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:** 15 July 2015

**Reviewer:** S. Patrick Kachur

**Reviewer’s report:**

Performance based incentives are increasingly considered by global health systems managers and researchers—largely based on published and anecdotal evidence that they improve the volume of health services delivered. There’s less evidence on their ability to enhance the quality of health services—and that is what makes this report noteworthy. In addition, the expansion of diagnostic testing as a universally recommended component of malaria case management is an issue that malaria control program managers are confronting across the endemic world. Most of the global expansion is being achieved through point-of-care rapid diagnostic tests, so the authors’ focus exclusively on blood slide microscopy is not as generalizable as it could be. Nonetheless, the findings almost certainly hold relevance for improving the quality of malaria case management.

The question posed by the authors is well defined and not broadly explored by previous investigations. For the most part, the authors’ methods are well described, and appropriate. There are a limited number of issues for which additional detail could be provided to ensure the work was replicable. The data appear sound and well-controlled and analytical methods are appropriate to the study design and research questions. All figures and tables appear to be genuine, and support the manuscript narrative. The authors consistently adhere to the relevant standards for reporting and data deposition. Their discussion and conclusions are well balanced and adequately supported with minor exceptions including potentially overlooking some key limitations. The title and abstract accurately convey what has been found and the writing is clear and concise, requiring only minor editing for style and consistency.

In short, this is a valuable study that merits publication in BMC Health Services with minor recommended revisions.

**Major Compulsory Revisions**

1. Page 4. Paragraph 3: describe the random sampling of health facilities more completely. Was it a simple random sample of health centers (what about hospitals or dispensaries where microscopy might have been available)? Was there any attempt to sample probabilistically (by utilization or catchment)? Were all facilities government operated or were NGO and or for profit facilities eligible? Later the random assignment is described at being deliberately balanced by
(former) district (Paragraph 5, same page and section)—so was selection stratified by (former) district? Otherwise, how could this have been achieved?

2. Page 5. Paragraph 3: “A systematic random sample of patients…” almost immediately calls to mind questions about the sampling strategy and sample size. But these aren’t answered well until the description of the power calculations and analytical approach. And even then the nature of the age stratification is hard to understand, especially since it is common for outpatient clinics to operate services for children under 5 independent from those for older children and adults. This would be more appropriately described in some detail at the top of the methods section with the selection of facilities (see previous comment).

3. Pages 9-11. Discussion. Acknowledge the limitations of the study design with respect to the following

a. All eligible facilities were health centers? Was there any non GoK facility included?

b. Non probabilistic sampling/ selection.

c. Focus is almost exclusively on microscopic diagnosis. What might the findings mean for widespread implementation of RDT based diagnosis (or what are the limits in generalizing from your data into a present or future where RDT diagnosis is more common than microscopy).

d. Limited if any effort had been made nationwide to implement 2010 diagnostic testing guidelines.

Minor Essential Revisions

4. Page 2. Abstract. Background: “We tested a P4P strategy that emphasized diagnosis and appropriate treatment…” could be restated as “We tested a P4P strategy that emphasized microscopic diagnosis and appropriate treatment…” to better distinguish this approach from one of clinical or syndromic diagnosis or malaria rapid diagnostic tests.


6. Page 2. Abstract. Methods: Consider replacing “a malaria test” with “a malaria blood slide” and “without a malaria test” to “without a blood slide”.

7. Page 3. Paragraph 2: Describe when the WHO (2010) and Kenyan MOH recommended universal diagnostic testing for malaria and relate this to the timing of your study.


9. Page 4. Paragraph 2: What is meant by the claim that AMPATH includes 17 districts in the west of Kenya? In Western Province? Western, Rift Valley, and Nyanza? The predevolution districts are no longer recognized administrative entities; it is probably important to identify the specific counties.

10. Page 5. Paragraph 3: Describe the introduction of the “AL-specific registers”
as part of the intervention description.

11. Page 5. Paragraph 4: Preventing AL stockouts should be described as a component of the intervention. Also, were there any measures taken by the study team to ensure diagnostic supplies and reagents? Even in the absence of complete stockouts, the history of actual or perceived scarcity in diagnostic or treatment commodities may have had an important impact on provider behavior. Is it possible to describe this as part of the background?

12. Page 7. Paragraph 8: Describe how high transmission areas and low transmission areas were defined for this study (probably belongs in the background or methods, unless this was based on a slide positivity rate during the data collection period).

13. Page 9. Paragraph 1 (and Page 15. Figure 2): It is not clear what measure of central tendency or other summary measure (Total? Mean? Median? Across how many facilities or time points?) is indicated by each point on the graphs.

Discretionary Revisions

14. Page 3. Paragraph 1: The references cited [1] and [2-7] all predate the WHO recommendation for universal diagnostic testing for malaria. Consider updating this with more current references, or explain that these were contemporaneous to your study period.

15. Page 4. Paragraph 5: Were there any attempts prior to the study to promote the 2010 National Guidelines? Describing these may be important.

16. Page 5. Paragraph 2: “…the amount…saved if over-use of ACT was curbed.” Curbed by what amount? Or do you mean eliminated entirely? Was that a realistic expectation?

17. Page 7. Paragraph 4: “One facility was excluded…after 6 months…” Were all the data collected from that facility (even in the first 6 mos) subsequently excluded…or were no data ever collected from that facility to begin with?

18. Page 8. Paragraph 7: What is “patient equipment”? Were any of the incentive funds used for malaria diagnostic reagents?

19. Page 9. Paragraph 3: “Overall, the training seems to have reduced clinical diagnosis…” Do you mean training AND incentives?

Quality of written English: Acceptable

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

Declaration of competing interests:

I declare that I have no competing interests.