Epidemiol. Infect. (1993), **111**, 503–523

Malaria transmission rates estimated from serological data

M. N. BURATTINI[1,2], E. MASSAD[1] AND F. A. B. COUTINHO[3]

[1] *Discipline of Medical Informatics, School of Medicine, the University of São Paulo, Av. Dr. Arnaldo 455, São Paulo, CEP 01246–903, Brazil and LIM 01, HCFMUSP, São Paulo, Brazil*

[2] *Discipline of Infectious and Parasitic Diseases, Escola Paulista de Medicina*

[3] *Institute of Physics of the University of São Paulo, CP 20516, 01498, São Paulo, Brazil*

(*Accepted 3 June 1993*)

SUMMARY

A mathematical model was used to estimate malaria transmission rates based on serological data. The model is minimally stochastic and assumes an age-dependent force of infection for malaria. The transmission rates estimated were applied to a simple compartmental model in order to mimic the malaria transmission.

The model has shown a good retrieving capacity for serological and parasite prevalence data.

Section 1. Introduction

The historical tradition of modelling malarial transmission based on entomological data has made such an approach a paradigm. Since the seminal work of Ross [1–4], through the developments of Macdonald [5–7] and until the celebrated Dietz-Molineaux-Thomas model [8], entomological parameters have been the centre of the majority of models dealing with the transmission dynamics of malaria. Indeed the two most important transmission concepts, namely Macdonald's 'Basic Reproductive Rate' [6] and Garret-Jones' 'Vectorial Capacity' [9] depend heavily on entomological data. To obtain the necessary data for their estimations, however, requires a huge amount of field work, making such an approach of poor practical feasibility. Counting mosquitoes, examining their stomachs for human blood and dissecting their salivary glands in the search for sporozoites are difficult and expensive procedures [10, 11]. A recent review on the entomological data required to estimate transmission, as well as the difficulties related to their estimation can be found in Dye [12].

Due to the difficulties related to entomological procedures, parasitological data have been chosen as a 'gold standard' to assess malarial transmission. However, such an approach has also limitations, such as the influence of the widespread usage of antimalarial drugs, the effect of concomitant immunity on parasitic levels, and the high seasonal variation of the vector population [11].

In fact, the so-called parasite rate, that is, the proportion of individuals with circulating parasites, which has been used as a gold standard in the measurement of malaria transmission, has poor sensitivity due to the great variation over time

in the parasite level. This variation reflects the relatively short period of parasitaemia. In addition, the indiscriminate use of antimalarial drugs, which decreases the quantity of circulating parasites, the seasonal variation in the population of the vector, and the immunity level of the affected individuals are additional causes of errors in the estimation of the actual parasite rate. In order to circumvent these limitations, the World Health Organization recommends the use of an averaged value taken repeatedly as many times as possible over the whole year [11]. To keep trained personnel for a whole year under field conditions, however, makes this kind of approach extremely expensive. Therefore a more practical way to estimate malarial transmission is still required [11].

With the development of reliable serological techniques the use of antibody prevalence data became an interesting candidate as an indicator of malarial transmission levels. In fact, it is possible to estimate the effective inoculation rate from serological data [13–15]. This parameter is usually defined as the number of infective bites which results in parasitaemia in a non-immune individual, per unit of time. Draper, Voller and Carpenter [13], probably the first investigators to apply serological data to a mathematical model in order to estimate the infection rate, pointed out that in serological surveys one is collecting *period* prevalence data, the total experience of malaria in a community, which is in contrast to the *point* prevalence derived from parasite data.

Draper and colleagues [13] developed a rough first model in which the serological data were plotted against age in an inverse logarithmic scale. From this relation they extracted the so-called R parameter, the probability of being infected each year. Van Druten's analysis [15] is further refined by applying a catalytic model to cross-sectional serological data. Both authors interpret this age variation in the effective inoculation rate as being solely due to changes of malaria transmission in time, reflecting a past situation.

In this paper we propose a model to assess the effective inoculation rate, h, from cross-sectional serological data. The model considers an age-dependent inoculation rate, the 'boosting effect' and the fading of antibodies in the absence of further exposure to plasmodia. We also propose a simple compartmental structure in order to provide a tool by which we can test the inoculation rate deduced as applied to real epidemiological data. This compartmental structure could also be used to predict in a simple, although still accurately enough way the age-related profile of the parasite and serological prevalence data.

This paper is organized as follows: In section 2 we present a method to estimate the inoculation rate from cross-sectional serological data, and also the rate of immunity loss, γ. Although the inoculation rate is generally age and time dependent, in this paper we consider estimations in which either one or the other of these dependencies can be neglected.

In section 2 we also present a four-compartment deterministic model designed to serve as a tool to check the accuracy of the estimates and to describe in a simple, though still accurately enough way, the malarial dynamics.

In section 3 we estimate the inoculation rate, h, and the rate of loss of antibodies in the absence of further inoculations, γ, from particular areas with distinct patterns of endemicity. These estimates are intended to exemplify the theory developed in section 2. Next we apply these estimates to the compartmental model in an attempt to retrieve field data from real areas.

In the appendices we elaborate the mathematical arguments involved in the estimation of the inoculation rate.

The results presented in this paper are encouraging in providing an alternative, feasible and sufficiently accurate way to quantify malarial transmission.

Section 2. The model

The basic aim of this paper is to estimate the effective inoculation rate (also called dependent happening [1–4]), $h(t, a)$ from serological data. Basically $h(t, a)\,da$ means the probability, at time t, that an individual may receive an inoculation while his age is between a and $(a+da)$. The inoculation is here defined as 'the infective bite which results in parasitaemia in a nonimmume individual'. We therefore, as in some previous works [8, 13, 15], do not take into account the cycle of transmission in the vector population.

We begin by considering the population divided into two states: (1) individuals who are negative to the serological test, i.e. those who have never had an inoculation and those who having had a previous inoculation have already lost their specific antibodies, and (2) individuals who are positive to the serological test. i.e. those with at least one previous inoculation and who have not yet lost their antibodies.

Our basic input is $P^+(t, a)$, the proportion of individuals with circulating antibodies to malaria parasites, detected in an age-related cross-sectional survey. We propose to estimate the classical 'happening' factor $h(t, a)$ [1–4], from $P^+(t, a)$.

Actually, malarial transmission is both time and age-dependent. However, in some endemic areas, either one of these dependencies may be neglected. When transmission has not changed dramatically over the period of time corresponding to the age of the eldest individual in the surveyed population, we have an equilibrium transmission situation and the time dependence can be neglected; on the other hand, there are areas in which the temporal variation has been so intense that age-dependence is negligible. For the purpose of this section we will consider a steady state situation and hence drop out the time dependence.

We assume that inoculations occur in a non-uniform Poisson fashion [16]. This implies a non-constant inoculation rate. Individuals are assumed to reach a variable immunity level after receiving an inoculation, which is sufficient to result in a positive response to the serological test. In addition, we are considering the possibility of a 'boosting effect', that is the sudden increase in antibodies levels in previously exposed individuals, and individual variation in the immune response. We are neglecting the effects of the sensitivity and the specificity of the serological test in the analysis.

In the Appendix I we show that the relationship between $P^+(a)$ and $h(a)$ is given by:

$$P^+(a) = \int_0^a h(a')\,da' e^{[-\nu(a)+\nu(a')]} f(a-a') \tag{1}$$

where

$$e^{[-\nu(a)+\nu(a')]} = e^{-\int_{a'}^{a} h(s)ds} \tag{2}$$

and $f(a-a')$ describes the population process of losing antibodies. Equation (2) gives the probability of not receiving any inoculation between ages a' and a.

Equation (1) means that the probability of being seropositive at age a is a

https://doi.org/10.1017/S0950268800057241 Published online by Cambridge University Press

product of the probabilities of having received the last inoculation at age a' and of not have had sufficient time for antibodies to decay, integrated for all ages.

Given $P^+(a)$ from the cross-sectional serological data, $h(a)$ can be determined from equation (1) once $f(a-a')$ is known.

We have assumed $f(a–a')$ as a function of the interval between the last inoculation and the instant of time of the serological survey. A simple form for $f(a-a')$ is:

$$f(a-a') = e^{[-1/\tau(a-a')]}, \tag{3}$$

where τ is a constant. Hence in the absence of any new inoculation, antibodies fade away and the proportion of positive people decay in an exponential fashion with the increasing interval between the last bite and the survey time.

Assuming (3) and solving (1) for $h(a)$ gives:

$$h(a) = \frac{[(P^+(a)/\tau) + (dP^+(a)/da)]}{[1-P^+(a)]}. \tag{4}$$

In order to provide a tool by which we could test $h(a)$ as applied to real epidemiological data we designed a compartmental model. This model is also intended to describe in a simple, although sufficiently accurate way, the dynamics of malaria. As we will see the model provides estimates of the parasite and serological prevalences, which reproduce field data.

In our model the total population is divided into four compartments, namely, 'susceptibles', 'infected but yet seronegatives', 'infected seropositives', and 'immunes', represented by $X(t,a)$, $Y(t,a)$, $Y'(t,a)$ and $Z(t,a)$, respectively, where a stands for age and t for time, both in years. The compartments $Y(t,a)$ and $Y'(t,a)$ represent the 'parasite positive' fractions of the population ('prevalence'), and $Y'(t,a)$ and $Z(t,a)$ represent the fraction of the population with 'malaria antibodies' detectable in a serological test ('seroprevalence'). Hence people in $Y'(t,a)$ are both 'parasite' and 'seropositive'. Figure 1 shows the block diagram representing the compartmental structure of the model.

The four compartments described in Fig. 1 can be identified under field conditions. However, their magnitudes are subject to errors, particularly in the determination of the parasite rate, i.e. the proportion of individuals positive to parasitological tests, as discussed above.

The model has a dynamic form described by the following set of partial differential equations:

$$\frac{\partial X}{\partial t} + \frac{\partial X}{\partial a} = -h(t,a)\,X(t,a) + rY(t,a) + \gamma Z(t,a), \tag{5}$$

$$\frac{\partial Y}{\partial t} + \frac{\partial Y}{\partial a} = h(t,a)\,X(t,a) - (r+\delta)\,Y(t,a), \tag{6}$$

$$\frac{\partial Y'}{\partial t} + \frac{\partial Y'}{\partial a} = \delta Y(t,a) - \phi Y'(t,a), \tag{7}$$

$$\frac{\partial Z}{\partial t} + \frac{\partial Z}{\partial a} = \phi Y'(t,a) - \gamma Z(t,a), \tag{8}$$

where $h(t,a)$, r, δ, ϕ and γ are the transition rates between the compartments described above.

https://doi.org/10.1017/S0950268800057241 Published online by Cambridge University Press

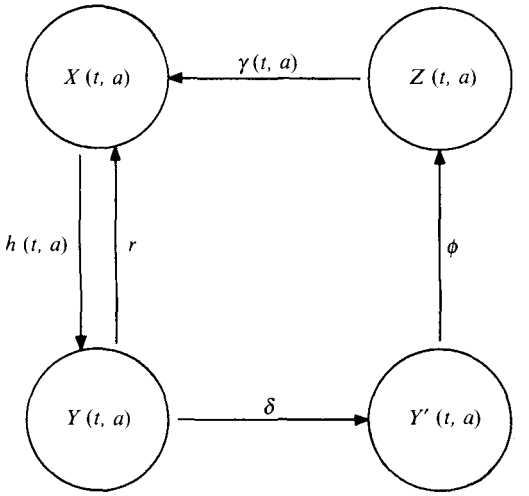

Fig. 1. Block diagram representing the compartmental structure of the model. X represents the susceptible fraction of the population; Y, the parasite positive but yet seronegative fraction; Y', the parasite positive already showing circulating antibodies; and Z, the seropositives recovered from parasitaemia. The rates $h(t, a)$, r, δ, ϕ and $\gamma(t, a)$ represent the transitions rates of the model as discussed in the text.

In the Appendix I we show how γ, the rate at which people leave the seropositive status, is derived. For the simple form chosen for $f(a-a')$, given by equation (3) we get:

$$\gamma = \frac{1}{\tau}. \tag{9}$$

This result follows from the simplicity of the assumption relating the way individuals lose antibodies in the absence of further inoculation. However, equation (9) can be solved for other forms of $f(a-a')$. Equations (4) and (9) will be applied to real epidemiological data.

In section 3 we solve numerically the above system of equations. For this we use $h(a)$ deduced from (4) and γ from (9), which of course are only approximations since they are deduced from a two compartment structure. However, by setting the spontaneous recovery rate, r, equal to zero we will see that our estimates are almost exact. In fact, though the spontaneous recovery rate is a biological possibility, its value is negligible for most endemic areas. We believe, and the results will show it to be correct, that this is a good approximation. Furthermore, the period of time an individual takes to acquire circulating parasites and thereafter become seropositive is very short (days) as compared to the time antibodies take to decay (years).

The transitions rates δ and ϕ are arbitrarily set constants and chosen in order to best fit the data. The limitations of this assumption will be discussed later.

Section 3. Testing the model

In order to check the retrieving capacity and reliability of the model, we fitted the data from areas of distinct levels of endemicity.

The first set of data is derived from the Bioco Island (Equatorial Guinea) survey, described by Merlin and colleagues [17]. This area was chosen as a typical

https://doi.org/10.1017/S0950268800057241 Published online by Cambridge University Press

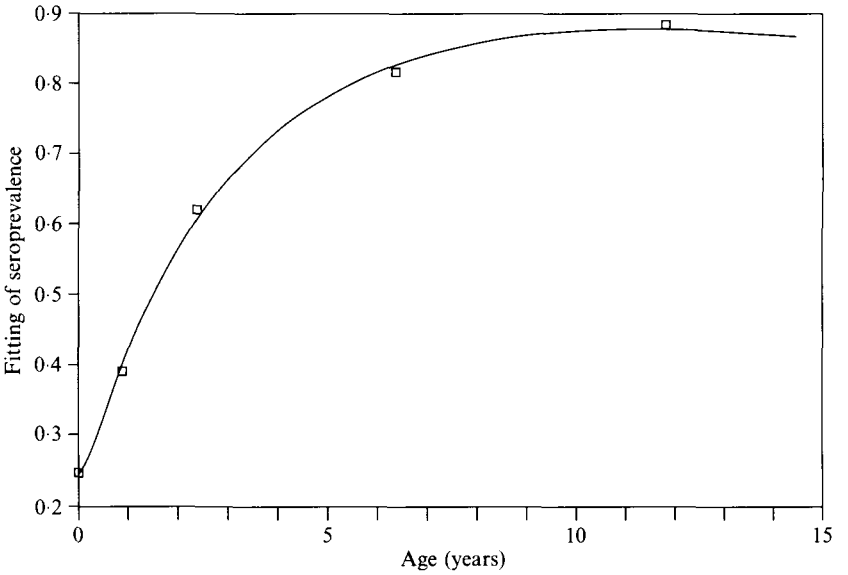

Fig. 2. Result of the fitting procedures for the serological data from the Bioco area. Squares represent the actual values whereas the continuous line represents the function fitted to data [equation (17)]. The fitting parameters are $b_0 = 0{\cdot}305$: $b_1 = 0{\cdot}190$; $b_2 = 0{\cdot}125$; and $b_3 = -0{\cdot}0438$.

situation in which to apply this analysis. It is characterized by a steady-state situation in malaria transmission with practically no control measures. Even the self medication index is reported as nil for this area.

The first step of the procedure is to fit the serological data to a continuous function.

For this particular area the fitting equation has the form:

$$P^+(a) = b_0 + b_1\, ln(a) + b_2[ln(a)]^2 + b_3[ln(a)]^3, \tag{10}$$

where b_i are the fitting parameter.

Equation (10) is a form of the well known logistic function, widely used in biology and acknowledged as the best shape for an age-structured serological profile [18]. We are well aware of the singularity in equation (10) at time zero. On the other hand, for the kind of analysis proposed it is meaningless to consider the period below 6 months of age due to the presence of maternally derived antibodies. Figure 2 shows the fitting accuracy of the above equation.

It should be noted that equation (10) fits the raw data with a high degree of accuracy.

Secondly we estimated $h(a)$ as in equation (4) for $P^+(a)$ as in equation (10), with τ set as equal to 10 years. This value of τ is reasonable when compared with data found in the literature [19–21]. Figure 3 shows the curve of $h(a)$.

The results obtained for the estimation of $h(a)$, shown in Figure 3, corresponds in its magnitude to a meso-to hyperendemic pattern of transmission, which is indeed the case for this area. It is also highly significant that the $h(a)$ profile found is in accord with the age-dependent entomological observations for *A. gambiae*, the main vector involved in malaria transmission for that area [22–24].

https://doi.org/10.1017/S0950268800057241 Published online by Cambridge University Press

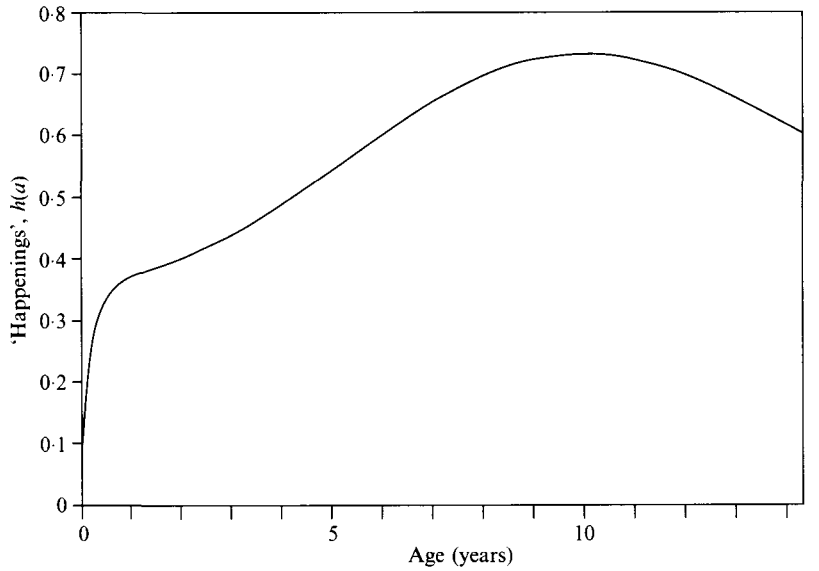

Fig. 3. Estimate of the age-dependent 'happening factor', $h(a)$, for the Bioco area.

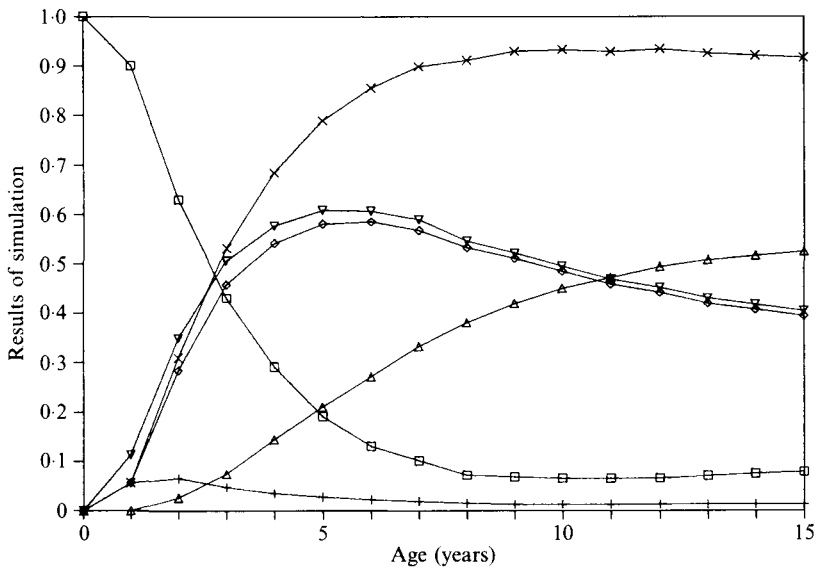

Fig. 4. Results of the numerical simulation of the model for the Bioco area, showing the proportion of individual in each compartment and the serological and parasitological levels. Squares represent the proportion of 'susceptibles', crosses the 'infected but yet seronegatives', diamonds the 'infected seropositives', triangles the 'seropositives recovered from parasitaemia', ×'s the 'seroprevalence', and inverted triangles the 'parasite prevalence'.

We then applied $h(a)$, and γ as estimated from equation (9), to the compartmental model given by equations (5–8). The two other rates, namely δ and ϕ. were set as constants equal to 4·0 and 0·15, respectively. The results of the integration of equations (5–8) can be seen in Figure 4, which shows all the compartments simultaneously, in order to illustrate the model outcomes.

https://doi.org/10.1017/S0950268800057241 Published online by Cambridge University Press

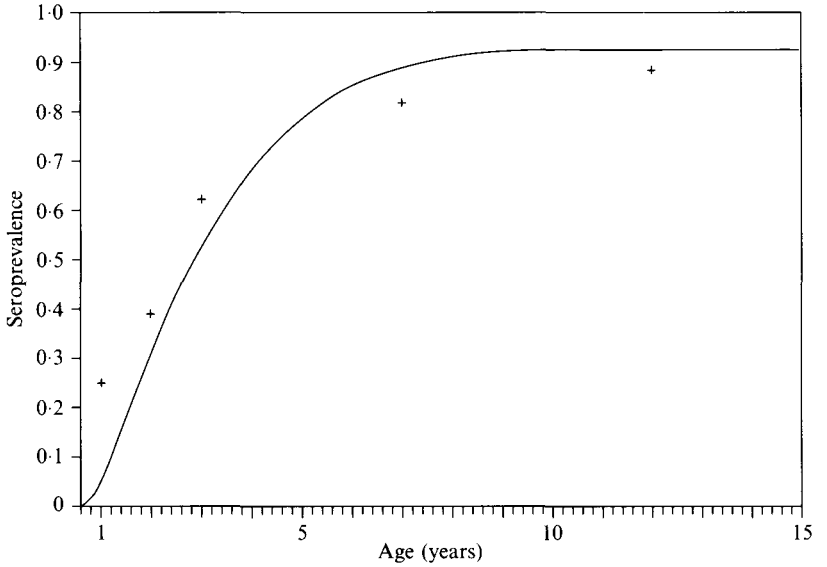

Fig. 5. Retrieved seroprevalence levels for the Bioco area. Crosses represent the actual values and continuous line the result of the numerical simulation of the model.

Figure 4 shows some interesting features whose interpretation could be applied in planning control strategies. The age variation in the proportion of susceptibles can be applied for designing mass chemoprophylaxis programmes. Since the susceptibles curve decreases up to the age of 8 years and remains stable at low levels for the older age groups, only those younger than that should be the target fraction of the population eligible for such a programme.

In addition, it can also be noted that the model provides estimates of the fraction of the population with parasitaemia and without a detectable immune response. If we assume that these individuals are those at major risk for developing severe illness, with higher mortality rates, then this will assist in the planning of the allocation of resources in the malaria control programmes.

Of course these comments should be interpreted with caution and are intended only to exemplify possible uses of the model. However, it should be stressed that the age profile and magnitudes of such estimates are in agreement with the morbidity and mortality rates due to malaria in areas with that endemic pattern.

Figures 5 and 6 show the retrieving of serological and parasite curves, respectively. It should be noted that the retrieved serological and parasitological curves are in good agreement with the epidemiological observations for the area.

Other examples of the retrieving capacity of the model are presented in the following analysis.

We fitted the data from the 1972 immunofluorescence antibody (IFA) survey to *P. falciparum* from the baseline period of the Garki Project [25], with equation (10). The transmission of malaria can be considered as in a steady-state, with virtually no variation in the period of time corresponding to the age of the eldest component of the studied community [25].

This is a particularly challenging area for sero-epidemiological analysis, due to the extremely high levels of endemicity. It was chosen in order to illustrate the

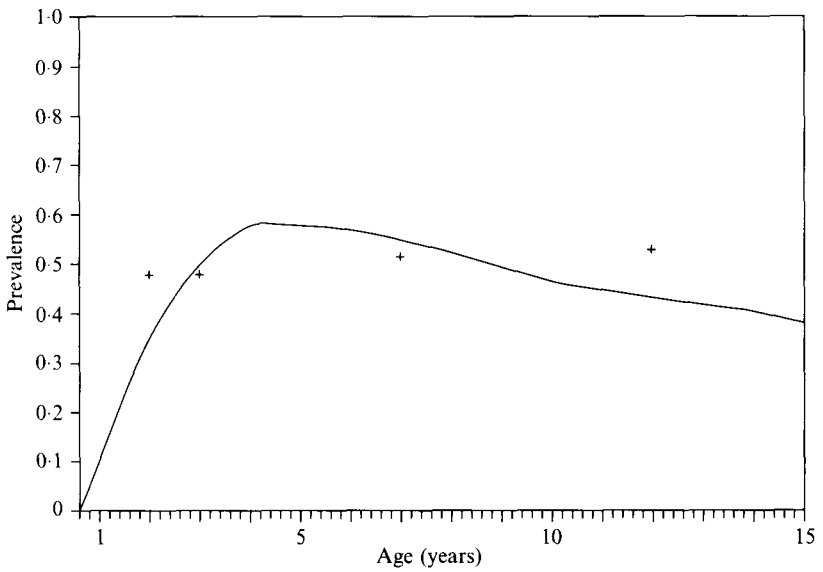

Fig. 6. Retrieved parasite prevalence levels for the Bioco area. Crosses represent the actual values and continuous line the result of the numerical simulation of the model.

difficulties of the method described in this paper when applied to areas with such levels of endemicity. Usually, when faced with 100% of seropositivity (which is indeed the case for this area), authors either dispose the point of [13] or declare arbitrarily one individual of the sample as negative [15]. We chose as another approximation to consider 99% of positives for those points. Any of the above treatments, however disputable, are epidemiologically justifiable depending on the area. In addition, for the Garki 1972 IFA survey there may be an overestimation of seropositivity due to the low threshold positive titre chosen, 20, which should be compared with the one used in the Bioco area, 100.

The age-dependent force of infection, $h(a)$, and the rate of immunity loss, γ, were estimated with equations (4) and (9) and were applied to the compartmental model given by equations (5–8). The transition rates δ and ϕ were assumed to be equal to 6·00 and 0·25, respectively. Figure 7 shows the curve of $h(a)$. The result of the estimations for $h(a)$ for this area is also in good agreement with the entomological results. For instance, our average h is 10·8 years^{-1}, which should be compared with the averaged entomological result for the effective inoculation rate, 8·3 years^{-1}, calculated with the parameters described in the Garki Project. Figures 8 and 9 show the retrieving of serological and parasite curves, respectively.

As can be noted in Figure 8 the serological profile retrieved by the model is in good agreement with the raw data. In contrast, for the parasitological curve, the model underestimates the real data below 20 years. This is probably due to the assumption related to the δ and ϕ rates, considered as constants. Actually, these rates are expected to increase with age. We believe that the observed differences would probably be less marked if we had taken these age-dependencies into account. The differences between this and the Bioco area concerning this point will be further considered in the discussion section.

https://doi.org/10.1017/S0950268800057241 Published online by Cambridge University Press

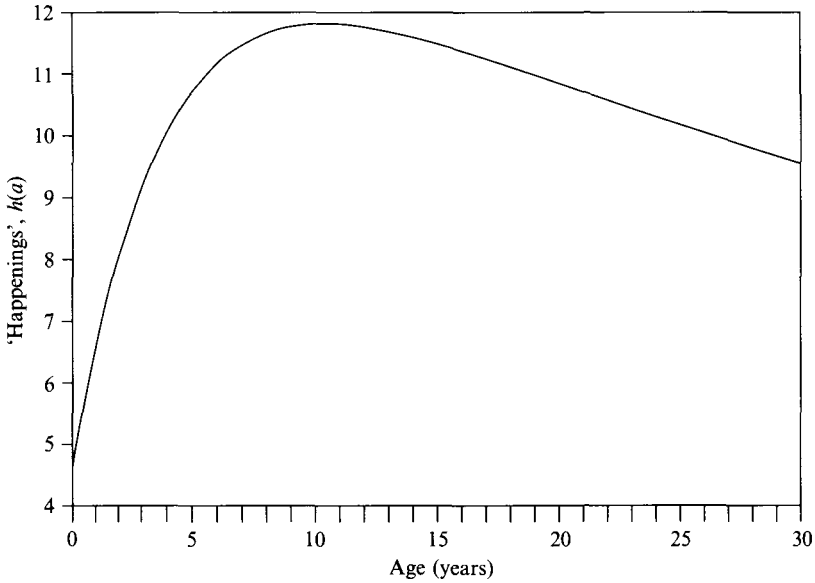

Fig. 7. Estimate of the age-dependent 'happening factor', $h(a)$, for the Garki area.

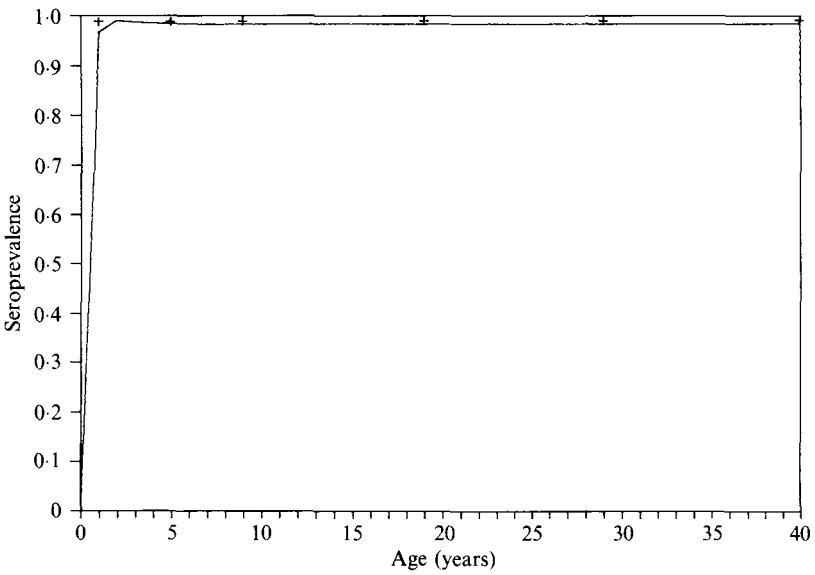

Fig. 8. Retrieved seroprevalence levels for the Garki area. Crosses represent the actual data and continuous line the result of the numerical simulation of the model.

The third set of data intended to exemplify the model's application involves an alternative interpretation. It is related to the Mauritius Island study [26]. This region is characterized by a successful control programme, with levels of transmission virtually nil at the time of survey. Hence, according to our assumptions, age-dependence can be neglected.

The serological data were fitted by a third degree polynomial with respect to time. As in the other areas, we calculated $h(t)$, shown in Figure 10, and applied it

https://doi.org/10.1017/S0950268800057241 Published online by Cambridge University Press

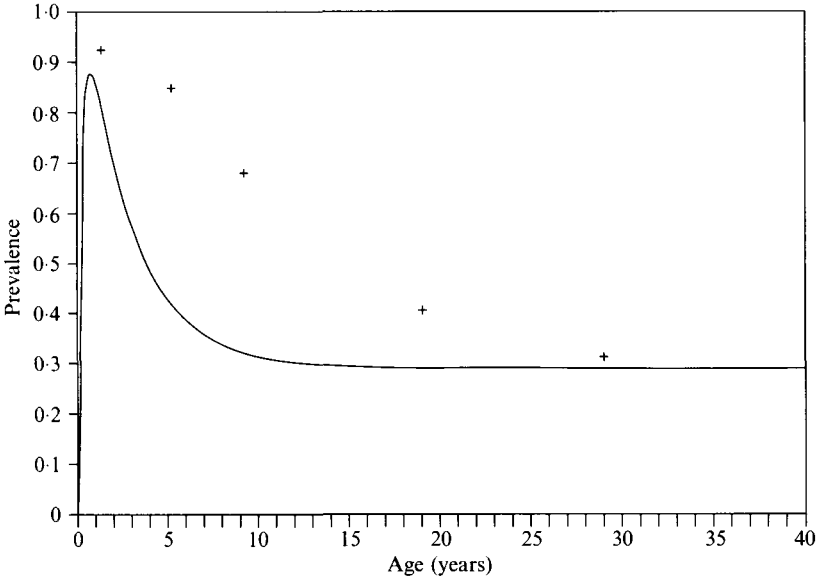

Fig. 9. Retrieved parasite prevalence levels for the Garki area. Crosses represent the actual data and continuous line the result of the numerical simulation of the model.

to the set of equations (5–8). The transition rates δ and ϕ were assumed to be equal to 1·00 and 0·20, respectively.

The retrieved serological profile, now interpreted as a time variation, is shown in Figure 11. Our results reflects a mesoendemic situation 40 years before the survey, turning to hypoendemic some 20 years and to virtually nil in the last 3 years before the 1972 survey. These findings are in agreement with the historical endemicity of malaria known for the area [26].

Section 4. Discussion

Basically, malarial transmission can be assessed directly through entomological data, and indirectly through serological and parasitological data.

As mentioned in Section 1, Introduction, intrinsic difficulties of the direct approach make it of poor practical use. Notwithstanding, the models based on entomological data have provided important insights into the comprehension of the dynamics of malarial transmission [10, 27]. Actually the best indicators of the transmission dynamics of malaria are the Basic Reproductive Rate, R_0 [6], the Vectorial Capacity, C [9], and the inoculation rate, h [6, 7]. All these indicators were deduced from entomological parameters. However, the need of continuous monitoring of entomological parameters have made those studies unfeasible in control practice. Therefore, alternative ways for the estimation of these transmission indicators are necessary.

The use of the parasite rate and spleen rate, the classical malariometric indexes, as alternatives to entomological data have also intrinsic limitations, which have been extensively discussed elsewhere [11, 28].

With the development of reliable serological techniques the use of cross-sectional antibodies prevalence data became an interesting candidate as an indicator of malarial transmission levels. This can be done either by identifying

https://doi.org/10.1017/S0950268800057241 Published online by Cambridge University Press

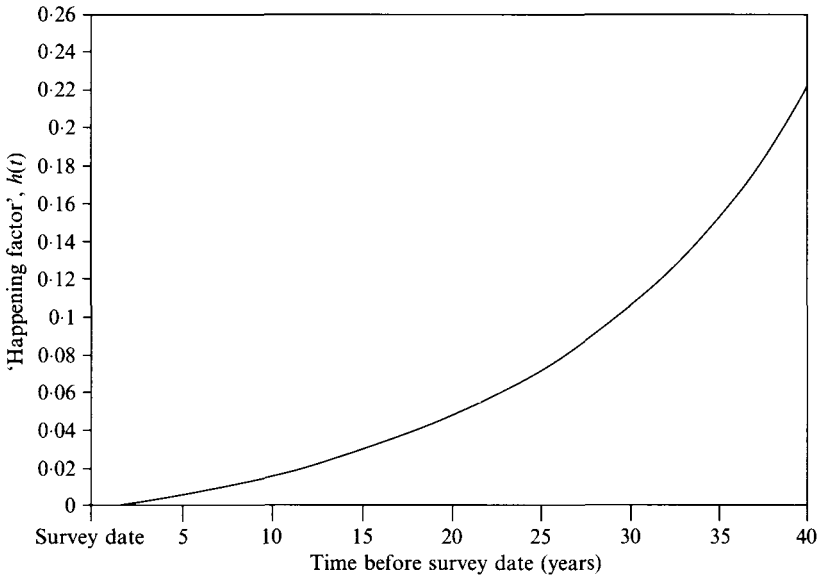

Fig. 10. Estimate of the time-dependent 'happening factor', $h(t)$, for the Mauritius area. Here time is counted backwardly as related to the survey time.

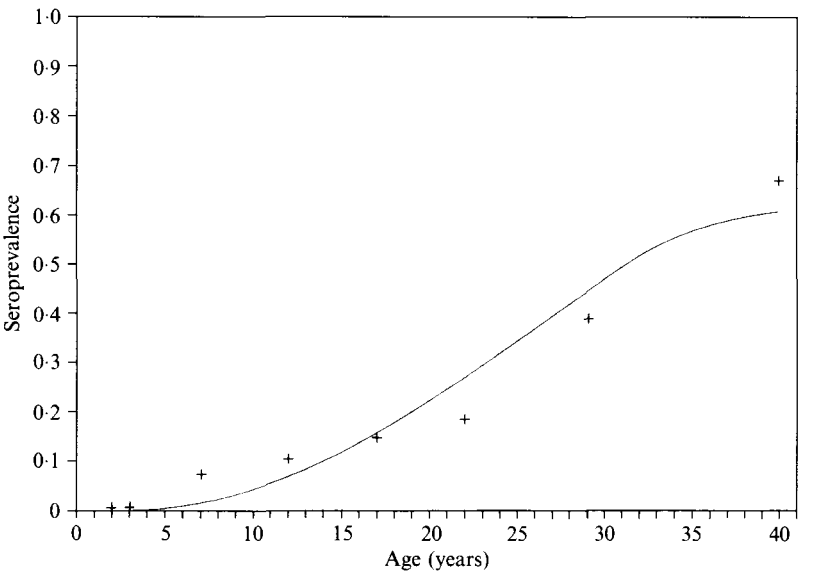

Fig. 11. Retrieved seroprevalence levels for the Mauritius area. Crosses represent the actual data and continuous line the result of the numerical simulation of the model.

stage-specific antigens that could be related to transmission intensity [29] or by estimating the effective inoculation rate from cross-sectional seroprevalence surveys [13–15].

Draper and colleagues [13] developed their model in which they estimate the probability of being infected each year. Van Druten's analysis [15] is further refined by applying a catalytic model to cross-sectional serological data, and attempts to deal with the loss of immunity. Both authors interpret the age

variation in h as been solely due to changes in time of malaria transmission, reflecting a past situation.

This paper is an attempt to improve the previous analysis of the above works by considering the age-dependence in malaria transmission, and the individual variation in the immune response to plasmodia, such as the 'boosting effect' and the fading of antibodies.

Our assumption that the inoculation rate is age dependent is supported by entomological data, which has been recognised since the mid 1950s [22–24]. Indeed, Carnevale [22] found a bite preference by mosquitoes three times higher in adults than in children, a result which roughly coincides with our findings (see Figure 3). Therefore, we consider h as both time and age dependent, but in this paper we analyse only situations in which either the time or the age-dependence can be neglected for practical purposes. As mentioned above, in the cases where the time dependence could be neglected our results concerning $h(a)$ are supported by previous entomological works on the feeding habits of anophelenes.

The individual variation in the immune responses were treated stochastically. The results obtained apply to a population rather than to individuals. This may seem crude. However, the lack of sufficient experimental data on the immune response justifies our approach. Appendix II shows a more detailed analysis of the above.

Estimating $h(t, a)$ according to equation (4), from serological data, we have actually considered a two compartments model, namely, seronegatives and seropositives. Notwithstanding, our results were applied to a four compartments model, chosen in order to mimic, as best as possible, and in a practical way, the actual clinical states of individuals. In addition, all the four compartments considered can be identified in the field. Moreover, as mentioned above, by setting $r = 0$ we guarantee that the use of $h(a)$, as deduced in equation (4), is almost exact. Furthermore, the average period of time necessary to raise a positive response to a serological test is negligible (days) when compared with the time antibodies take to decay (years). Also, the absence of superinfection is comprised by the two compartments assumption. However, as shown by Dietz [27], the prevalence curves retrieved by models with and without superinfections show practically no difference.

We are well aware of the existence of models with more, or less, compartments, including, for instance, subpopulations of individuals who are infected but are not infective to mosquitoes, individuals who have a slow recovery rate and a fast recovery rate, and others. Some of these models have shown a good accurate fit like that of Dietz, Molineaux and Thomas [8], which retrieved the parasite prevalence curve from entomological inputs. Another example is Aron's [30] model, which proposed a three compartments structure, also retrieving the trends in the parasite prevalence curves with good accuracy. However, none of the models proposed so far allows the retrieving of seroprevalence curves.

As far as the results are concerned, some further points should be made. Despite some over simplifications, the model has shown a good retrieving capacity of the original data. It should be expected that the rates δ and ϕ would increase with age, due to the cumulative experience with plasmodia. Actually Dietz (unpublished work cited by Nedelman [31]) has already shown that the recovery rate in his

https://doi.org/10.1017/S0950268800057241 Published online by Cambridge University Press

model increases with age. On the other hand, γ is expected to decrease with age. The age dependencies of these rates, if considered, would probably result in a better agreement of the parasite curve with the real situation. Such an effect is probably due to very intense contact with plasmodia, which triggers the individual immune response at an early age. Thus, this effect should be more evident in areas with particularly high levels of endemicity, such as Garki. The better retrieving capacity of the parasitological data observed for Bioco (a meso to hyperendemic area), as compared with Garki (a holoendemic area) strongly supports the hypothesis. Note that the values of the inoculation rate estimated for Garki are approximately 16 times greater than those for Bioco. Also, the classical definitions of holo versus hyperendemic endemicity of malaria implicity assumes these mechanisms. However, the lack of appropriate field data and the present stage of knowledge as far as the immunity process is concerned, make any attempt to include these transition rates variations premature.

We are well aware that the falls in the $h(a)$ curves after reaching their maximum for the Bioco and Garki data are, to a certain degree, dependent on the falls in the fitted curve of the serological data. However, we are sure that these shortcomings do not compromise our analysis.

The history of the mathematical approach to malaria is nearly as old as the discovery of its mode of transmission [27]. However, views on malaria control have changed since then and, to a certain degree, so has the emphasis on quantitative studies [12]. But the theory continues to be, in a sense, inappropriately applied, particularly the quantification of transmission based on entomological data. Therefore, as mentioned by Dye [12], a shift of emphasis is needed, and the proposed comparative approach suggested by that author provides more efficient solutions to the same range of questions related to malaria transmission. Molineaux [10] argued that dynamic malaria models are employed most successfully when asking comparative rather than absolute questions.

This work is an attempt to provide alternative ways to estimate malaria transmission that should be compared with the classical entomological and parasitological approaches.

An extension of this work dealing with the estimation of the inoculation rate to regions in which both time and age dependencies have to be considered [32] will be presented in a future paper.

In the accompanying paper [33] are the results of the method described here as applied to an area where the control of malaria has been impaired by the difficulties in the determination of the real pattern of transmission as assessed by the classical parasitological and splenic indexes [34–36].

APPENDIX I

In this appendix we show how to derive equations (1) and (9) of the main text.

Inoculations are assumed to occur as a non-uniform Poisson process [16]. The host population is assumed to be in a steady-state and the differential mortality due to malaria is neglected. Therefore the probability of receiving at least one inoculation at age a is:

$$P_a(n \geqslant 1) = 1 - e^{-\int_0^a h(s)ds}. \tag{A 1}$$

https://doi.org/10.1017/S0950268800057241 Published online by Cambridge University Press

Let us suppose that the fraction of the population aged a at the time of the serological survey received an inoculation in the period between ages a' and $(a'+da')$ with a probability determined by the Poisson distribution of inoculations:

$$h(a')\,da'. \tag{A 2}$$

The probability of not receiving any inoculation between $(a'+da')$ and a is:

$$e^{-\int_{a'}^{a} h(s)ds} = e^{[-\nu(a)+\nu(a')]}. \tag{A 3}$$

So the probability of having received the last inoculation between a' and $(a'+da')$ is:

$$h(a')\,da'e^{[-\nu(a)+\nu(a')]}. \tag{A 4}$$

Assuming that all individuals reach an immunity level above the positive threshold after an inoculation, and that in the absence of further inoculations that immunity level decays, the probability of still having detectable antibodies at age a is given by:

$$h(a')\,da'e^{[-\nu(a)+\nu(a')]}f(a-a'), \tag{A 5}$$

where $f(a-a')$ describes the population process of losing immunity.

Therefore the probability of an individual of age a being seropositive at the survey time is:

$$P^+(a) = \int_0^a h(a')\,da'e^{[-\nu(a)+\nu(a')]}f(a-a'), \tag{A 6}$$

Equation (A 6) is a 'Volterra Integral Equation' of second kind for $h(a)$ [37]. Given $P^+(a)$ from the cross-sectional serological data, $h(a)$ can be determined once $f(a-a')$ is known.

We have assumed $f(a-a')$ as a function of the interval between the last inoculation and the instant of time of the serological survey. A simple form for $f(a-a')$ is:

$$f(a-a') = e^{[-\frac{1}{\tau}(a-a')]}, \tag{A 7}$$

where τ is a constant. Hence in the absence of any new inoculation, antibodies fade away and the proportion of positive people decay in an exponential fashion with the increasing interval between the last bite and the survey time.

Assuming (A 7) and solving (A 6) for $h(a)$ gives:

$$h(a) = \frac{[(P^+(a)/\tau)+(dP^+(a)/da)]}{[1-P^+(a)]}, \tag{A 8}$$

which is equation (4) of the main text.

The next rate to be estimated is γ, the rate at which people leave the seropositive status.

Let us consider equation (A 6) which describes the proportion of people that are still seropositive at age a. The proportion of individuals that are still seropositive at age $(a-dx)$ is:

$$\int_0^a h(a')\,da'e^{[-\nu(a)+\nu(a')]}f(a-a'-dx). \tag{A 9}$$

https://doi.org/10.1017/S0950268800057241 Published online by Cambridge University Press

Note that the term in the exponential guarantees no further inoculation in the infinitesimal interval dx. So the proportion of individuals leaving the seropositive compartment between $(a-dx)$ and a is:

$$-\int_0^a h(a')\,da'e^{[-\nu(a)+\nu(a')]}\left[\frac{df}{da}\right]dx. \tag{A 10}$$

Hence the rate of individuals leaving the seropositive condition between $(a-dx)$ and a is:

$$\gamma dx = \frac{-\int_0^a h(a')\,da'e^{[-\nu(a)+\nu(a)]}\,[df/da]\,dx}{\int_0^a h(a')\,da'e^{[-\nu(a)+\nu(a')]}f(a-a')}. \tag{A 11}$$

It is illuminating to apply equation (A 11) to Aron and May's model of immunity [30, 38]. In their model an individual is assumed to remain seropositive for a period τ after the last inoculation. So $f(a-a')$ is the Heaviside function [39], which has the form:

$$f(a-a') = \theta_\tau(a-a') = \begin{cases} 1 \text{ if } (a-a') < \tau \\ 0 \text{ if } (a-a') \geqslant \tau \end{cases}.$$

They also assume that the inoculation rate is constant (h). Therefore, noting that the derivative of the Heaviside function is the Dirac's delta function [39], we have immediately:

$$\gamma = \frac{he^{-(h\tau)}}{1-e^{-(h\tau)}}$$

which is exactly the same equation (1) of Aron's paper [30] or equation (5·33) of [38].

Solving (A 11) for $f(a-a')$ as in equation (3) gives:

$$\gamma = \frac{1}{\tau}. \tag{A 12}$$

This is the result given by equation (9) of the main text.

APPENDIX II

Equation (1) of the main text hides a number of assumptions related to the mechanisms concerning the build up of the immune response against malarial parasites. Actually this process is not entirely known.

The purpose of this Appendix is to show how to include, as much as possible, some biological realities into the model.

Instead of calculating the probability of being positive or negative to the serological test, we begin by calculating the average age-related antibody concentration in the population as a function of $h(a)$.

Assume that an individual receives n effective inoculations randomly distributed along the interval between ages 0 and a. The first inoculation occurs at age a_1, the second at age a_2, and so on. Let us also assume that each inoculation contributes

https://doi.org/10.1017/S0950268800057241 Published online by Cambridge University Press

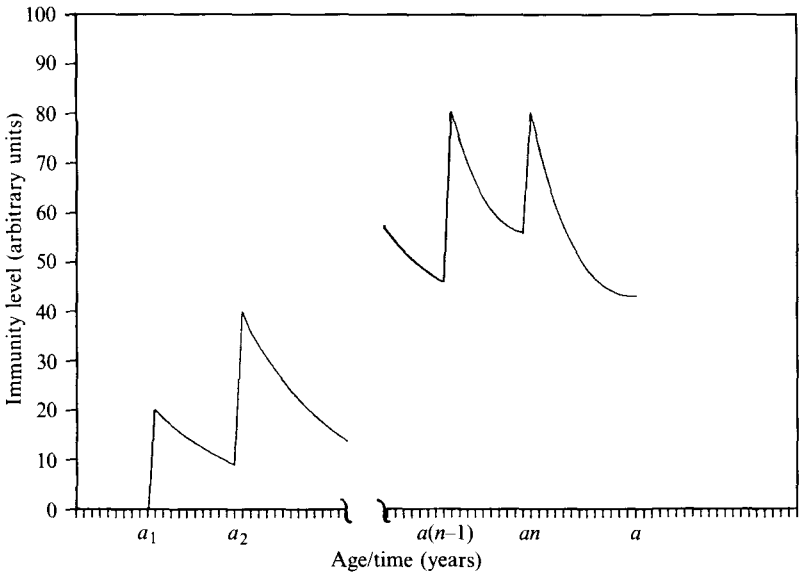

Fig. A1. Hypothetical immunity level, expressed as a function of the age distribution of inoculations and the time elapsed between these inoculations.

to the build up of the antibody level measured at age a. In addition, the level of antibodies reached after an inoculation is assumed to decay in the absence of further inoculations. This process is schematically illustrated by Figure A 1.

Let us assume that the antibody level measured at age a after the first inoculation is given by:

$$I^{(1)}(a, a_1) = I_c f(a-a_1) \qquad a_1 \leqslant a \leqslant a_2, \tag{A 13}$$

where I_c means the level of antibodies reached after a single inoculation in a non-immune individual and $f(a-a_1)$ is the function describing the fading of antibiotics in the absence of further inoculations.

The level of antibodies after the second inoculation is assumed to be given by:

$$I^{(2)}(a, a_1, a_2) = \left\{I_c + (I_{max} - I_c)\left[1 - \frac{2}{1+e^{\gamma_1 I^{(1)}(a_2, a_1)}}\right]\right\} f(a-a_2) \qquad a_2 \leqslant a \leqslant a_3. \tag{A 14}$$

Equation (A 14) means that the level of antibodies reached after the second inoculation is equal to I_c plus an extra amount (the boosting effect) given by the second term of equation (A 14). In this equation I_{max} represents the maximum level of antibodies attainable. The form chosen for the boosting effect is phenomenological and incorporates the current beliefs about the building up of immune response against malarial parasite. For instance, when inoculations are so sparse that the second one occurs when there is almost no antibodies left from the first inoculation the boosting effect is negligible and the antibodies concentration rises to I_c. On the other hand, when inoculations are very frequent the level of antibodies tends to a saturation level, I_{max}, that is, the boosting effect is also negligible.

https://doi.org/10.1017/S0950268800057241 Published online by Cambridge University Press

The parameter γ_i is intended to represent the factors involved in immune response that determine the magnitude of the boosting effect.

So the level of antibodies after n inoculations is given by:

$$I^{(n)}(a, a_n, \ldots, a_1) = \left\{ I_c + (I_{max} - I_c)\left[1 - \frac{2}{1 + e^{\gamma_{n-1} I^{n-1}(a_n, \ldots, a_1)}} \right] \right\} \qquad f(a - a_n), \tag{A 15}$$

Quite generally one can write:

$$I^{(n)}(a, a_n, \ldots, a_1) = \{ I_c + [I_{max} - I_c] F(I^{(n-1)}(a_n, \ldots, a_1)) \} \qquad f(a - a_n) \tag{A 16}$$

where F is some function that depends on $I^{(n-1)}(a_n, \ldots, a_1)$ as described after equation (A 2). Its exact form is not known.

This is a very general expression for the building up of an humoral response comprising the concepts of 'boosting effect' and fading of antibodies.

Let us now calculate the probability of receiving n inoculations as shown in Figure A 1. Let $h(a)\,da$ be the probability of receiving an inoculation between a and $a + da$. Then, the probability of the 'history' n inoculations given by Figure A 1 is:

$$p(a_1, a_2, \ldots, a_n a) = e^{-[\nu(a_1) - 0]}, h(a_1 da_1\, e^{-[\nu(a_2) - \nu(a_1)]} h(a_2)\, da \ldots h(a_n)\, da_n\, e^{-[\nu(a) - \nu(a_n)]}$$

$$= \prod_{i=1}^{n} h(a_i)\, da_i e^{-\nu(a)}. \tag{A 17}$$

So the average antibody level at age a due to n inoculations is:

$$I_n(a) = \int_0^a \ldots \int_0^a \prod_{i=1}^{n} h(a_i)\, da_i\, e^{-\nu(a)} I^n(a, a_n, \ldots, a_1)\, \theta(a_2 - a_1)\theta(a_3 - a_2) \ldots \theta(a_n - a_{n-1}), \tag{A 18}$$

where $I^{(n)}(a, a_n, \ldots, a_1)$ is obtained by solving equations (A 13–A 15).

The average level of antibodies at age a due to all possible histories of inoculations is:

$$I(a) = \sum_{n=1}^{\infty} I_n(a). \tag{A 19}$$

The sum of all histories described by equation (A 18) is similar to Feynman integral [40]. This integral can be computed analytically in a number of particular cases, as shown below:

Case 1

Assume that $I^{(n)}(a, a_n, \ldots, a_1) = nI$, that is, each inoculation adds I, a constant, to the antibody level, and there is no fading of antibodies. Then, by integrating equation (A 18) and substituting in equation (A 19), we have:

$$I(a) = \sum_{n=1}^{\infty} nI \frac{\left(\int_0^a h(a)\, da \right)^n e^{-\int_0^a j\, h(a)\, da}}{n!}. \tag{A 20}$$

https://doi.org/10.1017/S0950268800057241 Published online by Cambridge University Press

Case 2

Assume that the immunity level depends only on the last inoculation, that is $I(a, a_n, \ldots, a_1) = I\, exp\,[-c(a-a_n)]$. This is a purely Markovian assumption for the immunity process with fading of antibodies. Then, by integrating equation (A 18) and substituting in (A 19), we have:

$$I(a) = \int_0^a da' h(a') I\, e^{-[\nu(a)-\nu(a')]}\, e^{-c(a-a')}. \tag{A 21}$$

Note that this equation is very similar to equation (1) of the main text.

Case 3

Assume that the level of immunity is given by a constant I that is simply added to the pre-existing antibody level, also considering the fading process. This means, for example, that after two inoculations at ages a_1 and a_2, the level of antibodies at age a is:

$$[I\, e^{-c(a_2-a_1)} + I]\, e^{-c(a-a_2)}. \tag{A 22}$$

Then by integrating equation (A 18) and substituting in (A 19) we have:

$$I(a) = I\, e^{-ca} \int_0^a h(a')\, e^{ca'}\, da'. \tag{A 23}$$

As can be seen from the above analysis, the calculation of the average immunity level is far from simple. Furthermore, the above analysis does not take into account individual variations in the host and parasite populations. In order to circumvent these difficulties, we now ask a different question:

'What is the probability of finding an individual positive to a serological test after a history of n inoculations as shown in Figure A 1?'

This probability, $P_n^+(a)$, is given by:

$$P_n^+(a) = p(a, a_n, \ldots, a_1) P^+(a|a_1, \ldots, a_n), \tag{A 24}$$

where the second term is the conditional probability of being positive given that one has suffered n inoculations at ages $a_1, \ldots, a_n$.

The probabilities of being positive is given by the sun of histories:

$$P^+(a) = \sum_{n=1}^{\infty} \int_0^a \ldots \int_0^a \prod_{i=1}^{n} h(a_i)\, da_i\, e^{-\psi(a)} P^+(a|a_1, \ldots, a_n), \\ \times \theta(a_2-a_1)\, \theta(a_3-a_2) \ldots \theta(a_n-a_{n-1}) \tag{A 25}$$

The conditional probability appearing in (A 24) can be calculated by counting all histories that result in an antibody level greater than a certain threshold. However, this is very complicated although it can be done numerically, for instance by using Montecarlo techniques, or analytically in cases where the immune response is taken unrealistically simple. However, as mentioned above, one has to take into account individual variations in the immune response against the parasite. How important this may be is not known. So, it is reasonable to assume this condition probability independent of $a_1 \ldots a_{n-1}$, and equal to:

$$P^+(a|a_1, \ldots, a_n) = e^{\frac{1}{r}(a-a_n)}. \tag{A 26}$$

Substituting (A 26) in (A 25) we get equation (1) of the main text.

https://doi.org/10.1017/S0950268800057241 Published online by Cambridge University Press

The assumption implied in equation (A 26), that is, that the conditional probability of being positive given a history of inoculations as dependent only on the average period of time elapsed since the last inoculation may seem crude. It essentially states that the level of antibodies reached in a randomly selected individual in the population after an inoculation is a stochastic quantity, but the fading of antibodies is deterministic. Moreover, as mentioned in the main text the lack of sufficient data on malaria immunology justify this approach. In addition, it is possible to estimate the average duration of immune response to serological test in a given population [19–21].

REFERENCES

1. Ross R. The prevention of malaria, 2nd ed. London: John Murray, 1911.
2. Ross R. Some *a priori* pathometric equation. B M J 1915; **1**: 546–7.
3. Ross R. An application of the theory of probabilities to the study of *a priori* pathometry. I. Proc R Soc, A 1916; **92**: 204–30.
4. Ross R, Hudson HP. An application of the theory of probabilities to the study of *a priori* pathometry, III. Proc R Soc, A 1917; **93**: 225–40.
5. Macdonald G. The analysis of infection rates in diseases in which superinfection occurs. Trop Dis Bull 1950; **47**: 907–15.
6. Macdonald G. The analysis of equilibrium in malaria. Trop Dis Bull 1952; **49**: 813–28.
7. Macdonald G. The analysis of malaria epidemics. Trop Dis Bull 1953; **50**: 871–89.
8. Dietz K, Molineaux L, Thomas A. A malaria model tested in the African savannah. Bull WHO 1974; **50**: 347–57.
9. Garret-Jones C. The human blood index of malaria vectors in relation to epidemiological assessment. Bull WHO 1964; **30**: 241–61.
10. Molineaux L. The pros and cons of modelling malaria transmission. Trans R Soc Trop Med Hyg 1985; **79**: 743–7.
11. Bruce-Chwatt LJ. Essential malariology, 2nd ed. London: William Heinemann Books. 1985.
12. Dye C, Hasibeder G. Mosquito-borne disease dynamics: control of flies which bite some people more frequently than others. Trans Roy Soc Trop Med Hyg 1986; **80**: 69–77.
13. Draper CC, Voller A, Carpenter RG. The epidemiologic interpretation of serological data in malaria. Am J Trop Med Hyg 1972; **21**: 696–703.
14. Meuwissen JHETh. Introductory remarks on seroepidemiology and immunodiagnosis of malaria. Israel J Med Sci 1978; **14**: 687–9.
15. VanDruten JAM. A mathematical-statistical model for the analysis of cross-sectional data with special reference to the epidemiology of malaria. PhD Thesis. Nijmegen, Netherlands: The Katholiek Universiteit, 1981.
16. Papoulis A. Probability, random variables and stochastic processes. Tokyo: McGraw–Hill International Book Company, 1981.
17. Merlin M, Lettesram JY, Josse R, et al. Évaluation des indices cliniques, parasitologiques et immunologiques du paludisme dans la région de la baie Bonny en Afrique Centrale. Bull Soc Pathol Exot Fil 1986; **79**: 707–20.
18. Bailey NTJ. The mathematical theory of infectious diseases, 2nd ed. London, High Wycombe: Charles Griffin & Company Ltd, 1975.
19. Dragna A, Marinov R, Mihai M. Contribution a l'étude de l'immunité rèsiduelle aux paludisme en Roumanie. Bull WHO 1969; **40**: 753–61.
20. Luby JP, Collins WE, Kaiser RL. Persistence of malaria antibody. Findings in patients infected during the outbreak of malaria in Lake Vera, California 1952–1953. Am J Trop Med Hyg 1967; **16**: 255–7.
21. Collins WE, Skinner JC, Jeffery GM. Studies on the persistence of malarial antibody response. Am J Epidemiol 1968; **87**: 592–8.
22. Carnevale P, Frézil JL, Bosseno MF, LePont F, Lancien J. Etude de l'agressivité d'*Anopheles gambiae* en fonction de l'âge et du sexe des sujets humains. Bull WHO 1978; **56**: 147–54.

23. Muirhead-Thomson RC. The distribution of anopheline mosquito bites among different age groups. B M J 1951; **1**: 1114–7.
24. Clyde DF, Shute GT. Selective feeding habits of anophelines amongst Africans of different ages. Am J Trop Med Hyg 1958; **7**: 543–5.
25. Molineaux L, Gramicia G. The Garki Project. Geneva: World Health Organisation, 1980.
26. Bruce-Chwatt LJ, Draper CC, Konfortion P. Sero-epidemiological evidence of eradication of malaria from Mauritius. Lancet 1973; ii: 547–51.
27. Dietz K. Mathematical models for transmission and control of malaria. In: Wernsdorfer WH, McGregor I, eds. Malaria: principles and practice of malariology. London: Churchill Livingstone, 1988: 1091–133.
28. Molineaux L. The epidemiology of human malaria as an explanation of its distribution, including some implications for its control. In: Wernsdorfer WH, McGregor I, eds. Malaria: principles and practice of malariology. London: Churchill–Livingstone, 1988: 913–98.
29. Committee of Experts on Malaria. Report of the Scientific Work Group on Immunology of Malaria and on Applied Field Research in Malaria. Technical Report TDR/IMMAL-FIELDMAL/SWG(2)/88.3. Geneva: World Health Organization, 1988.
30. Aron JL. Mathematical modelling of immunity to malaria. Math Biosc 1988; **90**: 385–96.
31. Nedelman J. Introduction review: some new thoughts about some old malaria models. Math Biosc 1985; **73**: 159–82.
32. Burattini MN. Contribuição para o estudo da dinâmica da transmissão da malária a partir de inquérito sorepidemiológico transveral (port). [PhD. thesis]. São Paulo. Brazil: Escola Paulista de Medicina, 1989.
33. Burattini MN, Massad E, Coutinho FAB, Baruzzi RG. Malaria prevalence amongst Brazilian Indians assessed by a new mathematical model. Epidemiol Infect 1993; **111**: 525–537.
34. Burattini MN. Epidemiologia da Malária no Parque Indígena do Xingu (port). [MA thesis]. São Paulo. Brazil: Escola Paulista de Medicina, 1987.
35. Burattini MN, Baruzzi RG, Wucker S. Impact of DDT spraying in malaria transmission at an Indian reservation in Central Brazil. Proceedings of the International Congress for Infectious Diseases. Rio de Janeiro: International Society of Infectious Diseases, 1988: 153.
36. Burattini MN, Baruzzi RG, Sanchez-Ruiz MCA, Castelo A, Wey SB. Comparação entre índice esplênico e sorologia como índices malariométricos na área do Baixo Xingu. Rev Soc Bras Med Trop 1986; **19** (suppl): 88.
37. Morse PM, Feshbach H. Methods of theoretical physics, vol. I. New York and Tokyo: McGraw–Hill Book Company Inc, 1953.
38. Aron JL, May RM. The population dynamics of malaria. In: Anderson RM, ed. The population dynamics of infectious diseases: theory and applications. London and New York: Chapman and Hall, 1982; 139–79.
39. Arfken G. Mathematical methods for physicists. New York, San Francisco, London: Academic Press, 1970.
40. Feyman RP, Hibbs AR. Quantum mechanics and path integrals. New York: McGraw–Hill, 1965.