Author's response to reviews

Title: Diagnostic Accuracy of Quantitative PCR (Xpert MTB/RIF) for Tuberculous Pericarditis Compared to Adenosine Deaminase and Unstimulated Interferon-gamma in a High Burden Setting: A Prospective Study

Authors:

Shaheen Pandie (pandieshaheen@yahoo.com)
Jonathan G Peter (jonny.g.peter@gmail.com)
Zita Kerbelker (krbzit001@gmail.com)
Richard Meldau (richard.meldau@uct.ac.za)
Grant Theron (grant.theron@uct.ac.za)
Ureshnie Govender (ureshnie.govender@uct.ac.za)
Mpiko Ntsekhe (mpiko.ntsekhe@uct.ac.za)
Keertan Dheda (keertan.dheda@uct.ac.za)
Bongani M Mayosi (bongani.mayosi@uct.ac.za)

Version: 2
Date: 15 March 2014

Author's response to reviews: see over
Response to reviewers' comments
MS: 3024144681179374

Diagnostic Accuracy of Quantitative PCR (Xpert MTB/RIF) for Tuberculous Pericarditis Compared to Adenosine Deaminase and Unstimulated Interferon-gamma in a High Burden Setting: A Prospective Study

Dear Professor D’Souza:

We thank the reviewers for their useful comments which have helped to improve our manuscript. We have addressed all major and minor comments and provided an explanation of sample volumes used, level of detection experiments, test costing and the contextualisation of study findings for HIV-positive and negative groups. Please find our responses to all queries outlined below (the changes in the manuscript are highlighted in red text as well as being presented, where required, with each response).

We have expanded and clarified the informed consent process by including the following statement in the Methods section: “Informed consent was obtained from each patient prior to enrolment in the registry and the study protocol conforms to the ethical guidelines of the 2008 Declaration of Helsinki as reflected in a priori approval by the human research ethics committee of the University of Cape Town (HREC REF402/2005)”.

We also noted an error with reference 23 which has been replaced with the correct article. We trust that the revised manuscript is acceptable for publication in BMC Medicine.

Yours sincerely

Bongani M Mayosi, DPhil
Reviewer #1: Padmapriya Banada

General comments:

The above manuscript is well presented. This study has compared Xpert MTB/RIF assay to a newer modified uIFNr assay (IRISA, Antrum Biotech) and Adenosine Deaminase (ADA) assay. The authors tested 151 patients undergoing pericardiocentesis using the above diagnostic tests. The authors conclude uIFNr assay to be better in overall performance compared to the other two.

Comment 1: However, it is important to note that this comparison is between 1 ml of pericardial fluid in Xpert to 20 ml in uIFN.

Response 1: We compared diagnostic accuracy between Xpert MTB/RIF performed on 1ml, and Xpert MTB/RIF performed on a resuspended pellet of 3-20mls of centrifuged pericardial fluid, and uIFNγ performed on the supernatants of these centrifuged pellets. The biomarkers (uIFNγ and ADA) are dissolved evenly in the fluid, and as such, their concentrations (in pg/ml or IU/L respectively) should be similar whether measured in a 1ml or 20ml starting volume. By contrast, Xpert MTB/RIF is assessing the presence of nucleic acids in mycobacteria (i.e., particulate substance) whose yield may be increased by using a larger volume of fluid. We chose to present data on Xpert MTB/RIF on 1ml unprocessed pericardial fluid because with standard processing according to manufacturer’s instructions, no significant increase was noted in the sensitivity of Xpert MTB/RIF using 20mls of centrifuged pericardial fluid compared to 1ml unprocessed pericardial fluid, and in fact error rate increased.

We have now described clearly the starting volumes used for the different tests throughout the manuscript, as follows:

Methods:
Xpert MTB/RIF
Xpert MTB/RIF was performed using both 1ml of unconcentrated and unprocessed PF as well as 3-20 ml of centrifuged (3000g x 15 minutes) PF reconstituted to 1 ml with phosphate buffered saline (PBS)

uIFNγ

uIFNγ levels were measured in duplicate using supernatant attained from 3-20 ml of thawed and centrifuged (3000g for 15 minutes) PF

ADA

Adenosine deaminase assay (Diazyme, USA, http://www.diazyme.com/adenosine-deaminase-ada) was performed on 1-8mls PF samples

Table 2 footnote:

“Diagnostic accuracy measures presented in this table are for Xpert MTB/RIF performed on 1ml unconcentrated PF, uIFNγ (both the QFT-kit and Intergam-kits) performed on the supernatants on 3-20mls of centrifuged PF, and ADA performed on 1-8mls of unprocessed PF”

Comment 2: Although authors show the 20 ml volumes did not yield significance difference compared to 1 ml in Xpert, no efforts were made to optimize the sample processing to overcome the error rate.

Response 2: We agree that we did not attempt further sample processing optimisation steps beyond simple centrifugation and that these may help to improve Xpert MTB/RIF sensitivity and decrease error rate. However, we wanted to keep sample processing very simple and suitable for ease-of-use in resource-limited laboratories. Additionally, as there is very little published data on Xpert MTB/RIF performance in pericardial fluid we thought that a starting point was to evaluate performance in a large cohort of patients with very basic sample processing easily applied to both Xpert MTB/RIF and biomarker testing. As far as we are aware no existing data is available on optimising pericardial or pleural fluid samples to improve Xpert MTB/RIF performance and decrease error rates. We did explore the effect of PCR inhibition on Xpert MTB/RIF performance and noted no effect of pericardial fluid concentration. We agree that further studies are required to optimise the preparation of
pericardial fluid samples for Xpert MTB/RIF, and we have added your comment to our limitations section and adjusted our final study conclusions as follows:

Limitations section:
“This study did not optimise pericardial fluid sample processing beyond a simple centrifugation step thought applicable to resource-limited settings, and the use of additional or alternative approaches may have improved Xpert MTB/RIF sensitivity and/or decreased the high indeterminate rate found.”

Conclusions section:
“In conclusion, uIFN\(\gamma\), offers superior accuracy for the diagnosis of microbiologically confirmed TB pericarditis compared to the new Xpert MTB/RIF test and the established ADA assay, performed using available Xpert MTB/RIF testing protocols without fluid-specific optimisation beyond simple centrifugation.”

Comment 3: Specificity was considerably compromised with HIV positive patients in uIFN and ADA tests (Table S2) indicating, these tests may not be of use in those populations. Specificity is a major concern in diagnostics especially, when it is used for treatment decisions.

Response 3: We were interested in the comparative performance of biomarkers and Xpert MTB/RIF in HIV-positive versus negative patients. However, as noted in the limitations section our study had a limited number of patients that were considered reference standard: “non-TB”. With only 5 HIV-infected patients in this group we felt it difficult to draw any definite conclusions from this data which is why we did not place too much emphasis on it. Indeed the decreased specificity of biomarkers may reduce the rule-in utility of these tests in HIV-endemic settings. Consequently, we suggest the possibility of a two-test strategy which would involve initial biomarker testing followed by Xpert MTB/RIF testing in biomarker positive cases. Further studies are needed to test the utility and cost-effectiveness of this strategy. We have highlighted this issue further in an updated conclusions paragraph as follows:
“Studies are needed to test the utility and cost-effectiveness of a two-test strategy, which may be preferred in HIV-positive patients where biomarker specificity may be reduced”

Comment 4: Sensitivity of the Xpert MTB/RIF assay can be improved by optimizing the sample processing, sample volume etc., So the authors might want to consider rephrasing their conclusions.

Response 4: Thank you for this comment which we agree with and responded to in detail in Response 2. We have changed our conclusions in line with this comment as shown in Response 2.

Major Compulsory Revisions

Comment 5: Pages 8, 12, 16: Although it is agreeable that the LOD of Xpert assay in detecting TB in pericardial fluids was 75CFU/ml, it is statistically not well represented. The LOD should be calculated at 95% or 99% CI. It is not very clear from your methods, how many total replicates and experiments were run to establish LOD analytically!

Response 5: We agree that our analytics for LOD are not statistically well represented. We wanted preliminary data to show that there was not a major difference in LOD between pericardial fluid and sputum samples. Consequently, and with consideration of cartridge costs, we performed 4 replicates for each CFU category. We acknowledge that this is a low number of replicates compared to the large numbers used in other LOD studies for Xpert MTB/RIF in sputum and blood samples. However, we do feel that this proof-of-principle data is a useful addition to the manuscript despite the limited number of replicates. We have updated Figure 2 as you have suggested in your comment below (Comment 6) and made the following changes to the text:

We have updated the Methods and Results sections to improve clarity as follows:

“The limit of detection was determined in duplicate by spiking 0, 50, 75, 100 and 150 H37Rv colony forming units (CFU) to 1ml aliquots of pericardial fluid before dilution with sample buffer and subsequent Xpert MTB/RIF analysis. This experiment was repeated twice, thus providing four replicates for each CFU concentration.”
“Spiking experiments in PF demonstrated that the Xpert MTB/RIF assay was detected in 100% of samples spiked with ≥75 CFUs per millilitre of pericardial fluid (Figure 2, 4/4 replicates detected for 75, 100 and 150 CFUs/ml).”

We have updated the conclusions to indicate the low replicate numbers and preliminary nature of the LOD experiments as follows:

“Preliminary level of detection experiments suggest that the Xpert MTB/RIF assay could reliably detect pericardial fluid samples spiked with ≥75 cfus/ml of H37Rv, which is lower than the 131 cfu/ml limit of detection found in spiked sputum samples. Further studies with more replicates are required to confirm this finding.”

“A low number of replicates was performed in limit of detection experiments and these findings should be confirmed in further studies.”

Comment 6: Fig. 2: Curve should not be drawn here, since lower the cell numbers, the Ct will be higher not lower! This can get the reader confused! Instead you can draw a percent positive curve, to establish your LOD. But please indicate your replicates.

Response 6: Thank you for this suggestion for an improved figure layout. We have changed the figure in line with your comments and updated the figure legend to indicate the number of replicates for each CFU concentration tested.

The revised figure legend now reads:

“Figure 2. Limit of detection of mycobacteria in pericardial fluid. H37Rv was spiked into 1 ml pericardial fluid aliquots. The proportion of positive samples for each CFU concentration out of 4 samples tested per concentration is shown.”

Comment 7: Does uIFN levels alter with freeze thawing?
Response 7: Data is available on freeze-thaw for Quantiferon TB-GIT tests which measure stimulated IFN-gamma. No differences are noted if the supernatants are frozen and stored below -20 degrees. Studies have compared other IFN-gamma assay freeze-thaw and found no effects (Aziz et al. Clin Diagn Lab Immunol 1999). Thus, although specific data is pending on the impact of freeze-thaw for the Intergam test, it is unlikely have a major impact on measurements.

Comment 8: Detailed description of ADA assay is needed. How much of unconcentrated PF sample was used?

Response 8: We have now provided further details on the ADA test and referenced the package insert with a hyperlink. 1-8mls of unprocessed PF was used for testing. The updated methods section now reads as follows:

“Adenosine deaminase assay (Diazyme, USA, http://www.diazyme.com/adenosine-deaminase-ada) was performed on 1-8mls PF samples, collected in serum tubes, according to the manufacturers’ specifications by the National Health Laboratory Services, Groote Schuur, Cape Town (NHLS GSH). Samples were either processed immediately, or stored (at 2 to 4 °C) for processing within 24 hours.

The Diazyme ADA assay is based on the enzymatic deamination of adenosine to inosine, which is converted to hypoxanthine by purine nucleoside phosphorylase. The reagent is used at 37°C ±0.5°C, using an instrument that is capable of reading absorbance accurately at 540nm to 550nm. ADA activity was measured as units per litre (U/L), where one unit of ADA is defined as the amount of ADA that generates one micromole (µmol) of inosine from adenosine per minute at 37°C.”

Comment 9: Page 12: How can ‘20’ be median of 10 to 20?

Response 9: Thank you for highlighting this confusing statistic. The variable is massively skewed to the right with the majority of entries and the highest possible volume being 20mls. We have changed this sentence to indicate that the volume that was used ranged from 3ml to 20 ml.
Comment 10: Page 12: SD is not put in parenthesis, it is traditionally represented as mean±SD.

Response 10: Thank you for noting this. This has now been corrected.

Comment 11: Table S5: please indicate in the food note, what n and N represents.

Response 11: We have updated the footnote of table S5 to indicate what n and N represent.

Comment 12: Page 16, line 1: Without establishing a proper analytical LOD, this phrase would be inappropriate.

Response 12: We agree that this sentence needs to be revised. We have responded in detail to the LOD experiments and analysis in Response 5 above. In line with this we have updated page 16 line 1.

Comment 13: Page 16. Line 4: any possible reasons, why there was an increased ‘indeterminate’ rate, with increase in sample size?

Response 13: Thank you for this comment. It is likely that the increased error rate associated with the use of larger pericardial fluid volumes relates to the concentration of blood, inflammatory proteins and debris that is found in pericardial fluid exudates. We have updated the discussion with a comment of this and also two suggested optimisation methods with recent references as follows:

“The increased error rate may have resulted from reaction failure secondary to large amounts of pelleted blood and other inflammatory proteins found in pericardial exudates. Methods to further digest these proteins or the addition of a PCR-friendly blood lysis buffer may help to decrease error rates.”

New references added:


Comment 14: 10. Fig. 4: I recommend including the Xpert data as well in the ROC curve analysis, unless there is any other specific reason?

Response 14: The comparative ROC-curves are shown for the two biomarkers to primarily show the overall diagnostic accuracy over the spectrum of possible cut-points. ROC-curves tend to be most useful for tests with continuous outputs for which a cut-point selection is required to generate a binary classification of patients into test positive and negative groups. Xpert MTB/RIF is semi-quantitated and the test is reported primarily in a binary manner based on manufacturer decisions. Thus, we had not included Xpert MTB/RIF performance in Figure 4. We have now added the point performance onto the graph represented by a solid black triangle and added a footnote to this effect.

Minor revisions:

Comment 1: Please make sure the abbreviations are expanded in their first use in the text (in addition to the abbreviation list)

Response 1: Thank you for this suggestion. We have ensured that the abbreviations are expanded in full when first used in the text.

Comment 2: ADA- Expand and provide details for the first use of the term in the introduction.

Response 2: This has been updated in line with your comment above.
Comment 3: What are the total hands-on time, time to detection and volume of the fluid used in all the three assays?

Response 3: This study did not specifically compare the total hands-on time and time-to-detection using the different tests. However, a standard protocol was used with only the addition of a 15 minute centrifugation step for Xpert MTB/RIF testing and consequently results were available in just over two hours as the published time-to-detection for this assay. ADA is processed on an automated analyser in the biochemistry laboratory with results usually available in under an hour and very little hands on time. Antrum Biotech has been optimising the time-to-detection for the InterGam assay down to about 2 hours or just over that, which is quicker than the QFT (ELISA component) and rapid compared to a standard ELISA. They have made several modifications to get the time down, but this has not yet been formally assessed.

The starting volumes are provided in the methods section and Table 2 as follows:

Methods:
Xpert MTB/RIF
Xpert MTB/RIF was performed using both 1ml of unconcentrated and unprocessed PF as well as 3-20 ml of centrifuged (3000g x 15 minutes) PF reconstituted to 1 ml with phosphate buffered saline (PBS)

uIFN\(\gamma\)

uIFN\(\gamma\) levels were measured in duplicate using supernatant attained from 3-20 ml of thawed and centrifuged (3000g for 15 minutes) PF

ADA
Adenosine deaminase assay (Diazyme, USA, http://www.diazyme.com/adenosine-deaminase-ada) was performed on 1-8mls PF samples

Table 2 footnote:
“Diagnostic accuracy measures presented in this table are for Xpert MTB/RIF performed on 1ml unconcentrated PF, uIFN\(\gamma\) (both the QFT-kit and Intergam-kits) performed on the supernatants on 3-20mls of centrifuged PF, and ADA performed on 1-8mls of unprocessed PF”
Comment 4: Table S7 is not referenced in the manuscript and authors are requested to explain the significance of this table.

Response 4: Thank you for pointing this out. On considering the matter we have decided to remove Table S7.
Reviewer #2

Reviewer: Pierre Goussard

Reviewer's report:

Dear Editor

The authors address interesting and challenging issues in the confirming of the diagnosis of TB pericardial effusions. This is a well thought out and constructed study and well written article.

Major Compulsory Revisions

Comment 1: The limitations of this study may be the high number of HIV positive patients and the significant difference in age in the non tb group. The Xpert sensitivity was higher in HIV-positive compared to HIV-negative patients [74.6% (61.7-84.2) vs. 21.4% (7.6-47.6). The authors must elaborate on this in the discussion. Does this mean is only a valuable test in HIV positive patients. The HIV patients may have a higher load of bacilli than HIV negative patients? Do the authors have any proof of this?

Response 1: Thank you for this comment to highlight the higher Xpert MTB/RIF sensitivity in HIV-positive versus negative patients. The reviewer is correct in the fact that this is likely a result of higher pericardial fluid bacillary loads in HIV-positive versus –negative patient samples. This is evidenced by a shorter time-to-positivity of pericardial fluid liquid TB culture samples in HIV-positive versus –negative [Median (IQR) time-to-positivity (days) of liquid TB culture samples HIV- positive: 21 (17-29) versus HIV-negative: 25 (12-38), p<0.001]. These results do have implications for the use of the test in HIV negative populations. This is one of the main reasons that we have evaluated the use of Xpert MTB/RIF compared to or combined with other host-derived biomarkers. Our study was performed in a high HIV-TB setting and this does affect the generalisability of our findings to low burden settings. We have attempted to mitigate this to some extent by providing diagnostic comparisions at differing TB prevalences (table S6) but have now specifically raised the HIV issue in both the Results and Discussion sections as follows:
Results:

“The sensitivity was higher in HIV-positive compared to HIV-negative patients [74.6% (61.7-84.2) vs. 21.4% (7.6-47.6), p<0.001; see online supplementary Table S2], corresponding to higher bacillary loads in the pericardial fluids of HIV-positive patients [Median (IQR) time-to-positivity (days) of liquid TB culture samples HIV-positive: 21 (17-29) versus HIV-negative: 25 (12-38), p<0.001].”

Conclusions:

“Interestingly, unlike in sputum and pleural samples, no correlation was found between Xpert MTB/RIF-generated C_T-values and liquid culture time-to-positivity using pericardial fluid.[30] However, the sensitivity of Xpert MTB/RIF was found to be significantly higher in HIV-positive versus negative patients, and this was due to the higher bacillary loads, as measured by liquid culture time-to-positivity (TTP), found in the pericardial fluid of HIV positive versus negative TBP. This sensitivity difference may impact on the utility of Xpert MTB/RIF in low HIV prevalence settings.”

Limitations

“The study was conducted in a high TB and HIV burden setting, which may limit the generalisability of the findings. Performance may differ in a low TB burden setting and where HIV co-infection rates, and hence bacterial load, are lower, such as Europe and the US. However, the use of diagnostic accuracy measures that are less affected by prevalence, such as likelihood ratios, and generating estimates across varying TB prevalence rates helps to highlight potential performance differences between low and high burden settings and hence improve generalisability.”

Comment 2: Do the authors have any idea what volume of fluid must be used for Xpert?

Response 2: Thank you for this comment. Very limited data is available about the performance of Xpert MTB/RIF using pericardial fluid samples. To date, all studies including ours have primarily chosen to follow the manufacturer’s suggested and validated volume for use on sputum samples. i.e., 1ml with the addition of Sample Reagent at a ratio of 3:1. As
concentration of pleural and pericardial fluid has been shown to increase TB culture yields, we attempted to concentrate a 20mls of fluid with a simple centrifugation step with resuspension of the pellet into 1ml PBS and then use of the standard protocol. 20mls was selected as this was felt to i) be the amount of fluid that could reliably be attained during pericardial aspiration in the majority of cases, and ii) be easily handled in standard centrifuge equipment. As discussed in detail above in response to Reviewer #1 comments, we did not attempt further optimisation of sample processing in this study. We have now acknowledged this in the limitations section as follows:

“This study did not optimise pericardial fluid sample volumes or processing beyond comparing two volumes and a simple centrifugation step thought applicable to resource-limited settings. The use of different volumes or alternative processing methods may have improved Xpert MTB/RIF sensitivity and/or decreased the high indeterminate rate found.”

Comment 3: The biggest burden of TB is in the developing world, doing this combination of test may be very expensive and not affordable in most settings. Authors must state the cost of this test.

Response 3: We agree that the cost of the tests is important for sustainable use in low-income countries. Xpert MTB/RIF assay currently costs approximately US$20 per test/cartridge, and the WHO and FIND diagnostics continue to negotiate for preferred pricing. This should see the cost per cartridge reduce to around US$14 per test in the coming few years. South Africa has replaced smear microscopy with Xpert MTB/RIF for frontline TB diagnosis and consequently substantial government and donor funding (e.g. PEPFAR) is making the tests widely available. The cost of ADA measurement is less than US$0.1 test, while the cost for the measurement of uIFN using the novel intergam assay will likely be slightly more than that of smear microscopy, although the test is not yet commercially available.

We have added the following to the discussion section:

“Xpert MTB/RIF currently costs approximately US$20/test, while ADA measurement is less than US$0.1/test. Intergam kits are not currently commercially available so the cost is unknown but likely to be only slightly more than smear microscopy. Prospective studies of
the cost-effectiveness of diagnostic options are needed before it can be considered for clinical practice.”

Minor revisions

Comment 1: There are too many figures and tables. Table 3 does not add much.

Response 1: Thank you for this suggestion. We do feel that Table 3 and the current figures all provide additional information not fully explained in the text. However, there are a number of tables and figures and we have removed, such as Table S4 and S7, to decrease the amount of supplementary material.