Yakob, L; Lloyd, AL; Kao, RR; Ferguson, HM; Brock, PM; Drake-ley, C; Bonsall, MB (2017) Plasmodium knowlesi invasion following spread by infected mosquitoes, macaques and humans. Parasitology. pp. 1-10. ISSN 0031-1820 DOI: https://doi.org/10.1017/S0031182016002456

Downloaded from: http://researchonline.lshtm.ac.uk/3716606/

DOI: 10.1017/S0031182016002456

Usage Guidelines

Please refer to usage guidelines at http://researchonline.lshtm.ac.uk/policies.html or alternatively contact researchonline@lshtm.ac.uk.

Available under license: http://creativecommons.org/licenses/by/2.5/
**Plasmodium knowlesi** invasion following spread by infected mosquitoes, macaques and humans

LAITH YAKOB1*, ALUN L. LLOYD2, ROWLAND R. KAO3, HEATHER M. FERGUSON3, PATRICK M. BROCK3, CHRIS DRAKELEY4 and MICHAEL B. BONSALL5,6

1 Department of Disease Control, London School of Hygiene and Tropical Medicine, Keppel Street, London WC1E 7HT, UK
2 Department of Mathematics, Biomathematics Graduate Program, Center for Quantitative Sciences in Biomedicine, North Carolina State University, Raleigh, NC 27695, USA
3 Institute of Biodiversity, Animal Health and Comparative Medicine, College of Medical Veterinary and Life Sciences, University of Glasgow, Glasgow G61 1QH, UK
4 Immunology and Infection Department, London School of Hygiene and Tropical Medicine, Keppel Street, London WC1E 7HT, UK
5 Mathematical Ecology Research Group, Department of Zoology, University of Oxford, Oxford OX1 3PS, Oxfordshire, UK
6 St Peter’s College, Oxford OX1 2DL, Oxfordshire, UK

(Received 21 April 2016; revised 9 November 2016; accepted 9 November 2016)

**SUMMARY**

*Plasmodium knowlesi* is increasingly recognized as a major cause of malaria in Southeast Asia. *Anopheles leucosphyros* group mosquitoes transmit the parasite and natural hosts include long-tailed and pig-tailed macaques. Despite early laboratory experiments demonstrating successful passage of infection between humans, the true role that humans play in *P. knowlesi* epidemiology remains unclear. The threat posed by its introduction into immunologically naïve populations is unknown despite being a public health priority for this region. A two-host species mathematical model was constructed to analyse this threat. Global sensitivity analysis using Monte Carlo methods highlighted the biological processes of greatest influence to transmission. These included parameters known to be influential in classic mosquito-borne disease models (e.g. vector longevity); however, interesting ecological components that are specific to this system were also highlighted: while local vectors likely have intrinsic preferences for certain host species, how plastic these preferences are, and how this is shaped by local conditions, are key determinants of parasite transmission potential. Invasion analysis demonstrates that this behavioural plasticity can qualitatively impact the probability of an epidemic sparked by imported infection. Identifying key vector sub/species and studying their biting behaviours constitute important next steps before models can better assist in strategizing disease control.

Key words: invasion analysis, *Plasmodium knowlesi*, vector-borne disease, mathematical model, vector behaviour.

**INTRODUCTION**

The major human malaria species *Plasmodium falciparum* and *Plasmodium vivax* infect approximately 200 million people every year, killing nearly 600 000 (WHO, 2014). These parasites successfully established in human populations thousands of years ago following zoontic emergence from ape hosts in Africa (Liu et al. 2010, 2014). In 2004, a surprisingly high prevalence of *P. knowlesi* was found in humans in Malaysian Borneo when diagnostic microscopy was replaced by the more discriminatory method of nested PCR (Singh et al. 2004). This ground-breaking study identified that all blood samples from 208 people reporting atypical malaria infection in Kapit division of Malaysian Borneo were *P. knowlesi*-positive but misidentified as the morphologically similar *P. malariae* – a result subsequently corroborated by a larger, follow-up study conducted by the same group (Cox-Singh et al. 2008). Although long- and pig-tailed macaques are the natural hosts for this species, *P. knowlesi* has now been described in humans across several Southeast Asian countries and is the leading cause of human malaria in Malaysian Borneo (Singh and Daneshvar, 2013).

Mathematical models have been exploited in malaria research for a century and have produced considerable insight in both the epidemiology and control of infection (Smith et al. 2012). Model complexity has increased along with biological understanding and computational power; however, even the most complex ecological transmission models have fundamental elements that are identical, or analogous, to the original Ross–Macdonald formulations (Reiner et al. 2013). This family of models typically assume a single host species – an assumption that must be relaxed in the current context. Due to the relatively recent discovery of human infections with this species, and the correspondingly nascent...
understanding of infection processes, *P. knowlesi* models are relatively scarce and uncomplicated. The first published *knowlesi* malaria model expanded the Ross–Macdonald formula to account for heterogeneous biting of the vector (*Anopheles leucosphyrous* group) split between both macaque and human mammalian hosts (Yakob et al. 2010). A game theoretic approach to evolutionary invasion analysis of this deterministic system of ordinary differential equations was used to calculate the conditions under which a parasite might switch natural hosts from macaques to humans (Yakob et al. 2010).

Subsequent adaptations of this model were used to explore how vector control strategies could be optimized – both at larval and adult stages (Abdullahi et al. 2013); and, to explore how the basic reproduction number may be impacted by different ecological settings (Imai et al. 2014). Using a mathematical model, we build on this work to analyse the probability of successful parasite invasion into a host population following its introduction by an infected vector or host (either human or macaque).

Stochastic effects are known to be highly influential during the period immediately after the introduction of infection into a population (Bartlett, 1956), and are accounted for in calculating the probabilities of successful invasion of *P. knowlesi* introduced into susceptible populations (ranging from exclusively macaque to exclusively human). We also incorporate a flexible formulation that allows for qualitatively distinct host-selection vector biting behaviours because this aspect remains largely unknown for local vector species while also being: (1) critical to vector-borne disease epidemiology and control (Besansky et al. 2004); (2) likely to vary considerably (and not necessarily linearly) across differing proportionate representations of alternative mammalian hosts (Takk and Verhulst, 2013); and (3) also likely to vary according to local vector sibling species (Gillies, 1967). Insights gained into *P. knowlesi* epidemiology, including parasite invasion probabilities, are discussed along with proposed future research directions.

**Methods**

Figure 1 depicts the different epidemiological compartments in the model and their connections. Being a severely neglected tropical disease, there is a general absence of longitudinal studies detailing *P. knowlesi* malaria infection (Fornace et al. 2015). Consequently, a flexible and open-ended description of the transmission dynamics (Yakob, 2016a, b) is presented and used to calculate between-species parasite transmission numbers as well as invasion probabilities. Sensitivity analysis of the parameters underlying these thresholds will determine the aspects of unknown infection biology that might constitute priorities for future research.

Transmission dynamics

\[
\frac{dS}{dt} = \mu + yI + \tau R - m_Np_{11}b_{11}SZ - \mu S \tag{1}
\]

\[
\frac{dI}{dt} = m_Np_{11}b_{11}SZ - (\gamma + \varepsilon + \pi + \mu)I \tag{2}
\]

\[
\frac{dR}{dt} = \varepsilon I + \kappa A - (\tau + m_N\theta_{11})(I + \pi + \mu)R \tag{3}
\]

\[
\frac{dA}{dt} = \pi I + m_N\theta_{11}b_{11}Z - \kappa A - \mu A \tag{4}
\]

\[
\frac{dX}{dt} = \mu_v - (p_{11}b_{11}v(I + \sigma A))X - \mu_v X \tag{5}
\]

\[
\frac{dY}{dt} = (p_{11}b_{11}(I + \sigma A) + (1 - p_{11})b_{22}(I + \sigma A))X - (\zeta + \mu_v)Y \tag{6}
\]

\[
\frac{dZ}{dt} = \zeta Y - \mu_v Z \tag{7}
\]

\[
\frac{dS_N}{dt} = \mu_N + y_N I + \tau_N R_N - m_N(1 - p_{11})b_{22}N_SZ - \mu_NS_N \tag{8}
\]

\[
\frac{dI_N}{dt} = m_N(1 - p_{11})b_{22}N_SZ - (\gamma_N + \varepsilon_N + \pi_N + \mu_N)I_N \tag{9}
\]

\[
\frac{dR_N}{dt} = \varepsilon_N I_N + \kappa_N A_N - (\tau_N + m_N\theta_N)(1 - p_{11})b_{22}N_Z - \mu_N R_N \tag{10}
\]

\[
\frac{dA_N}{dt} = \pi_N I_N + m_N\theta_N(1 - p_{11})b_{22}N_Z R_N - (\kappa_N + \mu_N)A_N \tag{11}
\]

All variables depicting epidemiological categories are proportions. Susceptible humans (*S*) become infectious (*I*) following a bite from an infectious vector (*Z*). Infectious humans revert to susceptible at rate γ. Different parameterization of the clearance rate of symptomatic infection (ε), the rate of reversion to full susceptibility (τ) and the susceptibility to asymptomatic infections (θ) affects the temporal-ity of immunity. Human hosts can become asymptomatically infected (*A*) directly progressing from symptomatic infection when the rate termed π is greater than 0, or following on from recovery (*R*) and subsequent reinfection (θ > 0). Asymptomatic infection in macaques is assumed to be lifelong (by setting recovery from secondary infection, κN, to equal 0) whereas humans are assumed to be able to clear the parasites and recover at rate κ. Processes governing infection in the natural macaque hosts are denoted by subscript *N*. Susceptible vectors...
(X) become infected (Y) following a bite from an infectious host, and after the extrinsic incubation period (1/\(\zeta\)), become infectious (Z). The ratio of mosquitoes to hosts is denoted \(m\) (subscript H and N for human and non-human hosts respectively) and the vector mortality rate is \(\mu_V\). Transmission coefficients are denoted by \(b'\) with associated subscripts (these are distinguished by the host species involved should species-specific estimates arise in the future, e.g. \(b'_{HH}\) is the transmission coefficient from vectors to human hosts and comprises the bite rate per vector multiplied by the probability of parasite transmission per bite). However, because there are two alternative host species, bites must be further partitioned according to which host species actually receives the bite from a vector. This required the following framework to apportion these bites among alternative host species as determined by both their relative abundances and intrinsic vector preferences for specific host species.

**Functional responses in the human blood index**

The proportion of bites on humans is determined by a flexible formula that allows for a wide range of different functional responses depicting distinct vector-biting behaviours:

\[
p_{HH} = \frac{H}{H + \alpha(1 - H)\beta} \quad (12)
\]

Here \(p_{HH}\) is the ‘human blood index’ (Garret-Jones, 1964); \(H\) is the availability of humans relative to all other potential hosts; \(\alpha\) and \(\beta\) are parameters that shape the functional response of human bite proportion relative to all potential host species. Type I responses (\(\alpha = \beta = 1\)) assume bite distribution among alternative host species that is directly proportionate to their relative availability; Type II human blood index responses (\(0 \leq \alpha < 1\) and \(\beta \geq 1\)) are convex-up with increasing human availability relative to alternative hosts and describe an anthropophilic vector; Type III responses (\(\alpha \geq 1\) and \(\beta > 1\)) are s-shaped and depict a zoophilic vector that becomes increasingly anthropophilic with increased human encounters; Type IV responses (\(\alpha > 1\) and \(0 < \beta \leq 1\)) are convex-down and describe a zoophagic vector that only bites humans when there are few alternatives; and type V responses (\(0 \leq \alpha \leq 1\) and \(0 < \beta < 1\)) are s-shaped reflected in the \(y = x\) line and describe a negative prey-switching (Abrams et al. 1993) analogue, e.g. whereby anthropophilic vectors avoid a nuisance response. A fuller description of these functional responses can be found in (Yakob, 2016b).

**Figure 2** illustrates the shape of association between the human blood index and human host availability relative to all potential blood hosts. A complete range of host availabilities is displayed – from entirely macaque populations (0 on the x-axis) to entirely human populations (1 on the x-axis), and everything in between, e.g. at the half-way mark (0.5) of the x-axis, equal availability of humans and macaques is shown for a mixed population. This formula is used to assess the importance of different host availabilities (i.e. different environmental settings) and
different host-feeding behaviours in the resulting between-species transmission rates and invasion analysis.

**Calculation of the basic reproduction number: entries of the next generation matrix (NGM)**

Standard theory states that the basic reproduction number, $R_0$, can be calculated as the largest eigenvalue (i.e. the spectral radius) of the NGM, $K$ (Diekmann and Heesterbeek, 2000). In the present context, involving two types of hosts and one type of vector, $K$ is a $3 \times 3$ matrix. Entries of $K$, which we write as $K^{ij}$, depict the expected number of infections of each type (human host, macaque host or vector) that are directly produced by an infectious individual of each type (human, macaque or vector) when the system is at (or very near) the infection-free equilibrium. Standard theory shows how the $K^{ij}$ can be calculated by considering the linearized infected subsystems, decomposing each into two matrices (Diekmann *et al.* 2010): one depicting the infection transmission ($T$) and the other depicting all other transitions ($2$). Each $K^{ij}$ is calculated as the spectral radius of the NGM for that component of the system calculated from $-\Sigma^{-1}$ (Diekmann *et al.* 2010). For the present system, there are four non-zero entries of the NGM (whose derivations are shown below): the average number of human cases arising from an infected vector ($K^{VIH}$); the average number of macaque cases arising from an infected vector ($K^{VNI}$); the average number of vector infections arising from an infected human ($K^{HV}$); and the average number of vector infections arising from an infected macaque ($K^{NNV}$). These between-species transmission numbers and their sensitivities to the underlying model parameters are assessed in terms of the Spearman’s rank correlation coefficient calculated from 5000 iterations of a Monte Carlo multivariate sensitivity analysis (whereby all parameters were assumed to have triangular probability distributions $\pm 10\%$ about the median values described in Table 1). Global sensitivity analysis was used to ascertain the processes that are most instrumental in *P. knowlesi* transmission rates.

**Invasion probabilities**

For deterministic model formulations, if the average number of secondary infections arising from a primary infection exceeds unity, the successful invasion of the pathogen into the host population is guaranteed. New epidemics driven by the imports of small numbers of infected hosts or vectors are less certain than implied by determinism: for instance, an initial infective could, with some probability, recover or die before causing any secondary infections. Calculation of invasion probabilities requires a stochastic model, a framework that can be obtained by reinterpreting the rates of continuous movement between compartments in the deterministic differential equation model as rates (probabilities per unit time) at which discrete transition events occur in the stochastic model. Branching process theory has been used to calculate the extinction probability of (potential) epidemics sparked by the introduction of infected individuals (Athreya and Ney, 1972) and this has recently been expanded to calculate invasion probabilities for vector-borne disease systems allowing for two levels of host attractiveness (Lloyd *et al.* 2007). In line with these previous developments, invasion probabilities among the different host types are the same, in that an outbreak amongst one host type necessarily means ongoing infections amongst other host types, even if this is just a spill-over effect. To the best of our knowledge the current analysis constitutes the first to describe methods of invasion analysis for a real multi-host vector-borne disease system. This theory requires the calculation of probability generating functions, $G(i)$ that summarize the distributions of secondary infections of each type of species that results from the introduction of an infected vector, macaque or human. In these functions, secondary infections amongst vectors, macaques and humans are labelled using powers of $s_v, s_h$ and $s_m$, respectively. As in the deterministic analysis, all quantities are calculated at the infection free equilibrium. For the human host population, calculation of the probability generating function needs to account for the fact that an infectious human host in the $I$ compartment can move to the asymptomatic (A) compartment and continue to cause infections. This is achieved by calculating generating functions for infections produced while in the two compartments and combining them, accounting for the probability of making the infected ($I$) to asymptomatic ($A$) transition, to give the overall generating function for an infective human host. We remark that the branching process analysis does not need to consider the transition from recovered ($R$) to asymptomatic ($A$) (recovered individuals becoming re-infected) as the rate of this flow is negligible near the infection free equilibrium. The generating function for the number of secondary infections generated from the infected ($I$) class is

$$G_I(s_v) = \frac{1}{1 + R_I(1 - s_v)} \quad (13)$$

where $R_I = m_I(p_I b_{1IV}/(\gamma + \epsilon + \pi + \mu))$. The generating function for the asymptomatic ($A$) class is

$$G_A(s_v) = \frac{1}{1 + R_2(1 - s_v)} \quad (14)$$

where $R_2 = m_I(p_I b_{1IV}/(\epsilon + \mu))$. With $\phi$ denoting the probability that an infected ($I$) individual will become asymptomatic ($A$), i.e. $\phi = \pi/(\gamma + \epsilon + \pi + \mu)$,
Table 1. *Plasmodium knowlesi* mathematical model parameters, descriptions, median values and source

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Definition</th>
<th>Median values humans (macaques)</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>$b_{VH}$</td>
<td>Transmission coefficient (to humans); bite rate $\times$ transmission probability</td>
<td>0·1; 1/3 $\times$ 0·3</td>
<td>Rickman et al. (1990)</td>
</tr>
<tr>
<td>$b_{VN}$</td>
<td>Transmission coefficient (to non-humans); bite rate $\times$ transmission probability</td>
<td>0·1; 1/3 $\times$ 0·3</td>
<td>Rickman et al. (1990)</td>
</tr>
<tr>
<td>$b_{HV}$</td>
<td>Transmission coefficient (humans $\rightarrow$ vectors); bite rate $\times$ transmission probability</td>
<td>0·007; 1/3 $\times$ 0·02</td>
<td>Bonnet et al. (2003)</td>
</tr>
<tr>
<td>$b_{NV}$</td>
<td>Transmission coefficient (non-humans $\rightarrow$ vectors); bite rate $\times$ transmission probability</td>
<td>0·007</td>
<td>Bonnet et al. (2003)</td>
</tr>
<tr>
<td>$m_h$</td>
<td>Ratio of mosquitoes to human hosts</td>
<td>10 (but varied for invasion analysis)</td>
<td>Assumption</td>
</tr>
<tr>
<td>$m_N$</td>
<td>Ratio of mosquitoes to macaque hosts</td>
<td>10 (but varied for invasion analysis)</td>
<td>Assumption</td>
</tr>
<tr>
<td>$\gamma$</td>
<td>Recovery rate</td>
<td>0·07 (0) day$^{-1}$</td>
<td>Coatney et al. (2003)</td>
</tr>
<tr>
<td>$\varepsilon$</td>
<td>Clearance rate of symptomatic infection</td>
<td>0·07 (0) day$^{-1}$</td>
<td>Coatney et al. (2003)</td>
</tr>
<tr>
<td>$\kappa$</td>
<td>Clearance rate of asymptomatic infection</td>
<td>0·01 (0) day$^{-1}$</td>
<td>Franks et al. (2001)</td>
</tr>
<tr>
<td>$\pi$</td>
<td>Asymptomatic primary infection rate</td>
<td>0·14 (0·14) day$^{-1}$</td>
<td>Assumption</td>
</tr>
<tr>
<td>$\theta$</td>
<td>Susceptibility to secondary asymptomatic infection</td>
<td>1 (0)</td>
<td>Assumption</td>
</tr>
<tr>
<td>$\tau$</td>
<td>Full susceptibility reversion rate</td>
<td>0·0057 (0) day$^{-1}$; 1/(ln(2) $\times$ 3 years)</td>
<td>White et al. (2014)</td>
</tr>
<tr>
<td>$\sigma$</td>
<td>Adjustment factor for asymptomatic transmissibility to vector</td>
<td>0·25 (0·25)</td>
<td>Okell et al. (2012)</td>
</tr>
<tr>
<td>$\mu$</td>
<td>Birth and death rate of hosts (i.e. stable population)</td>
<td>3·4 $\times$ 10$^{-5}$ (2·7 $\times$ 10$^{-4}$ day$^{-1}$)</td>
<td>Anonymous (2010), Yanuar et al. (2009)</td>
</tr>
<tr>
<td>$\mu_N$</td>
<td>Birth (or maturation) and death rate of vectors (i.e. stable population)</td>
<td>0·1 day$^{-1}$</td>
<td>Yakob et al. (2010)</td>
</tr>
<tr>
<td>$\zeta$</td>
<td>Rate of parasite development within vector</td>
<td>0·1 day$^{-1}$</td>
<td>Collins (2012)</td>
</tr>
</tbody>
</table>

The generating function for the number of secondary infections generated after departure from the infected ($I$) class is given by

$$G_Z(s_V) = 1 - \phi + \phi G_A(s_V)$$

(15)

Making use of the fact that the generating function for the sum of two independent random variables is the product of their generating functions, we have that the generating function for the secondary infections resulting from an infected human host is given by $G_{IIV}(s_V) = G_I(s_V).G_Z(s_V)$ and hence

$$G(s_V) = \frac{1}{1 + R_I(1 - s_V)} \left( 1 - \phi + \phi \frac{1}{1 + R_Z(1 - s_V)} \right)$$

(16)

The generating function, $G_{NV}(s_V)$, describing the distribution of the number of vectors infected by an infectious macaque is obtained similarly. The generating function for the numbers of humans and macaques infected by an infectious vector is $G_V(s_h, s_n)$, where

$$G_V(s_h, s_n) = \frac{1}{1 + K_{VH}(1 - s_h) + K_{VN}(1 - s_n)}$$

(17)

As in Lloyd et al. (2007), extinction probabilities following an introduction of an infected vector, human or macaque ($s_V$, $s_h$ and $s_n$, respectively) are found by solving the set

$$G_V(s_h, s_n) = s_V$$
$$G_{IIV}(s_V) = s_h$$
$$G_{NV}(s_V) = s_n$$

(18)

This is most easily achieved by substituting the second and third of these equations into the first, leaving an equation for $s_V$ alone. This results in a fifth degree polynomial for which one root is $s_V = 1$, and thus leaves a quartic polynomial to solve for $s_V$. This equation can be solved numerically and $s_h$ and $s_n$ found by substitution. Standard theory shows that these invasion probabilities are all zero when the basic reproduction number, $R_0$, of the system is less than one and fall between 0 and 1 when $R_0$ is greater than one (i.e. invasion happens with some non-zero probability, but is not guaranteed).

Previous explorations of multi-host systems have assumed that the proportion of bites on alternative host species is directly proportional to their relative availability. Using the new formulation that allows for qualitatively different functional responses in
vector bite behaviours [equation (12)], the sensitivity of invasion probabilities to this neglected aspect of disease vector ecology was also assessed.

RESULTS

NGMs were used to calculate the expected number of infections of each type (human host, macaque host or vector) that are directly produced by an infectious individual of each type:

\[
K_{HV} = \frac{m b_{HV} p_{H}(\kappa + \mu + \pi \sigma)}{(\kappa + \mu)(y + \pi + \varepsilon + \mu)}
\]

\[
K_{NV} = \frac{m b_{NV}(1 - p_{H})(\kappa_{N} + \mu_{N} + \pi \sigma_{N})}{(\kappa_{N} + \mu_{N})(y_{N} + \pi_{N} + \varepsilon_{N} + \mu_{N})}
\]

\[
K_{VH} = \frac{b_{VH} p_{H} \zeta}{\mu_{V}(\mu_{V} + \zeta)}
\]

\[
K_{VN} = \frac{b_{VN}(1 - p_{H}) \zeta}{\mu_{V}(\mu_{V} + \zeta)}
\]

The resulting basic reproduction number, \( R_0 \), is calculated as:

\[
R_0 = \sqrt{(K_{HV} K_{VH} + K_{NV} K_{VN})}
\]

Using our baseline parameterisation, the numerical value of \( R_0 \) is calculated to be 79.9. Figure 3 describes the sensitivity of the parasite transmission numbers between species to the parameter values in the form of tornado plots. Across the different functional response Types, there is good qualitative consistency in the transmission numbers’ sensitivity to underlying parameters. Intuitively, both \( K_{VH} \) and \( K_{VN} \) are highly sensitive to the mosquito mortality rate – a parameter that is well understood to be strongly influential in classic models of vector-borne diseases (Macdonald, 1956). Both \( K_{IV} \) and \( K_{IN} \) are similarly sensitive to the transmission coefficients \( b \) and very insensitive to mammalian host longevity (inverse of their respective mortality rates, \( \mu \) and \( \mu_{N} \)) as per traditional malaria models. Of note is the considerable variation in transmission numbers in relation to the availability of humans relative to all alternative blood hosts, \( H \) whereby \( H \) was the most influential parameter for all transmission numbers under a Type III functional response (a zoophagic vector that becomes increasingly anthropophilic with increased human encounters) and of markedly lower significance under a type V response (negative prey-switching). This result is apparent from Fig. 2.

Sensitivity analysis was conducted at \( H = 0.5 \) (i.e. humans and macaques are equally available) because
this is where differences between the types are most pronounced. The gradient of the human blood index as a function of human availability relative to all blood meal hosts is steepest for type III and flattest for type V at this cross-section. This ranking in sensitivity will shift non-monotonically for the different functional types in vector biting behaviour across the range of alternative host availabilities.

**Figure 4** shows the invasion probabilities for *P. knowlesi* in relation to host availability and vector host-selection behaviours. General trends arise when comparing these probabilities across scenarios whereby the pathogen is introduced by vectors, humans and macaques: introduction of the pathogen by an infected vector is most likely to elicit an outbreak when macaques are the dominant blood host (i.e. when the human blood index and the human availability relative to all blood-hosts are between zero and 0.5); similarly, for all biting Types, the pathogen is least likely to invade when introduced by infected humans and most likely to invade when introduced by macaques. This can be explained by the assumed superiority of macaques as parasite hosts (they are assumed to remain infectious for life). Additionally, regardless of the introducing species, maximum invasion probabilities are achieved with a Type IV vector (a zoophilic vector that only bites humans when there are few alternatives). For most biting Types, parasite invasion is most likely for mid-level human availabilities and HBI (i.e. when there is a mix of blood host species). However, some important caveats emerge under specific biting Type scenarios. For Type IV and V mosquitoes (zoophilic or switching to zoophilic when human hosts dominate), an invasion driven by malaria imported by an infected macaque has the highest probability when the HBI approaches unity. These biting behaviours would be the most likely to ensure that the importing macaque is bitten and thereby transmits the parasite.

**DISCUSSION**

Malaria caused by *P. knowlesi* can be a highly debilitating and potentially fatal disease. To improve our understanding of this neglected tropical disease, we developed models to explore the probability of *P. knowlesi* invasion into different populations.

Multivariate sensitivity analyses highlight aspects of vector and pathogen life history that are most influential in disease transmission. Consistent with models of other malarias, disease transmission is critically sensitive to vector longevity. Accurate age-grading for natural anopheline mosquitoes remains a major hurdle and most estimates come from ovarian examination of the number of
gonotrophic cycles that females have undergone (Cook and Sinkins, 2010). Not even rough estimates produced through this indirect measuring method are yet available for members of A. leucosphyrous group. Additionally, this group is made up of several species that are morphologically impossible to distinguish (Sallum et al. 2005) and whose life histories, bite behaviours and thus contribution to P. knowlesi transmission are only just beginning to be uncovered (Vythilingam et al. 2006; Tan et al. 2008; Wong et al. 2015). Future modelling efforts incorporating entomological parameters will require allowing for considerable uncertainty – as incorporated here – until empirical information becomes available.

The current study constitutes the first endeavour in determining the probability of successful invasion following a P. knowlesi introduction into a susceptible population. This is particularly relevant for newly emerging infectious diseases because of their vulnerability to fade-out through random effects when infection numbers are low. To conduct this invasion analysis, it was assumed that the human hosts were immunologically naïve. In terms of P. knowlesi transmission, over 70% of infections are in individuals over the age of 20 years (W. Grigg et al. in prep.). This is not the epidemiological profile that would be expected if acquired immunity were an important transmission determinant locally. There is good evidence that P. knowlesi exhibits unstable transmission in humans (with a strong seasonal effect). Indeed, unstable transmission would be expected for a spill-over parasite. Together, these factors suggest that human populations that suffer from P. knowlesi infection do so through the repeat invasion of the parasite into humans from the macaque reservoir; and, that sustained transmission within humans over prolonged periods is seldom (if ever) experienced. Therefore, the assumption of an immunologically naïve human population with which to simulate P. knowlesi invasion currently seems appropriate.

The current study highlights vector biting behaviours that can profoundly impact the probability of successful P. knowlesi invasion (e.g. maximum invasion probability is three-fold higher for Type IV mosquitoes compared to a Type II vector). Critical in ascertaining the true threat that humans pose in transporting infection between different populations will be identification of the functional response in vector biting behaviour to variations in the availability of alternative blood hosts.

An in-depth analysis was conducted into how vectors respond to differing availabilities of alternative blood sources in terms of their host selection and how this impacts transmission. When non-linear responses are accounted for, quantitative differences arise in the parasite transmission numbers between species but qualitative differences emerge in the invasion probabilities. For example, when humans constitute the overwhelming majority of the available blood hosts, invasions sparked by infected macaques are completely precluded when spread by vectors exhibiting Type I, II or III responses. Establishing how local vector biting behaviour responds to a changing environment as humans increasingly encroach upon and supplant macaque habitats will be key to addressing the likelihood of P. knowlesi spread by human (or macaque) importation. Semi-field experiments using varied availabilities of alternative hosts and testing blood meals of fed mosquitoes could help improve understanding of this behaviour.

Following the precedents of the major human malaria species P. falciparum and P. vivax, P. knowlesi may be in the process of emerging as a substantive agent of malaria from primates into human populations – and recent field studies suggest that distinct parasite strains have invaded human populations (Ahmed et al. 2014; Divis et al. 2015; Pinheiro et al. 2015). This offers a unique opportunity to identify the environmental drivers behind the parasite’s evolution. To this end, the current study in which methods are developed to calculate invasion probabilities for multi-host malaria species a unique avenue for further investigation.

The present study highlights areas requiring further investigation. Biological understanding for P. knowlesi is germinal (although burgeoning) and currently dictates the appropriate level of complexity for disease models. Numerous host, parasite and environmental factors impact the epidemiology of all malarias and the coming years can be expected to better equip us in building upon this initial effort to simulate P. knowlesi invasion. For example, haemoglobinopathies are known to impact malaria epidemiology and (particularly beta thalassemia) occur at high rates in P. knowlesi-endemic populations. Currently, it is unknown whether/how these haemoglobinopathies affect susceptibility to P. knowlesi infection and these were consequently omitted from the current analysis. Additionally, given the overlapping endemicity with other malaria species in some regions, a future direction of the current work would be the exploration of the effects of P. knowlesi invasion in regions with P. falciparum and/or P. vivax already. However, much of our parameterization comes from studies in Sabah where levels of P. falciparum and P. vivax transmission are very low and unlikely to impact P. knowlesi invasion.

Another shortcoming arising from data paucity is the need to resort to parameter values gleaned from classic malaria entomological and epidemiological studies. Recent genetic analysis suggests a lack of clustering of parasite genotypes in humans or macaques, which may be suggestive of zoonotic rather than human-vector-human transmission (Lee et al. 2011; Divis et al. 2015). However, a similar result would be anticipated under the circumstance that
human outbreaks were limited in size, i.e. transmission chains were relatively short. A comprehensive multivariate sensitivity analysis allowed detection of the model parameters for which direct estimates were as yet unavailable and that were simultaneously highly influential in disease transmission. As described above, mosquito longevity is highly influential, but, so too is the vector biting behaviour. Additionally, seasonal effects on vector species’ (or sibling species’) abundance (absolute as well as relative to one another) have only recently been described for Anopheles balabacensis (Wong et al., 2015), and the integration of these new data into seasonally driven entomological models constitutes important future work.

Following a successful control campaign, malaria incidence in Malaysia has declined considerably in recent years and targets have been set for imminent elimination (Cotter et al., 2011). Unfortunately, the current endemicity of P. knowlesi threatens elimination in this region (William et al., 2013). While informing the epidemiology and control of a considerable public health threat, rapid knowledge development in the ecology of this newly emerging disease can also be expected to provide invaluable insight into the evolutionary processes underlying successful pathogen invasion into humans.

Acknowledgements

The authors would like to thank the many reviewers that contributed important ideas to the study.

Competing Interests

We have no competing interests.

Author Contributions

L.Y. and M.B.B. conceived the study; L.Y. produced the model; L.Y. and A.L.L. carried out model analysis. All authors interpreted model output; contributed important intellectual content; and gave their final approval of the version to be published.

Financial Support

R.R.K., H.M.F., P.B. and C.D. were supported by MRC ESEI grant G1100796/1.

References


