Author's response to reviews

Title: Test characteristics and potential impact of the urine LAM lateral flow assay in HIV-infected outpatients under investigation for TB and able to self-expectorate sputum for diagnostic testing

Authors:

Jonathan Peter (Jonny.Peter@uct.ac.za)
Grant Theron (Grant.Theron@uct.ac.za)
Duncan Chanda (duncanchanda@gmail.com)
Petra Clowes (pclowes@nimr-mmrc.org)
Andrea Rachow (arachow@mmrp.org)
Maia Lesosky (lesosky@gmail.com)
Michael Hoelscher (hoelscher@lrz.uni-muenchen.de)
Peter Mwaba (pbmwaba2000@yahoo.com)
Alex Pym (alexanderpym@hotmail.com)
Keertan Dheda (keertan.dheda@uct.ac.za)

Version: 7 Date: 12 April 2015

Author's response to reviews: see over
Response to reviewers’ comments

Test characteristics and potential impact of the urine LAM lateral flow assay in HIV-infected outpatients under investigation for TB and able to self-expectorate sputum for diagnostic testing

Dear Dr Nazareno,

We thank the reviewers for their useful comments and chance to improve our manuscript. We have incorporated many of the suggested edits and believe that we have substantially improved our manuscript to a publication-ready level. Please find our responses to all queries outlined below (actual changes in the manuscript are highlighted in red text as well as being presented, where required, with each response).

Best wishes,

Jonathan Peter, Duncan Chanda and Keertan Dheda
Overall summary
Test characteristics and potential impact of the urine LAM lateral flow assay in HIV infected outpatients under investigation for TB and able to self-expectorate sputum for diagnostic testing. This is an interesting, well-executed cross-sectional study designed to evaluate the role of urine LAM for TB diagnosis in the outpatient setting, which was nested within a randomized multi-center trial to evaluate the impact of Xpert MTB/RIF as a point-of-care test. Optimal diagnostic algorithms for TB remain elusive and currently need to be setting specific so the evaluation of a cheap, non-invasive POC test such as urine LAM in a primary care environment is of scientific and clinical importance. Other studies have suggested that using urine LAM with sputum microscopy or Xpert may add incremental benefit. The study demonstrates that urine LAM testing had poor sensitivity in HIV-infected outpatients who are able to expectorate sputum as part of TB diagnostic testing and that there was no incremental benefit over Xpert or smear microscopy. They also evaluated potential impact on clinical outcomes by assessing indicators such as same-day treatment initiation but did not conclude that LAM could increase same-day treatment initiation except potentially in settings without access to chest radiography. The study was adequately powered to detect differences in diagnostic accuracy between LAM, Xpert and smear microscopy. Based on the study findings, urine LAM should not be used as an adjunctive diagnostic tool for HIV positive outpatients with symptoms/signs of TB who are able to expectorate sputum in a high burden setting.

Major Compulsory Revisions - None

Minor Essential Revisions

C1: Were pregnant patients excluded?
R1: We did not specifically exclude pregnant patients as none of the TB diagnostics posed any risk to mother or fetus.

C2: Why were patients who refused HIV testing combined as a group with HIV positive patients to constitute the group to undergo LAM testing? I’m not sure whether this is standard practice but I would consider adding rationale described in the figure legend to the text as well. The number of patients who refused is likely too small to merit sub-group analysis but could potentially cause false results.

R2: Thank you for raising this point. We took this approach because i) published data supports a higher risk of HIV-seropositivity amongst patients refusing voluntary counselling and testing even in low prevalence settings (Jones J et al. Sex Transm Dis. 1993), and ii) we thought this the pragmatic approach for an HIV-endemic remote setting where the >50% of new TB cases are HIV co-infected. In addition, as the reviewer mentions this is a very small sub-group making sub-group analysis inappropriate. We have now added the following to the study population methods section:

"Patients refusing voluntary counselling and testing for HIV (3%) where considered “positive” and included in the LAM analysis as this would occur in routine clinical practice given the very high (>50%) incidence of HIV co-infection amongst new TB cases in these endemic countries.”
C3: Would suggest brief discussion regarding potential causes for false negative results.

R3: Thank you for this comment. A number of existing reviews and published studies, referenced in this manuscript, have discussed the likely reasons for the low sensitivity of urine LAM strip testing in outpatient, HIV uninfected or HIV infected patients with high CD4 counts. Consequently, we opted to discuss the reasons why our study, unlike preceding studies, failed to show an incremental benefit of urine LAM strip testing over sputum-based smear microscopy or Xpert MTB/RIF testing. This is the focus of paragraph two of the discussion and three possible reasons are given. Nevertheless, to be more explicit about what is thought to increase the detection of LAM in the urine we have altered the introductory sentence of paragraph two to read as follows:

“Published data, albeit limited, indicate that LAM sensitivity is increased with higher circulating LAM levels, occurring with higher mycobacterial disease burden, extrapulmonary TB, lower CD4 cell count and WHO clinical stage 3 and 4 in out- and in-patient settings [11,25-32].”

C4: Would suggest brief discussion regarding how freezing might specifically affect results in terms of effect on sensitivity/specificity.

R4: Thank you for this comment. We have now clarified this in the limitations paragraph of the discussion as follows:

“LAM was performed on frozen rather than fresh samples which could have reduced test sensitivity, however, meta-analysis data suggests no differences in diagnostic accuracy using frozen rather than fresh samples.”

C5: page 5, line 12 spelling of diagnostics

R5: We have corrected this.

C6: page 11, line 9, increase rather than increased

R6: We have corrected this.

C7: page 17, line 10, spelling of Nocardia

R7: We have corrected this.

Discretionary Revisions:

C8: It may be worth commenting on potential utility of LAM in children.

R8: Thank you for this comment. We were also hopeful about the potential utility of urine LAM testing in children. However, a recent large study from our setting showed disappointing sensitivity and particularly poor specificity in HIV uninfected and infected children (Nicol et al *Lancet Global Health*)
Consequently, we have opted to restrict the discussions to this specific out-patient adult population able to self-expectorate sputum.

C9: Revise discussion so there is less repetition between paragraphs 2 and 5.

R9: Thank you for this comment. We have recrafted paragraph two and five to reduce repetition between the two. Paragraph 2 is focused on the setting in which LAM has shown to add diagnostic value to sputum-based diagnostics and the reasons why we did not find that in this setting. In contrast, paragraph 5 is focused now on the consequences of reduced diagnostic value, namely little expected benefit on patient-important outcomes and perhaps even harm.

**Paragraph 2:**

Published data, albeit limited, indicate that LAM sensitivity is increased with higher circulating LAM levels, occurring with higher mycobacterial disease burden, extrapulmonary TB, lower CD4 cell count and WHO clinical stage 3 and 4 in out- and in-patient settings [11, 27-34]. Moreover, LAM and sputum smear microscopy identified non-overlapping sub-groups of culture-positive TB, thereby offering additive diagnostic value [11-14, 16, 35]. By contrast, we found no incremental benefit of LAM. There are a number of possible explanations. Firstly, sputum-scarce TB, smear-negative TB, and EPTB is more common in hospital-based and pre-ART screening cohorts and thus smear microscopy sensitivity is more likely to be reduced, which in turn would increase the incremental benefit offered by LAM testing. Secondly, differences in sputum smear microscopy staining and concentration methodologies across different studies affect smear microscopy sensitivity [13]. Thirdly, in contrast to other LAM-based studies [11, 12, 14] our study did not offer sputum induction to improve sputum sampling and thus, those unable to spontaneously produce two spot sputa were excluded. Sputum induction was not offered as the pragmatic study design of the parent study reflects the reality that sputum induction facilities remain unavailable in the majority of routine primary care clinic settings. Nevertheless, the inability of LAM testing to improve the diagnostic yield of a single sputum-based Xpert MTB/RIF, irrespective of declining CD4 cell count, is consistent with the findings of Lawn et al. and reflect the superior sensitivity of sputum-based Xpert MTB/RIF for the diagnosis of pulmonary TB [11, 36, 37].

**Paragraph 5:**

LAM offered limited incremental diagnostic benefit over sputum-based diagnostics. In contrast to studies showing incremental benefit of LAM in hospitalised patients with sputum-scarce TB and EPTB [14, 16] or those identifying patients with TB missed by empiric treatment initiation but identified by LAM [40]. Consequently, our study suggests LAM would have minimal potential impact on patient important treatment outcomes. In fact, test specificity was significantly lower when combining Xpert MTB/RIF with LAM for both a culture and composite reference standard with the potential to increase inappropriate treatment initiation. Thus, sputum-based diagnosis, especially where Xpert MTB/RIF is available, should be preferred in HIV-infected outpatients able to spontaneously provide sputa. However, LAM may potentially improve same-day treatment initiation in the clinic setting where only sputum smear microscopy is performed and no chest radiography facilities are available. In addition, LAM may still offer i) important added diagnostic benefit where the performance of sputum-based tests is reduced such as sputum-scarce TB, extrapulmonary TB, mycobacteremia [27], and/or renal TB [41], and ii) important prognostic and treatment monitoring utility [18].
C10: Since TB treatment monitoring remains challenging, it may be worth discussing whether you think the potential prognostic utility for LAM in terms of risk of mortality merits further investigation.

R10: Thank you for this interesting comment. We definitely think that the potential prognostic utility of LAM for mortality warrants further investigations. It has led many, including us, to the hypothesis that identifying LAM-positive patients early and initiating treatment may save lives. However, the reverse may be true with LAM-positive patients representing an “unsalvageable” population in whom early treatment may make little difference. This equipoise we have addressed with our recently completed LAM RCT of 2600 patients across these same four countries. We have examined if the use of urine LAM as a basis for early TB-treatment initiation in hospitalised HIV-infected patients can decrease 8-week all cause mortality. The results should be available shortly. Other groups are investigating the potential utility of urine LAM for treatment monitoring (Wood et al BMC Infectious Diseases 2012) or targeting of these patients to closer follow-up regimens.
The authors conducted a prospective study of various TB diagnostics, and conducted rapid LAM testing using frozen samples. Their conclusion is that urine LAM has limited diagnostic value where sputum Xpert and AFB smear microscopy is available, which may indeed be true.

Major comments

C1: My major concern relates to the conclusions that have been drawn. The Abstract Conclusion that ‘LAM offers no incremental value over Xpert or AFB smear’ is not entirely correct, according to their data. The authors reported a non-significant increase in diagnostic sensitivity when adding urine LAM to both tests, but this does not equate to “no incremental value”. The authors should also specify if they are referring to “diagnostic value”, since the value of a test also relates to accessibility and cost, where urine LAM is superior to AFB and Xpert. In this case, it would be most appropriate to change “value” to “diagnostic sensitivity”.

While urine LAM may not have had a significant increase for ‘same-day treatment initiation with CXR’, the number needed to test to identify one additional TB-infected person was only 12 people. For a 25-minute, urine-based, inexpensive test, that doesn’t seem so bad. Their conclusions that urine LAM would ‘only improve Rx initiation in settings without chest radiography’ also does not seem entirely supported by their data. The data suggest that LAM may improve Rx initiation in settings without chest radiography, sputum microscopy, or Xpert. The conclusion that “LAM is unlikely to impact TB-related morbidity and patient dropout” does not appear to be supported by the available/presented data. The authors should provide more justification or delete this phrase.

R1: Thank you for these important comments leading to a revision of the manuscript abstract. We agree that we should be more specific with the use of the term “incremental value” as we agree that there are many components encompassed by a test’s value. Consequently, we have now changed this phrase to read “…incremental diagnostic value…”, a more accurate term which we feel is supported by the manuscript data.

The parent RCT offered a very detailed look at same-day clinic diagnostics and treatment practices, also providing accurate data about patient dropout rates with different same-day sputum-based diagnostics. We thus thought it a unique opportunity to gain insight into the possible incremental benefit of incorporating urine LAM testing into the same-day clinic diagnostic armamentarium. We have acknowledged in the limitations that this is not prospectively collected data and have also drawn all conclusions cautiously, but nevertheless feel that conclusions can be drawn and based on the diagnostic accuracy data in this setting, are likely to be accurate. The data presented in support of this second abstract conclusion is figure 2 and results section: “Potential impact on patient-important treatment outcomes of adding point-of-care LAM to sputum-based Xpert MTB/RIF or smear microscopy”. It reads as follows:

“The potential impact on these outcomes of adding point-of-care LAM (in the manner described in methods section) is also shown in Figure 2. In the smear microscopy study arm with the use of empiric treatment based on CXR, use of POC LAM would neither have significantly decreased dropout [18% (16/89) to 12 % (11/89), p=0.3], nor significantly increased same-day treatment initiation [35% (31/89) to 44% (39/89), p=0.2]. However, in primary care settings without same-day CXR facilities, POC LAM would have significantly increased same-day treatment initiation [24% (21/89) to 44% (39/89, p=0.004]. Potentially unnecessary treatment initiation would have increased by 3% (4/116) due to LAM
'false-positives'. In the Xpert MTB/RIF study arm, LAM would neither have significantly decreased treatment dropout [11% (10/92) to 10% (9/92), p=0.8] nor significantly increased same-day treatment initiation [55% (51/92) to 57% (52/92), p=0.9]. However, potentially unnecessary treatment initiation would have increased by 7% (7/106) due to LAM 'false-positives'."

Although we feel that the above is sufficient data, we agree that extrapolating this to: "only improve Rx initiation in settings without chest radiography" is a stretch. We have thus reworded this sentence to be more specific and cautious as follows:

"In primary care clinic settings with chest radiography and current empiric TB treatment practices, LAM seems unlikely to improve same-day treatment initiation and patient dropout."

C2: The authors report a significant association for urine LAM to predict patient outcomes, which should be mentioned in the Abstract Conclusion.

R2: Thank you for this comment. We have now included this in the abstract conclusion as follows:

"In African HIV-TB co-infected outpatients able to self-expectorate sputum LAM had limited sensitivity even at low CD4 counts, and offered no significant incremental diagnostic yield over Xpert-MTB/RIF or smear microscopy. In primary care clinics with chest radiography and where empiric TB treatment is common, LAM seems unlikely to improve rates of same-day treatment initiation and patient dropout, however, the ability of LAM to identify patients at high risk of death or lost-to-follow-up may offer important prognostic value."

C3: Using a higher grade of the urine LAM test (>=2+) is likely to detect patients with a higher bacillary load, and may not be the most suitable threshold for a clinic-based test. At the clinic-based level using this threshold, one wouldn’t expect any increased sensitivity over Xpert or AFB smear, since they are also detecting a high bacillary-load infection. Using a lower grade (>=1+) LAM as positive, the rapid test may have more clinical utility in an ambulatory setting. Therefore, the authors should also present results when using the >=1+ grade for the urine LAM test, even if as a supplementary table.

R3: Thank you for this comment. We have now included a comparison of the urine LAM diagnostic accuracy using the grade 1 compared to the grade 2 cut-point as supplementary table S2. This is referenced in the LAM performance results section line 20. In addition, we have included the old and new reference card in the supplementary materials so that the use of the LAM grade-2 cut-point is clarified for readers.

C4: The study had 14 people did not have the reference gold-standard test (“contaminated culture or no available result”), this group should be eliminated from the analyses. The final cohort size should then be 569 people. Table 1 should then contain the description of this analyzed cohort (N=569), not the cohort of the parent study (N=1,095).

R4: Thank you for this comment. Figure 1 outlines the exclusion of HIV uninfected patients and the inclusion of HIV positive (n=564) and refused testing patients (n=19). This left a cohort of 583 on which the analyses were performed. Demographic data for this cohort is provided in Table 1. We opted to describe the demographics of the 583 patient cohort rather than the 569 patient cohort although we
report LAM diagnostic accuracy versus a culture-based reference standard because i) we provide
treatment data on the 583 patient cohort and ii) considered patient-important outcomes where culture
was only one aspect of diagnosis and treatment. Based on this comment we have check the
demographics for the 569 patient cohort and noted no major differences. Thus, if the reviewer feels
strongly about this we would be happy to change and report the demographics for only the 569 patients
as suggested.

C5: The authors’ statement that there is no available data on patient impact and mortality is not correct.
They should cite/reference the following reference and cite the available published data about the (1)
treatment response and (2) association with patient outcomes and mortality: 1. Drain PK, Gounder L,
Grobler A, Sahid F, Wilson D, Bassett IV, Moosa MYS. Point-of-Care Urine Lipoarabinomannan (LAM) for
Diagnosis and Treatment Response of Pulmonary Tuberculosis in Sputum Smear-Negative Suspects.
44th Union World Conference on Lung Health on November 2, 2013. 2. Drain PK, Gounder L, Grobler A,
Sahid F, Wilson D, Bassett IV, Moosa MYS. Rapid Urine Lipoarabinomannan (LAM) Testing after Two
Months of Tuberculosis Treatment Independently Predicts Mortality in a Resource-Limited Setting. 44th
Union World Conference on Lung Health on November 2, 2013.

R5: Thank you for this correction. We have now updated the relevant paragraph of the introduction and
cited the reviewers two conference abstracts. We have also cited the reviewers abstract in connection
with the potentially important use of urine LAM testing in treatment monitoring in the discussion.

C6: Since this is a longitudinal study, the authors should use Cox proportional hazards (rather than
simple logistic regression) to assess the relationship to clinical endpoints (mortality/LTFU).

R6: Thank you for this comment and we agree that Cox proportional hazards would be the more
statistically powerful analysis for exploring the time to mortality. However, we felt that it was also of
clinical interest to consider both mortality, and mortality combined with lost-to-follow up as a binary
outcome, especially given that lost-to-follow up data could not really be considered as a continuous
outcome variable. Consequently, we discussed the analysis with the study statistician and she felt that
using logistic regression analysis was appropriate.

C7: The study excluded those people unable to expectorate 2 sputum samples. The urine-based test may
be best utilized in those who do not have a productive cough (i.e. able to expectorate sputum).
Therefore, the authors should provide more discussion on how this exclusion may temper the potential
impact of the rapid urine LAM test.

R7: Thank you for this comment. We actually think that this is a very important and particularly unique
aspect of this study. Thus, we have highlighted this throughout both the abstract and main manuscript
methods and discussion. The study population is a very homogenous one in whom sputum-based
diagnostics can be performed. This has allowed for a focused investigation of urine LAM strip test
performance and potential utility in this patient sub-group, the dominant one presenting to primary
care clinics even in HIV endemic settings. Indeed, we expected urine Lam impact to be lowered in this
sub-group based on previous data, but we feel this study was able to generate a very clear picture to
guide National Tuberculosis Programme (NTP) policy. We have added to the discussion to make clear
that urine LAM may offer important benefit in those unable to perform sputum-based diagnostics, and
that urine Lam may offer important prognostic utility.
“LAM offered limited incremental benefit over sputum-based diagnostics. In contrast to studies showing incremental benefit of LAM in hospitalised patients with sputum-scarce TB and EPTB [14, 16] or those identifying patients with TB missed by empiric treatment initiation but identified by LAM [39], in our study LAM had minimal potential impact on treatment outcomes, or incremental benefit over either Xpert MTB/RIF or smear. This is likely explained by the nature of our inclusion criteria (sputum expectorating patients). In fact, test specificity was significantly lower when combining Xpert MTB/RIF with LAM for both a culture and composite reference standard. Thus, sputum-based diagnosis, especially where Xpert MTB/RIF is available, should be preferred in HIV-infected outpatients able to spontaneously provide sputa. LAM may potentially improve same-day treatment initiation in the clinic setting where only sputum smear microscopy is performed and no chest radiography facilities are available. In addition, LAM may still offer i) important added diagnostic benefit where the performance of sputum-based tests is reduced such as sputum-scarce TB, extrapulmonary TB, mycobacteremia [27], and/or renal TB [40], and ii) important prognostic and treatment monitoring utility {Drain, 2013 #3052}.”

C8: These analyses may have been better served to help guide primary-care clinics needing to introduce a rapid TB diagnostic test – should they implement AFB microscopy, urine LAM, or Xpert?

R8: Thank you for this comment. We hope that this data will form part of the growing body of data to guide NTPs and, in turn, primary-care clinics on the best use of urine LAM strip testing in the TB diagnostic algorithm for different patient sub-groups.

Minor Comments:

C1: In the Abstract Results, authors state “583/1095 patients”, but the 1,095 is unnecessary. Just state, “Among 583 participants, …”

R1: Thank you for this comment. We have updated the abstract as you have suggested.

C2: The commentary in the Abstract Results, “would have had little impact on patient dropout” is interpretation of data and would be better served in the Conclusion section.

R2: Thank you for this comment. We have changed this section of the abstract results to more accurately describe the data generated. Thus, we have changed the sentence to the following: “Clinic-based LAM, as adjunct to either smear microscopy or Xpert MTB/RIF same-day testing, would neither have decreased patient dropout, nor increased same-day treatment initiation in these clinic settings (with same-day chest radiography available).”

We have not included p-values here as this would make the abstract results section too long. We will be happy to include should the reviewer and editors indicate that this would be preferred.

C3: The authors state concerns about test sensitivity in the Abstract, but then do not comment on this in the Conclusions.
R3: Thank you for this comment. We only make reference to concerns about the specificity of the assay in the abstract introductions. We have now removed this entire sentence to shorten the abstract as we have added to the results and conclusions sections in response to previous reviewers’ comments.

C4: The grammar could be improved. For example, “Used at POC Xpert” is not proper English.

R4: Thank you for this suggestion to improve the grammar. We have gone through the manuscript to improve the grammar and spelling as instructed. For instance, “Used at POC Xpert” has been updated to “Used at the POC, Xpert...”

C5: The authors should be clear that although 2 sputa specimens were collection, only 1 specimen underwent mycobacterial culture testing, which can result in diagnostic misclassification – leading to a lower specificity.

R5: Thank you for this important comment. We have been careful to draw the readers’ attention to the fact that only one sputum culture was performed and we acknowledge this as a major limitation of this study, although secondary analysis using a composite reference was performed to somewhat mitigate against this bias and strengthen conclusions.

The following is in the methods:
“Patients randomised to the smear microscopy study arm received two same-day sputum smears for acid-fast bacilli and one arbitrarily selected specimen also underwent culture.”

The following is in the limitations paragraph of the discussion:
“Misclassification bias was a potential problem in our study as a single sputum culture can miss TB cases amongst HIV-infected patients.”

C6: The authors show an association between urine LAM + and mortality/LTFU. Does this outcome differ when stratified by those started on anti-TB therapy and those not started on treatment?

R6: Thank you for this interesting comment. LAM remains a significant predictor of mortality/LTFU irrespective of treatment in both all HIV infected patients and restricted to TB culture positive only. This may reflect how LAM is a poor prognostic sign irrespective of appropriate TB treatment initiation.

C7: The authors should also include citation of another LAM study in an outpatient setting: Drain PK, Losina E, Coleman SM, Giddy J, Ross D, Katz JN, Bassett IV. Value of Urine Lipoarabinomannan Grade and Second Test for Optimizing Clinic-Based Screening for HIV-Associated Pulmonary Tuberculosis. J Acquir Immune Defic Syndr. 2015 Mar 1;68(3):274-80.

R7: Thank you for this reference and we apologise for not citing it. We have now added the citation at the appropriate place in the discussion.

C8: This manuscript provides limited additional information from what has been previously published. However, as more data continued to be collected about urine LAM performance, these data will contribute to the conversation. I would recommend publication after major revisions.
Thank you for this comment. We agree that there is a growing body of published work on the performance of urine LAM strip testing. However, evidence is still limited and heterogeneous for optimal policy making. As the WHO STAG committee meeting approaches where the urine LAM strip test will be discussed, we believe this data will be informative and useful. Indeed we have already been approached to have it included in the meta-analysis under preparation.