RESPONSE TO DECISION LETTER

Dear Editor-in-chief,
We are grateful to the editors and reviewers for their time and constructive comments on our manuscript. We have implemented their comments and suggestions and wish to submit a revised version of the manuscript for further consideration in the journal. Changes in the initial version of the manuscript are either highlighted for added sentences or strikethrough for deleted sentences in the revised version. Below, we also provide a point-by-point response explaining how we have addressed each of the editors or reviewers' comments. We look forward to the outcome of your assessment
Yours sincerely,
On behalf of the co-authors
Joel FOKOM DOMGUE, MD, MPH

Editor(s)' decision and comments

Dear Dr. FOKOM DOMGUE

Manuscript ID BMJ.2014.024615 entitled "Alternative strategies for primary cervical cancer screening in sub-Saharan Africa: a systematic review and meta-analysis of tests performance for VIA, VILI and HPV testing"

Thank you for sending us your article and giving us the chance to consider your work. Your article was read by ten editors and two external reviewers. We enjoyed reading your article and are pleased to make a provisional offer of publication if you are able to revise the paper to address the following comments:

Our answer: Thanks for your appreciation and offer.

* Could you tell us more about the place of HPV testing here. You say data on HPV were limited, and also "Of the three tools, VILI seems to be the most sensitive method in the African continent. " Even though there was no difference in sensitivity between HPV and VILI? From the abstract: "Pooled sensitivity and specificity were similar for HPV testing vs. VIA (both p>0·23), and for HPV testing vs. VILI (both p>0·16)". Please expand on the reasoning here, with general readers in mind.

Our answer: Thank you for bringing this inconsistency to our attention. We have changed the wording of the relevant sentence to read "Among visual methods, VILI seems to be the most sensitive test in the African continent".

* Please provide more information about the overall strategy for screening and treatment with each test. How were they employed clinically in these studies? Positive screen followed by colposcopy and biopsy, followed by treatment ? (what treatment?). General readers won't be aware of what happens to screen positive women.

Our answer: Thank you for raising this point. Indeed, only three studies clearly reported information about the management of screened women. In five studies, the choice between punch biopsy and conization was based on the Reid colposcopic score, but further management of histologically confirmed lesions was not specified. In another study, women with histologically proven cervical lesions were referred for appropriate management according to standard local protocol (which was not specified). In the remaining ten studies, the management of screened women was not reported.
To address this concern, we have added a sub-section in the Methods Section (Page 7) entitled 'management of screened women' where it reads: "The management of women across selected studies when reported was based on the results of screening and reference tests. When colposcopic findings and/or biopsy results showed the presence of high grade cervical dysplasia, women were generally treated using cryotherapy (in recent studies), cauterization (in earlier studies) or conization (loop electrosurgical excision procedure or cold knife conization) as indicated. If the histological interpretation was suggestive of invasive cervical cancer, the patients were referred to a specialized Centre for disease assessment and appropriate treatment (hysterectomy, radiotherapy and / or chemotherapy when necessary). Women who screened negative and those presenting minor grade cervical lesions on colposcopy and/or biopsy were advised to repeat screening within the next three years19"

* Please move to the main body of the paper a section on the study quality (instead of in the appendix) as well as a brief discussion on the gold standard. We thought this would allow tying it up in the Discussion (particularly in relation to HPV).

Our answer: Thank you for this comment. We have moved the indicated sub-sections from the appendix to the Methods section of the main body on pages 6-7 (for the description of gold standard and definition of positive screening tests) and to the narrative section of the results on page 10 (for the study quality assessment).

* You end the abstract by saying "Confirmatory studies using random cervical biopsies' histology as reference are needed." In view of this, how definitive are your results, and how applicable clinically?

Our answer: Thanks for raising this point. We have reformulated the statement to read as indicated below. We feel indeed that impact rather than confirmatory studies are needed to validate the clinical utility of screening strategies in the African setting.
'Implementation studies are needed to assess the impact of these screening strategies on the incidence and outcomes

of cervical cancer in the region.'

* The analysis of "relative sensitivity" and "relative specificity" is not clearly specified. (It will take a lot of detailed reading to understand the relation to the cited paper, which is on Poisson time-series models in twenty US cities.) It seems that the relative sensitivities for VIA versus VILI are obtained from aggregate estimates across all studies that contribute any sensitivity for either VIA or VILI; that is, not restricted to studies that assessed both VIA and VILI. If so, it should be made clearer that this analysis is based on indirect rather than direct comparisons; and a sensitivity analysis using direct comparisons would be a useful comparator, even if underpowered.

Our Answer: Thank you for raising this point. Indeed, the primary analysis of relative sensitivity and relative specificity for either pair of screening tests was not restricted to studies that simultaneously assessed both tests. The text of the manuscript has been modified to clarify this (see the second and third paragraphs of the 'data synthesis' sub-section of the Methods section). As requested, we conducted a sensitivity analysis based on direct comparisons (by restricting the analyses to studies based on paired tests):

- Regarding relative sensitivity and specificity between VIA and VILI (n=8 studies assessed both VIA and VILI, while n=2 studies assessed VIA alone). When restricting the analysis to the 8 studies reporting on both VIA and VILI, the pooled estimates were 0.88 (95%CI 0.83-0.92) with p<0.001 for the relative sensitivity and 1.03 (95%CI 0.93-1.15) with p=0.52 for the relative specificity.
- Regarding relative sensitivity and specificity between VIA and HPV (n=9 studies assessed VIA alone, n=2 studies assessed HPV alone and n=1 study assessed both VIA and HPV). Based on direct comparisons, the estimate of the relative sensitivity was 0.77 (95%CI 0.63-0.89) with p=0.004 and the estimate of the relative specificity was 1.20 (95%CI 1.11-1.31) with p<0.001.
- Regarding relative sensitivity and specificity between HPV and VILI (n=3 studies on HPV alone, and n=8 studies on VILI alone), no study assessed simultaneously both HPV and VILI and the sensitivity analysis could not be conducted.
These estimates were robust except that of the relative sensitivity between VIA and HPV (0.94 with the indirect comparison and 0.77 with the direct comparison which is based on a single study). We have added these results in the Manuscript (see the 'sensitivity analyses and publication bias' sub-section of the Results section)

With regards to the cited paper on Poisson time-series models in US cities, we propose another reference which is more appropriate for the topic of our paper (reference number 27):
Jackson D, Riley R, White I.R. Multivariate meta-analysis: potential and promise. Statistics in Medicine 2011, 30:2481-2498.
However, we still want to keep the reference on Poisson time-series to show the R package we used for the analyses (reference number 32).

* Usually, a univariate analysis of sensitivity and specificity – as shown in Figure 2 – would be deprecated, and a bivariate model (Figure 3) preferred. However, looking at the results in Figure 3, it seems hard not to conclude that the bivariate model of dependence between sensitivity and specificity is a poor fit in this data, and a better conclusion would be that sensitivity is relatively constant across settings (~80%) compared to the great heterogeneity in specificity. Therefore, we thought it was justified to present Figure 2 as well as Figure 3.

Our Answer: We welcome the opportunity of presenting both figure 2 and figure 3, as we suspect that a significant proportion of readers will be more used to figure 2, and perhaps less figure 3. However, Figure 2 does not actually show univariate analyses. The pooled estimates shown in Figure 2 are those obtained with the bivariate model (proposed by Reitsma and obtained with the R package mada). We also estimated the pooled sensitivity and specificity using the R package mvmeta (with the same bivariate model) to obtain in addition heterogeneity statistics (not provided in the R package mada). It is of note that the pooled estimates obtained with both R packages were identical. In our opinion, both Figures are informative. To make it clearer to the reader, this information has been added in the legend of Figure 2.

* Presumably, the analysis of prevalence uses the reference standard rather than the index date. If so it is not clear why it is useful to stratify the results by index test (Table 2 and Supplementary Figure 2). Also, we were unclear on why it makes sense to produce summary estimates of prevalence across countries. It's clear (Supplementary Figure 2) that the prevalence differs greatly by country, so the summary diamonds in this Figure – which are also the main results in Table 2 – do not appear meaningful. The same appears to be true of the positivity rates (Table 2 and Supplementary Figure 3).

Our answer: Thank you for raising these points. The prevalence indeed is based on reference standard and not the index test, and it may not be necessarily relevant to stratify the prevalence by index test. It is however of note that measures of test performance such as positivity rate, predictive values, and to a lesser extent sensitivity and specificity are affected by the background prevalence of disease in the study population. Since those performance measures are subsequently shown by screening test, it makes sense to stratify also the prevalence by screening test to give the reader that information on the background disease prevalence in the population in which each screening method was applied. This is the reason why we are proposing to maintain Supplementary figures 2 (referred to as Supplementary figure 2a in the revised version), Supplementary Figure 4 (referred to as Supplementary figure 4a in the revised version) and Table 2 in our results. However, we have further reinforced the presentation by adding new supplemental figures showing the prevalence also by geographic regions (Supplementary Figures 2b and 4b), as it might be useful to draw estimates of predictive values of these screening tests (as shown in Figure 4 and Supplementary table 3).
With regard to presenting the overall estimates from meta-analysis, these are been presented only for the completeness of reporting, and not to emphasize on a particular meaning of those estimates considering the substantial and unexplained heterogeneity indeed. However, since meta-analyses in these studies were parts of pre-planned analyses, we feel that selectively reporting pooled estimates may rather looks like a fishing trip, as opposed to providing the readers with the full scope of the analyses and allowing them to make their own judgment.

The remark in the text (line 44 on page 9) "broadly of moderate methodological quality" could usefully be fleshed out a bit – for example by remarking that none of the studies were deemed to have avoided "imperfect gold standard bias".

Our answer: Thank you for pointing this out. We have moved from the appendix to the main body of the manuscript, the section on the studies quality (Page 10).

INFORMATION TO INCLUDE IN REVISION
Please would you also check that you have provided the following information

* Competing interest statement (in the style explained at www.bmj.com/about-bmj/resources-authors/forms-policies-and-checklists/declaration-competing-interests)

* Contributorship statement + guarantor
(see http://resources.bmj.com/bmj/authors/article-submission/authorship-contributorship)

* Copyright statement/ licence for publication (see http://www.bmj.com/about-bmj/resources-authors/forms-policies-and-checklists/copyright-open-access-and-permission-reuse)

Our answer: done

Reviewer(s)' Comments to Author:

Reviewer: 1

Recommendation:

Comments:
This is well analyzed and well written manuscript describing the results of a pooled analysis of cervical screening studies in Africa in which screening tests were provided by nurses and addresses a topic of major interest for cervical cancer control in sub-Saharan Africa.

Our answer: Thanks for your appreciation

There are no major issues in the statistical analysis of this pooled analysis. The authors may consider the following comments in revising their manuscript.

1. Methods: The authors state that the included studies should meet the criteria that the study population is not at risk of cancer. This may be confusing to the reader. The women in the age-group included in the studies are at risk for cervical cancer. Probably the authors meant that women already diagnosed with cervical cancer were excluded. Moreover, the HIV infection status of women included in many of the included studies was not known, since these women were not tested for HIV or if tested elsewhere, the test results were not documented in the study databases. Most women in tropical countries complain of symptoms such as discharge commonly.

Our answer: Thanks for raising these important points. We have replaced the criteria "women not at risk of cervical cancer" by "women not at particular risk of cervical cancer" to make it less confusing to the reader that studies based only on specific populations (studies focusing on HIV positive women and studies based only on women presenting with gynecologic symptoms) were excluded from our review" Following editor's comment, this information has been clearly mentioned in the second paragraph of the Methods section, were it reads: "(ii) the study population was not at particular risk of cervical cancer (studies focusing on HIV positive women or on women presenting with gynecological symptoms were excluded)"

2. Methods: the reference standard investigations ("gold standard" investigation for final diagnosis) used in the studies may be briefly described.

Our answer: We have moved from the appendix to the main body of the manuscript, the section on the description of gold standard.

3. The number of included studies in which quality assurance was carried out for reference standard (colposcopy, histopathology) should be mentioned.

Our answer: Thank you for pointing this out. Among selected studies, quality assurance of the reference test was carried out and clearly mentioned in six studies (all conducted under the auspices of the International Agency for Research on Cancer, Lyon, France) through an internal (regular on-site assessment and re-training of colposcopists and pathologists throughout the project) and external (review of slides in a reference laboratory) quality control.
. In eight studies, histological samples were shipped to and analyzed in Western countries (USA, Canada and Belgium) by skilled pathologists, but quality assurance was not mentioned except in one report where histological interpretation was performed independently by two pathologists. In the remaining six studies, information on quality assurance of the reference test was not provided. This information has been added in the 'Assessment of studies quality' sub-section of the Results section (see page 11 of the main body of the manuscript).

4. The sample size of included studies may be shown in the in their forest plots (Figure 2).

Our Answer: Thank you for raising this point. The sample sizes have been added to the forest plots in Figure 2.

5. A major concern in the included studies is the high accuracy reported in those studies. The possible reasons for this reported high accuracy should be discussed in detail. The reference investigations were largely applied by newly trained clinicians in techniques such as colposcopy and directing biopsy and their learning curves possibly evolved with the studies. In most locations where these studies were conducted, there was practically no or only limited prior screening and prior colposcopy experiences and these providers were mostly rapidly trained when the studies were initiated. Similarly the pathologists involved in the studies probably overcalled CIN 2 + lesions given their limited experience, recent reorientations and lack of adequate quality assurance. The correlation between the visual techniques might have been another contributory factor. These factors together might have contributed to overcall of true positive lesions. The lesion sizes detected might have been larger lesions given the fact these practically never screened populations.

Our answer: Thank you for pointing this out. As requested, we have added a paragraph in the Discussion section to discuss the possible reasons for these reported high accuracy estimates. Please see pages 13-14 of the main body of the manuscript.

6. The authors may emphasize in the discussion that VILI as a single test has never been implemented. It has always been evaluated following VIA. Even with all the blinding, some level of contamination cannot be eliminated. Given the uncertainties for VILI as a single test, the value and limitations using both visual tests in a sequential manner may be discussed.

Our answer: Thank you for bringing this point to our attention. We have added some phrases in the discussion section to address this issue as a limitation of our review (see Page 16 of the main body of the manuscript).

Reviewer: 2

Recommendation:

Comments:
I found this to be a strong, well-executed study/meta-analysis.

Our answer: Thank you for your appreciation

There are a few grammatical/spelling errors:
Line one of abstract - I would change to "clinical utility of CERVICAL visual inspection with acetic acid...."

Our answer: We have made this change in the abstract.

In the abstract, the authors could potentially strengthen the conclusion statement by highlighting that VILI had a higher sensitivity than the other tests (i.e. For primary screening of cervical cancer in SSA, VILI is a simple and affordable alternative to cytology that demonstrates higher sensitivity than VIA or HPV testing. Confirmatory studies comparing to histology from random cervical biopsies are still needed).

Our answer: Thank you for pointing this out. We have highlighted that VILI had a greater sensitivity than VIA in the abstract.

In the discussion, on page 12, 2nd paragraph, "colposcope" is misspelled.

Our answer: We have corrected this spelling error.

The authors, in highlighting the benefits of VILI over VIA, discuss the lack of a time gap with application of iodine. The next sentence discussion need for repeat soaking of the cervix, I assume, is meant to say repeated soaking of the cervix with acetic acid rather than iodine.

Our answer: We understand the reviewer's suggestion. Indeed in this sentence, we meant to say "reduced need to repeat for repeated soaking of the cervix with iodine", by comparison with acetic acid.

Further discussion of VILI in pre versus postmenopausal women would be an interesting addition.

Our answer: The last sentence of the second paragraph of the discussion on page 15 (of the main body of the manuscript) highlights the difficulties of interpreting VILI in postmenopausal women compared with premenopausal women.