Reviewer's report

Title: Time-to-infection by Plasmodium falciparum is largely determined by random factors.

Version: 3  Date: 9 October 2014

Reviewer: Lisa LJW White

Reviewer's report:

Major Compulsory Revisions:

1. This approach combines two levels of organisation (population level transmission dynamics and individual level infection dynamics) in a very elegant mathematical framework. It is used to consider the implications for interpreting the time to infection with Plasmodium falciparum malaria derived from follow-up data. Two methods of detection are considered simultaneously: microscopy and PCR where the difference in the detection thresholds for within-host parasite loads is accounted for. It is my opinion that with some adjustments to the style and a little more consideration of the context that this research will be of interest to a very diverse readership.

2. The title “Time-to-infection by Plasmodium falciparum is largely determined by random factors” is supported by the analysis in the manuscript. However, I would like to know whether this is only the case for natural immunity or whether this is also the case for vaccine-induced immunity. Natural immunity is characterised very nicely as having no infection blocking action but rather acts to reduce parasite growth upon infection. This is a very nice result, but are the authors concluding that this is the case for vaccines as well?

3. The method demonstrates that immunity (assumed to increase with age) most likely manifests as a reduction the parasite multiplication rate upon infection rather than being protective against infection. This then differs significantly from the assumption that the RTS,S vaccine provides some protective efficacy against infection even if short-lived. Moreover, the design of the study from which the data are derived differs fundamentally from the phase III RTS,S vaccine trial design in that participants are sampled weekly regardless of symptoms in the time to infection study whereas in the RTS,S phase III vaccine trial, individuals are only sampled if they report a clinical case. With this key difference in mind, my first suggestion would be for the authors to find a way to use the time-to-infection study to simulate passive case detection and explore the implications of the analysis of this type of data. This would be possible if they have access to information on clinical cases which they could use retrospectively as they did with the PCR data. I think that such an analysis will highlight how typical vaccine trial designs for malaria fall short of providing sufficient information to understand the way in which both the vaccine-induced and natural immunity differ and interact with each other. This should be a caveat by which current trial data is interpreted and a contribution to the design of future trials of
this and other malaria vaccines.

4. In the abstract the authors mention how important the identification of protective immune response is for vaccine development. I think that they could simulate a protective immune response, which is clearly not provided by natural immunity to demonstrate how the time to infection observations would change for detection by both microscopy and PCR. Many RTS,S vaccine trials demonstrate differences in time to infection (by both active and passive case detection) between the vaccine and control groups. It would be extremely useful if this approach were used to assist in the interpretation of such results.

5. I have reproduced the model results using a simple set of ordinary differential equations in the following format:

\[
\frac{d}{dt}(\text{susceptible}[1..N,1..2]) = -\lambda[i]*\text{susceptible}[i,j]
\]

\[
\frac{d}{dt}(\text{infected}[1..N,1..2]) = \lambda[i]*\text{susceptible}[i,j] - \delta[i,j]*\text{infected}[i,j]
\]

\[
\frac{d}{dt}(\text{detected}[1..N,1..2]) = \delta[i,j]*\text{infected}[i,j]
\]

\[
\text{out}[1..N,1..2] = 1 - \text{detected}[i,j]
\]

\[
\delta[1..N,1..2] = \logn(\text{PMR}[i])/(2*(\logn(C[i,j]) - \logn(C0)))
\]

\[
\lambda[1..N] = \text{if time}<\tau \text{then } 0 \text{ else } \text{EIR}*x[i]/365
\]

with \(N=4\) age groups and 2 detection methods being microscopy and PCR. Then the function \(\delta\) relates the rate of detection to the average parasite multiplication rate \(\text{PMR}\) for each age group. The detection thresholds in each age group, \(i\), and for each detection method, \(j\), are given by \(C[I,j]\). The initial parasite load is given by \(C0\). The rate of infection is switched on after a lag time, \(\tau\). This makes me very confident that the results presented are sound. Some diagrams of the process being modelled along with the provision of an additional file showing the steps towards obtaining the equations in the manuscript would be useful for a more general reader.

6. Throughout the manuscript, the authors should clarify which parameters and are fixed in each analysis and which are estimated directly from the data or from fitting models (and which models/equations were used in each case).

7. The authors should consider the positive correlation between age and biting rate and therefore infection rate and how including this relationship might affect their results.

8. The authors should consider also the relationship between age (and therefore blood volume) on detection thresholds for both microscopy and PCR. That is, for the same parasite load, the concentration of parasites in children would be higher than in adults leading to detection at lower absolute loads in children compared to adults.

9. In general I think that there are too many very busy figures. Some of these should be relegated to additional material for completeness and a select few very simplified versions should be used in the main paper to illustrate key results. For example, figure 5 A to E is not necessary. It would be much better to pick one age group and show 5 lines for increasing EIR on a single graph. Graphs F to J
could be summarised by plotting EIR on the x-axis and the corresponding mean early and late PMR values on the y-axis. Same for K to O. It is still not clear to me what P to T is showing me. The green area that is changing is really small compared to the blue and red lines.

Minor Essential Revisions:
10. The numbering of the equations should be corrected
11. Results first paragraph: it looks like the time to detection was derived directly from the data an then the rate of new infections was derived from fitting a simple exponential curve. Is that correct?
12. Line 255 to 265: was the equation on line 186 used to estimate the PMR in each case? If so, please state this more clearly.
13. Line 383 to 399: I think the authors are saying here that by fixing the infection rate at different assumed values will lead to different estimates of the PMR.
14. Figure 2: please state which values are derived from the data and which are estimated from the model (eg I think “delay in weeks” is derived from the data and PMR is estimated).
15. Figure 5: do these graphs represent the predictions using microscopy?
16. Figure 5 F to O: is dashed late infection and solid early infection?
17. Figure 4 C to F: It is not clear to me how PMR can be estimated separately for both PCR and microscopy as I understood from the first equation that this is calculated by comparing the time to infection within an individual as detected by each of these methods.
18. Line 579: change “control the” to “control than the”

**Quality of written English:** Acceptable

**Statistical review:** No, the manuscript does not need to be seen by a statistician.

**Declaration of competing interests:**

Do you have any non-financial competing interests in relation to this paper? Yes.

I am collaborating with some of the authors of this paper on a similar project for the analysis of P. vivax data.