AGE INCIDENCE IN ZYMOTIC DISEASES.

BY

JOHN BROWNLEE, M.A., M.D.,

PHYSICIAN SUPERINTENDENT, CITY OF GLASGOW FEVER HOSPITAL, BELVIDERE.

PRINTED FOR THE

ROYAL PHILOSOPHICAL SOCIETY OF GLASGOW,

BY CARTER & PRATT, 62 BOTHWELL CIRCUS.

1904.

[From the *Proceedings* of the Royal Philosophical Society of Glasgow.]

Age Incidence in Zymotic Diseases. (A contribution from the Sanitary and Social Economy Section). By JOHN BROWNLEE, M.A., M.D., (Glasg.), D.P.H. (Camb.), Physician Superintendent, City of Glasgow Fever Hospital, Belvidere.

[Read before the Society, 24th February, 1904.]

THE subject of this paper is a consideration of the meaning of the statistics of infectious diseases, as they occur at different ages, in the light of the mathematical developments of the theory of probability which have been divised in recent years, especially by Professor Karl Pearson. The needs of astronomers and surveyors long ago made the adoption of some means of reconciling discordant observations a practical necessity. It early came to be generally recognised that if an observation were made a certain number of times with equal care, and with equally perfect instruments, the most probable value of the quantity observed is the arithmetical mean or average of the values given by the observations. From this, what is commonly called the normal probability curve, was deduced as a necessary consequence. The normal probability curve, which is figured in diagram I., is the geometrical expression of the frequency with which any error of definite magnitude may be expected to occur in a number of observations. It is symmetrical about the middle line, which indicates that positive and negative errors are equally likely to occur. The ordinate or vertical distance of any point on this curve from the base line decreases in value as the error increases in magnitude, thus large errors occur with great infrequency, while small ones are much more numerous. The example which is most commonly chosen to illustrate the application of this law to practical problems is taken from rifle shooting. Given a good shot, using a rifle without bias, the atmosphere being clear and steady, out of a given number of shots bull's eyes will constitute the most numerous group ; inners will follow, then outers, while misses will be the rarest of all. As many shots will fall to the right of the bull's eye as fall to the left, and as many above as below. This kind of grouping extends very widely throughout nature. If in considering, for example, the stature of children at a definite age, the children are graded so that they fall into groups in accordance with their height, it will be found that there is a height which is more frequent than any other, and with each inch of difference above or below this, the numbers in the groups tend to grow smaller and smaller, while giants and dwarfs are exceedingly rare. When, however, this method is applied extensively it is quickly seen that this symmetry does not obtain to any very large extent in nature's groups, the numbers tending to be distributed more to one side of the line of maximum frequency than to the other. This gives the curves a character, which is called "Skewness" by Professor Karl Pearson. As an example of what is meant, the illustration of the rifleman can again be employed: he was referred to as shooting under ideal conditions, but let there be in addition a variable wind blowing across the range, and it is evident that the number of shots striking the one half of the target will somewhat exceed those striking the other. All the distributions which are considered in this paper, except one, belong to this class of "Skew" curves, few instances of an infectious disease grouping itself with regard to age incidence so as to approximate to the normal curve existing. For each disease there is a definite period of life at which the maximum susceptibility occurs, and above and below this the number of cases becomes gradually smaller and smaller though the fall in numbers may occur much more quickly on one side than on the other. Thus in scarlet fever the maximum frequency of attack is between four and six years, while under one year and over forty, few cases occur.

The manner in which the mathematical theory of probability is applied to such cases as this is briefly as follows. The form of the curves which represent distributions in the production of which chance is the main factor are first thoroughly investigated. For this Professor Pearson is practically the sole authority. The classes of statistics which it is desired to examine are then tested as to how far they conform with the distributions admitted by the theory. Discrepancy between the facts and the theory, if it exists, can be observed. Of the causes of this, inadequacy in the number of observations is one. It may also be due, and in the class of fever statistics is commonly due, to the presence of some factor in addition to the most important, namely the age susceptibility, acting specially at some age period. An obvious instance is given by the age distribution of persons suffering from measles: the susceptibility to this disease decreases with each year of life from four years upwards, but on the curve of age distribution a marked secondary rise takes place between the ages of seventeen to thirty,[1] which, when examined into, is found due to immigration to Glasgow of persons from country districts from which measles has long been absent, who consequently form a new population of susceptible persons at an age period at which among natives of Glasgow few such are to be found. It would seem that one of the chief advantages offered by the application of the method in our case is to determine at what ages such causes are specially to be looked for. The method by which the appropriate probability curve can be found from the observations given in the statistics cannot be explained here. It depends in principle upon obtaining the curve belonging to the group of probability curves which has the same moments about a vertical line through its centre of inertia as has the curve given by observations. The first moment about a vertical line is that obtained by multiplying each vertical strip by its distance from the fixed line as in the process of finding the centre of the centre of gravity, the second by the square of that distance as in finding the "moment of inertia," the third by the cube of the distance, and the higher moments in the same way.

In general in infectious disease the origin of the curve must be placed at birth. With this assumption the number of moments required is reduced from four to three. It is true that in a few diseases abortion takes place, and some of these cases of premature birth may be due to the fœtus having acquired the disease from which the mother is suffering. This especially applies to smallpox and to typhus fever, but with regard to all other infectious diseases such is either of great rarity or impossible. The point of origin of the curves in general must thus be taken at the date of birth. Two classes of curves satisfy an asymmetrical distribution which begins at zero, rises to a maximum and then declines. In one of these the upper limit is definite and in the other indefinite. With most infectious diseases the susceptibility becomes so small at high ages as to be almost nil, and although it is more logical to assume a definite limit to life yet the simplification in calculation involved in leaving the upper

---

[1] See Appendix.

limit undefined is so great, and the difference of the number of cases at high ages given by the two assumptions is so very minute, that I have not hesitated to chiefly use the latter. It is found by trial that the age distribution follows this theoretical grouping very closely when diseases are considered which chiefly affect adults; thus, enteric fever, typhus fever, and smallpox in vaccinated persons, show a very marked correspondence between the theoretical and the actual distribution of cases. Prof. Pearson has fully discussed the case of enteric fever. He takes the statistics of 8,689 cases of this fever treated by the Metropolitan Asylums Board Hospitals. The fit of the theoretical curves is here exceedingly close, but those given by considering the higher moments of the curves demand placing the origin of the curve at respectively 1.3 and 2.8 years before birth. When the origin is assumed to be at birth the fit is not quite so close. This, however, is what is a priori to be expected. Susceptibility to a fever cannot be expected to vary only according to only one law from birth to old age, even when the disease is essentially one affecting adults. To pass from enteric fever to typhus the statistics of the male patients suffering from this fever, and treated in Belvidere Fever Hospital, have been chosen for consideration. The diagram[1] shows the kind of correspondence which the theoretical curve bears to the actual statistics. It will be seen that the theoretical curve has a course such that the curve of actual observations zig-zags backwards and forwards across it. One interesting variation noticed here, and also when the statistics of London are similarly examined, is that at the age of from forty to fifty-five years, the actual numbers are in excess of the theoretical ones. These are the years at which habitual drunkenness has fixed its mark on its devotees, and it would seem to be indicated that such persons not only offer themselves more easy victims to this disease when attacked, but in addition are more susceptible to the infection. The same condition has likewise been noted clinically as predisposing to pneumonia. A similar diagram is given of the deaths among the same series of cases. Here on account of smaller numbers the actual curve is much more irregular than that noted among the cases, which number about eight times those of the deaths. The fit of the theoretical curve is thus not so close. When, however, the two are combined so as to form a theoretical mortality curve

---

[1] Diagram II.

it will be seen that, the same zig-zagging of the actual, through the theoretical again takes place, and at only two points is there a serious discrepancy. One of these exceptions is in the first five-yearly period (0·5 years), where it is known that the mortality among children under three years is specially high. This comes under a group of special infantile susceptibility which will be discussed later. The other exception is in the age group fifty-five to sixty-five years, where the number of cases and deaths is small and the liability to error consequently great. It is thus probable that the theoretical value is in this case much more accurate than the observed, a conclusion borne out when comparison is made with the tables of age mortality in other Hospitals.

It will lead to a more easy development of the question if the age incidence in the case of whooping cough[1] be considered next. This disease is almost absolutely one of childhood. It is true that in the Hospital statistics, a number of females occur above the age of twenty, but these are largely nurses or nursing mothers who are necessarily brought into such special contact with the disease as to constitute a completely new method of infection. The statistics of the male patients are therefore chosen as more likely to represent the actual grouping of the disease. It will be seen at once by reference to the appendix that a theoretical curve beginning at birth bears little or no resemblance to the actual distribution of the cases. If by the method of successive approximation an attempt is made to split up the actual curve into groups of cases a result such as is seen in diagram III. is easily arrived at. The method is as follows: at each age period where the theoretical number given by the first found curve is in defect of the actual, the former is chosen and a theoretical curve found for this new grouping. This process is repeated until two succeeding theoretical curves do not greatly differ, and it is found then that the remaining cases throw themselves into two groups, one beginning at birth and practically ending at three years, another beginning about four years and ending about nine years. When these curves are arranged and added, the result forms a close approximation to the form of the whooping cough curve. It is not pretended that this splitting up of the curve into these three factors is more than a rough approximation or the only method by which the obviously compound form of the actual curve may be represented. A similar result,

[1] See Diagram III.

however, is obtained when the problem is attacked by the method of least squares and a probability curve fitted to the figures representing the number of cases in each of the first twelve months of life, the residue being considered as a compound curve. This analysis appears to bear the following interpretation. There is a period during infancy of special susceptibility to whooping cough. The second group represents the normal susceptibility to the disease during the age of childhood. This has its maximum between the ages of three and four years. The third group which increases the number of cases at higher ages begins about four years of age, and has its maximum at six. There seems little doubt that this represents the amount of disease which is spread by infection in schools. An exactly similar splitting up of the compound curve which represents the age incidence in measles can be made. The special susceptibility under one year of age, however, is not so great, while the part played by schools is greater. This special infantile susceptibility may also be noted to a small extent in the case of scarlet fever.[1]

The age distribution of the cases of smallpox both in unvaccinated and vaccinated persons can now be analysed. Among the former there is a very great infantile susceptibility. A second period of susceptibility occurs as in measles between the ages of three and four years. When a probability curve is fitted so as to begin at birth, and apply chiefly to the years above six, it will be found to have a correspondence with the observations of a closeness shown by no other fever except enteric. This curve of which the figures are given in the appendix was found by trial. There is no evidence at all of schools spreading the disease, a fact which can readily be understood when it is considered that the great majority of scholars are protected against the disease by vaccination, and that the character of the disease in unvaccinated persons is of a severity such as is in general will prevent their attendance at school. When the case of vaccinated persons is considered in the same manner, they are also seen to form a homogeneous group. The chief variation is at early ages, and it is to be noted in support of the truth of the statistics, that the curve based on them considered as a whole gives a very close fit. In its points of difference it demands considerably fewer cases at the ages at from five to fifteen years than are given in the actual

[1] See Appendix.

numbers, showing that the tendency to classify cases occurring at these early ages as unvaccinated, which the anti-vaccinators constantly claim is done, can hardly have arisen in the minds of those responsible for the collection of the statistics.

The statistics of scarlet fever remain to be considered. Here again as with measles and whooping cough no theoretical curve with its origin at the point of birth can be obtained which bears any particular resemblance to the actual distribution of the cases. If the higher moments are considered so that no assumption is made as to the point of origin, the resulting theoretical curve thus obtained much less represents the facts. The age distribution of scarlet fever must thus be considered due to a variety of contributory causes, acting in different degrees at different ages, as have been already seen to be the case with whooping cough and measles. Spread by schools may from the very first be put out of count. Definite school infections are very rare in Glasgow, and the influence of schools in spreading the disease, if any, must be looked for in the steady supply of small numbers of cases due to the presence of the poison either in small amount or in attenuated form. The holidays in Glasgow usually comprise most of the last fortnight in June and extend through July. The incubation period of scarlet fever is rarely more than 8 days, so that the age distribution of cases of this disease in July must represent that existing when dissemination of the poison by schools does not exist. The schools resume work after the holidays early in August, and the usual rise of scarlet fever begins in the middle of that month. At this period a large number of children are sent to school for the first time, and these must include many susceptible to the disease. If then scarlet fever were spread by schools to any extent, the percentage of cases occurring at the ages of five to twelve ought to considerably exceed that occurring at these ages during the month of July alone. As a matter of fact the percentages occurring at each age period are practically identical.[1]

At higher ages (from fifteen years and upwards) another factor evidently comes into play, about half the cases occurring in persons of country extraction. Even, however, if the number of these is subtracted from the total numbers and a theoretical curve calculated for the residue no better fit is obtained. All the

---

[1] See Appendix.

curves with their origin at birth, which I have calculated by trials such as these demand a very much larger number of cases between the ages of one and four, and a much smaller number at ages over fifteen than are given by the actual figures. The best fit on the whole is that obtained by considering the figures from four to twelve years of age, as representing the normal distribution of the disease in childhood—which can be safely done so far, inasmuch as schools play little part in its dissemination—and calculating the corresponding theoretical curve by the method of least squares. The excess at the ages 0-5 of the theoretical values above the actual is not now so marked, though it still obtains, but the theoretical numbers are only a fraction of the actual ones at all ages above twelve. This difficulty may be explained partly by immigration to Glasgow of susceptible persons as already mentioned, but partly also by the fact that infection of parents or nurses is not unusual, and a new factor influencing the spread of the disease is thus introduced. I have not succeeded any further than this in fitting to the scarlet fever statistics a compound curve, which approaches the actual facts with any closeness.

In conclusion, it may be stated that there is a very close correspondence of the theoretical and actual curves when those diseases affecting adults chiefly are considered. The application of the method, however, to diseases of children where special infantile susceptibility spread by schools, etc., are definite factors is a matter of great difficulty.

APPENDIX TO ILLUSTRATE PAPER.

Age Distribution of the Infectious Diseases of Childhood. Belvidere, 1886-1900.

Measles.

Age-Periods.	Cases.
Months.	
0— 1	3
1— 2	7
2— 3	10
3— 4	19
4— 5	36
5— 6	42
6— 7	108
7— 8	113
8— 9	145
9—10	129
10—11	122
11—12	86
Years.	
1— 2	1,227
2— 3	1,545
3— 4	1,718
4— 5	1,600
5— 6	1,503
6— 7	1,235
7— 8	796
8— 9	398
9—10	214
10—11	110
11—12	60
12—13	47
13—14	28
14—15	23
15—16	29
16—17	33
17—18	34
18—19	63
19—20	91
20—21	113
21—22	117
22—23	111
23—24	88
24—25	79
25—26	78
26—27	52
27—30	77
30—35	52
35—40	12
40—50	7
50—60	2
All Ages,	12,362

Whooping-Cough

Age-Periods.	Cases.
0— 1	13
1— 2	16
2— 3	35
3— 4	59
4— 5	48
5— 6	66
6— 7	87
7— 8	87
8— 9	91
9—10	66
10—11	73
11—12	40
1— 2	619
2— 3	742
3— 4	779
4— 5	695
5— 6	585
6— 7	420
7— 8	228
8— 9	112
9—10	55
10—15	34
15—20	6
20—25	4
25—35	2
35—45	...
45—56	3
Total,	4,965

Whooping Cough.

Males, Age Period	Cas	Th. No.
0— 1	323	166
1— 2	304	426
2— 3	345	437
3— 4	330	390
4— 5	299	290
5— 6	266	228
6— 7	201	134
7— 8	103	84
8— 9	51	52
9—10	25	31
10—15	12	30
15—20	1	3

Scarlet Fever.

Age-Periods.	Cases.
0— 1	8
1— 2	7
2— 3	6
3— 4	11
4— 5	16
5— 6	20
6— 7	35
7— 8	34
8— 9	49
9—10	41
10—11	36
11—12	32
1— 2	700
2— 3	1,296
3— 4	1,839
4— 5	2,117
5— 6	2,151
6— 7	2,077
7— 8	1,903
8— 9	1,483
9—10	1,207
10—11	1,109
11—12	845
12—13	707
13—14	572
14—15	495
15—16	367
16—18	598
18—20	463
20—25	923
25—30	474
30—35	232
35—40	106
40—50	61
50—60	10
60—70	2
70—80	1
All Ages,	22,033

Age Distribution of Cases of Scarlet Fever per 10,000 Cases for Month of July alone, and for Months of Aug., Sept., and Oct.

Age-Periods.	July alone.	Aug., Sep., Oct.
0— 1	146	132
1— 2	335	314
2— 3	524	557
3— 4	962	813
4— 5	1,168	988
Total, 0—5,	3,135	2,804
5— 6	936	957
6— 7	1,082	920
7— 8	893	902
8— 9	584	689
9—10	524	551
10—11	403	555
11—12	386	412
Total, 6—12,	3,872	4,029
12—13	309	338
13—14	309	280
14—15	180	232
15—16	137	178
16—18	292	236
18—20	128	212
Total, 12—20,	1,355	1,476
20—25	292	371
25—30	197	184
30—35	146	91
35—	60	78
Total, 20— ,	695	724
Total number of Cases on which above figures are based,	1,164	7,915

Actual and Theoretical Distribution of Vaccinated Cases of Smallpox, Glasgow, 1900-02.

Sparse Eruption.

Age-Periods.	Cases.	Theoretical No.
0— 5	0	0
5—10	25	10
10—15	78	63
15—20	95	135
20—25	188	186
25—35	390	345
35—45	177	177
45—55	56	64
55—65	17	18
65—75	6	4·3
75—85	2	1
Total,	1,034	

Abundant Eruption.

Age-Periods.	Cases.	Theoretical No.
0— 5	2	0
5—10	6	5
10—15	17	11
15—20	32	47
20—25	82	95
25—35	264	260
35—45	186	169
45—55	55	63
55—65	11	16
65—75	5	3·3
75—85	1	·6
	661	

Confluent Eruption.

Age-Periods.	Cases.	Theoretical No.
0— 5	0	·3
5—10	1	1·0
10—15	4	4·4
15—20	9	11·2
20—25	28	23·8
25—35	95	90·5
35—45	88	91·5
45—55	34	33 7
55—65	7	4 7
65—75	1	2

Actual and Theoretical Distribution of Unvaccinated Cases of Smallpox, London, 1901-02.

Age-Periods.	Cases.	Theoretical No.
0— 1	163	74
1— 2	93	98
2— 3	88	106·2
3— 4	130	107·1
4— 5	125	104·3
5— 7	212	196·4
7— 9	179	170
9—11	163	151
11—13	131	129
13—15	109	109
15—20	191	201
20—30	203	203
30—40	87	75
40—50	24	27
50—60	12	9·7
60—70	4	3·5
70—80	1	1
Total,	1,905	

Vaccinated—*Cont.*

Hæmorrhagic Cases.

Age-Periods.	Cases.	Theoretical No.
0— 5	0	
5—10	0	
10—15	0	·24
15—20	0	1·04
20—25	7	4
25—35	15	16·02
35—45	17	16·5
45—55	8	10
55—65	5	4
65—75	2	1·2
75—		

NOTES ON APPENDIX.

IN the appendix, tables are given of the age distribution of measles, scarlet fever, and whooping-cough in such detail as may allow the observation of the various periods of special susceptibility. In the case of whooping-cough a special table is given showing the number of cases at each age period among males, and for comparison the numbers given by the corresponding theoretical curve whose first two moments are identical, and whose origin is at birth. This curve is referred to in the text.

A table is also given showing the age distribution of cases of scarlet fever during the holiday months and the months immediately succeeding the holidays for comparison.

Tables showing the actual and theoretical distribution among vaccinated and unvaccinated cases are added. The vaccinated persons are divided into four classes—those who have sparse, abundant, or confluent eruptions, and those in whom the disease was hæmorrhagic.

The equations of these curves are as follows :—

Vaccinated cases

Sparse eruption, $$y = 95{\cdot}9\left(1 + \frac{x}{9{\cdot}983}\right)^{5{\cdot}8056} e^{-x \times {\cdot}58155}$$
unit of $x = 2{\cdot}5$ years.

Abundant eruption, $$y = 64{\cdot}77\left(1 + \frac{x}{11{\cdot}869}\right)^{8{\cdot}6634} e^{-x \times {\cdot}72992}$$
unit of $x = 2{\cdot}5$ years.

Confluent eruption, $$y = 26{\cdot}013\, e^{-\frac{x^2}{(4{\cdot}096)^2}}$$ (the normal curve)
unit of $x = 2{\cdot}5$ years.

Hæmorrhagic Cases, $$y = 4{\cdot}5516\left(1 + \frac{x}{14{\cdot}2103}\right) e^{-x \times {\cdot}65494}$$
unit of $x = 2{\cdot}5$ years.

Unvaccinated Cases, $$y = 107{\cdot}24\left(1 + \frac{x}{3{\cdot}205}\right) e^{-x \times {\cdot}11013}$$
unit of $x = 1$ year.

The point of origin of all these curves is the age period of maximum frequency. It will be noticed how among the vaccinated class the distance of this from the zero ordinate, which represents birth, increases with the severity of the attack. Thus among cases with a sparse eruption the age of maximum frequency is twenty-five years ($2 \cdot 5 \times 9 \cdot 98$), among those with an abundant eruption 29·6 years, among those with a confluent eruption 35·2 years, and with a hæmorrhagic one 35·5 years. It is noticeable that those cases with a confluent eruption are distributed so as to nearly coincide with that given by the normal curve, and thus obey the law of error. The point of maximum frequency is 35.2 years, and it is evident from the curve that the protection afforded above this period by advancing age exactly corresponds to the loss by lapse of time of the protection afforded by vaccination between the period at which that operation was performed and the age of maximum frequency.

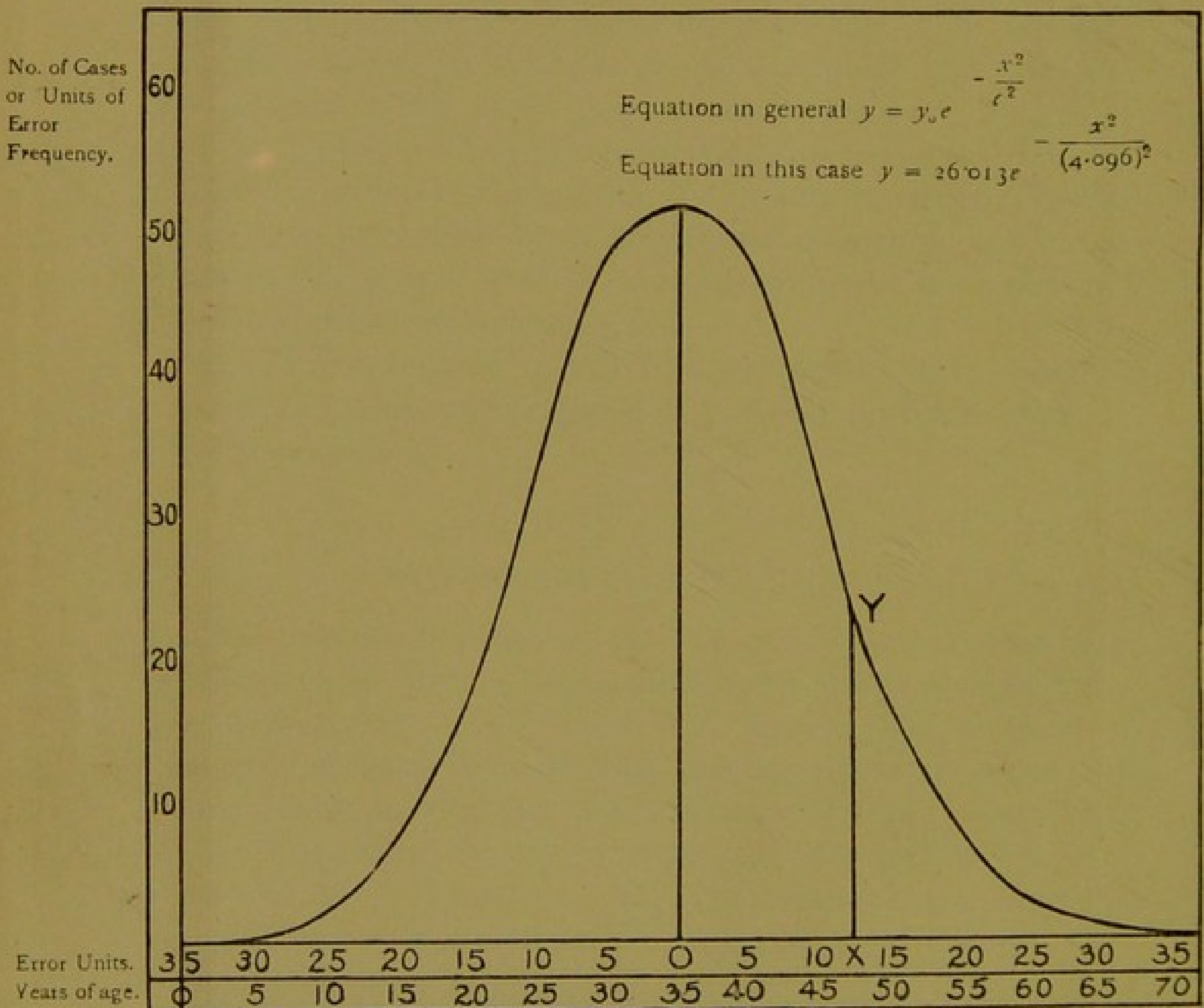
In contra-distinction to this it is noticeable that the first few years of life constitute the period of maximum attack rate among the unvaccinated.

NOTES ON THE DIAGRAMS.

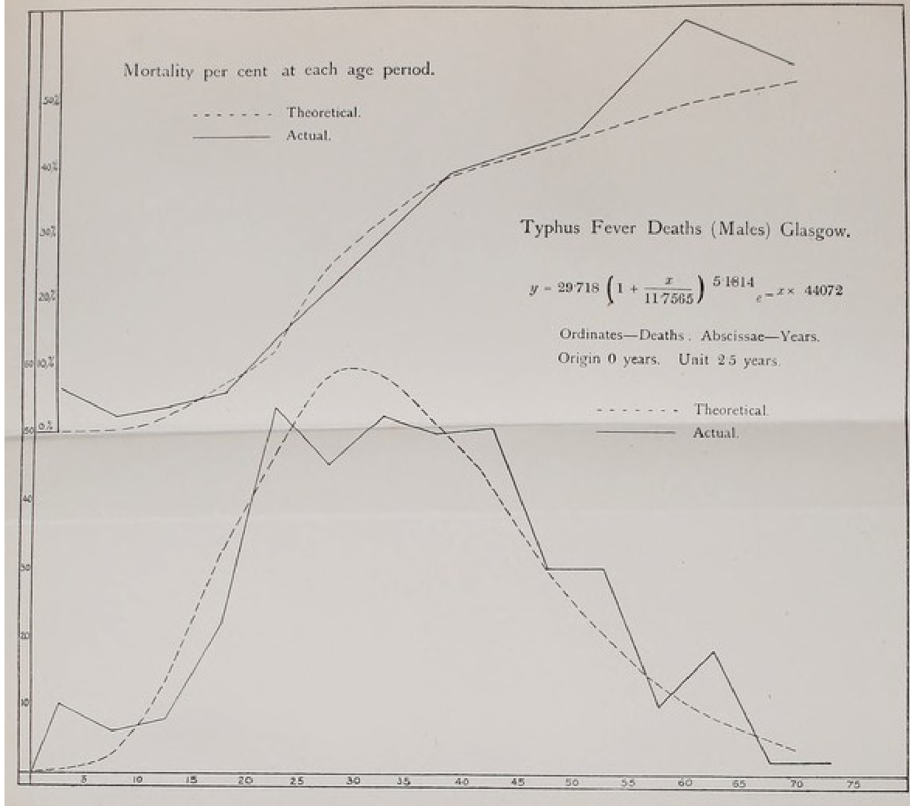
Diagram I.—The normal curve of error.—The abscissa O X measured from the middle point of the base line represents the size of the error. The ordinate X Y measures the frequency of that error so that it is easily seen that negative errors are equally common with positive, and that the larger O X is, the smaller is X Y. This curve very appropriately represents the distribution of confluent cases of smallpox, and is drawn to the scale as in the figures given for such cases in the appendix.

Diagram III.—Composite curve of whooping-cough. It is to be noted in this diagram that the second and third curves which go to make up the composite curve, which consists of heavy dots, are drawn to a slightly larger scale than is correct in order to avoid unnecessary confusion in the diagram. The fit of the composite curve is therefore considerably closer than is represented.

DIAGRAM I.



THE NORMAL CURVE.


Mortality per cent at each age period.
Theoretical.
Actual.
Typhus Fever Deaths (Males) Glasgow.
y = 29·718 (1 + x/11·7565)^5·1814 e^−x× 44072
Ordinates—Deaths. Abscissae—Years.
Origin 0 years. Unit 2·5 years.
Theoretical.
Actual.

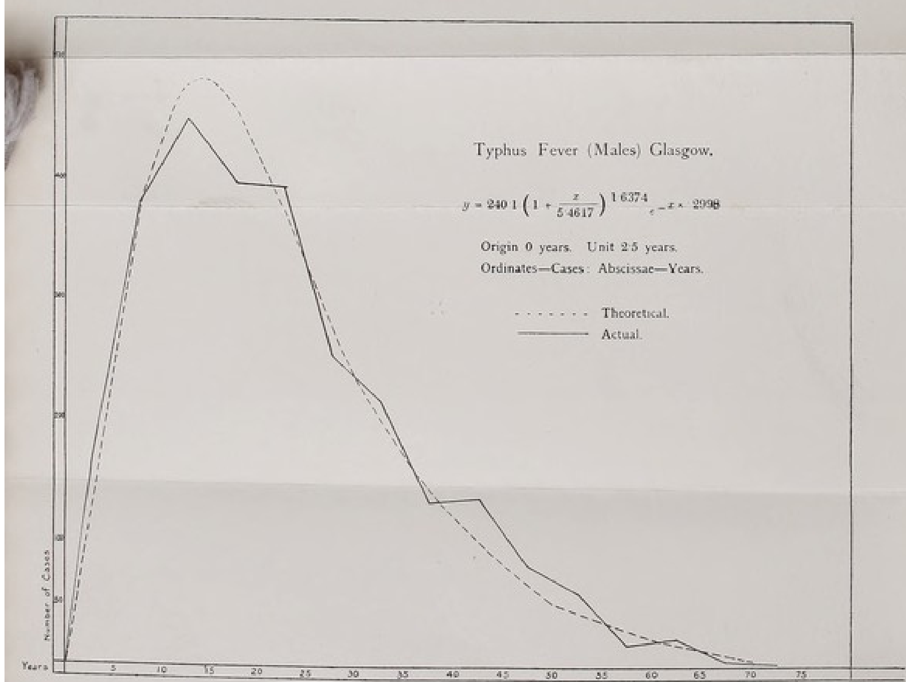
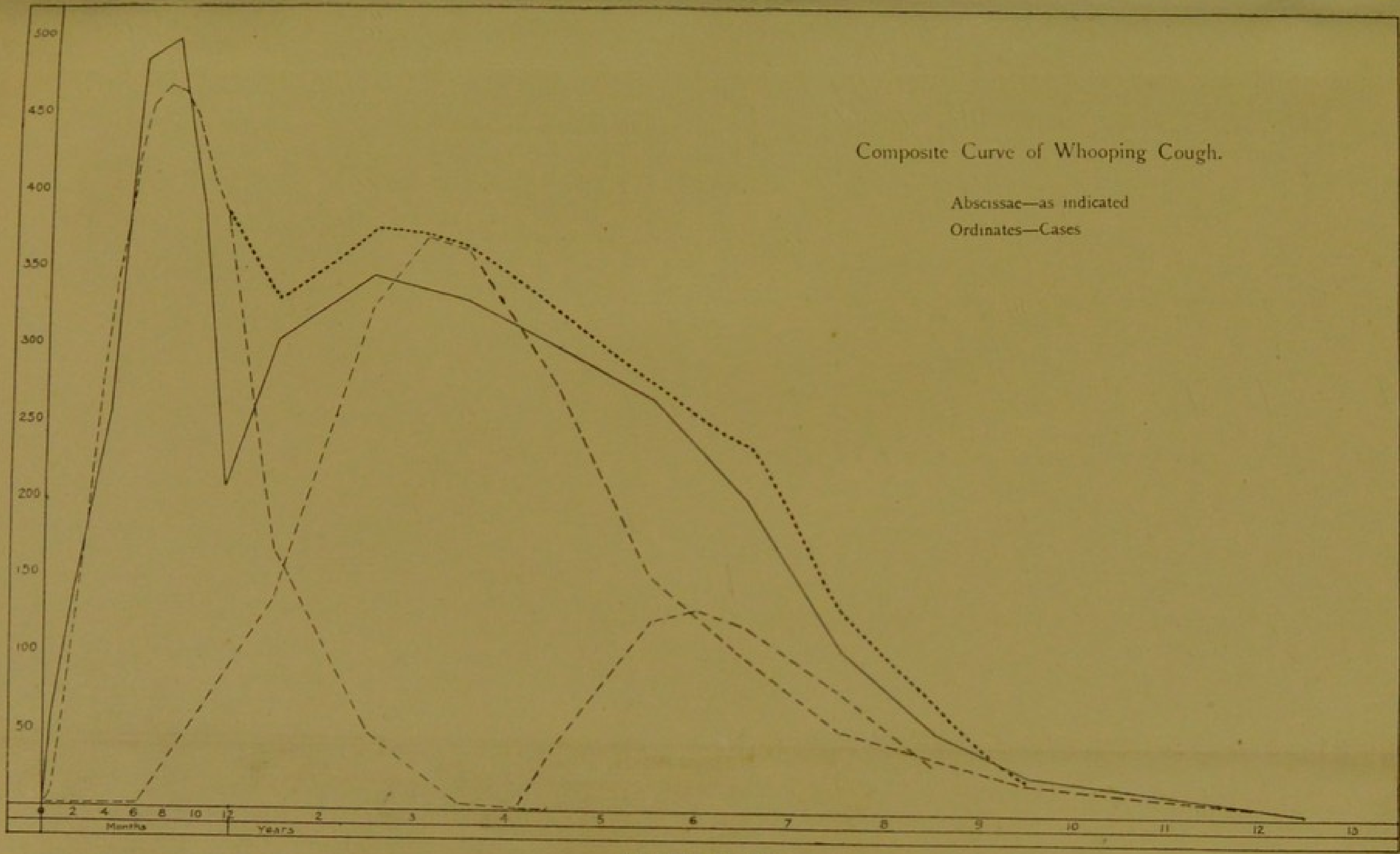

Typhus Fever (Males) Glasgow.
y = 240·1 (1 + x/5·4617)^1·6374 e^−x× 2998
Origin 0 years. Unit 2·5 years.
Ordinates—Cases: Abscissae—Years.
Theoretical.
Actual.
Number of Cases
Years

DIAGRAM II.



Composite Curve of Whooping Cough.
Abscissae—as indicated
Ordinates—Cases
500
450
400
350
300
250
200
150
100
50
2 4 6 8 10 12
Months
2 3 4 5 6 7 8 9 10 11 12 13
Years

DIAGRAM III.