A group of Manhattan residents was questioned weekly on the occurrence of acute respiratory symptoms. Incidence and prevalence rates of "common colds" were related to environmental variables. Findings are discussed.

HEALTH AND THE URBAN ENVIRONMENT

VIII—AIR POLLUTION, WEATHER, AND THE COMMON COLD

Donovan J. Thompson, Ph.D., F.A.P.H.A.; Michael Lebowitz, M.S.; Eric J. Cassell, M.D.; Doris Wolter, R.N., M.P.H., F.A.P.H.A.; and James McCarroll, M.D., F.A.P.H.A.

Introduction

Although the ubiquitous air pollution to which urban residents are constantly exposed has been widely condemned as a community health hazard, solid evidence of such an effect has been difficult to demonstrate. The notorious, but rare, acute air pollution episodes (Donora, 1948; London, 1952) took their greatest toll from the elderly—largely those with pre-existing cardiorespiratory disabilities. The two principal questions posed by such evidence are: (1) Does air pollution, at its usually prevailing lower levels, also affect health? (2) If so, are ordinary urban populations involved?

Studies on the behavior of airborne respiratory disease agents suggest that one mechanism through which air pollution may produce such an effect could be by altering susceptibility to, or transmission of, such agents in a population. Complicating any analysis of the relationship of levels of air pollution to incidence or prevalence of such infections are the intricate interrelationships between various meteorological factors and pollution levels. To assess the relative importance of meteorological and pollution factors in common upper respiratory illness, the following analysis was performed.

Materials and Methods

The study providing the data on which the analyses reported here are based has been described fully elsewhere (McCarroll, 1965; McCarroll, 1966). Briefly, the data on the "common cold" were obtained from a panel of families residing in the lower east side area of New York City through weekly interviews. The meteorology data came from the official U.S. weather station located in Central Park, about 4.5 miles from the center of the study area. The air pollution data were obtained from special monitoring equipment installed in the study area specifically for the study. The variables included in these analyses are:

"Common Cold"

1. Incidence rate/1,000/day
2. Prevalence rate/1,000/day

Environmental Variables

1. Pollutant indexes
   a. Particulate matter (COH units)
   b. Carbon monoxide (CO/ppm)
   c. Sulfur dioxide (SO_2/ppm)

2. Meteorological variables (24 hour averages)
   a. Temperature (°F)
   b. Relative humidity (%)
   c. Wind velocity (mph)
   d. Barometric pressure (inches Hg.)
   e. Solar radiation (gram cal./cm^2/sec)

In earlier analyses of these materials, various periodicities have been demonstrated; these may have been due either to the periodic nature of the interrogation regarding the health status of panel members (a weekly questionnaire), or to the variation introduced into meteorological and air pollution measurements by season of year and day of week (e.g., traffic patterns, industrial activity, and the like). The influence of season on these variables, and the levels recorded for New York City during this period, may be examined in Table 1. To eliminate these sources of variation from our analyses, aimed at studying the relation between the common cold and meteorological and pollutant variables, we have made each analysis specific for the consecutive nine seasons (winter, spring, summer, fall) during the study period for which panel membership and pollution data are sufficiently complete to warrant analysis. Within each season we have removed the variability attributable to the day of the week, by taking deviations from the day average and standardizing these deviations by division by the standard deviation. Thus, for the prevalence of common cold on a particular Wednesday during the spring of 1963, we have as a basic unit for analysis

$$C=\frac{x-\bar{x}}{s}$$

where x is the prevalence rate for the Wednesday in question, $\bar{x}$ is the average prevalence rate for all Wednesdays in the spring of 1963, and s is the standard deviation of these prevalence rates. (For another example of the use of this technique, see Sterling, 1966.)

Previous analyses of health data from this study have been restricted to the prevalence of a symptom or health-related condition on a given day. Since the role of meteorological and pollutant variables in influencing the common cold is poorly understood, we wished to be able to make separate analyses of the influence of these variables on the onset of the common cold (incidence), as well as the duration of the common cold once acquired (prevalence). Moreover, provision was made for examining the effect of these influences on the onset of a common cold for time lags up to 72 hours. In defining incidence, we have required only that an onset be preceded by one or more days, when the question concerning common cold was answered "no" (McCarroll, 1965). Some notion of the symptoms prompting a report of common cold may be gained by noting that such a report involved rhinitis 90 per cent of the time, cough 41 per cent, sore throat 34 per cent, feverishness 13 per cent, and gastrointestinal symptoms in 8 per cent of reported common colds.

All the findings reported here are based on stepwise multiple regression analyses of the standardized variables described above. If the incidence (or prevalence) of the common cold were linearly related to the total set of available meteorological and pollutant variables, the stepwise procedure would generate the coefficients a_i in the following linear equation:

$$C=a_o+a_1x_1+a_2x_2+\cdots+a_8x_8$$

In this equation, C is the standardized common cold incidence (or prevalence) measurement and x_1 to x_8 are the standardized measurements of the eight pollutant and meteorological variables. The stepwise procedure has the property that it enters the pollutant and meteorological variables one at a time in order of their ability to "explain" the variability in the common cold incidence,

Table 1—Mean and standard deviations of daily incidence and prevalence rates and environmental variables by season

	Winter 1962–63		Spring 1963		Summer 1963		Fall 1963		Winter 1963–64		Spring 1964		Summer 1964		Fall 1964		Winter 1964–65	
	$\bar{X}$	S	$\bar{X}$	S	$\bar{X}$	S	$\bar{X}$	S	$\bar{X}$	S	$\bar{X}$	S	$\bar{X}$	S	$\bar{X}$	S	$\bar{X}$	S
"Common cold" prevalence rate/1,000	88.0	23.4	69.9	33.8	45.6	18.3	107.4	17.5	106.7	18.6	74.3	35.7	31.2	10.8	86.6	24.2	95.3	18.6
"Common cold" incidence rate/1,000	10.5	7.6	8.6	6.0	10.3	8.8	12.8	7.7	11.2	3.6	7.2	4.1	3.4	3.7	10.6	4.3	11.7	7.0
Average no. of participants	336		270		190		545		644		654		533		477		373	
Pollutant indexes																		
Particulate matter (COH)	3.15	1.19	1.58	0.58	1.18	0.42	1.68	0.61	2.19	0.75	1.56	0.74	1.16	0.34	1.76	0.46	2.02	0.87
CO (ppm)	4.38	1.99	3.02	2.07	3.26	0.81	4.03	2.08	6.00	3.31	3.61	2.44	1.78	0.56	3.36	1.84	2.94	0.99
SO_2 (ppm)	0.17	0.13	0.10	0.04	0.17	0.10	0.15	0.09	0.24	0.19	0.14	0.09	0.09	0.04	0.16	0.05	0.25	0.07
Meteorologic variables																		
Wind velocity (mph)	10.3	3.8	11.2	4.8	8.2	3.4	9.5	4.1	11.9	3.9	9.8	3.5	8.3	1.5	8.2	2.7	9.0	2.9
Solar radiation (gram cal/cm^2/sec)	0.2	0.1	0.4	0.2	0.4	0.1	0.2	0.1	0.3	0.1	0.4	0.2	0.4	0.1	0.2	0.1	0.2	0.2
Temperature (°F)	33.2	7.1	57.9	9.9	71.8	5.8	48.6	12.2	33.9	7.1	57.2	11.7	72.2	7.0	48.6	12.6	35.9	6.8
Relative humidity (%)	64.7	15.2	58.9	15.3	62.5	11.6	57.8	16.4	55.5	16.2	61.9	19.8	65.6	13.7	57.8	16.6	63.2	14.7
Barometric pressure (mm. Hg.)	29.9	0.3	29.9	0.2	30.0	0.2	30.1	0.2	30.0	0.3	30.0	0.2	30.0	0.2	30.1	0.2	30.0	0.2

as measured by the square of the multiple correlation coefficient. The procedure terminates when none of the remaining unentered variables significantly increases the explained portion of the variability in the common cold incidence (or prevalence). Thus, in any given analysis of the relation between the health measurement and the environmental variables, the procedure could produce an equation with as many as nine terms on the right-hand side. The magnitude and sign of the coefficients in the equation determined by the procedure express the "strength" and "sense" of the contribution of that variable to the linear determination of the common cold incidence (or prevalence). Since standardized variables were used, the coefficients may be compared to one another to obtain relative contributions of each of the variables in the same and different seasons.

The size and composition of the respondent panel varied throughout the three years of the study period, averaging from 190 to 650 people providing information on a given day. This depended upon the willingness of those enrolled to continue, and the recruiting of new groups for participation. The number of common cold episodes ranged from 150 to 480 per season. The prevalence and incidence rates of common cold calculable from these data are therefore, at best, crude estimates of prevalence and incidence levels for the area in which the study was made. However, with respect to examining the relation between changes in the environmental variables and concurrent or subsequent changes in the health status of an individual panel member, the same criticisms of potential bias and lack of representativeness lose much of their force. If linear relations between the onset or duration of the common cold and the environmental variables included in the analysis exist, the panel should be able to demonstrate them.

Findings

The results of the regression of incidence of common cold on the environmental variables for nine seasons are summarized in Table 2. Perhaps the most interesting points emphasized by this table are that: (1) there is little consistency from one season to another in the variables involved in the regression; (2) all the meteorological variables that enter the regression (with the single exception for temperature in the winter of 1962-1963) enter negatively (inverse relations) and the pollutant variables enter positively (direct relations); and (3) only in the spring of 1964 do these variables "explain" a sizable fraction of the common cold incidence $R^2=(.70)^2=49$ per cent.

Since the influence, if any, of environmental variables of the type considered here on the onset of common cold might logically occur on days preceding the day of onset, we next looked at the same regressions with incidence lagged one, two, and three days. In Table 3 are shown the results for a lag of three days. Lags of one and two days produced very similar results to those in Table 3, but the relations were strongest for three days, as measured by the per cent of the variation in incidence explained by the regression.

While it is still the case in Table 3 that there is little consistency from season to season in the relation between cold incidence (lagged three days) and the environmental variables, the general improvement in the proportion of the variance explained by the regression is clearly in evidence. The quite remarkably high multiple regression coefficient for the spring of 1964, $R=.92$, indicates that over 85 per cent of the variance in cold incidence during that season is "explained" in terms of regression on the five variables indicated. A second general conclusion from Table 3 is that pollutant variables enter the re-

Table 2—Coefficients in the stepwise multiple regression analysis[1] of incidence of "common cold" on stated environmental variables by season—New York City

Season	Particulate matter	CO	SO_2	Wind velocity	Solar radiation	Tempera- ture	Relative humidity	Barometric pressure	R^2	R^2 (%)
Winter 1962–63			.31[2]			.24[1]			.37	14
Spring 1963			.13			−.22[2]		−.22[1]	.32	10
Summer 1963		.38[2]			−.24[1]	−.29			.47	22
Fall 1963										
Winter 1963–64										
Spring 1964						−.77[1]		−.29[2]	.70	49
Summer 1964					−.28[2]		−.38[1]		.28	8
Fall 1964										
Winter 1964–65							−.26		.26	7

[1] The order in which the two most important variables enter the regression is indicated by the superscript on the coefficients. R = multiple correlation coefficient.

Table 3—Coefficients in the stepwise multiple regression analysis[1] of incidence of "common cold" on stated environmental variables lagged 3 days by season—New York City

Season	Particulate matter	CO	SO_2	Wind velocity	Solar radiation	Tempera- ture	Relative humidity	Barometric pressure	R	R^2
Winter 1962–63	−.24		−.31[2]				.49[1]		.47	22
Spring 1963						−.29[2]	−.23	−.52[1]	.55	30
Summer 1963		.21		−.35[2]			−.32[1]		.56	31
Fall 1963										
Winter 1963–64						−.31			.29	8
Spring 1964		.27	.38[2]		.26		−.65[1]	−.11	.92	85
Summer 1964						−.49[2]	−.22[1]	−.40	.47	22
Fall 1964										
Winter 1964–65		.30	−.41			−.41[2]			.45	20

[1] See footnote to Table 2.

Table 4—Coefficients in the stepwise multiple regression analysis[1] of prevalence of "common cold" on stated environmental variables by season—New York City

Season	Particulate matter	CO	SO_2	Wind velocity	Solar radiation	Temperature	Relative humidity	Barometric pressure	R	R^2
Winter 1962–63	−.59			−.33			−.48[2]	−.57[1]	.53	28
Spring 1963						−.30[2]	−.35[1]	−.28	.58	34
Summer 1963	.34	.24	−.35[1]			−.46[2]			.56	31
Fall 1963	.31				.49[2]	−.71[1]		−.44	.71	50
Winter 1963–64			−.31		−.33[2]	.24		.52[1]	.53	28
Spring 1964	.30[2]	.11		.13		−.81[1]	−.20	−.18	.93	86
Summer 1964	−.16[1]			−.53			−.54[2]		.47	22
Fall 1964		.30[1]		.21[2]					.34	12
Winter 1964–65		−.30[1]						.26[2]	.42	18

[1] See footnote to Table 2.

gressions in more seasons when incidence is lagged three days, than when a time lag is not considered. However, in three cases of seven in which they enter, the pollutant variables carry negative signs indicating an inverse relation with lagged incidence.

In Table 4 we examine the regressions of prevalence of "common cold" on the same set of variables. Only in one season (spring, 1963) are these regressions at all similar to those for incidence. In particular, the difference between the regressions for prevalence and incidence for fall, 1963, is noteworthy. Although the incidence of common cold appears to have almost no linear relation with the environmental variables during this season, the prevalence has the second highest multiple correlation coefficient recorded in Table 4. Close inspection and comparison of Tables 3 and 4 also reveal that the variables entering the regressions differ markedly from one to the other.

Discussion

The finding that the incidence and prevalence of the "common cold" are linearly associated with environmental variables in markedly different ways, by season of the year, perhaps should come as no surprise. Whatever the influence of these environmental variables on the agents, or host, in initiating or prolonging an episode of illness, it is likely to be complex rather than simple. Among the most obvious possible reasons for the variability of the relations from season to season are: (1) the obvious differences in the measures of environment among seasons in the year, and the same season over different years; (2) the variation in the presence of agents capable of producing the common cold syndrome; (3) the major differences in opportunities for transmitting the agents to susceptible hosts occasioned by changes in school attendance, holidays, and so on. However, it seems unusual to the authors that a multiple linear regression model would account for over 85 per cent of the variability in cold incidence in one season and account for essentially none in another. A chance concurrence of a slowly spreading epidemic with the gradually moderating weather of spring might also be a prime suspect for producing the observed association—the two-time series (daily incidence and environmental measurements) simply may happen to be in phase with one another.* We will return to this possibility later.

The order of importance, the magnitude of the contribution of a given variable to the regressions, and the sign of the contribution provide interesting grounds for speculation. For example, in the five regressions in which humidity is included (Table 3), it is entered as the first variable in the regression in four of the cases; it makes the largest contribution to the regression found in the entire table in two cases (−.65 and .49); and contributes negatively in all but one case, i.e., the winter of 1962-1963. In general, the meteorological variables, when they enter, tend to have inverse contributions to the regressions (all but two cases in Table 3). On the other hand, the pollutant contributions are sometimes direct and sometimes inverse, perhaps indicating either a subservient role to the meteorological variables or greater complexity in their operation in the mechanisms of spread and duration of the common cold. These findings are essentially consistent for all three tables.

The finding of no regression of common cold incidence on the environmental variables for the fall seasons, and the confusing evidence of relationships for the winters, may be due to the different combinations of environmental factors which may produce an effect. It has often been said that the relationship between environment and common colds is a complex, multivariable interaction, so these findings are not surprising. Thus, in winter there are at least two environmental conditions which may affect common colds but in which the variables often operate in different directions: (1) inversions with high barometric pressure, low wind velocity, above average temperature, high levels of pollution, and so on; and (2) turbulent, stormy weather with lower than average temperatures, lower than average levels of pollution, high wind velocity, and so on (Cassell 1968; Lowry 1967).

In the summer seasons, some relatively strong relationships were found, although common cold incidence was at a low level in summer, 1964. It is in the spring seasons that the most consistent and strongest relationships of the environmental variables with common colds are seen. A major question, after finding such high, almost extreme, multiple correlations, is whether the relationships found were due to artifacts.

We examined (1) the characteristics of common colds in individuals and families, (2) the difference in interaction of the environmental variables, and (3) corroborating evidence of the presence of respiratory disease agents in the community, as demonstrated by (a) increased school absenteeism, and (b) isolation of viral agents. Evidence that an epidemic of respiratory infection was underway was suggested by the unusual number of isolations of adeno-

* The possibilities were also considered that (1) fitting a stepwise multiple regression with up to eight independent variables to only 90 points (3 months) might account for the occurrence of these large values of R, and that (2) over a number of seasons Rs of these magnitudes would occur by chance occasionally. However, neither possibility seems an attractive alternative explanation.

Table 5—Pairwise correlations of standard scores of environmental factors for four seasons

Spring 1964	Spring 1963
Winter 1963–64	Summer 1964

	CO		SO_2		Wind velocity		Solar radiation		Temperature		Relative humidity	
1. Particulate	.56	.49	.70	.70	−.42	−.34	−.33	−.28	−.24	.11	.20	.03
matter (COH)	.56	.24	.86	.17	−.59	−.17	−.27	.04	.27	.36	.28	.07
2. CO			.51	.30	−.11	−.22	−.32	.03	−.26	.16	.19	−.04
			.51	.14	−.32	−.27	−.18	−.10	−.18	.21	.14	.13
3. SO_2					−.31	−.37	−.25	−.07	.05	.16	.08	−.05
					−.64	−.05	−.11	−.04	.20	.50	.12	.20
4. Wind							.27	.15	−.06	−.39	−.41	−.33
velocity							.14	−.02	−.11	.00	−.16	.06
5. Solar									.17	.02	−.74	−.59
radiation									−.13	.31	−.73	−.69
6. Temperature											−.10	.28
											.17	−.13
7. Relative humidity												
8. Barometric	0	0	.09	.35	−.31	−.43	.14	.20	−.41	−.02	−.07	−.05
pressure (BP)	−.03	.04	.22	.13	−.46	−.26	.21	.17	−.25	−.18	−.39	−.13

virus type 5 (Brandt, Wassermann, and Fox 1966), no influenza epidemic (Widelock, et al., 1965), and a concomitant increase in school absenteeism in the area. The illnesses reported by the participants, and apparent spread within their families, conform to expected patterns of respiratory illness. The pairwise associations between both pollutant and meteorological variables are recorded in Table 5 for the spring of 1964, the previous spring (1963), and the preceding and following seasons (winter, 1963-1964; summer, 1964). While some associations are stronger during spring, 1964, than for the other seasons (temperature and barometric pressure, for example), in other cases the reverse is true (SO_2 and wind velocity). Although one or more of these factors may, perhaps, account for this unusually high correlation, our data do not permit any firm conclusion on this matter.

Finally, the relation of these findings to those of Ipsen (1967), Lidwell, et al. (1965), and Carne (1966) should be mentioned. Differences in methodology cloud the issue but, somewhat surprisingly, all these investigations reach the same conclusion as that obtained here—among the meteorological and pollutant variables studied, temperature (on the day of onset of illness, or a few days prior) seems to be most highly associated with the incidence of upper respiratory disease. Humidity (relative or absolute) seems to be next in importance, although its influence varies from inverse to direct in different seasons. Air pollutants, at best, seem to have a barely detectable association with the common cold.

Summary

A panel of Manhattan residents was queried weekly for the occurrence of acute respiratory symptoms. Utilizing stepwise multiple regressions, the incidence and prevalence rates of "common colds" were related to the set of environmental variables measured, both pollution and weather.

The regressions obtained differed markedly by season of the year and by whether incidence or prevalence of common cold was being examined. The meteorologic variables appeared to be more related to the common cold rates than the pollutant variables, although both entered the regressions. A high multiple correlation coefficient (.93) was found for the regression of common cold incidence lagged 72 hours on a set of five of the environmental variables for spring, 1964. Additional examination of the data from this season failed to "explain" the reasons underlying this finding.

BIBLIOGRAPHY

Brandt, C. D.; Wassermann, F. E.; and Fox, J. P. The Virus Watch Program IV. Recovery and Comparison of Two Serological Varieties of Adenovirus Type 5 (31530). Proc. Soc. Exp. Biol. & Med. 123:513, 1966.

Carne, S. J. Study of the Effect of Air Pollution Upon Respiratory Diseases in London in the Winters of 1962-63 and 1963-64. In Proc. (Pt. I). Internat. Clean Air Cong., London, Oct., 1966. London: National Society for Clean Air, 1966, pp. 257-266.

Cassell, E., et al. Air Pollution, Weather and Illness in Children and Adults in a New York Population. Presented at 9th Air Pollution Medical Research Conference, Denver, Colo., 1968.

Ipsen, J. Episodic Morbidity and Mortality in Relation to Air Pollution. Presentation at New York Academy of Science, 1967.

Lidwell, O. M.; Morgan, R. W.; and Williams, E. O. The Epidemiology of the Common Cold IV. The Effect of the Weather. J. Hyg. 63:427-439, 1965.

Lowry, W. Weather and Life. Oregon State University presentation, Corvallis, Ore., 1967.

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

McCarroll, J., et al. Health and the Urban Environment: Health Profiles Versus Environmental Pollutants. A.J.P.H. 56:266-275, 1966.

Widelock, D.; Schaeffer, M.; and Millian, S. J. Surveillance of Infectious Disease by Serologic Methods. I. Asian Influenza, 1957-64. A.J.P.H. 55:578-586, 1965.

Sterling, T. D., et al. Urban Morbidity and Air Pollution. Arch. Environ. Health 13:158-170, 1968.

The authors are affiliated with the Division of Environmental Health, Department of Preventive Medicine, University of Washington School of Medicine, Seattle, Wash. 98105

This work was supported by research contract No. CPA-122-69-3, National Center for Air Pollution Control, U. S. Public Health Service, and the Health Research Council of the City of New York (Contract U-1155).

The paper was presented before the Epidemiology Section of the American Public Health Association at the Ninety-Sixth Annual Meeting in Detroit, Mich., November 11, 1968.