CONTRIBUTIONS TO THE MATHEMATICAL THEORY OF EPIDEMICS

IV. ANALYSIS OF EXPERIMENTAL EPIDEMICS OF THE VIRUS DISEASE MOUSE ECTROMELIA

By W. O. KERMACK and A. G. McKENDRICK

From the Laboratory of the Royal College of Physicians, Edinburgh

(With 4 Figures in the Text)

During recent years we have been engaged in the study of the mathematical theory of the spread of an infectious disease in a community of susceptible individuals, and some of the earlier results have been published in a series of papers (Kermack & McKendrick, 1927, 1932, 1933, 1936). We have assumed that the recovery rate, death-rate, infection rate, etc., are represented by functions of the period of the disease, and these in particular cases may be constants. The results obtained even in the absence of definite *a priori* knowledge of the values of these functions or constants allow us to reach certain qualitative conclusions and to interpret qualitative observations. So far, however, it has been difficult to apply the theory in a quantitative sense, chiefly because of the absence of satisfactory data referring to actual diseases. Such data might refer either to human or animal communities under ordinary non-experimental conditions, or to the results of animal experiments under strict control. The data referring to animal and human communities are usually meagre and far from homogeneous. On the other hand the experimental study of epidemics is still in its infancy and it is only recently that the pioneer work of Greenwood *et al.* (1936) has made available a certain amount of experimental data which though limited in range is of the highest value. It is obviously desirable to find out whether the theory as developed in our previous communications can be applied to these data, and if so to work out the results in a quantitative sense.

The mathematical theory which we have developed accommodates the experimental fact that an individual who has recovered from a disease is frequently found to be relatively immune. In the special case where the *immunity is complete so that no second infection ever occurs the theory proves to be comparatively simple. This is especially so when we can assume that the* community does not reproduce itself by birth, but is recruited solely by immigration. It happens that one of the cases investigated experimentally by Greenwood *et al.* (1936) is of this simple type. This is the epidemic of the virus disease ectromelia in mice which is fully reported in a memoir by the

https://doi.org/10.1017/S0022172400034902 Published online by Cambridge University Press

above authors, to which the reader is referred for details. Two separate epidemics are actually reported upon, and referred to as ectromelia 1 and ectromelia 2. The purpose of the present paper is to discuss the data derived from these two experiments in the light of our general mathematical theory.

First of all we may recall one or two quantitative results of the theory which are found to be borne out by the experimental observations. These results are dependent only upon assumptions of a very general kind. A steady state condition in the absence of births is possible only when immigration is maintained at a rate in excess of a certain minimum which we call the *threshold* value; when it exceeds this threshold a steady state condition is always possible. It does not however follow at once that a steady state necessarily exists. Small deviations inevitably occur, and these may either die out or result in the system deviating permanently from the steady state condition. It is only when the first alternative holds that the steady state can actually exist. It follows that before we can predict that a steady state is really possible, it is necessary to determine whether or not the theoretical steady state is stable in the above sense.

In a recent paper (1936) we have found the condition which must be satisfied in order that the steady state should be stable. In the particular case now under discussion (no births, complete immunity after an infection), the condition is that all the real parts of the roots of the following equation should be negative:

$$1=\frac{\bar{K}_\alpha}{\bar{K}}\frac{\alpha+\bar{\pi}}{\alpha+\bar{\pi}+\bar{K}V}. \qquad \ldots\ldots(1)$$

(For the meaning of the symbols employed see Kermack & McKendrick (1933, 1936).) We note that for all values of α for which the real part $\geqslant 0$,

$$\left|\frac{\bar{K}_\alpha}{\bar{K}}\right|\leqslant 1, \quad \left|\frac{\alpha+\bar{\pi}}{\alpha+\bar{\pi}+\bar{K}V}\right|<1,$$

it is easily seen that the above equation cannot be satisfied by any value of α for which $R(\alpha)\geqslant 0$. Thus the condition for stability is satisfied. It follows that in the case of an epidemic of ectromelia or a similar disease a steady state is possible provided that the immigration rate exceeds a certain value. Such a steady state does in fact seem to have been reached in the case of ectromelia 1 during the period 1. iii. 32–31. viii. 32, and we shall therefore now consider the application of the theory to the steady state condition during this period.

In our present state of knowledge and with the available data the quantitative application of the general theory involving variable coefficients appears to be impossible. It is of the greatest interest and importance to find out from the experimental data just how far the simplified form of the theory, in which these rates are assumed to be independent of the period of the disease, gives an adequate account of an epidemic. We shall therefore attempt to fit the data to the theory in this form, that is, we shall assume a constant infectivity rate $\bar{k}$, a specific death rate d, and a recovery rate l.

In the Medical Research Council memoir the chance of surviving at

https://doi.org/10.1017/S0022172400034902 Published online by Cambridge University Press

different cage ages (denoted by l_x as is customary in actuarial practice) is tabulated for various periods of the epidemic, one of which fortunately covers the steady state period 1. iii. 32–31. viii. 32.

Here l_x gives the chance of a mouse introduced into the cage being still alive on the xth day. In order to calculate l_x in terms of the above constants it is necessary to recognize that a certain number of mice die from non-specific causes. As the mice when introduced into the cage were usually young mice not varying greatly in age, it is natural to assume that the chance of a mouse dying from non-specific causes can be adequately represented as a function of the cage age, henceforth denoted by η, in harmony with our previous notation. The mice within the cage are of three types: the uninfected $\bar{x}$, those suffering from ectromelia y, and the recovered x.

It is readily found that if $\bar{X}$, Y and X are the steady state values of $\bar{x}$, y and x, and if $\bar{X}_\eta$, Y_η and X_η denote the chances of a mouse introduced into the cage being alive on the ηth day as an uninfected, an infected, or a recovered animal, the following equations hold:

$$\left.\begin{aligned} \frac{d\bar{X}_\eta}{d\eta} &= -\lambda_2\bar{X}_\eta - \rho_\eta\bar{X}_\eta, \\ \frac{dY_\eta}{d\eta} &= \lambda_2\bar{X}_\eta - \lambda_1 Y_\eta - \rho_\eta Y_\eta, \\ \frac{dX_\eta}{d\eta} &= c_2\lambda_1 Y_\eta - \rho_\eta X_\eta, \end{aligned}\right\} \qquad \text{.....(2)}$$

where $\lambda_1 = d+l$, $\lambda_2 = \bar{k}Y$, $c_2 = l/\lambda_1$, and ρ_η is the non-specific death-rate.

Let $\bar{X}_\eta = \bar{X}_\eta' e^{-R_\eta}$, $Y_\eta = Y_\eta' e^{-R_\eta}$ and $X_\eta = X_\eta' e^{-R_\eta}$,

where $R_\eta = \int_0^\eta \rho_\theta\, d\theta$.

It is easily shown that

$$\left.\begin{aligned} \frac{d\bar{X}'}{d\eta} &= -\lambda_2\bar{X}', \\ \frac{dY'}{d\eta} &= \lambda_2\bar{X}' - \lambda_1 Y', \\ \frac{dX'}{d\eta} &= c_2\lambda_1 Y', \end{aligned}\right\} \qquad \text{.....(3)}$$

where the suffices η have for convenience been dropped.

If we note that $\bar{X}'_{\eta=0} = 1$, and $Y'_{\eta=0} = X'_{\eta=0} = 0$, it is readily shown that the above equations lead to the solution

$$\left.\begin{aligned} \bar{X}_1 &= e^{-\lambda_2\eta}, \\ Y' &= \frac{\lambda_2}{\lambda_2-\lambda_1}\,(e^{-\lambda_1\eta} - e^{-\lambda_2\eta}), \\ X' &= c_2\left[\frac{1}{\lambda_2-\lambda_1}\,(\lambda_1 e^{-\lambda_2\eta} - \lambda_2 e^{-\lambda_1\eta}) + 1\right]. \end{aligned}\right\} \qquad \text{.....(4)}$$

https://doi.org/10.1017/S0022172400034902 Published online by Cambridge University Press

If N_η is the chance of an animal being alive on the ηth day of cage life (this is the same as the l_x referred to above) then

$$\begin{aligned} N_\eta &= \bar{X}_\eta + Y_\eta + X_\eta \\ &= (\bar{X}' + Y' + X')\, e^{-R_\eta} \\ &= \left\{ \frac{c_1}{\lambda_2 - \lambda_1} (\lambda_2 e^{-\lambda_1 \eta} - \lambda_1 e^{-\lambda_2 \eta}) + c_2 \right\} e^{-R_\eta}, \qquad \text{.....(5)} \end{aligned}$$

where $$c_1 = 1 - c_2 = d/\lambda_1.$$

In the special case where $\lambda_2 = \lambda_1 = \lambda$, it is not difficult to show that these equations become

$$\left.\begin{aligned} \bar{X}' &= e^{-\lambda\eta}, \\ Y' &= \lambda\eta e^{-\lambda\eta}, \\ X' &= c_2 \{1 - (1 + \lambda\eta) e^{-\lambda\eta}\}, \end{aligned}\right\} \qquad \text{.....(6)}$$

and finally $$N\eta = \{c_1 (1 + \lambda\eta) e^{-\lambda\eta} + c_2\} e^{-R_\eta}. \qquad \text{.....(7)}$$

I. ECTROMELIA 1

In order to test our theory it is necessary to fit equation (5) to the values of l_x experimentally found. Theoretically the best method of carrying this out would be difficult and laborious, but the following semi-empirical procedure has been successfully adopted.

First of all we form an approximate estimate of λ_1 and λ_2. For this purpose we utilize the data provided by the mice which were observed in isolation for several weeks after the ectromelia 1 experiment was broken up. The following figures, derived from the table on p. 103 of the memoir, afford a maximum estimate of the chance of infection of an animal after it has been in the cage for a given length of time:

Cage age in days	Fraction uninfected
1–10	0·400
11–20	0·166
21–30	0·166
31–40	0·100
41–50	0
51–60	0·033
61–70	0·100
71–80	0·033
	0·998

It is easily shown that this fraction should be represented by the expression $e^{-\lambda_2\eta}$, as $\lambda_2\ (=\bar{k}Y)$ is the chance of infection of an individual mouse per unit time, which it is convenient to take as one day. We note that

$$\int_0^\infty e^{-\lambda_2\eta} d\eta = 1/\lambda_2,$$

so that we obtain a rough estimate of $1/\lambda_2$ by adding up the ordinates. This gives 0·998 and as the interval is 10 days, we obtain $1/\lambda_2 \doteqdot 10 \times 0{\cdot}998 = 9{\cdot}98$, hence λ_2 is probably of the order of 0·1. We are also given the time of death

https://doi.org/10.1017/S0022172400034902 Published online by Cambridge University Press

of those of the fifteen animals, introduced into the cage during the 5 days before the experiment was stopped, which died of ectromelia. The actual number of those surviving in an infected state after 7 days would be proportional to $e^{-\lambda_1 \eta}$, so that the chance of death on ηth day would be $de^{-\lambda_1 \eta}$, whence

$$\frac{1}{\lambda_1} = \frac{\int_0^\infty \eta de^{-\lambda_1 \eta} d\eta}{\int_0^\infty de^{-\lambda_1 \eta} d\eta} = \text{average number of days survived.} \quad \text{.....(8)}$$

The figures are as follows: 9, 10, 7, 10, 14, 9 and 9, whence $1/\lambda_1 \doteqdot 9{\cdot}7$ days, so that λ_1 is also of the order 0·1. It follows that for values of η greater than 100 the coefficient of c_1 becomes very small, and so for values of η greater than 250 it is quite safe to use the expression

$$N_\eta = c_2 e^{-R_\eta}.$$

We therefore proceed to fit the tail of the l_x curve to this expression. On examination it is found that for the range 150–500, the non-specific death-rate is proportional to η, so that $R_\eta \left(= \int_0^\eta \rho_\eta d\eta \right)$ may be assumed to be of the form $r\eta^2$, where r is a constant. Evaluation of the constants c_2 and r from these observed figures leads to the values $c_2 = 0{\cdot}2$ and $r = 0{\cdot}000012$. We now determine the other two constants λ_1 and λ_2. Theoretically this may be done by the method of moments, the zero and first moments being here sufficient. We find

$$N^0 = \int_0^\infty N_\eta d\eta = \frac{c_2}{2}\sqrt{\frac{\pi}{r}} + \frac{c_1}{\sqrt{2r}\,(\zeta_2 - \zeta_1)} (\zeta_2 \Delta_{\zeta_1} - \zeta_1 \Delta_{\zeta_2}), \quad \text{.....(9)}$$

where

$$\Delta_\zeta = e^{\zeta^2/2} \int_\zeta^\infty e^{-z^2/2}\, dz, \quad \zeta_1 = \frac{\lambda_1}{\sqrt{2r}} \text{ and } \zeta_2 = \frac{\lambda_2}{\sqrt{2r}}; \quad \text{.....(10)}$$

and

$$N^1 = \int_0^\infty \eta N_\eta d\eta = \frac{1}{2r} - \frac{c_1}{2r} \frac{\zeta_1 \zeta_2}{\zeta_2 - \zeta_1} (\Delta_{\zeta_1} - \Delta_{\zeta_2}). \quad \text{.....(11)}$$

In practice it is found that the determination of λ_1 and λ_2 from these two equations is laborious and awkward, one difficulty being that the values obtained are highly sensitive to small changes in N^0 and N^1. It is found more satisfactory to assume in the first place that $\lambda_1 = \lambda_2$ as suggested by the rough evaluations already given, and to apply the N^0 equation which in this case reduces to

$$N^0 = \frac{c_2}{2}\sqrt{\frac{\pi}{r}} + \frac{c_1}{\sqrt{2r}} \{(1 - \zeta^2)\, \Delta_\zeta + \zeta\}, \quad \text{.....(12)}$$

or to a first approximation

$$N^0 = \frac{c_2}{2}\sqrt{\frac{\pi}{r}} + \frac{2c_1}{\zeta\sqrt{2r}}, \quad \text{.....(13)}$$

where $\zeta = \zeta_1 = \zeta_2$.

Whence

$$\lambda = \zeta\sqrt{2r} \doteqdot \frac{2c_1}{N^0 - \frac{c_2}{2}\sqrt{\frac{\pi}{r}}}. \quad \text{.....(14)}$$

https://doi.org/10.1017/S0022172400034902 Published online by Cambridge University Press

In order to apply this formula it is necessary to obtain the value N^0 (which is readily shown to be the expectation of life at cage age zero). This we have done by evaluating the following sum

$$10\,(\tfrac{1}{2}l_0+l_{10}+l_{20}+\ldots+l_{190}+l_{200})$$

and then adding on $\int_{205}^{\infty} c_2 e^{-rx^2}\,dx$. This seems preferable to the more obvious method of adding the l_x's over the whole series, not only because it avoids considerable labour but also because it makes certain that the result will not be disturbed by errors arising from the curtailment of the series.

We thus find $N^0=81{\cdot}295$ a value not far removed from that given in the memoir for the expectation—namely 81·87, whence $\lambda=0{\cdot}05$. If we now calculate the values of N_η from equation (7) we find that the calculated figures are too high at 20 and too low at 50 and 60. This suggests that λ_1 and λ_2 instead of being taken as equal should differ somewhat in value. As a guide, we notice that the equation gives approximately

$$\frac{\sqrt{2r}N^0-c_2\sqrt{\frac{\pi}{2}}}{c_1}\doteqdot\frac{1}{\zeta_1}+\frac{1}{\zeta_2},\quad\text{whence}\quad\frac{1}{\zeta_1}+\frac{1}{\zeta_2}\doteqdot\frac{2}{\zeta}. \qquad \ldots\ldots(15)$$

We therefore try pairs of values, λ_1 and λ_2, such that their harmonic mean is approximately 0·05. After some trials it was found that a very good fit was given by $\lambda_1=0{\cdot}039$ and $\lambda_2=0{\cdot}09$. Comparatively small deviations from those values resulted in the fit being noticeably less good.

Comparison of the observed and calculated figures is shown in Table I and Figs. 1 and 2. It is appropriate to apply the χ^2 test to the d_x figures, that is the number of deaths occurring at different cage ages, as these, unlike the l_x's, are not correlated with their neighbours. To do this, however, they must all be multiplied by the total number of mice involved, namely 552. We find $\chi^2=34{\cdot}22$. There are thirty-one different observations, and four constants have been employed in fitting the curve so that we take $n'=31-4+1$, whence P is 0·16. Actually the greatest contributions to χ^2 come from the two periods 0–10 and 10–20 days. During the first 10 days l_x decreases less rapidly than the theory predicts, and this is quite clearly due to the fact that the disease has an incubation period of several days, so that there is a deficiency of deaths during the first few days of cage life. If we take the interval 0–20 days as one period we find $\chi^2=18{\cdot}56$, $n'=27$, $P=0{\cdot}85$. The fit is evidently an unusually good one.

We note that the equation (5) for N_η is symmetrical in λ_1 and λ_2. It follows that it is impossible from the above data alone to distinguish between these two quantities and to say which value is to be attributed to $d+l$, and which to $\bar{k}Y$. We have therefore two possibilities:

$$(a)\ \lambda_1=0{\cdot}039,\ \lambda_2=0{\cdot}09,$$
$$(b)\ \lambda_1=0{\cdot}09,\quad\lambda_2=0{\cdot}039.$$

https://doi.org/10.1017/S0022172400034902 Published online by Cambridge University Press

Table I

Ectromelia 1. Period 1. iii. 32–31. viii. 32

Cage age in days	l_x		d_x	
	Experimental	Theoretical	Experimental	Theoretical
0	10000	10000	489	939
10	9511	9061	2076	1636
20	7435	7425	1362	1579
30	6073	5846	1390	1136
40	4683	4710	807	886
50	3876	3824	571	633
60	3305	3191	477	448
70	2828	2743	358	316
80	2470	2427	218	230
90	2252	2197	142	179
100	2110	2018	174	121
110	1936	1897	146	105
120	1790	1792	51	87
130	1739	1705	63	77
140	1676	1628	92	70
150	1584	1558	45	67
160	1539	1491	92	64
170	1447	1427	16	62
180	1431	1365	34	62
190	1397	1303	35	62
200	1362	1241	71	61
210	1291	1180	37	60
220	1254	1121	55	60
230	1199	1060	74	58
240	1125	1002	39	57
250	1086	945	318	266
300	768	679	285	219
350	483	460	240	167
400	243	293	114	117
450	129	176	44	77
500	85	100	28	47
550	57	53	—	—

It is easy to show that in a steady state

$$\left.\begin{aligned} \bar{X} &= m\int_0^\infty \bar{X}_\eta d\eta = \frac{m}{\sqrt{2r}}\Delta_{\zeta_1}, \\ Y &= m\int_0^\infty Y_\eta d\eta = \frac{m}{\sqrt{2r}}\frac{\zeta_1}{\zeta_2-\zeta_1}(\Delta_{\zeta_1}-\Delta_{\zeta_2}), \\ X &= m\int_0^\infty X_\eta d\eta = mc_2\left\{\frac{1}{2}\sqrt{\frac{\pi}{r}}-\frac{1}{\sqrt{2r}(\zeta_2-\zeta_1)}(\zeta_2\Delta_{\zeta_1}-\zeta_1\Delta_{\zeta_2})\right\}, \end{aligned}\right\} \quad \text{.....(16)}$$

and the total population

$$N = \bar{X}+Y+X = m\left\{\frac{c_2}{2}+\frac{c_1}{\zeta_2-\zeta_1}(\zeta_2\Delta_{\zeta_1}-\zeta_1\Delta_{\zeta_2})\right\}. \quad \text{.....(17)}$$

Inserting the values of λ_1, λ_2, m, c_1, c_2 and r in these equations we have for case (*a*) $\bar{X}=33{\cdot}72$, $Y=73{\cdot}95$, $X=131{\cdot}82$, and $N=239{\cdot}49$, whence $\bar{k}=\lambda_2/Y=0{\cdot}00122$; also $c_1=0{\cdot}8$, $c_2=0{\cdot}2$, and so $l=0{\cdot}0078$ and $d=0{\cdot}0312$: and for case (*b*) $\bar{X}=75{\cdot}92$, $Y=31{\cdot}95$, $X=131{\cdot}82$, and $N=239{\cdot}49$, whence $\bar{k}=0{\cdot}00122$; also $c_1=0{\cdot}08$, $c_2=0{\cdot}02$, and so $l=0{\cdot}019$ and $d=0{\cdot}072$.

It is interesting to note that the number uninfected in case (*a*) is approximately equal to the number who were ill in case (*b*) and vice versa. The

https://doi.org/10.1017/S0022172400034902 Published online by Cambridge University Press

number who recover and also the infectivity rate $\bar{k}$ are unaltered. These results follow at once from the approximate formulae

$$\bar{X}=\frac{m}{\lambda_2},\quad Y=\frac{m}{\lambda_1}\quad\text{and}\quad X=mc_2\left(\frac{1}{2}\sqrt{\frac{\pi}{r}}-\frac{1}{\lambda_1}-\frac{1}{\lambda_2}\right),\qquad\text{.....(18)}$$

and

$$\bar{k}\text{ is approximately equal to }\frac{\lambda_1\lambda_2}{m},\qquad\text{.....(19)}$$

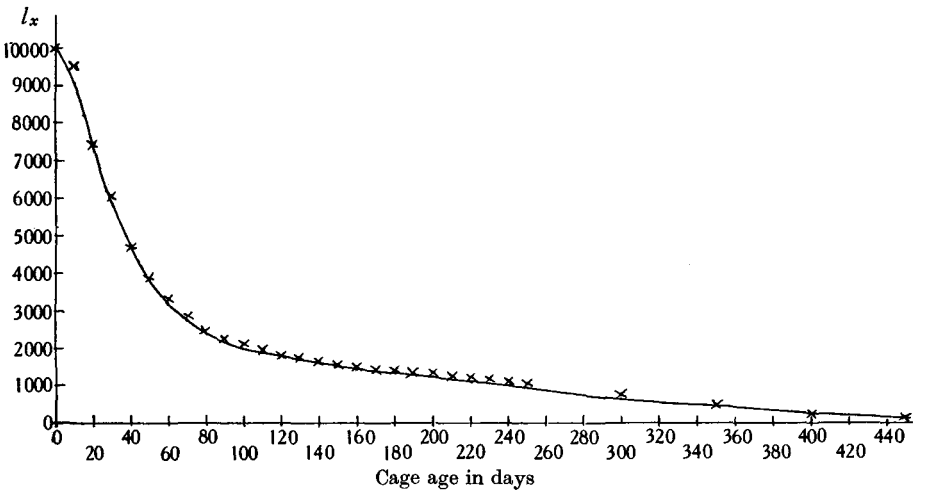

Fig. 1. l_x. Ectromelia 1. 1. iii. 32–31. viii. 32.

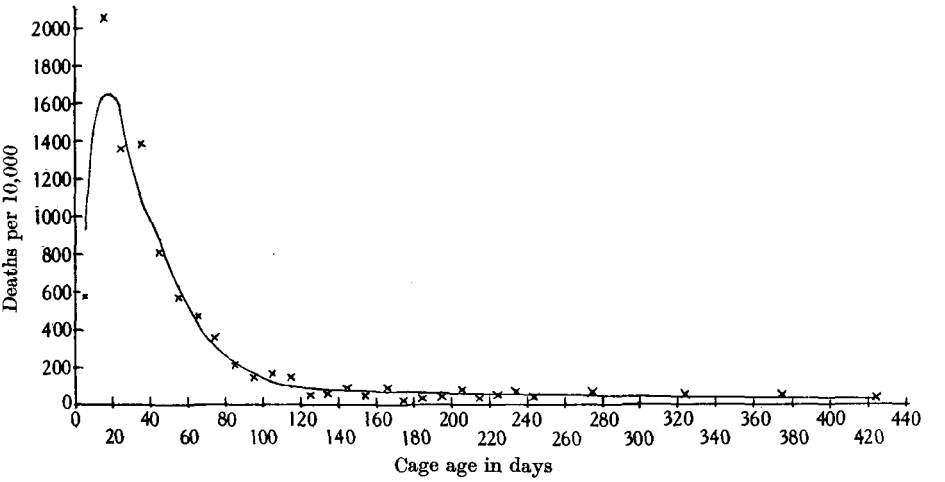

Fig. 2. d_x. Ectromelia 1. 1. iii. 32–31. viii. 32.

which are readily deduced if we replace Δ_ζ by its approximate value $1/\zeta$. The total number N (which in a steady state is m times the expectation N^0) is 239·49 in both cases. This is in good agreement with the observed value which during the period fluctuated between 225 and 254 and averaged 243·1. This

https://doi.org/10.1017/S0022172400034902 Published online by Cambridge University Press

agreement indicates that the population during this interval had in fact attained a steady state condition. It may readily be shown that $\frac{1}{\lambda_1}+\frac{1}{\lambda_2}$ gives to a first approximation the average cage age at death of those who die of the specific disease. This is a quantity which could be easily observed over any given period of time without disturbing the experiment. It does not however appear to be recorded in the memoir, but if available it would form a valuable check both on the above calculations, and on the constancy of the conditions of the experiment from time to time. One would expect that this quantity would be much less subject to error than $1/\lambda_1$ or $1/\lambda_2$ separately. It is evident in the above fitting of the life table that the values of λ_1 and λ_2, as deduced, are sensitive to relatively small changes in the slope of the curve during the period 20–80 days. One cannot therefore attach very great confidence to the accuracy of the separate values which have been obtained, but the quantity $\frac{1}{\lambda_1}+\frac{1}{\lambda_2}$ is much less dependent on chance irregularities in the curve. Of course data from a much larger epidemic carried on in a steady state over a longer period would allow of much more precise determination of both λ_1 and λ_2 separately. Moreover, we could, in theory at least, determine which value referred to λ_1, and which to λ_2, because the average cage age at death of the sick who died of ectromelia is symmetrical in λ_1 and λ_2 only to a first approximation, and therefore its exact value would afford a criterion for distinguishing the two possibilities, provided that the data were sufficiently accurate. The problem would be most simply settled by interrupting the epidemic during the steady state condition and observing the individual animals in isolation. In the case of ectromelia 1, the experiment was actually interrupted in February 1934, and the remaining mice, 110 in number, were isolated in separate cages and kept under observation. It is upon the data so obtained that the estimates of λ_1 and λ_2 on pp. 175 and 176 were based. However, in reference to the steady state period, these can only be regarded as very rough estimates, for evidently a very profound disturbance was introduced—presumably by the heat wave in August 1932—and the conditions obtaining in February 1934 were different from those present during the steady state period. The total number, for example, had fallen from about 250 to about 100. It is for this reason that we have employed these estimates only as a rough guide.

II. Fluctuations about steady state level

In discussing the fluctuations about the steady state level, it is sufficiently approximate to take the non-specific death-rate as constant and equal to the mean non-specific death-rate actually observed in the steady state condition over all cage ages. This procedure is justified by the fact that the non-specific death-rate is in any case small compared with either λ_1 or λ_2. The mean value in question, ρ, is readily shown to be approximately $2\sqrt{\frac{\bar{r}}{\pi}}$, i.e. about

https://doi.org/10.1017/S0022172400034902 Published online by Cambridge University Press

0·004. This is of the order of a tenth of the values of λ_1 and λ_2. It is not difficult to show that the course of a small disturbance about the steady state is given by the expression

$$\bar{x} = \bar{X} + A_1 e^{\alpha_1 t} + A_2 e^{\alpha_2 t} \quad \text{and} \quad y = Y + B_1 e^{\alpha_1 t} + B_2 e^{\alpha_2 t},$$

where A_1, A_2, B_1 and B_2 are constants adjusted in conformity with the initial disturbance, and α_1 and α_2 are the two roots of the quadratic

$$\alpha^2 + (\lambda_2 + \rho)\,\alpha + (\lambda_1 + \rho)\,\lambda_2 = 0. \qquad \ldots\ldots(20)$$

(It may easily be shown that equation (1) reduces to this quadratic in the case of constant coefficients.) In this quadratic it is sufficiently approximate to put $\rho = 0$. The period is then $\dfrac{4\pi}{\sqrt{4\lambda_1\lambda_2 - \lambda_2^2}}$, and the damping factor is $\lambda_2/2$. In the two cases considered above, the periods are 112·4 and 163·0 days and the damping factors 0·0195 and 0·045 respectively. These both represent highly damped vibrations, the amplitude in the course of one period being reduced to one-eighth and one-thousandth respectively of its original value.

It is interesting to examine the diagram given on p. 71 of the memoir in the light of this result. The graph shows the fluctuations in the 6-day death-rates of mice dying after a herd experience of 0–30 days, and on the same diagram are given the 6-monthly death-rates from measles of children from 0–5 years of age. Both curves exhibit recurring maxima, the ectromelia death-rates showing about 49 in 990 days and the measles 43 in 70 years. This gives for ectromelia an average period of about 20 days which is very much shorter than the theoretical period of either 112 or 163 days obtained above. Clearly the fluctuations shown in the figure are not to be explained as the intrinsic oscillations of the system. Not only does the observed period differ from the theoretical but the intrinsic oscillations would be highly damped and difficult to detect.

We think there is no doubt that the ectromelia fluctuations shown in the figure are essentially random in character, and that the relatively large amplitudes observed are a result of the smallness of the numbers involved. It can be shown that a random series satisfies certain criteria; thus, for example, on the average every third observation would be maximal, that is, greater than those immediately preceding and following. This question will be discussed more fully in a subsequent publication, but it may be stated here that the ectromelia death-rates, though possibly not completely random, are sufficiently nearly so, to enable us to exclude any periodic oscillation. This is in agreement with the views of Greenwood *et al.* (1936).

III. ECTROMELIA 2

It is natural to enquire whether similar results are obtained with the epidemic ectromelia 2. This experiment was carried out practically simultaneously with ectromelia 1, and is stated to have run a very similar course. Like ectromelia 1, ectromelia 2 would seem to have reached an approximately steady state between 1. iv. 32 and 1. viii. 32, but unfortunately the corresponding life table has been calculated not for this period but for the period

1. v. 32 and 20. x. 32; consequently the table includes the epoch of the heat-wave which profoundly disturbed the course of the epidemic. In fact the population decreased from about 350 to about 250 in approximately 8 weeks. It is not surprising therefore that an attempt to apply the theory to this life table does not lead to a satisfactory result. We find $r = 0{\cdot}000012$, $c_2 = 0{\cdot}166$ and $\frac{1}{\lambda_1} + \frac{1}{\lambda_2}$ approximately equal to 27·4 days, but we have not been able to arrive at two values of λ_1 and λ_2 which make the N_η curve agree at all closely with the observed l_x values. We therefore directed our attention to the preceding period 1. xi. 31 to 30. iv. 32. This period was evidently not one in which the population was in a steady state condition as the numbers increased from 200 to 310 during the 6 months, yet there is no indication of any abrupt change in the conditions during the period. When the l_x curve was examined by the method already described in detail it was found that $r = 0{\cdot}00000609$, $c_2 = 0{\cdot}3507$, $\frac{1}{\lambda_1} + \frac{1}{\lambda_2} = 22$ days, whence if $\lambda_1 = \lambda_2 = \lambda$, we find $\lambda = 0{\cdot}091$.

Calculating the values of N_η from the equation

$$N_\eta = \{c_1 (1 + \lambda\eta)\, e^{-\lambda\eta} + c_2\}\, e^{-r\eta^2}, \qquad \text{.....(21)}$$

we obtain the following figures for N_η $(= l_x)$, and for the deaths at the various cage ages d_x (Table II and Figs. 3 and 4). As before, we may compare the calculated and observed d_x's by the χ^2 test. The total number involved was

Table II

Ectromelia 2. Period 1. xi. 31–30. iv. 32

Cage age in days	l_x		d_x	
	Experimental	Theoretical	Experimental	Theoretical
0	10000	10000	678	1506
10	9322	8494	2843	2034
20	6479	6460	1662	1401
30	4817	5059	600	803
40	4217	4256	305	417
50	3912	3839	282	232
60	3630	3607	117	107
70	3513	3500	162	91
80	3351	3409	123	55
90	3228	3354	0	48
100	3228	3306	69	45
110	3159	3261	23	48
120	3136	3213	0	49
130	3136	3164	58	52
140	3078	3112	32	54
150	3046	3058	161	57
160	2885	3001	144	60
170	2741	2941	113	62
180	2628	2879	38	64
190	2590	2815	38	66
200	2552	2749	0	68
210	2552	2681	36	70
220	2516	2611	0	71
230	2516	2541	37	72
240	2479	2469	73	72
250	2406	2397	353	370
300	2054	2027	450	364
350	1603	1663	266	339
400	1337	1324	307	302
450	1030	1022	—	—

https://doi.org/10.1017/S0022172400034902 Published online by Cambridge University Press

543, and taking the period 0–20 as one we find $\chi^2 = 54{\cdot}0$, $n' = 28$, $P = 0{\cdot}004$. This value of P is not satisfactory but it is to be noted that a large contribution comes from the portion of the curve between 150–250 days. In fact the

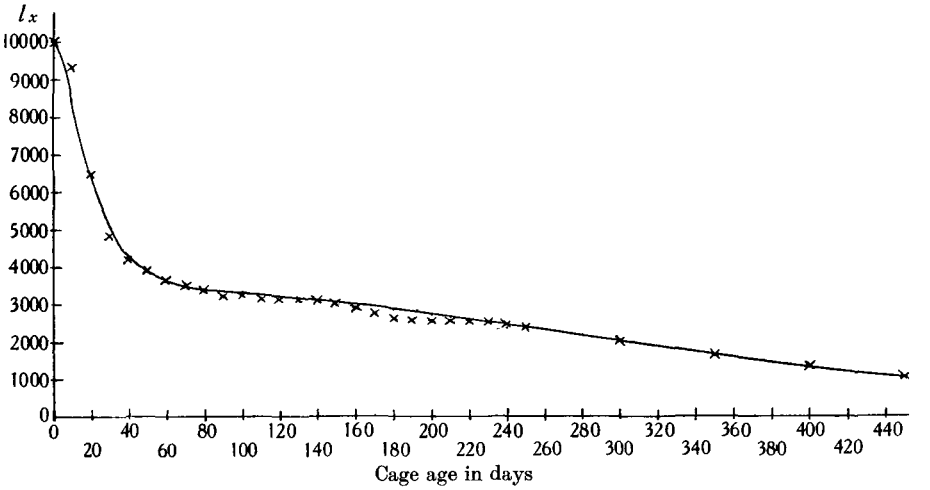

Fig. 3. l_x. Ectromelia 2. 1. xi. 31–30. iv. 32.

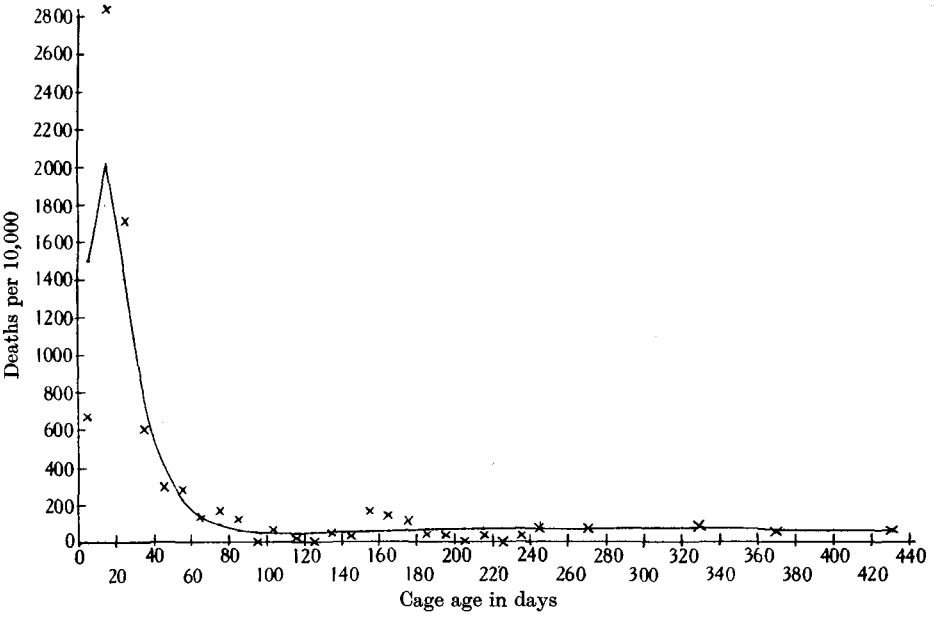

Fig. 4. d_x. Ectromelia 2. 1. xi. 31–30. iv. 32.

observed d_x figures over this period rise to a maximum and fall again in a way which it seems quite impossible to accommodate on any simple theory. It is probable that some disturbing factor must have intruded; if we omit this period and consider the other 300 days we find $\chi^2 = 25{\cdot}2$, $n' = 18$, $P = 0{\cdot}09$. The

N_η curve as shown in Fig. 3 closely follows the observed l_x curve and it does not seem probable that substantially better agreement would be obtained by taking unequal values for λ_1 and λ_2. The values $\lambda_1 = \lambda_2 = 0{\cdot}091$ give $\bar{X} = 32{\cdot}66$, $Y = 32{\cdot}82$, $X = 353{\cdot}04$, $N = 418{\cdot}52$, $\bar{k} = 0{\cdot}00277$, $l = 0{\cdot}0391$ and $d = 0{\cdot}0591$. The theoretical period is 79·4 days and the damping factor 0·045. It will be noticed that the value of N is much greater than the observed value (210 rising to 310). We have already pointed out that the theoretical value of N is in fact m times the expectation of life at zero cage age. This result holds generally in a steady state condition and is not dependent on a particular epidemiological theory. It follows that an expectation of life of 140 days necessarily leads to $N = 402$ days. (The expectation of life given in the memoir is 125·8 days, and this would correspond to $N = 377{\cdot}4$ days. The discrepancy between the two estimates is probably dependent upon different methods of treating the higher cage ages.) Whichever estimate be adopted, the resulting value of N is above the actual number of the population at any time within the period. This is obviously due to the circumstance that the experiment had not been continued for a sufficiently long time for a steady state level to establish itself, and this is borne out by the fact that the numbers continued to mount steadily throughout the period.

Discussion

In applying the above theory to the observed data we have found it desirable to treat the first 20 days of cage life as a single unit, because it is obvious from the trend of the l_x figures that the fall during the first 10 days is unduly slow. It is fairly certain that this discrepancy between observation and theory arises from the fact that the disease has a latent or incubation period of 4 or 5 days. It is thus impossible for specific deaths to occur during the first 4–5 days of cage life. The deaths which ought to have occurred according to theory during this period have been, as it were, postponed. It might have been better to have attempted to fit the curve from the fifth day onwards, but this procedure would appear to involve an undue degree of arbitrariness. Apart from the initial period the agreement between theory and observation in the case of ectromelia 1 is extremely satisfactory, whilst for ectromelia 2 it is as good as could be expected in view of the not altogether satisfactory nature of the data.

The memoir refers to the fact that in attempting to interpret epidemiological phenomena, the real difficulty is to be able to distinguish the factors that are really important. There is no doubt that in any actual epidemic the interplay of all the various effects is a highly complicated one. A complete theory which attempts to deal with the whole phenomena would be unworkable mathematically and too complicated to apply in practice. As is the case in practically all applied sciences, it is essential to guess or otherwise decide upon the really important factors and then to construct an idealized theory in which the factors which have been neglected may be subsequently

accommodated by *a posteriori* modifications. From this point of view it is a matter of considerable importance to find that the essential features of the cage experience of the individual mice can be adequately accounted for by the modification of the general theory in which the infectivity, specific death, and recovery rates are represented by constant coefficients. We do not suggest that the postulates of the theory which we have applied in the preceding pages are completely fulfilled in the experimental epidemics, we suggest however that they are approximately fulfilled. We find a satisfactory agreement on points which we can test; we therefore suggest that *other* conclusions from the deductive method which cannot be tested easily may also be approximately valid, and so we may be able to predict what *will* happen in cases which cannot be tested experimentally. These considerations justify the procedure of developing, at some length, the simplified theory based on constant coefficients and applying it to the explanation of the qualitative features of epidemics and even as far as may be to the calculation of the actual values of the coefficients in question.

It is one of the obvious criticisms of experimental work such as that embodied in the memoir that though exceedingly laborious and costly it nevertheless refers only to a particular species of animal and even then only to one or two diseases affecting the species. It is not at all immediately evident how far the results have any direct bearing on epidemics of human disease. It appears to us that this difficulty can only be met by developing the more theoretical side *pari passu* with the experimental, and analysing the latter in the light of the results obtained theoretically. It seems to us that possibly the most useful information of a general character which might be expected to accrue from the experimental work is the assistance it can give us in assessing the relative importance of the many complex factors. In particular the indication which it affords us of the substantial adequacy of the constant rate theory would seem to be most valuable. This result is all the more important because it is of a type which might be expected to hold generally, and not to be limited to the particular case.

We think it of interest to direct attention to the remarkable symmetry in respect of λ_1 and λ_2, which appears in the above results. This symmetry is in fact so complete up to first approximation that it has been found impossible to obtain from the life table any reliable criterion for distinguishing those constants from each other. Equation (19) gives $\frac{1}{\lambda_2}=\frac{\lambda_1}{m\bar{k}}=\frac{d+l}{m\bar{k}}$. Let us now imagine d and l to be fixed, then λ_2 will depend only on the product $m\bar{k}$. Thus the life table depending only upon λ_1 and λ_2 will remain unchanged if we vary m and $\bar{k}$ in such a way that their product remains constant. If now we consider this product as fixed we may interchange λ_1 and λ_2 without disturbing the life table. In other words for a constant $m\bar{k}$ there are two values of λ_1, that is, of $d+l$, which give the same life table. If we change from one value of $d+l$ to the other we approximately interchange the numbers of uninfected and sick,

the lower value of $d+l$ corresponding to a higher number of sick, in the steady state. However the average cage age of the sick at death remains approximately constant.

A few words may be added as to the significance of some of the constants which appear in the above theory. The deadliness of the disease is measured by $c_1 = \frac{d}{d+l}$, and not simply by d. The quantity λ_1, or $d+l$, represents as it were the rate of the disease process in the individual; in other words $1/\lambda_1$ measures the average period which elapses from the time when an individual becomes infected until the termination of his illness by either recovery or death. In comparing ectromelia 2 with ectromelia 1, it will be seen that in the former the general death-rate measured by r is lower, and that the deadliness of the disease process is less. This is in conformity with the statement in the memoir that "on the whole ectromelia 2 had a lower rate of mortality than ectromelia 1". On the other hand the time required for an individual entering the cage to become infected and then either to recover or die of the disease $\frac{1}{\lambda_1}+\frac{1}{\lambda_2}$ is shorter in the case of ectromelia 2. Ectromelia 2 would therefore

Table of numerical values

	Ectromelia 1 Period 1. iii. 32–31. viii. 32		Ectromelia 2 Period 1. xi. 31–30. iv. 32
	Case (*a*)	Case (*b*)	
λ_1	0·039	0·09	0·091
λ_2	0·09	0·039	0·091
c_1	0·8	0·8	0·6493
c_2	0·2	0·2	0·3507
$\bar{X}$	33·72	75·92	32·66
Y	73·95	31·95	32·82
X	131·82	131·82	353·04
N	239·49	239·49	418·52
$\bar{k}$	0·00122	0·00122	0·00277
l	0·0078	0·019	0·0391
d	0·0312	0·072	0·0591
N^0	79·83	79·83	139·51
$\frac{1}{\lambda_1}+\frac{1}{\lambda_2}$	36·5	36·5	22·0
T'	112·4	163·0	79·4
Damping factor r	0·0195 0·000012	0·045 0·000012	0·045 0·00000609

The chance that an individual may become infected by a particular infected individual in a day is $\bar{k}$.

The chance that an individual entering the cage may become infected in a day is $\bar{k}Y = \lambda_2$ (Dudley's "infection pressure").

The average period during which a mouse (ultimately infected) remains uninfected is therefore $1/\lambda_2$ days.

The average period during which an infected mouse carries the disease is approximately $\frac{1}{d+l}$, i.e. $1/\lambda_1$ days.

The deadliness of the disease, i.e. the chance that a mouse which becomes infected will die of the disease, is approximately $\frac{d}{d+l}$, that is, c_1, and that it will recover is approximately $\frac{l}{d+l}$, that is, c_2.

The average cage age at death of the mice which die of the disease is approximately $\frac{1}{\lambda_1}+\frac{1}{\lambda_2}$ days.

https://doi.org/10.1017/S0022172400034902 Published online by Cambridge University Press

appear to have been a more rapid but less lethal infection. These comparisons of course only refer to the periods actually analysed.

It is an obvious extension of the above work, to apply it to the other experimental epidemics dealt with in the memoir. These, however, refer to bacterial diseases in which presumably complete immunity is not conferred by an attack of the disease. Consequently equations (2) must be replaced by a somewhat more complicated set involving five unknown parameters instead of four, and the fitting of these to the data would be rather more difficult. However, it is hoped to deal with this problem in a subsequent communication.

SUMMARY

1. The experimental data obtained by Greenwood *et al.* (1936) relating to epidemics of the virus disease ectromelia in mice have been examined in the light of mathematical theory. Attention has been directed in particular to the life tables calculated from the observed data, which give the chance of survival and death after various periods of cage life.

2. The life table relating to ectromelia 1 during the steady state phase from 1. iii. 32 to 31. viii. 32 shows very close agreement over the range 0–550 days with that predicted by the theoretical equation which involves only four arbitrary constants. A slight discrepancy over the first few days is evidently due to the fact that representation of the death-rate and the recovery rate by constant coefficients does not accommodate an incubation period. The values of the constants obtained from the data give a measure of the essential characteristics of the epidemic.

3. General agreement, though not so complete, is also found when the theory is applied to ectromelia 2 during the period 1. xi. 31–30. iv. 32. During this phase however, although conditions were otherwise apparently uniform, a steady state had not actually been attained. On the other hand the equations do not apply so satisfactorily to the obviously inhomogeneous period 1. v. 32–20. x. 32.

4. In the present analysis the assumption of constant rates gives a satisfactory account of the progress of an infection in a susceptible community. This result suggests that for many purposes the assumption of constant rates may be adequate.

5. The short period fluctuations observed in the ectromelia epidemic were probably random in character and have no relationship to periodic fluctuations such, for example, as those detected by Brownlee in the case of measles.

REFERENCES

GREENWOOD, BRADFORD HILL, TOPLEY & WILSON (1936). *Medical Research Council Special Report* No. 209.
KERMACK & McKENDRICK (1927). *Proc. Roy. Soc.* A, **115**, 700.
—— —— (1932). *Ibid.* **138**, 55.
—— —— (1933). *Ibid.* **141**, 94.
—— —— (1936). *Proc. London Math. Soc.*, Series 2, **41**, 462.

(*MS. received for publication* 13. XI. 1936.—Ed.)