Xpert MTB/RIF Ultra for detection of *Mycobacterium tuberculosis* and rifampicin resistance: a prospective multicentre diagnostic accuracy study

Susan E Dorman*, Samuel G Schumacher*, David Alland, Pamela Nabeta, Derek T Armstrong, Bonnie King, Sandra L Hall, Soumitesh Chakravorty, Daniela M Cirillo, Nestani Tukvadze, Nino Bablishvili, Wendy Stevens, Lesley Scott, Camilla Rodrigues, Mubin I Kazi, Moses Joloba, Lydia Nakiyingi, Mark P Nicol, Yonas Ghebrekristos, Irene Anyango, Wilfred Munthi, Reynaldo Dietze, Renato Lyrio Peres, Alena Skrahina, Vera Auczyńska, Kamal Kishore Chopra, Mahmud Hanif, Xin Liu, Xing Yuan, Catharina C Boehme, Jerrold J Ellner, Claudia M Denkinger, on behalf of the study team†

**Summary**

**Background** The Xpert MTB/RIF assay is an automated molecular test that has improved the detection of tuberculosis and rifampicin resistance, but its sensitivity is inadequate in patients with paucibacillary disease or HIV. Xpert MTB/RIF Ultra (Xpert Ultra) was developed to overcome this limitation. We compared the diagnostic performance of Xpert Ultra with that of Xpert for detection of tuberculosis and rifampicin resistance.

**Methods** In this prospective, multicentre, diagnostic accuracy study, we recruited adults with pulmonary tuberculosis symptoms presenting at primary health-care centres and hospitals in eight countries (South Africa, Uganda, Kenya, India, China, Georgia, Belarus, and Brazil). Participants were allocated to the case detection group if no drugs had been taken for tuberculosis in the past 6 months or to the multidrug-resistance risk group if drugs for tuberculosis had been taken in the past 6 months, but drug resistance was suspected. Demographic information, medical history, chest imaging results, and HIV test results were recorded at enrolment, and each participant gave at least three sputum specimens on 2 separate days. Xpert and Xpert Ultra diagnostic performance in the same sputum specimen was compared with culture tests and drug susceptibility testing as reference standards. The primary objectives were to estimate and compare the sensitivity of Xpert Ultra test with that of Xpert for detection of smear-negative tuberculosis and rifampicin resistance and to estimate and compare Xpert Ultra and Xpert specificities for detection of rifampicin resistance. Study participants in the case detection group were included in all analyses, whereas participants in the multidrug-resistance risk group were only included in analyses of rifampicin-resistance detection.

**Findings** Between Feb 18, and Dec 24, 2016, we enrolled 2368 participants for sputum sampling. 248 participants were excluded from the analysis, and 1753 participants were distributed to the case detection group (n=1439) and the multidrug-resistance risk group (n=314). Sensitivities of Xpert Ultra and Xpert were 63% and 46%, respectively, for the 137 participants with smear-negative and culture-positive sputum (difference of 17%, 95% CI 10 to 24); 90% and 77%, respectively, for the 115 HIV-positive participants with culture-positive sputum (13%, 6-4 to 21); and 88% and 83%, respectively, across all 462 participants with culture-positive sputum (5-4%, 3-3 to 8-0). Specificities of Xpert Ultra and Xpert for case detection were 96% and 98% (–2·7%, –3·9 to –1·7) overall, and 93% and 98% for patients with a history of tuberculosis. Xpert Ultra and Xpert performed similarly in detecting rifampicin resistance.

**Interpretation** For tuberculosis case detection, sensitivity of Xpert Ultra was superior to that of Xpert in patients with paucibacillary disease and in patients with HIV. However, this increase in sensitivity came at the expense of a decrease in specificity.

**Funding** Government of Netherlands, Government of Australia, Bill & Melinda Gates Foundation, Government of the UK, and the National Institute of Allergy and Infectious Diseases.

**Copyright © The Author(s). Published by Elsevier Ltd. This is an Open Access article under the CC BY-NC-ND 4.0 license.**

**Introduction** An estimated 10·4 million new tuberculosis cases occurred in 2015, but only 5·1 million (59%) were diagnosed.1 That same year, an estimated 580 000 rifampicin-resistant cases occurred, but only 125 000 (20%) were identified.1 These diagnostic gaps are caused mostly by the lack of highly sensitive, rapid, accessible diagnostics.2 WHO recommended the Xpert MTB/RIF assay (Cepheid, Sunnyvale, CA, USA), an automated, integrated, cartridge-based molecular assay, as the initial test for tuberculosis to increase case detection and improve identification of rifampicin resistance directly from sputum.2,3 Xpert is used in tuberculosis programmes in more than 120 countries.4 However, Xpert’s sensitivity for tuberculosis detection is inadequate when few bacilli are present in a clinical specimen. This limits the usefulness of Xpert in patients with sputum smear-negative or extrapulmonary tuberculosis. This is particularly relevant for people with...
HIV and for children, in whom tuberculosis is often difficult to diagnose and morbidity can be high.13 One possible consequence of imperfect test sensitivity is lack of confidence in a negative test result, leading to empiric treatment and possibly overtreatment that might assure generalisability of the data to the tuberculosis epidemic worldwide. Our findings suggest that Xpert Ultra is substantially more sensitive than Xpert for detection of pulmonary tuberculosis, especially for paucibacillary specimens (ie, smear-negative specimens and specimens from HIV-positive individuals in whom available tests work best well). However, the increased sensitivity of Xpert Ultra came at the expense of a loss of specificity. For detection of rifampicin resistance, Xpert Ultra and Xpert performed comparably.

**Implications of all the available evidence**

The improved sensitivity of the Xpert Ultra assay relative to the Xpert assay should permit more evidence-based treatment decisions and at earlier stages of disease, even in people with HIV who can have high morbidity and mortality from tuberculosis despite relatively low bacillary burdens in sputum. On the basis of these findings, WHO has concluded that Xpert Ultra can be used as an alternative to Xpert for initial testing in adults with signs or symptoms of tuberculosis. Further research in different epidemiological settings and patient populations is needed to clarify the implications of the trade-off between increased sensitivity and decreased specificity and to determine the biological basis for Xpert Ultra-positive and culture-negative results.

**Methods**

**Study design and participants**

The primary objectives of this initial clinical diagnostic accuracy study were to estimate and compare the sensitivity of a single Xpert Ultra test with that of a single Xpert test of the same raw sputum specimen for detection of smear-negative tuberculosis and rifampicin resistance, and to estimate and compare Xpert Ultra and Xpert specificities for detection of rifampicin resistance. We hypothesised that the sensitivity of a single Xpert Ultra test for detection of smear-negative tuberculosis was non-inferior to that of a

**Research in context**

**Evidence before this study**

In 2010, WHO endorsed the Xpert MTB/RIF assay for initial diagnostic testing of individuals suspected of multidrug-resistant tuberculosis or HIV-associated tuberculosis. In 2014, WHO expanded this recommendation for use in all patients. The diagnostic accuracy of Xpert for pulmonary tuberculosis and rifampicin resistance has been assessed in Cochrane systematic reviews. The most recent update included studies described in any language until Feb 7, 2013. In 27 studies with nearly 10,000 participants, the pooled sensitivities of Xpert for pulmonary tuberculosis were 98% in those who were positive by sputum smear microscopy but only 67% in those who were negative by sputum smear microscopy. Pooled sensitivity was 79% in HIV-positive patients independent of sputum smear status, and pooled specificity was 99%. Performance characteristics for rifampicin resistance were 95% sensitivity and 98% specificity. The suboptimal detection of rifampicin resistance by Xpert in mixed populations containing rifampicin-resistant plus rifampicin-susceptible bacilli and some silent mutations and the consequent false determinations of rifampicin resistance have been confirmed in subsequent reports.

The Xpert MTB/RIF Ultra (Xpert Ultra) assay was developed to overcome the limited sensitivity of Xpert in the detection of pulmonary tuberculosis and limited accuracy of rifampicin resistance detection. We searched PubMed with the term “Xpert MTB/RIF Ultra” for articles in any language published until Oct 18, 2017. Other than two commentaries, we found two primary research articles describing the limit of detection and the performance of Xpert Ultra for detection of Mycobacterium tuberculosis in cerebrospinal fluid. Findings from analytical laboratory studies showed that Xpert Ultra had a lower limit of bacillary detection and was more accurate for detection of rifampicin resistance than Xpert. Xpert Ultra was also found to have higher sensitivity than Xpert and culture in paucibacillary specimens of cerebrospinal fluid.

**Added value of this study**

This is the first prospective study on the accuracy of Xpert Ultra for pulmonary tuberculosis. We did this study in eight countries with high burdens of tuberculosis or drug-resistant tuberculosis, and we applied a rigorous reference standard to assure generalisability of the data to the tuberculosis epidemic worldwide. Our findings suggest that Xpert Ultra is substantially more sensitive than Xpert for detection of pulmonary tuberculosis, especially for paucibacillary specimens (ie, smear-negative specimens and specimens from HIV-positive individuals in whom available tests work best well). However, the increased sensitivity of Xpert Ultra came at the expense of a loss of specificity. For detection of rifampicin resistance, Xpert Ultra and Xpert performed comparably.

**Added value of this study**

This is the first prospective study on the accuracy of Xpert Ultra for pulmonary tuberculosis. We did this study in eight countries with high burdens of tuberculosis or drug-resistant tuberculosis, and we applied a rigorous reference standard to assure generalisability of the data to the tuberculosis epidemic worldwide. Our findings suggest that Xpert Ultra is substantially more sensitive than Xpert for detection of pulmonary tuberculosis, especially for paucibacillary specimens (ie, smear-negative specimens and specimens from HIV-positive individuals in whom available tests work best well). However, the increased sensitivity of Xpert Ultra came at the expense of a loss of specificity. For detection of rifampicin resistance, Xpert Ultra and Xpert performed comparably.

**Added value of this study**

This is the first prospective study on the accuracy of Xpert Ultra for pulmonary tuberculosis. We did this study in eight countries with high burdens of tuberculosis or drug-resistant tuberculosis, and we applied a rigorous reference standard to assure generalisability of the data to the tuberculosis epidemic worldwide. Our findings suggest that Xpert Ultra is substantially more sensitive than Xpert for detection of pulmonary tuberculosis, especially for paucibacillary specimens (ie, smear-negative specimens and specimens from HIV-positive individuals in whom available tests work best well). However, the increased sensitivity of Xpert Ultra came at the expense of a loss of specificity. For detection of rifampicin resistance, Xpert Ultra and Xpert performed comparably.

**Added value of this study**

This is the first prospective study on the accuracy of Xpert Ultra for pulmonary tuberculosis. We did this study in eight countries with high burdens of tuberculosis or drug-resistant tuberculosis, and we applied a rigorous reference standard to assure generalisability of the data to the tuberculosis epidemic worldwide. Our findings suggest that Xpert Ultra is substantially more sensitive than Xpert for detection of pulmonary tuberculosis, especially for paucibacillary specimens (ie, smear-negative specimens and specimens from HIV-positive individuals in whom available tests work best well). However, the increased sensitivity of Xpert Ultra came at the expense of a loss of specificity. For detection of rifampicin resistance, Xpert Ultra and Xpert performed comparably.

**Added value of this study**

This is the first prospective study on the accuracy of Xpert Ultra for pulmonary tuberculosis. We did this study in eight countries with high burdens of tuberculosis or drug-resistant tuberculosis, and we applied a rigorous reference standard to assure generalisability of the data to the tuberculosis epidemic worldwide. Our findings suggest that Xpert Ultra is substantially more sensitive than Xpert for detection of pulmonary tuberculosis, especially for paucibacillary specimens (ie, smear-negative specimens and specimens from HIV-positive individuals in whom available tests work best well). However, the increased sensitivity of Xpert Ultra came at the expense of a loss of specificity. For detection of rifampicin resistance, Xpert Ultra and Xpert performed comparably.
single Xpert, and that the sensitivity and the specificity of Xpert Ultra for rifampicin resistance detection were non-inferior to those of Xpert. The study was done at ten reference laboratories in eight countries (South Africa, Uganda, Kenya, India, China, Georgia, Belarus, and Brazil). Eligible study participants were adults presenting at primary health-care centres and hospitals with pulmonary tuberculosis symptoms and who were willing to provide up to four sputum specimens at study enrolment. Participants were recruited prospectively into one of two groups: the case detection group or the multidrug-resistance risk group. Participation in the case detection group required willingness to attend study follow-up visits 42–70 days after enrolment and that no tuberculosis drugs had been taken in the past 6 months. Participants assigned to the multidrug-resistance risk group were at high risk of drug resistance on the basis of one or more of the following criteria: (1) microbiologically confirmed pulmonary tuberculosis with documented rifampicin resistance and tuberculosis treatment received for 31 days or less; (2) known pulmonary tuberculosis with suspected treatment failure; and (3) history of drug-resistant tuberculosis and off tuberculosis treatment for at least 3 months. The study protocol (appendix) was reviewed and approved by ethics committees at study sites and supervising organisations. Written informed consent was obtained from all study participants. Study participation did not affect the standard of care.

### Procedures

Demographic information, medical history, chest imaging results, and HIV test results (at sites where part of routine care) were recorded at enrolment. Participants were asked to provide minimum three sputum specimens on two separate days. Xpert and Xpert Ultra assays, smear microscopy, culture testing, and phenotypic drug susceptibility tests for rifampicin were done on site. Study-specified laboratory quality assurance included the use of external controls (positive and negative) and swab testing of the specimen processing area and of GeneXpert instrument surfaces.

Xpert and Xpert Ultra assays were done by adding sample reagent to the first collected sputum specimen in a 2:1 dilution, and 2.0 mL of the resulting mixture was added to one Xpert and one Xpert Ultra cartridge. Samples were analysed using standard four-module GeneXpert instruments with automated readouts for *M tuberculosis* detection (invalid [no internal assay control detected]; not detected; or detected [with semiquantitation]) and rifampicin resistance (detected, not detected, or indeterminate). The semiquantitative scale for Xpert Ultra results was trace, very low, low, medium, or high. The semiquantitative scale for Xpert results was very low, low, medium, or high.

For reference standard testing, the second and third sputum specimens were first digested with N-acetyl-L-cysteine and sodium hydroxide and concentrated using standard methods. Smear microscopy was done using Ziehl-Neelsen (Belarus site) or auramine-rhodamine staining (all other sites). 0.5 mL of the resuspended pellet was inoculated into liquid culture using mycobacteria growth indicator tube (MGIT) with a BACTEC 960 instrument (BD Microbiology Systems, Sparks, MD, USA), and 0.2 mL was inoculated on Löwenstein-Jensen medium. Cultures positive for growth of acid-fast bacilli underwent confirmation of *M tuberculosis* complex by MPT64/MPB64 antigen detection or line probe assays. Phenotypic drug susceptibility testing was done from the first positive *M tuberculosis* culture using the BACTEC MGIT 960 system and a rifampicin critical concentration of 1·0 µg/mL. Genetic drug susceptibility testing by Sanger DNA sequencing or pyrosequencing of the 81-bp *rpoB* core region was done for cultured isolates from all participants with discordant results between phenotypic drug susceptibility and Xpert Ultra readouts and for a subset of participants with concurrent results. Next-generation sequencing or pyrosequencing of IS6110, IS1081, and *rpoB* from the Xpert Ultra cartridge amplicon was done on specimens for which Xpert Ultra results were positive, but no culture was positive (appendix p 2).

Case definitions for the primary analyses were based on four culture results from sputum specimens two and three (figure 1). A culture-positive tuberculosis case was defined as a participant with at least one culture positive for *M tuberculosis*. Culture-positive cases were considered smear-positive if they had at least one positive smear (inclusive of scanty positive smears). A culture-negative participant had no culture positive for *M tuberculosis* and at least two cultures negative for *M tuberculosis*.

Staff doing Xpert or Xpert Ultra assays were blinded to results of other study tests through use of specimen codes and through staffing assignments. Data were captured through dedicated data-entry systems that were password-protected.

### Statistical analysis

Sample size was calculated by Monte-Carlo Simulation (appendix p 3). Sensitivity was defined as the proportion of patients testing positive with the reference standard who tested positive by the index test (Xpert Ultra) or comparator test (Xpert). Specificity was the proportion of patients testing negative with the reference standard who tested negative by the index test or comparator test. The primary analysis was based on results from initial testing of the first sputum specimen with Xpert and Xpert Ultra. Participants in the case detection group were included in all analyses, whereas participants in the multidrug-resistance risk group were only included in analyses of rifampicin-resistance detection. Patients were excluded from the analysis if culture contamination did not allow application of the case definition or if results of Xpert or Xpert Ultra were indeterminate or missing on initial testing. The proportion testing indeterminate is reported separately.

See Online for appendix
Results for simple proportions are presented with Clopper-Pearson 95% CI. The 95% CI around differences in proportions (for paired specimens in non-inferiority analyses) was computed using Tango’s score method.16,17 A non-inferiority endpoint, rather than a superiority endpoint, was selected for this initial clinical diagnostic accuracy study of Xpert Ultra because a superiority endpoint would have required a prohibitively large sample size of enrolled participants with smear-negative pulmonary tuberculosis. We reasoned that assurance from a diagnostic accuracy study that Xpert Ultra was at least as good as Xpert would be useful to clinicians and policy makers and provide rationale for a larger study to assess superiority. Superiority is demonstrated if it can be shown that sensitivity of Xpert Ultra is superior to Xpert beyond what could occur by chance alone. To

Figure 1: Specimen laboratory testing, participant flow, and exclusions from analysis eligibility

Eligible participants were asked to provide four sputum specimens (sputum 1–4) on 2 separate days. Xpert MTB/RIF Ultra assay (Xpert Ultra) on the first of sputum specimen was the index test, and Xpert MTB/RIF assay on the first sputum specimen was the comparator test. When possible, a fourth sputum specimen was obtained for additional solid and liquid cultures in cases with Xpert and Xpert Ultra discrepant results on sputum specimen 1. Sputum 4 results were only used for secondary analyses. NALC-NaOH=N-acetyl-L-cysteine and sodium hydroxide. MGIT=mycobacteria growth indicator tube. *Some reasons for exclusion overlap.
Of the 1439 participants in the case detection group, 694 (46%) met inclusion criteria and were included in the analyses. Between Feb 18, and Dec 24, 2016, we enrolled 1753 participants in the case detection and multidrug-resistance risk groups, 684 were culture-positive and had rifampicin resistance by Xpert or Xpert Ultra but provided interpretable rifampicin drug susceptibility test results. Xpert Ultra sensitivity and specificity are shown in table 2. The increase in sensitivity of Xpert Ultra relative to Xpert was larger than the remaining sensitivity gap between Xpert Ultra and a single liquid culture (appendix p 5). Xpert Ultra and Xpert sensitivities using phenotypic drug susceptibility testing (table 1). The comparison of sensitivity and specificity provided results for 580 participants (85%; appendix p 14). The comparison of sensitivity and specificity between Xpert and Xpert Ultra in the detection of rifampicin resistance is shown in table 2. Incorporating sequencing data for specimens that tested positive for rifampicin resistance by Xpert or Xpert Ultra but rifampicin-resistant by phenotypic drug susceptibility testing non-inferiority, the lower limit of the CI of the difference in sensitivity (Δ) was compared with the predefined non-inferiority margin; non-inferiority is achieved if the lower limit of the CI of Δ is no less than the non-inferiority margin. Non-inferiority margins for comparison between Xpert Ultra and Xpert were set at –7% for sensitivity to detect smear-negative tuberculosis, and at –3% for sensitivity and specificity to detect rifampicin resistance (appendix, p 3). A margin was not predefined for specificity of tuberculosis detection. We used Stata version 12 and R version 3.2.4 for statistical analyses. Data are median (IQR) or n/N (%). *For sites where HIV infection status was unknown for more than 50% of study participants, we show country-level HIV prevalence among tuberculosis cases. †HIV-infection status was unknown for 20 individuals; data are percentage of patients with known HIV status. ‡Numbers shown for study participants in the case detection group. §Data were missing for three patients at the Mumbai site. ¶At each site, study participants were enrolled in one of two possible (mutually exclusive) enrolment groups: the case detection group (based on suspicion of tuberculosis) or the multidrug-resistance risk group (based on suspicion of multidrug-resistant tuberculosis). **Calculated as percentage of the total number of culture-positive study participants in the case detection group and the multidrug-resistance risk group with available phenotypic drug-susceptibility test results.
testing gave specificity estimates of more than 99% for Xpert Ultra and Xpert, which were largely attributable to detection of mutations CTG533CCG, CAC526AAC, and CTG511CCG by Xpert Ultra and Xpert (appendix p 8) and only approached the specificity of those without a history of tuberculosis if the previous tuberculosis treatment was at least 7 years before enrolment.

19 (44%) of 43 participants with a positive Xpert Ultra test but no positive culture had an Xpert Ultra semiquantitative readout of trace. 15 (35%) participants with apparent false-positive Xpert Ultra results were also positive by Xpert (appendix p 9). Two (5%) participants with apparent false-positive Xpert Ultra results had *M tuberculosis* identified on a follow-up culture, and two (5%) participants were treated for tuberculosis on the basis of results from Xpert Ultra.

### Table 3: Test sensitivity and specificity depending on tuberculosis history and different approaches to the interpretation of semiquantitative trace-positive results for *Mycobacterium tuberculosis* detection by Xpert MTB/RIF Ultra (Xpert Ultra)

<table>
<thead>
<tr>
<th>Tuberculosis detection</th>
<th>Detection of rifampicin resistance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sensitivity: all culture-positive (95% CI; n/N)</td>
<td>Sensitivity (95% CI; n/N)</td>
</tr>
<tr>
<td>Sensitivity: smear-negative (95% CI; n/N)</td>
<td>Sensitivity: HIV-negative (95% CI; n/N)†</td>
</tr>
<tr>
<td>Sensitivity: HIV-positive (95% CI; n/N)‡</td>
<td>Specificity (95% CI; n/N)</td>
</tr>
<tr>
<td>Xpert</td>
<td>Xpert Ultra</td>
</tr>
<tr>
<td>83% (79–86; 383/462)</td>
<td>88% (85–91; 408/462)</td>
</tr>
<tr>
<td>46% (37–55; 63/137)</td>
<td>63% (54–71; 86/137)</td>
</tr>
<tr>
<td>90% (84–94; 143/159)</td>
<td>91% (86–95; 145/159)</td>
</tr>
<tr>
<td>77% (68–84; 88/115)</td>
<td>90% (83–95; 103/115)</td>
</tr>
<tr>
<td>98% (97–99; 960/977)</td>
<td>96% (94–97; 934/977)</td>
</tr>
<tr>
<td>95% (91–98; 167/175)</td>
<td>98% (96–99; 369/376)</td>
</tr>
</tbody>
</table>

Tables 2 and 3: Comparative accuracy for detection of tuberculosis and rifampicin resistance

<table>
<thead>
<tr>
<th>Sensitivity</th>
<th>Specificity</th>
</tr>
</thead>
<tbody>
<tr>
<td>All culture-positive (95% CI; n/N)</td>
<td>All culture-negative (95% CI; n/N)</td>
</tr>
<tr>
<td>Xpert</td>
<td>Xpert Ultra</td>
</tr>
<tr>
<td>83% (79–86; 383/462)</td>
<td>88% (85–91; 408/462)</td>
</tr>
<tr>
<td>46% (37–55; 63/137)</td>
<td>63% (54–71; 86/137)</td>
</tr>
<tr>
<td>98% (97–99; 960/977)</td>
<td>96% (94–97; 934/977)</td>
</tr>
<tr>
<td>98% (97–99; 715/727)</td>
<td>96% (95–98; 701/727)</td>
</tr>
<tr>
<td>98% (95–99; 244/249)</td>
<td>93% (89–96; 232/249)</td>
</tr>
</tbody>
</table>

Results were based on initial testing of the first sample with Xpert MTB/RIF and Xpert MTB/RIF Ultra (Xpert Ultra) assays. Uninterpretable results (contaminated cultures or non-determinate Xpert or Ultra results) were excluded from the analysis. Culture contamination averaged 4·3–7·8%, depending on sample and culture type. Non-determinate results (invalid, error, no result) are reported in the main text. Sensitivities of Xpert and Xpert Ultra for detection of smear-negative tuberculosis were 99% (95% CI 97–100) and 93% (97–100). *Accuracy for tuberculosis detection was estimated in study participants in the case detection group. Participants with unknown HIV-infection status are excluded from analyses stratified by HIV status but included in all other analyses. †Accuracy for detection of rifampicin resistance was estimated in all study participants with available drug susceptibility test results and valid rifampicin resistance results for both Xpert and Xpert Ultra. Data on HIV-infection status were not available for 188 culture-positive and 336 culture-negative study participants. Sensitivity of Xpert and Xpert Ultra in study participants with missing HIV status was 81% and 85%, respectively. Note that the estimate for pooled sensitivity of Xpert Ultra irrespective of HIV status does not fall between the estimates for HIV-infected and HIV-uninfected individuals. §Accuracy estimates are based on the reference standard as defined in the Methods section (using four cultures to define tuberculosis); using a less stringent reference standard with only one liquid and one solid culture (both from sputum sample 2), which is similar to the reference standard used in 21 of 22 studies included in the most recent Cochrane systematic review of the Xpert assay, resulted in Xpert sensitivity for smear-negative tuberculosis of 73% (Cochrane review pooled estimate 67%) and Xpert Ultra sensitivity of 86% (appendix p 5).
basis of clinical suspicion. Of the 24 (56%) participants who did not have culture-positive or Xpert-positive sputum and who had not started therapy, 18 participants gave 2-month follow-up information on symptoms. Symptoms had resolved in nine participants, improved in eight participants, and had not changed in one participant. Sequencing of the amplicons obtained from 14 cartridges (14 participants) with apparent false-positive results showed *M tuberculosis* DNA in 12 participants (appendix pp 10–12).

In a post-hoc analysis, we explored the effect of reclassifying Xpert Ultra trace-positive results as tuberculosis-negative on sensitivity and specificity for case detection (appendix p 7). Eliminating the trace-positive category and reclassifying all trace-positive results as tuberculosis-negative improved Xpert Ultra specificity but reduced its sensitivity (table 3). A conditional-trace approach (Xpert Ultra trace-positive results were reclassified as tuberculosis-negative only in participants with a history of tuberculosis) and a trace-repeat approach (participants with a trace-positive Xpert Ultra result for the first specimen were classified either as tuberculosis-negative if an Xpert Ultra test result of another sputum specimen was negative, or as tuberculosis-positive if an Xpert Ultra test result of another sputum specimen was positive) also improved Xpert Ultra specificity estimates (table 3). The conditional trace and trace-repeat approaches retained most of Xpert Ultra’s sensitivity in the smear-negative group. In a post-hoc analysis stratified by country-specific tuberculosis incidence, the specificity of Xpert Ultra was almost identical to that of Xpert in countries with low incidence (100 cases per 100 000 population or less), whereas the difference between Xpert Ultra and Xpert was greatest (–8% [95% CI –14 to –5] in favour of Xpert) in participants with a medical history of tuberculosis who were enrolled in countries with high tuberculosis incidence (more than 100 cases per 100 000 population; appendix p 13).

On initial testing of 2001 specimens, non-determinate readouts (invalid, error, no result) were obtained for 39 (2%) specimens with Xpert and for 79 (4%) specimens with Xpert Ultra. After excluding instrument-related errors, non-determinate readouts were obtained for 28 (1%) specimens with Xpert and for 64 (3%) specimens with Xpert Ultra. A single repeat test done on the same specimen that initially was non-determinate reduced the number of non-determinate results to four (<1%) specimens with Xpert and to ten (<1%) specimens for Xpert Ultra.

**Discussion**

Results of this multicentre diagnostic accuracy study show that the sensitivity of Xpert Ultra was superior to that of the standard Xpert for tuberculosis case detection in participants with sputum smear-negative pulmonary tuberculosis. Xpert Ultra also had superior sensitivity for tuberculosis case detection in HIV-infected participants and in all study participants. In clinical practice, the high sensitivity of Xpert Ultra could facilitate diagnosis of tuberculosis at earlier stages of disease and diagnosis of tuberculosis in patients with HIV and sputum smear-negative tuberculosis, a population with high mortality. Similarly, sensitivity gains could also be relevant for diagnosis of tuberculosis in children and for diagnosis of extrapulmonary forms of tuberculosis such as meningitis. These groups were assessed in separate studies.19,20

The increased sensitivity of Xpert Ultra came at the expense of a loss of specificity. For Xpert Ultra, we observed a difference in specificity between patients with and without a medical history of tuberculosis treatment. Xpert Ultra specificity increased with increasing time since completion of treatment since the preceding tuberculosis episode up to 7 years. Xpert specificity differed by tuberculosis treatment history only if the preceding treatment had been completed within the past 2 years. These results are in line with findings by Theron and colleagues21 that show that Xpert-positive, culture-negative results were more common in individuals with a history of tuberculosis. Extrinsic *M tuberculosis* from other specimens or the laboratory environment, or false-negative cultures from over-decontamination are possible explanations for a positive nucleic-acid amplification test result in participants with sputum cultures that are negative for *M tuberculosis*. However, in our study, over-decontamination is not sufficient to explain all of the specificity decrement for Xpert Ultra, and environmental contamination is an unlikely explanation because we implemented rigorous laboratory quality assurance and quality monitoring throughout the study. We speculate that in our study, most instances of Xpert Ultra-positive,
culture-negative results were caused by the presence of \( M \) tuberculosis DNA or intact \( M \) tuberculosis bacilli (either living or dead, originating from the participant’s lower respiratory system), or both in sputum. \( M \) tuberculosis mRNA has also been detected in sputum along with persisting PET thoracic lesion activity in some patients with tuberculosis 1 year after standard 6-month tuberculosis treatment. It remains to be seen whether apparent reductions in test specificity in patients with a history of tuberculosis will also be observed for other molecular tests for tuberculosis that aim to improve sensitivity through the detection of multicopy targets. Additional studies with longer follow-up that investigate the natural history of patients with Xpert Ultra-positive and culture-negative results are needed to understand the clinical relevance of these test results.

More than half of Xpert Ultra false-positive results in patients with a history of tuberculosis were trace-positive (the semiquantitative result corresponding to the lowest bacillary burden), so reclassification of these results as tuberculosis-negative could be considered for all patients, for patients with a tuberculosis history only, or on the basis of Xpert Ultra test results from another sputum specimen. These approaches mitigate some loss of Xpert Ultra specificity while maintaining some sensitivity gains over Xpert. The population-level effect of the sensitivity and specificity trade-off on patient-important outcomes would be expected to vary by setting. For Xpert Ultra, country-level tuberculosis incidence levels seem to affect test specificity. For example, in our study, Xpert Ultra specificity was 99% in participants without a history of tuberculosis treatment in study sites in countries where the tuberculosis incidence of 100 cases per 100,000 population or less, and 95% in patients without tuberculosis treatment history in countries where the tuberculosis incidence is more than 100 cases per 100,000 population. Modelling studies are underway and will allow more in-depth exploration of the trade-offs between increased numbers of patients correctly and falsely diagnosed under different epidemiologic scenarios.

For detection of rifampicin resistance, Xpert Ultra specificity was non-inferior to that of Xpert. The sensitivity point estimate for Xpert Ultra was slightly less than that of Xpert and the confidence interval was wide, such that non-inferiority criteria were not met. Additional studies including larger numbers of rifampicin-resistant specimens are needed to more precisely characterise Xpert Ultra accuracy for detection of rifampicin resistance. Patients belonging to the multidrug-resistance risk group were recruited mainly from four sites and, accordingly, most rifampicin-resistant cases come from these sites. Mutations such as Ile491Phe, which is not detected by Xpert or Xpert Ultra, might be more common in countries not included in this study, and inclusion of such sites could potentially have reduced sensitivity estimates. However, we found no evidence of bias given that our reported accuracy for rifampicin-resistance detection is equivalent to that reported in a Cochrane review of a broad group of studies and sites worldwide. Our estimates are also in line with WHO surveillance data on the frequency of rifampicin-resistance-conferring mutations obtained from resistance surveys.

In summary, Xpert Ultra holds promise as a rapid and highly sensitive test for tuberculosis case detection and simultaneous detection of rifampicin resistance. Its sensitivity gain compared with Xpert is most apparent in individuals with low sputum bacillary burdens. Implementation approaches will need to consider the effect of possible false-positive Xpert Ultra results.

**Contributors**

CMD, SED, DA, SGS, SC, CCB, JJE, and PN designed the study, and PN, DTA, BK, SGS, SLH, CMD, SED, and DA oversaw the trial conduct. NT, NB, WS, LS, CR, MK, MJ, LN, MPN, YG, IA, WM, RD, RLP, AS, VA, KKC, MH, XL, and XY coordinated individual trial sites. SGS, SED, CMD, DA, and DMC analysed data and developed the first manuscript draft. All authors contributed to data collection, interpretation of data, and revision of the Article and approved the final version of the Article before submission.

**Study team members**

Yukari C Maruabe (Johns Hopkins University School of Medicine, Baltimore, MD, USA); David Horn (Boston Medical Center and Boston University School of Medicine, Boston, MA, USA); Rusudan Aspindznelashvili (National Center for Tuberculosis and Lung Diseases, Tbilisi, Georgia); Anura David (National Health Laboratory Service, Johannesburg, South Africa); Utkarsha Surve (PD Hinduja Hospital and Medical Research Centre, Mumbai, India); Louis H Kamulejega, Sheila Nabwawambo (Infectious Disease Institute and Makerere University, Kampala, Uganda); Shireen Surtie, Nchimunya Hapeela (Division of Medical Microbiology and Institute for Infectious Diseases and Molecular Medicine, Cape Town, South Africa); Kevin P Cain, Janet Agaya, Kimberly D McCarthy (US Centers for Disease Control and Prevention, and Division of Global HIV and Tuberculosis, Kisumu, Kenya); Patricia Marques Rodrigues, Luiz Guillerme Schmitz Castellani, Pedro Sousa de Almeida Jr, Paula Poloni Lobo de Aguiar (Universidade Federal do Espírito Santo, Vitória, Brazil); Varvara Solodovnikova (National Reference Laboratory, Republican Scientific and Practical Centre for Pulmonology and Tuberculosis, Minsk, Belarus); Xianglin Ruan, Lili Liang, Guolong Zhang, Hong Zhu, Yingda Xie (Henan Provincial Chest Hospital, Zhengzhou, Henan Province, China).

**Declaration of interests**

SGS, PN, CMD, and CCB are employed by FIND. FIND is a not-for-profit foundation that supports the evaluation of publicly prioritised tuberculosis assays and the implementation of WHO-approved (guidance and prequalification) assays using donor grants. FIND has product evaluation agreements with several private sector companies that design diagnostics and related products for treatment of tuberculosis and other diseases. These agreements strictly define FIND’s independence and neutrality vis-a-vis the companies whose products get evaluated and describe roles and responsibilities. DA reports grants from Johns Hopkins University School of Medicine and Cepheid during the conduct of the study; grants from National Institutes of Health (NIH), the Foundation for Innovative New Diagnostics, and the Henry Jackson Foundation outside the submitted work; and US and European patents within the Rutgers University Molecular Beacon Patent Pool for non-competitive co-amplification methods, assays for short sequence variants, wavelength-shifting probes and primers, nucleic acid detection probes having non-fret fluorescence quenching and kits and assays including such probes, homogeneous assays multiplex screening assays and kits, and PCR primer sets and probes for \( M \) tuberculosis. WS is a board member of the Contract Project (for mines), the Task Team for Correctional Services, and the Antimicrobial Drug Resistance Group, the World Bank/Presidential Project (for mines), the Task Team for Correctional Services, and the
National HIV and Tuberculosis Drug Resistance Working Group. WS
declares consultancy paid to institution from the Bill & Melinda Gates
Foundation Grand Challenges Canada, consultancy from WHO (ID, CD4, 
drug resistance, POC. Xpert TB), consultancy paid to institution from 
Clintond Foundation (Point-of-Care), personal fees from National 
Health Laboratory Service, joint staff with University of the Witwatersrand, 
expert testimony for grant funders (NIH, US Centers for Disease Control 
and Prevention (CDC), Global Fund, grants from NIH, CDC, the Global Fund, 
Clintond Foundation, Bill & Melinda Gates Foundation, PATH, 
Grand Challenges Canada, London School of Hygiene & Tropical Medicine, 
South African Medical Research Council, and the UK Medical Research 
Council, expert training, teaching development, and speaker fees paid to 
institution from Cepheid, outside the submitted work. LS reports having 
a patent US 8709712 issued. SC now works for Cepheid and has a patent 
pending for some of the primers and probes used in the assay kit. All other 
authors declare no competing interests. The findings and conclusions in 
this report are those of the authors and do not necessarily represent the 
official position of the CDC.

Acknowledgments
We thank study participants, without whom this work would not have 
been possible, and the clinical and laboratory teams at the participating 
trial sites. We thank Kate Shearer for review of the analysis plan and 
statistical code—since FIND was involved in the development of the 
Xpert and Xpert Ultra assays, key sections of statistical code were 
completely re-written and re-run by Shearer, as an independent 
statistician, to ensure an independent analysis. We thank Martin Jones, 
Marie Simmons, Bob Kwiatkowski, Pam Johnson, and David Persing 
from Cepheid for their support during development assay and with 
logistics for the study. We thank Nora Champouillon from FIND for 
facilitating the project. This work was supported by grants from the 
Government of Netherlands (DSO/GA-521/10) for assay development; 
and grants from the Government of Australia (70975), the 
Bill & Melinda Gates Foundation (OPP1105925), the UK aid from the 
UK Government (204074-101); and by the National Institute of Allergy 
and Infectious Diseases, NIH, Department of Health and Human 
Services, USA (contract #HHSN272200090005C and K24AI104830 
(SED)) for clinical evaluation. Wendy Stevens and Lesley Scott were 
supported by funding received from the South African Medical 
Research Council with funds received from the South African National 
Department of Health, and the UK Medical Research Council, with 
funds received from the UK Government’s Newton Fund under the UK/ 
South Africa Newton Fund #015NEWTON TB. Investigational 
cartridges, Xpert MTB/RIF cartridges, and GeneXpert instruments were 
generously donated by Cepheid. Cepheid personnel had no role in study 
from Cepheid, outside the submitted work. LS reports having a 
patent pending for some of the primers and probes used in the assay kit. All other 
authors declare no competing interests. The findings and conclusions in 
this report are those of the authors and do not necessarily represent the 
official position of the CDC.

References
Organization, 2016
2. Uplekar M, Weil D, Lonroth K et al. WHO’s new end TB strategy. 
4. Steingart KR, Schiller I, Horne D, Pai MP, Boehme CC, 
Dendukuri N. Xpert MTB/RIF assay for pulmonary tuberculosis and 
rifampicin resistance in adults. Cochrane Database Syst Rev 
2014; 1: CD009593.
5. WHO. Xpert MTB/RIF assay for the diagnosis of pulmonary and 
extrapulmonary TB in adults and children: policy update. Geneva: 
Organization, 2014.
MTB/RIF test for the diagnosis of pulmonary tuberculosis in 
children admitted to hospital in Cape Town, South Africa: a 
MTB/RIF assay for the diagnosis of pulmonary tuberculosis in a 
high HIV prevalence setting. Am J Respir Crit Care Med 2011; 
tuberculosis incidence, high-resource setting: limitations in 
Do high rates of empirical treatment undermine the potential effect 
of new diagnostic tests for tuberculosis in high-burden settings? 
clinical effect of point-of-care Xpert MTB/RIF testing for 
tuberculosis in primary-care settings in Africa: a multicentre, 
False-positive rifampicin resistance on Xpert MTB/RIF caused by 
18: 1255–57.
resistant results with Xpert MTB/RIF version 4 assay in clinical 
samples with a low bacterial load. Diagn Microbiol Infect Dis 2016; 
MTB/RIF Ultra: improving detection of Mycobacterium tuberculosis 
and resistance to rifampicin in an assay suitable for point-of-care 
testing. MBio 2017; 8: e00812–17.
16. Rothmann MD, Wiens BL, Chan ISF. Design and analysis of 
17. FDA. Guidance for industry. Non-inferiority clinical trials. Silver 
Spring, MD: Food and Drug Administration, 2010.
18. Cleveland WS. Robust locally weighted regression and smoothing 
19. Zar HJ, Workman L, Nicol MP. Diagnosis of pulmonary tuberculosis in 
HIV-infected and uninfected children using Xpert MTB/RIF Ultra. 
International Conference of the American Thoracic Society. 
MTB/RIF Ultra for tuberculous meningitis in HIV-infected adults: 
a prospective cohort study. Lancet Infect Dis 2017; published online 
patients with previous tuberculosis: can we distinguish true from 
emission tomography lesion activity and Mycobacterium tuberculosis 
for detection of Mycobacterium tuberculosis complex DNA in clinical 
specimens in a low-incidence country. BMC Infect Dis 2014; 14: 59.
24. Hofmann-Thiel S, Molodtsov N, Antonenko U, Hoffmann H. 
Evaluation of the Abbott RealTime MTB and RealTime MTB 
INH/RIF assays for direct detection of Mycobacterium tuberculosis 
complex and resistance markers in respiratory and extrapulmonary 