Reviewer's report

Title: Thirty years after Alma-Ata: A systematic review of the impact of community health workers delivering interventions against malaria, pneumonia and diarrhoea on child mortality and morbidity in Sub-Saharan Africa.

Version: 1 Date: 19 February 2011

Reviewer: Christopher J Colvin

Reviewer's report:

Review of “Thirty years after Alma-Ata: A systematic review of the impact of community health workers delivering interventions against malaria, pneumonia and diarrhoea on child mortality and morbidity in Sub-Saharan Africa.”

Review by: Christopher J. Colvin, 19 Feb 2011

This manuscript presents findings from a systematic review of findings from randomised and non-randomised studies of the impact of community health worker (CHW) programs delivering interventions against malaria, pneumonia and diarrhoea on child mortality and morbidity in sub-Saharan Africa. This is an important topic and one, as the authors establish, that has been paid insufficient attention given the burden of these diseases, the global crisis in human resources for health, and the availability of efficacious interventions.

Overall, I found the review very well-designed and written up. I do, however, have a number of questions and points for clarification/response below. I don’t think any of these require major reworkings of the review and deal more with how the article presents its findings. But I do think they need to be clarified.

1. Page 5: the review include only those studies that involved curative care delivered by CHWs (and may or may not have included preventative services as well). It isn’t clear why the focus on curative interventions rather than including studies that involved only preventative interventions? Is it because we have curative interventions of known efficacy that need to be better delivered? Is it because delivering prevention and treatment services might involve different dynamics within CHW programs and should be assessed separately? Is it because there isn’t good evidence on the efficacy of prevention-only approaches?

2. On a related note, the introduction mentions antibiotics, ORS, ITNs and antimalarials (not clear whether this means chemoprophylaxis or chemotherapy) together as important interventions with known efficacy. In the table summarising the studies, the studies are all described as having “curative treatment” and then some, in addition to this, are listed as offering malaria chemoprophylaxis or ITNs (implying these weren’t treated as part of the curative treatments). In the findings and discussion, mention is made of the importance/impact of chemoprophylaxis and ITNs for reducing mortality. I read ITNs and chemoprophylaxis, however, as
preventative interventions, not curative ones. Which raises the question for me about why a systematic review set out to focus only on curative interventions but one of its main findings reported on the impact of two preventative interventions. If this was an unexpected finding (which it seems like it would be if they deliberately excluded interventions that would have only studied ITNs and chemoprophylaxis), it seems like they should call for another review that looks at these two interventions in particular (again, it appears their review didn’t look for studies that just used, for example, ITNs). Finally, also not clear on how the authors inferred that it was chemoprophylaxis and ITNs that caused the impact in the studies here (especially the Menon, Greenwood and Alonso ones) and not the accompanying curative treatments or health education (which are not specified/detailed but are listed as separate items in Table 1).

3. Page 2: when the authors say “A further review by Lehmann and Sanders identified no additional data on the effectiveness of CHW programmes on morbidity/mortality in sub-Saharan Africa”, it isn’t clear to me 1) in addition to what? (to the Lewin and Haines reviews mentioned in the sentence earlier?) and 2) what kind of data (Lewin and Haines reviewed RCTs...the phrasing seems to imply Lehmann and Sanders were looking for additional RCT data).

4. Page 3: might be useful to specify how evaluations of findings from non-randomised and/or weaker study designs can be helpful to policymakers. I assume they are arguing that in the absence of outcome data from randomised studies, policymakers should be able to assess findings from weaker studies. Or do they mean that non-randomised studies offer other, important kinds of information? Or both?

5. I’m not a clinician so the answer to this question may be obvious but I wanted some clarification on the distinction between ARIs and pneumonia. I think pneumonia is the most serious and common ARI in kids. And perhaps the studies they identified only consider pneumonia-specific interventions. But it wasn’t clear why the study definitions and inclusion criteria section specified ARIs and the rest of the article mentions just pneumonia.

6. Page 6: Could the 20 year time frame/limit have affected the results in some way? The time frame was justified in the methods in terms of the presence of 3 other studies that covered the time period before that. But I don’t think these three studies were systematic or comprehensive. Does ignoring data from before 20 year ago introduce any problems/bias? There might be other good reasons for limiting the time frame (social/political/economic contexts and global health policies have changed significantly and that might affect the comparability of the evidence and utility of the findings).

7. Page 6: I assume the paragraph beginning “Two reviewers (JC, AL)…” refers to the broader inclusion process and not only to issues of region and time. It looks, though, like it only refers to that sub-section.

8. Page 10: Last sentence of first paragraph should read: Seven studies, published between 1991 and 2005, met the review’s inclusion criteria. [with the
Otherwise, it looks like 1991 – 2005 is your timeframe under review.

9. Page 12: For the Pence study, it is not clear what “units” are. I am familiar with the terms “arms” and “clusters”. Page 15 also deals with this study and mentions arms, clusters, units and “areas”. The discussion of units and areas in this section is particularly confusing. I understand, I think, that the study had three interventions and thus four “arms” (including the control arm), that it picked one “cluster” per arm (a village, a district, something else) that was considered a cluster because it shared an internal homogeneity that might have made it different from the other clusters (??). Don’t understand units and areas, though. And it seems like the design may have been a stepped wedge design (given the mention of comparing areas that were and weren’t yet scaled up). Is the comparison of mortality here with a separate control arm or with pre-intervention “areas” (?) within clusters (which would imply four intervention??). Given the large and negative effect of this study, it would be useful to have clarity here.

10. Page 12: Gomoa study: Period of impact measurement is 0-3 years after onset but change in mortality is reported for 24-36 months. Is this just how they reported the findings or should 0-3 be 2-3 years? And shouldn’t a point estimate of reduced mortality be at a particular timepoint? Either 24 months or 36 months? The effect may have remained steady across these two timepoints but that seems an odd way to report it.

11. Page 12: Pahou study: confused about what 3 year period and 2 yrs after onset mean. Is three years how far they went back in interviews to assess CHW contact for individuals? Or the length of followup time in the study? Is 2 years after onset when followup commenced?

12. Page 13: Text says the number of CHWs included in studies ranged from 8 to 17. The Hill study, however, seems to indicate in Table 1 that it had 40 CHWs (1 per village and 40 villages).

13. Page 16: authors argue that unknown confounders and selection biases may have been present in the Velema study. Isn’t that always true? Is it any more likely in this study than in others?

14. Page 17: when the authors say “...the effect was larger when these programmes provided prophylaxis against malaria”, larger than what? What is the comparison group of studies/findings for this claim? I think I can see what they mean by looking at Table 1 but it isn’t clear.

15. Could the authors say a bit more about the Navrongo study. They say on page 21 that the authors of that study explained the negative effect was because of reduced healthcare-seeking from skilled providers. That is quite a disturbing finding and it would be nice to know more about the strength of the study authors' inference on this point and if there was something specific about the intervention design that led to the assumed change in health-seeking behaviour.

16. With respect to the PRISMA criteria for reporting in systematic reviews, I didn’t see mention of whether a protocol was available (#5) though the
appendices and the methods section together give quite a comprehensive picture. I also wondered if there was anything else to say about possible limitations of the review at the review level (#25). The authors describe the many limitations of the evidence under review and address the limitations/relevance of the evidence for policymakers but don’t really address limitations of the review design and implementation itself.

**Level of interest:** An article of outstanding merit and interest in its field

**Quality of written English:** Acceptable

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

**Declaration of competing interests:**

I have worked with one of the authors (Simon Lewin) on other projects. I did not work with him on this review or discuss the review and haven't worked with and don’t know the other authors.