J. Hyg., Camb. (1965), **63**, 427
Printed in Great Britain

The epidemiology of the common cold
IV. The effect of weather

O. M. LIDWELL
Central Public Health Laboratory, Colindale, London, N.W. 9
R. W. MORGAN
Unit of Biometry, 7 *Keble Road, Oxford*
AND
R. E. O. WILLIAMS
The Wright-Fleming Institute of Microbiology, London, W. 2

(*Received* 27 *March* 1965)

INTRODUCTION

It is a challenging fact that the most obvious feature of the epidemiology of the common cold, the seasonal variation in incidence, remains without any satisfactory explanation. Many attempts have been made to explain why these infections are so much more frequent in the winter months of the year in temperate climates. A general discussion of these arguments has been given recently by Andrewes (1964) but, as the subtitle of his review—'We do not yet understand how seasonal and other factors affect the incidence of colds and influenza'—shows, none of them is convincing. One difficulty in establishing any relation between the seasonal variation of an infection and climatic factors is that these factors, together with many aspects of human behaviour dependent on them, are all highly intercorrelated. It is, therefore, necessary to employ relatively complex methods of analysis in exploring such relationships.

We have accumulated a considerable series of records of the incidence of the common cold, over a period of 6 years in a group of offices in Newcastle upon Tyne, a total of over 2000 person-years of observations, and over a period of 4 years in offices in London, nearly 3000 further person-years (Lidwell & Williams, 1961*a*, *b*; Kingston, Lidwell & Williams, 1962). The availability of electronic computer methods now makes it possible to apply appropriate statistical procedures to this body of data.

METHODS OF ANALYSIS

In analyses of time series containing a number of intercorrelated variables it is usually helpful to eliminate the dominant trends, e.g. seasonal or yearly trends, and to investigate the association between the residual deviations from fitted trends. The data we are concerned with here, the incidence of colds and the meteorological variables, might show yearly differences, i.e. a given year might be consistently wetter or sunnier than usual, and they certainly exhibit seasonal variation, i.e. there are more colds in January than in May. By eliminating these trends we mean deducing a value for the variable in question, for example temperature, which can be regarded as the expected value for the particular day concerned allowing for the date in the year and the character of that year as warmer or cooler than usual in the set of data involved. These expected values can be derived from the daily records in many ways, in particular we might take the average values for each calendar date in the year and adjust these by adding the difference of the year's average from the overall average or we might fit a smoothed curve, using some suitable algebraic function, to the data. We adopted the second of these methods. Having arrived at the expected values in one way or other the differences between the actual values observed on a given day and the expected values can be obtained and the inter-relationships among these differences explored, e.g. are more colds than expected for the time of year recorded when the temperature is below expectation for that date?

This form of procedure is discussed by Quenouille (1952) and has been applied by Spicer (1959) to examine the relationship between meteorological factors and the incidence of poliomyelitis in England and Wales. An interesting feature of Spicer's analysis is that the regressions found for poliomyelitis incidence on temperature and relative humidity, based on the residuals of these quantities after the elimination of seasonal and annual trends and including correlations lagged by 1 or 2 months, are such as to predict remarkably closely the actual monthly incidence of the disease when applied to the raw weather data.

It appeared then that it might be useful to try a similar approach for exploring the relationship between the incidence of the common cold and weather. We had available daily records of the incidence of the common cold obtained in the offices by the methods described in the earlier papers in this series (Lidwell & Williams, 1961*a*, *b*; Kingston *et al.* 1962). These covered the 9 months from September to May inclusive for the 6 years 1951–57 at the offices at Newcastle upon Tyne and for the 4 years 1951–55 at the London offices.

The weather data were taken from the records of the nearest meteorological stations and included values for the month of August so that it was possible to compute correlations of disease incidence on the weather variables up to a lag of 30 days. The seasonal weather trends appeared to be adequately represented by fitting simple sine curves, one for London and one for Newcastle, to the averaged figures for each one-third of a month (10 days $\pm$ 1 day) for each of the variables recorded. There did not appear to be any significant systematic differences between the years. It was not possible to fit the data for the incidence of the common cold, averaged similarly for each one-third of a month, to such a simple sine curve over the whole year owing to the occurrence of a substantial peak in the incidence in the early autumn. A sine curve could, however, reasonably be fitted to the data for the months November to May inclusive. This was done and the autumn peak was fitted to a normal distribution superimposed on the calculated curve for November to May extrapolated back through October and September. The constants of the fitted curves are given in Table 1, which also shows the meteorological variables studied. The average seasonal variation in the cold incidence over the year is shown in Fig. 1. The differences between these smoothed curves and the actual day-to-day observed values were then tabulated to be used in the statistical analysis.

In addition we attempted to make use of an index of 'spreading' for the common cold in the form

$$r_i = \frac{3n_i}{n_{i-1}+n_{i-2}+n_{i-3}}$$

where n_i is the cold incidence observed on day i, i.e. the index is a way of representing the rate of increase or decrease of the disease in the community. In order to avoid infinite values and to produce a symmetrical function the quantity $r^* = (\tan^{-1} r)/45$ was computed. r^* is then unity if the incidence of new colds remains

Table 1. *Averaged values of cold incidence and weather data fitted to smooth curves*

Variable	Mean value	Swing ±	Trough or peak date
Newcastle 1951–57			
*Mean day temperature (° F.)	49·4	10·3	30 Jan., T
Day max.–night min. (° F.)	8·5	1·9	28 Dec., T
Water-vapour pressure, 9 a.m. (mb.)	9·1	3·1	3 Feb., T
Relative humidity, 9 a.m. (%)	79·3	6·8	12 Dec., P
Sunshine, for day (hr.)	3·8	2·7	25 Dec., T
Rainfall, 9 a.m.–9 p.m. (mm.)	0·90	0·21	7 Mar., T
Pollution index (smoke and SO_2)	0·17	0·08	6 Jan., P
†Chance of cyclonic or westerly weather	0·37	0·10	1 Apr., T
†Chance of anticyclonic or easterly weather	0·23	0·08	13 Apr., P
New colds/1000 at risk/day	6·9	3·8	22 Dec., P
Autumn peak	(Maximum value 4·5, S.D. 0·36 months)		26 Sept., P
London 1951–55			
*Mean day temperature (° F.)	52·4	13·8	21 Jan., T
Day max.–night min. (° F.)	13·8	5·0	25 Dec., T
Water-vapour pressure, 9 a.m. (mb.)	10·2	3·7	4 Feb., T
Relative humidity, 9 a.m. (%)	81·3	9·9	9 Dec., P
Sunshine, for day (hr.)	4·4	3·0	25 Dec., T
Rainfall, 9 a.m.–9 p.m. (mm.)	0·89	0·37	10 Mar., T
Pollution index (smoke and SO_2)	0·24	0·18	8 Jan., P
†Chance of cyclonic or westerly weather	0·37	0·10	1 Apr., T
†Chance of anticyclonic or easterly weather	0·23	0·08	13 Apr., P
New colds/1000 at risk/day	5·3	5·5	14 Jan., P
Autumn peak	(Maximum value 14·4, S.D. 0·44 months)		28 Sept., P

* The mean day temperature was taken as ½[maximum temperature (9 a.m.–9 p.m.) + minimum temperature (9 a.m.–9 p.m.)].
† The synoptic weather patterns were assessed for England as a whole.
Weather data fitted to sine curve, $y = a + d \sin x$.
Colds, Nov.–May, fitted to similar curve; end of Aug.–Oct. fitted to normal distribution superimposed on Nov.–May sine curve extrapolated back.

steady, with a minimum value of 0 and a maximum of 2. There were no apparent seasonal or annual trends in r^* so this quantity itself was used in the preliminary analyses, but as it was clearly less closely associated with the weather differences than the actual difference between the numbers of colds expected and observed on a given day it was omitted from the final computations and will not be referred to further in this paper.

https://doi.org/10.1017/S0022172400045319 Published online by Cambridge University Press

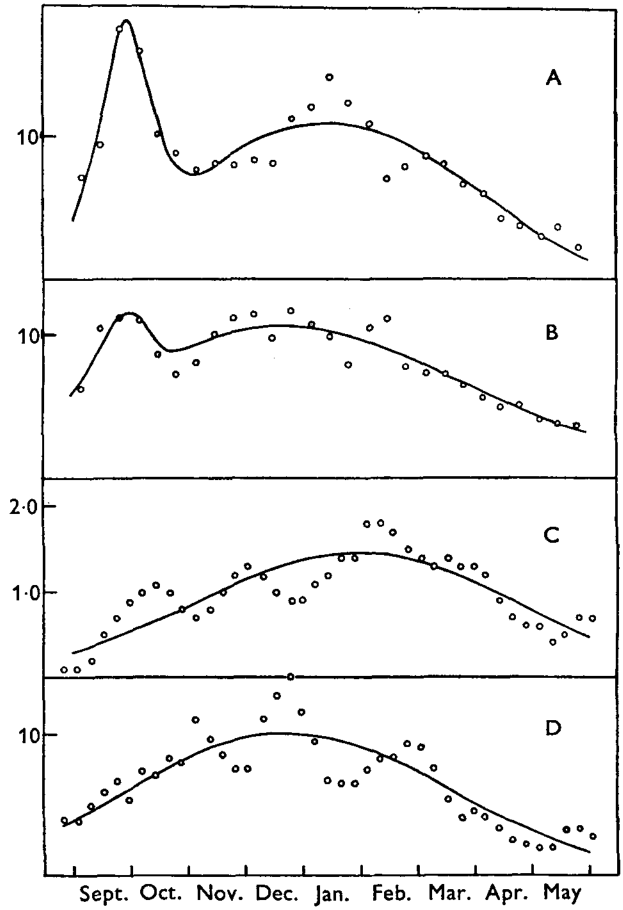

Fig. 1. Seasonal variation in incidence of colds. A. London offices: average one-third of a month (10 days ± 1 day) values for years 1951–55. Full line drawn according to fitted constants given in Table 1. B. Newcastle offices: average one-third of a month (10 days ± 1 day) values for years 1951–57. Full line drawn according to fitted constants given in Table 1. C. Cirencester: all upper respiratory symptoms, average weekly values for years 1954–56. Taken from Hope-Simpson (1958). Full line best fitted sine curve for values over the whole year. D. Chalke valley: average weekly values for colds in families for years 1948 and 1949. Taken from Lidwell & Sommerville (1951). Full line best fitted sine curve for values over the whole year. For A, B and D the ordinate represents the numbers of infections reported per 1000 person-days exposure. For C the figures are given as a fraction of the median weekly incidence.

CORRELATION ANALYSIS ON THE DIFFERENCES

We thus had available for analysis values of the difference between the expected weather and that observed for some 1200 days of observation spread over 4 years in London and for more than 1800 days over 6 years in Newcastle upon Tyne. In addition, we had values of the difference between the expected and recorded numbers of colds for about 1100 days in the London offices and for more than 1600 days in the Newcastle offices.

https://doi.org/10.1017/S0022172400045319 Published online by Cambridge University Press

The analysis of the differences was carried out at the Unit of Biometry, Oxford, using an Elliott 803 computer. This analysis fell into two parts, a correlation study, described in this section, and a regression analysis, which is described in the next section.

A correlation coefficient is a measure of the degree of association of two series of numbers. The methods of calculating correlation coefficients for time series are explained in Quenouille (1952), chapter 11, and several of the methods described in that chapter were used in the analysis. First, all eleven series of differences were tested for serial correlation and the effects eliminated in subsequent analyses by using partial correlation coefficients.

Initially the analysis was confined to two portions of the data, Newcastle 1951/2 and London 1951/2. Because of the irregular occurrence of colds in autumn each of these was split up into three periods of approximately equal length, August–October, November–January and February–May. For each period, each of the nine weather variable differences was compared with the expected–observed cold difference. Thirty correlation coefficients were calculated every time, on the weather for the day and each of the 29 preceding days.

Thus for any period we had a measure of the degree of association between colds and the weather on the same day, on any of the preceding 29 days and on the following 5 days. The lags were chosen to cover such a wide period because of the large difference in phase of the variables. An examination of the results revealed that four of the variables were clearly not important factors in relation to the incidence of colds. These were sunshine, pollution, day maximum minus night minimum temperature and anticyclonic weather. The analysis was then repeated including three additional portions of the data, Newcastle 1952/3, London 1952/3 and London 1953/4, omitting the above four variables. The extra results allowed us to eliminate relative humidity, rainfall and cyclonic weather, leaving mean day temperature and vapour pressure, both having highly significant association with colds.

Temperature and vapour pressure are highly correlated variables and this is true even when the main trend has been removed, i.e. there is still a high degree of correlation between the values of the differences between the observed and expected values of these two quantities on each day. The analysis for the above-mentioned five portions of the data was therefore repeated, using partial correlation to estimate the degree of association between colds and temperature and between colds and water-vapour pressure each independently of the other. These results showed clearly that the main association was related to temperature. Since almost all the significant contribution to the correlation was derived from the data for the periods November–January and February–May the final analysis for the 6 years 1951–57 at Newcastle and the 4 years 1951–55 in London made use of the colds reported between the beginning of November and the end of May only.

Owing to limitations in computer storage the serial correlations were not eliminated in these calculations. This will tend to depress the peak values in the series and will also exert a general smoothing effect. The values obtained, however, confirm the findings of the analysis based on the years 1951–53 at Newcastle and 1951–54 in London, and the two together show that similar effects are present in both cities over the whole period of the study. Table 2 gives the values of the first-order and of the partial correlation coefficients for the whole period. Smoothed curves derived from these are given in Fig. 2. On the same figure are also plotted curves for the

Table 2. *Correlation coefficients*

Lag (days)	Colds and temperature: Vapour pressure not eliminated	Colds and temperature: Vapour pressure eliminated	Colds and vapour pressure: Temp. not eliminated	Colds and vapour pressure: Temp. eliminated
−7	−0·0251	−0·0058	−0·0339	−0·0235
−6	−0·0425	0·0138	−0·0883	−0·0787
−5	−0·0646	−0·0286	−0·0695	−0·0384
−4	−0·0894	−0·0261	−0·1144	−0·0763
−3	−0·0989	−0·0402	−0·1118	−0·0660
−2	−0·0810	−0·0342	−0·0900	−0·0521
−1	−0·0842	−0·0359	−0·0931	−0·0536
0	−0·1087	−0·0401	−0·1286	−0·0798
1	−0·1397	−0·0774	−0·1308	−0·0596
2	−0·1616	−0·0969	−0·1419	−0·0575
3	−0·1468	−0·0861	−0·1313	−0·0551
4	−0·1573	−0·0936	−0·1391	−0·0573
5	−0·1327	−0·0668	−0·1337	−0·0688
6	−0·1219	−0·0673	−0·1146	−0·0528
7	−0·1340	−0·0813	−0·1164	−0·0463
8	−0·1341	−0·0778	−0·1214	−0·0527
9	−0·1151	−0·0645	−0·1074	−0·0494
10	−0·1083	−0·0673	−0·0921	−0·0354
11	−0·0874	−0·0468	−0·0844	−0·0409
12	−0·0697	−0·0275	−0·0806	−0·0490
13	−0·0740	−0·0494	−0·0579	−0·0176
14	−0·0411	−0·0109	−0·0547	−0·0377
15	−0·0542	−0·0365	−0·0420	−0·0124
16	−0·0468	−0·0360	−0·0301	−0·0031
17	−0·0185	−0·0181	−0·0665	0·0051
18	−0·0101	0·0035	−0·0221	−0·0200

The lag is given as the number of days by which the weather correlated antedated the reporting of the colds.

The 95 % confidence limits for the individual values in this table are approximately ± 0·05.

distribution of incubation times following nasal inoculations with common cold viruses and the distribution of serial intervals between colds in families. The correlations with temperature (Fig. 2A) rise to a peak value for a time interval of about 3 days after the weather difference. This is true both of the first-order coefficients and of the partial coefficients from which the effect of water-vapour pressure has been eliminated. As the coefficients are negative this means that an increased number of colds follow colder weather. The lag of about 3 days is very close to the median incubation period following nasal inoculation, namely 2–4 days, and to the median interval between presumed cross-infection in families, namely 2·8 days.

https://doi.org/10.1017/S0022172400045319 Published online by Cambridge University Press

The absolute values of the correlation coefficients are not large, probably, in part, owing to the considerable amount of random variation in the data. This aspect of the analyses is discussed later in connexion with the results of the regression analysis. The relative vertical scale of the correlation curve depends on the relationship

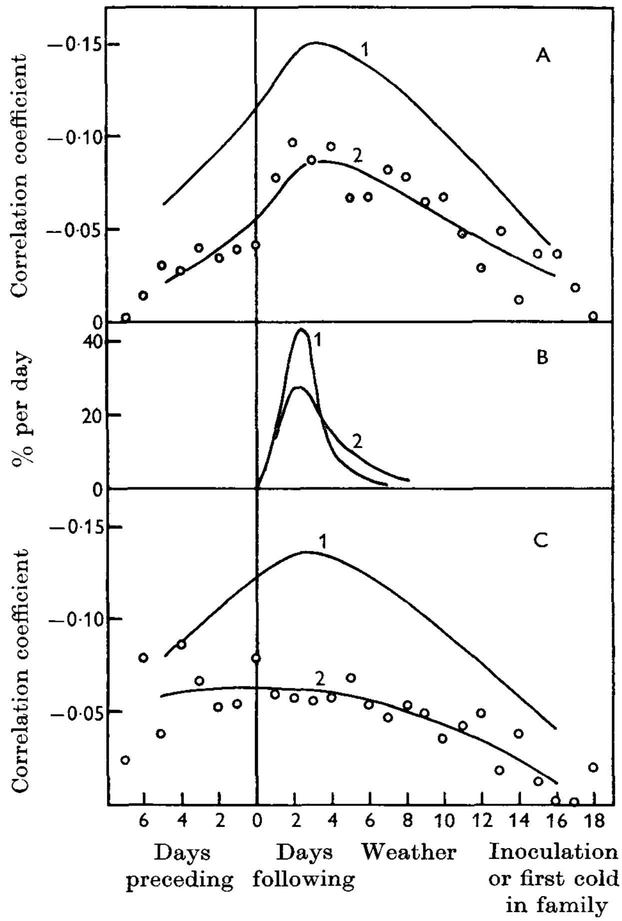

Fig. 2. Correlation between colds and the outside temperature, or the outside water-vapour pressure, on the days preceding and following the first reported day of symptoms. A. Correlation with outside mean day temperature. 1, 1st-order coefficients; 2, partial coefficients with outside water-vapour pressure eliminated. B. 1, The distribution of intervals between nasal inoculation and the appearance of symptoms (Sartwell (1950), based on the data of C. H. Andrewes); 2, the distribution of intervals between the onset of colds presumed transmitted from one member of a family to another (Lidwell & Williams, 1961 *b*). C. Correlation with outside water-vapour pressure. 1, 1st-order coefficients; 2, partial coefficients with outside temperature eliminated. The actual values of the 2nd-order partial correlation coefficients only are shown as points on graphs A 2 and C 2.

assumed between a correlation coefficient and the percentage of infections falling within a given day. Any assumption is more or less arbitrary so that no deductions can be based on comparisons of this aspect of the curves. The wider temporal

https://doi.org/10.1017/S0022172400045319 Published online by Cambridge University Press

spread of the curve of correlation coefficients is, to some extent, accounted for by the effects of serial correlation but this seems likely to be responsible for only a small part of it. Among other possible causes, the consequences of any increased number of infections at a particular time will persist through a series of person-to-person transmissions over, possibly, several generations or many days. The apparently random effects of a multiplicity of unrecorded influences operating on a relatively small number of events are, however, possibly the principal causes of this. The correlations with water-vapour pressure (Fig. 2C) show a different picture. As this variable is highly correlated with mean outdoor temperature, the first-order correlation coefficients are similar to those with temperature. The elimination of the effect of temperature, however, removes the peak at about a 3-day lag. The values of the partial correlation coefficients themselves are somewhat smaller and more irregular than the corresponding cold-temperature correlations. The smoothed values show a tendency to a broad maximum value at or about the same day as the day of reputed onset of colds. These coefficients are also negative, i.e. increased numbers of colds are associated with low water-vapour pressures.

REGRESSION ANALYSIS

The correlation analyses demonstrate the extent to which the incidence of colds, or rather the difference between the observed and expected number of colds, on a given day was associated with the differences between the observed weather and that expected on that day or on the preceding days. It is of interest to explore further the quantitative aspects of this association, e.g. how many more colds are reported for each degree difference between the observed and the expected mean day temperature, and to see how far the actual differences between summer and winter weather can account for the observed difference between the summer and winter incidence of colds. Limitations of storage space made it inconvenient to apply the standard methods of multiple regression analysis and the method of Woolf was employed (Woolf, 1951).

All the analyses were performed on the differences between the observed values on the day in question and the expected value for that day derived from the curves fitted to the averaged values.

A preliminary analysis was carried out for a single year using temperature and water-vapour pressure on the day and on each of the 16 previous days as independent variables with colds reported as the dependent variable. This showed that almost all the variation that could be accounted for by the regression was accounted for by the same day and the 7 preceding days.

With this result and those of the correlation analyses described in the previous section in mind, full analyses were carried out for the 4 years in the London offices and the 6 years at Newcastle omitting the autumn period, namely the period up to the end of October, in each year. The analyses each included forty-one variables; one cold variable and the value on the day of reporting and on each of the preceding 7 days, of mean day temperature, water-vapour pressure, relative humidity, cyclonic weather and atmospheric pollution. Temperature and vapour pressure were included as the variables showing the highest correlation with colds. Relative humidity was also included as being closely related to these. Air pollution was added in view of the considerable current interest in its relation to respiratory disease and finally cyclonic weather as a check since it had shown negligible correlation in the previous analyses.

The effects of relative humidity and cyclonic weather were not significant. A significant positive regression was found, both for London and for Newcastle, with values of atmospheric pollution on the same day only, although closer examination showed that this almost entirely derived from the data recorded for the year 1952/53 in both London and Newcastle. The regressions with mean day

Table 3. *Predicted seasonal swings derived from regression coefficients (reported colds/1000 person-days)*

Variables included		Newcastle, 1951–57	London, 1951–55
Mean day temperature on day	2 only	2·26	1·64
Mean day temperature on day	0	1·46	1·29
	0 and 1	2·02	1·60
	0–2	2·54	1·93
	0–3	2·81	2·02
	0–4	3·05	2·52
	0–5	3·14	2·67
	0–6	3·25	2·86
	0–7	3·53	3·10
Including also water-vapour pressure on day	0	3·71	2·94
	0 and 1	3·74	2·90
	0–2	3·74	2·87
	0–3	3·75	2·93
	0–4	3·76	2·84
	0–5	3·78	2·87
	0–6	3·78	2·98
	0–7	3·78	3·11
Including also atmospheric pollution on day	0	4·11	3·26

The 95 % confidence limits for the figures in the final row of the table are approximately $\pm 0{\cdot}7$ for Newcastle and $\pm 1{\cdot}0$ for London.

temperature and water-vapour pressure were negative and highly significant in both sets of data. As non-significant results were obtained with some of the variables this could not be attributed to the large number of degrees of freedom involved. The regression accounted for only about 10 % of the variance. The limited size of the populations studied, about 675 in London and 350 in Newcastle, with a daily average of about five colds reported in London and less than three in Newcastle, leads of itself to a substantial variance equal to at least one half of the gross variance. The variance accounted for by the regression represents therefore at least 20–25 % of the variance potentially explicable in terms of the independent variables.

Using the regression coefficients obtained in these analyses and applying to them the range of variation of the smoothed meteorological variables as given in Table 1, a predicted swing, or half-range of variation, for the seasonal variation in cold

https://doi.org/10.1017/S0022172400045319 Published online by Cambridge University Press

incidence was obtained. The figures are given in Table 3. With the method of analysis employed it is not possible to distinguish the contributions of the several variables independently. Table 3, therefore, serves only to demonstrate that the total seasonal effect accounted for by the regression appears to have reached a stabilized value with that number of variables which have been included in the computations.

For the Newcastle data the seasonal swing predicted in this way, $\pm 4 \cdot 11$, slightly exceeds that of the curve fitted to the averaged weather data, $\pm 3 \cdot 8$ (Table 1). For the London data the predicted value of $\pm 3 \cdot 26$ is appreciably less than that of the weather data curve, $\pm 5 \cdot 5$, but as this curve actually leads to negative values of incidence for the summer minimum the indicated swing of 5·5/1000 person-days would seem to be too large in any case. Taking the two sets of data together the average predicted swing is of the order of 80 % of that of the fitted curves, which represents very reasonable agreement, probably within the statistical errors of the calculations.

DISCUSSION

The analysis of time series is rarely straightforward and when, as in this case, we have no clear idea of the underlying model it is impossible to put the validity of the method beyond question. It is generally necessary to remove the trends from such series in order to minimize the association due to trends common to the several series. There is, however, no way of deciding what the real trends are so that the method of elimination used is a matter of judgement and once the main component of trend is removed it is impossible to distinguish between the remaining trend and serial correlation in the series. All we can say in this case is that the methods of trend elimination used leave residuals having no obvious trends.

Spicer (1959), in his discussion on the poliomyelitis analysis, had two main statistical worries, serial correlation which we have been able to allow for, and the fact that the strength of the relation between the meteorological data and the poliomyelitis incidence varied from year to year, even to the extent that the correlation in one year was negative. We have found some variation in the magnitude of the correlations from year to year but to a much smaller extent; the correlations are identical in sign for each year and, despite the variation, suggest a stable underlying relationship.

There is obviously a strong negative association between the numbers of colds reported as starting on a given day and the mean outdoor day temperature about that time. Our analyses show that this is present not only between the actual values of these quantities but also between their differences from their expected values for that particular day in the year and that this association with temperature is independent of the intercorrelation of temperature with the other meteorological variables studied, i.e. water-vapour pressure, relative humidity, the difference between the maximum day and the minimum night temperature, rain, hours of sunlight, atmospheric pollution, cyclonic or anticyclonic weather. Of these only water-vapour pressure shows a small but perhaps significant independent association with colds. The correlation reaches a maximum value between the colds caught and the temperature difference, i.e. its fall below the expected value for the time of year, about 3 days before. This time interval is so close to the median interval between inoculation and the development of symptoms as to suggest that some consequence of lowered external temperature exerts a direct influence on the transmission of the infection. This association is independent of any effect of water-vapour pressure and does not, therefore, support the suggestion that the seasonal variation in the numbers of colds is a consequence of change in indoor humidity which might exert drying effects on the mucous membranes of the upper respiratory tract.

If there is any independent association between colds and reduced water-vapour content of the air then it is probably maximal for water-vapour content and colds arising on the same day. There is also some evidence of an association between colds and increased atmospheric pollution on the same day. If this is real, the effect was largely confined to one winter although present in both places at that time. These two effects may reflect some exacerbation of symptoms which influences the day on which the disease becomes apparent.

The regression analysis shows that these associations are sufficient to account quantitatively for most, if not all, of the difference between the summer and winter incidence of colds. This analysis was carried out using the differences of the observed values, both of the numbers of reported colds and of the meteorological variables, from those expected for the time of year. When, however, the regression coefficients obtained in this way are applied to the actual range of temperature, humidity and atmospheric pollution values found over the year, then the predicted range for the numbers of colds reported corresponds remarkably closely to the observed seasonal variation.

The exception to this is the autumn peak in cold incidence. This was very marked in the results from London and moderate, but unmistakable, at Newcastle. Figure 1 shows the average seasonal course in these two places together with data from Cirencester and the Chalke valley, Salisbury. There is little evidence of any such autumn peak in the Cirencester records and none at all from the Chalke valley. The records from both these places, however, show apparent cyclic variations, with a period of about 6–8 weeks, superimposed on the seasonal movement. If real, these may indicate small distinct epidemics each tending to die out with exhaustion of the limited number of susceptible individuals in these small populations. It is interesting and, possibly, of significance that the magnitude of this autumn peak follows the size of the community concerned, which ranges from around 10^7 for London down to a few hundred only in the Chalke valley. No weather variable has a distribution remotely resembling this autumn peak in colds and this, together with its apparent association with community size, suggests strongly that it is related to the immunity state of the population at this time of the year.

The proportion of variation absorbed by the regression analysis is limited and might lead to suspicion that the effects revealed by the analysis are, themselves, of limited importance. Since, however, such a large proportion of the variance is certainly irremovable it may be unrealistic to expect to be able to absorb a much greater fraction from data of this kind.

The striking success of the correlation analysis in reproducing the general shape of the incubation curve and of the regression analysis in predicting the magnitude of the seasonal variation in the incidence of the disease is, however, the main justification for the whole procedure. At no point where we might doubt the validity of the analysis should we expect it to produce effects of this kind unless they were really due to strong association between the variables. The nature and force of this association are probably dependent on the community concerned so that it should occasion no surprise if studies in other climatic regions or in populations with different social habits should lead to results quantitatively or qualitatively different.

We referred in the introduction to the many, but unconvincing, explanations adduced for the seasonal variation in upper-respiratory infections in temperate climates. Our analysis does not, of itself, solve this problem. In so far as the indications of the correlation analysis are a reliable guide we have to look for some effect of low outdoor temperature which promotes transmission of the virus from person to person, or the development of overt disease. A number of possible ways in which cold weather might induce colds are discussed by Andrewes (1964). It has often been suggested that changes in room ventilation consequent upon the seasonal climatic changes might be responsible for a more widespread dissemination of the disease in the colder months of the year when windows are more usually closed. Experimental studies of the effect of artificially increased ventilation (Kingston *et al.* 1962) suggest that this could not produce the observed effects.

The virus might survive better in cold environments, but it seems very unlikely that outdoor survival plays any part in the transmission of the infection. Indoor temperatures on the other hand remain relatively constant.

The low indoor humidities found in cold weather are a consequence of the low absolute outdoor humidity associated with cold weather. This analysis has shown that it is the low outdoor temperature, independent of the humidity, which is associated with the increased number of winter colds. This contradicts any arguments based on virus survival in relation to indoor humidity or on a postulated damaging effect, due to drying, on the mucous membranes, predisposing to the initiation of infection when the indoor humidity falls in cold weather. It is possible that exposure to cold outdoor conditions produces physiological changes, in the membranes of the respiratory tract or elsewhere, which promote the transmission of the infection. This might take place either through increased dispersion of infected secretion from an infected individual, caused by increased volume of secretion, or increased tendency to sneezing, etc., or through some increase in susceptibility of the relevant sites so that infection is more easily initiated or perceived.

SUMMARY

An investigation has been made of the association between weather and the numbers of colds reported on a given day. The seasonal trends were eliminated by working with the differences between the observed values on any day and the expected values derived from smooth curves fitted to the averages for the time of year.

Examination of nine weather variables for the day on which the colds were reported and for each of the 29 preceding days showed that only two, mean day temperature and water-vapour pressure at 9 a.m., were significantly correlated with the numbers of colds. Partial correlation studies showed that the strongest association was with lowered mean day temperature between 2 and 4 days before the reported onset of symptoms.

Regression analysis demonstrated that the magnitudes of the associations were sufficient to account for the greater part of the seasonal variation in the incidence of the common cold in both London and Newcastle. A small effect of atmospheric pollution appeared in this analysis.

These results suggest that some effect of low outdoor temperature promotes transmission of the virus or the development of disease.

Our thanks are due to the Director General of the Meteorological Office for access to the weather records and for supplying the daily synoptic weather patterns. We should also like to repeat in this, the final paper of the series, our thanks to the Ministry of Pensions and National Insurance, to the Shell Petroleum Company and to those members of their staffs who collaborated with us in this investigation.

REFERENCES

ANDREWES, C. H. (1964). The complex epidemiology of respiratory virus infections. *Science*, **146**, 1274.

HOPE-SIMPSON, R. E. (1958). The epidemiology of non-infectious diseases. *R. Soc. Hlth J.* **78**, 593.

KINGSTON, D., LIDWELL, O. M. & WILLIAMS, R. E. O. (1962). The epidemiology of the common cold. III. The effect of ventilation, air disinfection and room size. *J. Hyg., Camb.*, **60**, 341.

LIDWELL, O. M. & SOMMERVILLE, T. (1951). Observations on the incidence and distribution of the common cold in a rural community during 1948 and 1949. *J. Hyg., Camb.*, **49**, 365.

LIDWELL, O. M. & WILLIAMS, R. E. O. (1961*a*). The epidemiology of the common cold. I. *J. Hyg., Camb.*, **59**, 309.

LIDWELL, O. M. & WILLIAMS, R. E. O. (1961*b*). The epidemiology of the common cold. II. Cross-infection and immunity. *J. Hyg., Camb.*, **59**, 321.

QUENOUILLE, M. H. (1952). *Associated Measurements*, chapter 11. London: Butterworth.

SARTWELL, P. E. (1950). The distribution of incubation periods of infectious diseases. *Am. J. Hyg.* **51**, 310.

SPICER, C. C. (1959). Influences of some meteorological factors in the incidence of poliomyelitis. *Br. J. prev. soc. Med.* **13**, 139.

WOOLF, B. (1951). Computation and interpretation of multiple regressions. *Jl R. statist. Soc.*, B, **13**, 100.