initial stage of this work. We are very grateful to the Ministry of Science and Technology for the funding support of research in big data analysis and the development of smart sensors. This work was supported by 106-2111-M-008-011 and 107-2111-M-008-027.

Author contributions

T.T.C. and K.Y.W. made equal contributions to the conception and design of the work, and interpretation of data. K.Y.W. was responsible for creation of the model, the detailed methodology design, and writing the software to run, calibrate and validate the model. K.Y.W. performed the experiments. T.T.C. collected and prepared the medical data. K.Y.W. collected and prepared the air pollution and meteorological data. T.T.C. and K.Y.W. jointly performed the design and the review of this work. K.Y.W. wrote the manuscript, and prepared all the figures and tables used in the manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information is available for this paper at <https://doi.org/10.1038/s41598-020-58721-0>.

Correspondence and requests for materials should be addressed to K.-Y.W.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2020