

Pulmonary function changes in asthmatics associated with low-level SO₂ and NO₂ air pollution, weather, and medicine intake

An 8-month prospective study analyzed by neural networks

Moseholm L, Taudorf E, Frøsig A. Pulmonary function changes in asthmatics associated with low-level SO₂ and NO₂ air pollution, weather, and medicine intake.

Allergy 1993; 48: 334–344. © Munksgaard 1993.

This paper presents the response in subjects with asthma to gaseous air-pollution levels, weather, and medicine intake as identified by principal-component analysis and neural network techniques. Pulmonary function measured by respiratory peak-flow rate in nonallergic asthmatics was associated with ambient, low-level, air-pollution concentrations of sulfur dioxide and nitrogen dioxide, temperature, relative humidity, and medicine intake. Results from 27 nonallergic asthmatics aged 18–60 years with well-characterized bronchial asthma and regular medical treatment were analyzed from two cities. During an 8-month period, each subject kept a diary table, which included symptoms, lung function (evening peak flow), medicine intake, and tobacco smoking.

High intake of medicine and high ambient temperatures corresponded to decreased peak flow. The changes in temperature did not occur in situations with low medicine intake. During frost periods, peak-flow values decreased independently of medicine intake and levels of SO₂. During other times, increased levels of SO₂ and NO₂ increased temperature, and increased intake of medicine, and low relative humidity corresponded to decreased peak flow. Increased levels of SO₂ and NO₂ corresponded synergistically to decreased peak flow at levels above 40 µg/m³.

L. Moseholm^{1,4}, E. Taudorf^{2,4}, A. Frøsig³

¹ Civil Engineering, University of Washington, FX-10, Seattle, WA 98195, USA; ² Department of Respiratory Medicine, Skive Hospital, DK-7800 Skive, Denmark; ³ Environmental Health, University of Washington, Seattle, WA 98195, USA; ⁴ Danish Air Pollution Group*

Key words: air pollution; asthma; climate; neural networks; NO₂; SO₂.

Dr L. Moseholm
Civil Engineering
University of Washington
FX-10, Seattle, WA 98195
USA

Accepted for publication 29 September 1992

During the last decade, asthma hospitalization among children and asthma mortality rates have increased (13, 35). From 1980 to 1987, the US national hospital discharge rates for asthma increased nearly threefold and the death rate from asthma increased by 31% (6). Lødrup et al. (21) found that the admission rate of children from birth to 14 years old with bronchial asthma in Oslo increased by a factor of 2.3 over the period 1980–9. This recent increase in prevalence of and mortality from asthma may be due to increased concentrations of airborne pollutants during the last decades (31).

There is epidemiologic evidence that air pollutants are associated with impaired lung function and respiratory symptoms, although not all studies systematically characterized subjects with respect to asthma. Ware et al. (33) found increased respiratory symptoms in children exposed chronically to particulate matter and sulfur dioxide (SO₂). Dockery et al. (7) reported that rates of chronic cough, bronchitis, and other chest illnesses in children were positively associated with particulate air pollution.

Experimentally, increased exposure to SO₂ and nitrogen dioxide (NO₂) have been shown to increase bronchial reactivity and decrease lung function in asthmatics (36). Other studies have found no readily apparent risk to human subjects from exposure to relatively low levels of air pollutants (10, 16).

There is great variability of sensitivity within the asthmatic population from subject to subject and within each subject over time, and symptoms are highly influenced by environmental variables (36).

* The Danish Air Pollution Group: Lars Moseholm PhD; Alix Rasmussen MSc, Lars S. Sørensen MSc, Bente E. Jørgensen MSc, Hans Buch MSc (Danish Meteorological Institute); Knud Ramkjær PhD (Council of Greater Copenhagen); Mikael Hansen MD (Roskilde County Hospital); Bruno N. Pedersen MD, Karsten M. Christensen (Copenhagen County Hospital, Gentofte); Ebbe Taudorf MD, Torben Evald MD (Bispebjerg Hospital, Copenhagen); Otto Davidsen MD, Preben Larsen MD, Helle Dahl Madsen MD (Odense Hospital); Mogens Dueholm MD, Carl Nielsen MD (Aalborg Hospital).

Empirically, asthmatics feel the influence of meteorologic factors such as fog and wind, and fluctuations in barometric pressure, thermal inversions, or cool air may trigger asthmatic events. The influence of weather conditions on asthmatic outbreaks is still a controversial subject (4, 24, 25).

In 1986, Copenhagen City Council proposed a new air-pollution warning system including a service for persons with respiratory diseases. Consequently, a prospective open study was initiated at five Danish hospitals in four cities to identify possible relationships between the relatively low Danish air-pollution levels and meteorologic parameters and illness in a population of nonallergic asthmatics. A total of 73 subjects participated in this study over an 8-month period (October–May 1987–8). Because of the lack of understanding of the possible dose-response relationships, neural networks were applied to recognize patterns in the asthmatic responses. Neural networks have been shown to provide efficient solutions to problems that are difficult to solve by traditional statistical techniques (29).

Methods

Subjects

Asthmatics from two of the four cities, Odense (12 subjects) and Aalborg (15 subjects), were used in this analysis. The two cities were very similar in size (100–150 000 inhabitants) with mixed industry. Domestic heating and traffic emissions caused most of the air pollution.

Outpatients aged 18–60 who had been diagnosed as having asthma the previous year at the chest clinics involved were invited to participate. Those who answered positively were screened for inclusion. All subjects had well-characterized, nonallergic, perennial asthma and were familiar to the clinics. The diagnosis of asthma was verified by the coexistence of symptoms (dyspnea, cough, or wheeze) and reversible airway obstruction for more than 3 months. All subjects had at least 20% variability in lung function measured by either FEV₁ or peak flow over short periods, and consumed a monthly average of at least one β_2 -agonist nebulizer (200 doses). The mean measured FEV₁ of predicted value was 60% for smokers and 85% for nonsmokers. Forty percent of the 27 subjects smoked, and the male/female ratio was 0.4. No significantly different characteristics were detected between the two city asthma panels.

No subject suffered from concomitant severe disease. Pregnant women and persons changing habits during the period were excluded. Smoking was allowed if held constant. Treatment with theophylline and inhaled steroids was also allowed. All subjects had negative skin prick tests to a panel of 10 com-

mon Scandinavian allergens; this avoided interactions with effects from indoor or outdoor airborne allergens (23).

At inclusion, the lung function (FEV₁ and FVC) of all subjects was thoroughly tested by international standard procedures. The daily measurements of peak flow were obtained with the Wright mini-peak-flow meter. The highest of three technically correct measurements was recorded in the diary.

The daily use of medication was recorded by the drug and dosage registration. In the analysis, this information was transformed into a standardized score by weighing the medicine doses; one score unit was one inhalation of spray, one 5-mg prednisone tablet was 10 points, etc. In each chest clinic the subjects were instructed in medication by oral and written information.

The fieldwork was performed by members of the Air Pollution Group of the Danish Asthma-Allergy Association.

Environmental variables

The main sources of air pollution in the two cities were heating and vehicle emissions. Data on the 24-h mean concentrations of SO₂ and NO₂ were available from the National Environmental Research Institute's measuring station network. Data were available from only one central site in each city. Values of the 12-h day-mean (0600–1800) relative humidity, and temperature were obtained from local airports near each city.

Statistical analysis

Multivariate analysis. Multivariate statistical analyses of the data obtained for each subject by the partial least squares method (12) revealed that 0–50% of the subject-specific asthma response (symptoms, peak flow, and medicine consumption) could be related to changes in environmental conditions, and that the relationships between the asthmatic response and variations in air pollutants and weather were extremely subject-dependent (22). Data plots further revealed nonlinear relationships. The observations for each variable were autocorrelated with a correlation coefficient of about 0.6–0.8, and the dependent and independent variables were correlated within each set (multicollinearity). Such characteristics significantly bias the estimation in least-squares regression models. When such models were applied, including ridge regression, indecisive results were obtained for the two independent data sets, one for each city.

The principal-component analysis is the classic multivariate analysis to find combinations of a set of correlated variables which produce components that are uncorrelated. In order to compare the results

from this multivariate analysis and the neural network analysis, we used the same input in both analyses. A random selection of 22% of the full data set (the test set; see below) over subjects and cities was used for the principal component and to test the neural network models. The following central variables were included:

- 1) the day's medicine intake normalized to personal average [Dmed(t)] in score values
- 2) the relative change in medicine intake from the previous day in score values [Dmed(t - 1)]
- 3) the day's and the preceding day's peak flow normalized to personal average in l/min [DPF(t), DPF(t - 1)]
- 4) day-mean temperatures in °C [Temp(t)]
- 5) day-mean relative humidity in % [Hum(t)]
- 6) the day's 24-h mean values of SO₂ and NO₂ in µg/m³ [SO₂(t) and NO₂(t)].

The definition of these variables was determined by the requirement of the input to the neural network.

The neural net

The network used was a "back-propagation" network. It learns by example and repetition. The network consisted of units (neurons): a layer of input units, layers of hidden units that have no direct input from or output to outside the network, and a layer of output units. The input layer sends signals to the next layer of neurons. Each hidden neuron receives signals from all the neurons in the preceding layer. A neuron sends a signal when the sum of the input signals has reached a certain level. Some input signals can be inhibitory; others can be excitatory. The function which maps the input of a unit to its output has a sigmoid logistic form. Connected to every unit connection in the neural network is a variable weight that enforces or diminishes the signal throughout the network. The strengths of the weights are determined by a learning algorithm, enabling the network to learn and reproduce particular output from a given input pattern. Once the system is trained and a new input pattern is submitted, the system reacts in a highly complex, but well-defined manner, according to the learned weights. The neural network used was the commercial back-propagation network software, Brainmaker (California Scientific Software, USA). The training algorithm is

$$(1) w_{ij}(t + 1) = w_{ij}(t) + n e_j x_i + m[w_{ij}(t) - w_{ij}(t - 1)]$$

where w_{ij} is the weight associated with the transfer from neuron i to neuron j , x_i is the output from neuron i , e_j is the error term for neuron j , n is the learning rate, and m is a smoothing factor.

In this study, characteristics for a given person with asthma were given to the input layer with information on levels of environmental variables. Eventually, a pattern was formed by the internal neurons that matched a previously learned pattern representing the response of an asthmatic person to environmental conditions. The network then used its output neurons to signal a given response, i.e. changes in peak flow. Under the same input conditions, the trained neural network always sends the same output.

After several attempts, the neural network topology was set up with 52 neurons in an input layer, two intermediate layers with 40 and 25 neurons respectively, and one single output neuron (peak flow). The input neurons were stimulated by three consecutive days of data in order to identify time-lag effects. All data sets were randomized before being presented to the network, and, randomly, 22% of the data sets were held back for testing purposes. A new network training procedure was developed during this study for the learning process of the network to diverge, i.e. by gradually decreasing the learning rate (9).

The network was trained for each city population and for the two populations combined, giving a total of three independently trained networks.

Input and output variables

Each data set ("fact") included the following information for a specific day:

- 1) the patient's sex, age, city
- 2) average smoking habits: cigarettes smoked on a given day and the 2 previous days
- 3) average medicine intake and its variance and standard deviation, and intake on a given day normalized to personal average
- 4) the change in medicine intake on a given day relative to the intake of the 2 previous days
- 5) average peak-flow value and its variance and standard deviation, the previous day's peak flow normalized to personal average, and the change in peak flow between the previous day and the day before it
- 6) number of days the patient is unable to go to work because of symptoms, and days at hospital because of asthma
- 7) days with frost, 3 days' day-mean temperatures, and 3 days' day-mean relative humidity
- 8) SO₂ and NO₂ 24-h mean values on a given day and the 2 previous days.

The observed peak-flow values on a given day, normalized to the personal average, were presented to the network during training. After training, the network forecast the value as dependent on the input values.

The total data set was 2295 (facts) for training.

The number of facts was 918 for the city of Odense and 955 for the city of Aalborg when all days with missing values were excluded. A total of 422 facts comprised the independent test set. One run was all data sets presented at one time to the network. The largest deviation value in percentage between observed and estimated peak-flow values is called the tolerance value. The network was trained by presenting the facts over and over to it. This training process was repeated until the network got all the facts correct within a given tolerance value.

Results

The 24-h mean SO_2 concentration range was 15–18 $\mu\text{g}/\text{m}^3$ in the two cities, with a common 95th percentile value of 41 $\mu\text{g}/\text{m}^3$, and peak values of 81 $\mu\text{g}/\text{m}^3$ and 57 $\mu\text{g}/\text{m}^3$, respectively. The 24-h mean NO_2 concentrations over the 8-month period were 39 $\mu\text{g}/\text{m}^3$ and 48 $\mu\text{g}/\text{m}^3$, respectively, with 95th percentile values 73–80 $\mu\text{g}/\text{m}^3$ and peak 24-h values 91–95 $\mu\text{g}/\text{m}^3$. These concentration levels are well below the WHO recommended air-quality guidelines (36).

The weather was relatively variable with a temperature between a few degrees below freezing and 20°C (mean 6°C). The relative humidity was 48–99% (mean 79%). Fog occurred only on a very few days throughout the 8-month period.

The average standard deviation in peak-flow values between subjects was 38% of the mean; and between days within subjects, 14% of the mean, ranging from 4.2 to 29.9%.

Multivariate analysis

The Pearson correlation matrix for the central variables, peak flow, medicine consumption, temperature, relative humidity, SO_2 level, and NO_2 level, is

shown in Table 1. The matrix indicated a positive linear relationship between a given day's peak flow, $\text{DPF}(t)$, and the previous day's peak flow, $\text{DPF}(t-1)$, but no significant linear relationships existed between the peak flow and environmental variables except for temperature.

A given day's peak flow [$\text{DPF}(t)$] was specifically negatively correlated to the intake of medicine [$\text{Dmed}(t)$], which was negatively correlated to the previous day's medicine consumption [$\text{Dmed}(t-1)$] and day-mean temperature [$\text{Temp}(t)$], and positively correlated to the NO_2 level. The environmental variables were correlated, as one would expect. The autocorrelation influenced the significance given for the correlations in Table 1. The correlation coefficients were biased; the variance within each variable was larger than if the autocorrelation was zero, and, consequently, the correlation coefficients may have been underestimated. However, because of the high number of observations, Table 1 may still give a correct picture of the Pearson correlation pattern.

Table 2 shows the result of the principal-component analysis on the test set: the proportion of the total variance accounted for by each component and the component's weights. The values in each column of the matrix in Table 2 give the weights of the linear combinations forming each principal component. The coefficients of each component have been standardized so that the sum of their squares is equal to 1.

The lack of correlation between principal components is a useful property because it means that the components measure different "independent dimensions" in the data. Table 2 orders the components with regard to the amount of the total variation that each component explains. When one does a principal-component analysis, there is always the hope that the variances of most of the components will be so low as to be negligible. First, it is seen from

Table 1. Pearson sample correlation matrix

	Dmed (t-1)	DPF (t)	DPF (t-1)	Temp (t)	Hum (t)	SO_2 (t)	NO_2 (t)
Dmed(t)	-0.295 #	-0.261 #	-0.300 #	-0.151 \$	(0.027)	(0.085)	0.139 \$
Dmed(t-1)	-	0.263 #	(-0.071)	(0.044)	-0.119	(-0.017)	(0.022)
DPF(t)		-	0.576 #	(-0.042)	(0.016)	(0.020)	(0.035)
DPF(t-1)			-	(-0.092)	(0.042)	(0.067)	(0.017)
Temp(t)				-	-0.290 #	-0.259 #	-0.498 #
Hum(t)						0.128 \$	0.165 #
SO_2 (t)						-	0.510 #
NO_2 (t)							-

Number of cases=422.

(XX) nonsignificant value.

XX P value < 0.05.

\$ P value < 0.01.

P value < 0.001.

Table 2. Principal-component analysis: variance explained and component weights

Variables	Component no.							
	1	2	3	4	5	6	7	8
	25.4%	23.2%	14.1%	11.4%	9.4%	7.8%	5.0%	3.8%
Dmed(t)	0.23	<i>0.45</i>	0.14	-0.28	<i>-0.56</i>	<i>0.52</i>	-0.07	0.23
Dmed(t-1)	-0.13	-0.28	-0.74	0.31	-0.24	0.21	-0.17	0.37
DPF(t)	-0.02	<i>-0.61</i>	0.08	-0.11	-0.33	0.37	0.16	<i>-0.57</i>
DPF(t-1)	0.03	<i>-0.57</i>	<i>0.43</i>	-0.27	0.00	-0.07	-0.12	<i>0.64</i>
Temp(t)	<i>-0.53</i>	0.08	-0.03	-0.19	<i>0.42</i>	<i>0.52</i>	<i>0.44</i>	0.17
Hum(t)	0.32	-0.04	0.35	<i>0.76</i>	-0.18	0.38	0.08	0.09
SO ₂ (t)	<i>0.47</i>	-0.08	-0.23	-0.33	<i>0.54</i>	0.33	<i>-0.43</i>	-0.12
NO ₂ (t)	<i>0.56</i>	-0.07	-0.27	-0.14	0.00	-0.16	<i>0.73</i>	0.15

Values greater than 0.40 are italicized.

Table 2 that the variation in the data set cannot be adequately described by a few components. Second, it is seen that components 1, 4, and 7 mainly identified variation within the environmental variables, and components 2, 3, and 8 mainly identified variation within the subjects. The remaining two components, which accounted for only 17% of the total variation, basically related medicine consumption and peak flow to the environmental variables in combinations that are not easily understood.

Fig. 1 shows the biplot for the first two principal components, which together explain 48.6% of the total variance. The lines representing the positive directions of the axes of the original variables intersect at (0.0), which is the mark for the means of all variables. The length of each vector is proportional to its contribution to the principal component. The correlation between any two variables is represented approximately by the means of the cosines of the angles between the corresponding axes; a right angle indicates uncorrelated variables (11).

The multivariate structure is dominated by the correlations between the environmental variables, and between a given day's and the previous day's peak flow and the medicine intake. The components indicate a rather complicated correlation structure between air pollutants and peak flow. Peak flow basically correlated with medicine consumption, and medicine consumption correlated with environmental variables.

Neural net analysis

The networks could be trained to tolerance values 22% (2141 runs) for the city of Odense, 18% (909

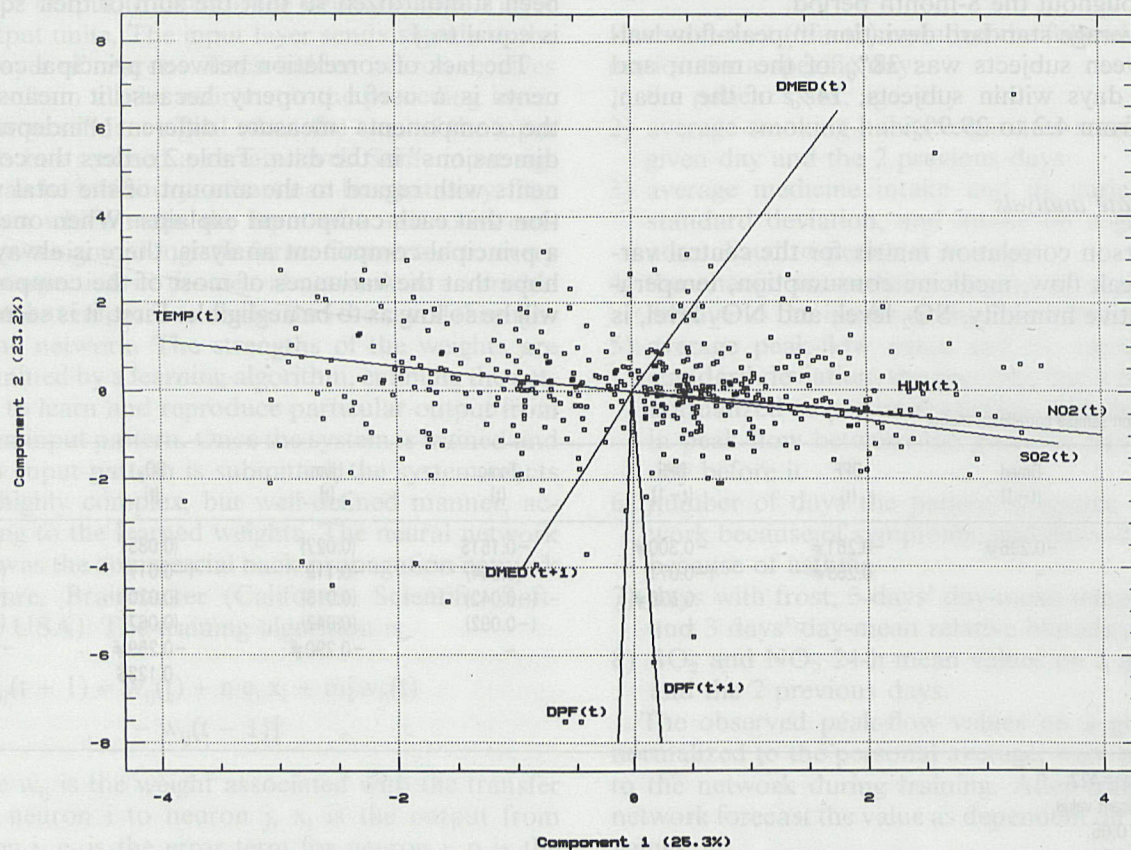


Fig. 1. Biplot for first two principal components, data test set.

Table 3. Sensitivity analysis of neural network's ability to forecast daily peak-flow changes in nonallergic asthmatics. Observed values are the independent test set for two cities

Predicted/observed number of cases	Increased peak flow	Decreased peak flow
Increased peak flow	78	32
Decreased peak flow	33	87
Predictive value of peak-flow increase	$78/(78+32)=0.71$	
Predictive value of peak-flow decrease	$87/(33+87)=0.73$	

runs) for the city of Aalborg, and 26% (714 runs) for the two data sets combined.

Testing the trained networks on the withheld 22% of the data (test sets) gave the following slopes for the linear orthogonal regression lines between observed peak-flow values and predicted values by the network: 0.77 (0.59–0.98) for Odense, and 0.81 (0.63–1.0) for Aalborg. The values are the mean estimate, and the upper and lower 95% significance values. The networks had never seen these results

before, and an ideal model has a correlation of 1 between the observed and predicted values. When the trained network for Aalborg was tested against the withheld data for Odense, and visa versa, the number of good facts, i.e. facts within a given tolerance value, decreased by less than 10%. Both models were therefore fairly universal. The number of good facts for the two data sets combined was 83.2% for the same tolerance value as obtained during the training. For unknown reasons, only two subjects were responsible for 29% of the bad facts. The remaining bad facts were randomly distributed among subjects and days.

The significance of the neural network models can be tested by calculating the Pearson correlation for the linear relationship between observed and predicted values in the test set.

The correlations were 0.528 for the "Odense" test set, 0.517 for the "Aalborg" test set, and 0.524 for the two test sets combined. For all three correlation values the *P* value was highly significant (less than 10^{-5}). The trained network models therefore are highly significant.

An important way to view the network's ability to

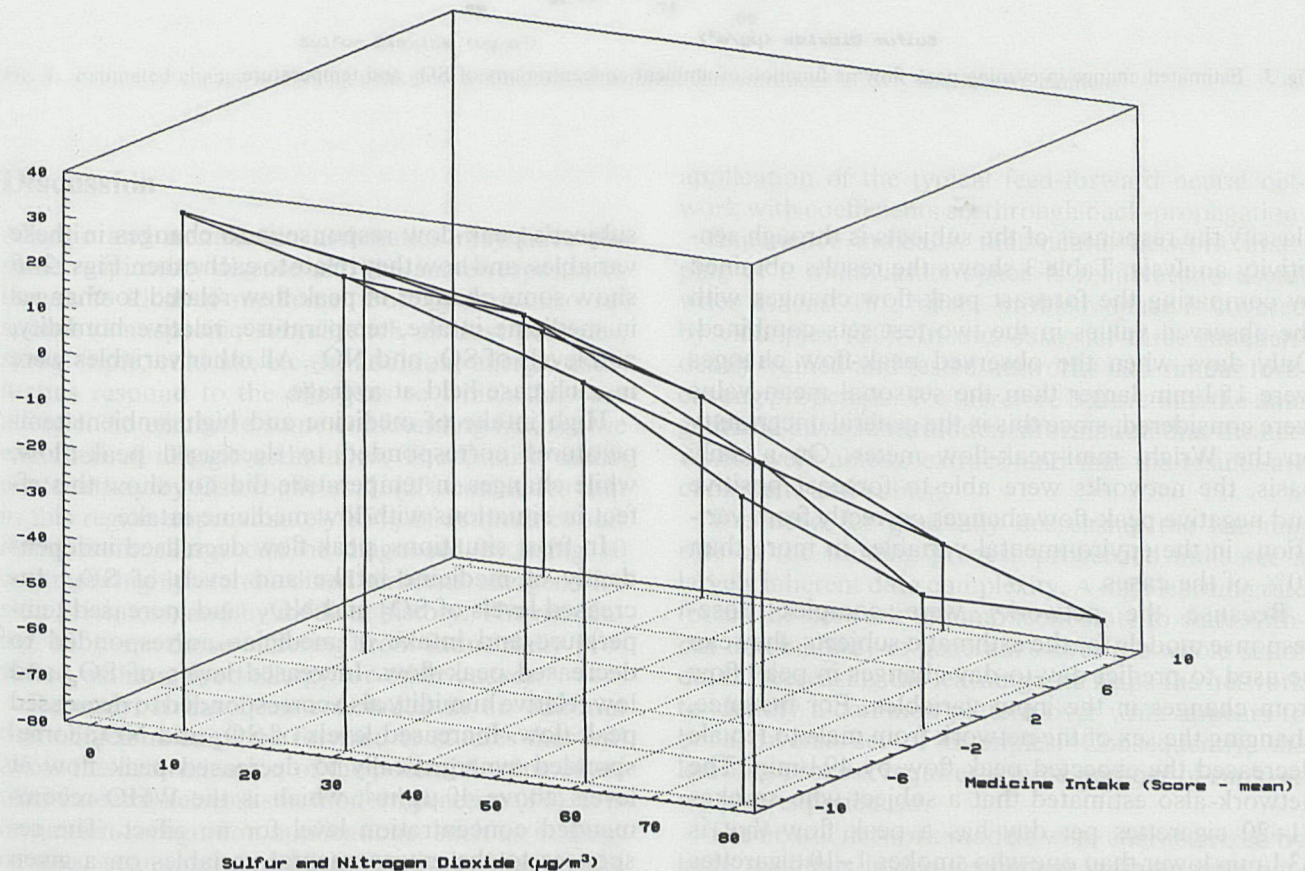


Fig. 2. Estimated change in evening peak flow as function of medicine consumption and ambient concentrations of SO_2 and NO_2 at equal levels.

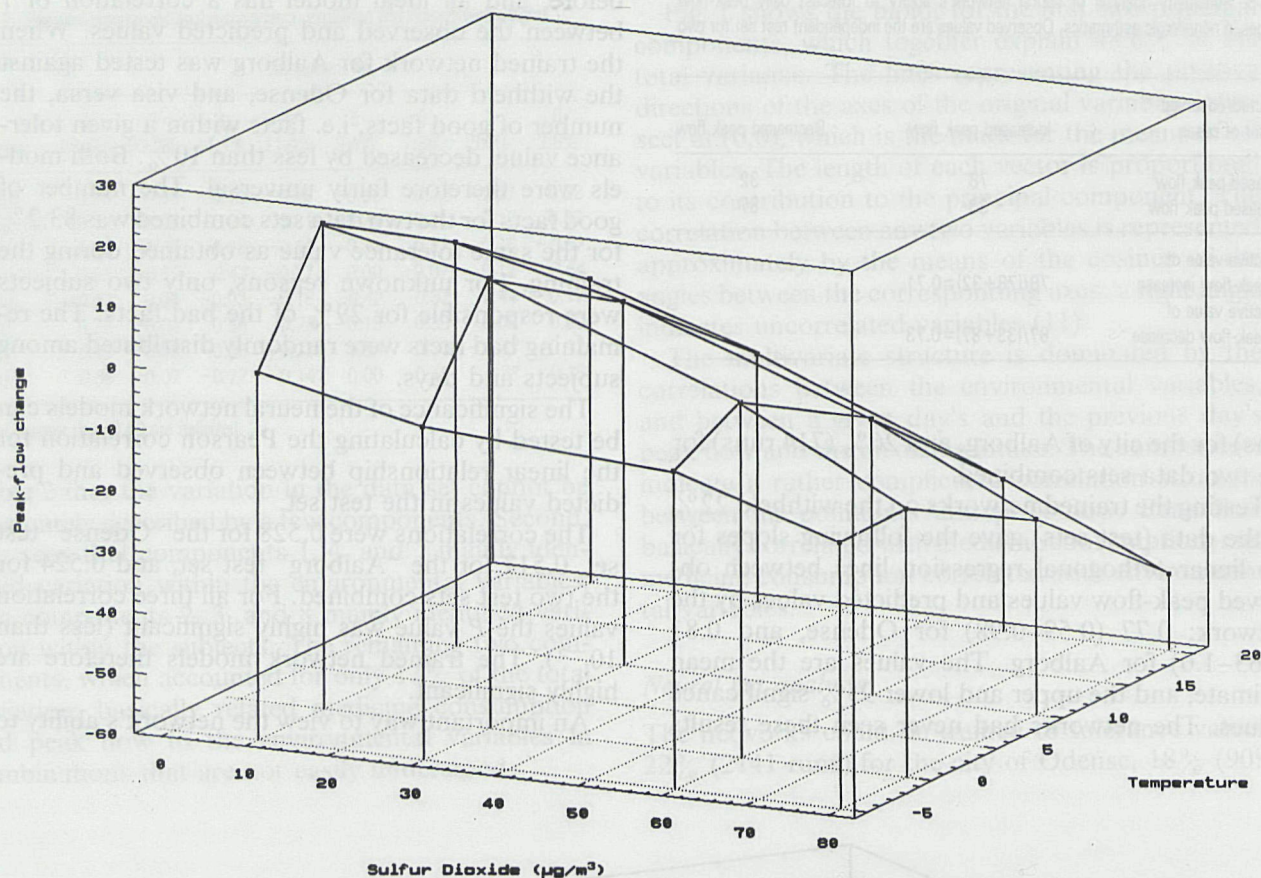


Fig. 3. Estimated change in evening peak flow as function of ambient concentrations of SO_2 and temperature.

classify the responses of the subjects is through sensitivity analysis. Table 3 shows the results obtained by comparing the forecast peak-flow changes with the observed values in the two test sets combined. Only days when the observed peak-flow changes were 15 l/min larger than the seasonal mean value were considered, since this is the general uncertainty on the Wright mini-peak-flow meter. On a daily basis, the networks were able to forecast positive and negative peak-flow changes correctly from variations in the environmental variables in more than 70% of the cases.

Because the networks were complex dose-response models for the asthmatic subjects, they can be used to predict day-to-day changes in peak flow from changes in the input variables. For instance, changing the sex of the network from male to female decreased the expected peak flow by 49 l/min. The network also estimated that a subject who smokes 11–20 cigarettes per day has a peak flow that is 63 l/min lower than one who smokes 1–10 cigarettes per day.

Changes of input variables show how sensitive the

subject's peak-flow response is to changes in these variables, and how they relate to each other. Figs. 2–5 show some changes in peak flow related to changes in medicine intake, temperature, relative humidity, and levels of SO_2 and NO_2 . All other variables were in each case held at average.

High intake of medicine and high ambient temperatures corresponded to decreased peak flow, while changes in temperature did not show this effect in situations with low medicine intake.

In frost situations, peak flow decreased independently of medicine intake and levels of SO_2 . Increased levels of SO_2 and NO_2 , and increased temperature and intake of medicine corresponded to decreased peak flow. Increased levels of SO_2 and low relative humidity also corresponded to decreased peak flow. Increased levels of SO_2 and NO_2 corresponded synergistically to decreased peak flow at levels above $40 \mu\text{g}/\text{m}^3$, which is the WHO recommended concentration level for no effect. The responses to the environmental variables on a given day also depended on the health status (peak flow and medicine consumption) of the 2 previous days.

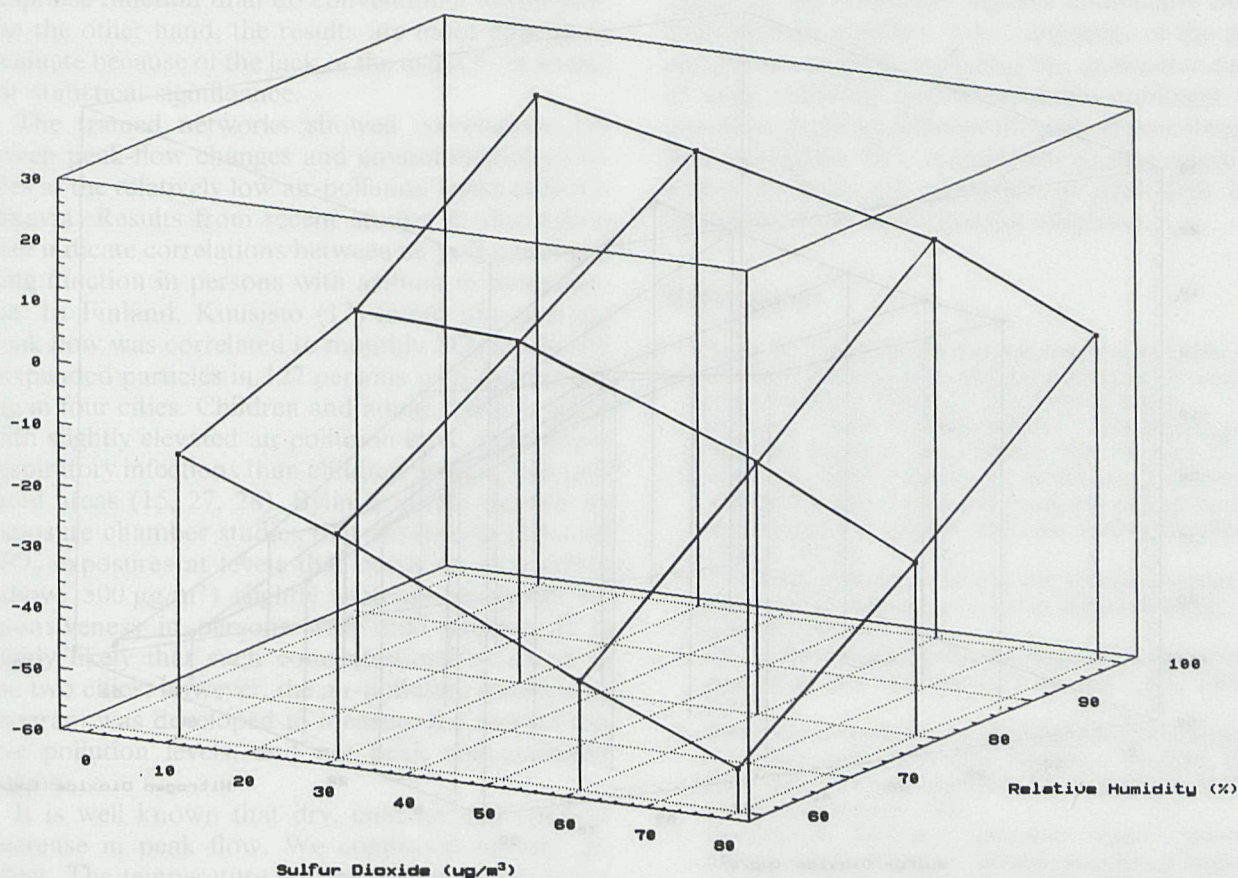


Fig. 4. Estimated change in evening peak flow as function of ambient concentrations of SO_2 and relative humidity.

Discussion

Allergic and nonallergic asthmatics may have the same respiratory disease and the same response to the same kind of medicine. Both groups also respond to unspecific stimuli such as dust particles, smog, fume, cold air, etc. In addition, allergic asthmatics respond to the allergens to which they are allergic. To eliminate some confounding factors, we excluded all allergic asthmatics. The Danish multicenter study by Østerballe et al. (23) concludes that in this region approximately 85% of asthmatics suffering from allergy can be diagnosed by using 10 common highly standardized airborne allergens for skin prick test, and by medical history. The remaining 15% are those suffering from industrially provoked asthma, food allergy, or other causes.

Recently, it has been concluded in an editorial that clinicians can look forward to studies that will allow the determination of the potential value of artificial neural networks in modern practice (14). It is in pattern-recognition applications, such as in diagnostics, scanning, and image analysis, that neural networks will make a significant contribution to clinical practice (1, 2, 8, 34). Our study represents an

application of the typical feed-forward neural network with coefficients set through back-propagation.

Experience indicates that neural network mapping of a multivariate space is reliable and useful when at least 10% of the problem space is covered by examples (20). In our case, all three independently trained and tested networks had similar forecasting properties. We therefore believe that the data gathered have some hidden information that the networks were able to extract, and that the results are consistent and robust.

The need for gradually decreasing the learning rate as the training process proceeded indicates a largely inherent data complexity. A high learning rate forces the network to jump from state to state without much ability to generalize over the whole set of data. Decreasing the learning rate helps the network to identify a common pattern over what appears to be a diverse set of examples. Consequently, the learning process requires a large number of runs, as we also experienced.

The neural network models were characterized by large residuals in the dependent variable with correlation coefficients of about 0.5–0.6 between observed and estimated peak-flow values in the test set.

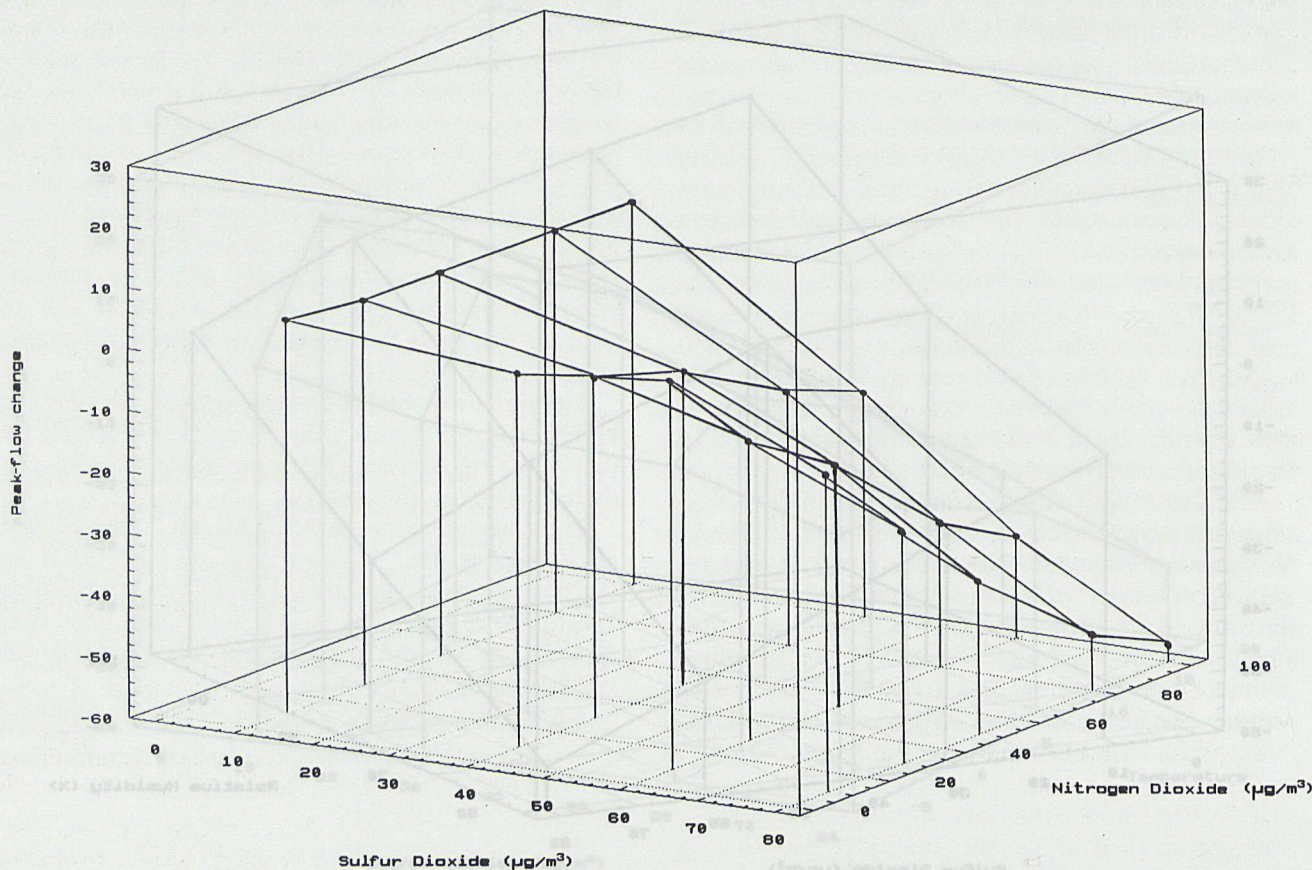


Fig. 5. Estimated change in evening peak flow as function of ambient concentrations of SO_2 and NO_2 .

However, the correlations are highly significant because of the high number of observations.

The large residuals may arise from a rather poor exposure assessment for each subject. The fixed-site monitoring of the pollutants does not allow for any change in spatial and, to some degree, temporal variability of the pollutant concentration field over the urban area within which people move. The representativeness of fixed monitors depends on the pollutant. Concentrations of SO_2 may be fairly representative, while concentrations of NO_2 may be highly variable, mainly as a result of traffic density. Concentrations of particles were not included in this initial analysis.

The testing of the trained nets on the independent test set (usually called cross-validation), i.e. to use only a portion of the database in training the network, and to use the rest of the database in assessing the capacity of the network to generalize, is a standard tool within statistical pattern recognition, and it is a hard test to pass (32). In our case, we also cross-validated the network developed for one city on the data test set for the other city, and vice versa. The multivariate models applied to the test set failed to have correlation coefficients above 0.3. Inclusion

of time-lag effects in the multivariate analyses did not improve this correlation. The fitting of (biased) least-square regression models to our data gave correlation coefficients of 0.4–0.6. Other researchers have found that correlation coefficients of about 0.4–0.6 between respiratory responses in asthmatics and environmental variables can be established by time series analyses (19).

There are alternative modeling strategies for longitudinal data on several subjects that consider random effects across individuals and autocorrelation in the within-subject errors. These models include, for example, linear models (5), binomial distributed response models (30), and models that compare the average change of response for two or more treatment groups (18). These models assume specific designs or dose-response relationships, and require specialized software for their implementation.

Our results showed that the network's efficiency at nonlinear fitting, and its learning ability, i.e. ability to ignore contradictory results and outliers, allowed it to generalize the generic pattern to the same extent as, or to a greater extent than, the classic methods. However, the network makes it possible to explore a far greater range of potential solutions to the dose-

response function than do conventional techniques. On the other hand, the results are more difficult to evaluate because of the lack of the usual F- or t-tests for statistical significance.

The trained networks showed correlations between peak-flow changes and environmental variables at the relatively low air-pollution levels in Scandinavia. Results from recent studies in this region also indicate correlations between air pollutants and lung function in persons with asthma in Scandinavia. In Finland, Kuusisto (17) found that evening peak flow was correlated to monthly SO_2 levels and suspended particles in 127 persons with asthma living in four cities. Children and adults living in areas with slightly elevated air pollution have more upper respiratory infections than children living in less polluted areas (15, 27, 28). Bylin et al. (3) showed by exposure chamber studies that short-term (30-min) NO_2 exposures at levels that occur on city streets (above $500 \mu\text{g}/\text{m}^3$) slightly increase bronchial responsiveness in persons with mild asthma. It is highly likely that such concentrations occurred in the two cities; however, the air-pollution monitoring program was developed to measure the representative pollution levels, and not peak concentration values.

It is well known that dry, cold air can cause a decrease in peak flow. We confirmed the dry air effect. The temperature and peak-flow relation was more complex. In situations (days) with temperature below 0°C , peak flow decreased (as expected) independently of medicine intake and levels of SO_2 . In situations with low medicine intake, increasing temperature (0 – 20°C) did not affect peak flow, which stayed high. In situations with high medicine consumption, increasing temperature (0 – 20°C) caused marked decreases in peak flow. This relationship between temperature and peak flow supports earlier findings by Pope et al. (26) using regression techniques.

Elevated air pollution caused deterioration of asthma (decreased peak flow and increased medicine intake), but it is surprising that asthmatics might be affected by air-pollution levels as low as those shown in Figs. 2–5. The changes in peak flow shown by the neural network models are, however, relatively small, about 10%. It must also be taken into consideration that the networks were trained on all subjects regardless of their sensitivity to air pollution and their management of their disease (medication and lifestyle changes).

How do levels of air pollutants and weather affect asthmatics in the streets and the home? Are there any synergistic, additive, or other effects in these situations? We found that increased levels of SO_2 and NO_2 correspond synergistically to decreased peak flow at levels above $40 \mu\text{g}/\text{m}^3$, and that the

effects of the pollutants depend nonlinearly on the characteristic weather. The complexity of the many parameters studied, including the qualitative nature of some variables and the multiple nonlinear relationships, makes it difficult to apply classical regression techniques. This example shows that neural networks do have the capability to deal with these variations in an epidemiology database.

References

1. BAXT W. Use of an artificial neural network for data analysis in clinical decision-making; the diagnosis of acute coronary occlusion. *Neural Comput* 1990; 2: 480–9.
2. BOONE J, GROSS G, GRECO-HUNT V. Neural networks in radiologic diagnosis. *Invest Radiol* 1990; 25: 1012–23.
3. BYLIN G, HEDENSTIERNA G, LINDVALL T, SUNDIN B. Ambient nitrogen dioxide concentrations increase bronchial responsiveness in subjects with mild asthma. *Eur Respir J* 1988; 1: 606–12.
4. CAREY M, CORDON I. Asthma and climatic conditions: experience from Bermuda, an isolated island community. *BMJ* 1986; 293: 843–4.
5. CHI E, REINSEL G. Models for longitudinal data with random effects and AR(1) errors. *J Am Stat Assoc* 1989; 84: 452–9.
6. DHC report. Guidelines for the diagnosis and management of asthma. National Asthma Education Program, Expert Panel Report. Washington, DC: US Department of Health and Human Services, 1991.
7. DOCKERY D, SPEIZER F, STRAM D, WARE J, SPENGLER J, FERRIS B Jr. Effects of inhalable particles on respiratory health in children. *Am Rev Respir Dis* 1989; 139: 587–94.
8. DUTCH H, WLED G. Artificial neural networks and their use in quantitative pathology. *Anal Quant Cytol Histol* 1990; 12: 379–93.
9. ENEVOLDSEN K, GHEORGHE A, MOSEHOLM L, SØRENSEN B. The effect of low-level air pollution and weather on asthma and chronic bronchitis patients, studied by neural network methods. International Conference on Artificial Neural Networks, 24–28 June 1991, Helsinki, Finland.
10. FREUDENTHAL P, ROTH H, HAMMERSTROM T, LICHTENSTEIN C, WYZGA R. Health risks of short-term SO_2 exposure to exercising asthmatics. *JAPCA* 1989; 39: 831–5.
11. GABRIEL K, ODOROFF C. Biplots in biomedical research. *Stat Med* 1990; 9: 469–85.
12. GELADI P, KOWALSKI B. Partial least squares regression: a tutorial. *Anal Chim Acta* 1986; 185: 1–17.
13. GERGEN P, WEISS K. Changing patterns of asthma hospitalization among children: 1979 to 1987. *JAMA* 1990; 264: 1688–92.
14. GUERRIERE M, DETSKY A. Neural networks: what are they? *Ann Intern Med* 1991; 115: 906–7.
15. JAakkola J, PAUNIO M, VIRTINEN M, HEINONEN O. Low-level air pollution and upper respiratory infections in children. *Am J Public Health* 1991; 81: 1060–3.
16. KULLE T, SAUDER L, SHANTY F, et al. Sulfur dioxide and ammonium sulfate effects on pulmonary function and bronchial reactivity in human subjects. *Am Ind Hyg Assoc J* 1984; 45: 156–61.
17. KUUSISTO P. Outdoor air and asthma. PhD thesis. University of Helsinki, 1990 (summary).
18. LAIRD N, WARE J. Random-effects models for longitudinal data. *Biometrics* 1982; 38: 963–74.
19. LEBOWITZ M, COLLINS L, HOLBERG C. Time series analysis of respiratory responses to indoor and outdoor environmental phenomena. *Environ Res* 1987; 43: 332–41.

20. LUISBERG C. Low-priced and robust expert systems are possible using neural networks and minimal entropy coding. Roskilde, Denmark: Risø National Laboratory, 1991.
21. LØDRUP K, CARLSEN K, LEEGAARD J, HOLT J, HALVORSEN K, SIVERTSEN B. Asthma hospital admission in children and the relationship to air pollution in two Norwegian cities. *Am Rev Respir Dis* 1990; 141: A144.
22. MOSEHOLM L, TAUDORF E. Effects of air pollution and weather on asthmatics and persons suffering from chronic bronchitis in 4 Danish cities. An 8-month prospective study. Status Report, May 1990, of the Air Pollution Group under the Asthma-Allergy Association in Denmark. Copenhagen, Denmark: COWIconsult, 1990.
23. ØSTERBALLE B, DIRKSEN A, WEEKE B, WEEKE ER. Cutaneous allergy in a Danish multicenter study [English summary]. *Ugeskr Laeger* 1981; 143: 3211-18.
24. PACKE G, ARCHER P, AYRES J. Asthma and the weather. *Lancet* 1983; 2: 281-3.
25. PICCOLO M, PERILLO G, RAMON C, DIDIO V. Outbreaks of asthma attacks and meteorologic parameters in Bahia Blanca, Argentina. *Ann Allergy* 1988; 60: 107-10.
26. POPE C, DOCKERY DW, SPENGLER J, RAIZENNE M. Respiratory health and PM₁₀ pollution. A daily time series analysis. *Am Rev Respir Dis* 1991; 144: 668-74.
27. PÖNKÄ A. Absenteeism and respiratory disease among children and adults in Helsinki in relation to low-level air pollution and temperature. *Environ Res* 1990; 52: 34-46.
28. PÖNKÄ A. Asthma and low-level air pollution in Helsinki. *Arch Environ Health* 1991; 46: 262-70.
29. RUMELHART D, MCCLELLAND J, eds. Parallel distributed processing. Vol. 1. Foundations. MIT Press, Cambridge, MA: 1986.
30. SCHWARTZ J, WYPIJ D, DOCKERY D, et al. Daily diaries of respiratory symptoms and air pollution: methodological issues and results. *Environ Health Perspect* 1991; 90: 181-7.
31. SCOTT M. Study links pollution with increase in asthma. *Environ Health* 1990; 29: 184-6.
32. TOUSSAINT G. Bibliography on estimation of misclassification. *IEEE Trans Inf Theor* 1974; 20: 472-9.
33. WARE J, FERRIS B Jr, DOCKERY D, SPENGLER J, STRAM D, SPEIZER F. Effects of ambient sulfur oxides and suspended particles on respiratory health of pre-adolescent children. *Am Rev Respir Dis* 1986; 133: 834-42.
34. WEINSTEIN J, KOHN K, GREVER M, et al. Neural computing in cancer drug development: prediction of mechanism of action. *Science* 1992; 258: 447-51.
35. WEISS K, WAGENER D. Changing patterns of asthma mortality. *JAMA* 1990; 264: 1683-7.
36. WHO. Air quality guidelines for Europe. WHO Regional Publications, European Series No. 23. Copenhagen: World Health Organization, 1987.

This document is a scanned copy of a printed document. No warranty is given about the accuracy of the copy. Users should refer to the original published version of the material.