ESTIMATING THE DURATION OF PLASMODIUM FALCIPARUM INFECTION FROM TRIALS OF INDOOR RESIDUAL SPRAYING

WILSON SAMA, GERRY KILLEEN, AND TOM SMITH
Department of Public Health and Epidemiology, Swiss Tropical Institute, Basel, Switzerland

Abstract. We reviewed the use of simple mathematical models to estimate the duration of Plasmodium falciparum infection after transmission has been interrupted. We then fit an exponential decay model to repeated cross-sectional survey data collected from three historical trials of indoor residual spraying against malaria: one from two contiguous districts in Tanzania–Kenya (Pare Taveta) carried out in 1954, the others in West Papua (1953), and the Garki project in northern Nigeria (1972–1973). A cross-sectional analysis of these datasets gave overall estimates of 602 days (95% confidence interval [CI] = 581–625) for the infection duration in Pare Taveta, 734 days (95% CI = 645–849) in West Papua, and 1,329 days (95% CI = 1,193–1,499) for Garki. These estimates are much greater than the most widely quoted figures for the duration of untreated P. falciparum infections. Although these may be exaggerated because some reinfections occurred despite intensive vector control, prevalence was still decreasing when all these projects ended. Longitudinal survival analysis of the Garki data gave much shorter estimates of duration (186 days, 95% CI = 181–191), but effects of imperfect detection of parasites by microscopy severely bias these estimates. Estimates of infection duration for different age groups showed considerable variation but no general age trend. There was also no clear relationship between malaria endemicity and infection duration. Analyses of successive sampling from the same individuals with parasite typing are needed to obtain more reliable estimates of infection duration in endemic areas. Periods of several years may be required to evaluate long-term effects of interventions on malaria prevalence.

INTRODUCTION

The duration of untreated malaria infections is an important determinant of the level of transmission in endemic areas and determines the time scale of the effects on prevalence of reductions in malaria transmission. Its reciprocal, the clearance rate of infections (r) is a parameter in many mathematical models of transmission and immunity. Several models, for instance,1–3 assume that an important effect of natural immunity is to increase r. Despite this importance, few studies have attempted to estimate the duration of Plasmodium falciparum infections in natural populations, and there is also a dearth of information on the effects of naturally acquired immunity (or even of age) on infection duration.

Most detailed studies of the duration of P. falciparum parasitemia refer to malaria infections deliberately used for treatment of syphilis (therapeutic malaria) and report average infection durations of 200–300 days.4–9 Data such as these convinced most malariologists that untreated infections would generally persist for periods of this order, although occasionally P. falciparum infections are reported in returned tourists and immigrants from endemic areas whose last exposure was much further in the past.

Most of these data deal with induced malaria in non-immune subjects, and while these may be applicable in areas of low endemicity subject to epidemics, this does not necessarily reflect the duration in endemic areas where people are repeatedly reinfected. The most widely quoted figure of 200 days for the total duration of infection is that derived by MacDonald,10 who analyzed weekly parasitemia data recorded for a small group of individuals in Puerto Rico.11 However reassessment of the original dataset 25 years later led to the conclusion that infections with P. falciparum might still be patent some 30 months after the original infection and possibly longer.12

It is difficult to see how, in the absence of parasite typing data, duration of infection could be reliably estimated from field studies in areas with ongoing reinfection. However, the decay in the parasite rate when transmission is interrupted can be used to estimate the average duration of infection. A seminal paper in this field was that of Macdonald and Gockel.13 Using cross-sectional data from a number of attempts at eradication, they fitted a simple model of constant clearance rate to the parasite prevalence P

\[
\frac{dP}{dt} = -rP
\]

with solution:

\[
r = \frac{1}{t} \log \left( \frac{P_0}{P} \right)
\]

where \(P_0\) is the prevalence at time \(t = 0\), immediately prior to the interruption of transmission, and claimed that the results were broadly consistent with a duration of 200 days. Here log refers to the natural logarithm and \(P\) to the prevalence at time \(t\).

In this report, we make use of this model to estimate the total duration of P. falciparum infection after transmission has been interrupted, but improve on the basic model by allowing for recruitment of new individuals and for changes in age. Modeling the natural duration of infection is considerably complicated when anti-malarial treatment is available. We therefore fitted the models to P. falciparum prevalence data from three historical datasets from malaria research projects that preceded the introduction of primary health care providing anti-malarial treatments: the Pare-Taveta scheme,14 a pilot project in West Papua (Metselaar D, 1957. A Pilot Project of Residual Insecticide Spraying in Netherlands New Guinea: Contribution to the Knowledge of Holo-Endemic Malaria. Ph.D Dissertation. Leiden, The Netherlands: Leiden University), and the Garki project.15 We also test whether the duration of infection depends on the age of the host in these studies.

METHODS

Data sources. The Pare-Taveta Malaria Scheme, Tanzania–Kenya. To find out whether malaria transmission could be...
interrupted by the adoption of a certain technique of residual spraying, a large-scale trial was conducted in the Taveta sub-district of Kenya and the Pare district of Tanzania. The first round of residual spraying with Dieldrin and DDT was begun in July 1955 and five additional spraying cycles were carried out at approximately eight-month intervals, with an interval of at least two months between the end of one spraying round and the beginning of the next. The ecologic contrasts in the different parts of the whole study area led to the division of the study area in five distinct zones: the South Pare swamp villages, the South Pare roadside villages, the South Pare mesoendemic area, the Taveta forest, and the North Pare hyperendemic area. Repeated cross-sectional surveys were carried out at different times for each of the different zones, with irregular survey periodicity within each study zone. The total number of surveys carried out was not the same (between 7 and 9) for all the study sites within the study area. The population included in the treated area was approximately 5,300. The number of people examined at each survey was given as intervals, which probably indicated the minimum and maximum number within each age group examined during the entire sequence of surveys for each study zone (Table 1). 

This report, we use the midpoints of these intervals to approximate the number of samples examined.

**Malaria Research in Netherlands New Guinea (West Papua).** A similar experiment was carried out in West Papua. The experimental area (known as the “Lake Area”) was situated southwest of Hollandia (now called Jayapura) in the basin of the Sentani Lake. It consisted of two different areas, a mesoendemic region and a holoendemic one. Spleens were examined and Giemsa-stained thick blood films were prepared from a representative sample of the population in the lake area before spraying. The results of the two complete surveys made one (1955) and two years (1956) after the first application of insecticide (DDT) in the lake area are summarized by age groups in the original report (Metselaar D, 1957. *A Pilot Project of Residual Insecticide Spraying in Netherlands New Guinea: Contribution to the Knowledge of Holo-Endemic Malaria.* Ph.D Dissertation. Leiden, The Netherlands: Leiden University). Unlike the Pare Taveta scheme, the number of samples examined for *P. falciparum* was clearly stated in the original report. However in Table 1, we simply show the minimum and maximum number of people in the different age groups sampled in the different surveys. Al-

### Table 1

**Age distribution of the number of samples examined at different surveys in the three study areas**

<table>
<thead>
<tr>
<th>Age groups (years)</th>
<th>South Pare swamp villages</th>
<th>Taveta forest</th>
<th>South Pare roadside villages</th>
<th>South Pare mesoendemic area</th>
<th>North Pare hyperendemic area</th>
</tr>
</thead>
<tbody>
<tr>
<td>&lt; 1</td>
<td>8–85</td>
<td>15–34</td>
<td>38–68</td>
<td>14–39</td>
<td>43–82</td>
</tr>
<tr>
<td>&lt; 2</td>
<td>15–63</td>
<td>16–40</td>
<td>21–66</td>
<td>7–36</td>
<td>50–68</td>
</tr>
<tr>
<td>5–9</td>
<td>36–262</td>
<td>64–105</td>
<td>144–328</td>
<td>69–165</td>
<td>224–349</td>
</tr>
<tr>
<td>15–19</td>
<td>48–121</td>
<td>27–42</td>
<td>41–115</td>
<td>18–49</td>
<td>44–121</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Age groups (years)</th>
<th>Holoendemic part</th>
<th>Mesoendemic part</th>
</tr>
</thead>
<tbody>
<tr>
<td>0–2 months</td>
<td>22–42</td>
<td>24–51</td>
</tr>
<tr>
<td>3–5 months</td>
<td>17–34</td>
<td>20–35</td>
</tr>
<tr>
<td>6–8 months</td>
<td>15–53</td>
<td>23–31</td>
</tr>
<tr>
<td>9–11 months</td>
<td>11–21</td>
<td>19–46</td>
</tr>
<tr>
<td>1 year</td>
<td>74–122</td>
<td>65–141</td>
</tr>
<tr>
<td>2 years</td>
<td>61–95</td>
<td>75–113</td>
</tr>
<tr>
<td>3–5 years</td>
<td>179–208</td>
<td>171–318</td>
</tr>
<tr>
<td>6–8 years</td>
<td>174–209</td>
<td>271–311</td>
</tr>
<tr>
<td>9–11 years</td>
<td>74–157</td>
<td>120–249</td>
</tr>
<tr>
<td>12–14 years</td>
<td>76–90</td>
<td>84–106</td>
</tr>
<tr>
<td>15–24 years</td>
<td>132–234</td>
<td>95–296</td>
</tr>
<tr>
<td>25–34 years</td>
<td>147–257</td>
<td>82–293</td>
</tr>
<tr>
<td>35–44 years</td>
<td>97–221</td>
<td>93–224</td>
</tr>
<tr>
<td>&gt; 44 years</td>
<td>77–168</td>
<td>66–154</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Number examined</th>
<th>&lt; 1</th>
<th>&lt; 2</th>
<th>2–4</th>
<th>5–9</th>
<th>10–14</th>
<th>15–19</th>
<th>20–39</th>
<th>≥40</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pare-Taveta</td>
<td>96</td>
<td>71</td>
<td>182</td>
<td>389</td>
<td>147</td>
<td>114</td>
<td>873</td>
<td>491</td>
</tr>
<tr>
<td>West Papua</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Garki</td>
<td>96</td>
<td>71</td>
<td>182</td>
<td>389</td>
<td>147</td>
<td>114</td>
<td>873</td>
<td>491</td>
</tr>
</tbody>
</table>

*The number of people examined for the Pare-Taveta and West Papua studies are given as intervals. For instance, 22–42 for the West Papua study indicates that of the four surveys carried out in this area, 22 and 42 subjects of the corresponding age group were examined at two of these surveys while at the remaining two surveys, the number of subjects lies in the interval 22–42. For Garki, only the numbers at baseline are given. There was not much variation in these numbers in the eight subsequent surveys.*
though the time interval between the surveys was also irregular, the surveys were carried out at the same time in the mesoendemic and holoendemic regions of the study area.

The Garki Project. The data used here were collected from the Garki project, an intensively monitored trial of malaria control in northern Nigeria carried out in 1969–1976. In contrast to the Pare-Taveta and West Papua datasets, in Garki the individuals were identified and parasitologic status could therefore be analyzed in the same individuals longitudinally. From April 1972 to October 1973, villages in three concentric areas were treated with one of three control strategies (A1, A2, and B), which are described in detail by Molineaux and Gramiccia.15 Since the objective of this paper is to estimate the duration of P. falciparum infections in untreated individuals, we analyze only data collected from the six sentinel villages in the intervention area where there was no mass drug administration (area B). Residual indoor spraying with insecticide propoxur for three or four rounds, at intervals of approximately two months, was applied to this area both before and during each of two main transmission seasons (1972 and 1973). Eight surveys of the entire population of these villages were carried out prior to the intervention, and an additional eight surveys were carried out during the intervention. We consider the data from survey 8 as comprising a baseline for our analysis, and analyze changes in parasitologic status during surveys 9–16. Table 1 shows the number of people in each age group that were examined at baseline. There were very little changes in these numbers in the subsequent surveys.

MODELS

The basic model is a differential equation model describing the rate of change of the proportion positive with time. The underlying assumptions are 1) that transmission was completely interrupted after the comprehensive application of residual insecticides, 2) that there was no drug treatment so loss of infections were due to spontaneous dying out of parasites or immune response of the host, and 3) that the proportion of those positive decreases in an exponential manner with time. For example, examination of Figure 1 shows little evidence for transmission after the beginning of the intensive vector control measures (which started around the third survey), and the general picture we get from the graphs is prevalence continually decreasing after the last pre-spraying surveys, so our assumption is quite reasonable.

We assume that there are k age groups at baseline with each individual being assigned into one of age groups i, l ≤ i ≤ k. Let P_{i,t} denote the proportion positive at time t among those in age group i at time t_0 (i.e., at baseline). We assume that this proportion is adequately described by the equation

\[ \frac{dP_{i,t}}{dt} = -r_i P_{i,t} \]

where \( r_i \) is the clearance rate at time t for individuals in age group i at baseline. If \( P_{i,0} \) is the proportion positive at baseline then it follows that

\[ P_{i,t} = P_{i,0} e^{-r_i(t-t_0)} \]

The total duration of infection was estimated using two different methods of analysis, namely, repeated cross-sectional and longitudinal survival analysis. Both methods were used to analyze the data from the Garki (area B) study, while only the repeated cross-sectional analysis method was applied to the other two datasets.

Analysis of infection duration from repeated cross-sectional data. Garki data. In the Garki data, the exact age of each individual was known and the individuals were identified and followed-up longitudinally. However, for the purpose of comparison with the other two datasets, we start by analyzing this dataset as if they were collected from unlinked cross-sectional surveys, and assign individuals to different age groups using the age groupings in the Pare-Taveta study. We fitted equation 4 with separate estimates for \( P_{i,0} \) and a common estimate for \( r_i \) (i.e., \( r_i = r \) for all the age groups). To allow for effects of age on infection duration, we also fitted equation 4 with separate estimates of \( P_{i,0} \) and of \( r_i \) for each age group and to test for a linear trend in the effect of age on r we fitted the model with \( r_i \) in (4) substituted with the following age dependent term

\[ r_i = r_0 + r_1 a_{i,0} \]

where \( a_{i,0} \) is the mid-age of group i at baseline.

To allow for random variation, we assume a binomial error function for the parasite prevalence, i.e.,

\[ X_{i,t} \sim \text{Bin}(n_i(t), P_{i,t}) \]

where \( X_{i,t} \) is the number of positive samples in age group i at time t and \( n_i(t) \) is the total number of samples examined in age group i at time t.

Pare-Taveta and West Papua data. We take into consideration the fact that the population sampled in each age group varies for each survey by attempting to capture in our model the proportions of the population within each age group at a later time that were in the different age groups at baseline. For this we make the additional assumption that the age distributions are approximately uniform. That is, the number of individuals within each age group is proportional to the width of the age group. We make use of this assumption in deriving equation 7.

First, we give a brief explanation, in non-mathematical terms, of the scenario that the equations developed below attempt to capture. Consider, for example, three distinct cohorts (or age groups) at baseline, i.e., at time \( t_0 \) (Figure 2). We assume that the prevalence in each age group decreases in an exponential manner with time (given by equation 4), with the possibility of a different rate of decrease for the different age groups. Since different cohorts were studied at different surveys, we wish to estimate for example at a later time \( t_1 \) in Figure 2) the proportion of a newly recruited cohort 2 who were in cohort 1 at baseline (\( t_0 \) given by equation 7). Looking at Figure 2 again, this proportion is given by the segment AB. We estimate this proportion and take it back to its original age group at baseline. It is clear that when the time interval between the baseline and any subsequent survey is long enough, some individuals within a specific cohort at baseline will move to two or more subsequent cohorts. From the example given in Figure 2, we find that at time-point \( t_2 \), a certain proportion (AC) of cohort 1 has moved to cohort 2 while another proportion (CD) has moved to cohort 3. We sum up these proportions, take them back to their original age groups (i.e., the age group they belonged to at baseline), and apply the corresponding exponential decrease that was as-
sumed for this group to start with. This is summarized by equation 8.

Let \( L_i \) and \( U_i \) denote respectively the lower and upper limits of age category \( i \). Then the probability that any individual, \( x \), will belong to age group \( j \) at time \( t \) given that he was in age group \( i \) at time \( t_0 \) (or equivalently the proportion of age group \( j \) at time \( t \) that were in age group \( i \) at time \( t_0 \)) is given by

\[
Pr(x \in j,t | x \in i,t_0) = \max \left( \frac{\min(U_i + t - t_0, U_j) - \max(L_i + t - t_0, L_j)}{U_j - L_j}, 0 \right)
\]

and it follows from this and equation 4 that the expected value of the prevalence in age group \( j \) at time \( t \), \( E(P_{j,t}) \) is

\[
E(P_{j,t}) = \sum_i Pr(x \in j,t | x \in i,t_0) P_{i,t_0}
\]

FIGURE 1. Observed changes in prevalence of Plasmodium falciparum parasitemia with time for different age groups within the Pare-Taveta project. The figure refers to four of the five sites in the project. The corresponding data for the fifth site are shown in Figure 5. The arrows indicate the onset of spraying. 0-9 years: 0–11 months (●), 12–23 months (■), 2–4 years (□), 5–9 years (●); ≥ 10 years: 10–14 (∇), 15–19 (◆), 20–39 (◇), ≥ 40 (▲).
being infected at time \( t \), for the parasite prevalence in both areas, i.e., report results obtained by assuming a binomial error function repeated cross-sectional analysis. Open circles which is a function of the unknown parameters \( r_i \), times (from one age group to another between survey rounds at different observations for each individual were independent of each other were either negative or absent. We assumed that the observations for each individual were independent of each other (i.e., an individual positive at two consecutive surveys was treated as two separate counts and so on) and we estimated the proportion \( P \) infected at time \( t + 1 \), \( (I_{t+1}) \), conditional on being infected at time \( t \), \( (I_t) \), as follows

\[
E(P_{jt}) = \sum \Pr(x \in j,t|I_t)P_{jt} \tag{8}
\]

which is a function of the unknown parameters \( r_i \), \( P_{jt} \). We report results obtained by assuming a binomial error function for the parasite prevalence in both areas, i.e.,

\[
P_{jt} \sim \text{Bin}(n_{jt}, P_{jt}) \tag{9}
\]

The estimates for the \( r \) values were obtained in a similar manner as described in this section. In addition, analyses were carried out separately for different sites within the study areas to test for effects of malaria endemicity on infection duration.

**Longitudinal survival analysis of infection duration.** Second, we fitted an exponential model to the survival times of the infections in the Garki dataset. We studied only changes in the parasitologic status of individuals who were present and positive at baseline (eighth survey) until the survey when they were either negative or absent. We assumed that the observations for each individual were independent of each other (i.e., an individual positive at two consecutive surveys was treated as two separate counts and so on) and we estimated the proportion \( P \) infected at time \( t + 1 \), \( (I_{t+1}) \), conditional on being infected at time \( t \), \( (I_t) \), as follows

\[
\Pr(I_{t+1} | I_t) = e^{-rt} \tag{10}
\]

We also took into account random variation by assuming a binomial error function as follows

\[
X_{i,j} \sim \text{Bin}(n_{i,j}, P) \tag{11}
\]

where \( X_{i,j} \) equals the total number positive at all surveys \( j \) \((8 \leq j < i)\) and positive at survey \( i \), while \( n_{i,j} \) equals the total number positive at all surveys \( j \) \((8 \leq j < i)\) and present at survey \( i \), \((9 \leq i \leq 16)\).

The model parameters were estimated using WinBUGS version 1.3, and assuming gamma priors for the \( r \) values. The quoted results are based on samples of 29,300 values from the posterior densities, following a burn-in of 30,000 iterations.

**RESULTS**

The model for repeated cross-sectional data gave good fits to the data (Figure 3). Overall estimates for the total duration of infection across all age groups were 602 days (95% confidence interval [CI] = 581–625) for Pare-Taveta, 734 days (95% CI = 645–849) for West Papua, and 1,329 days (95% CI = 1,193–1,499) for Garki (area B).

Prevalence appeared to decrease faster in the younger age groups than the older ones (Figure 1). However, this was due to the fact that newborns recruited at the subsequent surveys came in at a time when the insecticide spraying already had a massive effect in reducing transmission. Extension of the models to allow variation between age groups in the estimates of the duration of infection in each of the study areas did not indicate any clear age trend. The age pattern for the estimates in the West Papua was especially noisy, probably due to relatively small size of this dataset (only four surveys were conducted in this trial). However, the clearance rate increased modestly with age in all three areas \( (r_{i} = 0.0125 \text{ per year} [95\% \text{ CI} = 0.0075–0.01776] \text{ for West Papua, } r_{i} = 0.0043 [95\% \text{ CI} = 0.0021–0.0066] \text{ for Pare-Taveta, and } r_{i} = 0.0074 [95\% \text{ CI} = 0.0047–0.0103] \text{ for Garki}).

Figure 4 shows the fit of the model to the data and also the predicted prevalence at each time point in the cohort initially present at baseline. This is shown (Figure 4) explicitly for the eight different age groups in one of the sites (Pare swamp) in the Pare-Taveta study. There is a clear difference in the rate

**Figure 2.** Schematic representation of the movement of cohorts from one age group to another between survey rounds at different times \( t \). A-F are arbitrary ages.

**Figure 3.** Observed and predicted prevalence curves for the pooled data in the West Papua study and the Garki (area B) study using the repeated cross-sectional analysis. Open circles indicate observed values while solid lines indicate predicted values. a, predicted prevalence curve for Garki obtained by using equation 4. b, predicted prevalence curve for West Papua obtained by using equation 8.
of decrease in prevalence in the younger age groups (especially in the first age group of 0–11 months) because newborns were directly recruited into these age groups. The difference between these two curves become less visible as the age groups become older because at subsequent surveys the newborns were not old enough to attain these age groups and thus affect the observed prevalence within them. The model fit to the Garki data was also good (Figure 5).

We also obtained separate estimates for the total duration of infection for the different zones in the West Papua and
Pare-Taveta studies. The duration of infection for the pooled data in the holoendemic part of the study area in West Papua was estimated to be 699 days (95% CI = 604–824) and that for the pooled data in the mesoendemic part was 820 days (95% CI = 639–1,131). The estimates for the pooled data in each of the five different sites in the Pare-Taveta study are shown in Figure 6, while the predicted age prevalence curves for the five sites are shown in Figure 7. The results in Figure 7 suggest that Taveta forest was a highly endemic area, while the South Pare mesoendemic area was the least endemic area.

A comparison of Figures 6 and 7 shows that there is no general relationship between infection duration and malaria endemicity. This was confirmed by performing a Spearman’s rank correlation analysis, which gave a correlation coefficient $R$ of 0.30 ($P = 0.62$).

In contrast to the analyses of repeated cross-sectional data, the longitudinal survival analysis of the Garki data gave an overall estimate of 186 days (95% CI = 181–191), which is far lower than all the other estimates quoted above but closer to most of the values in the literature.

FIGURE 5. Observed and predicted prevalence curves for eight different age groups in the Garki (area B). Open circles indicate the observed values while the solid lines indicate the predicted values. Prediction is based on the repeated cross-sectional analysis (using equation 4).
DISCUSSION

We have revisited three historical datasets from trials of indoor residual spraying, two of them comprising unlinked repeated cross-sectional surveys, and the third, a longitudinal cohort study, and used an exponential decay model to estimate the duration of infection \((1/r)\) for untreated malaria infections in endemic areas.

The overall estimates for the duration of infection from the analysis of the repeated cross-sectional data are similar in the three areas, but are much higher than the most widely quoted values derived from analyses by Macdonald and Göckel two generations ago.\(^{10,11}\) There was no obvious relationship between infection duration and the endemicity of malaria (Figures 6 and 7). Our results suggest a waiting time of at least 2–3 years is needed before evaluating the effect of interventions that suppress transmission without actively clearing parasites such as insecticide treated nets, indoor residual spraying, and mosquito source reduction.

Among the sites for which Macdonald and Göckel\(^{13}\) presented curves for the decrease in parasite rates when transmission was interrupted, the Pare Taveta and West Papua sites showed the lowest rates of decrease. According to Macdonald and Göckel,\(^{13}\) this was because transmission was not completely interrupted in these sites. However, although there were some infections in children born after the start of spraying, the comprehensive application of insecticide ensured that these were very few in number (Figures 1 and 4). This indicates that post-intervention new infections were rare and cannot account for the persistently high prevalence (Figure 1).

A number of studies in endemic areas have reported that infections are of relatively short duration in very young children. For example, Walton\(^{17}\) found that the average duration of infection in infants in Freetown, Sierra Leone was slightly more than three months at a time when there was relatively little transmission there. More recently, it was estimated that the duration of infections with parasites belonging to the merozoite surface protein 2 FC27 allelic family increased with age using data collected from Tanzanian children 6–30 months old.\(^{18,19}\) An increase in duration with age during the first two years of life has also been reported in a study with Ghanaian children.\(^{20}\)

We chose the Pare Taveta, West Papua, and Garki data for re-analysis because all three allowed us to analyze the duration of infection by age. The Pare-Taveta and West Papua data showed a faster initial decrease in prevalence in the youngest age group than in the older ones. However, Macdonald and Göckel\(^{13}\) suggested that infection duration appeared shorter in young children only because the data had been analyzed inappropriately, and that a cohort analysis should have been carried out. This is because straightforward analysis of the decrease in parasite prevalence with time after transmission is interrupted fails to allow for the recruitment of new uninfected individuals (newborns). Our new analysis allows for this effect, and indeed we find that after this adjustment there is no indication in these datasets that parasites are cleared faster in the youngest children. Indeed, there seems to be a slow decrease in duration with age.

While we agree with Macdonald and Göckel\(^{13}\) about the biases due to recruitment of unexposed newborns, all our estimates of clearance rates, except those from the cohort analysis of the Garki data, remain much lower than those reported based on other datasets quoted by Macdonald and Göckel.\(^{22}\) Part of the reason for this may be that few of the datasets they used are from surveys in endemic communities in the absence of mass treatment. For instance, the data they refer to for the eradication of *Anopheles gambiae* from Brazil\(^{21}\) appear to be incidence figures for clinical cases.

The explanation of why the estimates of duration from unlinked data are lower than those from the cohort analysis is that the latter systematically underestimates duration by treating temporarily subpatent infections as though they had been cleared.\(^{22}\) This can be illustrated by a typical profile of parasite density during follow-up of a malaria patient (Figure 8). The patient was inoculated at time A and parasites were cleared at some time H after the last day G on which parasites were detectable. The true duration of infection is thus H-A. However, longitudinal studies that do not allow for
imperfect detectability, such as that of Macdonald\cite{10} and our own longitudinal Garki analysis, give estimates of the duration of parasitic episodes (either D-C or F-E). Estimates of durations from longitudinal data can be even shorter if sampling is frequent because of sequestration during the 48-hour cycle of the parasite. Using longitudinal microscopy data from blood smears collected at very frequent intervals among inhabitants of a single village in Papua New Guinea, Bruce and others obtained an estimate of 3–27 days for the duration of parasitic episodes of asymptomatic \textit{P. falciparum} infections.\cite{23}

In contrast, the unlinked analysis is not subject to this source of bias. If detectability is age- and time-independent and the true process follows the dynamics described by equation 1, then the observed process will also follow these dynamics. This implies that the model for unlinked data remains the same whether or not detectability is taken into account. It is therefore appropriate for estimating the total duration of infection (H-B in Figure 8).

Our study thus provides convincing evidence that malaria infections in endemic areas actually persist on average for much longer than Macdonald\cite{10} claimed and that it is the cohort analysis that can be biased. It also suggests that duration is not very dependent on exposure in highly endemic communities. However, studies where transmission is interrupted cannot tell us what is the infection duration when superinfection is occurring. Moreover, very few of the individuals were in the youngest age categories, so we cannot be confident that very young children and malaria-naïve individuals have similar duration to people with more exposure. A critical assumption of this paper was that residual transmission was negligible. A key question that could not be adequately addressed using these datasets is that by how much is the duration of infection overestimated due to the fact that reinfection is not accounted for. However, a graphic display of the raw data provides convincing evidence of a general exponential decay pattern, therefore indicating that the results are not greatly altered, although it is possible that a better fit could be obtained with more complex decay models. For example, in the Garki project, it was shown that after the spraying program, reinfection did continue at a level approximately one-sixth its original value.\cite{15} Using the cross-sectional analysis for the Garki data, we obtained an estimated infection duration of 1,329 days, suggesting that even if this low level of reinfection was accounted for, the estimated duration of infection will still be much longer than those that have been thought of in the past. The average durations that we estimated are also much longer than those seen in malarial therapy patients in whom the whole course of infection was observed.\cite{9}

Field-based parasite typing studies now make it possible to study duration of infection in the presence of superinfections. Thus, it was found\cite{24} that the mean duration of episodes of positivity (i.e., D-C or F-E in Figure 8) for the same \textit{P. falciparum} genotype to be approximately 60 days in Papua New Guinean children. Typing studies have found asymptomatic infections persisting for more than 12 months in eastern Sudan.\cite{23} It was also found that a single parasite genotype of \textit{P. falciparum} asymptomatic infections could persist for as long as 40 weeks.\cite{26} However, few typing studies have attempted to calculate population averages of overall duration.

Recently, a model that allows for imperfect detectability has been proposed\cite{19} that analyzes infection dynamics in longitudinal data where parasites have been typed by a polymerase chain reaction and in the presence of new inoculations. These analyses covered only a narrow age range of hosts. Further analyses of longitudinal parasite typing data are needed that cover the whole age distribution and from areas of different endemicity.

Received May 27, 2003. Accepted for publication October 22, 2003.

Acknowledgments: Wilson Sama is the recipient of a stipend from the Stipendiumkommission of the Amt für Ausbildungsbeiträge of the Canton of Basel. We thank Klaus Dietz and Louis Molineaux for their helpful comments on an earlier draft of the manuscript, and Penelope Vounatsou for assistance with the statistical software. David Bradley, Louis Molineaux, and Paulette Roze helped in the recovery of data from old records.

Authors’ address: Wilson Sama, Gerry Killeen, and Tom Smith, Department of Public Health and Epidemiology, Swiss Tropical Institute, Socinstrasse 57, Postfach, CH-4002 Basel Switzerland, Telephone: 41-61-284-8282 or 41-79-233-6748, Fax: 41-61-271-7951, E-mail: wilson.sama@unibas.ch.

REFERENCES


