Modelling the population dynamics of *Plasmodium falciparum* gametocytes in humans during malaria infection

Pengxing Cao[1, *], Katharine A. Collins[2], Sophie Zaloumis[3], Thanaporn Wattanakul[4,5], Joel Tarning[4,5], Julie A. Simpson[3], James S. McCarthy[6], James M. McCaw[1,3,7, *]

1. School of Mathematics and Statistics, University of Melbourne, Melbourne, Australia.
2. Department of Medical Microbiology, Radboud University Medical Center, Nijmegen, Netherland.
3. Centre for Epidemiology and Biostatistics, Melbourne School of Population and Global Health, University of Melbourne, Melbourne, Australia
4. Mahidol-Oxford Tropical Medicine Research Unit, Faculty of Tropical Medicine, Mahidol University, Bangkok, Thailand.
5. Centre for Tropical Medicine and Global Health, Nuffield Department of Medicine, University of Oxford, Oxford, United Kingdom
6. QIMR Berghofer Medical Research Institute, Brisbane, Queensland, Australia.
7. Epidemiology, Peter Doherty Institute for Infection and Immunity, Parkville, Victoria, Australia.

* Address correspondence to Pengxing Cao (pengxing.cao@unimelb.edu.au) or James M. McCaw (jamesm@unimelb.edu.au).

Abstract

Every year over two hundred million people are infected with the malaria parasite. Renewed efforts to eliminate malaria has highlighted the potential to interrupt transmission from humans to mosquitoes which is mediated through the gametocytes. Reliable prediction of transmission requires an improved understanding of *in vivo* kinetics of gametocytes. Here we study the population dynamics of *Plasmodium falciparum* gametocytes in human hosts by establishing a framework which incorporates improved measurements of parasitaemia in humans, a novel mathematical model of gametocyte dynamics, and model validation using a Bayesian hierarchical inference method. We found that the novel mathematical model provides an excellent fit to the available clinical data from 17 volunteers infected with *P. falciparum*, and reliably predicts observed gametocyte levels. We estimated the *P. falciparum*'s sexual commitment rate and gametocyte sequestration time in humans to be 0.54% (95% credible interval: 0.30-1.00) per life cycle and 8.39 (6.54-10.59) days respectively. Furthermore, we used the data-calibrated model to predict the effects of those gametocyte dynamics parameters on human-to-mosquito transmissibility, providing a method to link within-human host kinetics of malaria infection to epidemiological-scale infection and transmission patterns.

Malaria is a mosquito-borne parasitic disease caused by protozoan parasites of the *Plasmodium* genus. It is estimated to have caused approximately 219 million new cases and 435,000 deaths (primarily due to *Plasmodium falciparum*) in 2017[1]. New tools will be required to achieve the ambitious goal of malaria elimination. Among tools proposed are novel interventions to block transmission of infection from human hosts to vector mosquitoes[2]. Falciparum malaria is transmitted from the human host to the mosquito when terminally differentiated male and female sexual stages of the parasite, called gametocytes, are taken up by the female *Anopheles* mosquito during a blood meal[3,4]. The level of gametocytes in the blood, often referred to as gametocytaemia, is highly associated with the probability of human-to-mosquito transmission[5,6]. Gametocyte levels below a certain threshold (i.e. < 1000 per mL blood) minimise the probability for a mosquito to take up both a male and female gametocyte during a blood meal, necessary to propagate infection[7]. An accurate understanding of the kinetics of gametocytes in human hosts is essential to predict the probability of transmission. A mathematical model that accurately captures the processes that give rise to observed gametocyte kinetics would be an important predictive tool to facilitate the design of effective intervention strategies.

Nonetheless, there is significant uncertainty surrounding fundamental aspects of *Plasmodium falciparum* gametocyte population dynamics in humans. Parameters such as how many gametocytes are produced during each asexual life cycle, the period of time in which early gametocytes disappear from the circulation before mature gametocytes appear (referred to as sequestration), and the period in which gametocytes circulate are poorly quantified. These gaps in understanding are due to a range of technical and logistic limitations. The first is the relatively poor sensitivity of the standard diagnostic test, namely microscopic examination of blood films. *In vivo* estimates of gametocyte kinetic parameters have been primarily based on historical data from neurosyphilis patients whose illness was treated with so-called malariotherapy[8,9]. In these studies, the limit of quantification was approximately 10^4 parasites/mL blood, at least two orders of magnitude higher than that of current quantitative PCR (qPCR) assays[10]. This high limit of quantification prevents an accurate estimation of the onset of the emergence of both asexual parasites and mature gametocytes in peripheral blood. The second limitation is that the available estimates of gametocyte population dynamics parameters based on *in vitro* cultures[11,12] may not applicable to natural infection with *P. falciparum* gametocytes due to the *in vitro* conditions that may not mimic what occurs in the human host[3].

Recent advances in experimental medicine using volunteer infection studies (VIS), otherwise known as controlled human malaria infection (CHMI) studies, allow prospective study design and data collection with the explicit aim of collecting *in vivo* data[13], in particular an improved quantification of *P. falciparum* gametocyte kinetics by qPCR applied in a novel VIS[7]. Furthermore, the models and fitting methods used in the neurosyphilis patient studies have been superseded for parameter estimation by increasingly sophisticated within-host models[14] and improvements in computational algorithms for Bayesian statistical inference[15]. Therefore, there is an emerging opportunity to improve our quantitative understanding of the population dynamics of *P. falciparum* gametocytes in human hosts by combining the novel VIS data and advanced modelling approaches.

In this paper, we developed a novel mathematical model of gametocyte dynamics, fitted the model to the VIS data and estimated the gametocyte dynamics parameters using a Bayesian hierarchical inference method (details are given in the Materials and Methods). We demonstrate that the data-calibrated model can reliably predict the time course of gametocytaemia and thus should form an essential part of modelling studies of malaria transmission.

Results

Model fitting and validation

The outcome variable used in model fitting was the total parasitaemia (total circulating asexual parasites and gametocytes per mL blood) collected from a previously published VIS[7], with other measurements from the same study, such as the asexual parasitaemia (circulating asexual parasites per mL blood) and gametocytaemia (circulating female and male gametocytes per mL blood), used to validate model predictions.

The results of fitting the mathematical model to total parasitaemia data for all 17 volunteers are given in Figure 1 (upper panels) where 12 volunteers experienced recrudescence while the rest did not. The median of posterior predictions and 95% prediction interval (PI) were computed from 5000 model simulations based on 5000 samples from the posterior parameter distribution (see Materials and Methods). The results show that the model describes the total parasitaemia

data very well (median of posterior predictions vs. observed data). Furthermore, the narrow 95% PI indicates a low level of uncertainty in predicted total parasitaemia.

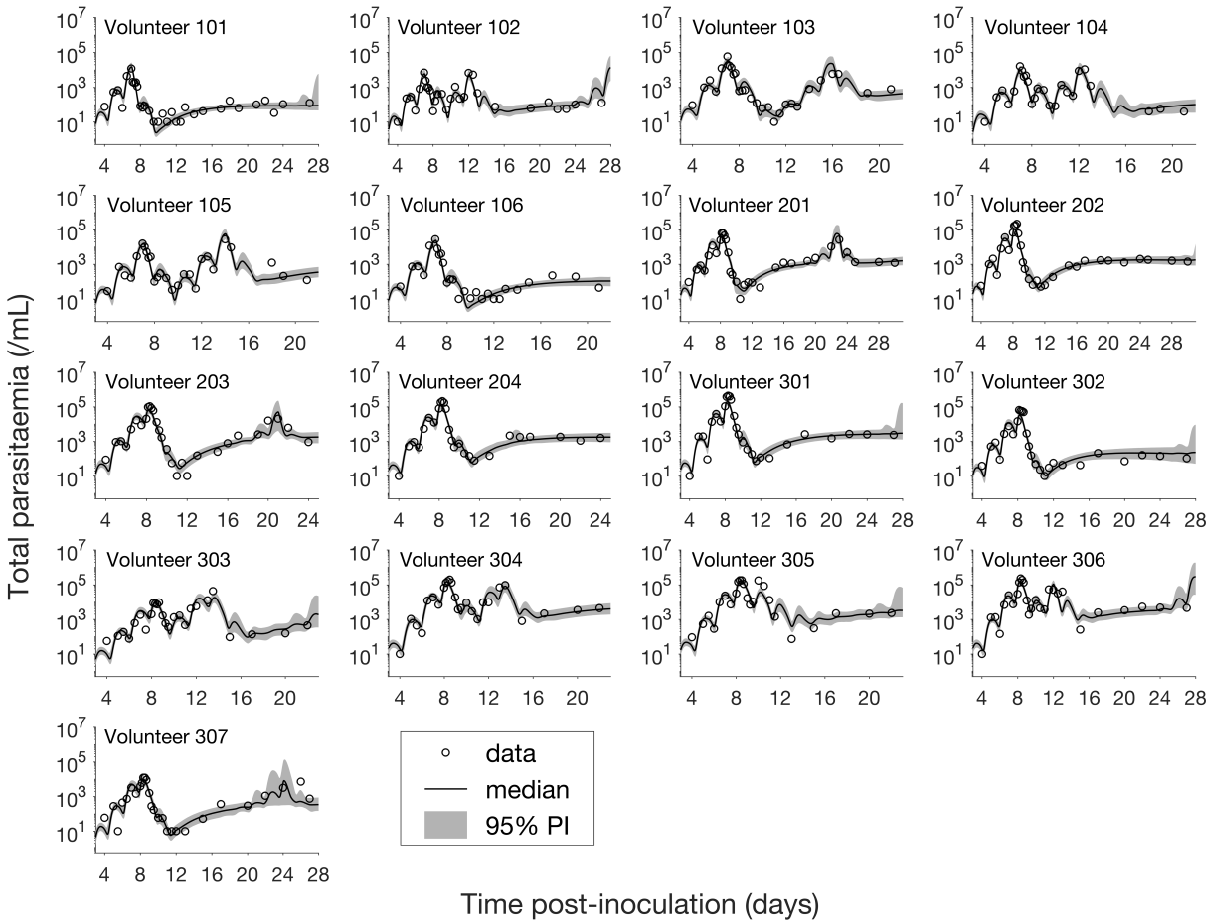


Figure 1: Results of model fitting for all 17 volunteers. Data are presented by circles. The median of posterior predictions (solid line) and 95% prediction interval (PI, shaded area) are generated by 5000 model simulations based on 5000 samples from the posterior parameter distribution (see the Materials and Methods).

Having calibrated the model against total parasitaemia, the 5000 posterior parameter sets were used to calculate the median of posterior predictions and 95% PI of the asexual parasitaemia and gametocytaemia versus time profiles. Model predictions of the asexual parasitaemia and gametocytaemia for all 17 volunteers are shown in Figure 2 and Figure 3 respectively (curves: median prediction; shaded areas: 95% PI) and are compared to the asexual parasitaemia and

gametocytaemia data (circles). We emphasise that this was not a fitting exercise, rather an independent validation of the calibrated model.

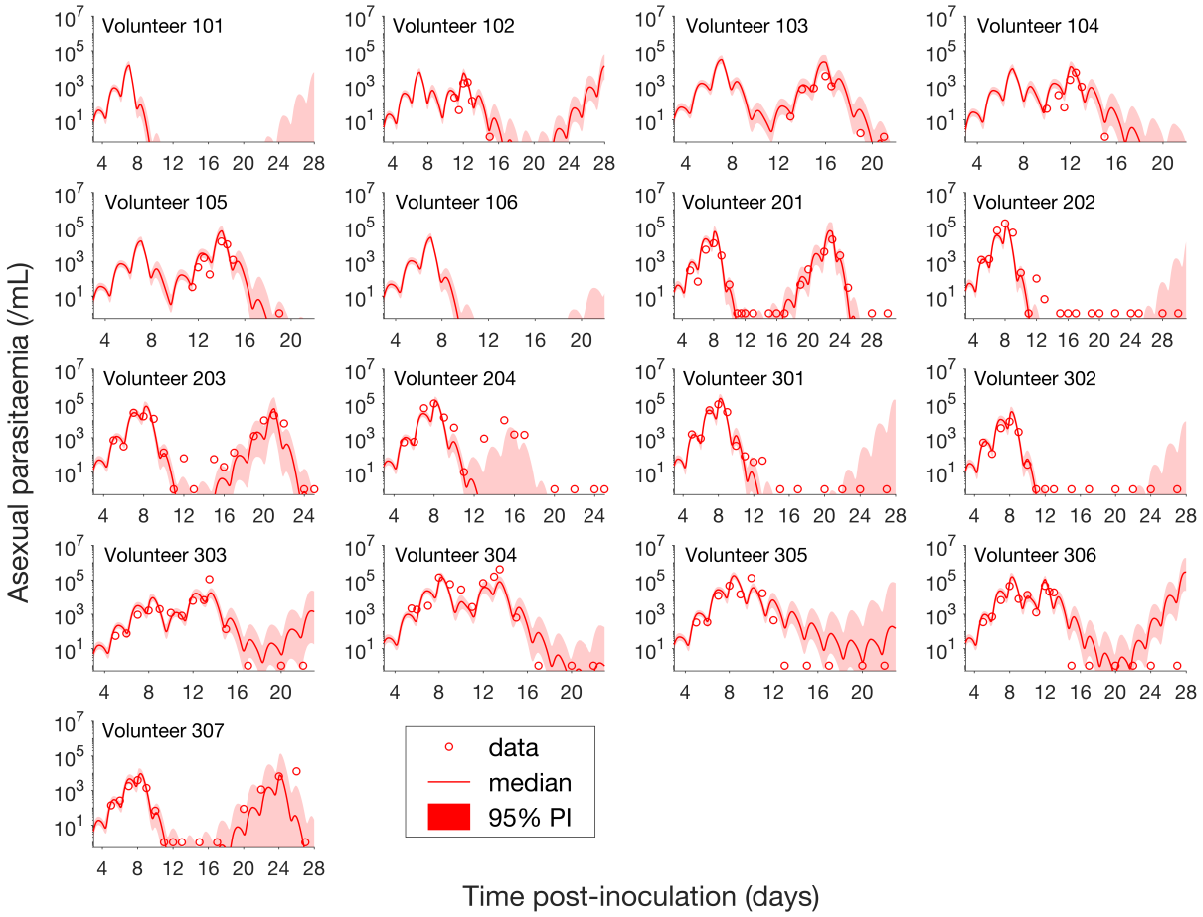


Figure 2: Comparison of model predictions and clinical data for the asexual parasitaemia for all 17 volunteers. Data are presented by circles. The median of posterior predictions (solid curve) and 95% PI (shaded area) are generated by 5000 model simulations based on 5000 samples from the posterior parameter distribution (see the Materials and Methods). Note that one asexual parasite/mL was the level of detection and no data are available for Volunteer 101 and 106 to validate the model predictions.

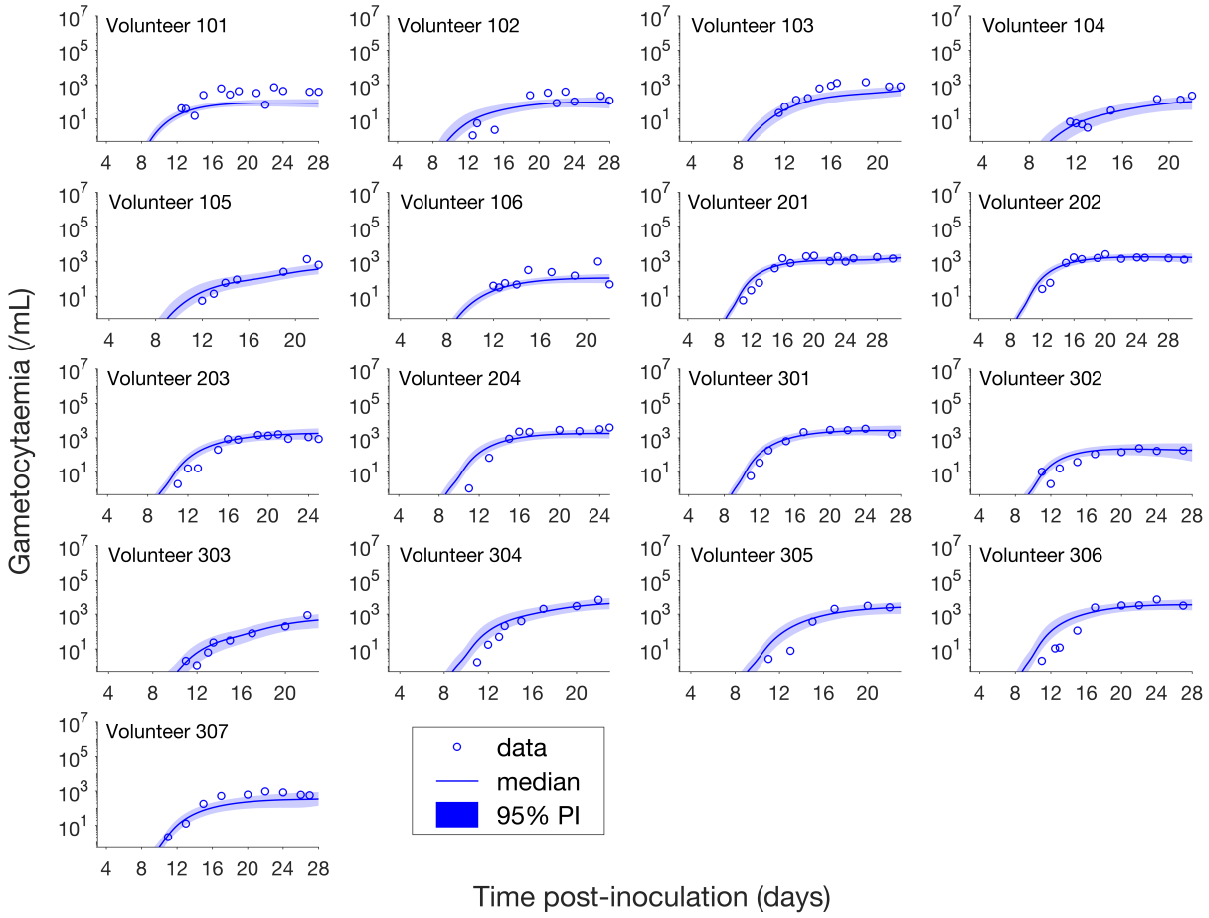


Figure 3: Comparison of model predictions and clinical data for the gametocytaemia for all 17 volunteers. Data are presented by circles. The median of posterior predictions (solid curve) and 95% PI (shaded area) are generated by 5000 model simulations based on 5000 samples from the posterior parameter distribution (see the Materials and Methods). Note that one gametocyte/mL was the level of detection.

For the majority of asexual parasitaemia data there is an excellent predictive capability for the model (see Figure 2), in particular the accurate predictions for both the recrudescent case where a portion of asexual parasitaemia data diverge from the total parasitaemia measurement (e.g. Volunteer 103, 104, 105, 201, 203, 304 and 307) and the non-recrudescent case where the posterior-median prediction curve (solid red curve) lies below the limit of detection (one asexual parasite/mL) (e.g. Volunteer 202, 301 and 302). However, there are some discrepant observations. The model under-predicts (see Volunteer 204) or over-predicts (see Volunteer 303, 305 and 306) a portion of the asexual parasitaemia data. Furthermore, for some volunteers such

133 as 202, 301 and 302, the 95% PI (red shaded area) extends into the detectable region again after
134 the asexual parasitaemia reaches below the detection limit, indicating that there was a small
135 chance that the patients may have suffered a recrudescence during the observation period (of
136 course, they did not) or after the observation period (although this predication cannot be
137 evaluated because artemisinin combination therapy was given immediately after the period).
138 Despite the above discrepant observations for asexual parasitaemia, the model predictions of
139 gametocytaemia are very consistent with the data for all 17 volunteers, as shown in Figure 3.

140 Estimation of gametocyte population dynamics parameters

141 The model calibration process provided the joint posterior density for the model parameters,
142 which were used to estimate some key biological parameters governing the population dynamics
143 of *P. falciparum* gametocytes (see the Material and Methods for details). As shown in Table 1,
144 the sexual commitment rate – the percentage of asexual parasites entering sexual development
145 during each asexual life cycle – is estimated to be 0.54% (95% credible interval (CI): 0.30%,
146 1.00%). This *in vivo* estimate is much lower than 11% that was estimated from *in vitro* data[11].
147 The gametocyte sequestration time is the average time that immature gametocytes (stages I—IV)
148 cannot be observed in the peripheral circulation and was estimated to be 8.39 days (95% CI:
149 6.54, 10.59 days). The estimate for the circulating gametocyte lifespan is 63.5 days, with a broad
150 95% CI (12.7 to 1513.9 days) resulting from the long-tailed posterior distribution (see Figure S1,
151 Supplementary Information) and is much longer than the previous *in vitro* estimate of 16—32
152 days[12] (note that our lower bound of the 95% CI is lower than the *in vitro* estimated range). The
153 wide estimate for the circulating gametocyte lifespan, and in particular the high upper bound of
154 the 95% CI, is primarily due to the limited observation time in the VIS as explored in more detail
155 in the Discussion.

156 **Table 1: Estimates of some key biological parameters and comparison with the literature.**
157 The estimates of the biological parameters (middle column) are shown as the median and 95%
158 credible interval (CI) of the marginal posterior parameter distribution (see Figure S1 in the
159 Supplementary Information). Estimates from the literature (third column) are shown in the
160 format of either "mean estimate (95% confidence interval)" or "mean estimate [minimum –

161 maximum estimate]" or simply "a low estimate – a high estimate". Note some quoted estimates
162 are from studies of *P. falciparum* (*Pf*) from either *in vivo* or *in vitro* studies.

Biological parameters (unit)	**Median estimate (95%CI)**	**Estimates in the literature**
Sexual commitment rate (%/life cycle)	0.54 (0.30, 1.00)	11 (6.2, 15.8)[11] (*Pf; in vitro*) 0.64 [0.03 – 13.51][9] (*Pf; in vivo*)
Gametocyte sequestration time (days)	8.39 (6.54, 10.59)	7.4 [4 – 12][9] (*Pf; in vivo*)
Circulating gametocyte lifespan (days)	63.5 (12.7, 1513.9)	16 – 32[12] (*Pf; in vitro*) 6.4 [1.3 – 22.2][9] (*Pf; in vivo*)
Parasite multiplication factor (/life cycle)	21.8 (17.6, 26.9)	10 – 33[16] (*Pf; in vivo*) 16.4 (15.1, 17.8)[17] (*Pf; in vivo*)

163

164 As shown in Table 1, there are both similarities and differences in parameter estimates for *P.*
165 *falciparum* based on our analysis and the analysis of historical neurosyphilis patient data[9]. We
166 found that they exhibited very similar *in vivo* sexual commitment rate (median 0.54% vs. mean
167 0.64% with overlapping plausible ranges) and gametocyte sequestration time (median 8.39 days
168 vs. mean 7.4 days with overlapping plausible ranges). On the other hand, we found that the
169 circulating gametocyte lifespan for *P. falciparum* was much longer than that estimated from the
170 neurosyphilis patient data[9] (median 63.5 days vs. 6.4 days with very different ranges).

171 Finally, we provided an estimate for the parasite multiplication factor, an important parameter
172 that quantifies the *net* replication rate of asexual parasites and also influences the rate of
173 gametocyte generation. Our posterior-median estimate is 21.8 parasites per life cycle (95% CI:
174 17.6, 26.9), consistent with previous estimates which lie in the range 10—33[16], and a bit larger
175 than an updated estimate calculated from a pooled analysis of parasite counts from 177
176 volunteers infected with the same *P. falciparum* strain (16.4 parasites per life cycle)[17].

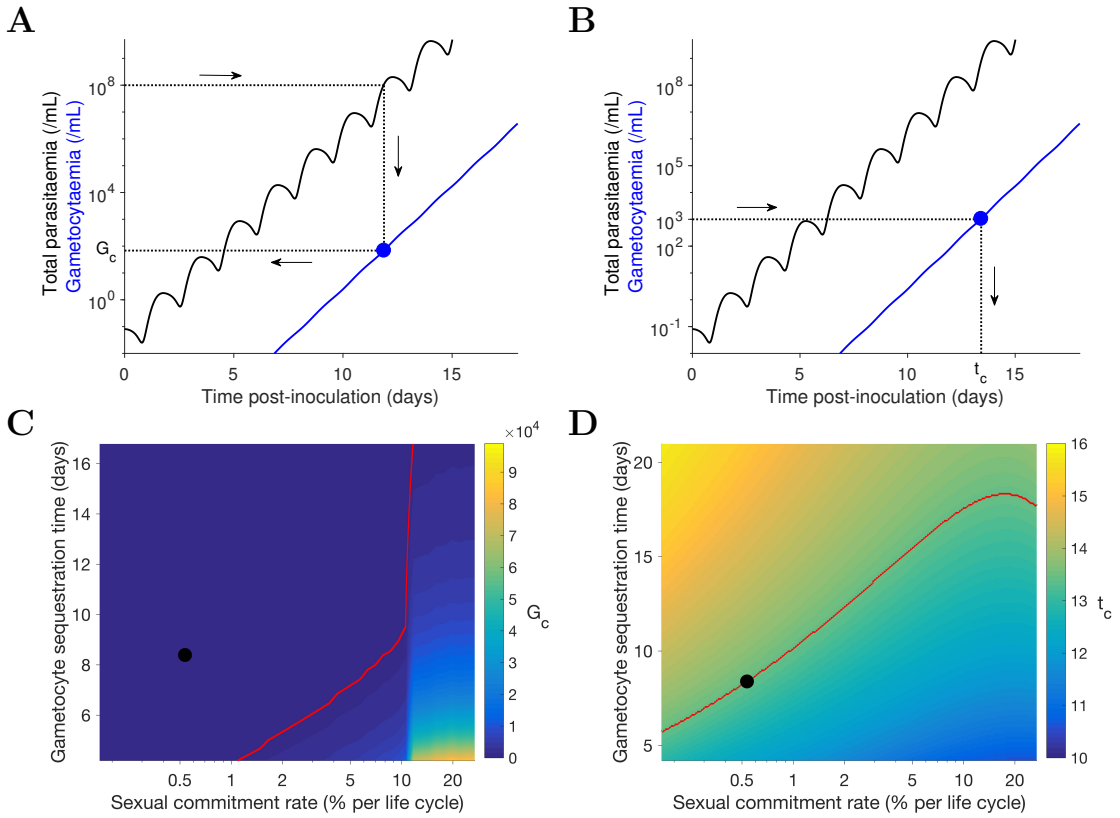
Predicting the impact of gametocyte kinetics on human-to-mosquito transmissibility

Having validated our mathematical model of asexual parasitaemia and gametocyte dynamics, we were able to predict how the gametocyte dynamics parameters influence the transmissibility of *P. falciparum* malaria from humans to mosquitoes in various epidemiological scenarios. In particular, we focused on the early phase of infection where the innate immune response is minimal and treatment has not been administered (in order to avoid complications that our mathematical model was not designed to capture). Two scenarios were considered:

- Predicting the potential infectiousness of newly hospitalised malaria patients for various values of sexual commitment rate and gametocyte sequestration time. In the model, gametocytaemia was assumed to be a surrogate of the potential infectiousness and newly hospitalised patients were assumed to exhibit a total parasitaemia of approximately 10^8 parasites/mL based on published clinical data from Cambodia and Thailand[18]. As illustrated in Figure 4A, we simulated the model (for different sexual commitment rates and gametocyte sequestration times) and looked at the critical gametocytaemia level (indicated by G_c) corresponding to the time when the total parasitaemia (wave-like black curves) first reached 10^8 parasites/mL (their associations are indicated by the dotted lines and arrows).

- Predicting the non-infectious period of malaria patients for various values of sexual commitment rate and gametocyte sequestration time. In the model, the non-infectious period was defined to be time from the inoculation of infected red blood cells to the time when the gametocytaemia reach 10^3 parasites/mL, which is a threshold allowing for human-to-mosquito transmission[7]. Note that this non-infectious period does not include the latent period due to the liver stage which should be considered if the starting time were taken as the inoculation by mosquito. As illustrated in Figure 4B, we simulated the model (for different sexual commitment rates and gametocyte sequestration times) and identified the critical time (indicated by t_c) when the gametocytaemia (blue curve) first reached 10^3

203 parasites/mL (their associations are indicated by the dotted lines and arrows).



204

205 **Figure 4: Simulation of two scenarios predicting the dependence of human-to-mosquito**
206 **transmissibility on the sexual commitment rate and gametocyte sequestration**
207 **time.** (A) illustration of the first scenario: predicting the critical gametocytaemia level
208 (indicated by G_c) at the time when the total parasitaemia reach 10^8 parasites/mL. (B)
209 illustration of the second scenario: predicting the non-infectious period (indicated by
210 t_c), which is defined to be time from inoculation of infected red blood cells to the time
211 when the gametocytaemia reach 10^3 parasites/mL (a threshold for human-to-mosquito
212 transmission[7]). (C and D) Heatmaps showing the dependence of the critical
213 gametocytaemia G_c and the non-infectious period t_c on the sexual commitment rate
214 and gametocyte sequestration time. The black dots represent the benchmark value
215 obtained by simulating the gametocyte population dynamics model using the median
216 estimates of the posterior samples of the population mean parameters (see the

Materials and Methods for details). The red curves indicate the cases corresponding to gametocytaemia of 10^3 parasites/mL.

Figure 4C shows that a higher sexual commitment rate or a lower gametocyte sequestration time leads to a higher gametocytaemia (G_c) at the time of hospitalisation. The red curve in Figure 3C indicates the level curve of 10^3 gametocytes/mL (i.e. the threshold for infectiousness as mentioned above) dividing the heatmap into two regions. To the left, G_c is below 10^3 gametocytes/mL, suggesting clinical presentation precedes infectiousness, while to the right G_c is above 10^3 gametocytes/mL and the converse applies. The G_c value obtained by model simulation using the median estimates of the population mean parameters (indicated by the black dot) is below 10^3 gametocytes/mL, suggesting that newly hospitalized malaria patients are less likely to be infectious, and thus efforts to identify and treat infections in a timely manner may have a substantial impact in terms of reduced transmission potential.

Figure 4D reinforces the result in Figure 4C using the non-infectious period (t_c). As the sexual commitment rate increases or the gametocyte sequestration time decreases, t_c decreases. However, for relatively large values of the sexual commitment rate (e.g. > 20%), we observed an increase in t_c as the sexual commitment rate increases (see the top-right corner of Figure 4D). This is because an increased fraction of sexual conversion can lead to both a decrease in the rate of asexual parasite growth and an increase in the number of sexual parasites, and the impact of the former more than counterbalances that of the latter.

Discussion

We have developed a novel mathematical model of gametocyte dynamics, that integrates both multi-state asexual parasite's life cycle and the development of gametocytes for the first time. Model parameters were estimated by fitting the model to data from 17 malaria-naïve volunteers inoculated with *P. falciparum*-infected red blood cells (3D7 strain). Compared to previous studies, our work is distinguished by three novel contributions: (1) the use of a prospectively planned clinical trial to collect more accurate quantitative data of parasite levels measured by qPCR; (2) the development of a novel population dynamics mathematical model which allows for robust and biologically-informed extrapolation and hypothesis testing/scenario analysis; and

245 (3) the use of a Bayesian hierarchical inference method for model calibration and parameter
246 estimation.

247 For gametocyte kinetic parameters, we found that our *in vivo* estimate of the *P. falciparum*
248 sexual commitment rate was similar to that found in the neurosyphilis patient data[9] but was much
249 smaller than previous *in vitro* estimates (see Table 1). Novel VIS data using biomarkers specific
250 to early sexual parasites (e.g. PfGEXP5[19]) may help to validate and improve our *in vivo* estimate.
251 Furthermore, our *in vivo* estimate of the circulating gametocyte lifespan is substantially larger
252 than previous *in vitro* estimates and is also larger than that found in the neurosyphilis patient
253 data[9] (see Table 1). One way to improve our *in vivo* estimate would be to use new *P. falciparum*
254 data with gametocytaemia measurements over a longer period of time to capture the natural
255 decay of circulating gametocytes, which was not observed in the current VIS data (see Figure 3).

256 We also predicted the effects of altered gametocyte kinetic parameters on the transmissibility
257 from human to mosquito, focusing on two scenarios: the infectiousness of newly hospitalised
258 malaria patients (i.e. the gametocytaemia when asexual parasitaemia first reaches a hospitalised
259 level of 10^8 parasites/mL in the model); and the non-infectious period of malaria patients (i.e. the
260 time from the inoculation of infected red blood cells to the time when the gametocytaemia reach
261 a transmission threshold of 10^3 parasites/mL in the model). We explored how the sexual
262 commitment rate and gametocyte sequestration time influenced the gametocyte level and the
263 non-infectious period. We would like to emphasize that human-to-mosquito transmissibility is
264 determined by both the gametocytaemia level in humans and the relationship between
265 gametocytaemia and the probability of transmission per bite. A reliable prediction of the former
266 is essential but not a sole determinant of transmissibility. Therefore, it is also important to
267 improved our quantitative understanding of the latter, which could be complicated and also
268 determined by the densities of female and male gametocytes[5,6,20].

269 Our study has some limitations. The gametocyte population dynamics model, that has been
270 shown to have sufficient complexity to reproduce the clinical observations, is still a rather coarse
271 simplification of the actual biological processes. For example, the model does not assume an
272 adaptive sexual commitment rate[21], nor does it consider the mechanisms of sexual
273 commitment[22]. Furthermore, the model assumes a constant gametocyte death rate but does not
274 consider other non-constant formulations as have been previously proposed[8]. Another limitation

275 is that we assumed a fixed duration for the asexual life cycle of 42 hours, while previous work by
276 our group suggests that the life cycle may be altered by up to a few hours in response to
277 antimalarial drugs (e.g. artemisinin[23,24]), although it is not known if piperaquine (which was
278 administered in this VIS) has a similar effect.

279 In conclusion, we have developed a novel mathematical model of gametocyte population
280 dynamics, and demonstrated that it reliably predicts time series data of gametocytaemia. The
281 model provides a powerful predictive tool for informing the design of future volunteer infection
282 studies which aim to test transmission-blocking interventions. Furthermore, the within human
283 host transmission model can be incorporated into epidemiological-scale models to refine
284 predictions of the impacts of various antimalarial treatments and transmission interventions.

285 Materials and Methods

286 Study population and measurements

287 The data used in this modelling study are from a previously published VIS[7] where 17 malaria-
288 naïve volunteers were inoculated with *P. falciparum*-infected red blood cells (3D7 strain). The
289 study was approved by the QIMR Berghofer Human Research Ethics Committee and registered
290 with ClinicalTrials.gov (NCT02431637 and NCT02431650). The volunteers were treated with
291 480 mg piperaquine phosphate (PQP) on day 7 or 8 post-inoculation to attenuate asexual parasite
292 growth and a second dose of 960 mg PQP was given to any volunteer for treatment of
293 recrudescent asexual parasitaemia. All participants received a course of artemether/lumefantrine
294 and, if required, a single dose of primaquine (45 mg) to clear all parasites at the end of the study.
295 Parasitaemia in the volunteers were monitored at least daily since the inoculation.

296 The data contain the measurements of total parasitaemia (total circulating asexual parasites and
297 gametocytes per mL blood), asexual parasitaemia (circulating asexual parasites per mL blood),
298 gametocytaemia (circulating female and male gametocytes per mL blood) and plasma
299 concentration of PQP, at multiple time points after inoculation (see Figure 1—3 and Figure S2 in
300 the Supplementary Information). Further details about the VIS are given in[7]. We emphasize that
301 the data used in model fitting is the total parasitaemia (from the first measurement to the time

before any treatment other than PQP) and the other data (i.e. asexual parasitaemia and gametocytaemia) are used to validate the model.

Gametocyte dynamics model

The model extends the published models of asexual parasite life cycle[18,25] by incorporating the development of gametocytes. The model is comprised of three parts describing three populations of parasites: asexual parasites (P), sexual parasites (P_G) and gametocytes (G). Note that biologically speaking the phrase "sexual parasites" includes gametocytes, but in order to model the sexual parasites which are indistinguishable from asexual parasites under a microscope, we treat those sexual parasites and gametocytes as two different populations. A schematic diagram of the development of those populations based on current knowledge[11,22] is shown in Figure 5.

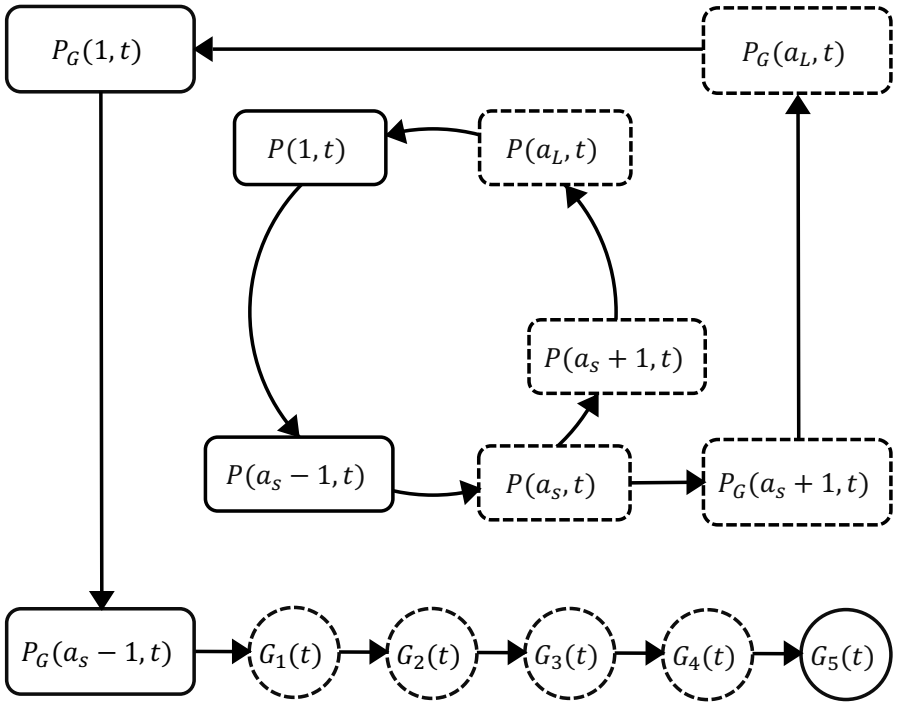


Figure 5: Schematic diagram showing the model compartments and transitions. Square compartments in the inner loop represent the asexual parasite population which follows a life cycle of maturation and replication every a_L hours. Sexual commitment occurs from age a_s and a fraction of asexual parasites become sexual (the bigger square compartments in the outer loop) and eventually enter the development of stage

I—V gametocytes (the round compartments). The compartments with a dashed boundary are sequestered to tissues and thus not measurable in a blood smear. The notation of each compartment is consistent with those in model equations and is explained in the main text.

Asexual parasites develop and replicate in the red blood cells (RBCs) till cell rupture at the end of each life cycle and the released free parasites (merozoites) can initiate new life cycles if successfully invading susceptible RBCs. At the time of inoculation (i.e. $t = 0$ hours in the model), we define the inoculum size to be P_{init} and assume the age distribution of inoculated parasites is Gaussian with mean μ and standard deviation σ. As time increments by one hour, the asexual parasites of age a at time t (denoted as $P(a,t)$) follow the iterative equation:

$$P(a,t) = \begin{cases} P(a-1,t-1)e^{-\overline{k_d}-\delta_P}, & a = 2,3,\dots,a_L \\ r_P P(a_L,t-1)e^{-\overline{k_d}-\delta_P}, & a = 1 \end{cases} \quad (1)$$

where $\overline{k_d}$ represents the average rate of asexual parasite killing by PQP and δ_P is the rate of asexual parasite death due to processes other than PQP. $\overline{k_d}$ is approximated by the average of $k_d(t-1)$ and $k_d(t)$ and $k_d(t) = k_{max}C(t)^\gamma/(C(t)^\gamma + EC_{50}{}^\gamma)$ where k_{max} is the maximum killing rate and EC_{50} is the PQP concentration at which half maximum killing is achieved. $C(t)$ is the PQP concentration which is simulated by a pharmacokinetic model introduced below. a_L is the length of each life cycle and r_P is the parasite replication rate indicating the average number of newly infected RBCs attributable to the rupture of a single infected RBC at the end of its life cycle. Sexual commitment is assumed to occur at the first age of the trophozoite stage (denoted to be a_s) and a percentage (f) of asexual parasites leave asexual life cycle and start sexual development in the next hour, which is modelled by

$$P(a_s+1,t) = (1-f)P(a_s,t-1)e^{-\overline{k_d}-\delta_P}, \quad (2)$$

$$P_G(a_s+1,t) = fP(a_s,t-1)e^{-\overline{k_d}-\delta_P}. \quad (3)$$

The first equation describes the proportion of parasites remaining in the asexual life cycle while the second equation describes the proportion of parasites becoming sexual parasites (P_G). According to Figure 5, the sexual parasites continue the rest of the life cycle and a part of the next life cycle (note that they appear indistinguishable from asexual parasites under microscopy) before becoming stage I gametocytes. The process is modelled by

$$P_G(a,t) = \begin{cases} P_G(a-1,t-1)e^{-\delta_P}, & a = 2,3,\dots,a_L \text{ except } a = a_s \text{ and } a = a_s+1 \\ r_P P_G(a_L,t-1)e^{-\delta_P}, & a = 1 \end{cases} \quad (4)$$

Note that no $\overline{k_d}$ for P_G because PQP does not kill sexual parasites and gametocytes. The changes of the sequestered stage I—IV gametocytes (G_1—G_4) are governed by difference equations

$$G_1(t) = G_1(t-1)e^{-(m+\delta_G)} + \frac{P_G(a_s-1,t-1)e^{-\delta_P}\left(1-e^{-(m+\delta_G)}\right)}{m+\delta_G}, \quad (5)$$

$$G_2(t) = G_2(t-1)e^{-(m+\delta_G)} + \frac{mG_1(t-1)\left(1-e^{-(m+\delta_G)}\right)}{m+\delta_G}, \quad (6)$$

$$G_3(t) = G_3(t-1)e^{-(m+\delta_G)} + \frac{mG_2(t-1)\left(1-e^{-(m+\delta_G)}\right)}{m+\delta_G}, \quad (7)$$

$$G_4(t) = G_4(t-1)e^{-(m+\delta_G)} + \frac{mG_3(t-1)\left(1-e^{-(m+\delta_G)}\right)}{m+\delta_G}, \quad (8)$$

where m is the rate of gametocyte maturation and δ_G is the death rate of sequestered gametocytes. Stage V gametocytes are circulating in bloodstream (and therefore can be measured from the peripheral blood film) modelled by

$$G_5(t) = G_5(t-1)e^{-\delta_{Gm}} + \frac{mG_4(t-1)\left(1-e^{-\delta_{Gm}}\right)}{\delta_{Gm}}, \quad (9)$$

where δ_{Gm} is the death rate of mature circulating gametocytes.

The total parasitaemia in the model is given by $\sum_{a=1}^{a=a_s-1}[P(a,t)+P_G(a,t)]+G_5(t)$, which was fitted to the VIS data. After model fitting, we simulated the asexual parasitaemia $\sum_{a=1}^{a=a_s-1}P(a,t)$ and gametocytaemia $G_5(t)$ and compared them with associated data for model validation. Table 2 presents all the model parameters and their units and descriptions.

Table 2: Details of the gametocyte dynamics model parameters. The parameter ranges of the prior distributions are chosen based on previous estimates[25] and biological plausibility. We assumed parasites younger than 25h are circulating and thus fix a_s to be 25h. For 3D7 strain, the life cycle is approximately 39—45h (based on *in vitro* estimates[26] and *in vivo* estimates by Wockner et al.[17]) and we fix a_L to be 42h. The boundaries of the uniform prior distribution (U) also indicate the lower and upper bounds of the corresponding parameters.

Parameter	Unit	Description	Prior distribution

P_{init}	parasites/mL	inoculation size	U(0, 10)
μ	h	mean of the initial parasite age distribution	U(0, 35)
σ	h	SD of the initial parasite age distribution	U(0, 20)
r_P	(unitless)	parasite replication rate	U(0, 100)
k_{max}	h^{-1}	maximum rate of parasite killing by PQP	U(0, 1)
EC_{50}	ng/mL	half-maximum effective PQP concentration	U(1, 100)
γ	(unitless)	Hill coefficient for PQP	U(0, 20)
f	(unitless)	sexual commitment rate (in percentage)	U(0, 1)
δ_P	h^{-1}	death rate of asexual and sexual parasites	U(0, 0.2)
m	h^{-1}	maturation rate of gametocytes	U(0, 0.1)
δ_G	h^{-1}	death rate of sequestered gametocytes	U(0, 0.1)
δ_{Gm}	h^{-1}	death rate of circulating gametocytes	U(0, 0.1)
a_s	h	sequestration age of asexual parasites	fixed to be 25
a_L	h	length of life cycle of asexual parasites	fixed to be 42

SD: standard deviation; h: hour.

Pharmacokinetic model of piperaquine (PQP)

In the within-host model, the killing rate $k_d(t)$ is determined by PQP concentration $C(t)$ which was simulated from a pharmacokinetic (PK) model introduced in this section. The PK model, provided by Thanaporn Wattanakul and Joel Tarning (Mahidol-Oxford Tropical Medicine Research Unit, Bangkok), is a three-compartment disposition model with two transit compartments for absorption (see schematic diagram in Figure S3, Supplementary Information). Based on Figure S3, the model is formulated to be a system of ordinary differential equations:

$$\frac{dD}{dt} = -k_T D, \tag{10}$$

$$\frac{dT_1}{dt} = k_T D - k_T T_1, \tag{11}$$

$$\frac{dT_2}{dt} = k_T T_1 - k_T T_2, \tag{12}$$

$$\frac{dC}{dt} = \frac{k_T T_2 + q_1 P_1 + q_2 P_2 - q_1 C - q_2 C - q_c C}{V_c}, \tag{13}$$

$$\frac{dP_1}{dt} = \frac{q_1 C - q_1 P_1}{V_1}\ , \tag{14}$$

$$\frac{dP_2}{dt} = \frac{q_2 C - q_2 P_2}{V_2}\ , \tag{15}$$

The model parameters were obtained by fitting the model output $C(t)$ to the PQP concentration data from the VIS in MATLAB (version 2016b; The MathWorks, Natick, MA) using *lsqcurvefit*. The MATLAB code (with detailed comments) are publicly available at https://doi.org/10.26188/5cde4c26c8201. Note that for volunteers whose number of data points is less than the number of parameters in the PK model (such that optimization fails), the parameter values are given by the estimates from another VIS (provided by Thanaporn Wattanakul and Joel Tarning). The best-fit curve and associated parameter values for all volunteers are provided in Figure S2 and Table S1 in the Supplementary Information.

Fitting the model to parasitaemia data

We took a Bayesian hierarchical modelling approach[27] to fit the gametocyte dynamics model to the data from all 17 volunteers. In detail, each volunteer has 12 model parameters (i.e. those in Table 2 except a_s and a_L; also called the individual parameters) and lower and upper bounds of the parameters are given in Table 2. If denoting the individual parameters to be θ_{ind} and lower and upper bounds to be b_L and b_U respectively, the following transformations are used to convert the bounded individual parameters to unbounded ones (denoted by φ_{ind}) in order to in order to improve computational efficiency[28,29]:

$$\varphi_{ind} = \ln\left(\frac{\theta_{ind} - b_L}{b_U - \theta_{ind}}\right), \tag{16}$$

φ_{ind} obeys a multivariate normal distribution $\mathcal{N}(\varphi_{pop}, \Omega_{pop})$ where

$$\varphi_{pop} = \ln\left(\frac{\theta_{pop} - b_L}{b_U - \theta_{pop}}\right) \tag{17}$$

and θ_{pop} is a vector containing 12 population mean parameters (hyperparameters) corresponding to the 12 gametocyte dynamics model parameters. Ω_{pop} is the covariance matrix. For more efficient sampling process, $\varphi_{ind} \sim \mathcal{N}(\varphi_{pop}, \Omega_{pop})$ was reparameterised to a non-centred form $\varphi_{ind} = \varphi_{pop} + \omega_{pop} L\eta$, where ω_{pop} is the diagonal standard deviation (SD) matrix whose diagonal elements are the 12 population SD parameters (hyperparameters); L is the lower

Cholesky factor of the correlation matrix; η obeys standard multivariate normal distribution. Note that $\Omega_{pop} = \omega_{pop} L L^T \omega_{pop}$ where LL^T is the correlation matrix. The prior distributions for the 12 population mean parameters θ_{pop} are given by uniform distributions with bounds given in Table 2. The prior distribution for the 12 population SD parameters is standard half-normal and the prior distribution for the lower Cholesky factor of the correlation matrix L is given by Cholesky LKJ correlation distribution with shape parameter of 2[29,30]. The distribution of the observed parasitaemia measurements is assumed to be a lognormal distribution with mean given by the model-simulated values and SD parameter with prior distribution of a half-Cauchy distribution with a location parameter of zero and a scale parameter of 5. The distribution for the observed parasitaemia measurements was used to calculate the likelihood function and the M3 method[31] was used to penalise the likelihood for data points below the limit of detection for the total parasitaemia (10 parasites/mL[7]).

Model fitting was implemented in R (version 3.2.3)[32] and Stan (RStan 2.17.3)[29] using the Hamiltonian Monte Carlo (HMC) optimized by the No-U-Turn Sampler (NUTS) to draw samples from the joint posterior distribution of the parameters including the individual parameters (12 parameters for each volunteers) and population mean parameters (12 hyperparameters). Five chains with different starting points (set by different random seeds) were implemented and 1000 posterior samples retained from each chain after a burn-in of 1000 iterations (in total 5000 samples were drawn from the joint posterior distribution). The marginal posterior and prior distributions of the population mean and SD parameters are shown in Figure S4 and S5 in the Supplementary Information. The marginal posterior distributions of the individual parameters for all 17 volunteers are shown in Figure S6—S17 (using violin plots) in the Supplementary Information. For each volunteer, the 5000 sets of individual parameters are used to simulate the gametocyte population dynamics model and generate 5000 simulated model outputs (e.g. 5000 time series of total parasitaemia, asexual parasitaemia or gametocytaemia). The posterior prediction and 95% prediction interval (PI) are given by the median and quantiles of 2.5% and 97.5% of the 5000 model outputs at each time respectively (see Figure 1—3 for example).

The estimates of some key biological parameters (Table 1) were calculated using the 5000 posterior draws of the 12 population mean parameters, i.e. median and 2.5%- and 97.5%-quantile (95% credible interval). The sexual commitment rate was calculated directly by f_{pop} (the

437 population mean parameter for f). Sequestered and circulating gametocyte lifespan were
438 calculated by $1/\delta_{G_{pop}}/24$ and $1/\delta_{Gm_{pop}}/24$ respectively (the factor of 24 converts hours into
439 days). Gametocyte sequestration time was calculated by $4/m_{pop}/24$ where 4 indicates four
440 sequestered state (stage I to IV) and 24 converts hours into days. Parasite multiplication factor is
441 calculated by $r_{P_{pop}}\exp(-\delta_{P_{pop}}a_L)(1-f_{pop})$ where the term $\exp(-\delta_{P_{pop}}a_L)(1-f_{pop})$ gives
442 the fraction of surviving asexual parasites after death and sexual conversion per life cycle.
443 The gametocyte dynamics model with parameters given by the median estimates of the
444 population mean parameters was used to simulate the two scenarios predicting the dependence of
445 human-to-mosquito transmissibility on the sexual commitment rate and gametocyte sequestration
446 time (Figure 4).
447 Final analysis and visualization were performed in MATLAB. All computer codes (R, Stan,
448 MATLAB), data and fitting results (CSV and MAT files) and an instruction document (note that
449 reading the document first will make the code much easy to follow) are publicly available at
450 https://doi.org/10.26188/5cde4c26c8201.

451 # References

452 1. The World Health Organization. *World Malaria Report 2018*. (2018).
453 2. malERA: An updated research agenda for diagnostics, drugs, vaccines, and vector control
454 in malaria elimination and eradication. *PLoS Med.* **14**, e1002455 (2017).
455 3. Bousema, T. & Drakeley, C. Epidemiology and infectivity of Plasmodium falciparum and
456 Plasmodium vivax gametocytes in relation to malaria control and elimination. *Clin.*
457 *Microbiol. Rev.* **24**, 377–410 (2011).
458 4. Josling, G. A. & Llinás, M. Sexual development in Plasmodium parasites: Knowing when
459 it's time to commit. *Nat. Rev. Microbiol.* **13**, 573–587 (2015).
460 5. Churcher, T. S. *et al.* Predicting mosquito infection from Plasmodium falciparum
461 gametocyte density and estimating the reservoir of infection. *Elife* **2**, e00626 (2013).
462 6. Bradley, J. *et al.* Predicting the likelihood and intensity of mosquito infection from sex
463 specific Plasmodium falciparum gametocyte density. *Elife* **7**, e34463 (2018).
464 7. Collins, K. A. *et al.* A controlled human malaria infection model enabling evaluation of
465 transmission-blocking interventions. *J. Clin. Invest.* **128**, 1551–1562 (2018).

466 8. Diebner, H. H. *et al.* Modelling the transition of asexual blood stages of Plasmodium
467 falciparum to gametocytes. *J. Theor. Biol.* **202**, 113–127 (2000).
468 9. Eichner, M. *et al.* Genesis, sequestration and survival of Plasmodium falciparum
469 gametocytes: Parameter estimates from fitting a model to malariatherapy data. *Trans. R.*
470 *Soc. Trop. Med. Hyg.* **95**, 497–501 (2001).
471 10. Rockett, R. J. *et al.* A real-time, quantitative PCR method using hydrolysis probes for the
472 monitoring of Plasmodium falciparum load in experimentally infected human volunteers.
473 *Malar. J.* **10**, 48 (2011).
474 11. Filarsky, M. *et al.* GDV1 induces sexual commitment of malaria parasites by antagonizing
475 HP1-dependent gene silencing. *Science.* **359**, 1259–1263 (2018).
476 12. Gebru, T., Lalremruata, A., Kremsner, P. G., Mordmüller, B. & Held, J. Life-span of in
477 vitro differentiated Plasmodium falciparum gametocytes. *Malar. J.* **16**, 330 (2017).
478 13. McCarthy, J. S. *et al.* A Pilot Randomised Trial of Induced Blood-Stage Plasmodium
479 falciparum Infections in Healthy Volunteers for Testing Efficacy of New Antimalarial
480 Drugs. *PLoS One* **6**, e21914 (2011).
481 14. Khoury, D. S. *et al.* Within-host modeling of blood-stage malaria. *Immunol. Rev.* **285**,
482 168–193 (2018).
483 15. Piray, P., Dezfouli, A., Heskes, T., Frank, M. J. & Daw, N. D. Hierarchical Bayesian
484 inference for concurrent model fitting and comparison for group studies. *bioRxiv* 393561
485 (2018). doi:10.1101/393561
486 16. Wockner, L. F., Hoffmann, I., O'Rourke, P., McCarthy, J. S. & Marquart, L. Comparison
487 of statistical models to estimate parasite growth rate in the induced blood stage malaria
488 model. *Malar. J.* **16**, 352 (2017).
489 17. Wockner, L. F. *et al.* In vivo growth rate of Plasmodium falciparum: analysis of parasite
490 growth data from malaria volunteer infection studies. *Submitted*
491 18. Saralamba, S. *et al.* Intrahost modeling of artemisinin resistance in Plasmodium
492 falciparum. *Proc. Natl. Acad. Sci.* **108**, 397–402 (2010).
493 19. Tibúrcio, M. *et al.* Specific expression and export of the Plasmodium falciparum
494 Gametocyte EXported Protein-5 marks the gametocyte ring stage. *Malar. J.* **14**, 334
495 (2015).
496 20. Da, D. F. *et al.* Experimental study of the relationship between Plasmodium gametocyte

density and infection success in mosquitoes; implications for the evaluation of malaria transmission-reducing interventions. *Exp. Parasitol.* **149**, 74–83 (2015).

21. Schneider, P. *et al.* Adaptive plasticity in the gametocyte conversion rate of malaria parasites. *PLOS Pathog.* **14**, e1007371 (2018).

22. Bancells, C. *et al.* Revisiting the initial steps of sexual development in the malaria parasite Plasmodium falciparum. *Nat. Microbiol.* **4**, 144–154 (2019).

23. Dogovski, C. *et al.* Targeting the Cell Stress Response of Plasmodium falciparum to Overcome Artemisinin Resistance. *PLOS Biol.* **13**, e1002132 (2015).

24. Cao, P. *et al.* A mechanistic model quantifies artemisinin-induced parasite growth retardation in blood-stage Plasmodium falciparum infection. *J. Theor. Biol.* **430**, 117–127 (2017).

25. Zaloumis, S. *et al.* Assessing the utility of an anti-malarial pharmacokinetic-pharmacodynamic model for aiding drug clinical development. *Malar. J.* **11**, 303 (2012).

26. Duffy, S. & Avery, V. M. Plasmodium falciparum in vitro continuous culture conditions: A comparison of parasite susceptibility and tolerance to anti-malarial drugs throughout the asexual intra-erythrocytic life cycle. *Int. J. Parasitol. Drugs Drug Resist.* **7**, 295–302 (2017).

27. Gelman, A. *et al. Bayesian Data Analysis, Third Edition (Texts in Statistical Science).* (Taylor & Francis, 2013).

28. Lesaffre, E., Rizopoulos, D. & Tsonaka, R. The logistic transform for bounded outcome scores. *Biostatistics* **8**, 72–85 (2007).

29. Stan Development Team. *Stan Modeling Language Users Guide and Reference Manual*. (2017).

30. Lewandowski, D., Kurowicka, D. & Joe, H. Generating random correlation matrices based on vines and extended onion method. *J. Multivar. Anal.* **100**, 1989–2001 (2009).

31. Ahn, J. E., Karlsson, M. O., Dunne, A. & Ludden, T. M. Likelihood based approaches to handling data below the quantification limit using NONMEM VI. *J. Pharmacokinet. Pharmacodyn.* **35**, 401–421 (2008).

32. R Core Team. R: A language and environment for statistical computing. http://www.R-project.org/. *R Foundation for Statistical Computing, Vienna, Austria* (2017).

Acknowledgments

We acknowledge useful conversations with David S. Khoury, Deborah Cromer and Miles P. Davenport (Kirby Institute, UNSW Australia, Sydney, Australia) and assistance in drafting the manuscript from Laura Cascales (QIMR Berghofer, Brisbane, Australia). We thank Jörg J. Möhrle, Head of Translation from Medicines for Malaria Venture for his support and permission to use the trial data. The work is supported by an Australian Research Council (ARC) Discovery project (DP170103076) and a National Health and Medical Research Centre (NHMRC) of Australia Project Grant (1025319) and supported in part by the Australian Centre for Research Excellence on Malaria Elimination, funded by the NHMRC (1134989). SGZ is funded by an ARC Discovery Early Career Researcher Award (170100785). JAS is supported by a NHMRC Senior Research Fellowship (1104975). This research was supported by use of the Nectar Research Cloud, a collaborative Australian research platform supported by the National Collaborative Research Infrastructure Strategy (NCRIS).

Author contributions

P.C., J.A.S., J.S.M. and J.M.M conceived the study; K.A.C. and J.S.M. provided the clinical data; P.C. analyzed the data with contributions from K.A.C., J.A.S., J.S.M. and J.M.M.; P.C. developed the mathematical model; S.Z. developed the statistical method; T.W. and J.T. provided the pharmacokinetic model for piperaquine; P.C. performed model fitting, simulation and analysis; P.C. wrote the first draft of the manuscript; All authors reviewed and commented on the manuscript.

Competing interests

We have no competing interests.