Tuberculous meningitis diagnosis and outcomes during the Xpert MTB/Rif era: a 6.5-year cohort study in Uganda [version 2; referees: 2 approved, 1 approved with reservations]

Previously titled: Can improved diagnostics reduce mortality from Tuberculous meningitis? Findings from a 6.5-year cohort in Uganda

Fiona V. Cresswell, Ananta S. Bangdiwala, Nathan C. Bahr, Emily Trautner, Edwin Nuwagira, Jayne Ellis, Radha Rajasingham, Joshua Rhein, Darlisha A. Williams, Conrad Muzoora, Alison M. Elliott, David B. Meya, David R. Boulware

Department of Clinical Research, London School of Hygiene and Tropical Medicine, London, WC1E 7HT, UK
Infectious Diseases Institute, Kampala, Uganda
Division Biostatistics, School of Public Health, University of Minnesota, Minneapolis, MN, 55455, USA
Division of Infectious Diseases, Department of Medicine, University of Kansas, Kansas City, KS, 66160, USA
University of Utah, Salt Lake City, UT, 84112, USA
Mbarara University of Science and Technology, Mbarara, Uganda
Doctors.net.uk, Abingdon, OX14 4SH, UK
Division of Infectious Disease and International Medicine, Department of Medicine, University of Minnesota, Minneapolis, MN, 55455, USA
Medical Research Council/Uganda Virus Research Institute and London School of Hygiene & Tropical Medicine Uganda Research Unit on AIDS, Entebbe, Uganda
College of Health Sciences, Makerere University, Kampala, Uganda

Abstract

Background: Tuberculous meningitis (TBM), a leading cause of meningitis in sub-Saharan Africa, is notoriously difficult to diagnose. In our Ugandan setting TB diagnostics have evolved rapidly in recent years, with introduction of Xpert MTB/Rif (Xpert) in 2011 and culture in 2013. We aim to describe the impact of improved TBM diagnostics at two Ugandan hospitals between 2010 and 2017.

Methods: Adults presenting with meningitis (headache and objective meningeal signs) were assessed for eligibility for enrolment in two consecutive trials investigating cryptococcal meningitis. Cohort one received cerebrospinal fluid (CSF) smear microscopy only (2010-2013). Cohort two received smear microscopy and Xpert on 1ml unprocessed CSF at physician discretion (2011-2013). Cohort three received smear microscopy, routine liquid-media culture and Xpert on large volume CSF (2013-2017) for all meningitis suspects with a negative CSF cryptococcal antigen (crAg). In a post-hoc analysis of three prospective cohorts, we compare rates of microbiologically confirmed TBM and hospital outcomes over time.

Results: 1672 predominantly HIV-infected adults underwent lumbar puncture, of which 33% (558/1672) had negative CSF crAg and 12% (195/1672) were treated for TBM. Over the study period, microbiological confirmation of TBM...
increased from 3% to 41% (P<0.01) and there was a decline in in-hospital mortality from 57% to 41% (P=0.27). Adjusting for definite TBM and antiretroviral therapy, and using imputed data, the odds of dying were nearly twice as high in cohort one (adjusted odds ratio 1.7, 95% CI 0.7 to 4.4) compared to cohort three. Sensitivity of Xpert was 63% (38/60) and culture was 65% (39/60) against a composite reference standard.

Conclusions: Since 2010, as TBM diagnostics have evolved, microbiologically-confirmed TBM diagnoses have increased significantly. There has been a non-significant decline in TBM in-hospital mortality but due to multiple possible confounding factors it is not possible to conclude what has driven this decline in mortality.

Keywords
Tuberculous meningitis, TBM, HIV, diagnosis, outcomes
Amendments from Version 1

Title
The title has been changed to: Tuberculous meningitis diagnosis and outcomes during the Xpert MTB/RIF era: a 6.5-year retrospective cohort study in Uganda

Methods:
As per reviewer suggestions we have expanded on the nature of the parent trials. To clarify who was eligible to be included in the diagnostic analysis and TBM cohort (n=195) we have added to the methods that "any patient from the 1672 patients screened who received testing for TBM (AFB smear, Xpert or culture) was eligible to be included in the diagnostic analysis and any patient who was treated for TBM was eligible to be included in the TBM cohort".

We have added a 2x2 table and a study schematic to make the flow of patients clearer as per the suggestion of Dr Wasserman and Dr Davis. It is not a classic Consort diagram as this population were not enrolled into a clinical trial.

We have added a detail to the methods section to define empiric TBM treatment "TBM treatment given in the absence of a positive microbiologic result, based on a high index of clinical suspicion, was defined as 'empiric TBM treatment'".

Results
The reviewers kindly pointed out a typographical error in the manuscript. The number of microbiologically confirmed cases is 74 (not 76) and the text has been amended accordingly.

Regarding the sensitivity of Xpert and Ultra against a composite reference standard a 2x2 table has been added to accompany Figure 2.

Discussion
In light of all three reviewer's comments we have expanded the discussion to further acknowledge the other potential confounding factors and concluding that it is not possible to draw a conclusion about whether Xpert has reduced TBM mortality in this setting.

See referee reports

Introduction
Tuberculous meningitis (TBM) is the second most common cause of adult meningitis in sub-Saharan Africa1,2, accounting for one to five percent of the 10.4 million tuberculosis (TB) cases reported worldwide in 2016. Despite treatment, TBM outcomes are poor with 19–28% mortality in HIV-uninfected persons and 40–67% mortality in HIV-infected patients in addition to long-term disability is frequent among survivors3-6.

Insidious symptom onset in persons with TBM leads to delay in seeking care and increasing disease severity at presentation correlates with higher mortality1. Further, the paucibacillary nature of TBM increases the difficulty in confirming diagnosis once care is sought, also contributing to high mortality1. Cerebrospinal fluid (CSF) smear microscopy for acid-fast bacilli (AFB smear) has poor sensitivity (~10–20%) in routine practice. Culture has improved sensitivity (~50-60%) but is not widely available in many resource constrained settings and commonly takes at least 2–3 weeks for liquid culture growth, which is too slow to guide decision-making at the time of presentation1.

In 2013, the World Health Organization endorsed the Xpert MTB/RIF (Xpert) assay (Cepheid, Sunnyvale, California, USA), a cartridge-based, polymerase chain reaction assay with a run time of 113 minutes, as the preferred initial test to investigate TB meningitis on the basis of a meta-analysis of 13 studies7. Of the two major studies included in the meta-analysis, Patel and colleagues reported 67% sensitivity against microbiologically proven TBM and 36% against consensus clinical case definitions, while Nhu and colleagues showed 59% sensitivity against the same case definitions8-10. Additionally, use of a larger volume of centrifuged CSF improves sensitivity of Xpert10,11. Yet, inadequate negative predictive value means that a negative Xpert result has limited influence on clinical decision making12.

There is evidence that use of Xpert for diagnosis of pulmonary TB reduces diagnostic delay, increases the rate of same day treatment, and decreases usage of empiric treatment13,14. However, for pulmonary TB, Xpert has not been shown to decrease mortality15-18. Yet, lessening diagnostic delay in persons with TBM may be more likely to lead to improved outcomes as compared to pulmonary TB given the high early mortality of TBM19. Whether routine use of Xpert for investigation of suspected TBM has made an impact on diagnosis or mortality has not yet been investigated.

Herein we describe TBM diagnosis and outcomes over a 6.5-year period in prospective cohorts at two Ugandan referral hospitals.

Methods
Study population
Uganda is a high burden HIV setting, with a prevalence is 6.2% among adults aged 15 to 64 years with an estimated 60% viral load suppression in 2017 among all HIV-infected adults20. Adults presenting with suspected meningitis (headache and neck stiffness +/- vomiting, fever, seizures, focal neurological deficits, or altered consciousness), to Mulago National Referral Hospital, Kampala, and Mbarara Regional Referral Hospital, were assessed for eligibility for enrolment in two consecutive randomised clinical trials investigating cryptococcal meningitis. The first trial Cryptococcal Optimal Antiretroviral Timing (COAT) investigated early versus delayed antiretroviral therapy in HIV-related cryptococcal meningitis (www.clinicaltrials.gov: NCT01075152) and the second, Adjunctive Sertraline for the Treatment of Cryptococcal Meningitis (ASTRO-CM) evaluated whether sertraline when added to standard amphotericin-based therapy for cryptococcal meningitis, lead to improved survival (NCT01802385). Screening began on 22nd November 2010 and continued until 28th May 2017. After an informed consent process for trial screening a diagnostic lumbar puncture was performed and baseline demographics and clinical information were recorded on all. Participants with non-cryptococcal meningitis were not enrolled into the clinical trials but followed until hospital discharge.

Any patient who received testing for TBM (CSF AFB smear, Xpert or mycobacterial culture) during this period was eligible to be included in the diagnostic accuracy analysis. Any patient who was ultimately treated for TBM was eligible to be included in one of the three TBM cohorts, from which data as used to compare rates of microbiological confirmation and outcomes. Cohort was determined by what type of TB testing they individual had undergone.
Microbiologically proven (definite) TB meningitis was defined as any positive AFB smear, culture or Xpert result from CSF testing. TBM treatment given in the absence of a positive microbiologic result, due to high index of clinical suspicion, was defined as ‘empiric TBM treatment’. Consensus uniform case definitions were used to categorise patients as definite, probable, possible or not TBM\(^{11}\). TBM treatment included 12 months of antituberculous therapy with 6–8 weeks of adjunctive corticosteroids as per Ugandan guidelines\(^{21}\).

Cohort definitions and diagnostic tests used
Cohort one (16th November 2010 until 28th May 2013) received only CSF AFB smear testing (Figure 1). If available, 1mL cryopreserved CSF was later tested with Xpert MTB/Rif when Xpert became available. Cohort two (1st April 2011 until 10th November 2013) underwent CSF AFB smear and Xpert MTB/Rif on a 1mL sample of uncentrifuged CSF. Testing was performed at physician discretion when there was lymphocytic pleocytosis and/or high degree of clinical suspicion. In the period of overlap of cohort one and two (April 2011–May 2013), Xpert testing was not being done on a routine basis; subjects were included in cohort two when Xpert was done in real-time and in cohort one if Xpert was not done, or only done at a later date on cryopreserved specimens.

In cohort three (11th November 2013 until 28th May 2017) all cryptococcal antigen negative (IMMY, Norman, Oklahoma, USA) patients were systematically investigated for the presence of TB meningitis, irrespective of physician discretion. Subjects had comprehensive testing for TBM with CSF AFB smear (Mulago Hospital only), Xpert MTB/Rif on large volume centrifuged CSF\(^{13}\) and CSF Mycobacteria Growth Inhibitor Tube culture (MGIT, Becton Dickinson, Franklin Lakes, USA). AFB smear was discontinued in Mbarara in 2013 as the sensitivity was deemed too low to justify further use. In patients with a confirmed diagnosis of cryptococcal meningitis (CM), if TBM co-infection was suspected, patients would be investigated for TBM at the physician’s discretion.

### Assessment of outcome
In-hospital outcome was determined from case report forms, hospital medical records or follow-up telephone calls with the patient or their surrogate where hospital outcome was unknown. The outcome was categorised as discharged alive, deceased prior to hospital discharge or unknown (i.e. self-discharged against medical advice in an imminently terminal patient, hospital outcome undetermined, transferred to another facility).

### Statistical methods
Comparisons of categorical and continuous demographic and clinical characteristics by cohort were performed using Fisher’s exact tests and Kruskal-Wallis tests, respectively. Sensitivity of Xpert MTB/Rif was evaluated against a composite reference standard (any positive CSF test - AFB smear, Xpert or culture i.e. definite TBM according to the uniform case definition)\(^{11}\). A separate analysis was conducted against the uniform case definition of probable or definite TBM\(^{11}\). Concordance between Xpert MTB/Rif and culture was evaluated with a kappa statistic and McNemar’s test. Invalid tests (e.g. culture contamination, Xpert error) were counted as negative results. Mortality was first compared by cohort for participants with a known outcome using Fisher’s exact test. Data for patients with unknown outcome was imputed to assume first that 50% within each cohort died, or that 75% died (both within the expected mortality range for this population). Odds ratios and 95% confidence intervals were computed from multivariable logistic regression models with these imputed data, adjusted for 1) ART status, and 2) ART status and definite TBM diagnosis. Imputations were repeated with new random assignments to confirm results. Analyses were conducted using SAS version 9.4 (The SAS Institute, Cary, NC) and p-values <0.05 were considered statistically significant.

### Ethics
Institutional review board approvals for the studies and the associated screening process were obtained locally in Uganda [ASTRO: Mulago Hospital Research Ethics Committee (approval

---

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>AFB smear</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AFB smear Xpert (1mL uncentrifuged CSF)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AFB smear Xpert (large vol centrifuged CSF)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MGIT culture</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Figure 1.** Timeline illustrating evolution of diagnostic testing.
number, MREC 429); COAT: Makerere University School of Medicine Research and Ethics Committee (approval number, REC Ref No. 2009-022)), from the University of Minnesota (USA), and by the Uganda National Council of Science and Technology. Written informed consent for screening or participation in the studies was obtained from all participants or from their surrogates (e.g. family member or guardian) where the patient had altered mental status and did not have the capacity to provide consent.

Results
Participant characteristics
Over the study period, 1672 patients with meningitis symptoms were assessed and underwent lumbar puncture: 1058 (63%) had a positive CSF cryptococcal antigen test, 558 (33%) had negative CSF cryptococcal antigen test (data missing, n=56). A total of 195 subjects were treated for TBM, see Figure 2. Overall 61% were male, median age was 35 years (IQR 30–42), 96% were HIV-positive, median CD4 count was 78 cells/μL (IQR 26–191) and the majority (69%) presented with British Medical Research Council severity grade II disease, see Table 1. Baseline characteristics were similar between cohorts with the exception of antiretroviral (ART) experience; 0% of participants were on ART in cohort one compared to 61% in cohort three (P<0.01).

Among the 74 cases of microbiologically proven TBM in this population with advanced HIV infection, 34% (25/74) had an acellular CSF (white cells <5 cells/μL) at presentation, and 4% (3/74) had a normal CSF profile (CSF cells <5 cells/μL, protein <45 mg/dL, and glucose >2.2mmol/l).

Method of diagnosis
Microbiological confirmation of TBM was made in 38% (74/195) of cases. The proportion of cases with microbiologically confirmed TBM (definite TBM) increased significantly, from 3% (1/33) in cohort one to 87% (13/15) in cohort 2 and 41% (60/147) in cohort 3 (P<0.01). Categorisation by uniform case definition is summarised in Table 2.

There was a marked difference in physician threshold for empiric TBM therapy between the two clinical sites. In cohort three, Mulago Hospital recorded 44 cases of which 77% (34/44) were microbiologically confirmed and 23% (10/44) were empirically treated, whilst Mbarara Hospital recorded 103 cases of which 25% (26/103) were microbiologically confirmed and 75% (77/103) were empirically treated.

Diagnostic accuracy of Xpert MTB/Rif
Xpert MTB/Rif was positive in 51 of 455 tested (11%), MGIT culture positive in 39 of 321 (12%) tested, AFB stain positive on 5 of 818 tested (1%), as summarised in Table 2. Diagnostic accuracy of Xpert and MGIT were analysed in cohort three, when both assays were done routinely, and 60 participants had a microbiologically confirmed diagnosis (composite reference standard). Sensitivity of Xpert was 63% (38/60) against the composite reference standard and 54% (38/71) against the uniform case definition (probable or definite TBM). Sensitivity of MGIT culture was 65% (39/60) against the composite reference standard of definite microbiologic-confirmed TBM.

Figure 2. Illustration of flow of patients from the screening population into the TBM cohort.
and 55% (39/71) against *uniform case definition* for probable or definite TBM.

Concordance between Xpert MTB/Rif and MGIT culture was analysed in the 118 with both Xpert and MGIT culture results available (Table 3). Either Xpert or MGIT culture was positive in 56 patients, of which only 30% (17/56) were positive by both modalities (kappa 0.23 95% CI [0.04, 0.41], p=0.01) (Figure 3). Neither method diagnosed significantly more cases than the other (p=0.42).

**Outcomes**

Hospital outcome was known for 142 participants, 53 had unknown outcomes or self-discharged against medical advice. Median time to death was 3 days (IQR 1–9 days) among those known to have died, and median length of hospitalization was 7 days (IQR 4–10 days) for participants known to have survived to hospital discharge. Among those with known outcomes, there was a non-significant decline in mortality from 57% in cohort one to 41% in cohort three (p=0.27) (Table 4). Assuming that 50% of those with unknown outcome died, and adjusting for ART status and definite TBM diagnosis at hospitalization, the odds of dying were approximately twice as high for cohort one (aOR 1.7 95% CI [0.7, 4.4]) and cohort two (1.8 [0.6, 5.6]) as compared to cohort three. Assuming that 75% of those with unknown outcome died, adjusted odds of death increase further, cohort one (4.0 [1.5, 10.9]) and cohort two (2.0 [0.6, 6.7]) compared to cohort three (Table 4, Figure 4).

**Discussion**

Rapid molecular diagnostics have been predicted to reduce TB-related mortality but little is reported about the impact of Xpert on TB-related mortality. Here we report diagnosis and clinical outcomes among hospitalized Ugandans treated for TB meningitis over a 6.5-year period. In-hospital mortality was high in the cohort overall (44% 95% CI [36,52%]), similar to other research settings with high HIV prevalence. The adjusted model found that odds of in-hospital mortality were almost two-fold higher in the earliest cohort, tested by CSF smear microscopy only, compared to that of the most recent cohort in which Xpert (and culture) were routinely performed. Though severity of TBM at presentation was similar over the study period.

### Table 1. Demographics, HIV details and outcomes of cohort.

<table>
<thead>
<tr>
<th>N with data</th>
<th>Cohort 1 Nov 2010 to May 2013</th>
<th>Cohort 2 Apr 2011 to Nov 2013</th>
<th>Cohort 3 Nov 2013 to May 2017</th>
<th>Total</th>
<th>P-value*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Diagnostics used</td>
<td>AFB smear</td>
<td>AFB smear Xpert</td>
<td>AFB smear Xpert Culture</td>
<td></td>
<td></td>
</tr>
<tr>
<td>N in TBM case cohort</td>
<td>33</td>
<td>15</td>
<td>147</td>
<td>195</td>
<td></td>
</tr>
<tr>
<td>Demographics</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sex</td>
<td>195</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>18 (55%)</td>
<td>8 (53%)</td>
<td>92 (63%)</td>
<td>118 (61%)</td>
<td>0.58</td>
</tr>
<tr>
<td>Age</td>
<td>195</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Median (IQR)</td>
<td>33 (29, 38)</td>
<td>35 (29, 40)</td>
<td>35 (30, 43)</td>
<td>35 (30, 43)</td>
<td>0.33</td>
</tr>
<tr>
<td>HIV details</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HIV status, n (%)</td>
<td>195</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HIV-positive</td>
<td>32 (97%)</td>
<td>15 (100%)</td>
<td>141 (96%)</td>
<td>188 (96%)</td>
<td>1.00</td>
</tr>
<tr>
<td>ART status, n (%)</td>
<td>179</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>On ART</td>
<td>0 (0%)</td>
<td>4 (27%)</td>
<td>80 (61%)</td>
<td>84 (47%)</td>
<td>&lt;0.01</td>
</tr>
<tr>
<td>ART naive</td>
<td>32 (100%)</td>
<td>11 (73%)</td>
<td>52 (39%)</td>
<td>95 (53%)</td>
<td></td>
</tr>
<tr>
<td>CD4</td>
<td>131</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Median (IQR)</td>
<td>12 (7, 121)</td>
<td>148 (54, 169)</td>
<td>78 (26, 206)</td>
<td>78 (26, 191)</td>
<td>0.30</td>
</tr>
<tr>
<td>TBM details</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MRC severity grade, n (%)</td>
<td>191</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>I</td>
<td>9 (27%)</td>
<td>4 (31%)</td>
<td>20 (14%)</td>
<td>33 (17%)</td>
<td>0.13</td>
</tr>
<tr>
<td>II</td>
<td>22 (67%)</td>
<td>7 (54%)</td>
<td>102 (70%)</td>
<td>131 (69%)</td>
<td></td>
</tr>
<tr>
<td>III</td>
<td>2 (6%)</td>
<td>2 (15%)</td>
<td>23 (16%)</td>
<td>27 (14%)</td>
<td></td>
</tr>
</tbody>
</table>

*P-values from Fisher’s exact tests for categorical variables and Kruskal-Wallis tests for continuous variables.*
Table 2. Methods of Diagnosis.

<table>
<thead>
<tr>
<th></th>
<th>Cohort 1</th>
<th>Cohort 2</th>
<th>Cohort 3</th>
<th>Total</th>
<th>P-value *</th>
</tr>
</thead>
<tbody>
<tr>
<td>AFB smear</td>
<td>471</td>
<td>71</td>
<td>1130</td>
<td>1672</td>
<td></td>
</tr>
<tr>
<td>Xpert</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Culture</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

All meningitis patients screened

<p>| | | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Total number</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cryptococcal Antigen positive</td>
<td>269</td>
<td>31</td>
<td>758</td>
<td>1058</td>
<td></td>
</tr>
<tr>
<td>Cryptococcal Antigen negative</td>
<td>187</td>
<td>38</td>
<td>333</td>
<td>558</td>
<td></td>
</tr>
</tbody>
</table>

TBM diagnostic tests performed

<p>| | | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>CSF AFB smear microscopy*</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>N AFB performed</td>
<td>466</td>
<td>71</td>
<td>281</td>
<td>818</td>
<td></td>
</tr>
<tr>
<td>N AFB positive</td>
<td>1 (0%)</td>
<td>0 (0%)</td>
<td>4 (2%)</td>
<td>5 (1%)</td>
<td></td>
</tr>
<tr>
<td>CSF TB culture</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>N TB culture performed</td>
<td>0</td>
<td>0</td>
<td>321</td>
<td>321</td>
<td></td>
</tr>
<tr>
<td>N TB culture positive</td>
<td>0</td>
<td>0</td>
<td>39 (12%)</td>
<td>39 (12%)</td>
<td></td>
</tr>
<tr>
<td>CSF Xpert MTB/Rif</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>N Xpert performed (realtime)</td>
<td>0</td>
<td>71</td>
<td>384</td>
<td>455</td>
<td></td>
</tr>
<tr>
<td>N Xpert positive</td>
<td>0</td>
<td>13 (18%)</td>
<td>38 (10%)</td>
<td>51 (11%)</td>
<td></td>
</tr>
</tbody>
</table>

Uniform case definition

<p>| | | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Definite</td>
<td>1 (3%)</td>
<td>13 (87%)</td>
<td>60 (41%)</td>
<td>74 (38%)</td>
<td>&lt;.01</td>
</tr>
<tr>
<td>Probable</td>
<td>5 (15%)</td>
<td>2 (13%)</td>
<td>11 (7%)</td>
<td>18 (9%)</td>
<td></td>
</tr>
<tr>
<td>Possible</td>
<td>22 (67%)</td>
<td>0 (0%)</td>
<td>53 (36%)</td>
<td>75 (38%)</td>
<td></td>
</tr>
<tr>
<td>Not</td>
<td>5 (15%)</td>
<td>0 (0%)</td>
<td>23 (16%)</td>
<td>28 (14%)</td>
<td></td>
</tr>
</tbody>
</table>

Prior to November, 2013 any patient not prospectively tested with Xpert was considered in Cohort 1. *AFB smear was initially performed on all meningitis patients regardless of CSF Cryptococcal antigen result. From October 2013, it was only performed on those with a negative Cryptococcal antigen, and was later stopped altogether in Mbarara. *P-value from Fisher’s exact test

Table 3. Summary of concordance between Xpert MTB/Rif and MGIT culture results.

<table>
<thead>
<tr>
<th>Diagnostic Test</th>
<th>Xpert MTB/Rif</th>
<th>Total</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>positive</td>
<td>negative</td>
<td></td>
</tr>
<tr>
<td>MGIT culture</td>
<td>17</td>
<td>22</td>
<td>39</td>
</tr>
<tr>
<td></td>
<td>17</td>
<td>62</td>
<td>79</td>
</tr>
<tr>
<td>Total</td>
<td>34</td>
<td>84</td>
<td>118</td>
</tr>
</tbody>
</table>

P-value from McNemar’s test
N=118 (with both Xpert and culture results)
Figure 3. Venn diagram illustrating the overlap of positive MGIT culture and Xpert test results in the n=118 samples tested with both assays. A total of 118 adults were tested with both MGIT culture and Xpert, of which 22 were positive by MGIT culture, 17 by Xpert and 17 by both tests. Neither test performed better than the other, p=0.423 by McNemar's. A kappa statistics value of 0.23 95%CI [0.04, 0.41], p=0.01, suggests only slight agreement of the two assays.

Table 4. Hospital outcomes.

<table>
<thead>
<tr>
<th></th>
<th>Cohort 1 Nov 2010 to May 2013</th>
<th>Cohort 2 Apr 2011 to Nov 2013</th>
<th>Cohort 3 Nov 2013 to May 2017</th>
<th>Total</th>
<th>P-value*</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>N in TBM case cohort</strong></td>
<td>33</td>
<td>15</td>
<td>147</td>
<td>195</td>
<td></td>
</tr>
<tr>
<td><strong>Outcome of hospitalization</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Unknown</td>
<td>26 (79%)</td>
<td>4 (27%)</td>
<td>23 (16%)</td>
<td>53 (27%)</td>
<td></td>
</tr>
<tr>
<td>Known</td>
<td>7 (21%)</td>
<td>11 (73%)</td>
<td>124 (84%)</td>
<td>142 (73%)</td>
<td></td>
</tr>
<tr>
<td>Discharged Alive</td>
<td>3 (43%)</td>
<td>4 (36%)</td>
<td>73 (59%)</td>
<td>80 (56%)</td>
<td>0.27</td>
</tr>
<tr>
<td>Died</td>
<td>4 (57%)</td>
<td>7 (64%)</td>
<td>51 (41%)</td>
<td>62 (44%)</td>
<td></td>
</tr>
</tbody>
</table>

**Odds Ratio (Mortality) and 95% CI (on imputed data)**

**Assuming 50% of unknowns died**

<table>
<thead>
<tr>
<th></th>
<th>Adjusted for ART status</th>
<th>Adjusted for ART status and confirmed TBM</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1.5 (0.6,3.6)</td>
<td>2.0 (0.7,6.2)</td>
</tr>
<tr>
<td></td>
<td>1.7 (0.7,4.4)</td>
<td>1.8 (0.6,5.6)</td>
</tr>
</tbody>
</table>

**Assuming 75% of unknowns died**

<table>
<thead>
<tr>
<th></th>
<th>Adjusted for ART status</th>
<th>Adjusted for ART status and confirmed TBM</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>3.3 (1.3,8.4)</td>
<td>2.5 (0.8,7.8)</td>
</tr>
<tr>
<td></td>
<td>4.0 (1.5,10.9)</td>
<td>2.0 (0.6,6.7)</td>
</tr>
</tbody>
</table>

Overall median (IQR) time in hospital was 7 (4, 10) days among those who were known to be discharged alive, and 3 (1, 9) days among those who were known to have died in hospital.

*P-value from Fisher’s exact test comparing KNOWN discharged alive vs KNOWN died; Odds ratios are the odds of being discharged alive, assuming 50% and 75% of those with unknown outcome died.

...and TBM treatment recommendations have not changed for Uganda, there are multiple other potential confounding factors due to the nature of the study. Improved access to ART, strengthening in healthcare services, increased awareness of TBM amongst communities and healthcare workers or changes in employment of empiric treatment all may have occurred over the study period and if present, may have impacted our findings. Due to these potential confounding factors, it is not possible to draw a conclusion about whether Xpert has reduced TBM mortality in this setting.
The proportion of ART experienced subjects increased significantly over time with improved access to ART treatment in Uganda and because the parent trial in cohort one enrolled only ART naïve subjects. Although ART status was not associated with mortality, we adjusted for ART in multivariable models due to the large discrepancy in ART status between cohorts. Nonetheless the impact of ART use is likely to have played a role in the observed decline in mortality.

Despite a non-significant decline in mortality, a current case-fatality rate of 41% remains unacceptably high and highlights the remaining work required to achieve the WHO goal of reducing TB-related deaths by 90% by 2030. Initiating treatment in the early stage of disease is the single most important factor in improving outcomes. Earlier presentation to the hospital is essential for prompt diagnosis and treatment initiation, yet, 83% of our cohort presented with MRC grade II or III disease.

Once the patient presents to care, an affordable, rapid, and reliable test that can effectively confirm or rule out TBM is crucial for prompt diagnosis. In this predominantly HIV-positive TBM cohort, sensitivity of Xpert was 63% against the composite reference standard. Thus, even though results were available rapidly, Xpert missed over one in three cases. The next generation assay Xpert MTB/Rif Ultra has an analytic limit of detection of 15 colony forming units (CFU/ml), compared to 113 CFU/ml for Xpert. Ultra appears to be significantly more sensitive than Xpert or culture for the diagnosis of TBM.

**Figure 4. Illustration of odds of dying in cohort one and two compared to cohort three in a multivariate model.** Odds ratios (and 95% confidence intervals) for death by the end of hospitalization comparing cohorts 1 and 2 to cohort 3, computed from multivariable logistic regression models with imputed data, adjusted for (1) ART status, and (2) ART status and definite TBM diagnosis. Data for patients with unknown outcome was imputed to assume that 50% within each cohort died, or that 75% died. In all models, neither ART nor definite TBM status had a significant association with in-hospital mortality, adjusted for cohort.
versus 45% and 45% respectively, P<0.001)28. Whether Ultra can reduce diagnostic delay and improve outcome from TBM requires further prospective evaluation.

Where both Xpert and MGIT had been done, less than a third (23%, 17/74) of confirmed cases were positive by both modalities. This is consistent with prior findings and is likely due to the relatively higher sensitivity of culture versus Xpert, and the ability of Xpert to detect non-viable TB bacilli damaged by host-immune response and/or antituberculous drug therapy13,28. Neither test performed better than the other (P=0.42) and currently CSF culture continues to earn a place in TBM diagnostics testing algorithms.

Until a rapid test with suitable negative predictive value is widely available, there is likely to be on-going heterogeneity in clinical practice regarding frequency of empiric TBM therapy (TB treatment in the absence of a positive result). Here, Mulago Hospital participants were treated for TBM on an empiric basis in under one quarter of cases as opposed to over three quarters of cases at Mbarara Hospital, despite the same diagnostic armamentarium being available in both settings. In our study settings, it is likely that individual or departmental clinical practice and risk thresholds for TB treatment initiation dictate this variation. Though empiric TBM therapy is potentially life-saving, significant risks such as side effects, drug-interactions and adjunctive steroids in an already immunosuppressed population need to be considered. Ideally, a rapid, accurate test allows therapy for TB meningitis to be started promptly only in those who actually have TBM. Overall, the proportion with microbiologically confirmed TBM increased significantly from 3% in cohort one to 41% in cohort three (P<0.01). In cohort two, selection bias could have impacted on results as Xpert was only performed in cases where there was extremely high index of suspicion and empiric treatment was given only twice in those with a negative Xpert (4%, 2/56). The low number of empiric diagnoses during this period were likely due to over-confidence in Xpert’s ability to rule-out TBM. As understanding regarding the limitations of Xpert for the diagnosis of TBM became known, empiric TBM treatment rose24.

Limitations of this study relate to the nature of available data namely missing data on hospital and long-term outcomes, the time to starting TB treatment, unbalanced numbers in each cohort, and other un-adjusted confounding factors and potential sources of bias. When imputing data in the model we assumed that either 50% or 75% of patients with unknown outcome actually died, which is a clinically reasonable judgment for this population6.

Here we present important data on diagnostic confirmation and TBM mortality during a period of TB diagnostic evolution. There has been a significant increase in microbiological confirmation and a modest, albeit non-significant, decline in mortality since introduction of Xpert and culture in our study setting. An on-going multifaceted approach is needed to further reduce death and disability from TBM.

**Consent**

Written informed consent for publication of the anonymised data was obtained from the participants or their surrogates.

**Data availability**

The database contains individual level data and as such is not available through an open-access data repository. The database is stored on a secure server at University of Minnesota. Researchers interested in accessing the data can contact the corresponding author (FVC), the last author (DRB) or the Division of Biostatistics at the University of Minnesota. Data access will be granted to active researchers in the field with the agreement of the authors.

**Competing interests**

No competing interests were disclosed.

**Grant information**

FVC is supported by the Wellcome Trust [210772/Z/18/Z]. ASB, DBM, and DRB are supported by Fogarty International Center and National Institute of Neurologic Diseases and Stroke [R01NS086312].

The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

**Acknowledgements**

FVC is an honorary fellow of the Makerere University – Uganda Virus Research Institute Centre of Excellence for Infection and Immunity Research and Training (MUII-plus). MUII-plus is supported through the DELTAS Africa Initiative (grant no. 107743). The DELTAS Africa Initiative is an independent funding scheme of the African Academy of Sciences (AAS), Alliance for Accelerating Excellence in Science in Africa (AESA), and supported by the New Partnership for Africa’s Development Planning and Coordinating Agency (NEPAD Agency) with funding from the Wellcome Trust (grant no. 107743) and the UK Government.

**References**


Open Peer Review

Current Referee Status: ✔️❓✔️

Version 2

Referee Report 12 July 2018
doi:10.21956/wellcomeopenres.15994.r33441

Angharad Davis ¹, Sean Wasserman ²
¹ National Hospital for Neurology and Neurosurgery, University College London Hospitals, London, UK
² University of Cape Town, Cape Town, South Africa

I have read through the comments. Many thanks to the authors. We are happy that the points have been sufficiently addressed. We are happy that the paper should be accepted and no more changes are required.

**Competing Interests:** No competing interests were disclosed.

We have read this submission. We believe that we have an appropriate level of expertise to confirm that it is of an acceptable scientific standard.

Version 1

Referee Report 15 June 2018
doi:10.21956/wellcomeopenres.15907.r33229

Tom Boyles
Wits Reproductive Health and HIV Institute, University of the Witwatersrand, Johannesburg, South Africa

I am satisfied that the authors have adequately addressed the points I made in my review.

**Competing Interests:** No competing interests were disclosed.

**Referee Expertise:** Clinical management of patients with HIV and meningitis. The interpretation of tests and thresholds for treatment.

I have read this submission. I believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard.
Many thanks for requesting us to review this article which addresses an important unanswered question in TB diagnostics. We have made specific points regarding each section of the manuscript as well as giving an overall impression.

This prospective cohort studies observes differences in diagnosis and hospital outcomes related to TBM across three cohorts of patients recruited over a 6.5 year period in Uganda. The study aims to measure the impact on diagnosis and hospital outcomes with the addition of Gene Xpert testing or Gene Xpert and routine liquid media culture to standard diagnostic testing (CSF smears) over time.

Abstract

The study design is unclear: is this a retrospective or post-hoc analysis of data collected from other studies? The nature of the parent studies should be mentioned.

The denominator is not clearly reported: the inclusion criteria for the current analysis appears to be “all meningitis suspects with a negative CSF cryptococcal antigen.” Were 558 patients with negative CSF cryptococcal antigen test included in the analysis?

The conclusion that “in-hospital mortality has declined” seems exaggerated because this trend towards lower mortality over the study periods was not significant, and may have been due to other factors unrelated to changing diagnostics.

Methods

The inclusion criteria should be clearly defined and justified.

Differences between the cohorts could have introduced bias towards improved outcomes (type 1 error): cohorts 1 and 2 at physician discretion, cohort 3 were systematically investigated. This should be addressed in the discussion.

Please explain how the performance of a test (eg Xpert) can be compared to a reference standard that includes the test itself (“any positive CSF test - AFB smear, Xpert or culture”)?

Were the variables in the multivariate model selected a priori?

Was ethics permission obtained for this sub-study/analysis?

Results

The denominator is not clearly reported: the inclusion criteria for the current analysis appears to be “all meningitis suspects with a negative CSF cryptococcal antigen.” Were 558 patients with negative CSF cryptococcal antigen test included in the analysis, or was it 195 subjects that were treated for TBM? This should match the inclusion criteria, which appear to be any patient with suspected meningitis and negative CrAg testing.

A large proportion of patients in cohorts 2 and 3 were on ART at the time of presentation: this could alter their prognosis and bias the results. Although the model adjusted for this, it should be emphasised more in the discussion.

There is a discrepancy in the number of microbiologically proven cases:

- “Among the 76 cases of microbiologically proven TBM”
- “Microbiological confirmation of TBM was made in 38% (74/195) of cases”

Why was the number of confirmed cases so high in Cohort 2?

The paragraph describing empiric treatment practices requires clarification: how was ‘empiric treatment’ defined? The numbers appear to suggest that no patients with a confirmed diagnosis
were empirically treated (i.e., empiric treatment = no microbiological confirmation). Suggest changing the terminology to make this clearer – 1. Treated based on positive result; 2. Treated without a positive result. It seems that in Mulago Hospital a higher proportion of treated patients had positive results, and therefore did not require ‘empiric treatment.’ So, the important question seems to be why there was such a bug discrepancy in the number of confirmed cases between the two sites, allowing ‘empiric treatment’ to be avoided?

- The denominator is again unclear for the diagnostic accuracy section: where do the “455 tested” come from if only 195 cases were included in the study (or is it 558?). Suggest including a consort diagram to explain the patient populations for this study.
- Suggest including a 2 x 2 table to demonstrate the performance of diagnostic tests. Although the sensitivity of Xpert is 63% compared to a reference standard, this this standard is only present in 40% of cases (60/147) in Cohort 3.

**Discussion**

- See comments related to discussion above
- The main (and probably only data-driven) conclusion is that more confirmed TBM cases were detected with the use of Xpert and culture. The authors did not convincingly show an improvement in outcomes. Any observed trends could be secular, or related to other confounding factors resulting from different populations, selection criteria, management practices, etc.

The study observes three cohorts recruited from two parent trials. Although the data provides an important insight into the unanswered questions above, the suitability of the cohorts as comparators to one another does pose some limitation to how the results can be interpreted. Namely there is large variation between the observed cohorts in a number of aspects namely the size of cohorts, ART status, and availability of follow up data. Although ART status had been accounted for in analysis, the impact of absent follow up data which is much more prominent for cohort 1 (79%) compared to cohort 3 (16%) should be made clearer in the discussion. Also, how were the patients allocated to cohort 1 and 2 during the overlapping study period (1st April 2011 until 28th May 2013)? Were those where there was more diagnostic doubt allocated to receive testing with GeneXpert (and therefore subsequently included in cohort 2 than cohort 1)? could this have affected outcomes observed between the two groups?

It is possible that apart from the mentioned variables, that over time, other aspects may have influenced hospital outcome in these patients over a period where there has naturally been some advances made in TBM care besides the diagnostics discussed. For example the use of imaging diagnostics, time to treatment, availability of drugs, better understanding of supportive measures in TBM. I feel that these should form part of the discussion here and although they may not be quantifiable in this setting, should be acknowledged as possible factors for the observed differences over time. In this setting specifically were there changes in resource allocation? Did hospital facilities change?

**Is the work clearly and accurately presented and does it cite the current literature?**

Yes

**Is the study design appropriate and is the work technically sound?**

Partly

**Are sufficient details of methods and analysis provided to allow replication by others?**

Partly

**If applicable, is the statistical analysis and its interpretation appropriate?**
Yes

Are all the source data underlying the results available to ensure full reproducibility?
No source data required

Are the conclusions drawn adequately supported by the results?
Partly

**Competing Interests:** No competing interests were disclosed.

We have read this submission. We believe that we have an appropriate level of expertise to confirm that it is of an acceptable scientific standard, however we have significant reservations, as outlined above.

Author Response 26 Jun 2018

Fiona Cresswell, London School of Hygiene & Tropical Medicine, UK

Thank you Dr Wasserman and Dr Davis for your many comments, questions and suggestions, which we will address this individually:

Abstract

1. The study design is unclear: is this a retrospective or post-hoc analysis of data collected from other studies? The nature of the parent studies should be mentioned.

The nature of the parent studies has been expanded in the methods section of the revised manuscript. We have also added further information to the abstract and clarified that this is a post-hoc analysis of prospective cohorts.

2. The denominator is not clearly reported: the inclusion criteria for the current analysis appears to be “all meningitis suspects with a negative CSF cryptococcal antigen.” Were 558 patients with negative CSF cryptococcal antigen test included in the analysis?

The main analysis for microbiological confirmation and mortality relates to the 195 people within the three TBM prospective cohorts. Of the 1672 screened any patient who received any testing for TBM (CSF AFB smear, Xpert or mycobacterial culture) during this period was eligible to be included in table 2 and the diagnostic accuracy of Xpert and culture analysis. Any patient who was ultimately treated for TBM (n=195) was eligible to be included in the three prospective TBM cohorts, and it was this group on whom the rate of microbiological confirmation and outcomes were compared. We have added a schematic diagram to the manuscript which i hope clarifies the nature of the study population.

3. The conclusion that “in-hospital mortality has declined” seems exaggerated because this trend towards lower mortality over the study periods was not significant, and may have been due to other factors unrelated to changing diagnostics

You are absolutely correct. We have changed to text in the abstract to read “Since 2010, as TBM diagnostics have evolved, microbiologically-confirmed TBM diagnoses have increased and there has been a non-significant decline in TBM in-hospital mortality. Due to multiple possible confounding factors it is not possible to conclude what has driven this decline in mortality”.


Methods
1. The inclusion criteria should be clearly defined and justified

The following text has been added to the revised manuscript:
"Adults presenting with suspected meningitis (headache and neck stiffness +/- vomiting, fever, seizures, focal neurological deficits, or altered consciousness), to Mulago National Referral Hospital, Kampala, and Mbarara Regional Referral Hospital, were assessed"  
"Participants with non-cryptococcal meningitis were not enrolled into the clinical trials but followed until hospital discharge".  
"Any patient who was ultimately treated for TBM was eligible to be included in one of the three TBM cohorts, on which rates of microbiological confirmation and outcomes were compared. Cohort was determined by what type of TB testing they had undergone".

2. Differences between the cohorts could have introduced bias
This is a valid observation. Text stating "selection bias could have impacted on results" has been added to the discussion.

3. Please explain how the performance of a test (eg Xpert) can be compared to a reference standard that includes the test itself ("any positive CSF test - AFB smear, Xpert or culture")?

Whilst we acknowledge this is an imperfect reference standard there is not a suitable reference standard that can be used in TBM diagnostic accuracy studies. Importantly, each of the tests is know to have a high specificity so the chance of false positives is extremely low, especially in a symptomatic population with a high disease prevalence. Latent class analysis would be a potential statistical approach that could be used in such cases which lack a perfect reference standard, and we would be interested to do such an analysis in future.

4. Were the variables in the multivariate model selected a priori?

Yes

5. Was ethics permission obtained for this sub-study/analysis?

Yes, the informed screening consent process sought approval for both storage of specimens for future research and use of data for analyses / research relating to meningitis.

Results
1. The denominator is not clearly reported: the inclusion criteria for the current analysis appears to be “all meningitis suspects with a negative CSF cryptococcal antigen.” Were 558 patients with negative CSF cryptococcal antigen test included in the analysis, or was it 195 subjects that were treated for TBM? This should match the inclusion criteria, which appear to be any patient with suspected meningitis and negative CrAg testing.

I hope the response to comment on abstract has clarified this.

2. A large proportion of patients in cohorts 2 and 3 were on ART at the time of presentation: this could alter their prognosis and bias the results. Although the model
adjusted for this, it should be emphasised more in the discussion.

This has been emphasised in the discussion of the revised manuscript.

3. There is a discrepancy in the number of microbiologically proven cases

Thanks for spotting this typo. This has been corrected to 74.

4. Why was the number of confirmed cases so high in Cohort 2

The review by Dr Hamers also made this comment. Please see response to his comment for a potential explanation.

5. The paragraph describing empiric treatment practices requires clarification: how was ‘empiric treatment’ defined? The numbers appear to suggest that no patients with a confirmed diagnosis were empirically treated (ie empiric treatment = no microbiological confirmation). Suggest changing the terminology to make this clearer.

Thank you. We have defined this more clearly in the methods.

6. So, the important question seems to be why there was such a big discrepancy in the number of confirmed cases between the two sites, allowing ‘empiric treatment’ to be avoided?

Both hospitals have the same available diagnostics so I believe the difference in empiric treatment relates to local or personal thresholds applied to the initiation of TB treatment in the absence of a confirmatory test. You will no doubt have observed different treatment thresholds amongst colleagues, especially faced with a critically ill patient. As Mbarara Hospital is a relatively small hospital with good longterm retention of staff, especially in the clinical research setting, it is possible that a handful of clinicians may have seen the majority of TBM cases and been more willing to initiate TB treatment despite negative tests than in Kampala where there is a bigger team and clinicians rotate regularly.

7. Suggest including a consort diagram to explain the patient populations for this study.

Thanks for this suggestion. We have included a schematic of patient flow from the screening population into the cohort. It is not a classic consort diagram since this population were not part of an RCT.

8. Suggest including a 2 x 2 table to demonstrate the performance of diagnostic tests.

We have added a 2x2 table.

Discussion

1. The main (and probably only data-driven) conclusion is that more confirmed TBM cases were detected with the use of Xpert and culture. The authors did not convincingly show an improvement in outcomes. Any observed trends could be secular, or related to other confounding factors resulting from different populations, selection criteria, management practices, etc.
Indeed. We acknowledge the limitations of the data available to us and have tempered the conclusion accordingly in the revised manuscript.

2. For example the use of imaging diagnostics, time to treatment, availability of drugs, better understanding of supportive measures in TBM. I feel that these should form part of the discussion here and although they may not be quantifiable in this setting, should be acknowledged as possible factors for the observed differences over time. In this setting specifically were there changes in resource allocation? Did hospital facilities change?

This a good thoughts and we concede that many potential confounders exist. However, we could show from our limited data that severity at presentation across cohorts was the same. Whilst the health services in the private sector have evolved in this time period the government facilities remain under resourced without access to routine blood tests and imaging. Thankfully access to TB medication has been stable for many years. Healthcare worker practices may certainly have changed and we recognise this as a potential confounder.

Thank you again for your detailed review. I hope the responses and amendments are satisfactory.

**Competing Interests:** nil
- The authors note no major changes in treatment practices over time, although they also report substantial differences between the 2 hospitals in ability to confirm diagnosis and start empirical treatment, which may have impacted on outcomes (if varied with time).
- Referral practices and general awareness among care providers may have improved in the study setting (the authors note they do not have data on time to Rx initiation), especially given that the authors have implemented 2 large clinical trials in the same period.
- The patient case mix may have changed over time in many other ways (e.g. better nutrition, better education, lower TB incidence, better HIV diagnosis and treatment, etc)

My recommendation is to reword the title, to avoid misleading the readership.

2. The authors have (deliberately) done separate analyses on the 3 cohorts, and combining historical cohorts may have its limitations. Nonetheless, the combined cohort offers opportunities to attempt answering the main question. In this respect, it would be interesting if the authors could undertake a multivariate analysis of the combined cohort data to identify what are the main determinants of mortality (adjusting for time period). This could help to establish whether indeed the diagnostics influenced outcomes, or that the survival gain was mainly driven by other factors (e.g. ART). In my view, adding this analysis could strengthen the paper.

Minor points

1. The % of microbiological confirmation is 3% in cohort #1, 87% in cohort #2, and 41% in cohort #3. It would be relevant to learn why the latest % seems to be lower than the middle one (despite better diagnostic protocols).

2. The degree of agreement between Xpert and MGIT is very low (kappa 0.23). I would like to see a better explanation for this discrepancy, and how this knowledge can be applied in recommendations for testing algorithms in practice.

Is the work clearly and accurately presented and does it cite the current literature?
Yes

Is the study design appropriate and is the work technically sound?
Partly

Are sufficient details of methods and analysis provided to allow replication by others?
Yes

If applicable, is the statistical analysis and its interpretation appropriate?
Yes

Are all the source data underlying the results available to ensure full reproducibility?
Yes

Are the conclusions drawn adequately supported by the results?
Yes

**Competing Interests:** No competing interests were disclosed.
I have read this submission. I believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard, however I have significant reservations, as outlined above.

---

**Author Response 26 Jun 2018**

**Fiona Cresswell**, London School of Hygiene & Tropical Medicine, UK

We thank Dr Hamers for his insightful and helpful comments.

1. **The title suggests that the aim of the study was to investigate if better TBM diagnostics can lead to improved survival. After reading the article, this question cannot be answered.**

   In light of your observation we have changed the title to "Tuberculous meningitis diagnosis and outcomes during the Xpert MTB/Rif era: a 6.5-year retrospective cohort study in Uganda". We hope this better reflects what can actually be concluded from this retrospective data analysis.

2. **Indeed, many other factors could have influenced the differential outcomes across the 3 historical cohorts that span the time period of 2011-2017, and I feel the authors could present these complex factors in a more structured and comprehensive manner in the Discussion.**

   Thank you for this suggestion. We have taken on board and revised the discussion to acknowledge the multiple potential unadjusted confounding factors and the inability to draw a conclusion about the impact of diagnostics on outcomes from the data we have available.

3. **The authors have (deliberately) done separate analyses on the 3 cohorts, and combining historical cohorts may have its limitations. Nonetheless, the combined cohort offers opportunities to attempt answering the main question. In this respect, it would be interesting if the authors could undertake a multivariate analysis of the combined cohort data to identify what are the main determinants of mortality (adjusting for time period). This could help to establish whether indeed the diagnostics influenced outcomes, or that the survival gain was mainly driven by other factors (e.g. ART). In my view, adding this analysis could strengthen the paper.**

   Thank you for this suggestion. Assessing risk factors for mortality was not the primary intent of the paper but we agree this would make an interesting analysis in future. The TBM testing cohorts in this study generally represent testing time period and cohort is the primary variable of interest for this study. Since ART use changed over time and definite TBM is associated with increased risk of mortality, these covariates were chosen to be included in adjusted analyses. Mortality was unknown for 27% of participants, so we conducted a series of models with imputations assuming that 50% within each cohort died, or that 75% died (both within the expected mortality range for this population). Six multivariable models of in-hospital mortality were run and all cohorts were included in all models. Cohort 3 was chosen as the reference category for presenting the odds ratios of interest since it represents the most current testing era. In all models, neither ART nor definite TBM status had a significant association with in-hospital mortality. All models run include all cohorts and the odds ratios comparing cohorts from the adjusted models are presented in Figure 4 in the (revised) manuscript.

4. **The % of microbiological confirmation is 3% in cohort #1, 87% in cohort #2, and 41% in**
cohort #3. It would be relevant to learn why the latest % seems to be lower than the middle one (despite better diagnostic protocols).

This is an interesting observation and we believe relates to two factors:

a) In cohort 2, during the initial period after Xpert introduction, it was only being used on cases with an extremely high clinical index of suspicion so there was a selection bias in population tested.
b) The % of confirmed case amongst the total number treated for TBM in cohort 2 appears spuriously high as the amount of empiric treatment was very low during the cohort. One potential reason is that there was an over confidence in Xpert's ability to be able to rule out TBM. In subsequent years a number of papers showed that the negative predictive value of Xpert is such that a negative test result should not deter TB treatment and empiric treatment was used more frequently again in out setting.

5. The degree of agreement between Xpert and MGIT is very low (kappa 0.23). I would like to see a better explanation for this discrepancy, and how this knowledge can to be applied in recommendations for testing algorithms in practice.

This is another interesting observation, and a somewhat surprising finding to us also. We have not elaborated in the discussion in the interest of brevity but possible explanations are mentioned here:

**Positive on Xpert whilst negative on culture:**
We lack data about the number of cases in the cohort who were on TB treatment at the time of lumbar puncture. It is plausible that if TB treatment had been initiated in the days or weeks prior to the CSF analysis the mycobacterium may have been rendered non-viable for culture. In our setting many patients present with a disseminated TB picture, some of whom have been coughing for several weeks prior to development of neurological symptoms and already initiated on antituberculous therapy from outpatient settings. We hope to capture this data more comprehensively in prospective studies.

**Positive on culture whilst negative on Xpert:**
Culture has a lower limit of detection that Xpert so patients with bacillary burdens on the <100CFU/ml region would likely be identified by culture only.

**Considerations for future testing algorithms:**
It is hopeful that the next generation assay Xpert MTB/Rif 'Ultra' will have a limit of detection similar to that of culture, so given that it can also detect non-viable bacilli, it is likely to perform better the culture in the diagnosis of TBM (as was the case in Bahr N. Lancet ID. 2018).
Testing algorithms must be customised and take into consideration the setting, HIV-prevalence, cost and the volume of CSF available for testing, amongst other things. Splitting a small CSF sample between multiple assays is likely to be counterproductive. In our setting we would recommend Ultra as the initial CSF test for TBM (in patients with a negative CSF cryptococcal antigen lateral flow assay). Culture remains an important adjunctive test especially in patients where there is a risk of drug resistance.

We have submitted a revised manuscript and hope these changes and responses are to Dr Hamer's satisfaction.

**Competing Interests:** nil
Tom Boyles

Wits Reproductive Health and HIV Institute, University of the Witwatersrand, Johannesburg, South Africa

Thank you for the opportunity to review this article which describes the investigation, treatment and outcomes of a cohort of mostly HIV infected adults in Uganda who were suspected of having TB meningitis. It is well written and easily reproducible by others.

The title asks the question of whether improved diagnostics can reduce mortality in TBM. It might be better to say ‘Have improved diagnostics improved outcomes?’ My view of the data is that it provides very limited evidence that improved diagnostics have improved outcomes of TBM in this cohort. The small improvements in outcome over time are statistically non-significant and due to the observational design are likely to be influenced by numerous unmeasured confounding factors. For example, the experience of clinician may well have changed over time, we know that empiric treatment was used differently in the 2 centres but it might also have changed over time. Part of the work-up of a patient with suspected TBM is a search for extra-neural TB and we do not know how this changed over time although it is likely that Xpert MTB/RIF was used for non-CSF samples in later cohorts compared to earlier, the same might be said for urine LAM.

My view as a reader is therefore that this data does not provide convincing evidence that the change in diagnostics was the driver of the small changes in mortality. It must be remembered that new diagnostics also have the potential to worsen outcomes, particularly if clinicians miss-interpret negative tests as ruling out the condition as might well occur with Xpert MTB/RIF and TBM.

I think that this well written work definitely deserves to be published but feel that the conclusions should reflect a greater level of uncertainty.

Minor points:

The abstract says that all 1672 patients were HIV infected but in the results section is says 96% were HIV infected so there is a discrepancy.

Last line, para 1 of introduction, probably remove word ‘to’.

Last line. Para 3 of introduction- It is not that Xpert cannot substitute for clinical judgement- Xpert is used to enhance clinical judgement, it’s just that the poor sensitivity means that when negative it has limited influence on decision making

Statistical methods- the uniform case definition is probable or definite TBM- do the authors think this is a reasonable reference standard for other tests? My view is that the threshold for treating TBM is very low and that patients with possible/probably or definite TBM should receive treatment and therefore this would be a more appropriate reference standard- what are your views on that?

Methods- Not clear if there were 76 or 74 microbiologically proven cases

Discussion- Para 1, repeat of ‘that of’

Discussion- Para 4, First sentence implies that tests are the only answer to the diagnosis of TBM, what
about clinical prediction rules relying on clinical data—although so far they have not been very successful, neither have tests so a robust CPR may negate the need for tests.

Discussion- Para5, Do the authors really think there are dead bacilli in CSF? Some of the authors have argued the opposite in their recent paper on Xpert Ultra so you can’t have it both ways.

Is the work clearly and accurately presented and does it cite the current literature?
Yes

Is the study design appropriate and is the work technically sound?
Yes

Are sufficient details of methods and analysis provided to allow replication by others?
Yes

If applicable, is the statistical analysis and its interpretation appropriate?
Yes

Are all the source data underlying the results available to ensure full reproducibility?
Yes

Are the conclusions drawn adequately supported by the results?
Partly

Competing Interests: No competing interests were disclosed.

Referee Expertise: Clinical management of patients with HIV and meningitis. The interpretation of tests and thresholds for treatment.

I have read this submission. I believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard, however I have significant reservations, as outlined above.

Author Response 26 Jun 2018
Fiona Cresswell, London School of Hygiene & Tropical Medicine, UK

Dear Dr Boyles,

Thank you for providing a considered and constructive of our paper. I’d like to respond to each comment in turn.

1. The findings provide limited evidence that improved diagnostics have improved outcomes of TBM in this cohort and due to the observational design are likely to be influenced by numerous unmeasured confounding factors. I think that this well written work definitely deserves to be published but feel that the conclusions should reflect a greater level of uncertainty.

We acknowledge the limitations of the retrospective data we present and the inability to account for
a multitude of other confounding factors. The title of the manuscript has been changed to "Tuberculous meningitis diagnosis and outcomes during the Xpert MTB/Rif era: a 6.5-year retrospective cohort study in Uganda". In the conclusion of the revised manuscript we given more emphasis to uncontrolled confounding factors the fact that the data on impact of improved diagnostics on outcomes in inconclusive.

2. Part of the work-up of a patient with suspected TBM is a search for extra-neural TB and we do not know how this changed over time although it is likely that Xpert MTB/RIF was used for non-CSF samples in later cohorts compared to earlier, the same might be said for urine LAM.

Indeed, sputum samples may be undergone testing with Xpert to assist in the diagnosis of TBM in cohort 3. TB-LAM has only become available in 2018 in Uganda so this will not have impacted our study which concluded in May 2017.

3. It must be remembered that new diagnostics also have the potential to worsen outcomes, particularly if clinicians miss-interpret negative tests as ruling out the condition as might well occur with Xpert MTB/RIF and TBM.

I completely agree. The delay in waiting for a diagnostic test result must also be recognised, which can be several days in many hospital settings. Bring on an accurate bedside POCT.

4. The abstract says that all 1672 patients were HIV infected but in the results section is says 96% were HIV infected so there is a discrepancy.

Thanks for pointing out this oversight. I have added the word predominantly to the abstract.

5. Last line, para 1 of introduction, probably remove word ‘to’.

Well spotted. Thank you.

6. Last line. Para 3 of introduction- It is not that Xpert cannot substitute for clinical judgement- Xpert is used to enhance clinical judgement, it’s just that the poor sensitivity means that when negative it has limited influence on decision making.

The text of the revised manuscript has been changes to read "a negative Xpert result has limited influence on clinical decision making".

7. Statistical methods- the uniform case definition is probable or definite TBM- do the authors think this is a reasonable reference standard for other tests? My view is that the threshold for treating TBM is very low and that patients with possible/probably or definite TBM should receive treatment and therefore this would be a more appropriate reference standard- what are your views on that?

This is a good discussion point. Where is the correct place to draw the line? The HIV co-infection makes this particularly challenging as CMV meningoencephalitis, toxoplasmosis, PML etc can muddy the water further. In Uganda, in our HIV/TBM cases 1/3rd have CSF WBC<5 which can mean the points scored in the CSF category are fewer. Furthermore in resource constrained settings access to neuroimaging and extra neural sampling can be limited which again affects the
ability to fully characterise the case. It would be completely acceptable to include ‘possible’ in the reference standard but we chose ‘probable and definite’ as we were concerned that the ‘possibles’ may include other HIV-related neuropathologies.

8. Methods- Not clear if there were 76 or 74 microbiologically proven cases

Thanks for pointing out this typo. The test has been corrected to 74.

9. Discussion- Para 1, repeat of ‘that of’

Well spotted again. Thank you.

10. Discussion- Para 4, First sentence implies that tests are the only answer to the diagnosis of TBM, what about clinical prediction rules relying on clinical data- although so far they have not been very successful, neither have tests so a robust CPR may negate the need for tests.

I agree there is seldom a perfect test and clinical prediction rules can be extremely useful in areas where access to tests is limited. I look forward to seeing a CPR that can accurately distinguish TBM from other HIV-related brain infections / pathology and commend researchers in this pursuit.

Discussion- Para5, Do the authors really think there are dead bacilli in CSF? Some of the authors have argued the opposite in their recent paper on Xpert Ultra so you can’t have it both ways.

The presence of MTB DNA in the CSF of a immunocompromised adult with clinical meningitis in a TB endemic setting almost certainly represents TBM. However it may not always be possible to culture organisms as they may have been rendered non-viable by recent TB treatment, inflammatory response or delays in processing.

Thanks again for your comments. We hope the revised manuscript is to your satisfaction.

\textit{Competing Interests:} nil