VOL. 49, No. 4 DECEMBER 1951

OBSERVATIONS ON THE INCIDENCE AND DISTRIBUTION OF THE COMMON COLD IN A RURAL COMMUNITY DURING 1948 AND 1949

By O. M. LIDWELL AND T. SOMMERVILLE

From the Medical Research Council's Common Cold and Environmental Hygiene Research Units, Harvard Hospital, Salisbury

(With 1 Figure in the Text)

INTRODUCTION

The development of means for the prevention and cure of the common cold is greatly hampered by the lack of precise information on the natural history of the disease. Field observations and experiments with human volunteers have strongly suggested that the common cold is a communicable disease caused by a virus or group of viruses, but there are many unsolved problems concerning the sources and modes of spread of the infection.

Although experiment probably offers the greatest hope of a rapid advance in knowledge, epidemiological studies may make a useful contribution. Observations have therefore been made of the incidence and distribution of the common cold, or rather of the group of upper respiratory infections commonly so called, in a small rural community during the years 1948 and 1949. The village of Bowerchalke lies in Wiltshire on the upper reaches of the river Ebble about 10 miles west of Salisbury. The scattered population of about 200 is mainly composed of agricultural labourers and their families, most of them living in detached three- or four-roomed cottages, often with the typical Wiltshire thatched roof. During the period under discussion the economic condition of the village was good and there were no signs of poverty, malnutrition or overcrowding. The population changed little during this time. The elementary school, three churches and a hall for social purposes provide opportunity for frequent collections of persons within the village, and good bus services to both Salisbury and Shaftesbury afford regular contact with these larger centres of population. Many of the villagers visit Salisbury once a week.

Additional observations were made concurrently in the village schools of Coombe Bissett at the eastern end of the Chalke valley, Bishopstone which lies in the same valley between Coombe Bissett and Bowerchalke, Bowerchalke and Ebbesbourne Wake at the western head of the valley. These villages are all on the Salisbury to Shaftesbury bus route at intervals of 2–3 miles. Each of these four elementary schools is run by a schoolmistress and her assistant. Each has two rooms, one for younger children aged 5–7 years and the other for the older children aged 8–15 years. The average numbers of children attending these schools were 55, 30, 45 and 25 respectively.

Although the experience reported is small it does focus attention on the school as an important source of the common cold and confirms the significance of household transmission of the disease.

METHODS

Weekly house-to-house visits were made by two observers, each covering about one-half of the village. The information was usually obtained from the womenfolk who reported on their husbands and children also, since these were often at work or at school when the visits were made. Generally, the same person provided the required information each week. At first there was some natural suspicion, but after a few months the weekly visit of the 'cold doctor' or 'cold matron' was accepted as a part of village life and co-operation was willingly given. In only one household was objection raised and this was omitted from the investigation. The diagnostic criteria were necessarily those generally accepted by the lay public and cannot therefore be rigorously defined. It was, however, repeatedly emphasized that it was the 'runny' nose type of cold which was of interest, and when a cold was reported an inquiry into the symptoms was made and sometimes the patient was seen. Although only a small proportion of the total colds recorded were actually seen the general interrogation as to symptoms and the occasional examination, especially the scrutiny of the results of nose-blowing, helped to maintain some uniformity in diagnosis.

Visits were usually made on a Friday afternoon, and the results of the visit entered at the time in a register. The villagers were listed by name at the left-hand side of the page, the remainder of which was divided into columns each headed by the date of a visit. Four alternative entries were made for each occasion and person: +, indicating symptoms of a cold during the week preceding the visit; −, indicating no symptoms of a cold during the week preceding the visit; 0, signifying that the family had not been seen for three weeks, experience showing that the memory of the reporters was unreliable after more than two weeks; H, signifying absence from the village.

In the elementary schools the diagnosis was made and records kept by the schoolmistress with a random check by the observers. It will be seen later that there is some evidence of a major difference in diagnosis at one of the schools.

Limitations of the data

Lack of uniformity in diagnosis is a major source of uncertainty in the assessment of these data. The absence of any specific sign in the disease and our inability to recognize the causal agent make a precise diagnosis impossible. The 'nose-blow' test has been found valuable in the examination of human volunteers who have been experimentally infected with the common cold, and for that reason stress was laid on the production of nasal discharge when blowing the nose into a handkerchief. Inevitably we are dealing in this inquiry with a group of related upper respiratory infections, not all of which will necessarily follow the same epidemiological pattern. An afebrile illness accompanied by a runny nose was, however, the commonest upper respiratory syndrome observed.

In a normal community the method of domiciliary visiting appears to us to be far more reliable than any system of posted returns filled in by the head of the household. It brings the observer into intimate contact with the reporter, thereby considerably assisting uniformity in diagnosis as well as bringing to light many otherwise overlooked details. Attempts to obtain information about the day of onset and the sequence of infection in the various members of a household did not produce reliable data. The reporters were often uncertain about dates although quite confident about the occurrence of colds in members of the household.

ANALYSIS

Particular attention has been paid in the analysis to the extent to which different classes of persons introduced the infection into their households and the pattern of the resulting cross-infection within the households. The nature of the data and the lack of information about times of onset of infection in different members of a household have made a statistical treatment the only possible one and have required a somewhat elaborate mathematical treatment. In order to keep the general line of argument clear the course of the analysis and the general results obtained will be described first. The detailed argument and tables then follow.

Incidence rates

Incidence rates in terms of numbers of new colds per 100 person-weeks of exposure have been computed for the years 1948 and 1949 separately for adult males, adult females, schoolchildren and infants under school age. These groups

Table 1. *Incidence rates in various groups. New colds per 100 person-weeks exposure to risk*

Household type	Male adults (MA)	Female adults (FA)	Schoolchildren (B)	Infants (C)
Adults only (A)	1·90 (0·18)	2·55 (0·13)	—	—
Adults with schoolchildren only (AB)	4·70 (0·62)	4·81 (0·77)	7·10 (0·12)	—
Adults with infants only (AC)	2·34 (0·41)	3·37 (0·10)	—	4·86 (0·19)
Adults with schoolchildren and infants (ABC)	3·97 (0·62)	6·17 (1·47)	6·85 (1·54)	10·97 (1·49)

In brackets, the standard error of the rate, based on four estimates.

were further subdivided according to the size and composition of the household. The eight household groups employed were: *households of adults only* (*a*) with one or two members, (*b*) with three or four members; *households of adults and schoolchildren* (*a*) with two to four members, (*b*) with five to eight members; *households of adults and infants* (*a*) with three or four members, (*b*) with five or six members; *households of adults, schoolchildren and infants* (*a*) with four or five members, (*b*) with seven to nine members. These groups were determined by the desire to separate larger from smaller households and the necessity of aggregating the limited numbers of households actually occurring into reasonably sized groups. These incidence rates are summarized in Tables 1 and 2.

The experience and the effect of schoolchildren is most noticeable. Their own attack rate is high, about three times that of an adult living in a household without children, and their presence appears approximately to double the attack rate among the adults who live with them. The effect of infants below school age on the experience of adults is comparatively small, but their susceptibility is high.

Female adults appear on the average to acquire nearly 50 % more colds than males; owing to difficulties with small numbers this sex difference was not explored among the children included in these household observations. While reporting of colds is perhaps likely to be more complete for females, the sex difference closely resembles that observed in inoculation experiments on volunteers at Harvard Hospital (unpublished data). The ratio FA/MA for these experimental inoculations was 1·32, for about 100 infections in each sex with an overall average infection rate of about 50 %.

In addition 3-weekly moving averages have been calculated for a few groups. They show the expected seasonal variation with a maximum incidence around the turn of the year and a minimum at or about midsummer. There is little evidence of any significant difference between the two years. The week-by-week experience of adults living in households with schoolchildren is similar to that of the schoolchildren themselves. In view of the apparent nature of the intra-household cross-infection process, discussed later, this would be expected. The experience of adults living apart from schoolchildren, while following the same general seasonal pattern, shows more difference from the group of schoolchildren. There is some indication of a periodicity of about a month in the weeks of peak incidence. This can be seen in all the groups of Fig. 1, and particularly in the experience of schoolchildren and of the adults living with them.

Table 2. *Ratio of incidences in comparable groups*

Ratio	Mean value	Standard error of mean	No. of estimates of ratio
Female adults to male adults (*FA/MA*)	1·44	0·13	16
Schoolchildren to male adults (*B/MA*)	1·80	0·16	8
Infants to male adults (*C/MA*)	2·62	0·29	8
Larger households/smaller households	1·06	0·024	24
1949/1948	1·08	0·10	24

The exposures to risk in these tables can be seen from the figures given in Table 5.

The data obtained from the schools have been used to obtain incidence rates in these communities also; the values are given by school term in Table 3, together with the figures for the schoolchildren included in the household survey. The incidence rates follow a similar pattern and lie within a single range of variance except for the recorded rates for school no. 3. The mean incidence rate for this school, which was that attended by the schoolchildren included in the household survey, is about 50 % above that for the others and the seasonal variation is not followed. There appears to be some particular cause, which was not determined, for an excess of reported cases from this school during the middle part of the period.

A higher incidence among girls is found in all the schools, although the ratio of the mean incidence rate among the girls to that for the boys, 1·17 with a standard error of 0·09, is less than that observed for adults in the household survey. We hoped that these observations might show whether 'waves' of common cold infection progressed along the valley, but if there is any such effect it is completely masked by the irregularities of the infection and the frequency of minor epidemic waves in the schools.

https://doi.org/10.1017/S0022172400066699 Published online by Cambridge University Press

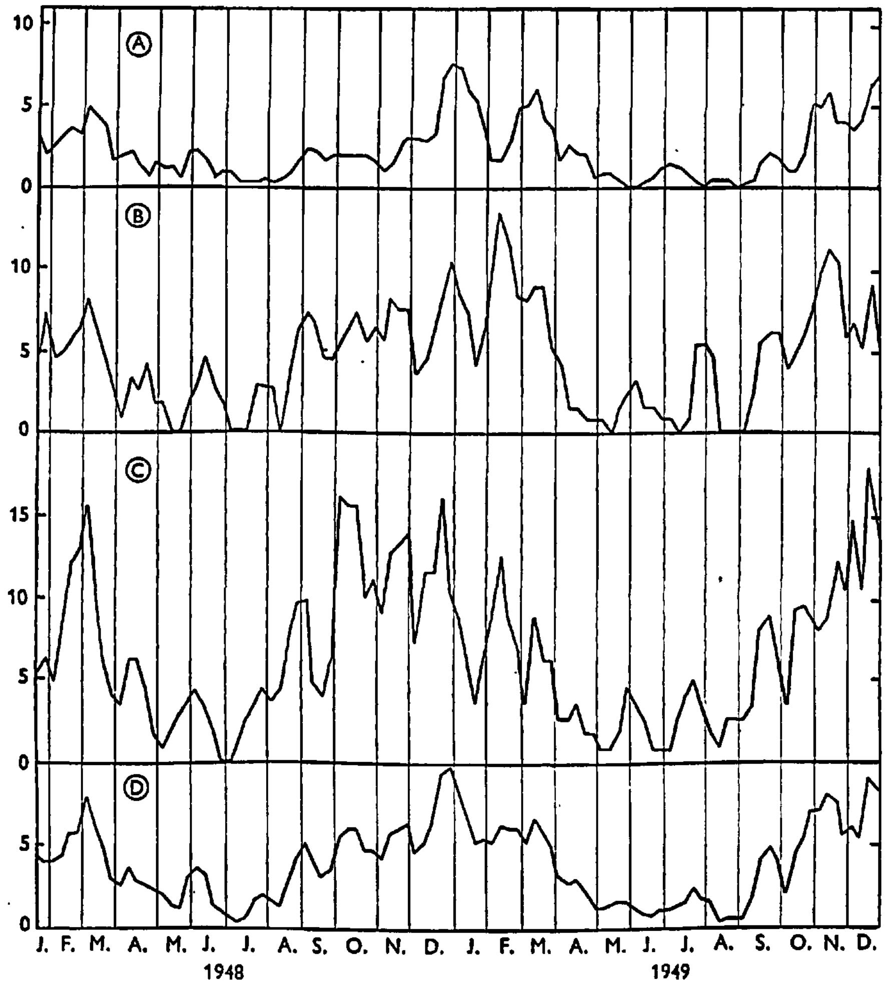

Fig. 1.

Fig. 1. Incidence rates through 1948 and 1949. Three-weekly moving averages (infections per 100 person-weeks exposure) for four groups: A, adults living in households of adults only; B, adults living in households with schoolchildren; C, schoolchildren; D, all persons.

Table 3. *Incidence rates among schoolchildren. New colds per 100 person-weeks exposure to risk*

School	Approx. no. on roll	1948 Term 1	1948 Term 2	1948 Term 3	1949 Term 1	1949 Term 2	1949 Term 3	All
No. 1 (Coombe Bissett)	55	—	6·9	8·6	11·6	7·9	9·6	8·9
No. 2 (Bishopstone)	30	9·1	2·7	8·7	10·1	7·3	10·0	8·1
No. 3 (Bowerchalke)	45	7·2	8·7	18·7	20·5	12·7	11·3	13·2
No. 4 (Ebbesbourne Wake)	25	9·1	6·4	11·7	8·4	5·7	10·0	8·5
Schoolchildren included in the household survey (*B*)	35	10·0	2·9	11·2	6·6	2·7	11·0	7·4

https://doi.org/10.1017/S0022172400066699 Published online by Cambridge University Press

Introduction rates and intra-household cross-infection

The infections occur in the households in groups which may comprise only a single infection or may involve any number of the members of the household. We assume that the chance of an individual acquiring an infection outside the household, and hence introducing it into the household, is not affected by the composition of the household in which he or she lives. The rate at which infections would be expected to be introduced into any given type of household can then be expressed in terms of (i) the composition of the household and (ii) the rates at which the different kinds of persons comprising the household acquire infections outside the household. If each household group of infections, or incident, was due to a single infection introduced from outside, then the rate at which such incidents occurred in a given type of household would be simply that at which infections were being introduced into that type of household from outside. In practice a correction is necessary to allow for more than one infection being acquired outside the household by its members during the course of a given incident. Introduction rates for the various classes of persons calculated in this way give a consistent picture.

Table 4. *Summary of class introduction rates and the values of the simple cross-infection risk*

Household type	Adults (*A*)	Schoolchildren (*B*)	Infants (*C*)	Weighted mean of *A*, *B* and *C*
		Introduction rates (per person-week of exposure)		
All types	0·0171	0·0565	0·0303	0·0249
		Cross-infection risk (per person-incident)		
Adults only (*A*)	0·240	—	—	0·240
Adults with schoolchildren only (*AB*)	0·282	0·234	—	0·272
Adults with infants only (*AC*)	0·230	—	0·436	0·290
Adults with schoolchildren and infants (*ABC*)	0·241	0·093	0·602	0·273
All types	0·257	0·167	0·540	0·269

Deduction of the number of infections acquired outside the household from the number of infections actually occurring in the various groups of individuals yields an estimate of the number of cross-infections acquired within the household. For this purpose the number of infections acquired outside the household (called 'Introductions' in Table 5 and elsewhere) is calculated from the introduction rates derived in the manner described above. A cross-infection risk is then obtained as the fraction of occasions on which exposure to infection in a household incident resulted in a cross-infection. Table 4 gives a summary of the introduction rates and cross-infection risks obtained in this way. More detailed figures are given in Table 5. The most notable points in the two tables are, the high introduction rates for schoolchildren, over three times the adult rate, and the high cross-infection risk

Table 5. *Introduction rates and the simple cross-infection risk*

Household type	Class	Household weeks of exposure	Person-weeks of exposure	Incidents	Introductions	Introduction rate (per person-week)	Infections	Cross-infections	Fraction of infections acquired in household	Exposure to cross-infection risk (person-incidents)	Cross-infection risk (per person incident)
$(A)_{1 \text{ and } 2}$	A	2848	5339	98	102	0·0191	118	16	0·136	78	0·205
$(A)_{3 \text{ and } 4}$	A	1287	4404	62	67	0·0152	104	37	0·356	143	0·256
(A)	A	4135	9743	160	169	0·0174	222	53	0·239	221	0·240
$(AB)_{2,3 \text{ and } 4}$	Both A and B	721	2303	70	84	0·0364	128	44	0·344	150	0·294
$(AB)_{5,6,7 \text{ and } 8}$	Both A and B	433	2636	60	83	0·0315	159	76	0·478	292	0·260
AB	Both A and B	1154	4939	130	167	0·0348	287	120	0·418	442	0·272
AB	A only	1154	2719	130	45	(0·0166)	129	84	0·651	288	0·282
AB	B only	1154	2220	130	122	(0·0548)	158	36	0·228	154	0·234
$(AC)_{3 \text{ and } 4}$	Both A and C	733	2518	46	51	0·0203	93	42	0·452	115	0·365
$(AC)_{5 \text{ and } 6}$	Both A and C	208	1087	18	21	0·0193	34	13	0·372	75	0·173
AC	Both A and C	941	3605	64	72	0·0200	127	55	0·433	190	0·290
AC	A only	941	2435	64	39	(0·0160)	70	31	0·443	135	0·230
AC	C only	941	1170	64	33	(0·0283)	57	24	0·421	55	0·436
$(ABC)_{4 \text{ and } 5}$	All classes	377	1535	57	75	0·0488	113	38	0·336	157	0·242
$(ABC)_{7,8 \text{ and } 9}$	All classes	254	1919	41	59	0·0308	131	72	0·550	246	0·293
ABC	All classes	631	3454	98	134	0·0388	244	110	0·451	403	0·273
ABC	A only	631	1237	98	22	(0·0182)	63	41	0·651	170	0·241
ABC	B only	631	1460	98	87	(0·0601)	100	13	0·130	140	0·093
ABC	C only	631	757	98	25	(0·0322)	81	56	0·691	93	0·602

The number and composition of households varied during the course of the experiment. The average number of families in any group and their composition can be deduced from the household and class-exposure columns in this table, remembering that the duration of the experiment recorded here was 103 weeks. The values for the introduction rates given in brackets are in the same ratio as the 'best' values derived in the text; the absolute values are chosen so as to yield the computed total number of introductions for the given household type.

A represents an adult, B a child of school age, C an infant below school age. The groups of letters indicate the composition of the households and the subscript numbers the total number of individuals included in them.

for infants, about twice the apparent rate for adults. The cross-infection risk for schoolchildren appears somewhat lower than for adults. If this is significant it may be that the high risk of acquiring infection outside the household leaves only the less susceptible open to household infection.

Further examination of Table 5 shows that while adults living with other adults appear to acquire only about one-quarter of their infections in the household, when living with schoolchildren they appear to acquire nearly two-thirds of their infections in this way. The schoolchildren themselves appear to acquire only about one-sixth of their infections in the household. Both infants and the adults living with them in households without schoolchildren acquire rather less than half their infections in the household, but infants living with schoolchildren acquire over two-thirds of their infections in the household. The low proportion of cross-infection shown by adults living with adults is partly a reflexion of the smaller size of such households (mean household sizes, adults only, 2·36; adults with schoolchildren only, 4·28; adults with infants under school age only, 3·83; adults with schoolchildren and infants under school age, 5·47). In adult households of three and four members (mean household size, 3·43) rather over one-third of the infections appear to be acquired in the household, a figure which is not significantly different from that of rather under half for adults and infants living in households without schoolchildren.

The distribution of multiple infections in households

A more detailed analysis of the groups of infection in the various households suggests that those who escape the disease on its first appearance in the household have as great a chance of acquiring it from secondary or subsequent cases as from the primary case. The risk of cross-infection appears to be relatively low, about 1/5 from each infected member of the household to whom the individual is exposed. These indications suggest that avoidance of infection on any given exposure is usually due to factors which have a chance variation rather than to a stable high level of immunity. These conclusions, though tentative, are of considerable interest in the epidemiology of the common cold, and in general agree with the observations on colds induced in volunteers by nasal installation of washings, and with the studies of experimental cross-infection (Andrewes, 1949; Andrewes *et al.* 1951).

The variability in individual experience

The numbers of colds experienced by individuals during a given year have been obtained and the distribution tabulated (Table 6) for four classes of persons and four types of household. No abnormalities appear in the distributions, which, however, show a distinctly broader spread than would simple Poisson distributions with the same mean or modal values. This suggests real differences in individual susceptibility, in addition to differences in experience deriving from a random risk of infection.

The experiences of the same individual in the two years have been compared in a further attempt to assess the magnitude of the variation in individual susceptibility. The numbers available for analysis are not very large, but a person experiencing more colds than the median experience of comparably situated persons in one year appears likely to suffer more colds in the other year also. The data are shown in Table 7.

Recurrence of infection

Any tendency to spontaneous recurrence of a cold in an individual after the infection had apparently subsided should be shown by the unduly frequent occurrence of a particular time interval between recorded infections.

Table 6. *The numbers of colds experienced by individuals in a single year*

Household type	Class	No. of colds experienced in one calendar year										
		0	1	2	3	4	5	6	7	8	9	10
Adults only (*A*)	Male adults (*MA*)	26	18	18	2	1	0	0	0	0	0	0
	Female adults (*FA*)	32	27	16	9	3	2	1	0	0	0	0
Adults with schoolchildren only (*AB*)	Male adults (*MA*)	2	8	2	4	4	2	1	0	0	0	0
	Female adults (*FA*)	1	5	5	7	0	1	1	0	0	0	0
	Schoolchildren (*B*)	3	4	3	10	7	8	3	0	0	0	0
Adults with infants only (*AC*)	Male adults (*MA*)	9	6	2	3	1	0	0	0	0	0	0
	Female adults (*FA*)	6	4	6	4	1	1	0	0	0	0	0
	Infants (*C*)	4	2	6	5	1	0	1	1	0	0	0
Adults with schoolchildren and infants (*ABC*)	Male adults (*MA*)	1	2	5	2	1	0	0	0	0	0	0
	Female adults (*FA*)	0	2	2	4	2	0	0	1	0	0	0
	Schoolchildren (*B*)	2	6	3	3	4	2	3	2	0	0	0
	Infants (*C*)	0	0	1	1	4	1	2	0	3	0	1

The table gives the number of individuals having the indicated experience.

Table 7. *The correlation between the numbers of colds experienced by the same individuals in* 1948 *and* 1949

			1948 Less affected 50%	1948 More affected 50%	χ^2	*P*
Adults (*A*)	1949	Less affected 50%	30·8	16·2	9·09	<0·01
		More affected 50%	16·2	30·8		
Schoolchildren (*B*)	1949	Less affected 50%	6·8	4·2	1·23	>0·10
		More affected 50%	4·2	6·8		
Infants (*C*)	1949	Less affected 50%	4·0	2·0	—	0·28
		More affected 50%	2·0	4·0		

The fractional figures in the table are derived by apportioning border-line cases into the four groups. The values of *P* are for a single tailed test and that for infants (*C*) has been derived by the 'exact' expression given by Fisher (1946).

The data for colds in adults living without children have been analysed to examine the possibility. Both the complications due to household cross-infection and the incidence rates themselves are lowest in this group, so that any such recurrence phenomenon would be most apparent in it. Table 8 shows the observed distribution of intervals together with that calculated from the overall incidence rate on the assumption that the infections are randomly distributed in time and that spontaneous recurrence does not occur. It will be seen that there is no evidence of such recurrence.

https://doi.org/10.1017/S0022172400066699 Published online by Cambridge University Press

A similar analysis of the intervals between infection in adults living with schoolchildren is also shown in the table. This shows a significant excess, at the 1/100 level of intervals in the 4–6 weeks range. This excess, which amounts to about 10% of the total number of observed intervals, might be due to re-infection from within the household or to a tendency to a periodicity of this order in the epidemic cycle in the schools. It will be remembered that the weekly incidence rates (Fig. 1) give some indication of a periodicity of about a month. This is especially noticeable in the data for schoolchildren. Whether this period is derived from local factors or whether it is related to an average duration of some form of relative immunity is a question which cannot be answered from the data at present available.

Table 8. *Distribution of intervals between successive infections in the same individual; for adults of two household types*

	Duration of interval (weeks)					
Household type	1	2	3	4–6	7–10	More than 10
Adults only (A)	5 (5)	5 (5)	5 (4)	10 (13)	16 (17)	173 (170)
Adults with schoolchildren only (AB)	3 (6)	2 (5)	7 (5)	25 (13)	11 (15)	67 (71)

The figures in the body of the table are the number of times intervals of the duration indicated at the head of the columns were observed. In brackets are the calculated numbers on the basis of a random distribution of infections in time. In all cases the interval is taken as the number of weekly records of 'no new cold' intervening between two new infections. Grouping the first three intervals together: for household type A, $\chi^2 = 0{\cdot}87$ and for household type AB, $\chi^2 = 13{\cdot}4$. For two degrees freedom these correspond to values of P of 0·5–0·7 and $<0{\cdot}01$ respectively.

DETAILED ANALYSIS

In the argument and in some of the tables the various classes of persons and household types have been represented by capital letters and groups of letters. A = adult, B = child of school age, i.e. over 5 years old, C = infant under school age, i.e. less than 5 years old, M and F are used as prefixes when it is required to indicate male and female individuals separately. Groups of letters are used to indicate the composition of households, with numerical suffixes, when needed, to denote the number of individuals comprising the household. Over the period of observation the average number of persons included was approximately 210, forming on the average 67 households.

(1) *The rates of incidence in various groups (Tables 1 and 2)*

Incidence rates were computed separately for the calendar years 1948 and 1949 for the classes MA, FA, B and C in each of the groups of households types

$$(A)_1+(A)_2,\quad (A)_3+(A)_4,\quad (AB)_2+(AB)_3+(AB)_4,$$
$$(AB)_5+(AB)_6+(AB)_7+(AB)_8,\quad (AC)_3+(AC)_4,$$
$$(AC)_5+(AC)_6,\quad (ABC)_4+(ABC)_5,\quad (ABC)_7+(ABC)_8+(ABC)_9.$$

The grouping of household types was determined by the need for adequate numbers in each group. From these 48 incidence rates mean values for the incidence rates for each class of persons in the household types, A, AB, AC and ABC, have been derived together with standard errors of the mean values based on the variance over the four estimates, namely, those for larger and for smaller families in 1948 and 1949. These variances therefore include the variances in respect of the years 1948, 1949 and of the larger and smaller families, but, as is shown in Table 2, these make only a small contribution.

(2) *The weekly rate of incidence* (*Fig.* 1)

Since the number of infections occurring per week was often too low for a simple weekly average, 3 weekly moving averages were computed for the following groups: (*a*) adults living in households without schoolchildren, (*b*) adults living in households with schoolchildren, (*c*) schoolchildren, (*d*) all persons together.

(3) *Incidence rates as recorded in the schools* (*Table* 3)

These rates were calculated for each school for each term. For comparison the incidence rates recorded for the schoolchildren included in the household survey were also computed for the same periods. While there is some reason to believe that mild colds which were not reported in the household visits were recorded at the schools, especially during the summer months, we know of no reason to account for the major discrepancy between the incidence as reported from school no. 3 and the recorded experience of the schoolchildren obtained from the households. These latter children all attended this particular school and formed over three-quarters of the pupils.

(4) *Introduction rates and the simple cross-infection risk* (*Tables* 4 *and* 5)

In order to discuss introductions into the household, and cross-infection within it, some further definitions are necessary. An incident is defined as a group of infections in one household limited to a period such that there is no week without a new case with the further condition that no individual may count more than once in a single incident. On two occasions this last condition leads to an arbitrary separation of a pair of incidents; on all other occasions an individual's second infection has been regarded as a solitary infection comprising a second incident.

It cannot be assumed that each incident, as defined above, derives from a single introduction into the household; there is a finite chance that two or more infections may be acquired simultaneously outside the household by the individuals who comprise the household. It is, however, possible to make an approximate correction for this chance. If $p=(1-q)$ is the chance of an individual in a household of n members acquiring an infection outside the household in a given week, then the expected number of incidents which will start per week in N such households is given by $Nq^n(1-q^n)$. This formula is exact, if the chance of an individual introducing an infection in a given week is unrelated to his state in the previous week, if the chance, p, is constant for all the individuals over the whole period under examination and if cross-infections within the household do not lead to introductions falling into the same incident which would otherwise belong to separate incidents. None of these conditions can be regarded as strictly fulfilled, but the errors caused by neglecting them do not appear likely to be individually large and the several causes operate in different directions. The above formula has therefore been used to compute the introduction rates and the numbers of introduction given in Table 5.

An attempt has been made to derive introduction rates for each class of person, from those for the household types on the assumption, which seems reasonable, that an individual's chance of acquiring an infection outside his household is independent of the make-up of his household. Taking male and female adults together we have the four following equations from which to derive the introduction rates.

$$9743p_A + 0p_B + 0p_C - 169 = d_1, \quad (1a)$$
$$2719p_A + 2220p_B + 0p_C - 167 = d_2, \quad (1b)$$
$$2435p_A + 0p_B + 1170p_C - 72 = d_3, \quad (1c)$$
$$1237p_A + 1460p_B + 757p_C - 134 = d_4, \quad (1d)$$

where p_A, p_B, p_C are the introduction rates for adults, schoolchildren and infants respectively and d is a residual error. Minimizing Σd^2 we derive the following normal equations:

$$109{\cdot}8p_A + 7{\cdot}840p_B + 3{\cdot}785p_C = 2{\cdot}437, \quad (2a)$$
$$7{\cdot}840p_A + 7{\cdot}060p_B + 1{\cdot}105p_C = 0{\cdot}5664, \quad (2b)$$
$$3{\cdot}785p_A + 1{\cdot}105p_B + 1{\cdot}942p_C = 0{\cdot}1857. \quad (2c)$$

Whence the 'best' values for the introduction rates are given by

$$p_A = 0{\cdot}0171, \quad p_B = 0{\cdot}0565, \quad p_C = 0{\cdot}0303,$$

with $\Sigma d^2 = 112{\cdot}5$ and a root-mean-square error in the determination of the household introductions, by assuming the introduction rates for each class to be independent of household type, of 5·30 or about 4 %. This is less, though not disturbingly less, than the error to be expected, about 9 %, on the basis of random sampling.

A simple cross-infection risk can be computed for the household types in Table 5 by dividing the number of apparent cross-infections, i.e. the total infections occurring during the period less the computed number of introductions, by the exposure to risk, Σ {$n \times$ number of incidents in household of n members} less the computed number of introductions.

A similar cross-infection risk for each class of person can be obtained as follows. Introduction rates are assumed to be proportional to those just obtained, i.e. λp_A, λp_B, λp_C. When these introduction rates are multiplied by the 'person-weeks of exposure' for each category and summed for the household type the result must be equal to the 'total household introductions'. Hence λ is determined and the 'total household introductions' may be apportioned between the categories (the exact values cannot be used since, owing to error, they would not lead to the observed aggregate number of household infections). The total Σ ($n \times$ incidents) in the given household type may be apportioned according to the number of person-weeks of exposure in each category. Then for each category the cross-infection risk is obtained by dividing the total number of infections less the computed introductions by Σ ($n \times$ incidents) less the computed introductions.

(5) *The distribution of multiple infections in households* (*Table* 9)

If multiple infection in a household be assumed to result from exposure to a single individual introducing the infection the frequencies of 0, 1, ..., n cross-infections in a household of $n+1$ individuals will be given by the successive terms of the expansion of $(q+p)^n$, where $p=(1-q)$ is the cross-infection risk. (This assumes equal susceptibility among all the individuals comprising the household.)

It was obvious from inspection of the data that multiple infections involving a large fraction of the household occurred much more frequently than was compatible with this hypothesis.

If the assumption be made that subsequent infection in the household exposes the remaining, as yet uninfected, members of the household to the same risk as the primary case, then the chance of i cross-infections in a household of $n+1$ individuals following a single primary case is given by

$$ {}_nP_i = \frac{n.\overline{n-1}\ldots\overline{n-i+1}}{i!}\, p^i q^{2(n-i)}\,(F_i+1), \tag{3} $$

where $$F_i = \sum_{j=1}^{j=i-1} \frac{i.\overline{i-1}\ldots\overline{i-j+1}}{j!}\, q^{n-j}(F_j+1)$$

and $$F_1 = 0 \quad \text{and} \quad {}_nP_0 = q^n.$$

These formulae are essentially the same as those given by Greenwood (1931).

Alternatively, we may assume that the risk of cross-infection is the same *vis-à-vis* each infected individual, e.g. if there are m secondary cases then the risk of cross-infection presented, at this stage, to each of the $n-m$ as yet uninfected members of the household is $(1-q^m)$. On this basis:

$$ {}_nP_i = \frac{n.\overline{n-1}\ldots\overline{n-i+1}}{i!}\, q^{(i+1)(n-i)} G_i, \tag{4} $$

where $$G_i = 1 - \sum_{j=0}^{j=i-1} {}_iP_j,$$

and $${}_nP_0 = q^n.$$

In Table 9 the observed distribution for those groups in which a reasonable number of incidents were recorded are compared with the distribution calculated from the simple binomial $(p+q)^n$ and from equations (3) and (4). The calculations have been carried out assuming each incident to derive from a single introduction. The correction for the effect of multiple introductions is complex and would not materially affect the general conclusions. Values of the cross-infection risk approximately corrected for this factor are included in Table 9. It will be seen that both equations (3) and (4) give rise to distributions which approximate to those observed much more closely than do the distributions derived from the simple binomial. From this is might be inferred that most individuals who fail to be infected by exposure to a given infected individual do so because of lack of effective contact with the infection rather than a difference in their immunity state. Owing

Table 9. *The distribution of multiple infections in household incidents*

Household type	Formula	p	p'	Frequency of indicated no. of cross-infections: 0	1	2	3	4	5	6	D.F.	χ^2	P
$(A)_2$	0	—	—	62	20	—	—	—	—	—	—	—	—
	1	0·244	0·205	62	20	—	—	—	—	—	0	—	—
$(A)_3$	0	—	—	20	9	9	—	—	—	—	—	—	—
	1	0·355	0·329	15·8	17·4	4·8	—	—	—	—	1	8·83	<0·01
	2, 3	0·294	0·273	18·8	11·4	7·8	—	—	—	—	1	0·76	0·4
$(AB)_3$	0	—	—	20	6	6	—	—	—	—	—	—	—
	1	0·281	0·234	16·6	12·9	2·5	—	—	—	—	1	0·29	<0·01
	2, 3	0·238	0·201	18·5	9·0	4·5	—	—	—	—	1	1·63	0·2
$(AC)_3$	0	—	—	9	5	4	—	—	—	—	—	—	—
	1	0·361	0·343	7·4	8·3	2·3	—	—	—	—	1	2·92	0·1
	2, 3	0·299	0·284	9·0	5·1	3·9	—	—	—	—	1	0·00	>0·9
(All types)$_3$	1	—	—	—	—	—	—	—	—	—	3	21·0	<0·01
	2, 3	—	—	—	—	—	—	—	—	—	3	2·39	0·5
$(A)_4$	0	—	—	15	5	2	2	—	—	—	—	—	—
	1	0·208	0·186	12·0	9·3	**2·4**	**0·2**	—	—	—	1	3·39	0·05
	2	0·162	0·148	13·9	6·0	**3·2**	**0·9**	—	—	—	1	0·26	0·6
	3	0·159	0·146	14·3	5·8	**2·7**	**1·2**	—	—	—	1	0·15	0·7
$(AB)_4$	0	—	—	12	6	7	6	—	—	—	—	—	—
	1	0·408	0·346	6·4	13·3	9·2	2·1	—	—	—	2	16·7	<0·01
	2	0·288	0·250	11·1	6·9	8·0	5·0	—	—	—	2	0·51	0·8
	3	0·282	0·246	11·4	7·1	6·8	5·7	—	—	—	2	0·18	0·9
$(AC)_4$	0	—	—	10	9	2	7	—	—	—	—	—	—
	1	0·404	0·375	5·9	12·0	8·2	1·8	—	—	—	2	23·3	<0·01
	2	0·285	0·268	10·2	6·3	7·1	4·4	—	—	—	2	6·35	0·05
	3	0·280	0·263	10·4	6·4	6·2	5·1	—	—	—	2	4·64	0·1
$(ABC)_4$	0	—	—	26	15	3	9	—	—	—	—	—	—
	1	0·302	0·218	18·0	23·4	10·1	1·5	—	—	—	2	50·8	<0·01
	2	0·222	0·169	24·5	13·4	10·5	4·6	—	—	—	2	9·64	0·01
	3	0·220	0·166	25·4	12·8	9·2	5·7	—	—	—	2	6·11	0·05
$(AB)_6$	0	—	—	9	1	5	3	3	1	—	—	—	—
	1	0·336	0·284	2·9	7·2	7·3	3·6	**0·9**	**0·1**	—	3	28·0	<0·01
	2	0·200	0·174	7·2	3·7	4·2	3·6	**2·4**	**0·9**	—	3	2·83	0·4
	3	0·186	0·164	7·9	4·0	3·3	2·8	**2·5**	**1·6**	—	3	3·29	0·3
$(AB)_7$	0	—	—	5	6	4	1	1	0	5	—	—	—
	1	0·386	0·319	1·2	4·4	6·9	5·9	2·8	**0·7**	**0·1**	4	41·9	<0·01
	2	0·206	0·176	5·5	2·9	3·5	3·7	3·4	**2·3**	**0·8**	4	8·29	0·1
	3	0·184	0·161	6·5	2·8	2·9	2·6	2·8	**2·7**	**1·7**	4	6·52	0·3
$(ABC)_7$	0	—	—	6	8	4	4	2	4	1	—	—	—
	1	0·356	0·309	2·1	6·9	9·5	7·0	2·9	**0·6**	**0·1**	4	38·6	<0·01
	2	0·192	0·171	8·1	4·1	4·5	5·0	3·8	**2·5**	**1·0**	4	6·00	0·3
	3	0·174	0·157	9·2	4·5	3·4	3·5	3·5	**2·9**	**2·0**	4	4·65	0·5
(All types)$_{4-7}$	1	—	—	—	—	—	—	—	—	—	18	202·7	<0·01
	2	—	—	—	—	—	—	—	—	—	18	33·9	0·01
	3	—	—	—	—	—	—	—	—	—	18	25·5	0·1
(All types)$_{all}$	1	0·331	0·313	—	—	—	—	—	—	—	21	223·7	<0·01
	2	0·239	0·211	—	—	—	—	—	—	—	21	36·3	0·02
	3	0·233	0·206	—	—	—	—	—	—	—	21	27·9	0·15

The values of p, the cross-infection risk, which have been used to calculate the distributions, have been chosen so as to give agreement with the observed data in respect of the total number of cross-infections. The calculations have been carried out according to the following formulae, taking no account of multiple introductions: rows 0, the observed data; rows 1, the simple binomial; rows 2, the chain binomial equation (3); rows 3, the chain binomial equation (4).

The values of p', the cross-infection risk corrected for multiple introductions, given in rows 2 and 3 have been derived from those in rows 1 by means of equations (3) or (4) respectively.

The values of p and p' given in the last rows of the table are average values, unweighted, for all the groups included in the table.

In calculating the values of χ^2 the pairs of figures given in black have been aggregated.

The household types included in this table account for about 84 % of the incidents and of the apparent cross-infections in households of three or more members.

to the small numbers of some of the cells the statistical tests are subject to appreciable error, but it would appear that while the agreement with equation (3) is technically rather poor, that with equation (4) is moderate. Greenwood (1949) has pointed out a number of causes which may lead to a failure of this type of formula; unfortunately the further tests which he proposes cannot be applied to these data. Precise agreement could not be expected, since the risk of cross-infection, p, must vary to some extent from time to time and in the different households even when these are nominally similar in composition.

One further point needs discussion in view of the picture of the cross-infection process elaborated here. The incidence rates in the larger households (Table 2) are very little larger than those found in the smaller households. Calculation on the basis of the introduction rates computed above and a uniform cross-infection risk show that this ratio would be expected, to have the value 1·37 using the simple binomial, 1·60 using the chain binomial formulae of equation (3) and 1·68 using equation (4). These differ considerably from the value 1·06 ($\pm$ 0·024) given in Table 2. It would appear that the risk of cross-infection from one individual to another (p above) is not constant but diminishes with increasing household size, as may be inferred from the values given in Table 9; a consequence which might result from less close contact between all the members of a larger household.

(6) *The distribution of the number of colds experienced by individuals in a single year* (*Table* 6)

Since only completed years at risk could be included in this table the number of infections included is less than in Tables 1 and 2. A year was accepted so long as it did not lack more than 5 weeks. The rates per week computed from this table for a 52-week year will generally be found slightly lower than those given in Table 1. The two sets of figures are not directly comparable owing to the omission of fractional years of experience from Table 6 and the allowances made for missed weeks in Table 1.

(7) *The correlation between the number of colds suffered by the same individuals in* 1948 *and* 1949 (*Table* 7)

The necessity for two full years' experience in the same household type seriously reduces the numbers available. In order to obtain statistically significant figures it is necessary to aggregate groups. Since the incidence rates differ in different groups it is not possible to use the actual number of colds experienced as a basis for dividing the data. Instead, individuals of a given class in each year have been grouped according to whether they were among the more affected or less affected half of that class in their household type. The aggregated 2×2 contingency tables, necessarily symmetrical, are given in Table 7.

In all classes there is an apparent positive correlation between the experience in the two years. The magnitude of this correlation is such that the chance of an individual falling in the same half of the population in respect of his or her cold experience in successive years is approximately 2/3. This chance appears to be similar in all the three categories, but only in the adult category are there sufficient numbers to give a significant correlation at the 5 % level.

(8) *Recurrence of infection in an individual* (*Table* 8)

The distribution of intervals between successive infections in the same individual is given in the table. In computing the expected distribution for a random sequence of infections allowance must be made for the fact that the definition of an 'infection' which has been adopted excludes two or more infections being recorded in successive weeks, i.e. zero interval cannot appear in the distribution.

DISCUSSION

The amount of information which it has been possible to obtain from a survey of this kind has been very encouraging.

The most striking conclusion has been the major role of schoolchildren in introducing the infection into their households. This suggests that the school, which is their principal place of association, is the principal source of infection in this village.

The results obtained apply only to a small rural community. The relationships within such a community are likely to be simpler than in a more elaborate urban environment. It should, however, be possible among a larger population to study sufficiently large samples from the various classes of persons and household types, including certain additional differentiations, such as use of public transport and kind of occupation, to obtain statistically significant results. In particular, it would be of considerable interest to see whether the school still appeared to be the principal source from which infection was introduced into the household or whether the many additional contacts between adults in their daily occupations would provide such opportunities for cross-infection as to make them the primary means of spread among the working population.

SUMMARY

A study of the occurrence of the common cold in a Wiltshire village during the years 1948 and 1949 showed that, in this community, schoolchildren experienced about three times as many colds as adults living in households without schoolchildren, and that the presence of schoolchildren in the household approximately doubled the numbers of colds experienced by both adults and infants under 5 years of age.

More detailed analysis suggests that the schoolchildren acquired infections outside the household three times as frequently as did the adults and nearly twice as frequently as the infants, but that the infants were more than twice as susceptible as schoolchildren or adults to cross-infection within the household. The distribution of multiple infections in households conforms to that which would be expected if subsequent infections were as likely to infect the remaining uninfected members of the household as the first infected individual introduced into the

household. The risk of such cross-infection by exposure to infections within the same household appears to be about 1/5.

We wish to thank Miss M. T. Gamble, S.R.N., and Miss J. B. Chester McDonald, S.R.N., S.C.M., for help in the collection of the data, and the school teachers of the village schools in the Chalke Valley, the inhabitants of Bowerchalke and the Medical Officer of Health for Wiltshire for helpful co-operation.

REFERENCES

Andrewes, C. H. (1949). *Lancet*, 1, 71.
Andrewes, C. H., Lovelock, J. E. and Sommerville, T. (1951). *Lancet*, 1, 25.
Fisher, R. A. (1946). *Statistical Methods for Research Workers*, 21.02, 10th ed. Edinburgh: Oliver and Boyd.
Greenwood, M. (1931). *J. Hyg., Camb.*, 31, 336.
Greenwood, M. (1949). *Biometrika*, 36, 1.

(*MS. received for publication* 8. III. 51.)