Infection Age as a Predictor of Epidemiological Metrics for Malaria

John M Henry (henry529@uw.edu)
Institute for Health Metrics and Evaluation https://orcid.org/0000-0001-8684-2003
Austin Carter
University of Washington
David L Smith
University of Washington

Research

Keywords: Malaria, Data, Infection Age, Fever, Transmission Efficiency

Posted Date: August 2nd, 2021

DOI: https://doi.org/10.21203/rs.3.rs-746347/v1

Version of Record: A version of this preprint was published at Malaria Journal on April 7th, 2022. See the published version at https://doi.org/10.1186/s12936-022-04134-5.

RESEARCH

Infection Age as a Predictor of Epidemiological Metrics for Malaria

John M Henry[1,2*†], Austin Carter[2] and David L Smith[2]

[*]Correspondence:
henry529@uw.edu
[1] College of the Environment, University of Washington, 1492 NE Boat St, 98105, Seattle, USA
[2]Institute for Health Metrics and Evaluation, University of Washington, 3980 15th Ave. NE, 98195, Seattle, USA
Full list of author information is available at the end of the article
[†]Equal contributor

Abstract

Background: Accurate estimation of the burden of *Plasmodium falciparum* is essential for strategic planning for control and elimination. Due in part to the extreme heterogeneity in malaria exposure, immunity, other causes of disease, direct measurements of fever and disease attributable to malaria can be difficult. This can make a comparison of epidemiological metrics both within and between populations hard to interpret. An essential part of untangling this is an understanding of the complex time-course of malaria infections.

Methods: We reanalyzed malaria therapy infections in which individuals were intentionally infected with malaria parasites. In this analysis, we examined the age of an infection as a covariate describing aggregate patterns across all infections. We performed a series of piecewise linear and generalized linear regressions to highlight the infection age dependent patterns in both parasitemia and gametocytemia, and from parasitemia and gametocytemia to fever and transmission probabilities, respectively.

Results: The observed duration of untreated patent infection was 130 days. As infections progressed, the fraction of infections subpatent by microscopy increased steadily. The time-averaged malaria infections had three distinct phases in parasitemia: a growth phase for the first 6 days of patency, a rapid decline from day 6 to day 18, and a slowly declining chronic phase for the remaining duration of the infection. During the growth phase, parasite densities increased sharply to a peak. Densities sharply decline for a short period of time after the peak. During the chronic phase, infections declined steadily as infections age. Gametocytemia was strongly correlated with lagged asexual parasitemia. Fever rates and transmission efficiency were strongly correlated with parasitemia and gametocytemia. The comparison between raw data and prediction from the age of infection has good qualitative agreement across all quantities of interest for predicting averaged effects.

Conclusion: We established age of infection as a potentially useful covariate for malaria epidemiology. Infection age can be estimated given a history of exposure; accounting for exposure history may potentially provide a new way to estimate malaria-attributable fever rates, transmission efficiency, patent fraction, and more in immunologically naïve individuals such as children and people in low-transmission regions. Understanding how immune responses modify these statistical relationships is key for being able to apply these results more broadly.

Keywords: Malaria; Data; Infection Age; Fever; Transmission Efficiency

Introduction

Despite great progress in recent decades, malaria from *Plasmodium falciparum* infection continues to claim approximately 435 thousand lives each year [1, 2].

Deaths represent only part of the overall burden, as an estimated 194 million cases occurred in 2017 alone [1, 2]. Assessments of potential targeted interventions to efficiently reduce prevalence depend on detailed knowledge of the epidemiology in the region of interest. To this end, studies of *P. falciparum* epidemiology typically rely on asexual parasite counts, clinical incidence, prevalence, serology, and rates of fever to compare age-specific patterns of disease and transmission among populations [3, 4]. Patterns in routine clinical surveillance data can be difficult to interpret due to ambiguity in the causes of observed trends. Parasite densities are highly variable between and within populations, as well in the same individuals over time. Therefore it is important to characterize patterns which appear in first infections in order to determine how they may be altered with different levels of past exposure.

Previous research on infection dynamics has described the highly volatile trajectories of parasitemia in a single individual over time, whose counts can jump orders of magnitude over the course of a day [5, 6]. Studies focused on infection durations have estimated that the average time to the last observed patent infection ranges from around 100 days to over 1000 days [7, 8, 9], with at least one confirmed infection persisting for over a decade [10]. Patterns in the spikes and troughs of parasite counts have been investigated for evidence of patterns that suggest the impact of VAR gene switching in parasite densities [9], or blood cell age preferences of the parasites [11]. Insight derived from these studies are valuable, but difficult to interpret in context or scale to country or continent level estimates of burden. On the other end, asexual parasitemia is a standard covariate for estimating the malaria attributable fraction of fever [3] and the heterogeneous relationship between gametocytemia and transmission efficiency [12, 13, 14] at the population level, but there remain lingering issues of identifiability regarding the impact of immunity.

Due in part to widely varying histories of past exposure among individuals in a population, observational studies of the relationship between quantities such as prevalence, malaria attributable fever, and per capita transmission rate may be misleading if they do not take into account the heterogeneous effects of immunity. Innate immune responses may differ in individuals depending on exposure, age, and other possible underlying conditions, and adaptive immune responses may differ based on an individual's history of past exposure. The effects of immunity limit an infection, reducing parasite population growth and eventually clearing infections. The impact of adaptive immunity on subsequent infections can further be decomposed into five categorical effects: pre-erythrocytic immunity, anti-(asexual) parasitic immunity, parasite tolerance, anti-gametocytic immunity, and transmission blocking immunity. Pre-erythrocytic immunity slows or prevents the establishment of an infection before or in the liver. Anti-parasitic immunity acts to reduce the blood stage parasitemia, which is correlated with disease. Parasite tolerance modifies the relationship between parasitemia and disease; it is measured as a reduction in the likelihood of fever and other clinical symptoms for a given parasitemia. Anti-gametocytic immunity reduces gametocyte densities. Gametocytes infect mosquitoes, and their densities are correlated with transmission efficiency. Transmission blocking immunity reduces transmission efficiency for a given gametocytemia, analogous to parasite tolerance [13]. This is mediated through immune effectors which target gametocyte-specific antigens, compounding the impact of antiparasite immunity which reduces the number of asexual parasites which produce

gametocytes. Malaria epidemiology and immunity can therefore be understood as a set of cascading consequences of parasitemia (Figure 1). As these different modalities will impact fever, patency, and transmission rates, a direct translation from parasitemia or prevalence to other epidemiological measures without previous exposure taken into account may be difficult to establish at best and lead to spurious patterns at worst.

Understanding how *P. falciparum* infections develop in the absence of acquired immunity is thus key to understanding and interpreting malaria data in which immunity has modified baseline patterns. Baseline data are difficult to obtain in malaria-endemic areas, but data describing some malaria infections in non-immune individuals is available from the historical data describing carefully monitored, deliberate malaria infections used to induce a fever to treat neuro-syphilis, called malaria therapy [5, 15, 16, 17, 18, 19]. The malaria therapy data used in this study was initiated through either intravenous injection with asexual parasites directly or mosquito bite with sporozoites. They have previously been used to study the duration of single infections [20]. This is difficult to estimate from longitudinal studies due to the relatively common occurrence of superinfection, in which individuals are infected with multiple cohorts of parasites simultaneously, and unknown past exposure. Some recent studies have used genetic data to follow individual infections and estimate the multiplicity of infection [21, 22] but such studies can only detect genetically distinct strains sporadically, and infections with the highest densities at the time of measurement mask the presence of lower density infections. Some malaria therapy patients were infected several times, and their records have been used to study the effects of adaptive immunity by comparing the difference between homologous and heterologous challenge; effects of immunity to homologous challenge appear to be present after one or two infections, but it may take more exposure for strain transcending immunity to occur [23].

In this study, we have focused on population-averaged patterns in parasite densities over the time course of the infection in individuals with no prior exposure as a reference for uncomplicated malaria with no effects of previously acquired immunity in adults using a sample from the malaria therapy patient data. Further, we establish statistical relationships between average parasitemia and epidemiological measures such as fever rates and transmission efficiency in relation to recent exposure in immunologically naïve individuals (Figure 1). Given an infection has not been cleared, it can either be patent or subpatent; given patency, individuals will have some measurable asexual parasitemia. This parasitemia is used as a measure of severity of disease, and therefore risk of symptoms such as fever. Asexual parasites also produce gametocytes after some maturation period, previously estimated to be 9-12 days [24, 25], which persist with a short half-life [26]. In turn, gametocytemia is used as a predictor for transmission efficiency [13, 14]. We use this framework as a lens through which to statistically view the aggregate data, we demonstrate that the age of an infection can be used as a potentially powerful surrogate for estimating these hard to measure and dynamic quantities.

Methods

We used the age of infection to model population averaged features of the malaria therapy data set, in which 316 individuals with neurosyphilis were purposefully

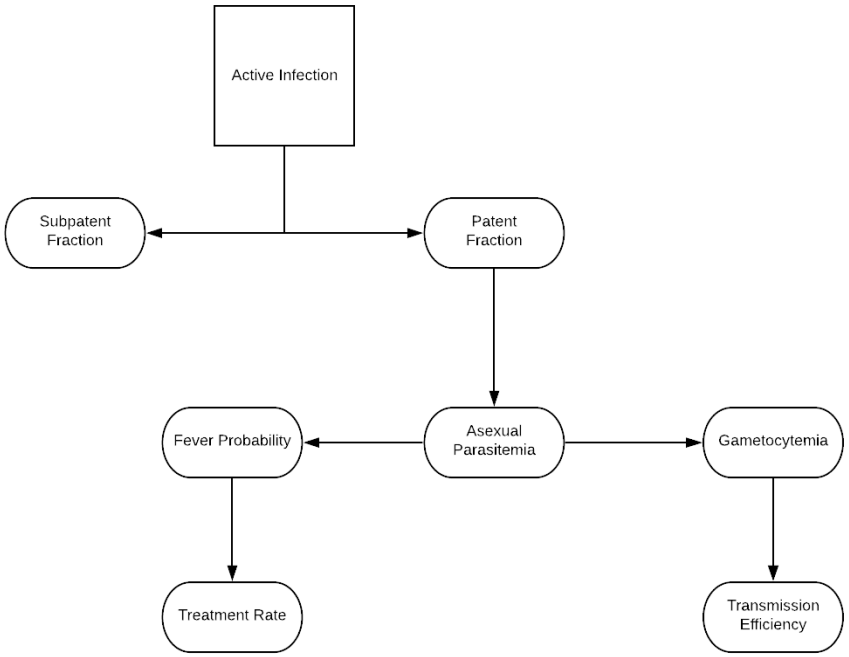

Figure 1 Schematic breakdown of infections. Given infections are active, some fraction will be subpatent (ie, undetectable by light microscopy) and the rest will be patent. Given the parasitemia is patent, we can estimate the parasitemia which will inform fever probability today and gametocytemia roughly 9 days in the future. Fever rates presumably are correlated with treatment rates as symptomatic individuals are much more likely to seek treatment, and gametocytemia is positively correlated with the transmission efficiency per mosquito bite.

infected with one of several strains of *falciparum* malaria. Infection age was estimated from the first day of patency of an infection. The data used consists of patient records between 1941 and 1954 in the American south. The participants were patients with neurosyphilis being treated with malaria parasites, which were intended to induce a fever and effective immune response against spyrochetes to mitigate outcomes of neurosyphilis. The infections were initiated through injection with either sporozoites to induce a liver infection or merozoites to directly induce blood stage infection. Little clinical difference in outcomes aside from latency before first measurement was noticed, so both types of exposure were included here. Once patent by microscopy, daily measurements were taken of asexual and sexual stage parasites, and body temperature if there was an apparent fever. If symptoms of malaria became severe or if parasite densities were too high, patients were given treatment inadequate to cure malaria but sufficient to reduce parasitemia. When there were detectable gametocytes in the blood, mosquito feedings were performed to determine the transmission efficiency from human to mosquito. Once the infection had been subpatent for some time, full treatment was given to clear the parasites entirely. In a subset of patients, secondary infections were initiated through either homologous or heterologous challenge.

For the estimate of duration of patent infection, we excluded treated cases (treated n=189). For all other estimation, we included infections that were treated until the

day they were first treated, where they were truncated (n=299). On any day that an individual had patent gametocytemia, a mosquito feeding was performed and after the estimated extrinsic incubation period the mosquitoes were dissected to determine the fraction which became sporozoite positive (n=2029 observations). The subsequent infection challenges that occurred in some patients were excluded, as we were interested in the course of first infections.

The primary features of interest included infection duration, patent fraction over time, asexual parasitemia over time, gametocytemia over time, fever risk associated with parasitemia, and transmission efficiency associated with gametocytemia. The fits of the relationships between all the observable quantities are summarized in table 1. For patency, asexual parasitemia, and gametocytemia, piecewise linear or generalized linear fits were performed and summarized. The distributions of asexual parasitemia and gametocytemia were also represented in violin plots aggregated by month to show general trends. Daily means and variances appeared to have a relationship, so a power law was fit. Logistic regressions were performed to translate daily average parasitemia to fever risk, and smoothed gametocytemia to transmission efficiency. The degree of zero inflation in the transmission efficiency and the beta-fitted histograms of transmission efficiency across binned levels of gametocytemia were also plotted to emphasize the overdispersion of the relationship.

Results

The average duration of the infection, restricted to the subset of patients who were untreated during the entire infection (n=110), was estimated to be 130 days (Figure 2b). We defined the duration as the age of infection on the last day with patent asexual parasitemia by microscopy, which was followed by a sequence of parasite negative observations and the cessation of measurement. We compared exponential, gamma, weibull, and lognormal survival curves, with delta AIC (23.3,0,1.7,6.8, respectively) and delta BIC (20.6,0,1.7,6.8, respectively) confirming exponential as a poorest fit and gamma and weibull as being the best candidates. The plotted blue curve is the gamma survival curve, with shape parameter 2.058 and rate parameter .015 per day.

For the analyses that follow, we included all infections, including observations of all infections until they were ended by treatment. This allowed us to reduce the bias of removing all infections with higher parasitemia which tended to be treated at higher rates.

Asexual parasite densities followed a three-phase pattern (Figure 2). For the first six days of a patent infection, parasite densities increased geometrically. This was followed by a sharp decline between days 6-20. After that, which we have called the chronic phase, parasite densities and the fraction patent on a given day slows to a shallower nearly linear trend in time. Among patent infections, the log10 transformed daily average declined linearly (Figure 2a). In the chronic phase, infections sometimes had subpatent periods before spikes in parasitemia occurred again. The proportion of peristent infections which were patent on a given day declined steadily (Figure 2c). Older infections tend to spend a signficant fraction of time at submicroscopic densities in the blood with occasional bouts above that theshold of detectability, somewhere around 88 parasites per cmm of blood [27]. The increasing

variance in points around the fitted average is due in part to the sample size in the daily averages decreasing as individuals either receive treatment or recover. Additional variation occurred to to a smaller fraction of those with persisting infection remaining patent, so many of the later points are a small number of individuals with late spikes in parasitemia.

We found a strong relationship between asexual parasitemia and gametocytemia. The log10 of the average parasitemia and gametocytemia appeared to be shifted and scaled versions of one another (Figure 3a). Lagged average parasitemia as a linear predictor of gametocytemia was therefore explored. The optimal lag was determined to be 9 days, which minimized the standard deviation of the residuals (see the trough of figure 3c). The flat nature of the standard deviation as a function of the lag around days 8-12 shows that other lags may be nearly as good of fits, which is consistent with the estimated maturation period of gametocytes [24, 25]. Gametocyte densities in the chronic phase also declined linearly (Fig 3b). After accounting for the lag, gametocyte densities were approximately 10-fold lower than parasite densities).

In addition to patterns in average asexual parasitemia and gametocytemia, we investigated patterns in the distribution across all individuals on a given day (Figure 4). Monthly violin plots of the asexual parasitemia and gametocytemia after log-transformation appeared to maintain their shape while shifting down as infections aged. We quantified a power law relationship between mean and variance of both asexual parasitemia and gametocytemia for patent infections (Fig 4b,d). Power laws are often known to exist in higher density regions [28], although a decrease in variance may occur in measurements near the threshold of detectability by light microscopy as lower measurements are likely to be recorded as subpatent and therefore not included in daily measurements.

Log10 daily measurements of parasitemia were then fit as a predictor of the proportion of individuals with fever, conditioned on patent parasitemia. The best fit was a sigmoid function from the GLM with logit link (Figure 5). There was a different relationship between parasite densities and fever in the first five days of patency compared to the rest of the infection. Note the five purple points laying above the sigmoid, with the leftmost two in particular being strong outliers. These points represent the first five days of infection, with the days being ordered from left to right. This implies parasitemia is a poor predictor in the first few days of infection, and in particular fever may come days before high parasitemia. Unsurprisingly, fever was also a function of the age of the infection. The fit appears to follow the data very well, though underestimates the first few days of infection as expected.

Finally we quantified the relationship between gametocytemia and transmission efficiency. Transmission efficiency was measured as the fraction of mosquitoes that developed sporozoites after feeding on individuals with patent gametocytemia. The relationship between gametocytemia and transmission efficiency was consistent but noisy. To fit the data, we modeled the data generating process as a mixture process, a zero-inflated beta-binomial distribution. Smoothing individual measurements across log10 gametocytemia by averaging over measurements with similar gametocytemia resulted in the blue points in figure 6a. A logistic regression was performed on the blue points, with weights proportional to the number of measurements used in the

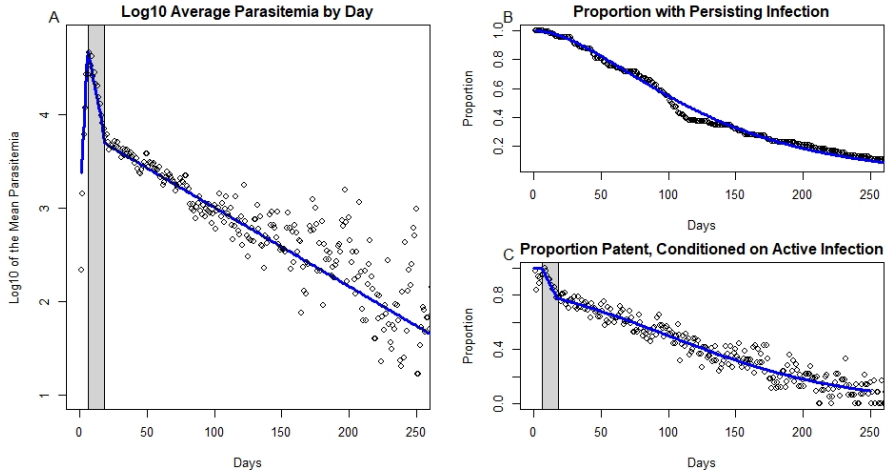

Figure 2 Plots of parasitemia, duration, and patency. Panel A shows a scatter plot of the log10 of the daily average parastemia among patent infections conditioned on continued infection. The blue lines are a three piecewise linear fit in three parts, with a gray shaded region representing the middle of the three. Panel B shows a plot of the empirical survival function of infections that had positive measures of parasitemia on or after that day of infection. The blue curve represents the best fit gamma survival curve, that is the complement of the corresponding gamma CDF. Finally panel C is a plot of the proportion patent conditioned on continued infection. The gray shaded region is the same highlighted in panel A, showing that during the initial growth phase nearly all infections remain patent; then patency drops about 20 percent over a short time, then it slowly decays to around 10 percent by day 250.

average. Analogous to figure 5b, figure 6b shows the log linear fit of gametocytemia filtered through the sigmoid in green.

The beta-binomial interpretation allows us to quantify overdispersion, with zero-inflation added due to the large abundance of zeroes. Interestingly, the amount of zero inflation appears to decrease with increasing gametocytemia, as shown in figure 6c. Conditioning on a nonzero number of mosquitoes counted, we then plotted the histograms in binned gametocytemia with beta distributions fit in figures 6d-i. As expected, increasing gametocytemia shifts likelihood to higher transmission efficiency but with a large amount of variability. Therefore, even highly gametocytemic patients often infected less than expected mosquitoes, or even none at all, during a particular feeding. This could be partially explained by a result of fairly small samples of mosquitoes feeding successfully per patient per day, though another mechanism cannot be ruled out such as differences in sex ratios of gametocytes or strain-specific differences that are accentuated in transmission more than disease states [29].

Discussion

We have shown have shown that asexual parasite densities is strongly predicted by the age of infection, and we have also quantified the relationship between parasite densities and fever and gametocytemia and the relationship between gametocytemia and infectiousness. In particular, we are able to estimate expected asexual parasitemia as a function of age from which every other quantity can be estimated. The diagram in figure 1 shows pathways through intermediate quantities to translate from infection age to outcomes of interest. As we also have a relationship between

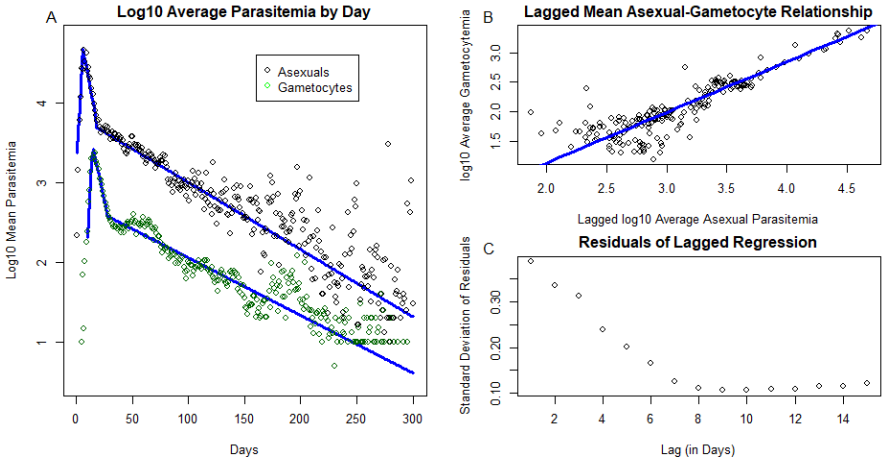

Figure 3 Panel A shows two time series of points, the top representing the log10 of the daily mean parasitemia on a given day across all continued patent infections as in figure 2 panel A and the bottom green time series represents the corresponding log10 daily mean gametocytemia. A fit to predict gametocytemia from lagged parasitemia was performed across many lags, and the optimal lag was determined by the minimum of the standard deviation of the lagged residuals shown in panel C. The linear fit at the optimal lag of 9 days in presented in panel B.

the mean and variance of asexual parasitemia and gametocytemia, if we fit a family of distributions we can also get estimates of full distributions of fever rates and transmission efficiencies as well. Therefore this represents a potentially powerful framework for future estimation.

This analysis suggests it may be possible to translate knowledge of a history of exposure to epidemiologically important quantities which are strongly correlated to the age of the infection if acquired immunity can also be estimated. Averaging these conditional rates on the probability an individual has had an infection for some duration across all present infection ages weighted by the fraction of the population who has had an infection for that duration gives expected population-level metrics in a given transmission setting. The basic idea follows the law of the unconscious statistician. Suppose we want to estimate some observable X which depends on the concentration of the pathogen p,

$$E[X_p] = \int_0^\infty P(p = p_0) E[X_{p_0} | p_0] dp_0$$

Examples of such X include fever rates and transmission efficiency, as each would reasonably increase on average with pathogen density. However, we would have to know the probability that someone in the population has a pathogen concentration of p. This is rarely known and difficult to measure directly, as surveys often only measure incident cases which tend to have higher parasitemia, and detection rates themselves depend on pathogen densities. However due to the observation that malaria appears to exhibit strong average pathogen concentration patterns with respect to the age of an infection, we can parameterize p (and therefore X) through the infection age α. This gives

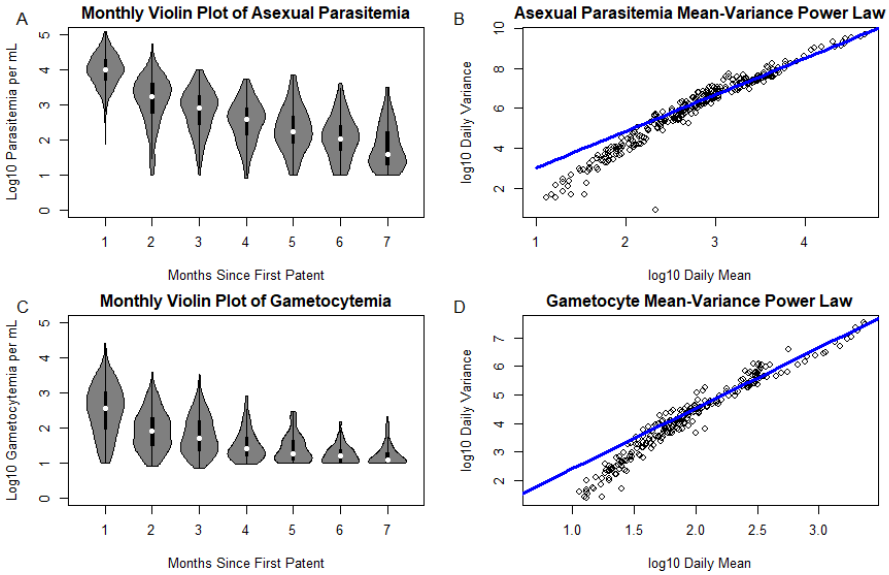

Figure 4 Panel A and C represent violin plots of the distributions across individuals of respectively the log10 daily parasitemia and log10 daily gametocytemia across all individuals, aggregated across 30 day periods for compactness. Despite being log transformed, the distribution shapes appear to shift down with infection age but maintain the same general shape, suggesting a power-law relationship between the mean and variance of parasitemia. Power laws are then fit to log10 transformed daily mean and variances in panels B and D, restricted to those above the estimated sensitivity of light microscopy (88 parasites per cmm blood).

$$E[X_{p(\alpha)}] = \int_0^\infty P(\alpha = \alpha_0) E[X_{p(\alpha_0)} | \alpha_0] d\alpha_0$$

The advantage to this formulation is we can estimate the probability of having an infection of age α through the use of standard age of infection dynamics given estimates of exposure [30]. The resulting distribution can give a reasonable estimate on the unknown probability given patterns in rates of exposure, allowing us to compute the desired expectation.

Past work on malaria attributable fever [3], which estimated malaria attributable fraction of total fever based on parasitological survey data of children, is analogous to the parasitemia-to-fever risk regression done here. Their interpretation allows for the logistic regression to assign a probability to any child with parasite density measurements and a fever to determine how likely it is that the fever they have is attributable to malaria. The data was restricted to relatively young children, so the effects of adaptive immunity could be largely ignored as they were here. However to apply those results to an entire population, the regression would need to be reworked across measurements of all age groups as their past exposure and developed immunity will modify parasite densities and fever tolerance. Additionally, exposure history may vary dramatically from location to location, so their method would require an enormous amount of data and each location with its unique history of exposure could have a very different estimate from even other locations with the same current day prevalence. The analysis presented here shows that if we are instead able to estimate how long ago individuals were most recently infected based

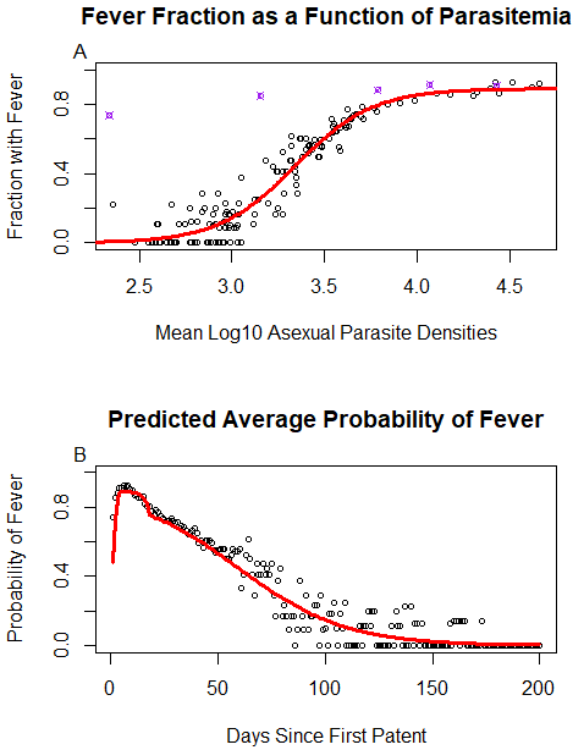

Figure 5 Panel A shows a logistic fit of the log10 of daily mean asexual parasitemia to the fraction of individuals with objective fever. The five points in purple above the logistic curve are the first 5 days of infection, suggesting that many febrile individuals got primary fevers before high density infections. Panel B shows a time series of the daily fraction of individuals with active infections who have fever. The red curve represents the transformed piecewise fit of parasitemia transformed through the fitted logistic curve.

on a history of exposure from historical prevalence data, we may be able to estimate a malaria attributable fraction of fever in the absence of parasite density surveys given a reasonable model of immunity.

Although parasitemia appears to be a very good predictor of fever after the first week or so of patent parasitemia, the fever rate is consistently higher than predicted by parasitemia in the early days of patency. This suggests the notion of a difference between primary fever, caused at the beginning of an infection, and a secondary fever, correlated strongly with parasitemia and occurring later in the infection. This could be a consequence of the inflammatory cascade early in infection which is subsequently tempered by anti-inflammatory responses as the immune response has matured. An uncontrolled early inflammatory response is often found in severe cases of malaria [31, 32], so mortality in cases may be closely correlated with this early stage of infection.

Analogous to the fever and asexual parasitemia relationship, the infection reservoir and its impact on estimation of the human-to-mosquito transmission potential in environments with seasonal transmission may be largely impacted by the additional heterogeneity presented here. In addition to the overdispersion shown in the translation from gametocytemia to transmission efficiency, the infection-age depen-

dent patterns of infectivity may be leveraged to improve our understanding of which locations may be a "source" or "sink" for malaria transmission, and how a location may switch from one to another based on the history of recent exposure and the age group of individuals in question [33, 34].

The fit for transmission efficiency as a function of gametocyte densities appears to be consistently above the data. This is due to a technical difference in the fitting procedure compared to fever. While fever was predicted from daily average parasitemia, transmission efficiency was predicted from a function of the gametocytemia measurements and not to the averaged time series data directly. Combined with the fact that most of the gametocytemia measurements in the time series are from the top half the fitted sigmoid and therefore filtered through a concave function, Jensen's inequality guarantees that plugging in the mean to the function rather than taking an average of the data filtered through the sigmoid is expected to be an overestimate. However either through computing the conditional expectation mentioned above or simulating draws of gametocytemia on a given day then converting through the sigmoid function, this problem is avoided completely.

The power law relationships demonstrated between the mean and the variance, which suggests evidence of the well-known Taylor's Law from ecology [35], can be used for practical computations here. This relationship is often seen between the means and variances of populations across different spatial regions, but applies here as each human host can be imagined as an independent habitat for the parasite populations. One of the significant advantages of it is it allows us to use the relatively simple pattern in mean parasitemia over time to obtain a similar pattern in variance over time, and therefore parametrically describe a wide class of two-parameter distributions for parasitemia as a function of infection age through moment matching. This allows us to propagate uncertainty of our estimates through the relationships in a way which circumvents the issues of Jensen's inequality mentioned above.

It is crucial to highlight several limitations to the extensibility of these observations. The atypical immunological states of the patients considered here are a clear concern. Naturally questions can be asked about the application of trends found in adults with neurosyphilis to otherwise healthy individuals in endemic settings. Further, all subjects were adults with presumably fully developed immune systems, and therefore their response may differ from children in endemic settings as well. However, none of these individuals have previously had exposure to malaria and therefore represent a sort of baseline for trends in first exposure, even if the exact values of the parameter values are not perfectly representative.

A conscious decision was made here to not limit the analysis to a single strain of *P. falciparum*. Strain-specific differences may play a large role in overall transmission dynamics [9], but often genetic information of strain diversity is limited in a particular setting. Additionally, no simple mapping between strain and pathogenicity or specific parasitemia profiles exists. Inflammatory signalling also changes in response to recent exposure [31], possibly altering the baseline relationship between fever and infections. Often in endemic settings individuals will have multiple infections simultaneously, so any mapping would also need to account for pairwise interactions or work on an assumption of independence. Handling the possibility of superinfection on this age of infection relationship needs further investigation in the future.

The five modalities of immunity as well as any age dependent trends will vary over time and impact all of the statistics presented here. With exposure, parasitemia and gametocytemia will decline; separate from that, higher parasitemias can be tolerated before a fever develops, and transmission given a set gametocytemia may decline. For these reasons, a static mapping from prevalence to fever rates or transmitting fraction in the absence of information on the history of exposure may be a poor representation of the epidemiological reality. Patterns of exposure (seasonality, source/sink dynamics, human travel patterns, etc) should play a large role in developing a dynamic mapping, which coincides with the understanding that malaria is a very heterogeneous disease by location. Given an understanding of the patterns of exposure, it may be possible to estimate the likely distribution of immune states in the population and take that into account for estimation of quantities of interest.

In light of these statistical relationships, if we are given an infection age distribution we can obtain estimates of fever rates and transmission potential in the absence of immunity. If a direct measure of these quantities is available, this may be able to act as a counterfactual for measuring the impact of immunity; if a model of immunity is included, we can obtain estimates of the quantities. In both cases, an infection age distribution is a crucial piece. Therefore, it places particular emphasis on the importance in determining such infection age distributions. Subsequent work is aimed to provide a model-based approach for constructing reasonable families of distributions of the age of infection given an exposure history.

Difficulty interpreting data arises in part from the extreme range of unknown previous exposure history across locations. Exposure has been measured at levels varying from no bites to more than a thousand bites by infectious mosquitoes, per person, per year; transmission efficiency, detection, and clinical manifestations of malaria depend on previous exposure and acquired immunity; exposure is often seasonal, and highly heterogeneous across individuals; and acquired immunity to malaria develops slowly, varies by exposure, protects poorly, and has poor memory [36, 23]. This prompts many studies to focus on the prevalence and outcomes in children, who can be reasonably assumed to have had little previous exposure [37]. Often patterns in children appear to more closely match trends in exposure intensity, but this leads to an identifiability problem: differences may arise due to the acquisition of immunity directly, or due to differences in the developing immune systems of children. Observed age-dependent patterns are likely a combination of both of these effects. Integrating the statistical patterns demonstrated here with mechanistic models of immunity may help to address this.

Conclusions

We have quantified patterns relating the age of malaria infections to patency, parasitemia, gametocytemia, fever rates, and transmission efficiency. These patterns can be leveraged in population-level models of disease transmission to obtain dynamic estimates of each of these quantities. Future investigations can use these models to determine the impact of public health interventions on fever and transmission.

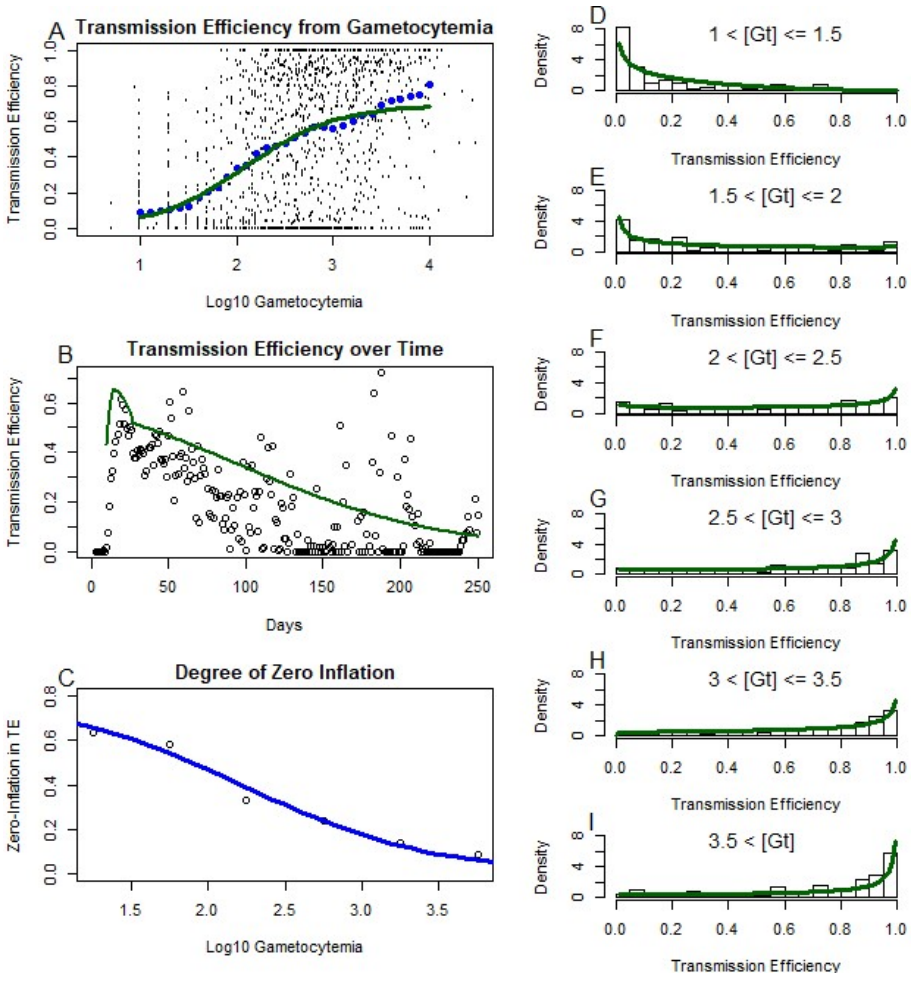

Figure 6 The black points in panel A represent the transmission efficiency, measured through the fraction of fed mosquitoes who developed sporozoites, against the log10 gametocytemia measured on that day for each individual on every day they were gametocyte positive. Treated infections were not included in this analysis. Blue points are a rolling average of transmission efficiency as a function of log10 gametocytemia, and the green curve is a sigmoid curve fit to these points with weights given by the number of points included in their rolling average. Panel B shows the daily average transmission efficency as a function of infection age, with the green curve representing the measured log10 daily average gametocytemia composed with the sigmoid fit in panel A, showing good qualitative agreement. As the relationship between gametocytemia and transmission efficiency is highly heterogeneous and zero-inflated, zero-inflated beta distributions were fit to binned values of gametocytemia to quantify this. Panel C shows the degree of zero inflation, that is the fraction of mosquito feedings resulting in no infections at all, as a function of log10 gametocytemia. The blue sigmoid curve was fit to these points, showing the apparent zero-inflation decreases with increasing gametocytemia. Finally, panels D-I are histograms of transmission efficiency for a given range of gametocytemia conditioned on nonzero measurements, with beta distributions fit to each. Despite the high degree of heterogeneity, density can be seen to aggregate to the right with increasing gametocytemia.

Table 1 Fitted Relationships between Infection Age and Quantities of Interest

Quantity of Interest	Fitted Equation of Kernel
Patent Fraction, D(α)	$\begin{cases} 1 & \text{if } \alpha \leq 6 \\ 1.12 - .02\ \alpha & \text{if } 6 < \alpha \leq 18 \\ \left(1 + e^{-1.52 + 0.0151\ \alpha}\right)^{-1} & \text{if } \alpha > 18 \end{cases}$
$\log_{10}$ Asexual Parasitemia, P(α)	$\begin{cases} 0, & \text{if } \alpha \leq 0 \\ 3.10 + .278\ \alpha, & \text{if } 0 < \alpha \leq 6 \\ 5.12 - .0743\ \alpha, & \text{if } 6 < \alpha \leq 18 \\ 3.85 - .00843\ \alpha, & \text{if } \alpha > 18 \end{cases}$
Fever Probability, F(P(α))	$\frac{.859\ e^{3.45\ P(\alpha)}}{58200 + e^{3.45\ P(\alpha)}}$
$\log_{10}$ Gametocytemia, G($L_9P(\alpha)$)	$-0.684 + .892\ L_9P(\alpha), \quad L_9P(\alpha) = P(\alpha - 9)$
Transmission Efficiency, c(G(α))	$\frac{.683\ e^{2.14\ G(\alpha)}}{131 + e^{2.14\ G(\alpha)}}$

Patent fraction consisted of two piecewise-linear fits and a generalized linear fit at each of the transitions mentioned in the text. Log parasitemia likewise is piecewise linear. Fever probability and transmission efficiency are logistic functions of their predictors, which are linear functions of infection age.

Declarations

Ethics approval and consent to participate

Not applicable; these data have been analyzed and published before [5, 15, 16, 17, 18, 19].

Consent for publication

Not applicable.

Availability of Data and Materials

All code used during this study are at the following location:

https://github.com/jmh227/Malaria-Therapy-Data-Analysis

The datasets used and/or analysed during the current study are available from the corresponding author on reasonable request.

Competing interests

The authors declare that they have no competing interests.

Author's contributions

All code, statistical fitting and preparation of the manuscript were performed by JMH. All authors contributed equally in the design of the analysis and editing of the manuscript.

Acknowledgements

The authors acknowledge and thank Klaus Dietz and William E. Collins for giving us access to these data. The authors also acknowledge funding from the Bill & Melinda Gates Foundation (OPP1110495, OPP1159934). JMH acknowledges funding from the J.W. Conquest Endowment Fund.

Author details

[1] College of the Environment, University of Washington, 1492 NE Boat St, 98105, Seattle, USA. [2]Institute for Health Metrics and Evaluation, University of Washington, 3980 15th Ave. NE, 98195, Seattle, USA.

References

1. Weiss, D.J., Lucas, T.C.D., Nguyen, M., Nandi, A.K., Bisanzio, D., Battle, K.E., Cameron, E., Twohig, K.A., Pfeffer, D.A., Rozier, J.A., Gibson, H.S., Rao, P.C., Casey, D., Bertozzi-Villa, A., Collins, E.L., Dalrymple, U., Gray, N., Harris, J.R., Howes, R.E., Kang, S.Y., Keddie, S.H., May, D., Rumisha, S., Thorn, M.P., Barber, R., Fullman, N., Huynh, C.K., Kulikoff, X., Kutz, M.J., Lopez, A.D., Mokdad, A.H., Naghavi, M., Nguyen, G., Shackelford, K.A., Vos, T., Wang, H., Smith, D.L., Lim, S.S., Murray, C.J.L., Bhatt, S., Hay, S.I., Gething, P.W.: Mapping the global prevalence, incidence, and mortality of *Plasmodium falciparum*, 2000–17: a spatial and temporal modelling study. Lancet **394**(10195), 322–331 (2019)
2. World Health Organization: World malaria report 2020: 20 years of global progress and challenges (2020)
3. Smith, T., Schellenberg, J.A., Hayes, R.: Attributable fraction estimates and case definitions for malaria in endemic areas. Stat. Med. **13**(22), 2345–2358 (1994)
4. Dalrymple, U., Cameron, E., Bhatt, S., Weiss, D.J., Gupta, S., Gething, P.W.: Quantifying the contribution of *Plasmodium falciparum* malaria to febrile illness amongst african children. Elife **6** (2017)
5. James, S.P.: Some general results of a study of induced malaria in england. Trans. R. Soc. Trop. Med. Hyg. **24**(5), 477–525 (1931)
6. James, S.P., Nicol, W.D., Shute, P.G.: A study of induced malignant tertian malaria. Proc. R. Soc. Med. **25**(8), 1153–1186 (1932)
7. Eyles, D.E., Young, M.D.: The duration of untreated or inadequately treated *Plasmodium falciparum* infections in the human host. J. Natl. Malar. Soc. **10**(4), 327–336 (1951)
8. Sama, W., Killeen, G., Smith, T.: Estimating the duration of *Plasmodium falciparum* infection from trials of indoor residual spraying. Am. J. Trop. Med. Hyg. **70**(6), 625–634 (2004)
9. Eckhoff, P.: *P. falciparum* infection durations and infectiousness are shaped by antigenic variation and innate and adaptive host immunity in a mathematical model. PLoS One **7**(9), 44950 (2012)
10. Ashley, E.A., White, N.J.: The duration of *Plasmodium falciparum* infections. Malar. J. **13**(1), 500 (2014)
11. McQueen, P.G., McKenzie, F.E.: Age-structured red blood cell susceptibility and the dynamics of malaria infections. Proc. Natl. Acad. Sci. U. S. A. **101**(24), 9161–9166 (2004)
12. Churcher, T.S., Trape, J.-F., Cohuet, A.: Human-to-mosquito transmission efficiency increases as malaria is controlled. Nat. Commun. **6**, 6054 (2015)
13. Ouédraogo, A.L., Eckhoff, P.A., Luty, A.J.F., Roeffen, W., Sauerwein, R.W., Bousema, T., Wenger, E.A.: Modeling the impact of *Plasmodium falciparum* sexual stage immunity on the composition and dynamics of the human infectious reservoir for malaria in natural settings. PLoS Pathog. **14**(5), 1007034 (2018)
14. Slater, H.C., Ross, A., Felger, I., Hofmann, N.E., Robinson, L., Cook, J., Gonçalves, B.P., Björkman, A., Ouedraogo, A.L., Morris, U., Msellem, M., Koepfli, C., Mueller, I., Tadesse, F., Gadisa, E., Das, S., Domingo, G., Kapulu, M., Midega, J., Owusu-Agyei, S., Nabet, C., Piarroux, R., Doumbo, O., Doumbo, S.N., Koram, K., Lucchi, N., Udhayakumar, V., Mosha, J., Tiono, A., Chandramohan, D., Gosling, R., Mwingira, F., Sauerwein, R., Paul, R., Riley, E.M., White, N.J., Nosten, F., Imwong, M., Bousema, T., Drakeley, C., Okell, L.C.: The temporal dynamics and infectiousness of subpatent *Plasmodium falciparum* infections in relation to parasite density. Nat. Commun. **10**(1), 1433 (2019)
15. Jeffery, G.M., Eyles, D.E.: The duration in the human host of infections with a Panama strain of *Plasmodium falciparum*. Am. J. Trop. Med. Hyg. **3**(2), 219–224 (1954)
16. Collins, W.E., Jeffery, G.M.: A retrospective examination of sporozoite-and trophozoite-induced infections with *Plasmodium falciparum*: development of parasitologic and clinical immunity during primary infection. Am. J. Trop. Med. Hyg. **61**(1_Supplement), 4–19 (1999)
17. Collins, W.E., Jeffery, G.M.: A retrospective examination of sporozoite- and trophozoite-induced infections with *Plasmodium falciparum* in patients previously infected with heterologous species of plasmodium: effect on development of parasitologic and clinical immunity. Am. J. Trop. Med. Hyg. **61**(1 Suppl), 36–43 (1999)
18. Collins, W.E., Jeffery, G.M.: A retrospective examination of the patterns of recrudescence in patients infected with *Plasmodium falciparum*. Am. J. Trop. Med. Hyg. **61**(1 Suppl), 44–48 (1999)

19. Collins, W.E., Jeffery, G.M.: A retrospective examination of sporozoite- and trophozoite-induced infections with *Plasmodium falciparum*: development of parasitologic and clinical immunity following secondary infection. Am. J. Trop. Med. Hyg. **61**(1_Supplement), 20–35 (1999)
20. Bretscher, M.T., Maire, N., Chitnis, N., Felger, I., Owusu-Agyei, S., Smith, T.: The distribution of *Plasmodium falciparum* infection durations. Epidemics **3**(2), 109–118 (2011)
21. Zhong, D., Koepfli, C., Cui, L., Yan, G.: Molecular approaches to determine the multiplicity of plasmodium infections. Malar. J. **17**(1), 172 (2018)
22. Zhong, D., Lo, E., Wang, X., Yewhalaw, D., Zhou, G., Atieli, H.E., Githeko, A., Hemming-Schroeder, E., Lee, M.-C., Afrane, Y., Yan, G.: Multiplicity and molecular epidemiology of *Plasmodium vivax* and *Plasmodium falciparum* infections in east africa. Malar. J. **17**(1), 185 (2018)
23. Doolan, D.L., Dobaño, C., Baird, J.K.: Acquired immunity to malaria. Clin. Microbiol. Rev. **22**(1), 13–36 (2009)
24. Hawking, F., Wilson, M.E., Gammage, K.: Evidence for cyclic development and short-lived maturity in the gametocytes of *Plasmodium falciparum*. Trans. R. Soc. Trop. Med. Hyg. **65**(5), 549–559 (1971)
25. Sinden, R.E.: Sexual development of malarial parasites. Adv. Parasitol. **22**, 153–216 (1983)
26. Smalley, M.E., Sinden, R.E.: *Plasmodium falciparum* gametocytes: their longevity and infectivity. Parasitology **74**(1), 1–8 (1977)
27. Joanny, F., Löhr, S.J.Z., Engleitner, T., Lell, B., Mordmüller, B.: Limit of blank and limit of detection of plasmodium falciparum thick blood smear microscopy in a routine setting in central africa. Malar. J. **13**, 234 (2014)
28. Anderson, R.M., Gordon, D.M., Crawley, M.J., Hassell, M.P.: Variability in the abundance of animal and plant species. Nature **296**(5854), 245–248 (1982)
29. Tadesse, F.G., Meerstein-Kessel, L., Gonçalves, B.P., Drakeley, C., Ranford-Cartwright, L., Bousema, T.: Gametocyte sex ratio: The key to understanding *Plasmodium falciparum* transmission? Trends Parasitol. **35**(3), 226–238 (2019)
30. Brauer, F.: Age of infection in epidemiology models. In: Electronic Journal of Differential Equations, Conference, vol. 12, pp. 29–37 (2005)
31. Portugal, S., Moebius, J., Skinner, J., Doumbo, S., Doumtabe, D., Kone, Y., Dia, S., Kanakabandi, K., Sturdevant, D.E., Virtaneva, K., Porcella, S.F., Li, S., Doumbo, O.K., Kayentao, K., Ongoiba, A., Traore, B., Crompton, P.D.: Exposure-dependent control of malaria-induced inflammation in children. PLoS Pathog. **10**(4), 1004079 (2014)
32. Clark, I.A., Budd, A.C., Alleva, L.M., Cowden, W.B.: Human malarial disease: a consequence of inflammatory cytokine release. Malar. J. **5**(1), 85 (2006)
33. Ruktanonchai, N.W., Smith, D.L., De Leenheer, P.: Parasite sources and sinks in a patched Ross-Macdonald malaria model with human and mosquito movement: Implications for control. Math. Biosci. **279**, 90–101 (2016)
34. Ruktanonchai, N.W., DeLeenheer, P., Tatem, A.J., Alegana, V.A., Caughlin, T.T., Zu Erbach-Schoenberg, E., Lourenço, C., Ruktanonchai, C.W., Smith, D.L.: Identifying malaria transmission foci for elimination using human mobility data. PLoS Comput. Biol. **12**(4), 1004846 (2016)
35. Taylor, L.R.: Aggregation, variance and the mean. Nature **189**(4766), 732–735 (1961)
36. Crompton, P.D., Moebius, J., Portugal, S., Waisberg, M., Hart, G., Garver, L.S., Miller, L.H., Barillas-Mury, C., Pierce, S.K.: Malaria immunity in man and mosquito: insights into unsolved mysteries of a deadly infectious disease. Annu. Rev. Immunol. **32**, 157–187 (2014)
37. Rodriguez-Barraquer, I., Arinaitwe, E., Jagannathan, P., Kamya, M.R., Rosenthal, P.J., Rek, J., Dorsey, G., Nankabirwa, J., Staedke, S.G., Kilama, M., Drakeley, C., Ssewanyana, I., Smith, D.L., Greenhouse, B.: Quantification of anti-parasite and anti-disease immunity to malaria as a function of age and exposure. Elife **7**, 35832 (2018)

Additional Files

Additional file 1 — Malaria Therapy Patient Data
.txt file containing individual anonymized records including strain of parasite, method of introduction (mosquito bite vs. injection of asexual parasites), asexual parasitemia, gametocytemia, recorded oral temperature, treatment, and proportion of fed mosquitoes with established infection.

Additional file 2 — R Code
R code to read and analyze the malaria therapy data and produce the data-based figures in the manuscript.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- bmcarticle.bib
- bmcartbiblio.sty
- dls.bib