# Health and the Urban Environment

*V. Air Pollution and Illness in a Normal Urban Population*

JAMES McCARROLL, MD; ERIC J. CASSELL, MD;
DORIS W. WOLTER, RN, MPH; COL JOSEPH D. MOUNTAIN, USAFR (Ret);
JUDITH R. DIAMOND, BS; AND ISABEL M. MOUNTAIN, PhD, NEW YORK

THE DIFFICULTY in demonstrating a meaningful relationship between usual urban levels of air pollution and illness in "normal" populations has been a frustrating experience for investigators in this area.

Much of this difficulty results from the fact that the host response customarily studied—some symptom of disease noted by a subject—may also be the result of many other factors. That portion of symptoms in a population which is related to air pollution must somehow be separated from the background count of the same symptoms resulting from infection, allergy, and numerous other causes. Our problem is to determine if there is a regular relationship, either immediate or delayed, between the appearance of an illness symptom in a population and some measure of air pollution.

## Materials and Methods

The population studied was a reasonable cross section of New York City residents described in detail elsewhere.[1] All 1,822 persons lived within a half-square-mile area of reasonably homogeneous air pollution. A daily record of the appearance of each of 21 symptoms of disease (as well as much other health information) was obtained by weekly interview. Each participant was expected to remain in the study for a full year and the average length of time each person participated in fact was 48 weeks. This analysis is based on the daily illness experience of the 1,090 adults over the age of 15. These represent 35,400 person-weeks or 247,800 person-days of observation on this population.

A monitoring station was established in the center of our study area to supply appropriate meteorologic and air pollution data.[1] To determine whether the levels of air pollution as measured at our station might be related to variations in health in our population, we carried out the following analysis. We selected two symptoms—cough and eye irritation—which might be expected to have some relation to air pollution. Persons with chronic coughs were excluded from this analysis as were all children under age 15. The presence of either symptom for each of the 1,054 days of the study was determined by a positive answer to the questions "Did you have a cough?" and "Did you have itching, burning or tearing of your eyes?" The population was analyzed in three groups: heavy smokers (those smoking 20 or more cigarettes daily), moderate and nonsmokers (using 0-19 cigarettes daily), and total adults (age 15 and over).

As indicators of the levels of air pollution to which the population was exposed, we used sulfur dioxide as measured by the Davis conductivity instrument expressed in parts per million and particulate density as measured by the A.I.S.I. sampler and expressed in COH units. For the purpose of this analysis 24-hour averages were used.

Correlation coefficients between each symptom

*Cross-Correlations Between Two Symptoms and Two Pollutants in One-Month Periods With Proportion of Cross-Correlations Showing an Increase or Decrease From Lag of Zero to One Day (Adults, 15 Years and Over)*

| | Symptoms | | | |
|---|---|---|---|---|
| | Cough | | Eye Irritation | |
| | Increase | Decrease | Increase | Decrease |
| Pollutants | | | | |
| $SO_2$ | 23/29 | 6/29 | 11/29 | 18/29 |
| Particulate density | 20/29 | 8/29 * | 14/29 | 15/29 |

$$\text{Each Ratio} = \frac{\text{No. showing increase or decrease}}{\text{Total No. of cross-correlations}}$$

* One correlation between cough and particulate density showed no difference between value for lag 0 and lag 1.

Submitted for publication April 7, 1966; accepted July 22, 1966.

Read before the Eighth Annual AMA Medical Research Conference, Los Angeles, March 2-4, 1966.

From the Division of Epidemiologic Research of the Department of Public Health, Cornell University Medical College, New York.

Reprint requests to Department of Preventive Medicine, School of Medicine, University of Washington, Seattle 98105 (Dr. McCarroll).

tions, A Report of the Conclusions of CIBA Guest Symposium, *Thorax* **14**:286, 1959.

12. Report of Expert Committee on Chronic Cor Pulmonale, WHO Technical Report Series, p 213, 1961.

13. Definitions and Classification of Chronic Bronchitis, Asthma and Pulmonary Emphysema, Statement of the Committee on Diagnostic Standards for Nontuberculous Respiratory Disease, *Amer Rev Resp Dis* **85**:762, 1962.

14. Reid, L.: Measurement of the Bronchial Mucous Gland Layers: A Diagnostic Yardstick in Chronic Bronchitis, *Thorax* **15**:132, 1960.

15. Thurlbeck, W.M., and Angus, G.E.: A Distribution Curve for Chronic Bronchitis, *Thorax* **19**:436, 1964.

16. Mitchell, R.S., et al: The Significance of Morphologic Chronic Hyperplastic Bronchitis, *Amer Rev Resp Dis* **93**:720, 1966.

17. Newman, J.K.; Reiss, O.K.; and Vatter, A.E.: Investigations of the Black Pigment in Human Lung, *Med Thorac* **22**:347, 1965.

18. Wright, R.R.: Elastic Tissue of Normal and Emphysematous Lungs, *Amer J Path* **39**:355, 1961.

19. Blumenthal, B.J., and Boren, H.G.: Lung Structure in Three Dimensions After Inflation in Fume Fixation, *Amer Rev Tuberculosis* **79**:764, 1959.

20. Feinstein, A.R.: Boolean Algebra and Clinical Taxonomy: I. Analytic Synthesis of the General Spectrum of Human Disease, *New Eng J Med* **269**:929, 1963.

## DISCUSSION

DR. BOREN: There is a study by Dr. Reid in which intra-alevolar installation of various mucins is followed by the development of dark-containing pigment within alveolar macrophages. My question to Drs. Thurlbeck and Myrvik, since the question of pigment has repeatedly been discussed for the last several years, is whether or not this might not represent failure of removal of surfactant or else retrograde flow of mucin to the alveolar level.

DR. THURLBECK: I'm not precisely sure what this question was. You wish to know in what way is the pigment that Tom is describing related to surfactant?

DR. BOREN: You don't know what this pigment is related to. Is there a possible origin of this pigment from phagocytosis of mucin or such substances at the alveolar level in the disease state?

DR. THURLBECK: Well, to be honest, nobody can tell you the answer insofar as the pigment is not known, but if one is allowed to make an inspired or uninspired guess, the answer is no.

DR. ANDERSON: Dr. Myrvik, have you an inspiration?

DR. MYRVIK: I know of no information that would suggest that it might originate from surfactant. There are various types of pigments that seem to accumulate in alveolar macrophages, but we have no information how long they might persist.

DR. KILBURN: I'd like to ask Dr. Petty if cough did not decrease, did sputum production decrease in the people who stopped smoking?

DR. PETTY: It is very difficult, to quantitate both cough and sputum production. It is our impression that sputum production parallels cough, but we really have no measured data on this. I really don't know.

DR. ANDERSON: In the pre-print, there is a statement that you have recently ascertained that chronic cough and expectoration during life correlate only roughly with the histological lesions, mucous gland hyperplasia found in the postmortem examination. I wonder if you might comment further.

DR. PETTY: I would be happy to. Chronic cough and expectoration, the cardinal symptoms of chronic airway obstruction, have been carefully recorded in our patients from the emphysema registry, many of whom eventually come to autopsy. Chronic cough and expectoration correlate much better with emphysema than with mucous gland hyperplasia.

CHAIRMAN: Dr. Thurlbeck, would you agree?

DR. THURLBECK: It depends on exactly whose numbers you want to believe. I know of two studies. One is Dr. Reid's original paper and one is my own. What you describe is a good measurement, almost a perfect one. In our hands, using essentially the same technique the figure, I believe, is something like 30% of the people with very large mucous glands, did not form spit, but again this is retrospective from clinical charts, and what is this worth really?

DR. PETTY: There is a real problem with the accuracy of histories from clinical charts which may explain some of the differences.

DR. BOUCOT: Dr. Petty, did you attempt to correlate the presence of pigment with occupation?

DR. PETTY: We did not, however, we did exclude patients that might have a great degree of anthracosis from the coal or hardrock mining industry in Colorado.

These patients were excluded from the pigment study. We found no other correlation with occupation.

DR. WINKELSTEIN: I wonder if Dr. Petty and some of the other pathologists might comment on the problem of sampling in autopsies. One of the aspects which comes up repeatedly in epidemiological studies is the difficulty of getting a representative sample of deaths for autopsy. Could this be improved if instead of asking for complete autopsies, studies were designed which would go after particular information.

For example, it might be possible to get one slice of the lung of a random sample of deaths in a community. Families might be willing to give permission for partial autopsy where they wouldn't be willing to give permission for total autopsy.

Could this be perhaps an approach to improving our understanding of the epidemiology of some of these conditions?

DR. PETTY: Fortunately, we are able to obtain autopsies in our hospitals 92% of the time. Thus we can go to the autopsy room and get a lung from almost every case.

We obtained a single lung from each of these cases, losing very few. Regarding localized autopsies, these may be an advantage to some places.

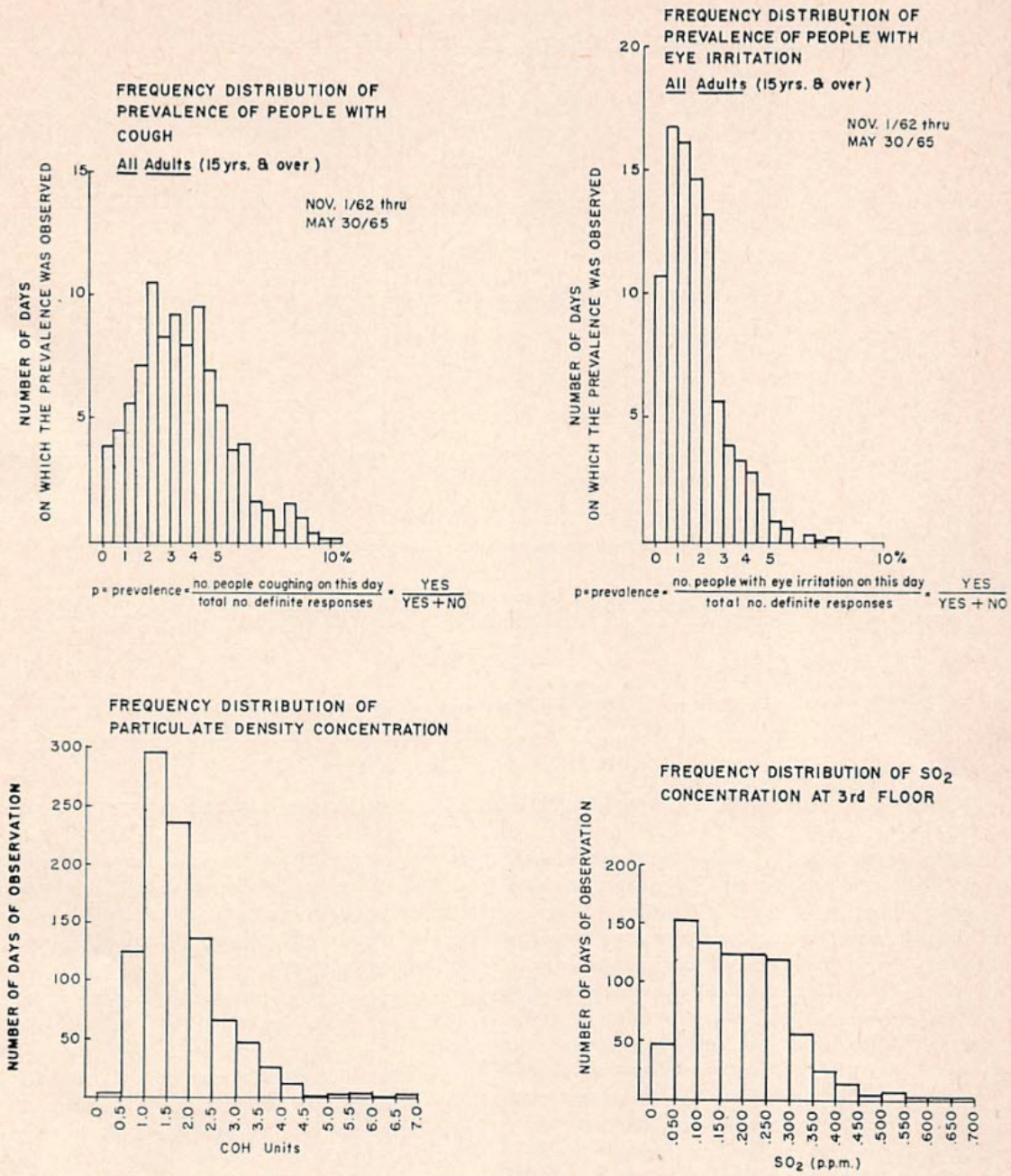


Fig 1.—Frequency distribution of the symptoms in the defined population and the frequency distribution of the two measures of pollution.

and each measure of pollution for the three population groups were then calculated. Because the appearance of a symptom on a given day might well be related to levels of pollution at some point in the past, cross-correlations between each symptom and each measure of pollution were calculated with a time lag of 0 to 28 days. In addition to the cross-correlations between each symptom and each measure of pollution, autocorrelation for each symptom and each measure of pollution were also calculated between the value for a given day and for each preceding day back to 28 days. The autocorrelations of each symptom are in effect measures of the persistence of that symptom in a population once it has appeared. The autocorrelations of each measure of pollution are similarly measures of the persistence of a pollutant in the environment. The frequency distribution of the symptoms in the defined population as well as the frequency distribution of the two measures of pollution used are graphed in Fig 1.

Since it is difficult to determine small differences in a mass of correlation coefficients listed in tabular form it was decided to plot the values as a correlogram for each symptom and pollutant in the different populations for each period over which we examined this pair of values. A total of 516 such correlograms were plotted. All data were processed on an IBM 7040 computer. Because the system under study is a dynamic one the time series is not stationary and thus the calculation of correlation coefficients over the entire study period

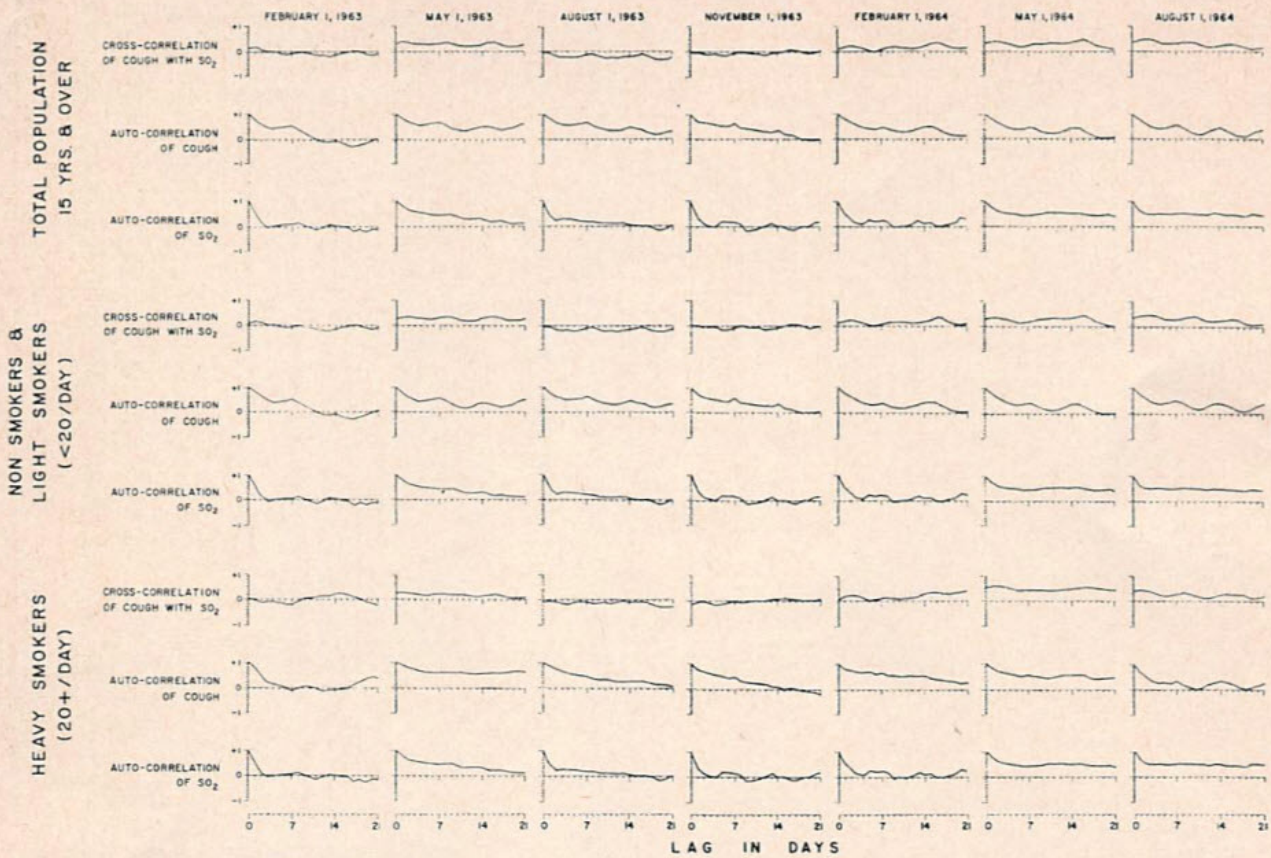


Fig 2.—Correlations of $SO_2$ and cough with 0 to 21 day lags for overlapping six-month periods.

of nearly three years (1,054 days) might well obscure significant relationships. We therefore chose to examine separately shorter lengths of time so that possible significant correlations would not be submerged. In the technique of time-series analysis as expounded by Blackman and Tukey,[2] it is recommended that for examination of correlation coefficients in which lags of varying degrees are introduced, the basic period of observation should be a minimum of ten times the maximum lag examined. To examine our data by autocorrelation and cross-correlation with a maximum lag of 28 days it was necessary to pool a minimum of 280 days of observations. Pooling data for such long periods of time might tend to submerge some valid correlations and we therefore decided to examine shorter time periods and reduce the length of time lag accordingly. We calculated our correlation coefficients in four different ways: first, with lags of 0 to 28 days based on three successive one-year periods; second, the same lags based on the total period; third, with lags of 0 to 21 days based on overlapping six-month periods; and finally, with lags of 0 to 3 days based on successive one-month observation periods. In this way we were able to sharpen the correlations for the first few days of a lag period as well as examine the seasonal variations in our data. In this technique one must compromise between the calculation of longer lag periods which must be based upon similarly longer periods of observation in which some factors may be hidden, or settle for shorter periods of observation which may permit identification of many seasonal and other factors, but at the sacrifice of the length of lag period which can be reasonably calculated.

In employing a correlation coefficient method of analysis it is important to establish that the variables are approximately normally distributed. Examination of daily averages of particulate density showed the frequency distribution to be approximately normal in the gaussian sense although slightly skewed to the right. Similar examination of the $SO_2$ data showed a curve somewhat flatter than the typical bell-shaped distribution and which was also somewhat skewed to the right. In future analyses of these data we will include a log transformation of the values of the pollutants in order to derive a normal distribution of these observations. The graphs of frequency distribution of the symptoms show that cough tends to be normally distributed, whereas eye irritation is distributed on a curve more markedly skewed to the right, reflecting the fact that the mean prevalence is relatively low (Fig 1).

## Results

The results obtained in this analysis are presented in a series of correlograms. Most of these are based on overlapping six-month

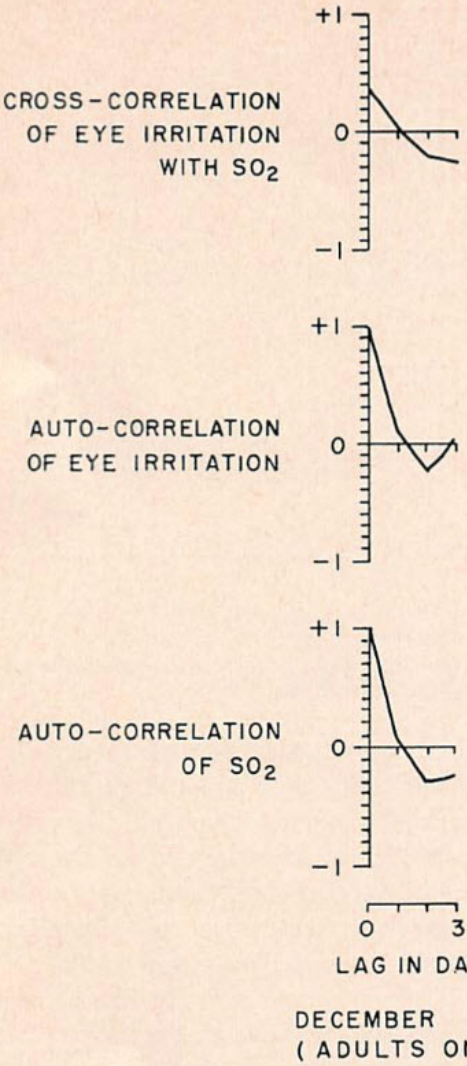


Fig 3.—Correlations for typical one-month period showing rapid drop in cross-correlations of eye irritation and $SO_2$.

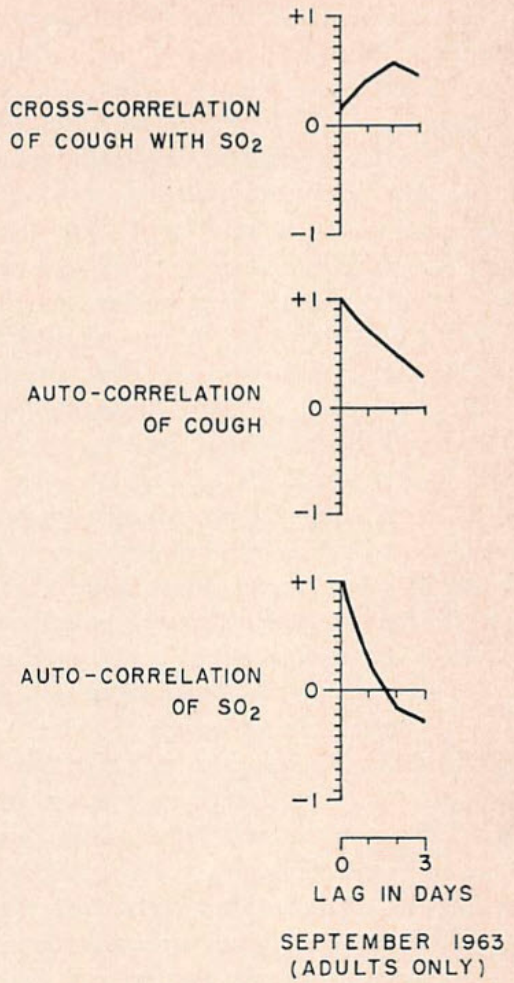


Fig 4.—Correlations for typical one-month period showing delay in appearance of cough correlated with $SO_2$.

periods of observations starting three months apart throughout the three years of the study. Three correlograms are presented in each figure. The first is a cross-correlation between a symptom (cough or eye irritation) and a measure of pollution ($SO_2$ or particulate density). The correlation coefficients are calculated for six-month periods based on the level of pollutant on any given day correlated with the presence or absence of a particular symptom on that and each preceding day for 21 days.

The second correlogram on each figure is an autocorrelation of the symptom with a time lag of up to 21 days. Obviously at lag zero the correlation is one. The correlation of the prevalence of a symptom on each day with the prevalence on each of the preceding 21 days for a six-month period is diagrammed. The third correlogram of each figure is the autocorrelation of the pollutant indicating the relation of the level of pollutant on each day to the level of the pollutant on each of the preceding 21 days. Examples of correlograms of correlation coefficients for both cross-correlations of symptoms and pollutants, and for autocorrelations of each symptom and pollutant are illustrated in Fig 2. These are a portion of the correlograms calculated with up to 21 days of lag, and are based on overlapping six-month periods of observations. Other correlograms were prepared for successive one-month periods with three-day lags, and for one-year periods using 0 to 28-day lags.

Inspection of the correlograms reveals differences between cough and eye irritation, and between the environmental influences represented by sulfur dioxide compared with

those represented by particulate density. It is apparent from Fig 3 and 4 that cough and eye irritation behave quite differently in this population when correlated with $SO_2$ levels. The cross-correlation of $SO_2$ and eye irritation characteristically shows a drop between the value for lag 0 and lag 1; whereas, the cross-correlation between $SO_2$ and cough typically shows an increase between lag 0 and lag 1. A different pattern emerged when the cross-correlations of the same symptoms with particulate density were examined: those with cough show a similar increase between values for lag 0 and lag 1, whereas the cross-correlations between particulate density and eye irritation show no consistent relationship. This seems to indicate that the environmental influences represented by $SO_2$ have an immediate effect in the production of eye symptoms, and that there is a delay in the production of cough. Similarly particulate density appears to make no contribution to the production of eye symptoms, but does appear to contribute to the production of cough. This is consistent with other evidence that the production of eye irritation is an immediate effect of some pollutants, but that a time lag may be necessary in the production of any but irritative cough. That the cough in question is more than an immediate and transient mucous membrane irritation is suggested by the delayed fall-off from lag 0

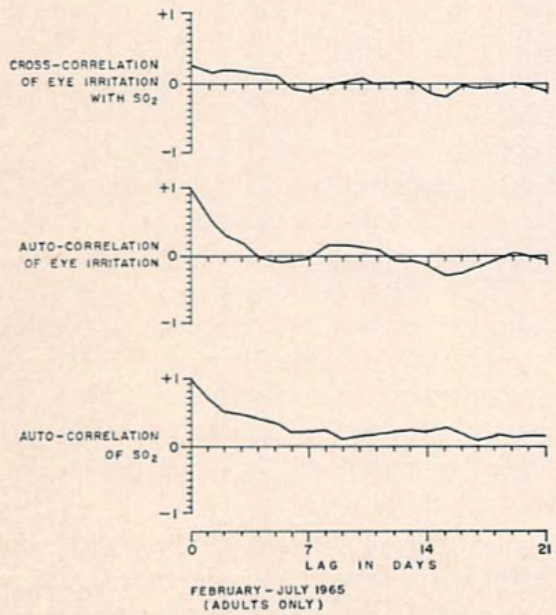


Fig 5.—Correlations for typical six-month period showing lack of persistence of eye irritation.

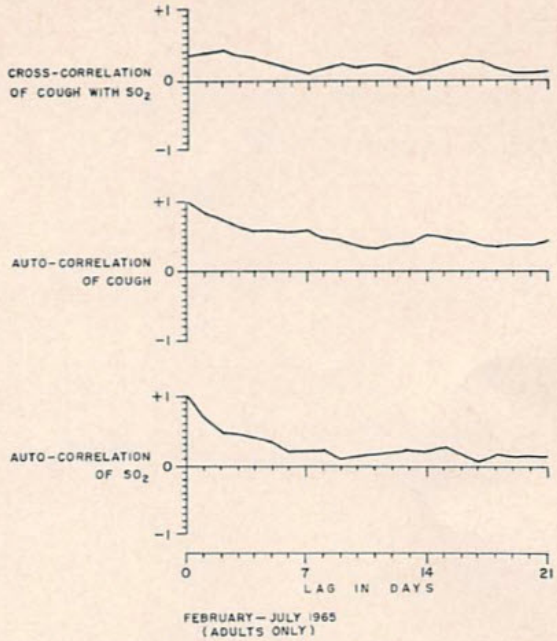


Fig 6.—Correlations for typical six-month period showing persistence of cough.

in the autocorrelation of cough. Indeed, it is reasonable that cough as a symptom of disease is rarely present for only a day or two. By contrast, the correlogram of the autocorrelation of eye symptoms shows a sharp fall-off from lag 0 and that, too, is consistent with clinical experience (Fig 5 and 6).

If the environmental influences represented by $SO_2$ and particulate density were the sole cause of cough, then the correlation coefficient would approach unity. Although

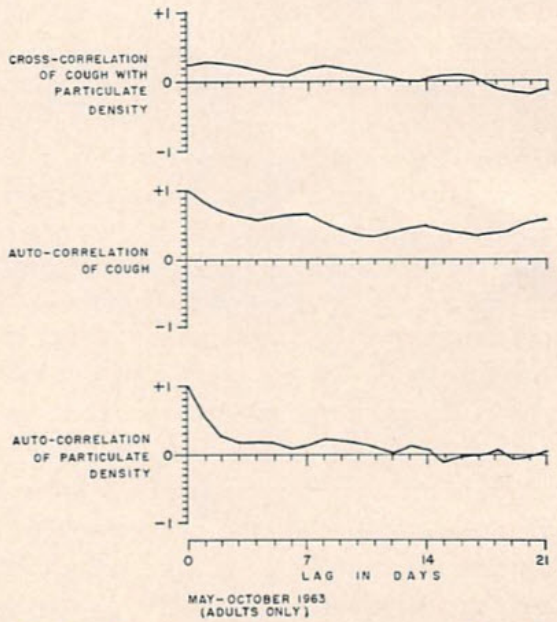


Fig 7.—Correlations for typical six-month period of cough and particulate density with suggestion of possible periodicity.

characteristically positive, they do not approach one. If these environmental influences in no way contribute to the production of cough, the relationship between the correlation coefficient on lag 0 and lag 1 would be random. In fact, however, the relationship between $SO_2$ and both cough and eye irritation is strong. For particulate density and eye irritation, however, the relationship between the lag 0 and lag 1 values is random. When these data are examined by one-month periods and the directions of the curves between lag 0 and lag 1 are summated, the impressions obtained by visual inspection of the correlograms are confirmed. The actual numbers of increases and decreases for each cross-correlation are listed in the Table. It would therefore appear from the use of time-series analysis that whatever is represented by $SO_2$ is not the sole cause but does contribute to the presence of cough and eye irritation in this adult population.

Similarly, whatever is represented by particulate density determinations is not the sole cause, but does appear to contribute to the presence of cough, and does not contribute to the presence of eye irritation. The difference between the measurements of $SO_2$ and particulate density in relation to these symptoms as indicated in the cross-correlations is particularly striking and suggests it may be possible to identify specific components or sets of conditions which may be related to appearance of specific symptoms. From these conclusions one should be able to predict that the correlation coefficient between both $SO_2$ and particulate density and cough will be higher than between $SO_2$ and eye irritation and that is, indeed, true. Similarly, the correlation coefficient of particulate density and eye irritation should range closely around zero and that also is true.

It is reassuring that the symptom expression in this population behaves in an expected manner. Cough, once it appears, tends to persist for considerably longer periods than does eye irritation which is known to be a more transient response. It is intriguing that repeatedly through these correlograms there appears a suggestion of periodicity which obviously warrants further investigation. Fig 7 illustrates this particularly well with suggestions of peaking at 7 and 14 days which is difficult to explain on the basis of our present knowledge.

## Conclusion

In applying the techniques of time-series analysis to large volumes of health and air pollution data over long periods, it is possible to demonstrate a consistent relationship between certain measures of environmental pollution and the presence of symptoms in a "normal" urban population.

This work was supported by the Health Research Council of the City of New York (contract U-1155) and the Division of Air Pollution of the US Public Health Service (AP 00266).

## REFERENCES

1. McCarroll, J., et al: Health and the Urban Environment: Air Pollution and Family Illness: I. Design for Study, *Arch Environ Health* **10**:357-363 (Feb) 1965.

2. Blackman, R.B., and Tukey, J.W.: The Measurement of Power Spectra From the Point of View of Communications Engineering, New York: Dover Publications, Inc., 1959.

## DISCUSSION

DR. ANDERSON: Clearly, this is one way of attempting to handle mass data. I suggest that we look at the method since the authors have really very few results to show to us yet.

DR. GOLDSMITH: In the first place, I am very impressed by this application of an important method. I am a little reluctant to comment on the statistical aspects of the problem in the presence of such eminent statisticians who will speak later, but perhaps if an epidemiologist does discuss this, it will reduce the number of comments.

First, the autocorrelation approach does not seem likely to be relevant to the number of variables which might affect these phenomena, but only to the possibility of an association between the two variables chosen.

There will be some random variation and other variables, but it does not necessarily follow that because this type of result is observed, that this indicates the absence of many other variables or of any specific variable.

The second thing is that in the group of correlograms, there are some weekly accentuations in correlation. In particular, I note a difference in the autocorrelation between smokers and nonsmokers in the presence of a weekly cycle which also seems for some of the periods to be associated with a weekly cycle of $SO_2$.

This is extraordinarily interesting. Of course, one would like to know a little bit more concerning the weekly cycle of autocorrelations in $SO_2$. The full power spectrum analysis should give some better reflection than this simple autocorrelation; perhaps in time, this will be available because there are no doubt other trends and perhaps there are 30-day cycles. I imagine they will not be quite as pronounced, but the point is that the method permits them to be detected.

It occurred to me that the people who were asked so many questions this frequently would surely figure out some way to get the interviewer to stop bothering them by giving stereotyped answers, and now I am convinced that at least they did not give stereotyped answers; I am sure, of course, that this is not true in all other respects.

Some of the autocorrelations are undoubtedly due to the rather low values present on many of these initial days. I tried to construct in my mind what the correlation matrix was for each one of these points, and I suspect that a number of observations cluster around the origin, that is, around zero, and that there are just one or two out in the fairly high area, rather than the values being uniformly distributed along the frequency of the particular phenomenon: this might be elaborated later.

Finally, I would like to point out that some of the assumptions about the distribution of data which are discussed in the paper are not altogether consistent with the histograms shown, which, in fact, are Poisson distributions; there are some very powerful statistical methods suitable for dealing with distributions which ought to be Poisson in nature and which might be very fruitfully applied to some of these data. They consist of choosing samples of days with certain attributes, and seeing whether the Poisson distribution for those days differs in fact from the Poisson distribution of days without those attributes. It is a very sensitive method and one that would be quite appropriate for these data.

DR. ANDERSON: That specifically relates to eye irritation, which is very definitely Poisson distributed.

DR. CASSELL: Periodicity was extremely interesting to us and as Dr. Goldsmith knows one of the strengths of the method is its pulling out periodicity.

We are somewhat suspicious. It is too neat, and anything that neat just cannot have been a gift from the gods. So we are examining it very closely before we accept that there is indeed true periodicity in the data.

DR. ANDERSON: There was a periodicity in the interviewing schedule also, though.

DR. CASSELL: There are weekly interviews but they do not all occur on the same day. The interviews are spread out so that if that is indeed the reason two things should be true, periodicity should be equal in every symptom and it is not. If it is solely the interviewing, the peaks should be broader than they appear to be and they are not. However, we are extremely skeptical about it, and we are looking at it closely.

---

## ADULT IMMUNIZATION

Immunization programs sponsored by pediatricians and school health authorities have given members of our younger generation a high degree of protection against infectious and communicable diseases. For adults, however, the situation is quite different. Except for those in the military services and those who have had occasion to travel abroad, only a small percentage has had more than an original vaccination against smallpox, and most have not been immunized against tetanus. Crash programs of smallpox vaccination can be instituted on short notice, but in order to make them effective, the public must know that successful vaccination in childhood does not insure life-long protection. Tetanus is completely preventable. Every adult should be urged to receive an initial course of tetanus toxoid and subsequent booster doses at appropriate intervals. Increased emphasis also should be placed on the importance of immunizing elderly people and those with chronic diseases against influenza and the more common types of pneumococcus pneumonia.—Ernstene, A.C.: The Internist and Preventive Medicine, *Bull Amer Coll Physicians* **7**:155-157, (May) 1966.