J. Hyg., Camb. (1966), **64**, 431
Printed in Great Britain

The mathematical analysis of concurrent epidemics of yaws and chickenpox

By JOHN J. GART
National Cancer Institute, Bethesda, Maryland 20014, *U.S.A.*
AND J. L. DE VRIES*
Yaws Control Section, Department of Health, Government of Netherlands New Guinea

(*Received* 21 *May* 1966)

INTRODUCTION

Mathematical analysis of factual epidemiological information is only slowly gaining recognition as a useful, and frequently preferable, method in the interpretation of epidemiological phenomena. In spite of a large amount of information on the epidemiology of yaws becoming available from large-scale eradication campaigns against yaws, much remains unclear. This is particularly true with respect to precise epidemiological parameters, such as the infection rate, which are vital for the planning of long-term surveillance of yaws in places where the prevalence of the disease has been brought down to a low level.

Periodic epidemic spread of endemic disease or long-term persistence of infectious diseases in populations in which otherwise the endemicity level would appear to be below the threshold level constitutes a puzzling aspect in epidemiology.

In this paper a mathematical analysis is employed to elucidate one epidemiological factor in yaws in a tropical area and to demonstrate the importance of a seemingly unrelated disease in changing the epidemiological pattern.

One of the authors (J.L.d.V.) investigated in 1957 an epidemic of yaws among children 10 years of age and younger† in a village of Netherlands New Guinea. The epidemic was of particular interest in that during it an epidemic of chickenpox broke out among the same population. Furthermore, it occurred a year and a half after a control survey during which total mass treatment (Hackett & Guthe, 1956; Report, 1960) had been applied in a WHO–UNICEF assisted yaws control campaign. Moreover, the spread of the disease was much faster than that normally found in yaws in Netherlands New Guinea (Kranendonk, 1958); there being a spectacular increase in the number of yaws cases in the month following the chickenpox epidemic. It was inferred that the chickenpox predisposed these children to become infected with yaws. This conjecture has some medical validity since:

(1) The incubation period of yaws is 20–30 days, usually assumed as 3 weeks

* Present address: RACD, WHO Regional Office for South East Asia, World Health House, Indraprastha Road, New Delhi 1, India.

† In holo-endemic areas, the percentage of infected children at age 5 is 95% (Kranendonk, 1958) and in general the susceptibles are exhausted at age 6 (Soetopo & Wasito, 1953).

(Turner & Hollander, 1957; Hackett, 1963), a time-gap consistent with the interval between the peaks of the chicken-pox and yaws epidemics;

(2) The spirochete causing yaws (*Treponema pertenue*) requires a portal of entry (Hackett, 1960) which might be provided by the skin eruptions caused by the chickenpox;

(3) The infection with chickenpox could cause increased susceptibility to yaws by stimulation of steroid production (Turner & Hollander, 1957).

This paper is concerned with advancing statistical and mathematical arguments relating to two specific aspects of these epidemics, namely:

(1) The children who had chickenpox were more likely to contract yaws in the month following the chicken-pox epidemic.

(2) The chicken-pox epidemic accelerated the yaws epidemic; that is, more cases of yaws were found than would otherwise have been the case.

The latter point is investigated using a mathematical model for a simple deterministic epidemic. A wider analysis relating to more general aspects of the epidemic is in preparation.

THE YAWS EPIDEMIC

Between August 1956 and February 1957 there was an epidemic of yaws among children 10 years of age and younger of Jongsu Besar in Netherlands New Guinea. In October of 1956 an epidemic of chickenpox occurred among the same population. The incidence of yaws by month was recorded retrospectively by one of the authors (J.L.d.V.) and is given in Table 1. In this table the children are also divided into two groups as to their previous history of yaws. This history is based on the information from accompanying parents and on the presence or absence of scars from infectious lesions, the latter being the arbiter when present. Absence of detectable scars and positive history is noted as positive, mainly because early infectious yaws does not always leave scars on healing.

A marked peak of the yaws epidemic is noted in November, the month following the chickenpox epidemic. Later the epidemic gradually builds up in an epidemio-

Table 1. *Course of yaws epidemic according to yaws history*

	New yaws cases		
Month	Negative history of yaws	Positive history of yaws	Total
Aug. 1956	1	0	1
Sept. 1956	1	0	1
Oct. 1956	0	0	0
Nov. 1956	11	2	13
Dec. 1956	2	0	2
Jan. 1957	3	1	4
Feb. 1957	8	2	10
Unknown but after Nov. '56	3	2	5
Total cases	29	7	36
No. uninfected	26	41	67
Totals	55	48	103

https://doi.org/10.1017/S0022172400040730 Published online by Cambridge University Press

logical pattern typical for yaws to a second peak in February. At this point mass treatment with long-acting penicillin (PAM) brought an end to the epidemic.

Table 2 records the monthly incidence of yaws among children 10 years and younger by the chickenpox infection status. It is noted that 12 of the 13 cases of yaws in November had chickenpox in October while of the 14 cases in January and February only 3 were of the chickenpox group.

Table 2. *Course of yaws epidemic in children aged* 10 *years or less according to chickenpox history*

Chickenpox status	Total no. of children	Yaws cases in: Aug.–Oct.	Nov.	Dec.	Jan.–Feb.	Unknown	Total Yaws
Yes	50	2	12	1	3	1	19
No	49	0	1	1	11	1	14
Unknown	4	—	—	—	—	3	3
Total	103	2	13	2	14	5	36

Except for five individuals among the total population of the village, all the chickenpox cases occurred in October. The five exceptional cases involved only two children who were 10 years of age or younger. One of these had chickenpox in September and already had a positive history of yaws. The second had chickenpox in November and had a negative history of yaws. None of the five contracted yaws during 1956 or 1957.

PRELIMINARY STATISTICAL ANALYSES

We shall consider first whether the chickenpox group had a significantly higher incidence of yaws in November than did those without chickenpox. There were 48 susceptibles among the former group in November of which 12 contracted yaws, for an incidence rate of 25% (see Table 2). On the other hand only 1 of 49 susceptibles in the non-chickenpox group contracted yaws in November for a rate of less than 2%. This difference is highly significant ($\chi_1^2 = 11{\cdot}02$, $P < 0{\cdot}001$). Turning to the incidences in January and February, we find 8·6% incidence of yaws among the chickenpox group (3 of the 35 remaining susceptibles), but 23·4% incidence of yaws among the non-chickenpox group. This difference approaches significance ($\chi_1^2 = 3{\cdot}12$, $P = 0{\cdot}08$). These results indicate a strong short-term association between chickenpox and yaws. These statistical tests should be adjusted to take into account the third variable of classification, the yaws history. Since the history status turns out to be distributed about proportionately between the other two classes, these conclusions are sustained by more sophisticated statistical tests which take into account the yaws history.

We turn to Table 1 to consider whether the distribution of cases could be described by a random process rather than the more mathematical model which we shall later consider. The simplest random model would be a Poisson process, for which the mean and variance are the same. We find for the distribution of cases by month

(excluding the unknown category) that the sample mean $\bar{x} = 31/7 = 4{\cdot}43$, while the sample variance $s^2 = 25{\cdot}62$. The Lexis ratio is $s^2/\bar{x} = 5{\cdot}78$, a highly significant departure from randomness ($\chi^2_6 = 6(s^2/\bar{x}) = 34{\cdot}71$, $P < 0{\cdot}001$). Since we are concerned particularly with the month of November, we shall compare the incidence of yaws in that month with the preceding months. The rate of cases per month for August to October inclusive is 2/3, while for November it is 13. Assuming Poisson variation for these months, we may test this difference by an F statistic (Cox, 1953) for which we find $F(5, 27) = 16{\cdot}2$, $P < 0{\cdot}001$. Taken together these results indicate these data cannot be adequately described by a Poisson process and that the month of November is exceptional when compared to the previous course of the epidemic.

REVIEW OF THE MATHEMATICAL MODEL

Bailey (1957) discusses a deterministic model for a simple epidemic with no removal which is 'approximately applicable to the sort of situation where (*a*) the disease is highly infectious but not sufficiently serious for cases to be withdrawn by death or isolation, and (*b*) no infective becomes clear of infection during the main part of the epidemic'. Yaws qualifies as a suitable disease for this model.

Let $X(t)$ be the number of susceptibles in the population at time t and let $Y(t)$ be the corresponding number of infectives. The basic model assumes the incidence is proportional to the product of the number of infectives and the number of susceptibles (i.e. the number of possible pairwise contacts). This is the 'homogeneous mixing' postulate and is mathematically written

$$\frac{dY(t)}{dt} = \beta\, Y(t)\, X(t), \tag{1}$$

where β is a constant infection rate. We solve equation (1) subject to the boundary conditions, $Y(0) = 1$ and $X(0) = n$; that is, we assume initially one infective and n susceptibles in a population of size $n+1$, which remains constant during the course of the epidemic. The resulting solution in terms of $Y(t)$ is the logistic (or autocatalytic) curve,

$$Y(t) = \frac{n+1}{1+ne^{-(n+1)\beta t}}. \tag{2}$$

This equation may also be written in the form*

$$L(t) = \frac{1}{n+1} \ln\left[\frac{nY(t)}{n+1-Y(t)}\right] = \beta t. \tag{3}$$

$L(t)$ will be called the logit at time t. We note that the logit graphs linearly against time and thus provides a simple test of the model. It is also of interest to note that $dY(t)/dt$, the epidemic curve, is a bell-shaped curve with maximum at

$$t_{\max.} = (\ln n)/[\beta(n+1)], \tag{4}$$

the time at which the incidence of the disease is the greatest.

* ln denotes natural logarithm.

https://doi.org/10.1017/S0022172400040730 Published online by Cambridge University Press

THE NUMBER OF SUSCEPTIBLES

In order to apply this model we must define a population size for the number of susceptibles. This size is not obvious and ascertainment of it is a difficulty associated with fitting other epidemic models (cf. Abbey, 1952). We cannot limit this population to those with a negative history of yaws since, in fact, seven cases also occurred in the positive history group. We consider two approaches to determining an 'effective population size'.

Assume that the positive history group can be divided into subpopulations of two discrete types. One subpopulation is wholly resistant to yaws and this does not contribute to the population of susceptibles. The second subpopulation is of the same degree of susceptibility as the negative history group in determining the effective population size. We estimate this size of this second subpopulation by assuming that the proportion having yaws in this subpopulation should be the same as that in the negative history group, that is, $\frac{29}{55}$ or 52·7%. All seven yaws cases would have occurred among the subpopulation and so we can solve back and find its size, namely $7 \times \frac{55}{29}$ or about 13. Thus, on the basis of the assumption, the effective population of susceptibles will be estimated at $55+13 = 68$.

A second argument assumes that two different infection rates β_1 and β_2 operate among the negative and positive history groups respectively. If we assume that $\beta_1/\beta_2 = \rho$ is large (compared to one) and that the number of cases among the negative group is small, it is possible to describe the epidemic by a modified form of (1) wherein $\beta_1 = \beta$ and the effective population size is

$$n+1 = n_1+\frac{n_2}{\rho}+1, \tag{5}$$

where n_1+1 and n_2 are the sizes of the negative and positive history groups respectively. The ratio ρ may be estimated by

$$\hat{\rho} = \frac{\ln(x_1/n_1)}{\ln(x_2/n_2)}, \tag{6}$$

where x_1 and x_2 are the number of remaining susceptibles in the two populations at the end of the epidemic.* For this example we find from (6) that

$$\hat{\rho} = \frac{\ln 54-\ln 26}{\ln 48-\ln 41} = 4{\cdot}636$$

and the effective population size estimated from (5) to be

$$n+1 = 55+\frac{48}{4{\cdot}636} \sim 65.$$

As can be shown to be the case under certain conditions, these two arguments lead to approximately the same results. We shall arbitrarily use sixty-eight as the effective population size although in the following analyses the qualitative results are not substantially affected by which of the two numbers is actually used.

* The mathematical details of this argument have been derived by one of the authors (J.J.G.) and will be published elsewhere.

https://doi.org/10.1017/S0022172400040730 Published online by Cambridge University Press

APPLICATION OF THE MODEL

In order to apply the model we must distribute the 5 unknown cases in Table 1 among the months after November. We arbitrarily assigned 1 of these to December, 2 to January and 2 to February, arriving at the totals given in Table 3. We shall see that the critical part of the analysis is invariant with respect to the specific allocation of these 5 cases among these 3 months. In Table 3 we have denoted the end of August by $t = 0$, the end of September by $t = 1$, etc. The last column of Table 3 gives monthly estimates of the infection rate given by taking differences among the adjacent logits, that is, $L(t) - L(t-1) = \hat{\beta}_t$, a result following from (3). We note that the infection rate for November is 0·03286 or over three times that

Table 3. *The adjusted data and the analysis using the model*

Month	t	New yaws cases (adj. totals)	$Y(t)$	$n+1 - Y(t)$	$L(t)$	$\hat{\beta}_t$
Aug. 1956	0	1	1	67	0·00000	—
Sept. 1956	1	1	2	66	0·01041	0·01041
Oct. 1956	2	0	2	66	0·01041	0·00000
Nov. 1956	3	13	15	53	0·04327	0·03286
Dec. 1956	4	3	18	50	0·04681	0·00354
Jan. 1957	5	6	24	44	0·05292	0·00611
Feb. 1957	6	12	36	32	0·06357	0·01065

for any other month. The arithmetic mean of the rates excluding November is 0·00614, while for the first 2 months it is 0·00520 and for the last 3 months it is 0·00677. It is of interest to note that the mean infection rates before and after the chickenpox epidemic are comparable. We also note that the mean infection rate between times t and $t+a$ may be alternatively calculated from the formula $\bar{\beta} = [L(t+a) - L(t)]/a$. Thus the mean infection rate for the months following November is invariant with respect to the assignment of the unknown cases among these months.

We now consider the projection of the epidemic from conditions prevailing before the chickenpox epidemic apparently accelerated the yaws epidemic. We substitute the mean infection rate for the first 2 months (0·00520) in (2) and find

$$Y(t) = \frac{68}{1 + 67e^{-0\cdot 354t}}. \tag{7}$$

The theoretical and observed epidemics are graphed in Fig. 1, this theoretical curve being the solid curve. It is noted that the projected epidemic would result in a total of about eight cases by the end of February, while in fact thirty-six were observed. These twenty-eight may be regarded as the excess due to the chickenpox epidemic. We may also use (4) to find the peak of the projected epidemic,

$$t_{\max.} = \frac{\ln 67}{68(0\cdot 00520)} = 11\cdot 9,$$

or about the end of August 1957. The epidemic was actually ended in March 1957 by the intervention of treatment. Moreover, August 1957 would have produced

https://doi.org/10.1017/S0022172400040730 Published online by Cambridge University Press

only eight new cases, while in fact peaks of thirteen and twelve cases were reached in November 1956 and February 1957 respectively.

We may also project the equation for the model from conditions prevailing at the end of November, after which time the effects of the chickenpox epidemic presumably no longer affected the course of the yaws epidemic. We solve (1) subject to the boundary conditions $Y(3) = 15$, $X(3) = 53$ and find

$$Y(t) = \frac{68}{1 + (\frac{53}{15}) e^{-68\beta(t-3)}}. \tag{8}$$

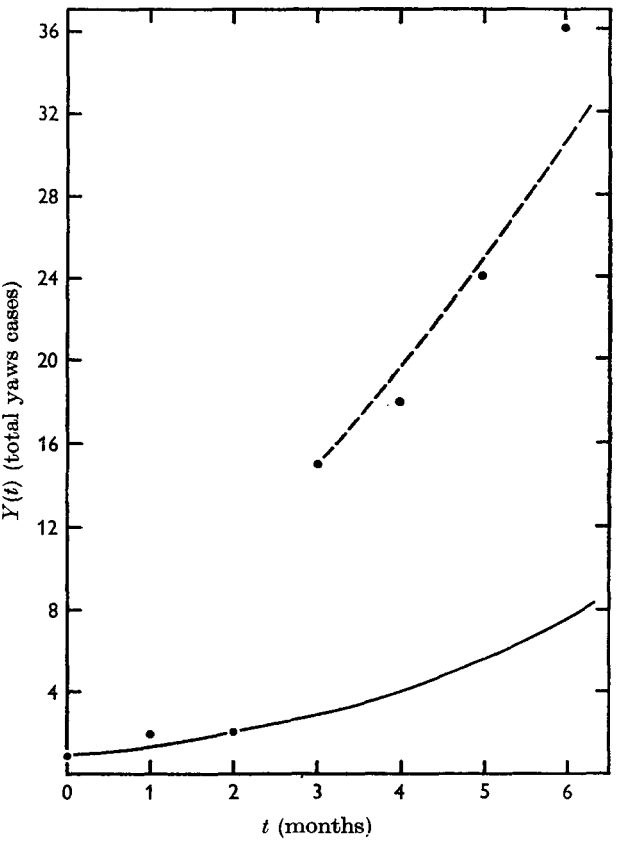

Fig. 1. The observed and projected courses of the yaws epidemic.
—, Projected from $t = 2$ (equation 7). - - -, Projected from $t = 3$ (equation 8).

We let β in (8) be the mean calculated from the first 2 months (0·00520) and find the theoretical curve given by the broken line in Fig. 1. We see that this curve approximately follows the actual course of the epidemic for this period. This reflects the fact that the average infection rates before and after November are roughly comparable.

The model as represented by (1) assumes that the population in question is subject to homogeneous mixing. We count all the possible contacts between infectives and susceptibles and make the incidence proportional to this number. This essentially weights all pairwise relationships between the individuals equally.

https://doi.org/10.1017/S0022172400040730 Published online by Cambridge University Press

Actually we would expect the individuals in the same household to have more contact than those, for example, living on opposite ends of the village. The first two cases in the epidemic of yaws illustrate the case. The second case was a twin sibling of the first case. The next thirteen cases in November were outside this household. Thus the observed course of the epidemic might then be used as an argument against the assumption of homogeneous mixing.

Let us consider this more carefully. The infection rate between twins would probably be the maximum of any situation. The rate between households (October and November) should be smaller, but in fact we observe a much larger rate for November, the month after the chickenpox epidemic. This would argue even more forcefully for the proposed hypothesis.

Because of the small numbers of cases in the first few months, the projections based on the model are not too accurate. One of the referees has suggested that the numbers of cases in the months of November to February inclusive might be considered to arise from a long, rather flat peak. Testing for homogeneity the number of cases in these months by using the Lexis ratio test on adjusted data of Table 3 (13, 3, 6, 12) we find $\chi_3^2 = 8{\cdot}12$, $P < 0{\cdot}05$. However, an arbitrary assignment of the unknown cases could equally well have led to the pattern of cases 13, 5, 6, 10. Then $\chi_3^2 = 4{\cdot}82$, $P = 0{\cdot}18$. In this case the contrast of the months of November and February versus December and January leads to

$$\chi_1^2 = \frac{[(13+10)-(5+6)]^2}{4(8{\cdot}5)} = 4{\cdot}24.$$

Thus the contrast between the two apparent peaks and the intervening months accounts for a very large part of the variation observed. The P value associated with the test of this contrast is 0·04, although it must be noted that this contrast, like the arbitrary assignment, was chosen after the fact.

All these results seem to point to the chickenpox epidemic as a factor in the acceleration of the yaws epidemic.

SUMMARY

The relationship between chickenpox and yaws epidemics occurring among children in a village in Netherlands New Guinea is analysed using the mathematical model for a simple deterministic epidemic. It is shown that the yaws epidemic accelerated significantly in the month following the chickenpox epidemic, but that it reverted to its previous rate in the succeeding months. The number of yaws cases attributable to the influence of the chickenpox is estimated from the projected course of the yaws epidemic. It is statistically verified that those children contracting chickenpox were more likely to become yaws cases in the subsequent month.

These results point to the danger of yaws's spread being much more rapid among a population which has recently been subject to an epidemic of chickenpox.

We are grateful to Prof. Philip Sartwell and a referee for useful suggestions.

REFERENCES

ABBEY, H. (1952). An examination of the Reed–Frost theory of epidemics. *Hum. Biol.* **24**, 201–33.

BAILEY, N. T. J. (1957). *The Mathematical Theory of Epidemics*, pp. 20–2. London: Charles Griffin and Co., Ltd.

COX, D. R. (1953). Some simple approximate tests for Poisson variates. *Biometrika* **40**, 354–60.

HACKETT, C. J. (1960). Some epidemiological aspects of yaws eradication. *Bull. Wld Hlth Org.* **23**, 739–61.

HACKETT, C. J. (1963). On the origin of the human treponematoses. *Bull. Wld Hlth Org.* **29**, 7–41.

HACKETT, C. J. & GUTHE, T. (1956). Some important aspects of yaws eradication. *Bull. Wld Hlth Org.* **15**, 869–96.

KRANENDONK, O. V. (1958). Serological and epidemiological aspects in yaws control. Thesis, University of Amsterdam.

REPORT (1960). Fifth report of the expert committee on venereal infections and treponematoses. *Tech. Rep. Ser. Wld Hlth Org.* no. 190, p. 36.

SOETOPO, M. & WASITO, R. (1953). Experience with yaws control in Indonesia. In First International Symposium on Yaws Control. *Monograph Ser. W.H.O.* no. 15, p. 283.

TURNER, T. B. & HOLLANDER, D. H. (1957). Biology of the treponematoses. *Monograph Ser. W.H.O.* no. 35.