Reviewer's report

Title: Using pay for performance incentives (P4P) to improve management of suspected malaria fevers in rural Kenya; a cluster-randomized controlled trial

Version: 1 Date: 8 July 2015

Reviewer: Joshua Yukich

Reviewer's report:

This manuscript describes the results of a CRCT of a pay for performance intervention delivered based on indicators of quality malaria diagnosis and prescribing behavior in Kenya. Overall, the study is interesting and the topic is likely to be of interest to readers of this journal. There are however a number of issues with the manuscript which should be addressed prior to publication.

Major Compulsory Revisions:

1) In the abstract the number of randomization units differs from that stated in the methods. I'll discuss later the need to identify if what is presented is an ITT analysis of a per-protocol analysis but I tend to believe that the abstract should clearly focus on the primary outcome in an ITT analysis. Its not clear that it does.

2) The primary outcome is described both in the abstract and throughout the manuscript as "at the end of the intervention period." its not clear what this means, or if it is the appropriate analysis. Based on the methods and the power calculation it would seem that the intended analysis would be over the entire intervention period, but this is perhaps not what was done. The authors need to clarify and present the results of the trial with the intended primary analysis. This needs to coincide with the description of the primary outcome, which should be clear in the abstract.

Intro/Background:

3) There are no citations for the statement that the prescription of ACTs to malaria negative patients will jeopardize the efficacy of these drugs or accelerate the time course to development of resistance. (I realize that this is "received wisdom" in this area) but it would be good to see the authors back this story up with strong evidence that this is likely to be the case, otherwise, we just repeat stories that we hear (this is also a minor comment).

4) I think the authors need to emphasize that this intervention is not particularly going to change incentives at the clinician level (except by changes in incentives for facility managers) and definitely not at the patient level. The concept of how the intervention should change these behaviours could be better deliniated in the background section.

5) The very end of the intrduction makes a distiction between outcomes
measured at the cluster level and those at the individual level (but it does not make clear which are which).

Methods:

6) again a minor comment, but would be nice to have some citations for the description of the study sites. Especially as the authors make a distinction between high and low transmission areas, but provide little info on how this is made.

7) The inclusion exclusion criteria clearly must have excluded hospitals and low level facilities (at least if they had diagnostic capacity). This should be clearly stated. It is not.

8) I find the dropping of one site which appears to have met the inclusion criteria strange. I understand why the authors did this, but it appears to compromise the randomization and the analysis if this is intended to be an ITT analysis. The authors need to address this limitation much more clearly, I would prefer to see an analysis with this information included as well. Further no including a facility that met inclusion criteria but then didnt adhere to the intervention as desired appears to preclude the possibility of a ITT analysis. The authors need to clearly state this in the manuscript up front.

9) the authors state in the methods that the incentives were designed to be offset by ACT savings, but there is mixed evidence on this when the ACT savings come as the result of changes in diagnostics (e.g. introduction of RDTs, or in this case rollout of a QA program). These costs were clearly not included and these statements are misleading. I will come back to this again in the discussion.

10) The use of routine reporting both for outcome assessment and as a part of the intervention (determination of the incentive payments for intervention groups) is highly concerning. This limitation is inadequately addressed in the discussion which I will come back to as well.

11) The intervention included tremendous amounts of external support besides the P4P incentive, QA for microscopy additional supervision. Because these were offered to both arms they should compromise the identifiability of the P4P effect, but they do compromise the external validity of this study, especially affect statements about cost which are important and laced throughout the manuscript.

12) the authors need to make explicity clear what is menat by "end of intervention period" are they analyzing only the data form the last quarter, last month, last day, the entire period ?

13) Im surprised that the ICC assumed in power calcs were so low, what was the source of this assumption.

14) what does sampled every nth patient mean?

15) why does the primary endpoint used differ in the manscript results and abstract from the power calculation?
16) What does it mean that models were chosen on fit? using what criteria? AIC, pseudo R2... Please clarify.

Discussion:

17) First sentence of discussion is way to strong. There are no clinical outcomes, they have not assessed the effect of P4P on case management only on provider prescribing and testing.

18) IN the 2nd paragraph of discussion the authors present the results of a new subgroup analysis. This needs to be adequately described and moved to the results and methods sections, not presented in the discussion.

19) in the 3rd paragraph of discussion the authors assert that the effect would have been larger in higher transmission areas, but earlier (unless i misread state that the intervention had larger effect size in the sub-group analysis for low transmission areas). These things would work against each other, and without an analysis of the interaction of intervention with transmission, or a look at the predictions from the model for absolute changes with intervention its not clear what wins out, the reduction in effect in higher transmission, or the higher proportions of patients who could be treated with ACT

20) last paragraph on page 10 - what is adherence to testing in malaria negative patients? this sentence doesn't make sense.

21) The Discussion asserts that the incentives could be offset by ACT cost savings, again this is a new result presented in the discussion not backed up in the methods or results and clearly not backed up by an honest costing of the incentive intervention, which would also need to include the cost of the additional supervision, the QA system and all other components. This cannot go in as stated as its clearly misleading.

22) There is no discussion of the potential for facilities to manipulate reporting to enhance incentive payments, clearly this could happen, is a real risk to this trial given that the intervention and its assessment are conducted through the same mechanism (routine reporting) and the discussion does not adequately assess this.

Other comments:

In table 2: Should one not report the following items n, SE or CI and not simple means and SD.

Table 3: There is neither in the methods nor in this table a description of how the covariates were created/coded. It make the coefficients difficult to interpret. e.g. what is an adjusted odds ratio for "all ages" is this the AOR for the intervention, or is it the AOR for age? if so how is age coded. I think this table should just be cleaned up a bit and the covariates better defined in the manuscript.

Figure 2: Would changes in the type of lines or possibly colors make more
readable? I find it a bit difficult to distinguish which line is supposed to be which. Further the y-axis labels do not adequately specify the outcome and conflict with the legend.

**Quality of written English:** Acceptable

**Statistical review:** Yes, but I do not feel adequately qualified to assess the statistics.

**Declaration of competing interests:**

I declare that I have no competing interests.