

CONTRIBUTIONS TO THE MATHEMATICAL  
THEORY OF EPIDEMICS

V. ANALYSIS OF EXPERIMENTAL EPIDEMICS OF MOUSE-  
TYPHOID; A BACTERIAL DISEASE CONFERRING  
INCOMPLETE IMMUNITY

BY W. O. KERMACK AND A. G. MCKENDRICK

*From the Laboratory of the Royal College of Physicians, Edinburgh*

(With 5 Figures in the Text)

In a recently published paper (Kermack & McKendrick, 1937) the observational data relating to epidemics of ectromelia in populations of mice maintained under experimental conditions (Greenwood *et al.* 1936) has been analysed in the light of a mathematical theory of epidemics developed by us during recent years (Kermack & McKendrick, 1927, 1932, 1933, 1936). It was shown that the life table giving the chance of mice surviving for various lengths of time in infected communities is very closely represented by a formula calculated on the assumption that the various rates—infection rate, recovery rate, death rate, etc.—are constants. It is, of course, realized that this simplifying assumption can only be regarded as approximately true. It renders the application of the general theory practicable, and the result of the investigation justifies its use, in so far as the theory so simplified does actually conform to the experimental results.

The ectromelia epidemics, however, are of rather special type in that they refer to a disease in which the immunity conferred by an attack is almost, if not quite, complete. The disease is also peculiar in having a very short incubation period. In the simplified theory with constant coefficients no allowance at all was made for an incubation period, and it was to be expected that in the case of a disease with a longer incubation period, some special method would have to be devised to accommodate it.

It was therefore especially desirable that an attempt should be made to extend the work already done in relation to the virus disease ectromelia to the case of a typical bacterial disease, in which immunity was not so complete, and the incubation period was somewhat longer. It was suggested by Prof. Greenwood that, in view of the relatively large numbers of mice involved, the epidemic denoted by the symbol  $B_6$  (Greenwood *et al.* 1930), or by  $A_6$  (Greenwood *et al.* 1936) would be a suitable experiment to analyse. The epidemic in question was one caused by *Bact. aertrycke*, in which observations were continued for 365 days, six fresh mice being added to the cage each day, a total of 2226 mice being employed.

We find from the laboratory graph, kindly placed at our disposal by Prof. Greenwood, that, at the beginning of the experiment, the cage contained forty-two mice; after 40 days the number had increased to about 250, and it remained approximately at this level (the highest value being 310) until the end of the experiment. A slight complication arises from the fact that towards the end of the experiment twenty mice were withdrawn at intervals of about a month. This would doubtless tend to cause the average population to be somewhat below the steady state level. Unfortunately the life table given for this epidemic (Greenwood *et al.* 1936) has been calculated over the whole experiment. For the purpose of present analysis it would have been better if it had referred only to the period from the 90th day onward, when an approximately steady state had evidently been established. However, in the absence of such data it is necessary to apply the theory to the available tables, it being remembered that the theory is at best only approximate, and that the chief interest is to see how far it accounts for the main features of the observed facts. It is not likely that the relatively short initial period of population increase would profoundly affect the figures given in the life tables.

It is first of all necessary to develop expressions for the  $l_x$  and the  $d_x$  of the life table in the case of a disease which confers only a partial immunity, corresponding to the expressions given in the previous paper for a disease which confers complete immunity. We shall next examine to what extent the figures for the epidemic  $B_6$  conform to these formulae. It will be found that special allowance has to be made for the somewhat extended incubation period (reckoned by Greenwood *et al.* at 9–14 days), but that when this is taken into account, the theoretical expressions can be fitted to the figures, the agreement being perhaps as satisfactory as might reasonably be expected.

#### MATHEMATICAL THEORY OF EPIDEMIC OF DISEASE, WITHOUT COMPLETE IMMUNITY, BUT WITH CONSTANT INFECTION, DEATH AND RECOVERY RATES

We shall first consider a system similar to that discussed in our previous paper (1937) except that in this case those who have recovered  $x$  are now liable to become reinfected. The sick  $y$  are recruited both from the entirely uninfected (virgin) mice  $\bar{x}$  and from the recovered  $x$ . As before, the specific death and recovery rates are assumed to be constant and equal to  $d$  and  $l$ , so that the number who die or recover in unit time are  $dy$  and  $ly$  respectively. Also the numbers of the virgin and of the recovered which become infected in unit time are assumed to be  $k\bar{x}y$  and  $kxy$  respectively. As will be seen later, the observations in the experiment under review are consistent with the assumption that the non-specific death rate  $\rho$  is constant; in the ectromelia epidemic, it will be remembered, this rate was approximately proportional to the cage age. The situation is in fact that represented as case (4) on p. 110 of a previous communication (Kermack & McKendrick, 1933) with the modification that  $\bar{\pi} = \pi = \rho$ .

Reference to Fig. 1 will make clear the various stages through which the mice can pass. The arrows indicate the possible movements of the animals, and the symbols attached to them represent the amount of movement per unit time.

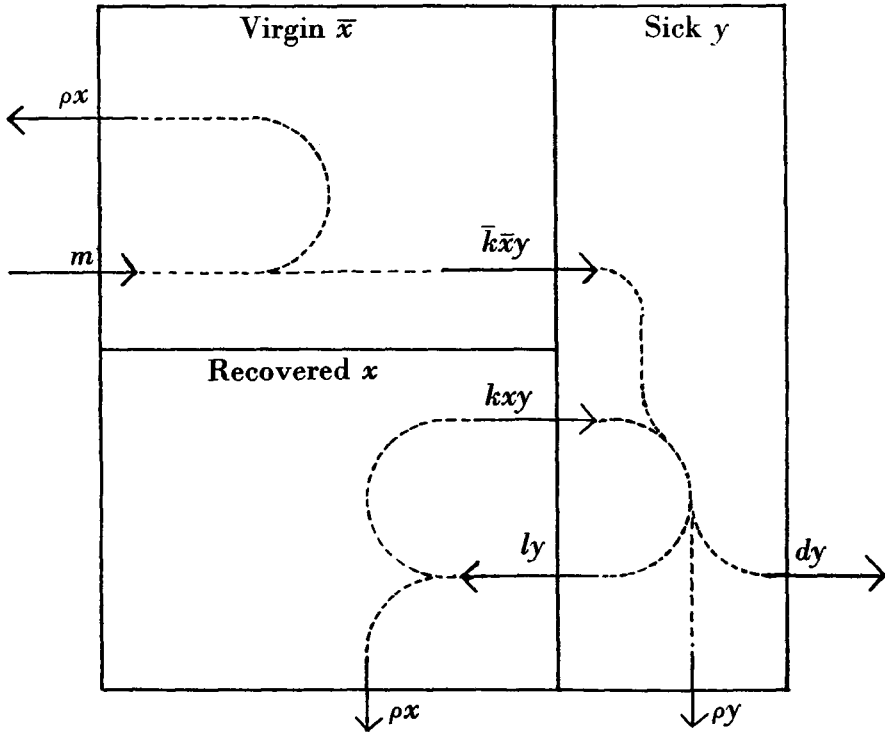


Fig. 1.

We are now in a position to write down the equations which must hold if a steady state condition is to exist, that is to say if the numbers of virgin, sick and recovered are to remain constant. We have simply to equate the number of mice entering a compartment per unit time to the number leaving it. Thus for the virgin compartment we have

$$m = \bar{k}\bar{X}Y + \rho\bar{X},$$

for the recovered

$$lY = kXY + \rho X,$$

for the sick

$$\bar{k}\bar{X}Y + kXY = dY + lY + \rho Y,$$

where capital letters indicate that the values in question refer to the steady state. It follows from these equations that

$$\left. \begin{aligned} \bar{X} &= \frac{m}{\bar{k}Y + \rho}, \\ Y &= \frac{\bar{k}\bar{X}Y + kXY}{d + l + \rho}, \\ X &= \frac{lY}{kY + \rho}. \end{aligned} \right\} \dots\dots(1)$$

It is necessary to determine  $\rho, d, l, \bar{k}, k$  and  $Y$  from the life table. With the same nomenclature as that used in the previous paper (1937) we have the following equations, where  $\eta$  is the cage age.

$$\left. \begin{aligned} \frac{d\bar{X}_\eta}{d\eta} &= -\bar{k}Y\bar{X}_\eta - \rho\bar{X}_\eta, \\ \frac{dY_\eta}{d\eta} &= \bar{k}Y\bar{X}_\eta + kYX_\eta - (d+l+\rho)Y_\eta, \\ \frac{dX_\eta}{d\eta} &= -kYX_\eta + lY_\eta - \rho X_\eta. \end{aligned} \right\} \dots\dots(2)$$

Writing

$$\left. \begin{aligned} d+l &= \lambda_1, \quad \bar{k}Y = \bar{\lambda}_2, \quad kY = \lambda_2, \quad c_1 = \frac{d}{d+l}, \quad c_2 = \frac{l}{d+l}, \quad (c_1 + c_2 = 1), \\ \bar{X}'_\eta &= \bar{X}_\eta e^{\rho\eta}, \quad Y'_\eta = Y_\eta e^{\rho\eta} \quad \text{and} \quad X'_\eta = X_\eta e^{\rho\eta}, \end{aligned} \right\} \dots\dots(3)$$

the above equations become

$$\left. \begin{aligned} \frac{d\bar{X}'_\eta}{d\eta} &= -\bar{\lambda}_2\bar{X}'_\eta, \\ \frac{dY'_\eta}{d\eta} &= \bar{\lambda}_2\bar{X}'_\eta + \lambda_2X'_\eta - \lambda_1Y'_\eta, \\ \frac{dX'_\eta}{d\eta} &= -\lambda_2X'_\eta + c_2\lambda_1Y'_\eta. \end{aligned} \right\} \dots\dots(4)$$

The quantities  $\bar{X}'_\eta, Y'_\eta$  and  $X'_\eta$  are obviously those obtained when the life table is calculated from specific deaths alone, that is, from animals which have died from the disease.

If we make use of the facts that  $\bar{X}'_0 = 1, Y'_0 = 0,$  and  $X'_0 = 0,$  we find that the solution of (4) is given by

$$\left. \begin{aligned} \bar{X}'_\eta &= e^{-\bar{\lambda}_2\eta}, \\ Y'_\eta &= \bar{\lambda}_2 \left\{ \frac{(\beta_1 - \lambda_2) e^{-\beta_1\eta}}{(\beta_1 - \beta_2)(\bar{\lambda}_2 - \beta_1)} + \frac{(\beta_2 - \lambda_2) e^{-\beta_2\eta}}{(\beta_2 - \bar{\lambda}_2)(\beta_1 - \beta_2)} + \frac{(\bar{\lambda}_2 - \lambda_2) e^{-\bar{\lambda}_2\eta}}{(\bar{\lambda}_2 - \beta_1)(\beta_2 - \bar{\lambda}_2)} \right\}, \\ \text{and} \\ X'_\eta &= \lambda_2 \left\{ \frac{(\beta_1 - \lambda_1)(\beta_1 - \lambda_2) e^{-\beta_1\eta}}{(\beta_1 - \beta_2)(\beta_1 - \bar{\lambda}_2)} + \frac{(\beta_2 - \lambda_1)(\beta_2 - \lambda_2) e^{-\beta_2\eta}}{(\beta_2 - \bar{\lambda}_2)(\beta_2 - \beta_1)} + \frac{(\lambda_1\lambda_2 - \beta_1\beta_2) e^{-\bar{\lambda}_2\eta}}{(\bar{\lambda}_2 - \beta_1)(\bar{\lambda}_2 - \beta_2)} \right\} \end{aligned} \right\} \dots\dots(5)$$

where  $\beta_1$  and  $\beta_2$  are the roots of the quadratic

$$\beta^2 - (\lambda_1 + \lambda_2)\beta + c_1\lambda_1\lambda_2 = 0. \dots\dots(6)$$

It follows that

$$\begin{aligned} N'_\eta &= \bar{X}'_\eta + Y'_\eta + X'_\eta \\ &= \frac{1}{\lambda_2} \left\{ \frac{\beta_2\bar{\lambda}_2(\beta_1 - \lambda_2) e^{-\beta_1\eta}}{(\beta_1 - \beta_2)(\bar{\lambda}_2 - \beta_1)} + \frac{\bar{\lambda}_2\beta_1(\beta_2 - \lambda_2) e^{-\beta_2\eta}}{(\beta_2 - \bar{\lambda}_2)(\beta_1 - \beta_2)} + \frac{\beta_1\beta_2(\bar{\lambda}_2 - \lambda_2) e^{-\bar{\lambda}_2\eta}}{(\bar{\lambda}_2 - \beta_1)(\beta_2 - \bar{\lambda}_2)} \right\}. \end{aligned} \dots\dots(7)$$

(It is to be noted that the expression for  $N'_\eta$  is symmetrical in  $\beta_1, \beta_2$  and  $\bar{\lambda}_2.$ )

Further, the expectation

$$E_0' = \frac{\int_0^\infty N_\eta' d\eta}{N_0'}$$

which after considerable reduction gives the simple and symmetrical expression

$$E_0' = \frac{1}{\beta_1} + \frac{1}{\beta_2} + \frac{1}{\lambda_2} - \frac{1}{\lambda_2} \tag{8}$$

In the case in which  $\beta_1 = \lambda_2$ ,

$$N_\eta' = W e^{-\beta_2 \eta} + (1 - W) (1 + \kappa \eta) e^{-\lambda_2 \eta} \tag{9}$$

where

$$W = \frac{\lambda_2 \beta_1 (\beta_2 - \lambda_2)}{\lambda_2 (\beta_2 - \lambda_2) (\beta_1 - \beta_2)} \tag{10}$$

and

$$\kappa = \lambda_2 + \frac{\beta_2 W}{(1 - W)} \tag{11}$$

The problem is now to obtain values for  $\beta_1, \beta_2, \lambda_2$  and  $\lambda_1$  which on substitution into equation (7) give a series of values approximating to those in the life table as calculated by Greenwood *et al.* from the specific deaths.

*Calculation of  $\rho, E_0', W$  and  $\beta_2$*

From the equation  $N_\eta = N_\eta' e^{-\rho \eta}$  it follows that

$$\rho \eta = \log N_\eta' - \log N_\eta$$

whence

$$\rho = \frac{1}{\eta} (\log N_\eta' - \log N_\eta) \tag{12}$$

In the life tables as published we are given life tables calculated (1) from specific deaths and (2) from total deaths. In the following table are given values of  $\rho$  calculated by equation (12) corresponding to different values of  $\eta$ .

$\eta = 100$	$\rho = 0.001236$	$\eta = 160$	$\rho = 0.001147$
110	0.001128	170	0.001193
120	0.001089	180	0.001250
130	0.001158	190	0.001184
140	0.001131	200	0.001124
150	0.001105		

Although the values vary considerably they seem to be arranged at random about the mean, which is therefore chosen as the best value of  $\rho$ . Thus  $\rho = 0.00116$ .

In order to find the value of  $E_0' = \frac{\int_0^\infty N_\eta' d\eta}{N_0'}$ , we may replace the integral

by the sum  $5N_0' + 10(N'_{10} + N'_{20} + \dots)$ .

The difficulty arises that the life table as given in the paper does not extend beyond  $\eta = 200$ . As in the previous paper, we must therefore make allowance for the continuance of the curve. We observe that if  $\beta_2$  be the smallest of the

three values  $\bar{\lambda}_2, \beta_1$  and  $\beta_2$ , then for large values of  $\eta, N'_\eta$  is approximately equal to  $W e^{-\beta_2 \eta}$ . We therefore find the values of  $W$  and  $\beta_2$  which best represent the curve for values of  $\eta$  between 100 and 200. By applying the method of least squares to the curve

$$\log N'_\eta = \log W - \beta_2 \eta$$

for values of  $\eta = 100, 110, 120$ , etc. the following constants are found,

$$W = 0.2293 \quad \text{and} \quad \beta_2 = 0.0062684.$$

We may now replace the actual tail of the curve by the theoretical tail as calculated from  $W e^{-\beta_2 \eta}$ , and thus we have

$$\begin{aligned} E'_0 &= 5N'_0 + 10(N'_{10} + N'_{20} + \dots + N'_{200}) + \int_{205}^{\infty} W e^{-\beta_2 \eta} d\eta, \quad \dots\dots(13) \\ &= 5N'_0 + 10(N'_{10} + N'_{20} + \dots + N'_{200}) + \frac{W}{\beta_2} e^{-205\beta_2}. \end{aligned}$$

Substitution of the appropriate values gives

$$\begin{aligned} E'_0 &= 46.17360 + 10.13847 \\ &= 56.31207. \end{aligned}$$

As there are four unknown parameters,  $\beta_1, \lambda_1, \lambda_2$  and  $\bar{\lambda}_2$ , and three determined values,  $\beta_2, E'_0$  and  $W$ , one of the unknown parameters (e.g.  $\beta_1$ ) may be selected arbitrarily. It must, however, be chosen so that the fit of the curve is the best possible. On trying a few values for  $\beta_1$ , it is soon found that all the calculated curves deviate markedly from the observed, for they fall too rapidly initially, and compensate for this by being too high later on.

It is not difficult to show that the minimum rate of initial fall is obtained when  $\beta_1 = \bar{\lambda}_2$ . When this condition is satisfied the values of all the constants are fixed by the three equations already obtained (8, 10, 11), and we find that  $\beta_1 = \bar{\lambda}_2 = 0.079$ . Substituting this in equation (9) we obtain

$$N'_\eta = 0.2293 e^{-0.0062684\eta} + 0.7707 (1 + 0.1019\eta) e^{-0.079\eta}.$$

On evaluating  $N'_\eta$  for  $\eta = 10, 20, 30$ , etc., the fit is found to be unsatisfactory. The trouble arises evidently as a consequence of the existence of a latent period.

*Modification necessitated by latent period*

In order to take a latent period into account, the most straightforward way would be to assume that  $d_\theta$  had the value zero from the time of infection until the  $a$ th day of the disease, and that it thereafter assumed a constant value  $d$ . The mathematics which results however from this assumption is complicated and intractable.

The following scheme results in an easier method of treatment and brings

the case of the latent period into simple relation with that already treated. We assume that animals when they are infected enter a category  $y_1$  from which they may either recover or become moribund. The chances of recovery and of entering the moribund condition are supposed to be constant and to be  $l$  and  $d$  respectively. The moribund animals belong to a category  $y_2$ , and the assumption is made that they remain for a definite period of time  $a$  in this group, if they do not die previously from natural causes. It will be seen that this system gives a latent period in the sense that no animal dies of the specific disease in less than  $a$  days after it has been infected. The animals can recover only during the earlier period of their illness, and they have no chance of overcoming the infection after they have entered the moribund category. In the case of mouse aertrycka the existence of a moribund period of this type is not perhaps entirely at variance with the facts of experience, though of course the system proposed is admittedly a somewhat artificial one. All the animals which are ill, both those in category  $y_1$  and in category  $y_2$ , are regarded as being infective.

We shall now consider the modification which the assumption of this moribund group of animals necessitates in our theory. It is easy to see that the effect on the life table for specific deaths of the existence of this category of moribund animals is merely to post date each death by a period  $a$ , that is to say, the whole curve is displaced to the right by this amount. We have therefore to fit the theoretical curve to the actual curve obtained by taking the origin at the point  $\eta = a$ , at which point we assume that  $N'_a = 1$ , as the theory implies that  $N'_\eta$  remains at this value between  $\eta = 0$  and  $\eta = a$ . The modification which must be introduced into the steady state conditions, results from the fact that the number of infecting animals is not  $Y_1$  but  $Y_1 + Y_2$ .

We readily find

$$\left. \begin{aligned} m &= \{\rho + \bar{k}(Y_1 + Y_2)\} \bar{X}, \\ (\bar{k}\bar{X} + kX)(Y_1 + Y_2) &= (d + l + \rho) Y_1, \\ lY_1 - kX(Y_1 + Y_2) &= \rho X. \end{aligned} \right\} \dots\dots(14)$$

Now the number who enter the moribund category each day is  $dY_1$ . If there were no deaths from non-specific causes in this category, then each animal remains in it for  $a$  days, and the number on any day would be  $adY_1$ . Hence

$$Y_2 = adY_1.$$

It is not difficult to show that the effect of non-specific deaths is to reduce the average duration of stay in the moribund category from  $a$  to  $\bar{a}$ , where

$$\bar{a} = \frac{1 - e^{-\rho a}}{\rho}, \dots\dots(15)$$

so that

$$Y_2 = \bar{a}dY_1. \dots\dots(16)$$

It is easy to see that as  $\rho \rightarrow 0$ ,  $\bar{a} \rightarrow a$ .

Returning to the life-table, it is not difficult to see that although the process is the same as in the previous case, apart from the change of origin, the interpretations of  $\bar{\lambda}_2$  and  $\lambda_2$  are slightly different. As the mice in both categories are infective we have

$$\bar{\lambda}_2 = \bar{k} (Y_1 + Y_2) = \bar{k} Y_1 (1 + \bar{a}d), \tag{17}$$

and 
$$\lambda_2 = k (Y_1 + Y_2) = k Y_1 (1 + \bar{a}d), \tag{18}$$

where  $\bar{a}$  is known and  $d = \lambda_1 c_1$ .

Equations 14, 16, 17 and 18, give six equations from which the values of  $\bar{X}$ ,  $Y_1$ ,  $Y_2$  and  $X$ ,  $k$  and  $\bar{k}$  may be found. We thus find,

$$\left. \begin{aligned} \bar{X} &= \frac{m}{\bar{\lambda}_2 + \rho}, \\ Y_1 &= \frac{m \bar{\lambda}_2 (\lambda_2 + \rho)}{(\beta_1 + \rho) (\beta_2 + \rho) (\bar{\lambda}_2 + \rho)}, \\ X &= \frac{m \bar{\lambda}_2 (\lambda_1 \lambda_2 - \beta_1 \beta_2)}{\lambda_2 (\beta_1 + \rho) (\beta_2 + \rho) (\bar{\lambda}_2 + \rho)}, \\ Y_2 &= \frac{m \bar{a} (\lambda_2 + \rho) \beta_1 \beta_2 \bar{\lambda}_2}{\lambda_2 (\beta_1 + \rho) (\beta_2 + \rho) (\bar{\lambda}_2 + \rho)}, \end{aligned} \right\} \tag{19}$$

and  $N = \bar{X} + Y_1 + X + Y_2$

$$= \frac{m \left\{ \beta_1 \beta_2 + \beta_2 \bar{\lambda}_2 + \bar{\lambda}_2 \beta_1 + (\beta_1 + \beta_2 + \bar{\lambda}_2) \rho + \rho^2 - \beta_1 \beta_2 \bar{\lambda}_2 \left( \frac{1 - \bar{a} (\lambda_2 + \rho)}{\lambda_2} \right) \right\}}{(\beta_1 + \rho) (\beta_2 + \rho) (\bar{\lambda}_2 + \rho)}. \tag{20}$$

*Values of constants and fitting of Life Table*  
(Greenwood *et al.* 1930, Table V B)

In fitting the curve to the life table it is found by trial and error that the best results are given by taking  $a = 13$ .

From equations (8) and (10) we obtain

$$\frac{1}{\lambda_2} = \frac{1}{\beta_2} + \frac{E^* G}{1 - G}, \tag{21}$$

where 
$$G = W \left( 1 - \frac{\beta_2}{\beta_1} \right), \tag{22}$$

and 
$$E^* = E'_0 - \frac{1}{\beta_1} - \frac{1}{\beta_2}. \tag{23}$$

Now, as we are considering the curve as commencing at the point  $\eta = 13$ , we have to replace  $E'_0$  by  $E'_{13}$ . According to the model which we have in mind we must assume that no deaths take place before the 13th day. Thus in the calculation of  $E'_{13}$  it must be assumed that  $N'_{13}$  has the value unity. For convenience we may group the deaths which occur after  $\eta = 15$  in 10-day



periods centred round the 20th, 30th, 40th, ... days. We have then approximately

$$\begin{aligned} E'_{13} &= 2N'_{13} + 10N'_{20} + 10N'_{30} + \dots \\ &= 2 + 10N'_{20} + 10N'_{30} + \dots \\ &= 2 + E'_0 - 5N'_0 - 10N'_{10} \end{aligned}$$

(since by equation (13),  $E'_0 = 5N'_0 + 10N'_{10} + 10N'_{20} + \dots$ )

$$\begin{aligned} &= 2 + E'_0 - 5 - 10N'_{10} \\ &= E'_0 - 10N'_{10} - 3 \\ &= 43.921. \end{aligned}$$

At the same time  $W$  must be replaced by  $W_{13} = We^{-13\beta_2} = 0.2106$ . Thus  $\lambda_2 = 0.007736$ .

Also 
$$\frac{1}{\lambda_2} = \frac{1}{\beta_2} + \frac{E^*}{1-G}, \dots\dots(24)$$

whence  $\lambda_2 = 0.3289$ .

By trial and error it is found that 0.1 is a reasonable value for  $\beta_1$ . Hence by equation (6)

$$\begin{aligned} \lambda_1 &= \beta_1 + \beta_2 - \lambda_2 \\ &= 0.098533, \end{aligned}$$

and

$$c_1 = \frac{\beta_1 \beta_2}{\lambda_1 \lambda_2} = 0.8225,$$

also

$$c_2 = 1 - c_1 = 0.1775.$$

Also

$$l = c_1 \lambda_1 = 0.01749,$$

and

$$d = c_2 \lambda_1 = 0.08104.$$

These values give the equation

$$N'_{\eta+13} = 0.2106 e^{-0.0062684\eta} + 1.147 e^{-0.1\eta} - 0.3526 e^{-0.3289\eta}. \dots\dots(25)$$

(The 13 is inserted because of the change of origin to  $\eta = 13$ ; thus putting  $\eta = 7$  in the above gives  $N'_{20}$ .)

Table I shows the calculated values of  $N'_\eta$  (or  $l_x$ )—i.e. for specific deaths—as well as the figures given in the record of the experiment (Greenwood *et al.* 1930). Table II shows the corresponding values for “all deaths”, calculated by the equation  $N_\eta = N'_\eta e^{-\rho\eta}$ . Both tables give also the  $d_x$ 's, i.e. the number of deaths in each period. A slight difficulty arises in relation to the deaths which occurred earlier than the 20th day. In the theoretical model, all these deaths necessarily take place between the 13th and the 20th day, whilst in the actual experiment there is no sharp onset of deaths at the 13th day, but the deaths are scattered throughout the whole of the initial period. It seems reasonable to compare the deaths observed up to the 20th day with the theoretical number up to the same date, and in Figs. 3 and 5 these values of  $d_x$  have been somewhat arbitrarily centred at the 15th day. A change of position would not materially affect the comparison.

Table I.  $B_6$  experiment

*Specific deaths* (Greenwood *et al.* 1930, Table V<sub>B</sub>)

Cage age in days	$l_x$		$d_x$	
	Calculated	Observed	Calculated	Observed
20	7358	7217	2642	2783
30	3976	3778	3382	3439
40	2549	2403	1427	1375
50	1954	1955	595	448
60	1673	1763	281	192
70	1511	1615	162	148
80	1396	1457	115	158
90	1304	1372	92	85
100	1220	1261	84	111
110	1147	1170	73	91
120	1077	1086	70	84
130	1011	1016	66	70
140	950	980	61	36
150	892	929	58	51
160	838	842	54	87
170	787	801	51	41
180	739	764	48	36
190	694	699	45	66
200	652	670	42	29

Table II.  $B_6$  experiment

*All deaths* (Greenwood *et al.* 1930, Table V)

Cage age in days	$l_x$		$d_x$	
	Calculated	Observed	Calculated	Observed
20	7189	6930	2811	3070
30	3837	3544	3352	3386
40	2433	2224	1404	1320
50	1844	1805	589	419
60	1561	1618	283	187
70	1393	1460	168	158
80	1272	1301	121	159
90	1175	1226	97	75
100	1086	1115	89	111
110	1010	1034	76	81
120	937	953	73	81
130	869	873	68	80
140	808	837	61	36
150	750	787	58	50
160	696	702	54	75
170	646	654	50	48
180	600	611	46	43
190	557	559	43	52
200	517	535	40	24

For convenience the  $d_x$  value corresponding to the interval 0 to 20 is inserted opposite the cage age 20, that for the interval 20 to 30, opposite the cage age 30, and so on.

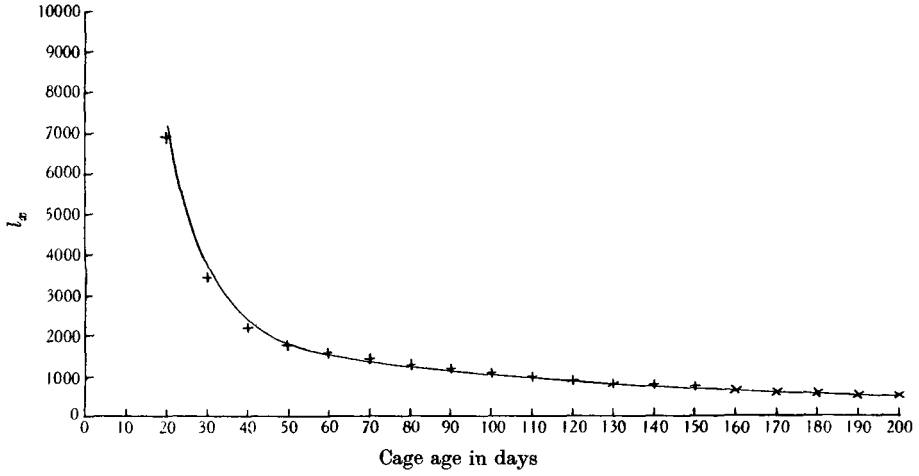


Fig. 2.  $l_x$ .  $B_6$ . Specific deaths.

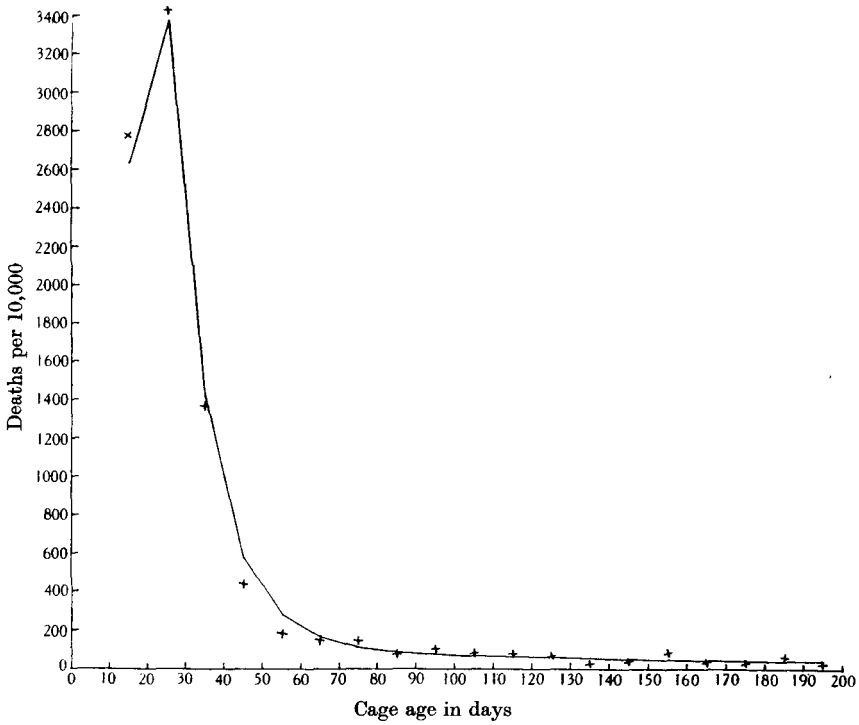


Fig. 3.  $d_x$ .  $B_6$ . Specific deaths.

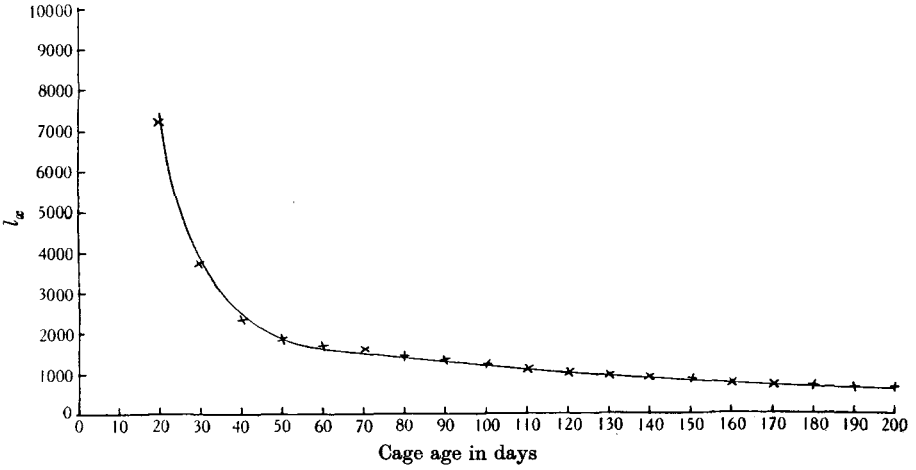


Fig. 4.  $l_x \cdot B_6$ . All deaths.

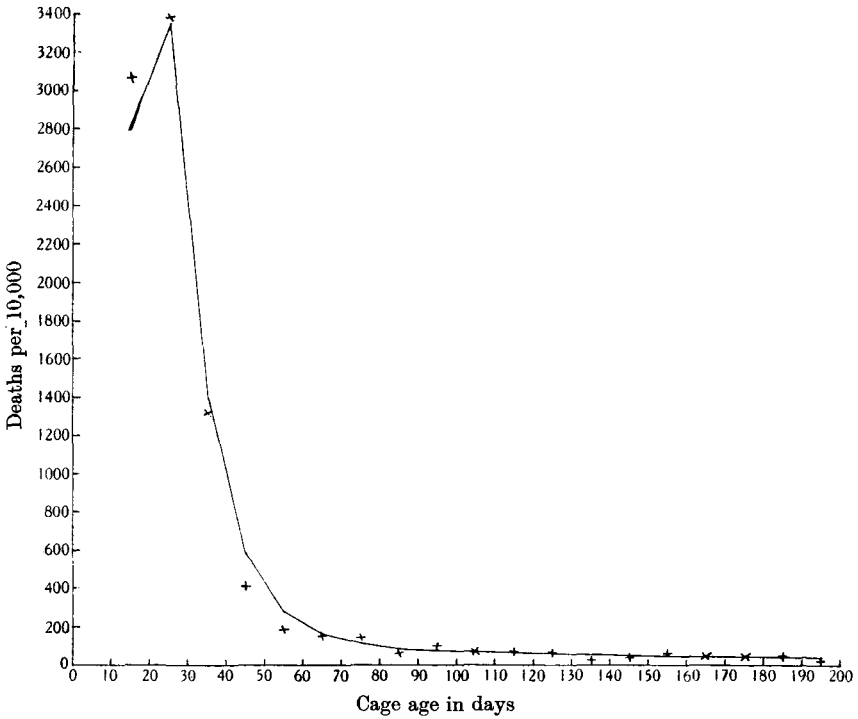


Fig. 5.  $d_x \cdot B_6$ . All deaths.

In order to compare the calculated with the observed, both sets are multiplied by the total number of deaths, namely 1766 and 2226 for "specific" and "all deaths" respectively, and the  $\chi^2$  test has been applied to the numbers so obtained. For "specific deaths" (Table I),  $\chi^2=28.9$ ,  $n=19-6=13$ , whence  $P=0.009$ . For "total deaths" (Table II)  $\chi^2=38.2$ ,  $n=13$ ,  $P=0.0002$ . This shows that whilst the specific deaths show deviations such as might possibly occur by chance (1 in 100), the total deaths appear to show very significant deviations from the expected. The somewhat irregular distribution of the observed numbers of deaths suggests that the effect of disturbing factors is not entirely negligible, and that part of the discrepancy at least is due to such cause. It must be remembered that the theory is definitely only an approximate one, and that no claim is made that all deviations observed are due to sampling error alone.

The constants already obtained, and the values deduced by equations (19), (17), (18) and (3) are given in the first column of the accompanying table.

	1st interpretation	2nd interpretation
$\lambda_1$	0.09853	0.3274
$c_1$	0.8225	0.8139
$c_2$	0.1775	0.1861
$X$	18.18	59.31
$Y_1$	70.77	21.51
$X$	139.06	147.39
$Y_2$	74.00	73.99
$N$	302.01	302.20
$\bar{k}$	0.002271	0.001047
$k$	0.00005342	0.00008099
$l$	0.01749	0.06095
$d$	0.08104	0.2666

*Second interpretation*

As the expression for  $N_\eta'$  (7) is symmetrical in  $\beta_1$ ,  $\beta_2$  and  $\bar{\lambda}_2$  it follows that we may interchange the numerical values of these three constants without altering the values of  $N_\eta'$ . The equation for  $E_0'$  (8) shows that  $\lambda_2$  will also remain unchanged, and it then follows from the quadratic (6), that, as  $\beta_1 + \beta_2 = \lambda_1 + \lambda_2$ ,  $\lambda_1$  will remain unchanged when  $\beta_1$  and  $\beta_2$  are interchanged,  $\bar{\lambda}_2$  remaining fixed. This means that of the six possible permutations of  $\beta_1$ ,  $\beta_2$  and  $\bar{\lambda}_2$  there are only three sets which are distinct from each other. Furthermore as  $\frac{\bar{\lambda}_2}{\lambda_2} = \frac{\bar{k}}{k}$ , and this fraction must be greater than unity because recovered mice are less easily infected than virgin mice, it follows that  $\bar{\lambda}_2 > \lambda_2$ . This reduces the number of possible sets to two, namely  $\beta_1=0.1$ ,  $\beta_2=0.0062684$  and  $\bar{\lambda}_2=0.3289$  (as found above) and  $\beta_1=0.0062684$ ,  $\beta_2=0.3289$  and  $\bar{\lambda}_2=0.1$ .

The second allocation of values is given in the second column of the above table.

By the symmetry of the equations (19) (no. 4) and (20),  $Y_2$  and  $N$  remain unchanged. Evidently it is impossible to distinguish between the two interpretations, from observations on the total numbers of mice in the steady state.

In order to effect an identification, it would be necessary to have special information about the relative numbers of infected and uninfected, or, for example, some estimate of the average period of time which elapsed between entry into the cage and the first signs of infection. However, we have not been able from the published papers to obtain any information on this point, which is sufficiently definite and precise to enable us to draw conclusions as to which interpretation is the more likely.

It is perhaps of some interest to find that in this case of a disease with incomplete immunity as well as in the simpler case previously treated (Kermack & McKendrick, 1937) of a disease with complete immunity, it is possible to give two distinct interpretations, both completely consistent with the facts as revealed by the cage-age life table. According to the first interpretation the disease is relatively highly infectious to the new entrants into the cage, but the course of the malady in the animals, once they have become infected, is slower. The ratio of the chances of death and recovery ( $d/l$ ) is approximately the same for both cases, but if recovery takes place the animals according to the first interpretation will have attained a higher degree of immunity. In the first case the population in the steady state condition contains a relatively smaller number of virgins, but a correspondingly larger number of diseased, the number of recovered being nearly equal in the two cases.

*Effect of variation of rate of immigration*

It is of interest to work out the effect on the course of the epidemic of variation in the rate at which animals are introduced into the cage, and to employ the numerical values obtained from experiment  $B_6$  to calculate the course of experiments  $B_3$  and  $B_1$  in which the daily additions to the cage were three mice and one mouse respectively. If the cage area were the same in each experiment, all that would be necessary would be to substitute the already calculated values of  $\bar{k}$ ,  $k$ ,  $\lambda_1$ ,  $c_1$ , etc., in the quadratic for  $Y_1$ , obtained from equations (14), i.e. in

$$\frac{\bar{k}m}{\bar{k}Y_1(1+d\bar{a})+\rho} + \frac{klY_1}{kY_1(1+d\bar{a})+\rho} = \frac{\lambda_1+\rho}{1+d\bar{a}}, \quad \dots\dots(26)$$

and thence find  $Y_1$ , and from it  $\bar{X}$ ,  $X$ ,  $Y_2$ , etc.

However, the situation is rather more complicated because, as stated on p. 11 of the memoir (Greenwood *et al.* 1936), cage units were added as the population increased, so as to provide one cage for every group of twenty-five mice or fraction thereof. As the total populations in each experiment underwent fluctuations it would appear that the number of cage units altered slightly even during the approximately steady state period. It seems justifiable to assume that the numbers of cage units in the different experiments were in the ratio of the average total populations during the course of each experiment. As these averages were in the ratios 237.2 : 103.1 : 32.7 in experiments  $B_6$ ,  $B_3$  and  $B_1$  respectively, it is assumed that the cage areas on the average were in

the ratios 7.25 : 3.15 : 1. Now in the mathematical theory the constants  $\bar{k}$  and  $k$  depend upon the unit of area chosen in measuring  $\bar{X}$ ,  $X$ ,  $Y_1$  and  $Y_2$ , which, it will be remembered, refer to numbers per unit area. If the total area available is altered, then, if we take  $\bar{X}$ ,  $X$ ,  $Y_1$  and  $Y_2$  as the actual numbers, it is necessary to make corresponding alterations in the constants  $\bar{k}$  and  $k$ . If the area is reduced in the ratio of  $\alpha$  to 1, then the net effect is to increase  $\bar{k}$  and  $k$  to  $\alpha\bar{k}$  and  $\alpha k$  respectively. Thus for  $m=1$ ,  $Y_1$  is given as the solution of the quadratic (26) where  $d$  and  $l$  have their previous values, and  $\bar{k}$  and  $k$  have their former values multiplied by 7.25.

If we consider case 1 this gives

$$Y_1^2 - 10.3706 Y_1 - 17.7703 = 0$$

whence  $Y_1 = 11.87$ ,  $\bar{X} = 2.50$ ,  $X = 19.57$  and  $Y_2 = 12.41$ .

The fact that the expression for  $N$  is symmetrical in  $\beta_1$ ,  $\beta_2$ , and  $\lambda_2$  suggests at first sight that the effect of alteration of  $m$ , which occurs explicitly as a factor of  $N$ , will be the same in both interpretations. It is to be remembered however that  $\beta_1$ ,  $\beta_2$ , and  $\lambda_2$  are different functions of the fundamental constants  $\bar{k}$ ,  $k$ ,  $l$ ,  $d$ ,  $\rho$  and  $m$ , so that they are altered in different ways when  $m$  is changed.

By employing the alternative interpretation of the constants (case 2) we obtain another set of values. Two analogous sets of values are likewise obtained for  $m=3$ . These values are collected in Table III.

Table III

	$m=6$		$m=3$		$m=1$	
	Case 1	Case 2	Case 1	Case 2	Case 1	Case 2
$\bar{X}$	18.2	59.3	8.0	25.7	2.5	8.2
$Y_1$	70.8	21.5	35.2	10.8	11.9	3.6
$X$	139.1	147.4	61.5	65.2	19.6	20.8
$Y_2$	74.0	74.0	36.8	37.2	12.4	12.4
$N$	302.1	302.2	141.5	138.9	46.4	45.0
$k$	0.002271	0.001047	0.005226	0.002409	0.01646	0.007591
$k$	0.00005342	0.00008099	0.0001229	0.0001864	0.0003873	0.0005873
$d$	0.08104	0.2666	0.08104	0.2666	0.08104	0.2666
$l$	0.01749	0.06095	0.01749	0.06095	0.01749	0.06095

It is evident that as far as the total steady state populations go the two interpretations lead to almost exactly the same values. The differences are too small to allow of conclusions being drawn as to which hypothesis is the correct one. According to the above theory the steady state levels for the three experiments  $B_6$ ,  $B_3$  and  $B_1$  should have been 302.2, 141.3 and 46.4 in the case of the first interpretation, and 302.2, 138.8 and 45.0 in the case of the second. Unfortunately the values of the observed steady state levels are not available, so that direct comparison of theory and experiment is impossible.<sup>1</sup> However, the average numbers of mice present throughout the experiments are available. These numbers are of course lower than the corresponding

<sup>1</sup> Even in the case of  $B_6$ , the daily records of which were at our disposal, the ascertainment of the steady state level was complicated by the fact that batches of mice were withdrawn at intervals from the cages.

steady state levels, but it might be anticipated that they are approximately proportional to the latter. They are in fact in the ratios 7.25 : 3.15 : 1, whilst the steady state levels are in the ratios 6.71 : 3.04 : 1 for the first interpretation, and 6.71 : 3.08 : 1 according to the second. It must be remembered of course that the observed ratios of the average numbers of mice present throughout the experiments were employed in calculating the steady state levels, a fact which perhaps renders the comparison less convincing, but the result does at least show that the theory does not lead to conclusions inconsistent with the facts. A comparison of the observed and calculated expectations of life in the three experiments gives results in harmony with those obtained from the steady state levels.

#### DISCUSSION

As remarked in our previous paper, attempts such as the present to interpret the course of experimental epidemics in terms of constant infectivity, death, and recovery rates must be looked on merely as a first approximation to the truth. Any agreement between theory and experiment which may be achieved only shows that of the vast number of factors which influence such an epidemic, the relatively important ones have been singled out, and their effect adequately allowed for. This does not necessarily mean that an accurate mathematical description of these effects has been presented. Too great an insistence on accuracy in the fundamental assumptions may result in a mathematical model which proves of little practical use because of the technical difficulties involved. The mathematical model we have postulated in the foregoing analysis is, then, to be looked upon as a first approximation, capable of refinements when applied to more perfect data, but adequate for the interpretation of the principal characteristics of the experimental epidemics here examined, and capable perhaps of extension to more complex cases in which the experimental method becomes difficult or impossible.

The chief result which emerges from the above analysis of the mouse typhoid epidemics is that the simple constant rate theory does not, as in the case of the ectromelia experiment, provide an adequate interpretation of the life tables. It is necessary to introduce a modification to allow for the existence of a latent period. The modification proposed is perhaps somewhat artificial, but it has the advantage of permitting the problem to be treated in a comparatively simple manner which does not involve undue mathematical complication. This simple theoretical model, as we have seen, permits the cage-age life table to be interpreted quantitatively in terms of five constants,  $\bar{k}$ ,  $k$ ,  $d$ ,  $l$  and  $\rho$ .

Some of the characteristics of this life table remarked upon by Greenwood *et al.* in their paper (1930) follow at once from the form of the curve. They point out, for example, that  $q_x$  for specific deaths rises to a maximum, and then falls to a constant level. In our notation  $\frac{1}{N_\eta} \frac{dN_\eta'}{d\eta}$  is equivalent to  $q_x$ , and it is



not difficult to show that after its initial maximum this function falls and approaches a constant value  $\beta_2 = 0.0062684$  in  $B_6$ . On the other hand,  $\frac{dN'_\eta}{d\eta}$  also reaches a maximum and then falls approximately following an exponential curve for values of  $\eta$  above 100, as remarked upon by Greenwood *et al.* with regard to their equivalent function  $d_x$ .

Another semi-quantitative point of agreement between theory and experiment is that in both the steady state level is approximately proportional to the immigration rate, as shown by the comparison of experiments  $B_6$ ,  $B_3$  and  $B_1$ , whilst the expectation of life at entry in terms of cage age remains approximately constant even though the immigration rate is varied. However, as explained above, the limitations of the experimental data preclude exact comparisons being made between theory and experiment in regard to these values.

#### CONCLUSIONS

1. An analysis has been made of the progress of an epidemic of mouse typhoid (previously described by Greenwood *et al.*) in a relatively large herd of mice. This epidemic differs from that of ectromelia analysed in our previous paper, in that it is caused by a bacterium (*Bact. aertrycke*), and not by a virus. In this case the immunity resulting from an attack of the disease is only partial, and the incubation period is somewhat longer, being about 13 days as compared with 3 or 4 in the case of ectromelia.

2. Although the existence of partial instead of complete immunity leads to more complicated mathematical expressions involving an extra constant, it does not introduce any essential difficulty. It is found however that the longer incubation period requires special treatment.

3. The life tables which have been calculated by Greenwood *et al.* for the epidemic  $B_6$ , have been fitted on the basis of the present theory. Reasonably good agreement between theoretical and observed values is obtained, provided that the theory is suitably modified so as to accommodate the incubation period. As a result of the symmetry of the equations involved, two alternative interpretations of the parameters, in terms of the constants which characterize the disease, are admissible. From the available data we have not been able to determine which of the two interpretations is the correct one.

4. When the immigration rate is three mice or one mouse per day, the steady state levels, according to the published data, appear to be approximately one-half, or one-sixth respectively of that found in the case where the immigration rate is six mice per day. This result is in agreement with the present theory.

5. The above results indicate that the present theory using constant coefficients is adequate to explain the main features of the mouse typhoid epidemics, provided that allowance be made for the somewhat prolonged incubation period.

In this and in previous work, we are much indebted to Prof. M. Greenwood and to Dr Irwin for helpful criticism and suggestions. Prof. Greenwood has also supplied us with certain facts relating to the experimental data, which were not available in the published papers.

## REFERENCES

- GREENWOOD, NEWBOLD, TOPLEY & WILSON (1930). *J. Hyg., Camb.*, **30**, 240.  
GREENWOOD, BRADFORD HILL, TOPLEY & WILSON (1936). *Spec. Rep. Ser. med. Res. Coun., Lond.*, no. 209.  
KERMACK & MCKENDRICK (1927). *Proc. roy. Soc. A*, **115**, 700.  
——— (1932). *Proc. roy. Soc. A*, **138**, 55.  
——— (1933). *Proc. roy. Soc. A*, **141**, 94.  
——— (1936). *Proc. Lond. Math. Soc. series 2*, **41**, 462.  
——— (1937). *J. Hyg., Camb.*, **37**, 172.

(*MS. received for publication 27. III. 39.—Ed.*)