Reviewer's report

Title: Temporal changes in haematocrit following artemisinin-based combination treatments of uncomplicated falciparum malaria in children

Version: 2  Date: 16 March 2015

Reviewer: Sue Lee

Reviewer's report:

Discretionary Revisions (which are recommendations for improvement but which the author can choose to ignore)

1. Should the line numbers be consecutive to the entire document rather than starting from 1 on each page?

Minor Essential Revisions (such as missing labels on figures, or the wrong use of a term, which the author can be trusted to correct)

1. Could the authors please provide an explanation as to why it has taken 4 years to publish these results?

2. P. 2, Line 27. “Increases in haematocrit is common…” should read “Increases in haematocrit ARE common…”

3. P. 3, Line 12. For “…haematocrit in the individual…”, omit “the”.

4. p. 5, Line 10. Include “treatment” between “before” and “and”.

5. P. 6, line 28. Please provide a reference for this.

6. P. 7, lines 10-11. It is my understanding that the Mann Whitney-U test and the Wilcoxon rank-sum test are actually the same thing?

7. Figure 1. Please add the n for each pattern on this figure.

Major Compulsory Revisions (which the author must respond to before a decision on publication can be reached)

1. Table 2a and corresponding results in text (p. 8). Please make clear (as a footnote in the table and also in the text) what is being compared here. Was the p-value obtained by comparison across all there groups? In the text, it implies that children from pattern 6 were compared against children from patterns 1 and 2 combined. The results state that the children in pattern 6 weighed less but the data presented in the table suggest that children in pattern 2 actually had the mean lowest weight. This discrepancy should be checked/corrected. Also, as the text refers to a comparison by proportion of those with hyperpyrexia, it would be useful to include the actual percentages in the table as well (rather than just the numbers). What purpose does the “ALL” column serve? In my opinion, this
column is not useful and is not needed. And finally, it would be useful to report the percentages for the parasitaemia > 100,000uL and again, also to specify which test was used to obtain the p-value... presumably it was ANOVA? Though it seems surprising that there was no significant difference between the groups as the children in group 6 appear to have consistently higher GM parasitaemia.

2. Figure 2 and corresponding results in text (p. 8). It appears that these graphs have been made by simply summarizing all values (within each group) at each time point. However, as these data are longitudinal and each patient has repeated measurements, the estimates should be adjusted for autocorrelation. Therefore, the DAFH should be calculated for each patient as the change in HCT from baseline or the fractional change in HCT from baseline.

3. Table 3. Add a column indicating which pattern group the patient comes from. Perhaps this table and the corresponding Figure 3 can go into supplementary materials? In any case, there are a large number of tables and figures and Figure 3 could be dropped.

4. p. 9, lines 6-22. This section needs more clarification please. The section is entitled “Comparison of children with fall in haematocrit ≥5 units and < 5 units following treatment” so why has day 7 been excluded from the table? How is the “fall” calculated? Is the “fall” calculated between each time point or from baseline? If calculated between each time point, are some patients included more than once in this table, i.e., if a patient dropped more than 5 units between day 7 and 14 and then again between days 14 and 21? Or, if calculated from baseline, is the patient only included at the first time point that they fell 5 or more units? On line 12, I get a different chi-squared value (and slightly different p-value, though still highly significant) using the numbers provided in the text (STATA command: tabi 57 55 \ 165 158, chi e). I also get a different statistic for day 14 (tabi 29 32\ 23 33, chi e) so please double check these calculations. If children are represented more than once in this table, then the data are not independent and it may not be appropriate to use the chi-squared statistic.

5. Table 5. Why does this table include only n=50 in the ≥5 unit change group and n=38 in the <5 unit change group? According to the text, it seems there should be n=122 and n=213, for each group, respectively, if counting days 14 to 28. Otherwise, according to table 4, the totals should be sum of each row (though it remains unclear why day 7 has been excluded if, in fact, you are summarizing the “Clinical, parasitological and other characteristics of patients with falls in haematocrit # 5 and < 5 units following treatment with artesunate- amodiaquine or artemether-lumefantrine”). It seems that the results from this table could be summarized in the text and the table dropped.

6. P. 10, lines 7-9. Ideally this time to recovery from anemia is best analysed using a log-rank test. Medians should reported as the distribution tends to be skewed. In addition, as means are reported here, should the reader assume that the time to recovery was analysed using a ttest? There appears to be greater variation in the late monophasic fall group which might violate the assumption of equal variances required for a valid ttest.
7. p. 10, lines 15-16. Also please check the equality of variances here if the t-test was used for comparison.

8. p. 11, line 9. Suggest drop Figure 7… it should be enough to report the rho and p-value.

9. p. 11, lines 10 – 16. While the text states that the limits of agreement were “narrow”, -9.8 days to 12 days does not seem narrow to me (nor does -11 to 11 days)… nor do I understand how to interpret a “negative” day. Please clarify. The Bland-Altman method depends on some assumptions about the data: Was the distribution of the differences checked by plotting a histogram? Also that the mean and SD of the differences are constant throughout the range of measurements. From the plots in Figure 8, this does not seem to be the case.

10. Please discuss some limitations of the study in the discussion section and how these limitations might impact the findings reported herein.

Level of interest: An article whose findings are important to those with closely related research interests

Quality of written English: Acceptable

Statistical review: Yes, and I have assessed the statistics in my report.

Declaration of competing interests:

I declare that I have no competing interests.