The Familial Aggregation of Infectious Diseases *

WADE H. FROST, M.D., F.A.P.H.A.

Professor of Epidemiology, School of Hygiene and Public Health, Johns Hopkins University, Baltimore, Md.

ONE of the most characteristic features in the epidemiology of the common acute communicable diseases is the grouping of cases in time and space; and this is especially apt to be noted in the group which constitutes a household, people in close contact with each other, sharing a common environment, mostly of close kinship, and usually under the eye of at least one medical or lay observer whose observation encompasses the whole group. It might be supposed, then, that we would have unusually full knowledge and exact records of intrafamilial concentration of infectious diseases. And it is true that much of what we know about the communicability of certain diseases, and the special conditions related to exposure, susceptibility, and immunity is derived from observation on their spread —or failure to spread—in attacked families. It is also true that it has long been customary, even in rudimentary discussions of the epidemiology of communicable diseases, to give some account of the frequency and sequence of multiple cases within the family. But notwithstanding the tremendous mass of observations and literature bearing upon the concentra-

tion of communicable diseases within the household, our knowledge of the subject is still, for the most part, diffuse and unorganized.

This is due partly, perhaps chiefly, to the fact that the observations which have been reported have been presented in a variety of different forms, usually incomplete, and not readily convertible to a common basis. For instance, one observer will report the number or proportion of families in which more than 1 case occurred, while another will report the proportion of all cases in the community which occurred in families with more than 1 case. Both of these forms of statement are obviously incomplete, and neither is convertible into the other. A better, though less common, form of statement is the number of families having 1, 2, 3, . . . etc. cases; but this is of limited significance unless the material is further classified to show the distribution of families by size, and the presentation then becomes cumbrous.

Such systematic knowledge of the subject as we have today we owe largely to Dr. Charles V. Chapin, to whom we are indebted not only for furnishing the most extensive data as yet available on three communicable diseases—scarlet fever, diphtheria, and measles—but even more for having de-

veloped and introduced the simplest, most direct, most widely applicable and useful method so far devised for assembling the essential facts. The development of this method is only a detail in Dr. Chapin's classic studies of the epidemiology of communicable diseases; but it is a detail which is so intimately related to the whole plan and purpose of his work, and has been so widely useful, that it has seemed to me of interest to follow his approach to this particular way of summarizing experience.

The method which Dr. Chapin used for quantitative description of familial aggregation, a method which, so far as I have been able to ascertain, was introduced by him,* is the so-called " secondary attack rate "; that is, the attack rate, within a specified time, in members of attacked families exclusive of the " primary " case, time being reckoned from the onset of this case. The method itself is familiar; but before discussing its history it may be appropriate to call attention to its special merits—simplicity, and adaptability to a variety of uses.

The facts required are simple. They are, essentially: the date of onset of the first case in each family; enumeration of the other persons present in the household at the time; and record of the dates of onset of such additional cases as occur among them within a specified time from onset (or recognition) of the first case. It is useful and customary to add further facts about the primary cases, their household associates, and the cases among them, in order that the secondary attack rate may be related to such circumstances as the character and man-

agement of the first case, the sex, age, and previous history of others in the household, various details of their habits, movements and environment, and perhaps their kinship to the primary case. In addition to being simple and readily obtainable with fair accuracy, these facts are all of the kind essential for guidance of administrative authorities in their management of cases of communicable diseases.

Given the essential facts, their compilation for derivation of secondary attack rates is equally simple. It is necessary only to set up an array of the exposed persons, that is, in each family, all those present exclusive of the primary case, classified by sex, age, or other conditions which it is desired to take into account; to set up parallel arrays of the cases occurring in each group within specified time limits dating from onset of the first case in each family; and to calculate the secondary attack rate as a simple ratio between cases and exposed persons in each class. It is, of course, necessary, in any particular study, to adopt certain definite rules which may have to be more or less arbitrary, or to follow established conventions, in such matters as the definition of a " primary " case, the period of time for which the secondary attack rate is to be calculated, etc. These, however, are merely the elaboration in detail of what is essentially an altogether simple procedure; and the same may be said of the devices which must be employed in application of the method to chronic diseases or to frequently recurring infections such as the common cold.

The secondary attack rate has the further merit that it yields information in terms which are simple and directly related to the practical problems of the health officer. The question or series of questions which it answers directly is: given a case of communicable disease in the family, what is the

---

* It is of course impossible to be certain that this method had not been used and presented in some earlier publication; but the evidence available to me indicates pretty clearly that Dr. Chapin devised the procedure to meet his own needs, and it is certain that its general use in this country derives from his work.

risk of attack borne by others in the same household within specified periods of time, and how does this risk vary with the character and management of the original case and with the sex, age, past history, and various other circumstances of those exposed? It is upon the answer to these questions that the health officer must base his decisions as to the kind and duration of protective and restrictive measures which are required for the protection of the family and the public.

In addition to its quite obvious usefulness in giving a direct answer to practical questions of this kind, the secondary attack rate has broader significance in contributing to better understanding of the natural history of communicable diseases, especially the directly transmitted endemic infections for which the human host is obligatory. For such diseases the secondary attack rate gives the best approximation we have to the limits of incidence of the clinically recognized disease—or equally of the sub-clinical infection where it can be independently observed—in a group subjected to almost certain physical exposure to the infective agent. More extended knowledge of these limits is of fundamental importance, not only for comparison of one disease with another, but equally for comparing the manifestations of the same disease in different areas and different periods of time.

It is also of interest to ascertain the ratio of the secondary familial attack rate to the currently prevailing attack rate of the same disease in the general population; and this is readily done because the two rates may readily be reduced to the same time basis and are otherwise directly comparable. The interpretation of this ratio is not simple, but it clearly is an index of the focalization of disease or infection around the recognized clinical case, and has interesting if somewhat elusive im-

plications with respect to the relative frequency of clinically manifest cases of the disease as compared with concealed carriers of infection at large in the community.

Those who have found the secondary attack rate ready-made, having become familiar with the procedure before they have encountered the problems to which it applies, are apt to think that because it is so simple and direct it must have suggested itself immediately in the study of infection in families. But in epidemiology, as in other fields, the procedures which are at the same time simplest and most efficient are not often the ones which suggest themselves first. The direct approach to any problem is only one of many possible approaches, most of which are indirect; and there is an adage which reminds us that the appropriate method is most likely to be invented when it is needed to achieve a definite result. This appears to be the history of Dr. Chapin's development of this particular method. So far as may be judged from the methods which he actually used in assembling his data, it appears that he first used this particular form of statement when his investigations had raised a quite definite and specific question which could be answered only in terms of the secondary attack rate.

The record from which this is inferred, fortunately an unusually full record, and the only one to which I shall refer, is the series of annual reports rendered by Dr. Chapin as Superintendent of Health of the City of Providence, R. I., beginning with the first year of his incumbency, 1884, and followed through 1905, beyond which, for present purposes, it is not necessary to go. In these reports Dr. Chapin gives and discusses, year by year, a series of tables presenting his observations on various aspects of the epidemiology of scarlet fever and

diphtheria. These, taken in sequence, give an objective record of the specific problems taken up and of the data and methods used in attacking them.

In the first paragraph of the first report of the series, for the year 1884, following a comment to the effect that health departments probably are devoting too large a share of their attention to the abatement of nuisances, Dr. Chapin expresses the belief that " there is no doubt that more attention might well be given to the prevention of the more common infectious diseases and more time devoted to their scientific investigation," and following this, he notes the need of a medical inspector for this purpose. The report for the next year announces the appointment of Dr. Gardner T. Swarts as Medical Inspector, and among his duties, listed in detail, specifies:

To visit every case of contagious disease which is reported to this office . . .
To obtain as much information in regard to the cause of the disease as possible and to keep a record of all cases on blanks which, when returned to this office, are there kept on file.

It is of interest that the report of the Medical Inspector for this year gives the following somewhat unusual table, which includes in one group cases of scarlet fever, diphtheria and typhoid fever:

Whole number of cases reported   387
Number of premises inspected   321
Extra number of cases on same
  premises   66

Beyond the suggestion which this table gives of a special interest in the spread of contagious disease within the family, there is little in the first three reports to indicate the particular lines of investigation to be pursued. These are revealed, however, in the report for 1887. Here Dr. Chapin mentions that he has " collected some facts in regard to cases of scarlet fever occurring during the latter part of last year,

which, perhaps, are not of much value in themselves but may be when the number of cases can be increased." This marks the beginning of systematic tabulations of certain facts about scarlet fever, continued in substantially the same form for more than 20 years, adding to the numbers each year. Dr. Chapin's discussion, in the 1887 report, of the data on scarlet fever collected that year, indicates pretty clearly the kind of facts collected and the specific questions which they were intended to answer. These, however, are indicated as clearly and somewhat more succinctly in the table given below which first appeared in the report for 1888 and was continued thereafter for many years in nearly the same form. For convenience of reference the table as reproduced here is divided into 5 numbered sections. There is no such subdivision in the original, but in all other respects the table is precisely as shown here.

Each section of the table states by implication at least one definite question having reference to the frequency of the extension of scarlet fever to other members of the same family, or to other families in the same dwelling. Excepting the last question, which refers to frequency of attack in (susceptible) *individuals* following return to their homes, the most direct answers which the data give to the questions are in terms of the proportion of *families* in which the extension occurs. The secondary attack rate in " susceptible children " (without further differentiation by age) may be derived very easily from the data given in section (1) of the table and would probably yield the most satisfactory answer to the general question there indicated. I think it will be granted, however, that given sufficiently extensive observation—which Dr. Chapin subsequently accumulated—all the specific questions proposed in the table

SUMMARY OF OBSERVATIONS ON SCARLET FEVER

From the

Sixth Annual Report of the Superintendent of Health of the City of Providence for the Year Ending December 31, 1888

(The table as printed is not divided into the numbered sections shown here.)

	1887	*1888*	*Total*
1			
Number of families in which there was more than one susceptible child	232	244	476
Number of these in which there was no second case	102	97	199
Number of susceptible children in all the above families	986	827	1,813
Number of these children who were attacked	452	511	963
2			
Number of additional families with susceptible children in the same house where the disease appeared	112	128	240
Number of susceptible children in these families	381	354	735
Number of these additional families attacked	27	16	43
Number of children in these families attacked	58	21	79
3			
Number of families where inunction was practiced	87	99	186
Number of instances in the above families where the disease was not confined to one case	44	64	108
Number of susceptible children in these families	218	496	714
Number of these children who were attacked	148	319	467
4			
Number of tenements which were fumigated	51	63	114
Number of instances where fumigation was done, and where the disease spread to other families in the house	5	5	10
5			
Number of instances where susceptible children were at once removed	24	18	42
Number of instances where they were attacked on their return	3	1	4

A susceptible child is defined as " one who has not had the disease "; and a footnote to the table explains that " two months is allowed to elapse before the family account is considered closed."

may be answered satisfactorily by analyses in the forms indicated. Use of the secondary attack rate would have been preferable in some of the analyses; but in none is it obligatory.

The same report, for 1888, gives, for the first time, in incomplete form, a table which, in various more complete forms, is continued thereafter for many years. It refers to attacked families, and shows, by single years of age, up to and including age 20 (all above this age being grouped together as adults):

(1) " Number exposed," *i.e.,* the whole number of persons of indicated age present in the families, including both primary and secondary cases

(2) " Number of cases," without distinction between primary and secondary

(3) " Ratio of cases to number exposed," expressed as a percentage

Inclusion of the primary cases among the number of " exposed " is unfortunate.* It makes the ratio of cases to people " exposed " exceedingly difficult to interpret, and in fact Dr. Chapin himself appears at times to

---

* The persons first attacked in their families have of course been " exposed " to the specific infection of scarlet fever; but, having come into the table *because* of having scarlet fever, they are not exposed to the risk of contracting it within the time to which the table refers. The failure to distinguish between these two meanings of " exposure " makes the table confusing.

have misinterpreted it. Nevertheless, this form of tabulation has a definite and useful place in the sequence leading to subsequent use of the secondary attack rate. It requires only dissection of the basic data to yield secondary attack rates; and, as it stands, it gives useful information as to age distribution.

Beginning in 1889, the reports include tabulations for diphtheria equivalent to those which have been mentioned for scarlet fever; and from 1896 separate tables are given for diphtheria in families where the diagnosis was bacteriologically confirmed.

In 1903 an important addition is made to the series, namely, a table relating to scarlet fever and a similar one for diphtheria, showing, for a considerable period of years, the number of secondary cases distributed according to time of onset in days from onset of the primary case in the family. The age distribution of secondary cases is not given in these tables.

It is in the report for 1903 that Dr. Chapin first makes use of the secondary attack rate. Since 1891 an increasing proportion of cases of diphtheria and scarlet fever had been removed each year to the hospital for contagious diseases; and the question which brings in the use of the secondary attack rate is that of the resultant benefit to other members of the families from which the primary cases had been removed.

Referring first to diphtheria, Dr. Chapin gives a table showing, by age, the number of persons left at home in the families from which primary cases were removed to the hospital. Among the 1,420 persons included in this group, 55 were attacked subsequent to removal of the primary case. The specific question asked is: how many cases of diphtheria might have been expected in these 1,420 people if the primary cases had remained at home? This, as he notes, requires the calcu-

lation of age-specific secondary attack rates—*i.e.*, attack rates in members of the family *exclusive* of the primary case—in families where the primary case remained at home. No table was available giving the full data required for calculating these rates, namely, (1) the age distribution of persons in attacked families exclusive of the primary cases, and (2) secondary cases at each age. However, a table (A), which has been mentioned heretofore, gives the aggregate number of secondary cases (1896–1903), not separated by age; and another (B) gives, by age, the number of people present in attacked families, the number of cases, and the ratio of cases to people. Referring to these two tables, Dr. Chapin observes that the secondary cases constitute about 40 per cent of all that occur, hence he takes 40 per cent of the ratios shown in table (B) as representing approximately the secondary attack rates. Applying these rates to the corresponding age groups of the population left at home, and making an ingenious adjustment for the varying intervals from onset of the primary case to removal from home, he estimates that had these patients remained at home their familial contacts might have been expected to suffer an incidence of about 120 cases instead of the 55 actually recorded. The same method is then applied in estimating the " expected " as contrasted with the observed incidence of diphtheria in well persons who removed from their homes subsequent to the onset of the primary cases; and similar estimates are made with respect to scarlet fever.

Two years later, in the report for 1905, appears the first table giving age-specific secondary attack rates for males, females, and both sexes combined. It refers to diphtheria, including only families for which the diagnosis was bacteriologically confirmed, and includes data for the full

decade 1896–1905. As the table is given here it is necessary to refer to another table for some basic data used in calculating the rates, but with this reference it is complete. Later, the form and arrangement of the table were altered at various times, bringing the data together in more convenient form. Beginning in 1908, tables are added showing secondary attack rates in scarlet fever; and in various subsequent reports special tables are given showing separately the secondary attack rates in families where the patient lived (that is, recovered) and remained at home; in families where the primary case was removed to a hospital, and in families employing servants or nurses. In general, these tables refer back to dates considerably earlier than the dates of the reports in which they first

appear, and the figures given are large.

This review has followed only one particular line of Dr. Chapin's studies, the one which led to use of the secondary familial attack rate. No attempt has been made to follow through the many other lines along which his investigations of scarlet fever and diphtheria extended; and space does not permit even their enumeration here. Most of the specific questions investigated did not require use of the secondary attack rate and had already been answered by other methods and in other terms before this form of statement had been adopted. When a specific question eventually arose definitely requiring this method, the essential data had already been provided, and the method itself cannot have been far below the surface.

National Foundation for Infantile Paralysis

FOLLOWING his recent announcement of the establishment of the National Foundation for Infantile Paralysis, President Roosevelt has issued a statement of the appointment of 34 citizens of various parts of the United States to serve as trustees.

It is hoped that a fund of from 7 to 10 million dollars will be raised in the next 5 years, with which the Foundation " will make every effort to ensure that every responsible research agency in this country is adequately financed to carry on investigations into the cause

of infantile paralysis and the methods by which it may be prevented. . . ." It also aims to develop means of enabling those already crippled by poliomyelitis to become economically independent in their own communities.

All funds collected through the celebration of the 5th annual " birthday ball," held January 29, 1938, in honor of the President's birthday January 30, will be given to the new Foundation, instead of being divided as heretofore between the Georgia Warm Springs Foundation and local communities.