PHILOSOPHICAL TRANSACTIONS B

royalsocietypublishing.org/journal/rstb

Research

Cite this article: Alonso D, Dobson A, Pascual M. 2019 Critical transitions in malaria transmission models are consistently generated by superinfection. *Phil. Trans. R. Soc. B* **374**: 20180275.
http://dx.doi.org/10.1098/rstb.2018.0275

Accepted: 11 March 2019

One contribution of 15 to a theme issue 'Modelling infectious disease outbreaks in humans, animals and plants: approaches and important themes'.

Subject Areas:
health and disease and epidemiology, theoretical biology

Keywords:
malaria dynamics, critical transitions, alternative steady states, malaria superinfection, backward bifurcation, hysteresis in malaria model

Author for correspondence:
Andy Dobson
e-mail: dobson@princeton.edu

Electronic supplementary material is available online at https://dx.doi.org/10.6084/m9.figshare.c.4450349.

THE ROYAL SOCIETY PUBLISHING

Critical transitions in malaria transmission models are consistently generated by superinfection

David Alonso[1], Andy Dobson[2,3] and Mercedes Pascual[3,4]

[1]Theoretical and Computational Ecology, Center for Advanced Studies (CEAB-CSIC), Blanes, Spain
[2]Ecology and Evolutionary Biology, Eno Hall, Princeton University, NJ 08540, USA
[3]Santa Fe Institute, Hyde Park Road, Santa Fe, NM, USA
[4]Ecology and Evolutionary Biology, University of Chicago, Chicago, IL, USA

DA, 0000-0002-8888-1644; AD, 0000-0002-9678-1694; MP, 0000-0003-3575-7233

The history of modelling vector-borne infections essentially begins with the papers by Ross on malaria. His models assume that the dynamics of malaria can most simply be characterized by two equations that describe the prevalence of malaria in the human and mosquito hosts. This structure has formed the central core of models for malaria and most other vector-borne diseases for the past century, with additions acknowledging important aetiological details. We partially add to this tradition by describing a malaria model that provides for vital dynamics in the vector and the possibility of superinfection in the human host: reinfection of asymptomatic hosts before they have cleared a prior infection. These key features of malaria aetiology create the potential for break points in the prevalence of infected hosts, sudden transitions that seem to characterize malaria's response to control in different locations. We show that this potential for critical transitions is a general and underappreciated feature of any model for vector-borne diseases with incomplete immunity, including the canonical Ross–McDonald model. Ignoring these details of the host's immune response to infection can potentially lead to serious misunderstanding in the interpretation of malaria distribution patterns and the design of control schemes for other vector-borne diseases.

This article is part of the theme issue 'Modelling infectious disease outbreaks in humans, animals and plants: approaches and important themes'. This issue is linked with the subsequent theme issue 'Modelling infectious disease outbreaks in humans, animals and plants: epidemic forecasting and control'.

1. Introduction

Critical transitions occur when natural systems drastically shift from one state to another. They are currently receiving considerable attention in ecology, geophysics, hydrology and economics [1]. In epidemiology, critical transitions are of relevance to the emergence of new pathogens and the re-emergence of old ones. They may also be central to the within-host dynamics particularly when pathogens evolve or mutate in ways that allow them to escape control by the immune system. Critical transitions often underlie and potentially enhance (or undermine) attempts to control and eliminate infectious pathogens. Following an intervention, the trajectory of the host–pathogen systems may cross a critical transition where pathogen prevalence drops to apparent eradication, the robustness of which is strongly determined by the structure of the transition. Tipping points associated with the coexistence of alternative equilibria are of particular interest, as small changes in a driving parameter can lead to large shifts from low to high levels of prevalence (or *vice versa*). Continuous external pressure on critical transmission parameters, or seasonal variation in vector abundance, can also lead to hysteresis, whereby a delayed response of the system would effectively keep it trapped longer in either the endemic or disease-free state.

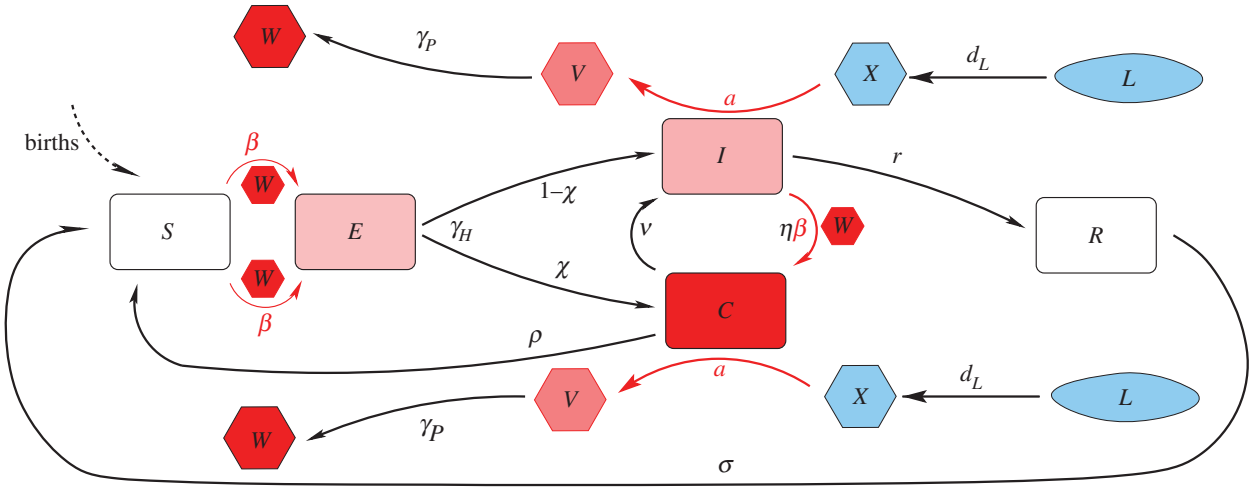

Figure 1. The human–mosquito SECIR–LXVW coupled model. The human stages of infection 'SECIR' are depicted in the central part of the figure progressing from *S* to *E* to *I* from left to right. The different stages of the mosquito population, from larva (*L*) to infectious adult mosquito (*W*), are drawn in the upper and bottom of the figure progressing from uninfected *L*, larvae and adults (blue) to incubating (pink) and infected adults (red), from right to left (LXVW). *Per capita* death rates (respectively, δ_H and δ_M, and δ_L in the equations) affect all mosquito and human classes, and for simplicity are not included in the diagram. (Online version in colour.)

Evidence for the existence of alternative steady states in infectious disease dynamics remains limited [2–4], although they have been proposed as a possible explanation for the observation that malaria often fails to re-invade local regions that have achieved elimination. One potentially important pre-condition for the existence of alternative steady states in malaria is superinfection: the infection of a single host by concurrent multiple strains of the pathogen. Malaria infections are not fully immunizing, and multiplicity of infection (MOI) is common in endemic regions as a consequence of additional infectious bites by the vector before the host has cleared a prior infection. In endemic regions, a large fraction of the human population carries the malaria parasite asymptomatically in non-apparent infections that can contribute to transmission [5,6]. Under these conditions, significant levels of superinfection can create a positive feedback within the intensity of transmission that has the potential to generate multiple alternative equilibria and associated tipping points.

A large number of model formulations have been proposed to capture the complexity of malaria epidemiology [7,8]. The consequences of repeated infectious bites have been represented in models of re-infection by a formulation that only allows a host to re-acquire infection after it has cleared a previous exposure. Two such recent models suggest that this might cause malaria dynamics to exhibit alternative steady states [9,10]. Here, we develop a complementary, but more general, analysis that provides a formulation for superinfection that explicitly allows infections to occur concurrently without interfering with each other. We initially present a semi-analytical approach to identify alternative equilibria in models for vector-transmitted diseases. We then apply these methods to a vector-borne disease model that was used to understand the origins of environmentally driven fluctuations of malaria in epidemic regions [11]. We then broadly consider superinfection in a series of hierarchical formulations of increasing complexity (commencing with the original Ross–MacDonald equations [12–15]). We demonstrate that irrespective of the details, superinfection consistently creates tipping points that can generate hysteresis in responses to control efforts (as well as seasonal variation in vector abundance). We argue that because models for re-infection and superinfection effectively bracket a continuum of different assumptions about within-host dynamics with multiple concomitant infections, then models that include these vital details of malaria biology should consistently exhibit tipping points in their response to control regardless of model formulation. Models that fail to include these effects may be misleading, or of limited utility when used to examine transitions towards low rates of transmission in response to control of vectors or of the pathogen. Although the model is framed in the context of malaria in humans, the results should be more generally relevant to the dynamics of the disease in other, non-human, hosts and to other vector-borne diseases, when immunity is partial and parasite antigenic diversity allows repeated infection.

2. The model

The model is formulated as a system of ordinary differential equations that describe the transmission of the disease between the mosquito vector and the human host. It differs from the standard Ross–McDonald model by explicitly considering the vital dynamics of the two populations. Figure 1 presents the flow diagram of the initial model whose equations are described in detail in the electronic supplementary material, §1. (Simplifications of this model are considered later in the paper.) The mosquito population is subdivided into larval and adult stages, and adult individuals can be susceptible, infected or infectious. The human population has a constant size and deaths are assumed to exactly balance the birth rate; infected hosts are subdivided into two classes to differentiate individuals whose clinical infection leads to treatment from those whose natural recovery leads to the acquisition of immunity. These two classes map, respectively, to symptomatic individuals who are detected by the health system and to symptomatic or asymptomatic individuals who are not. The latter are crucially important as clinical infections typically represent only a small fraction of the total incidence in endemic regions [16–18]. Dietz *et al.* [19] and Aron & May [20] scale the

Table 1. Derived parameters are those expressed as a function of basic parameters and variables of the SECIR–LXVW model (figure 1). Their definition is summarized here, together with their appearance in the electronic supplementary material or main text. The full list of basic model parameters with their definitions and symbols is given in table S1 (electronic supplementary material, §1.1).

mosquito parameters		
rate of infectious bites per human	$\Lambda = a\frac{W}{N}$	EIR, when expressed as infectious bites per person per year
force of infection	$\lambda = c\,a\,\frac{I+C}{N}$	equation (2.2)
human parameters		
fraction of infectious humans	$y = \frac{I+C}{N}$	
loss of immunity rate	$\sigma = \frac{\Lambda}{\exp(\Lambda/\sigma_0)-1}$	equation (S7), electronic supplementary material (section 1.1)
recovery rate	$r = \frac{\Lambda}{\exp(\Lambda/r_0)-1}$	equation (S8), electronic supplementary material (section 1.1)
force of infection	$\beta = b\Lambda + \beta_e$	equation (2.2) and equation (S10), electronic supplementary material (section 1.3)
human–mosquito parameters		
basic reproduction ratio	R_0	equation (S43), electronic supplementary material (section 2.2)
dominant period	T_0	equation (S67), electronic supplementary material (section 4)

effective recovery rates, r by transmission intensity (the rate of infectious bites per human). Under the assumptions that (i) infectious bites arrive at a constant rate, (ii) the individual infections within a host proceed independently, and (iii) a constant period, $1/r_0$, Dietz *et al.* [19] derived the following expression for the effective *per capita* recovery rate:

$$r(\Lambda) = \frac{\Lambda}{\exp(\Lambda/r_0) - 1}, \tag{2.1}$$

where Λ denotes the rate of infectious bites per human, and r_0 is the basal recovery rate when disease transmission is very low (more precisely, in the limit of the infectious mosquito population tending to zero). Thus, the higher the rate of infectious bites per human host, Λ, the slower the disease clearance rate. When this rate Λ is measured per year, it is usually called the entomological inoculation rate (EIR). Finally, we assume that once individuals recover naturally from malaria (including the superinfection state with multiple concurrent infections), they enter a refractory period during which they are transiently immune and cannot re-acquire infection until they return to the susceptible class. Those who receive treatment as a result of detected symptomatic infection either clear infection and move back to the susceptible class, or fail to clear the parasite and flow into the I class. All derived parameters such as the recovery rate and EIR whose expression is a function of basic parameters (and variables) of the model are summarized in table 1.

The couplings between human and mosquito dynamics are given by the force of infection β, the *per capita* rate at which a susceptible human contracts the disease from infectious bites, and by the *per capita* rate λ at which adult mosquitoes acquire the pathogen from infectious humans. Under reasonable assumptions [21], these two rates can be written as

$$\beta = b\frac{W}{N}a + \beta_e \quad \lambda = c\,a\,y, \tag{2.2}$$

where b and c are the respective probabilities that humans develop infection once bitten and that mosquitoes acquire the parasite from biting an infectious human host. Moreover, $(W/N)a$ is the number of infectious bites per human host (see definition of Λ above), β_e is an external force of infection and y is the fraction of infectious humans (see electronic supplementary material, §§1.2 and 1.3, for details).

Our model can be seen as an extension of the classical formulation introduced by Ross [22] and MacDonald [23] (RM) (see also [7,21,24] for reviews). Although the original RM model has given rise to a multiplicity of malaria models [8] spanning different degrees of complexity that range from delayed ordinary differential equations (delayed ODEs) to SIR (Susceptible–Infected–Recovered)-like structures, we show that these models are all related to the RM formulation through a set of sequentially simplifying assumptions. In particular, all converge to the RM model, which is based on two strong simplifications: (1) constant mosquito (M) and human populations (N), and (2) only two classes (susceptible and infectious) representing vectors and humans (see electronic supplementary material §7 and appendix A1, for details). We start with the analysis of alternative steady states in the full model of figure 1, and end with the consideration of their more general existence in a suite of models of decreasing complexity. Model parameter space has been explored between maximum and minimum values (around two different parameter combinations—A and B; see electronic supplementary material, table S1 in §1) that provided a good fit to data in a previous study from an epidemic region [11].

3. Results

To identify the stationary points of the coupled system, we present a semi-analytical method that consists of first finding the equilibria of the two submodels, namely the expression for infectious mosquitoes as a function of infectious humans and *vice versa*. The fixed points are then obtained by calculating the intersection of these two curves (figure 2). The generality and feasibility of this method rest on the linearity of the human and mosquito submodels when considered separately. This means that, for a given number of infectious mosquitoes ($W^\star$), the human submodel becomes a linear ODE system. And, likewise, for a given fraction of infectious humans (y), the mosquito submodel becomes a linear ODE system as well. We summarize the main steps of the approach below (full details are given in electronic supplementary material, §3). Once the stationary points have been fully characterized, we can map several quantities of interest onto the parameter space, including the model reproductive number R_0, the

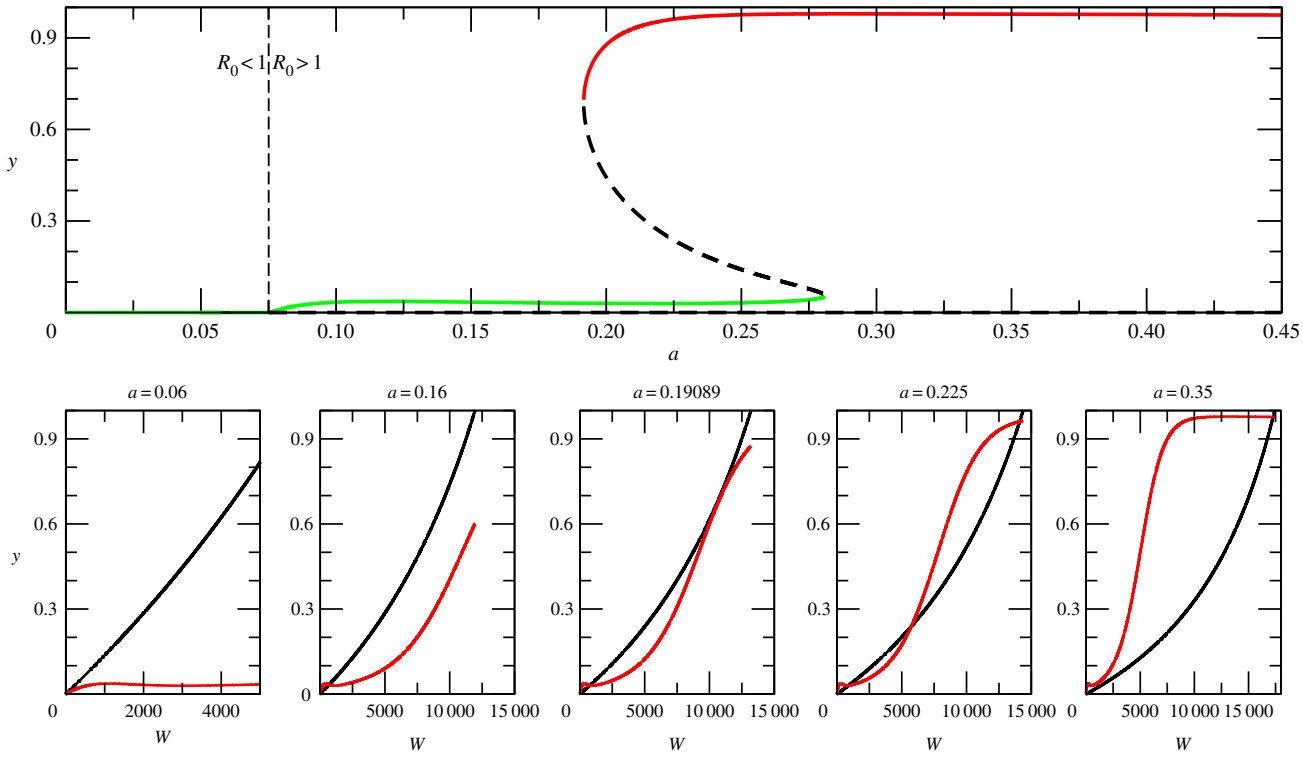

Figure 2. Relationship between saddle-node bifurcation and mosquito biting rates. When biting rates are low, the curves intersect at zero and malaria fails to establish, resulting in a disease-free equilibrium ($R_0 < 1$). As the biting rate a increases, the two curves intersect initially at one single point ($a = 0.16$) corresponding to a stable endemic equilibrium. When the biting rate crosses a critical value ($a_C = 0.19089$), a tangential bifurcation arises. Further increases of the biting rate end up producing two alternative stable states corresponding to the coexistence of two endemic equilibria characterized by high and low malaria incidence ($a_C < a < 0.28$). Eventually, only the upper equilibrium is left ($a > 0.28$). The full bifurcation diagram shown in the upper plot is calculated with parameter combination A (but with $\beta_e = 0$; see electronic supplementary material, table S1). It represents a cross section of figure 4*b* following the horizontal broken line shown there (but here starting at $a = 0$ rather than at $a = 0.15$) for incremental increases in biting rate. (Online version in colour.)

expected number of secondary cases produced by an initial single infection in a completely susceptible population, and the dominant period, T_0, associated with the damped oscillations in the transients towards stationarity.

(i) Mosquito submodel: stationary points

To obtain the expression for infectious mosquitoes at equilibrium, one first needs to calculate the fixed point for total mosquito abundance. It is easy to show that the mosquito population model has only one globally stable point (electronic supplementary material, §3):

$$M^* = K_0 \frac{d_L}{\delta_M} \frac{f - \delta_M(1 + \delta_L/d_L)}{f} \quad \text{and}$$
$$L^* = \frac{\delta_M}{d_L} M^\star = K_0 \frac{f - \delta_M(1 + \delta_L/d_L)}{f}. \tag{3.1}$$

A single condition controls when this point is a feasible stable point for the dynamics:

$$\frac{f}{\delta_M} \geq 1 + \frac{\delta_L}{d_L}. \tag{3.2}$$

Biologically, this condition means that the mosquito population will be locally maintained in a given area when the number of larvae recruited per female during their adult average life time compensates for the larval losses during their development stage. This threshold condition arises naturally in the full expression for R_0 (see electronic supplementary material, equations S28 and S43), whose value is positive only when M^* is also positive.

(ii) Human–mosquito model: stationary points

The equilibrium number of infectious mosquitoes $W^\star$ is derived by solving the mosquito ODE subsystem at equilibrium when the fraction of infectious humans y (table 1) is considered as a fixed parameter. We obtain the curve

$$W^\star = \frac{acy^\star}{\delta_M + acy^\star} \left[\frac{n_P \gamma_P}{n_P \gamma_P + \delta_M}\right]^{n_P} M^\star. \tag{3.3}$$

A second curve is then calculated by solving the human submodel at equilibrium for a given level of infectious mosquito population; this makes β a constant parameter (table 1). This curve for the stationary fraction of infectious humans is given by

$$y \equiv \frac{I + C}{N}$$
$$= \beta \frac{\nu + (1 - \chi)(\rho + \delta_H) + \eta\beta + \chi(r + \delta_H)}{q(1 + z)} [\theta_H]^{n_H}, \tag{3.4}$$

where the total force of infection, β, depends on the number of infectious mosquitoes per human, with q and z standing for composite parameters. Full expressions for these parameters and fixed points are given in electronic supplementary material §3, equations (S61)–(S62).

(a) Coexistence of alternative stable states

Figure 2 shows that the intersection of the curves represented by equations (3.3) and (3.4) can produce more that one fixed point. When intersections are present, the corresponding stable points can be computationally calculated with a

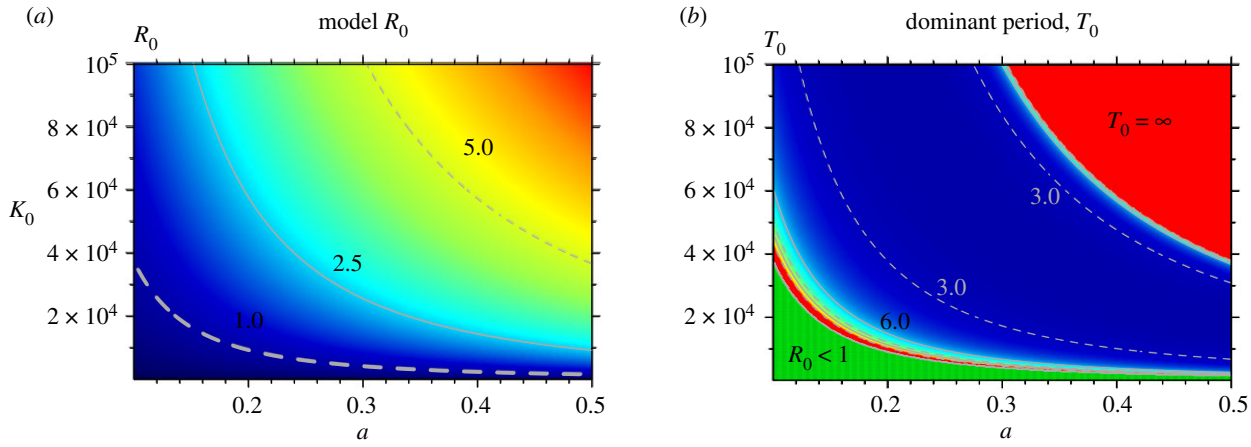

Figure 3. (*a*) The reproductive number R_0 (see electronic supplementary material equation S43, §2.2.2) as a function of mosquito carrying capacity K_0 and biting rate a. Values increase from low (blue) to high (red) values. The contour line $R_0 = 1$ is the thickest broken line delimiting the region where $R_0 < 1$ on the bottom-left corner. (*b*) The dominant period of the damped epidemic oscillations T_0 in the same parameter plane (see electronic supplementary material §4). In the region where bi-stability is found, the natural period was calculated for the damped oscillations to the lower stable equilibrium. This period was computed from the complex eigenvalues of the Jacobian evaluated at this lower equilibrium. Within the whole area where damped oscillations are found, the colours go from blue (low values of T_0), in the centre of the region, towards red (high values of T_0), at the boundaries of the region. Selected contour lines representing periods of 3 and 6 years are also shown for guidance. The red region in the upper-right corner corresponds to stable endemic equilibria that are approached monotonically with no transient oscillations ($T_0 = \infty$). The region in green is the parameter space where the system exhibits global disease-free equilibria. For consistency, we have kept the same convention of red and green areas as in figure 4. In panel (*a*), $\chi = 0$, and in (*b*), $\chi = 0.2$. In both panels, other parameter values are those of parameter combination B, with $\beta_e = 0$ (see electronic supplementary material table S1, §1). (*a*) Model R_0, (*b*) dominant period, T_0.

cobweb procedure (see electronic supplementary material, figure S2). Although this procedure cannot converge to the intermediate unstable equilibrium, this point can be calculated with a standard bisection method [25]. The pair $(W^\star, y^\star)$ defines the stationary state and the full solution can then be unfolded from it (see electronic supplementary material §3 for details). Figure 2 also shows that as the biting rate a increases, the system undergoes several bifurcations. Each of these represents the emergence of a new stationary point.

The first bifurcation is transcritical [26] and gives rise to a transition from a free-disease situation ($R_0 < 1$) to an endemic stable equilibrium ($R_0 > 0$). This kind of bifurcation is typical of infectious disease models and is found in Susceptible-Infectious-Susceptible (SIS), SIR and Susceptible-Exposed-Infectious-Recovered (SEIR) systems [27]. The second transition corresponds to a saddle node bifurcation (also called a tangential or fold bifurcation). The tangential intersection of the two curves defines a critical biting rate ($a_C = 0.19089$) at which a pair of resting points suddenly appears, consisting of a saddle node and a second stable point with higher disease incidence. The previous stable point remains for the lower incidence endemic equilibrium. As a result, two basins of attraction coexist, each consisting of initial conditions that lead to one of the two alternative stable states. These basins are separated by the intermediate unstable state. Thus, for the same parameter values and depending on initial conditions, the system can end up in one of two possible stable equilibria for high and low disease incidence, respectively.

The emergence of the first transition (from disease-free to endemic equilibrium) can be characterized by calculating the reproductive number R_0. The transcritical bifurcation corresponds to $R_0 = 1$ (see vertical broken line in the upper panel of figure 2, and thick broken line in figure 3*a*). We calculated this quantity as the dominant eigenvalue of the next-generation matrix [28] (see electronic supplementary material §2.2).

In addition, we mapped R_0 onto the sub-parameter space defined by the carrying capacity of the vector K_0 and its biting rate a (figure 3*a*). These two parameters determine the intensity of transmission. The expression underlying R_0 is the same for the model with and without superinfection since this quantity depends only on the basal rates r_0 and σ_0 (see electronic supplementary material §2, equation S43). Moreover, for the regions where damped oscillations occur, the dominant period of those oscillations can be also calculated (see electronic supplementary material §4). In the region of the parameter space we studied (see electronic supplementary material, table S1, §1), we found dominant periods between about 2.5 and 11 years (figure 3*b*).

Identifying the different regimes of the system requires us to determine the local stability of the different fixed points. The Jacobian matrix J^* is evaluated at the corresponding fixed points and its eigenvalues are obtained as the roots λ of the characteristic polynomial:

$$p_J(\lambda) = \det(J^* - \lambda\, 1) = 0, \tag{3.5}$$

where 1 is the identity matrix. Because this polynomial can have as many roots as the dimension of the dynamical system, a fixed point is asymptotically locally stable if and only if its associated eigenvalues λ_i have all strictly negative real parts. In that case, if the imaginary part of the dominant eigenvalue is not null, damped oscillations are observed in the approach to the fixed point. We determined different properties of the stationary state, such as the possibility of coexisting fixed points, parameter-induced bifurcations and the presence of endogenous (damped) oscillations. The resulting possible dynamical regimes are mapped onto different subregions of the parameter space in figure 4. Comparison of the upper and lower panels shows that the

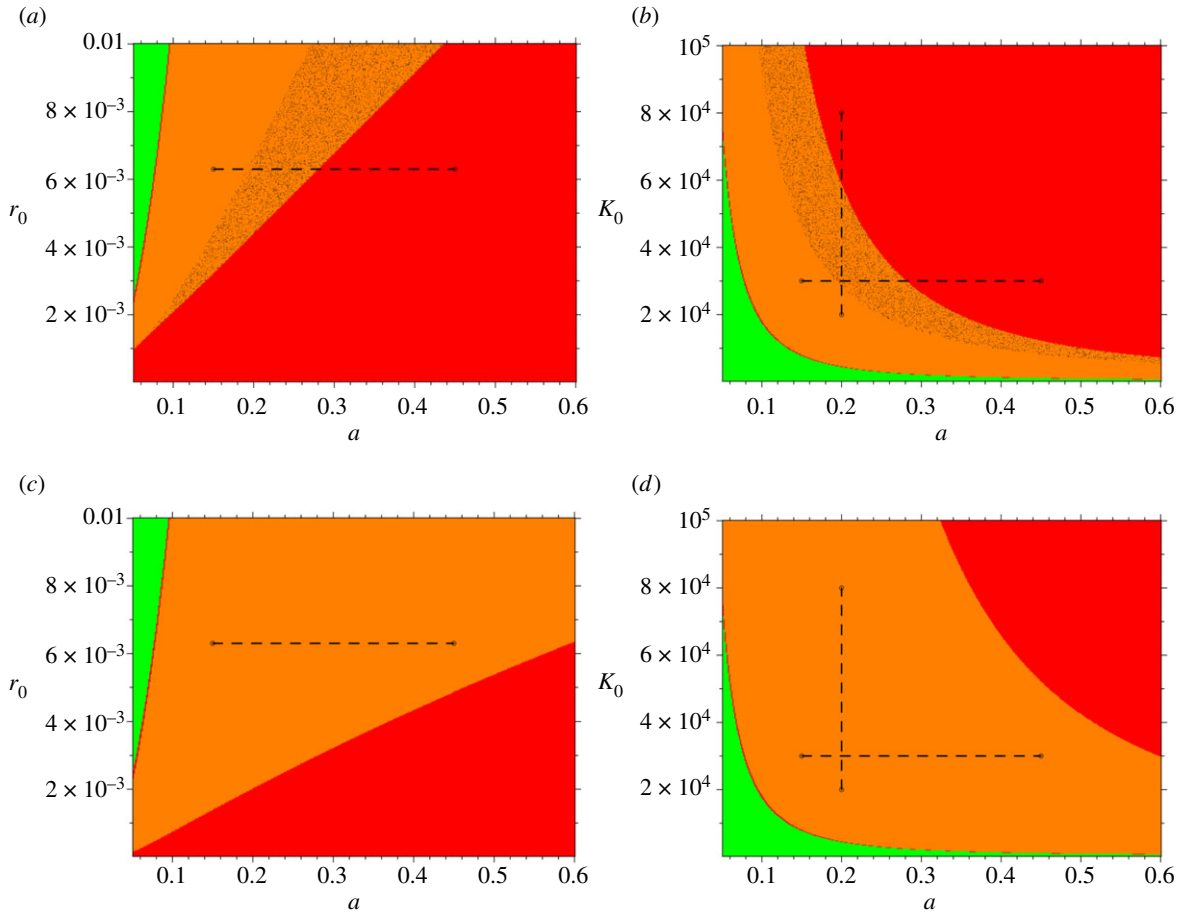

Figure 4. Dynamic regimes. Relationship between key parameters in the model and the emergence of bifurcations. In all cases, the *x*-axis depicts the biting rate a and the *y*-axis, either the basal recovery rate r_0 (*a*,*c*) or the carrying capacity K_0 (*b*,*d*). The coloured regions delineate different domains of dynamics behaviour. The orange regions correspond to non-trivial disease equilibria that are locally stable and are approached with damped oscillations. Within these regions, two stable equilibria coexist in the shaded area (panels (*a*) and (*b*)). The red regions correspond to stable endemic equilibria that are approached with no oscillations. The green regions represent the region of parameter space where the system exhibits global disease-free equilibria. The fraction of infectious humans exhibits a bifurcation in the section illustrated by the broken lines in the upper panels. These lines cross a region where a single-stable equilibrium gives rise to two coexisting stable equilibria as the value of either of the bifurcation parameters increases (see also figure 5*c* and electronic supplementary material, figure S2, panel *d*). The lower panels present phase diagrams for the same model structure and parameter combinations but without superinfection (parameter combination A, with $\beta_e = 0$, from electronic supplementary material §1, table S1). These panels are representative of the more widely used model, whose rates of recovery ($r = r_0$) and loss of immunity ($\sigma = \sigma_0$) are constant. Clearly, these simplifications fail to capture the tipping points and multiple equilibria seen in the upper figures.

coexistence of alternative steady states critically depends on including superinfection in the model.

Our procedure to calculate fixed points is completely general and can be applied to any coupled vector–human model. It is independent of specific assumptions about the force of infection, including the existence of an external source of infection, as well as other possible ways in which both the mosquito and human submodels might be defined, as long as both of these are linear ODE systems when considered separately. In the electronic supplementary material (see §§1 and 7), we add further realism and apply the approach to a system where the exponential distribution of disease incubation times in both human and mosquito submodels has been replaced by the more general (gamma) distribution, which effectively reduces the variance of the incubation period without changing its average value [29]. The gamma distribution has lower variance than the exponential distribution. Here, these distributions are implemented for humans and vectors, respectively, with a chain of n_H and n_V compartments. The higher their number, the lower the variance of the resulting distribution. The stationary solution reduces to the case illustrated in figure 4 when $n_H = n_V = 1$ (see also, electronic supplementary material, table S1).

(i) Hysteresis

Coexistence of stable equilibria can underlie sharp transitions between levels of disease prevalence if a sufficient perturbation is applied to the system to move it from one basin of attraction to the other. The responses to more gradual changes in parameters can instead lead however to hysteresis [1], which is of key interest to malaria control, since the dynamic memory of the system can delay responses and allow the persistence of endemicity beyond the bifurcation point at which we would expect elimination to occur. A related further possible effect is an asymmetry in the temporal trajectories from endemicity to elimination, and from elimination to re-emergence. These hysteresis effects are illustrated in figure 5 for slow changes in the biting rate a; these might occur during the transition between dry and rainy seasons. We consider first changes from an initial biting rate $a = 0.1$ at t_0, to a value of 0.5 at the intermediate time t_i, and then

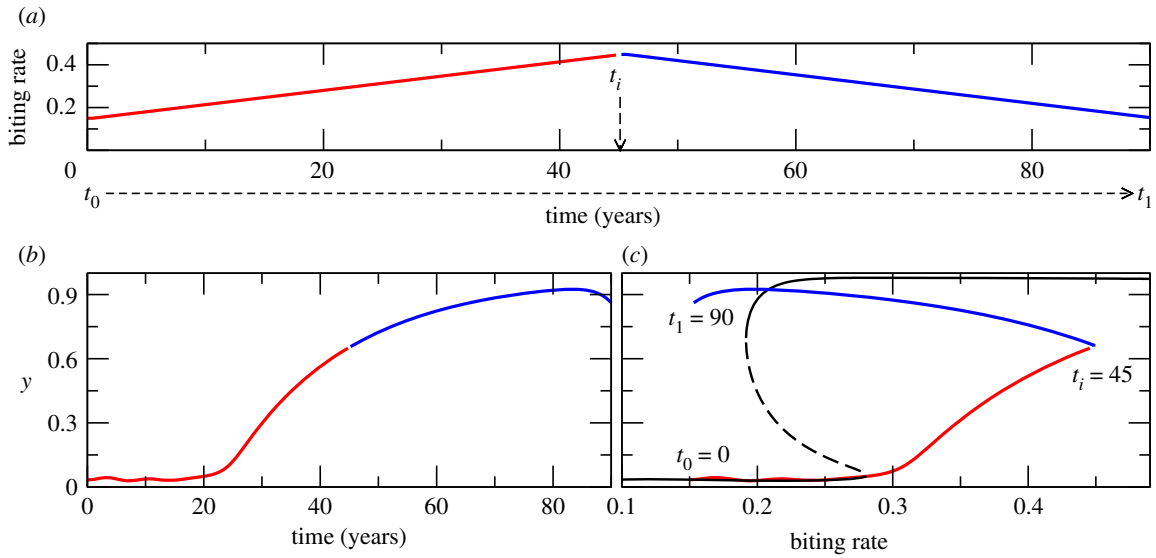

Figure 5. Hysteresis cycles. Response of the fraction of infectious humans to a slowly increasing biting rate *a*, from low to high values and back (*a*). Other parameter values are given by set A of electronic supplementary material, table S1. We plot the time evolution of the system when considering superinfection. The blue and red curves correspond to dynamics with a biting rate that increases and decreases respectively as a linear function of time. These diagrams takes us across changes in *a* given by the horizontal broken line in figure 4*a*. Because of the slowing-down effect, as biting rates go from low to high and back in a linear way (*a*), disease incidence (*y*) shows a highly nonlinear response (see panel *b*), which can also be represented as a hysteresis cycle in the bifurcation diagrams (see panel *c*). (Online version in colour.)

back to its initial value at time t_1. The response of the system is illustrated for the fraction of infectious humans ($y = (C + I)/N$), a measure of disease incidence, where N is total human population (figure 5*b*,*c*). The response of the system to a linear increase of the biting rate is a nonlinear increase in disease incidence (figure 5*c*). This can be represented as a hysteresis cycle on the corresponding bifurcation diagrams (panel *c*). In fact, extensive initial increases in *a* leave disease levels almost unchanged up to a certain threshold. In addition, there is clear asymmetry in the two opposite control trajectories: the identical linear trajectory back to the initial low levels of the biting rate does not return the system to the same initial levels of disease incidence (figure 5). Beyond the threshold, the system has been pushed towards a higher incidence state from which it is hard to return. Although decreasing *a* back to its initial low values would eventually lead the system to settle down at the initial low incidence equilibrium determined by imported infections (for $\beta_e = 0$), the transient trajectory to this final state can be very long. Provided everything else remains constant, the amplitude of hysteresis cycles depends on how fast the driving parameter changes. The same behaviours can be illustrated by measuring system response in terms of the EIR (electronic supplementary material, figure S2 in §5).

(ii) Model robustness

In order to quantify how the presence of bi-stability in malaria models depends on their degree of complexity, we evaluated the relative size of the region of parameter space where this behaviour arises for different formulations (figure 6). We specified 16 different human–mosquito coupled ODE models of increasing realism. The models are labelled according to the number of classes considered for the human and mosquito populations: four possibilities for the former (SE_nCIR, SE_nIR, SE_nI and SI), and four for the latter (LXV_nW, X_KV_nW, XV_nW and XW). Without exhausting possible variations, this scheme results in a total of 16 different combinations, which are labelled as SE_nCIR–LXV_nW, SE_nCIR–X_KV_nW (see electronic supplementary material, §7, for a detailed description of each model). When a model increases in complexity by the addition of new model parameters (encoding particular processes), the fraction of parameter space leading to a given dynamical regime can increase, remain unchanged or decrease. The fraction of parameter space where a given regime is found is calculated by randomly drawing parameter sets within minimum and maximum values (as given in electronic supplementary material, table S1) and evaluating the fraction of those draws leading to a given dynamic regime. Figure 6 compares models from the perspective of the relative size of regions with endemic equilibria or bi-stability. The exercise suggests that the coexistence of alternative stable states is robust to model simplification. This is shown here by representing the probability of bi-stability as model complexity decreases, from the SECIR–LXV_nW, a model with 21 parameters, to the SI–LXV_nW, a model with only 13 parameters (figure 6*b*). Even the simplest mosquito–human coupled model (SI–XW), with only eight parameters, corresponding to the RM formulation, exhibits this feature as long as the positive feedback introduced by the slowing down of the human recovery rate under repeated infectious bites is maintained (see electronic supplementary material, §6 and figure S4). For each of these model combinations, the removal of the slowing-down effect in the recovery rate that is produced by superinfection eliminates bi-stability (as shown in figure 4*c*,*d*). We also show the effect of one added parametric dimension, the introduction of an external force of infection, in models that combine the mosquito subcomponent LXV_nW with human submodels of decreasing model complexity (figure 6*b*). These 'open models ($\beta_e = 0$)' show a relative decrease in the fraction of the parameter space showing bi-stability, mainly owing to the linearizing effect of external infections.

In summary, superinfection introduces a slowing-down effect on recovery rates as infectious bites increase, which is

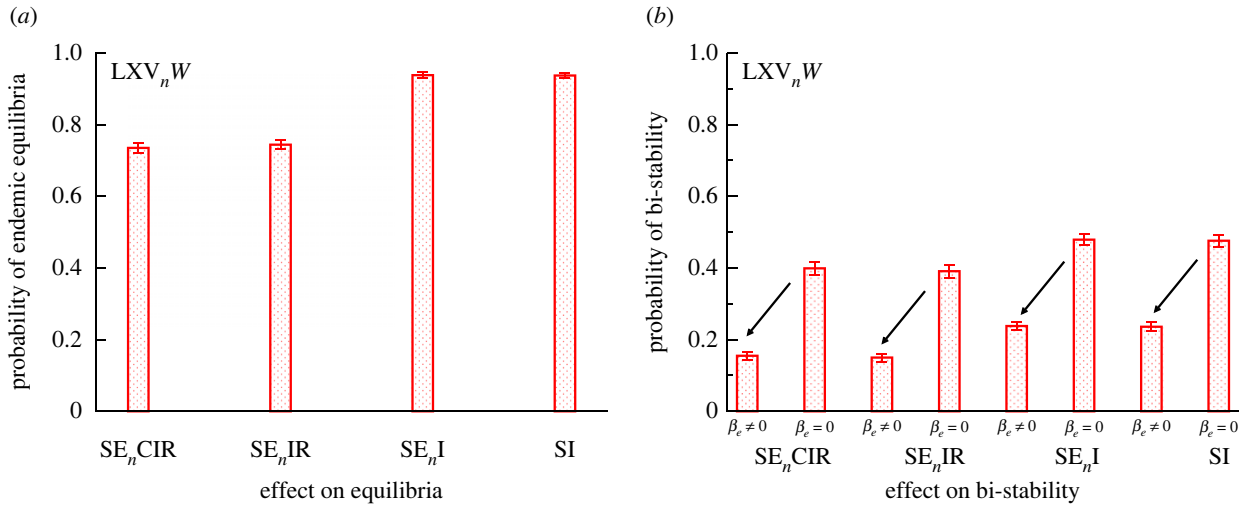

Figure 6. Effect of model complexity resulting from the addition of human and mosquito classes (and the corresponding increase in the number of parameters) on the fraction of the full parameter space corresponding to either the existence of a single endemic equilibrium (*a*) or the presence of bi-stability (*b*). (*a*) Effect on equilibria and (*b*) effect on bi-stability. (Online version in colour.)

a necessary condition for the presence of bi-stability. Bi-stability only arises, however, in certain regions of parameter space. Under superinfection, further refined necessary conditions for a given parameter can be derived. The presence of bi-stability can be studied in any vector–human coupled model by comparing the slopes of the curves represented in figure 2 at the origin for $W = 0$. By using this strategy, for instance for the model SI–XW, we derived a necessary condition for the onset of bi-stability in terms of b, the probability of infection of a susceptible human upon receiving an infectious bite. This condition $b_c < b < 0.5$ indicates that b has to be lower than 50% but higher than a certain critical threshold b_c determined by the other model parameters (see derivation and threshold value in electronic supplementary material §6.1).

4. Discussion

The results described here demonstrate that superinfection introduces the possibility of alternative steady states of low and high prevalence, with important implications for malaria control programmes. These findings resonate with insights gained from studies of other pathogens whose population dynamics exhibit nonlinear regime shifts, for example in [30] for the hepatitis B virus (HBV). Mathematical models for vector-borne diseases have tended to adopt representations of immunity from classical SIR formulations and to assume that coinfection with different strains of the same pathogen is a rare or unusual event. Yet immunity to malaria clearly differs from that in standard epidemiological models assuming full protection upon recovery (classical SIR models), and MOI is common in endemic regions where antigenic diversity is extremely large. Naturally acquired immunity involves diverse immunity responses to the different phases of the parasite within humans [31], and the major antigen of the blood stage of infection (PfEMP) is encoded by a multicopy gene family exhibiting extreme variation [32–34]. The antigenic complexity of the parasite in high transmission regions underlies superinfection and goes hand-in-hand with the existence of a vast reservoir of asymptomatic infections in individuals of all ages [5,6,35,36].

Similar epidemiological patterns of high prevalence and widespread population immunity characterize other vector-borne infections of non-human hosts in tropical regions [37]. Here too, a high degree of antigenic variation underlies superinfection, as shown for *Anaplasma marginale*, the most globally prevalent vector-borne pathogen of livestock infecting both wild and domestic ruminants [37]. Thus, our results might apply to other complex pathogens transmitted by vectors in human and non-human hosts whose immune responses must contend with large repertoires of antigenic variants.

In *falciparum* malaria, the additional nonlinearities created by important details of acquired immunity mean that bi-stability may not be uncommon in its empirical dynamics. Its presence would most likely express itself when control pushes the prevalence of the disease to lower levels, or more disconcertingly when malaria is expanding its range and establishing or re-establishing in new areas. It may have been ignored in linear statistical analyses focused on the impact of climate change on vector dynamics, or the quality of data available may have been gathered at insufficient temporal resolution or extent to detect sudden transitions in levels of prevalence. A recent study of highland malaria using long-term data from Kericho in Kenya presents evidence for the critical slowing down in malaria dynamics in response to different control strategies (M Harris, S Hay S, J Drake 2019, personal communication).

Dynamical consequences of bi-stability are numerous for outcomes of intervention and emergence/re-emergence. Small changes in parameters (for instance, biting rate a or mosquitoes' carrying capacity K) could give rise to large changes in incidence. Progressive control efforts may see no clear decrease of incidence until a sudden effect finally occurs. To achieve elimination, these efforts would require pushing transmission intensity to lower levels than under more standard transcritical transitions because of hysteresis. Drastic intervention may thus be required to move the system into the basin of attraction of the low prevalence equilibrium, and in so doing, avoid the transient effects of hysteresis with a protracted persistence of high prevalence. In the opposite direction, elimination states may be more

robust than in the case of standard transcritical transitions. Empirical evidence for the plausible robustness of the elimination state has been discussed in [38,39]. Also, the progressive relaxation of control efforts in endemic regions could generate sudden transitions from low to high incidence. Similarly, the range expansion and invasion of disease-free regions will tend to achieve low or high prevalence in a discontinuous fashion depending on the local conditions determining transmission intensity. Only beyond the upper threshold for which a single equilibrium exists (for $R_0 \gg 1$) will invasion establish high prevalence. For lower transmission intensity and $R_0 > 1$, the importation of cases would be trapped into the low prevalence equilibrium, even in areas where conditions previously allowed high prevalence. Finally, concerning variability, sudden shifts from low to large fluctuations in incidence may follow in epidemic regions from environmental conditions such as temperature warming driving the system across a critical threshold [11].

A positive feedback is the key ingredient for the strong nonlinearities that underlie regime shifts and associated alternative equilibria, as also illustrated by a number of ecological systems [1]. In our model, a positive feedback is introduced by the slowing-down effect of transmission intensity (rate of arrival of infectious bites) on the recovery rate (r) and therefore, the lengthening of the duration of infection. Thus, higher transmission intensity sustains and exacerbates higher transmission itself. Interestingly in HBV, higher prevalence of carriers leads through higher transmission rates to a younger age of infection, which in turn results in higher probability of carriage [30]. These are both examples of the reinforcement of transmission intensity through positive feedback. A second effect we considered in our model is on the lengthening of the duration of immunity (through σ). This effect is not sufficient for the occurrence of alternative steady states. This observation finds a clear explanation in the contrasting effects of the two assumptions: a lengthening of infectious periods produces a positive feedback on infection duration and transmission rate. With longer infections, a higher number of vectors acquire the parasite, increasing the rate of infectious bites, and concomitantly further slowing the recovery rate. By contrast, a deceleration of the loss of immunity slows down the return of resistant hosts to the susceptible class, and the resulting higher number of immune individuals decreases the number of infected vectors. In brief, this implements a negative feedback on transmission intensity.

Our derivation for the implications of superinfection for a critical transition is robust to consideration of a wide range of malaria models spanning different levels of complexity. Even the simple RM model exhibits bi-stability when the recovery rate slows down with increasing transmission intensity.

Previous malaria models showing bi-stability differ from ours in that they assume re-infection, and therefore only allow a host to re-acquire the parasite after complete clearance of infection and its return to the susceptible state [9,10]. When our findings are taken together with those of these earlier models, they underscore the broad generality of critical transitions in malaria models regardless of specific biological details on multiple infections and within-host dynamics. This generality follows from the observation that re-infection and superinfection (in the sense implemented here) effectively bracket a continuum of possible assumptions on the outcome of repeated exposure to infectious bites. In the former, repeated infections completely interfere with each other; in the latter, they do not 'see' each other at all. Neither complete interference (re-infection) nor a complete lack of interference (superinfection) is likely. Reality is likely to fall somewhere in between and closer to superinfection given that in endemic regions MOI is the rule rather than the exception, and that re-infection formulations cannot capture it.

A different kind of mechanism independent from repeated exposure has been proposed for alternative steady states in malaria dynamics. This mechanism combines density-dependent biting rates and disease-induced mortality [2]. This mechanism assumes that increasing disease levels would decrease total human population through disease-induced mortality, which then makes biting rates higher since they depend on the density of infectious mosquitoes per human, which, in turn, would raise infection levels in the human population. We believe this positive feedback is less important and considerably less general than the one described here, although it could work in conjunction with the main mechanism of repeated exposure described here, either through re-infection or superinfection.

Multiplicity of infection and the large antigenic diversity of the *Plasmodium* parasite in endemic regions complicate the acquisition of immunity in ways that can only be captured phenomenologically in standard transmission models. Whereas our abstraction is more suitable for the question addressed and the associated analytical approaches presented here, more complex representations could explore the generality of bi-stability. These representations include the explicit consideration of antigenic/strain variation and MOI in compartmental and individual-based models (e.g. [40–43]) and the age of hosts in age-structured (partial differential equations) models [44]. Stage-structured models could also be used to represent increasing levels of immunity acquired through repeated exposure, beyond the two-stage models considered so far [9,44]. Multiscale models representing both between-host and within-host dynamics (e.g. [45]) could investigate the consequences for critical transitions of relaxing our assumption of completely independent overlapping infections. We expect bi-stability to also occur in these more complex models depending on parameters.

An important next step is to confront these kinds of dynamics with data from surveillance efforts across changing control and environmental conditions (e.g. [44]). On-going developments in ecology and epidemiology include early warning systems related to critical transitions [46–49] and the prediction of bifurcations based on the monitoring of large perturbations [50]. These efforts should be brought together with sustained surveillance efforts over time and space, and with theoretical findings on the behaviour of mathematical models for the population dynamics of malaria and other vector-borne infections.

Data accessibility. This article has no additional data.

Competing interests. We declare we have no competing interests.

Funding. No funding has been received for this article.

References

1. Scheffer M. 2009 *Critical transition in nature and society*. Princeton, NJ: Princeton University Press.
2. Chitnis N, Cushing JM, Hyman JM. 2006 Bifurcation analysis of a mathematical model for malaria transmission. *SIAM J. Appl. Math.* **67**, 24–45. (doi:10.1137/050638941)
3. Lavine JS, King Aa, Bjørnstad ON. 2011 Natural immune boosting in pertussis dynamics and the potential for long-term vaccine failure. *Proc. Natl Acad. Sci. USA* **108**, 7259–7264. (doi:10.1073/pnas.1014394108)
4. Lavine JS, King AA, Andreasen V, Bjørnstad ON. 2013 Immune boosting explains regime-shifts in prevaccine-era pertussis dynamics. *PLoS ONE* **8**, e72086. (doi:10.1371/journal.pone.0072086)
5. Jenkins R, Omollo R, Ongecha M, Sifuna P, Othieno C, Ongeri L, Kingora J, Ogutu B. 2015 Prevalence of malaria parasites in adults and its determinants in malaria endemic area of Kisumu County, Kenya. *Malar. J.* **14**, 263. (doi:10.1186/s12936-015-0781-5)
6. Ghansah A *et al.* 2017 Seasonal variation in the epidemiology of asymptomatic *Plasmodium falciparum* infections across two catchment areas in Bongo district, Ghana. *Am. J. Trop. Med. Hyg.* **97**, 199–212. (doi:10.4269/ajtmh.16-0959)
7. Smith DL, Battle KE, Hay SI, Barker CM, Scott TW, McKenzie FE. 2012 Ross, MacDonald, and a theory for the dynamics and control of mosquito-transmitted pathogens. *PLoS Pathog.* **8**, e1002588. (doi:10.1371/journal.ppat.1002588)
8. Reiner RC *et al.* 2013 A systematic review of mathematical models of mosquito-borne pathogen transmission: 1970–2010. *J. R. Soc. Interface* **10**, 20120921. (doi:10.1098/rsif.2012.0921)
9. Keegan LT, Dushoff J. 2013 Population-level effects of clinical immunity to malaria. *BMC Infect. Dis.* **13**, 428. (doi:10.1186/1471-2334-13-428)
10. Águas R, White LJ, Snow RW, Gomes MGM. 2008 Prospects for malaria eradication in sub-Saharan Africa. *PLoS ONE* **3**, e1767. (doi:10.1371/journal.pone.0001767)
11. Alonso D, Bouma MJ, Pascual M. 2011 Epidemic malaria and warmer temperatures in recent decades in an East African highland. *Proc. R. Soc. B* **278**, 1661–1669. (doi:10.1098/rspb.2010.2020)
12. Ross R. 1911 *The prevention of malaria*. 2nd edn. London, UK: J. Murray.
13. Ross R. 1916 An application of the theory of probabilities to the study of *a priori* pathometry. *Proc. R. Soc. Lond. A* **92**, 204–230. (doi:10.1098/rspa.1916.0007)
14. MacDonald G. 1952 Ross's *a priori* Pathometry—a perspective. *Trop. Dis. Bull.* **49**, 813–829.
15. Fine PEM. 1975 A challenge for epidemiology: Ross's *a priori* Pathometry—a perspective. *Proc. R. Soc. Med.* **68**, 547–551.
16. Lindblade KA, Steinhardt L, Samuels A, Kachur SP, Slutsker L. 2013 The silent threat: asymptomatic parasitemia and malaria transmission. *Expert Rev. Anti Infect. Ther.* **11**, 623–639. (doi:10.1586/eri.13.45)
17. Nkumama IN, O'Meara WP, Osier FHA. 2017 Changes in malaria epidemiology in Africa and new challenges for elimination. *Trends Parasitol.* **33**, 128–140. (doi:10.1016/j.pt.2016.11.006)
18. Ghansah A *et al.* 2017 Seasonal variation in the epidemiology of asymptomatic *Plasmodium falciparum* infections across two catchment areas in Bongo District, Ghana. *Am. J. Trop. Med. Hyg.* **97**, 199–212. (doi:10.4269/ajtmh.16-0959)
19. Dietz K, Molineax L, Thomas A. 1974 Malaria model tested in African savannah. *Bull. World Health Org.* **50**, 347–357.
20. Aron JL, May RM. 1982 The population dynamics of malaria. In *Population dynamics and infectious diseases* (ed. RM Anderson), pp. 139–179. London, UK: Chapman and Hall.
21. Brauer F, Castillo-Chavez C, Mubayi A, Towers S. 2016 Some models for epidemics of vector-transmitted diseases. *Infect. Dis. Model.* **1**, 79–87. (doi:10.1016/j.idm.2016.08.001)
22. Ross R. 1910 *The prevention of malaria*. London, UK: J. Murray.
23. MacDonald G. 1957 *The epidemiology of malaria*. London, UK: Oxford University Press.
24. Smith DL, McKenzie FE, Snow RW, Hay SI. 2007 Revisiting the basic reproductive number for malaria and its implications for malaria control. *PLoS Biol.* **5**, e42. (doi:10.1371/journal.pbio.0050042)
25. Burden RL, Faires DJ. 1985 *Numerical analysis*. 3rd edn. Boston, MA: PWS Publishing Company.
26. Iooss G, Joseph DD. 1980 *Elementary stability and bifurcation theory*. New York, NY: Springer.
27. Anderson RM, May RM. 1991 *Infectious diseases of humans. Dynamics and control*. Oxford, UK: Oxford University Press.
28. Diekmann O, Heesterbeek JAP, Metz JAJ. 1990 On the definition and the computation of the basic reproduction ratio R_0 in models for infectious diseases in heterogeneous populations. *J. Math. Biol.* **28**, 365–382. (doi:10.1007/BF00178324)
29. Lloyd AL. 2001 Realistic distributions of infectious periods in epidemic models: changing patterns of persistence and dynamics. *Theor. Popul. Biol.* **60**, 59–71. (doi:10.1006/tpbi.2001.1525)
30. Medley GF, Lindop N, Edmunds W, Nokes D. 2001 Hepatitis-B virus endemicity: heterogeneity, catastrophic dynamics and control. *Nat. Med.* **7**, 619–624. (doi:10.1038/87953)
31. Struik SS, Riley EM. 2004 Does malaria suffer from lack of memory. *Immunol. Rev.* **201**, 268–290. (doi:10.1111/j.0105-2896.2004.00181.x)
32. Su XZ, Heatwole VM, Wertheimer SP, Guinet F, Herrfeldt JA, Peterson DS, Ravetch JA, Wellems TE. 1995 The large diverse gene family *var* encodes proteins involved in cytoadherence and antigenic variation of *Plasmodium falciparum*-infected erythrocytes. *Cell* **82**, 89–100. (doi:10.1016/0092-8674(95)90055-1)
33. Gupta S, Trenholme K, Anderson RM, Day KP. 1994 Antigenic diversity and the transmission dynamics of *Plasmodium falciparum*. *Science* **263**, 961–963. (doi:10.1126/science.8310293)
34. Day KP *et al.* 2017 Evidence of strain structure in *Plasmodium falciparum var* gene repertoires in children from Gabon, West Africa. *Proc. Natl Acad. Sci. USA* **114**, 201613018. (doi:10.1073/pnas.1708236114)
35. Geiger C, Agustar HK, Compaoré G, Coulibaly B, Sié A, Becher H, Lanzer M, Jänisch T. 2013 Declining malaria parasite prevalence and trends of asymptomatic parasitaemia in a seasonal transmission setting in North-Western Burkina Faso between 2000 and 2009–2012. *Malar. J.* **12**, 27. (doi:10.1186/1475-2875-12-27)
36. Gerardin J, Ouédraogo AL, McCarthy K, Eckhoff P, Wenger E. 2015 Characterization of the infectious reservoir of malaria with an agent-based model calibrated to age-stratified parasite densities and infectiousness. *Malar. J.* **14**, 231. (doi:10.1186/s12936-015-0751-y)
37. Futse JE, Brayton KA, Dark MJ, Knowles DP, Palmer GH. 2008 Superinfection as a driver of genomic diversification in antigenically variant pathogens. *Proc. Natl Acad. Sci. USA* **105**, 2123–2127. (doi:10.1073/pnas.0710333105)
38. Smith D *et al.* 2013 A sticky situation: the unexpected stability of malaria elimination. *Phil. Trans. R. Soc. B* **368**, 1623. (doi:10.1098/rstb.2012.0145)
39. Baeza A, Bouma MJ, Dhiman RC, Baskerville EB, Ceccato P, Yadav RS, Pascual M. 2013 Long-lasting transition toward sustainable elimination of desert malaria under irrigation development. *Proc. Natl Acad. Sci. USA* **110**, 15 157–15 162. (doi:10.1073/pnas.1305728110)
40. Gupta S, Maiden MCJ, Feavers IM, Nee S, May RM, Anderson RM. 1996 The maintenance of strain structure in populations of recombining infectious agents. *Nat. Med.* **2**, 437–442. (doi:10.1038/nm0496-437)
41. Artzy-Randrup Y, Rorick MM, Day K, Chen D, Dobson AP, Pascual M. 2012 Population structuring of multi-copy, antigen-encoding genes in *Plasmodium falciparum*. *eLife* **1**, e00093. (doi:10.7554/eLife.00093)
42. He Q, Pilosof S, Tiedje KE, Ruybal-Pesántez S, Artzy-Randrup Y, Baskerville EB, Day KP, Pascual M. 2018 Networks of genetic similarity reveal non-neutral processes shape strain structure in *Plasmodium falciparum*. *Nat. Commun.* **9**, 1817. (doi:10.1038/s41467-018-04219-3)
43. Karl S, White MT, Milne GJ, Gurarie D, Hay SI, Barry AE, Felger I, Mueller I, Marinho CR. 2016 Spatial effects on the multiplicity of *Plasmodium falciparum* infections. *PLoS ONE* **11**, 1–20. (doi:10.1371/journal.pone.0164054)
44. Águas R, White LJ, Snow RW, Gomes MGM. 2008 Prospects for malaria eradication in sub-Saharan Africa. *PLoS ONE* **3**, 1–6. (doi:10.1371/journal.pone.0001767)
45. Echhoff P. 2013 Mathematical models of within-host and transmission dynamics to determine effects of

malaria interventions in a variety of transmission settings. *Am. J. Trop. Med. Hyg.* **88**, 817–827. (doi:10.4269/ajtmh.12-0007)

46. Chaves LF, Pascual M. 2007 Comparing models for early warning systems of neglected tropical diseases. *PLoS Negl. Trop. Dis.* **1**, e33. (doi:10.1371/journal.pntd.0000033)
47. Scheffer M *et al.* 2009 Early-warning signals for critical transitions. *Nature* **461**, 53–59. (doi:10.1038/nature08227)
48. Kéfi S *et al.* 2014 Early warning signals of ecological transitions: methods for spatial patterns. *PLoS ONE* **9**, 10–13. (doi:10.1371/journal.pone.0092097)
49. Brett TS, Drake JM, Rohani P. 2017 Anticipating the emergence of infectious diseases. *J. R. Soc. Interface* **14**, 20170115. (doi:10.1098/rsif.2017.0115)
50. D'Souza K, Epureanu BI, Pascual M. 2015 Forecasting bifurcations from large perturbation recoveries in feedback ecosystems. *PLoS ONE* **10**, 1–19. (doi:10.1371/journal.pone.0137779)