

SUPPLEMENTARY TEXT 1

Part 1

The modelling of the effects of growth rate on infection dynamics.

In this supplement we derive the predicted infection curve for patients with different growth rates of parasites in blood. We assume normally distributed growth rates (expressed as parasite multiplication rate (PMR)) in each age group.

The model assumes a standard dynamic of infection. After the infective bite malaria parasites travel to the liver where they infect liver cells where they undergo asexual multiplication. After 7 days [1] the infected liver cells rupture, releasing approximately 20,000 merozoites per each infected liver cell [2]. Merozoites then invade red blood cells (RBC) initiating the blood stage infection. After approximately 2 days the infected RBCs rupture releasing new merozoites and the blood stage cycle repeats. In this study we use the PMR, which reflects the fold-increase in number of infected RBCs over the two-day life cycle. The blood stage parasites can only be detected some time after the initiation of the blood stage infection, when parasite concentration reaches the detection threshold.

The model of blood stage immunity accounts for the exponentially distributed waiting-time to the initiation of blood stage infection and the delay that is needed for parasites to reach the detection threshold (assuming initial geometric parasite growth). The model assumes a normal distribution of PMR in the age groups with means m_i ($i=1,\dots,4$); and standard deviation proportional to the mean by the parameter b (which is the same for all age groups). The rate of initiation of new blood stage infections (k) is the same for all age groups. The detection threshold for microscopy (T_{micro}) data is set to 40 and is not a free parameter, as this number corresponds to the minimal value in the microscopy dataset and conforms to the published values of 20-50 parasites per microlitre [2, 3]. There is a greater variation and uncertainty in the PCR detection threshold in the literature (0.02 to 3 parasites/microlitre) [3, 4] so we set the PCR detection

threshold as a free parameter T_{PCR} . This model assigns the same three parameters to all age groups (b, k, T_{PCR}) and fits one mean growth rate to each age group (m_1-m_4). Thus the model has seven parameters and assumes the variation between age groups arises from differences in blood stage growth rates.

Thus, we assume that the infective mosquito bites occur continuously and independently of each other. Consequently, the waiting time until an infective bite, and therefore the time until a successful initiation of blood stage, will be distributed exponentially. We denote the cumulative density function (CDF) of an exponential distribution by $F_E(.)$ with the average rate of successful initiations of blood stage infections per day equal to k .

We also assume that the parasite multiplication rate (PMR) has a normal distribution, with the probability density function (PDF) and the cumulative density function (CDF) denoted by $f_N(.)$ and $F_N(.)$ respectively. The mean of the normal distribution is m and the standard deviation is σ . Such a distribution of PMR will produce the distribution of delays between the initiation of blood stage infection and the detection of parasites. A PMR <1 implies that, for each currently infected RBC, less than one newly infected RBC will be produced in the next round of infection and the parasite will not grow.

From these assumptions it follows that the final infection function will be a convolution of an exponential infection function representing the time to initiation of blood stage infection, and the distribution function of the delays between the initiation of blood stage infection and the detection of parasites.

To find the distribution of delays, let us consider how the concentration of parasites changes with time. The initial geometric growth of parasites in blood can be approximated by the formula (1.1).

$$C(t) = Ar^{t/2}, \quad (1.1)$$

where $C(t)$ is the concentration of parasites at a time t after emerging from the liver, r is the average PMR, and constant A is the concentration of parasites in blood at the beginning of the blood stage. $A=a/V$, where a is the initial number of parasitized RBCs, and V is the blood volume.

We used Chart 1 in reference [5] in order to find the average blood volumes in age groups $V_i, i=1,\dots,4$ ($V_1 = 1.1 \times 10^6 \mu\text{l}$, $V_2 = 2 \times 10^6 \mu\text{l}$, $V_3 = 3.3 \times 10^6 \mu\text{l}$, $V_4 = 5 \times 10^6 \mu\text{l}$). The initial number of parasitized RBCs for a single bite we estimated as 5.6×10^4 , assuming that after 5 simultaneous bites the initial number of infected RBC was 28×10^4 [2].

Knowing the PDF of the PMR and the relationship between PMR and the delay to detection, we can find $g(t)$ – the PDF of the delay, for the fraction of population of $1 - F_N(x)$ which has the PMR > 1 . To do this we used the formula for the distribution of a function of a random variable. It requires the inverse function to the delay function of r . The inverse function $r(t) = t(r)^{-1} = (T / A)^{2/t}$, was found from the equation

$$C(t) = T, \quad (1.2)$$

where T is the detection threshold.

Its first derivative being $r'(t) = 2(T / A)^{2/t} \ln(T / A) / t^2$. Thus the distribution function of the delay will have following form

$$g(t) = \frac{f_N(r(t)) |r'(t)|}{1 - F_N(1)} = \frac{2(T / A)^{2/t} \ln(T / A) f_N\left((T / A)^{2/t}\right)}{t^2 (1 - F_N(1))}, \quad (1.3)$$

Now we can find the infection function $S(t)$, which includes the convolution of the CDF of the exponential distribution and the distribution of the delays.

$$S(t) = \begin{cases} 1 - (1 - F_N(1)) \int_{\tau}^t F_E(t-x-\tau) g(x-\tau) dx, & t > \tau, \\ 1, & 0 < t \leq \tau. \end{cases} \quad (1.4)$$

The time constant τ days, is the first possible moment of blood stage infection following to treatment. We set $\tau = 7$ days, since the blood concentration of lumefantrine (reported in [6, 7]) even by day 7 post treatment is high enough (more than 280 ng/ μ l) to kill a relatively small number of parasites released from liver .

Reducing the numerator and the denominator in (1.4) by $1 - F_N(1)$, we obtain a simpler expression for the infection function:

$$S(t) = \begin{cases} 1 - \int_{\tau}^t F_E(t-x-\tau) \hat{g}(x-\tau) dx, & t > \tau, \\ 1, & 0 < t \leq \tau. \end{cases} \quad (1.5)$$

where $g(t) = \frac{2(T/A)^{2/t} \ln(T/A) f_N((T/A)^{2/t})}{t^2}$.

In order to reduce the number of parameters, we assume that the standard deviation σ_i is proportional to the mean m_i in each group (group index $i=1, \dots, 4$, corresponds to C1, C2, C3, A) by the same constant b ($\sigma_i = m_i b$). Thus in our model biting rate (k) and b are the parameters shared among age groups and between the PCR and microscopy datasets; m_i ($i=1, \dots, 4$) are different for each group, but the same for PCR and microscopy data for the same age group. Detection threshold for microscopy T_{micro} is fixed and equal to 40 parasites per microlitre, since this was the minimal concentration in the analysed microscopy data table, which completely conforms to the range of 20-50 parasites/microlitre reported in [2, 3]. Detection threshold for PCR T_{PCR} is a free parameter because we found a wide range of previously published PCR detection threshold values [2-4, 8-12]. Thus, we must fit only 7 parameters. For fitting, we used

NonlinearModelFit function in Wolfram Mathematica®, Wolfram Research, Inc, Champaign, IL.

The best fit parameters of the model are in the Table 1 and the graphs of the best fit infection functions are in the main part of the manuscript. The distribution of PMRs for the best fit model, and the distribution of the delays to detection are shown in the Figure 1A-C.

Part 2

Liver stage immunity model with a distribution in strength of liver stage immunity.

We can also use the a similar approach to that in supplement 1 to test whether the data would better explained by a model where age groups differ in the rate of initiation of new blood stage infections as a result of liver stage immunity. This model is similar to the previous model, except that the growth rate of parasites in blood (m) is assumed to be the same for all age groups, and age groups differ by the mean rate of initiation of new blood stage infections k_i , $i=1,\dots,4$ (having standard deviation proportional to the mean by parameter β). This liver-stage immunity model has the same number of parameters as the blood stage immunity model.

In this model we also assume that the infective mosquito bites occur continuously and independently of each other, so that the waiting time until an infective bite and therefore the time until successful initiation of blood stage will have an exponential distribution. However, in contrast to the blood stage immunity model, the rate of initiation of blood stage immunity has a distribution within each age group with a different mean for each group, while the PMR is the same in all groups.

We assume that the probability that an infection will not be blocked by liver stage immunity has a normal distribution with the standard deviation proportional to the mean (k) by constant β . In another words, the rate of initiation of blood stage infection has a normal distribution within each age group. We denote the probability density function (PDF) and the

cumulative density function (CDF) of a normal distribution with parameters k and βk by $f_N(\cdot)$ and $F_N(\cdot)$ respectively. We assume the same PMR (r) for all groups.

To account for the distribution of initiation of blood stage infection in the infection function, we need to integrate the exponential CDF with a random parameter λ (the rate of initiation of blood stage infections) and the distribution function of this parameter ($f_N(\lambda)$). Since we do not integrate from minus infinity to infinity, we have to normalize the integral, dividing it by the fraction of values of the parameter λ in the interval between 0 and some maximal λ (λ_{\max}).

$$S(t) = \begin{cases} 1 - \int_0^{\lambda_{\max}} (1 - e^{-\lambda(t-\tau-\theta(r))}) f_N(\lambda) d\lambda / (F_N(\lambda_{\max}) - F_N(0)), & t > \tau + \theta(r), \\ 1, & t \leq \tau + \theta(r), \end{cases} \quad (2.1)$$

where τ is the duration of the liver stage; $\theta(r) = 2 \log_2 T / A$ is the delay between the initiation and the detection of the blood stage as a function of PMR (can be found from equation S1.1 in Part 1). The definition and the values of all constants are the same as in the Part 1.

The model was fitted to each age group with the same β and r in all age groups, and a different value of k for each age group (we denote them by k_i , group index $i=1, \dots, 4$ corresponding to groups C1, C2, C3, A). Similar to the blood -stage immunity model, the detection threshold for microscopy T_{micro} is fixed and equal to 40 parasites per microlitre and the detection threshold for PCR T_{PCR} is a free parameter. The fitting was done using NonlinearModelFit function in Wolfram Mathematica®, Wolfram Research, Inc, Champaign, IL. The best-fit parameters are in the Table 1, section A.

We used the Akaike Information Criteria (AIC, a method that takes into account the number of parameters as well as the goodness of fit to compare models [13]) to compare the liver stage immunity model with the blood stage immunity model. The fit of the liver stage immunity model (AIC = -105.184) was significantly worse than of the blood stage (AIC = -197.7). The liver stage immunity model also cannot capture the increase in the shoulder and convexity of the

curves, because its primary effect is reduction of the rate of successful initiation of blood stage infections that leads to flattening of the slope of the infection curves. The best fit infection curves and the distribution of initiation of blood stage infections are in the Figure 1D-G, the best fit parameter of the model are in the Table 1, section A .

Part 3

Models of the impact of liver and blood stage immunity without a distribution.

General approach.

In order to explore the potential roles of liver and blood stage immunity, we also compared a number of models combining liver and blood stage immunity. Similar to the models described in detail in Part 1 and Part 2, we assume that the waiting time until an infective bite has an exponential distribution, with the parameter k as the rate of initiation of blood stage infections. The initial geometric growth of parasites in blood is described by the formula (1.1) in the Part 1. Thus after the initiation of the blood-stage infection, the detectable parasitaemia will be observed after the time interval equal to $\theta(r) = 2\log_r T / A$, where r is PMR, T is the detection threshold and A is the initial concentration of parasites in blood. Therefore the infection function will be an exponential decay function with the delay due to liver stage of infection (τ) and due to the time needed for the parasites in blood to reach the detection threshold.

$$S(t) = \begin{cases} e^{-k(t-2\log_r T/A-\tau)}, & t > 2\log_r T / A + \tau \\ 1, & t \leq 2\log_r T / A + \tau. \end{cases} \quad (3.1)$$

In order to investigate the relative contribution of liver and blood stages of immunity (described in details in the main part of the paper), we developed 3 models that are based on the formula (3.1). We fitted these models simultaneously to PCR and microscopy data assuming that the detection threshold for microscopy (T_{micro}) is 40 parasites per microlitre and the detection

threshold for PCR assay (T_{PCR}) is a free variable (see Part 1 for the explanation). The values of all constant are taken from the literature and presented in the Part 1.

Simple liver stage immunity model.

We assume the same PMR (r) for all age groups and a different rate of initiation of the blood stage infection for each age group k_i , $i=1\dots 4$. This model has 6 free parameters. The best fit parameters are in the Table 1, section B and the graphs of the best fit infection functions are in the Figure 1H-I

Simple blood stage immunity model.

We assume the same rate of initiation of the blood stage infection for all age groups (k) and different PMR for each age group r_i , $i=1\dots 4$. This model has also 6 free parameters. The best fit parameters are in the Table 1, section C and the graphs of the best fit infection functions are in the Figure 1J-K

A simple model with combined liver and blood stage immunity.

Here, we assume a different rate of initiation of the blood stage infection for each age group (k_i , $i=1\dots 4$) and a different PMR for each age group r_i , $i=1\dots 4$. This model has free 9 parameters. The best fit parameters are in the Table 1, section D and the graphs of the best fit infection functions are in the Figure 1L-M

Part 4

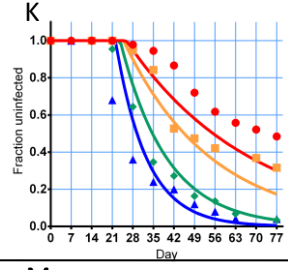
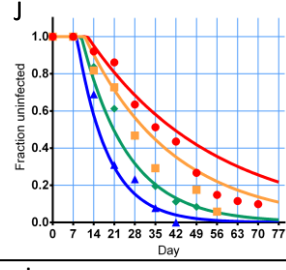
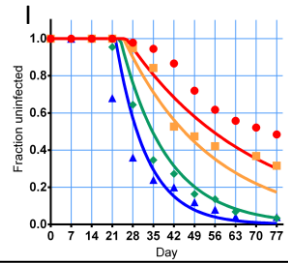
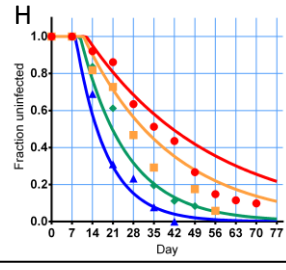
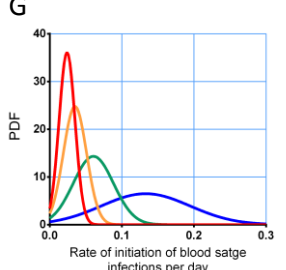
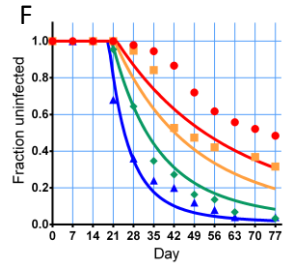
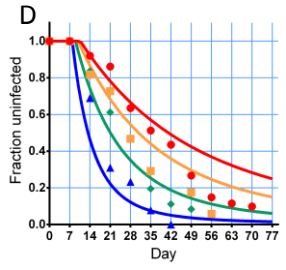
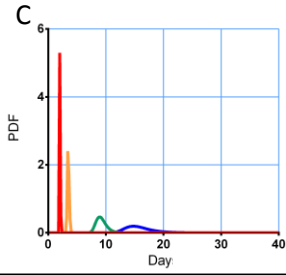
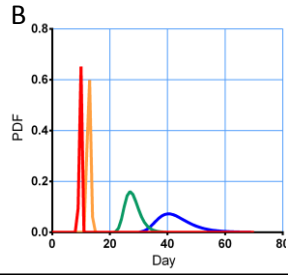
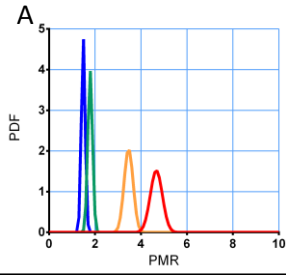
Does addition of liver stage immunity give a better fit to the model of variable blood stage immunity?

In supplements 2 and 3 we looked at various models of liver stage immunity, and whether they could accurately fit the data. In this section, we take our best-fit model of blood stage immunity (described in Part 1) and ask whether addition of liver stage immunity can provide a better fit than blood stage immunity alone. The assumption regarding the blood stage immunity of this model is the same as for the model in the Part 1. However, here we assume that in addition to the level of blood stage immunity, the level of liver stage immunity can also differ for different age groups. We can then ask the question “do we see any evidence for any effect of liver-stage immunity” that would improve our fits. The average rate of initiation of the blood stage infections (the stronger the liver stage immunity the lower rate of initiation of the blood stage infections) for the different age groups is denoted by k_i , $i=1,\dots,4$. The infection curve for the i -th age group is set by formula (4.1).

$$S(t) = \begin{cases} 1 - \int_0^{r_{\max}} (1 - e^{-k_i(\text{Max}(t-\tau-\theta(r),0))}) f_N(r, m_i, \beta m_i) dr / F_N(r_{\max}, m_i, \beta m_i), & t > \tau, \\ 1, & t \leq \tau, \end{cases} \quad (4.1)$$

where r_{\max} is maximal PMR (we took 32 parasite per cycle), m_i is the mean PMR for the i -th age group, β is the variable that relates standard deviation of the distribution of PMR to the group mean in the following way $sd_i = \beta m_i$, τ is the first possible moment of blood stage infection; $\theta(r) = 2 \log_r T / A$ is the delay between the initiation and the detection of the blood stage as a function of PMR, $f(r, m, sd)$ and $F(r, m, sd)$ are the PDF and CDF of normal distribution with mean m and standard deviation sd . The values and the meaning of all constants of the model are described in the Part 1. The graphs of the best fit infection function of the model are shown in the Figure 1N-P and the best fit distribution of PMR and the delays to detection are shown in the Figure 1Q-R, the best fit parameters of the model are in the Table 1, section E.

This model is more complex than the initial blood-stage immunity model, with three additional parameters for rate of infection in each age group. Thus, we expect a better fit to the data. However, comparing the AIC of the model in Part 1 with that shown in Table 1, section E, we see that although the inclusion of different infection rates for different age groups produces a slightly better fit compared to the blood-stage immunity model alone (AIC = -202.1 vs. -197.7 respectively), the difference in AIC (-4.4) is not at a level that would usually be considered significant. In addition, we see an unusual pattern where the estimated infection rate would actually need to be higher for the adults than the two intermediate age groups of children, which seems unexpected. These results suggest that the inclusion of age-specific differences in liver-stage immunity does not improve the fit of the model, and is unlikely to be a contributor to the observed dynamics of infection in this cohort.



- ▲ C1 (1-4 y.o),
- ◆ C2 (5-9 y.o),
- C3 (10-14 y.o)
- A (15+ y.o.)

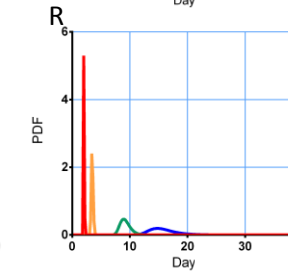
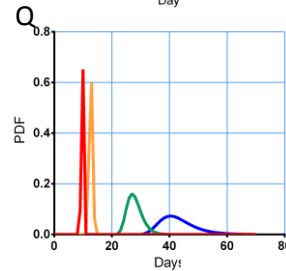
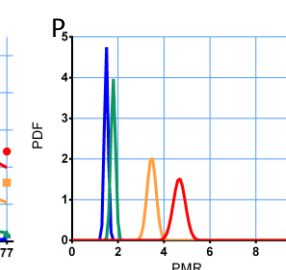
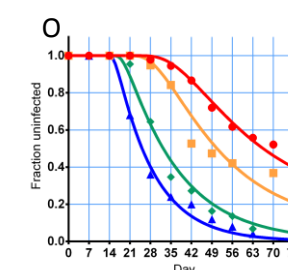
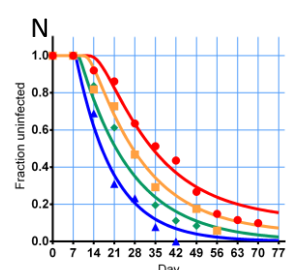
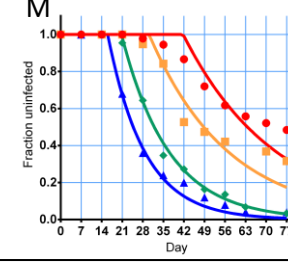
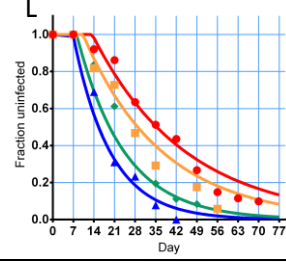


Figure 1. Comparing different models .

The figure shows the fitting of different models to the infection proportions of individuals of different ages. Age groups are indicated as; blue triangle and blue line- C1 (age 1-4 years), green diamond and green line - C2 (age 5-9 years), orange square and orange line - C3 (10-14 years), red circle and red line – A (>15 years).

Panels A-C. The distributions of PMR and delays to detection for the best-fit blood stage immunity model. (described in supplement 1 Part 1). The the best fit of the infection curves to the survival proportions of the different age groups are shown in the main text (Figure 5), as are the best fit parameters (Table 1 in the main text). Here we show the distributions of parameters for the different age groups. A. The distributions of PMR. B. The distributions of delays to detection for light microscopy at a threshold of 40 parasites/microlitre. C. The distributions of delays to detection for PCR assay at a threshold of 0.124 parasites/microliter (the best fit value from the model).

Panels D-F. The best fit infection curves of a model with a distribution in the level of liver stage immunity. (described in Supplement 1 Part 2). D. Fitting to infection curve observed in PCR data. E. Fitting to infection curve observed in microscopy data. F. The best fit distribution of the rate of initiation of blood stage infection per day.

Panels H, I. The best fit infection curves of a liver stage immunity model (without a distribution). (described in Supplement 1 Part 3). A. Fitting to infection proportions observed in PCR data. B. Fitting to infection proportions observed in microscopy data.

Panels J, K. The best fit infection curves of blood stage immunity model (without a distribution). (described in Supplement 1 Part 3). C. Fitting to infection proportions observed in PCR data. D. Fitting to infection proportions observed in microscopy data.

Panels L, M. The best fit infection curves of a combined liver and blood stages immunity model (without a distribution). (described in Supplement 1 Part 3). E. Fitting to infection proportions observed in PCR data. F. Fitting to infection proportions observed in microscopy data.

Panels N-R. The effect of adding liver-stage immunity to the best-fit model using a distribution of blood stage immunity. (described in Supplement 1 Part 3).

H. Fitting to infection proportions observed in PCR data. I. Fitting to infection proportions observed in microscopy data.

J. The distributions of PMR. K. The distributions of delays to detection for light microscopy threshold of 40 parasites/microlitre. L. The distributions of delays to detection for PCR assay at a threshold of 0.0947 parasites/microliter (the best fit threshold from the model).

A. Liver stage immunity model with a distribution in strength of immunity (AIC-105.2)				
Paramete	Estimate	Units	95 % CI	
k_1	0.1336	Blood inf./day	0.0594	0.2079
k_2	0.0605	“	0.0393	0.0817
k_3	0.0350	“	0.0258	0.0441
k_4	0.0240	“	0.0187	0.0294
r	2.9347	Inf . RBC /cycle	2.4218	3.4475
β	0.4601		0.0000	0.9740
T_{PCR}	0.0812	Par./ μ l	0.0000	0.2246
B. A simple liver stage immunity model. (AIC -115.3)				
r	2.3903	Inf. RBCs /cycle	2.0667	2.7140
k_1	0.0947	Blood inf./day	0.0663	0.1230
k_2	0.0630	“	0.0474	0.0786
k_3	0.0335	“	0.0267	0.0403
k_4	0.0234	“	0.0192	0.0277
T_{PCR}	0.0932	Par./ μ l	0.0059	0.1923
C. A simple blood stage immunity model (AIC -128.7)				
r_1	11.4341	Inf. RBCs /cycle	0	27.2632
r_2	3.3418		2.2988	4.3849
r_3	1.8103	“	1.6858	1.9349
r_4	1.4463	“	1.4032	1.4893
k	0.0516	Blood inf./day	0.0457	0.0574
T_{PCR}	0.1012	Parasites/ μ l	0.0506	0.1518
D. A simple model with combined liver and blood stage immunity (AIC -168.8)				
r_1	4.2976	Inf. RBCs /cycle	0.0645	0.0948
r_2	2.8827	“	0.0523	0.0718
r_3	1.9565	“	0.0310	0.0436
r_4	1.6015	“	0.0267	0.0366
k_1	0.0796	Blood inf./day	2.4115	6.1838
k_2	0.0621	“	2.4370	3.3283
k_3	0.0373	“	1.7356	2.1773
k_4	0.0317	“	1.5309	1.6721
T_{PCR}	0.0509	Parasites/ μ l	0.0203	0.0814
E. Liver and blood stage immunity model (AIC -202.1)				
m_1	4.6678	Inf. RBCs /cycle	2.9775	6.3581
m_2	3.4628	“	2.6114	4.3142
m_3	1.7909	“	1.5861	1.9956
m_4	1.4890	“	1.32	1.658
β	0.2426		0.186	0.2992
k_1	0.0785	Blood inf./day	0.0652	0.0918
k_2	0.0559	“	0.0468	0.065
k_3	0.0589	“	0.0348	0.083
k_4	0.0625	“	0.0189	0.106
T_{PCR}	0.0947	Parasites/ μ l	0.0377	0.1517

Table 1. The best fit parameters for the models.

References

1. Simpson JA, Aarons L, Collins WE, Jeffery GM, White NJ. Population dynamics of untreated *Plasmodium falciparum* malaria within the adult human host during the expansion phase of the infection. *Parasitology* **2002**; 124.
2. Bejon P, Andrews L, Andersen RF, et al. Calculation of liver-to-blood inocula, parasite growth rates, and preerythrocytic vaccine efficacy, from serial quantitative polymerase chain reaction studies of volunteers challenged with malaria sporozoites. *J Infect Dis* **2005**; 191:619-26.
3. Andrews L, Andersen RF, Webster D, et al. Quantitative real-time polymerase chain reaction for malaria diagnosis and its use in malaria vaccine clinical trials. *Am J Trop Med Hyg* **2005**; 73:191-8.
4. Veron V, Simon S, Carne B. Multiplex real-time PCR detection of *P. falciparum*, *P. vivax* and *P. malariae* in human blood samples. *Exp Parasitol* **2009**; 121:346-51.
5. Seckei H. Blood volume and circulation time in children. *Arch Dis Child* **1936**; 11:21-30.
6. Ezzet F, van Vugt M, Nosten F, Looareesuwan S, White NJ. Pharmacokinetics and pharmacodynamics of lumefantrine (benflumetol) in acute falciparum malaria. *Antimicrob Agents Chemother* **2000**; 44:697-704.
7. White NJ, van Vugt M, Ezzet F. Clinical pharmacokinetics and pharmacodynamics of artemether-lumefantrine. *Clin Pharmacokinet* **1999**; 37:105-25.
8. Coleman RE, Sattabongkot J, Promstaporm S, et al. Comparison of PCR and microscopy for the detection of asymptomatic malaria in a *Plasmodium falciparum/vivax* endemic area in Thailand. *Malar J* **2006**; 5:121.
9. Witney AA, Doolan DL, Anthony RM, Weiss WR, Hoffman SL, Carucci DJ. Determining liver stage parasite burden by real time quantitative PCR as a method for evaluating pre-erythrocytic malaria vaccine efficacy. *Mol Biochem Parasitol* **2001**; 118:233-45.
10. Rider MA, Byrd BD, Keating J, Wesson DM, Caillouet KA. PCR detection of malaria parasites in desiccated *Anopheles* mosquitoes is uninhibited by storage time and temperature. *Malar J* **2012**; 11:193.
11. Adam A, Witney RMA, and Daniel J. Carucci. Quantitation of Liver-Stage Parasites by Automated TaqMan[®] Real-Time PCR.
12. Taylor BJ, Martin KA, Arango E, Agudelo OM, Maestre A, Yanow SK. Real-time PCR detection of *Plasmodium* directly from whole blood and filter paper samples. *Malar J* **2011**; 10:244.
13. Motulsky HJ, Chistopoulos A. Fitting models to biological data using linear and nonlinear regression. A practical guide to curve fitting.: GraphPad Software Inc., San Diego CA, **2003**.