WHO Meeting Report of a Technical Expert Consultation: Non-inferiority analysis of Xpert MTB/RIF Ultra compared to Xpert MTB/RIF

World Health Organization

2017

WHO/HTM/TB/2017.04

Background

The development of the Xpert® MTB/RIF assay (Cepheid, Sunnyvale, USA) was a major step forward for improving the diagnosis of tuberculosis (TB) and rifampicin resistance detection globally. However, Xpert MTB/RIF sensitivity is imperfect, particularly in smear-negative and HIV-associated TB and some limitations also remain in its determination of rifampicin resistance. The Xpert® MTB/RIF Ultra assay (Ultra) has been developed by Cepheid as the next-generation assay to overcome these limitations, and uses the same GeneXpert® platform as Xpert MTB/RIF.

To improve assay sensitivity for the detection of *M. tuberculosis*, the Ultra assay incorporates two different multi-copy amplification targets (IS*6110* and IS*1081*) and a larger DNA reaction chamber than Xpert MTB/RIF (50μl PCR reaction in Ultra versus 25 μl in Xpert MTB/RIF). Ultra also incorporates fully nested nucleic acid amplification, more rapid thermal cycling, and improved fluidics and enzymes. This has resulted in Ultra having a limit of detection (LOD) of 16 bacterial colony forming units (cfu) per ml (compared to 114 cfu per ml for Xpert MTB/RIF). To improve the accuracy of rifampicin resistance detection, the Ultra incorporates melting temperature-based analysis instead of real-time PCR. Specifically, four probes identify rifampicin resistance mutations in the rifampicin resistance determining region of the *rpoB* gene by shifting the melting temperature away from the wild type reference value.

Mycobacterium tuberculosis (MTB) complex detection with Ultra is defined as: one or both of the probes that detect the multi-copy targets are positive with cycle thresholds (Cts) less than 37 cycles and at least two *rpo*B probes with Cts less than 40 cycles. Ultra uses the same semi-quantitative categories used in the Xpert MTB/RIF assay (high, medium, low and very low) as well as the addition of a new semi-quantitative category "trace" that corresponds to the lowest bacillary burden for MTB detection. The "MTB Detected, trace", henceforth referred to as "trace call" was added to improve sensitivity for MTB detection. For a "trace call", one or both of the probes for the multi-copy targets are positive with Cts less than 37 cycles and no more than one *rpo*B probes have a Ct less than 40 cycles. MTB is reported as NOT detected if neither of the multi-copy target probes are positive and the sample processing control is positive with a Ct less than 35 cycles.

For rifampicin resistance detection, rifampicin resistance is considered absent if MTB is detected (not trace) and all four *rpo*B probes have identifiable melt temperature (Tm) peaks in the wild type profile. Rifampicin resistance is detected if MTB is detected (not trace) and all four *rpo*B probes have identifiable Tms and at least one of the *rpo*B probes has a Tm mutant profile. If MTB is detected with a "trace call", then no interpretation can be made regarding rifampicin resistance and results are reported as MTB detected, trace, RIF indeterminate.

Technical Expert Consultation

The Global TB Programme of WHO convened a Technical Expert Consultation via webinar on 20 January 2017 to assess a multi-centre non-inferiority diagnostic accuracy study conducted by FIND of the Ultra assay compared with the Xpert MTB/RIF assay for the diagnosis of TB and the detection of rifampicin resistance. Details of the Technical Expert Group (TEG) membership and declarations of interest are given in Annex 1. The TEG evaluated the findings from the main study which was a prospective multi-centre diagnostic accuracy study in adults with signs and symptoms of pulmonary TB (Full report available at: https://www.finddx.org/publication/ultra-report/).

Xpert MTB/RIF and Ultra were performed from the same specimen and accuracy was determined with four cultures as the reference standard for TB detection (two MGIT tubes + two LJ slants, performed on two specimens obtained on separate days). Phenotypic drug-susceptibility testing as well as sequencing was performed for rifampicin

resistance detection. In parallel to the main study, data from several retrospective studies were available to assess the performance of Ultra in extrapulmonary (EPTB) and paediatric samples as well as in low TB burden settings. Modelling was used to assess the trade-offs of the assay based on the performance seen in the main study (Annex 2).

Findings

In the main study, 1,520 persons with signs and symptoms of TB were enrolled. Overall, sensitivity of the Ultra was 5% higher than that of Xpert MTB/RIF (95%CI +2.7, +7.8) but specificity was 3.2% lower (95%CI -2.1, -4.7). Sensitivity increases were highest among smear-negative culture-positive patients (+17%, 95%CI +10, +25) and among HIV-infected patients (+12%, 95%CI +4.9, +21). Specificity-decreases were higher in patients with a history of TB (-5.4%, 95%CI -9.1, -3.1) than in patients with no history of TB (-2.4%, 95%CI -4.0, -1.3).

Reclassifying "trace calls" as MTB-negative either in all cases, or in those with a history of TB, mitigated most of the specificity losses (specificity –1.0% and -1.9% if "trace calls" were reclassified for all cases or for cases with TB history, respectively) while maintaining some of the sensitivity gains over Xpert MTB/RIF (sensitivity +7.6% and +15% respectively). The important observation was that repeat testing of a fresh specimen from persons whose initial specimen had a "trace call" result would mitigate some of the loss in specificity similar to "No trace", but retain most of the gain in sensitivity. Ultra performed similarly well as Xpert MTB/RIF in detection of rifampicin resistance and the specificity of both assays was close to 100% when sequencing data was used to resolve discordant results with phenotypic DST. The use of melting temperature-based analysis with Ultra instead of real-time PCR analysis with Xpert MTB/RIF allows Ultra to better differentiate silent mutations (such as Q513Q or F514F) from resistance conferring mutations. The number of patients with rifampicin resistance enrolled in the study was, however, insufficient to confirm the initial

analytical results that suggested a superior performance of Ultra for rifampicin resistance detection.

Additional retrospective studies demonstrated that in low TB burden settings where there is very limited TB transmission the specificity of Ultra is very high (99.3%, 95%CI 96-99). For EPTB and paediatric TB, studies highlighted the benefit of the increased sensitivity (primarily due to the 'trace call') with a sensitivity of 95% for Ultra versus 45% for Xpert MTB/RIF for detection of TB meningitis using cerebrospinal fluid and 71% for Ultra on respiratory samples from children versus 47% for Xpert MTB/RIF.

Modelling demonstrated that Ultra is expected to improve pulmonary TB case detection and outcomes. Depending on the patient population, Ultra could detect 2 to 9 additional TB cases per 1000 individuals evaluated for presumptive TB, and prevent one additional TB death per 700 to 30,000 individuals evaluated. However, the increase in case detection comes at a cost: one false TB diagnosis and unnecessary treatment per 40 to 80 individuals evaluated and 10 to 500 unnecessary treatments per TB death prevented. The acceptable level of unnecessary treatments per prevented death (or per additional or earlier diagnosis) is likely to vary between settings. A similar trade-off exists regardless of whether the trace call is used.

Conclusions of the Technical Expert Consultation

The TEG agreed that the Ultra is non-inferior to the Xpert MTB/RIF assay for the detection of MTB and for the detection of rifampicin resistance.

The TEG recognized that Ultra has a higher sensitivity than Xpert MTB/RIF particularly in smear-negative culture-positive specimens and in specimens from HIV-infected patients with at least as good accuracy for rifampicin resistance detection. However, as a result of the increased sensitivity, Ultra also detects

non-replicating or non-viable bacilli present particularly in patients with recent history of TB, reducing the overall specificity of Ultra in high-burden TB settings. In low TB burden settings and in the testing of specimens for the diagnosis of extrapulmonary TB and paediatric TB, false positive results were not a major concern. In some isolated cases Ultra's sensitivity may also be superior to the current reference standard (liquid culture) depending on specimen quality and methods used for specimen handling and processing.

The TEG also recognized that much of the increase in sensitivity for MTB detection with the Ultra assay was attributed to the "trace calls" but these represented less than 1% of all results in the study. The group agreed that the greatest benefit was in the increased yield for the detection of MTB in smear-negative culture positive specimens, paediatric specimens, extra-pulmonary specimens (notably cerebrospinal fluid) and especially for HIV positive individuals whose specimens are frequently paucibacillary. The group recognized that the impact of increased sensitivity results in decreased specificity for TB detection (as with any other test) and becomes a trade-off between increased diagnosis and overtreatment. Among persons with HIV, given the widespread use of empiric treatment, the TEG agreed that "trace calls" should be considered as true positives in the clinical work-up of patients.

The TEG noted that amongst all persons with signs and symptoms of TB, excluding the use of the "trace call" resulted in higher specificity but lower sensitivity. The sensitivity increase among smear-negative culture-positive specimens using the Ultra with the "trace call" was 17% compared with the Xpert MTB/RIF, and this increase was reduced to around 8% when not utilising the "trace call". At the same time while the specificity was markedly improved when the trace call was excluded, it was still lower than Xpert MTB/RIF. The persistent reduced specificity was considered likely to be due to imperfect confirmation of TB history and possibly self-cured TB (i.e. incipient TB that resolved without treatment).

The TEG also noted that in some patient populations the exclusion of prior TB treatment may not always be reliable. In some instances, patients may hide their prior TB status due to fear of stigma and discrimination or have concerns regarding legal status for migrants, or the prior history is not adequately ascertained by the health care workers.

The TEG concluded that a "trace call" positive result was sufficient to initiate therapy in those with known or suspected HIV infection, children and for extrapulmonary samples from persons. The TEG noted that performing a repeat Ultra test on a fresh sputum specimen from adults with signs and symptoms of TB and not at risk for HIV infection that initially tested as "trace call" positive would contribute to increasing the specificity of Ultra without losing much of the benefits of increased sensitivity. The TEG agreed that the result of the second Ultra test could be used for clinical decisions and patient follow-up. While clinical and available radiological information should always be considered in the diagnosis of tuberculosis, a second "trace call" positive was felt sufficient to make a diagnosis of pulmonary TB. The TEG concluded that should a second test be negative, further clinical and radiological assessment should be made before initiating treatment for tuberculosis.

No information regarding rifampicin resistance is available from Ultra results that are reported as "MTB trace detected" as the rifampicin result is always reported as indeterminate for "trace calls". The TEG agreed that culture and drug susceptibility testing for rifampicin resistance be performed to confirm or exclude resistance while acknowledging the limitations of performing phenotypic DST for the detection of rifampicin resistance.

The TEG observed that the trade-off between potential overtreatment and increased diagnosis of TB and decreased mortality associated with TB treatment varies substantially between different settings with

variable populations determined by HIV, prior TB history, and prevalence.

Ultra implementation considerations

The Ultra assay is non-inferior to the current Xpert MTB/RIF assay for the diagnosis of MTB and the detection of rifampicin resistance and can be used as an alternative to the latter in all settings. Cepheid plans to gradually phase out and replace the current Xpert MTB/RIF assay with the Ultra assay. The current WHO recommendations for the use of Xpert MTB/RIF[1] also apply to the use of Ultra as the initial diagnostic test for all adults and children with signs and symptoms of TB and in the testing of selected extrapulmonary specimens (CSF, lymph nodes and tissue specimens).   The following implementation considerations apply to Ultra:

- The interpretation of Ultra results for MTB detection are the same as for Xpert MTB/RIF with the exception of "trace calls".
- Ultra has high sensitivity for MTB detection and incorporates a new semi-quantitative category "trace call" that corresponds to the lowest bacillary burden for MTB detection.   Interpret "trace calls" as follows:
  - Among persons with HIV, children and extrapulmonary specimens "trace calls" should be considered to be true positive results for use in clinical decisions and patient follow-up;
  - Among persons not at risk for HIV, with an initial "trace call" positive result, a fresh specimen from the patient should undergo repeat testing and the result of the second Ultra test be used for clinical decisions and patient follow-up;

---

[1] Automated real-time nucleic acid amplification technology for rapid and simultaneous detection of tuberculosis and rifampicin resistance: Xpert MTB/RIF assay for the diagnosis of pulmonary and extrapulmonary TB in adults and children. Policy update. WHO/HTM/TB/2013.16 Geneva, World Health Organization, 2013

- While clinical and available radiological information should always be considered in the diagnosis of tuberculosis, a second "trace call" positive is sufficient to make a diagnosis of pulmonary TB unless there is a recent history of TB;
- Among all persons that test "trace call" positive additional investigations are needed to confirm or exclude resistance to rifampicin.
- Ultra has both high sensitivity and specificity for rifampicin resistance detection.
  - The use of melting temperature-based analysis with Ultra instead of real-time PCR analysis with Xpert MTB/RIF allows Ultra to better differentiate silent mutations from resistance conferring mutations;
  - In persons at low risk for rifampicin resistance (e.g. new TB cases not at risk for MDR-TB) a positive rifampicin resistance result should be repeated using a fresh specimen. Repeat testing is recommended to control for any pre-analytical and/or post-analytical errors.
- All persons with rifampicin resistance, identified by Ultra should undergo further testing as per current WHO policy guidance to determine if there is additional resistance to the class of fluoroquinolones and/or the group of second-line injectable drugs.
- Ultra can be used on all GeneXpert instrument platforms and is suitable for use at central or national reference laboratory level, regional and district levels. GeneXpert has the potential to be used at the peripheral level, provided uninterrupted electricity supply and temperature conditions can be ensured.

Research needs

WHO plans to update policy recommendations for the use of Ultra in 2018.

Further operational research on Ultra should focus on addressing the following research priorities in different epidemiological and geographical settings and patient populations:

- Conduct additional studies assessing the value of performing a second Ultra assay when the initial test was MTB not detected, as a way to improve the sensitivity of the testing algorithm;
- Evaluation of the role of repeat testing of a fresh specimen from a patient with signs and symptoms of TB whose initial specimen was "trace call" positive, as a way to improve the specificity of the testing algorithm;
- Perform studies using Ultra for the detection of TB in children and extrapulmonary specimens (including faecal specimens from children);
- Conduct additional studies to assess the use of the "trace call" result when Ultra is used for systematic screening, active case finding and for prevalence surveys;
- Assess the performance of Ultra as a reference standard method for rifampicin resistance detection;
- Gathering more evidence on the impact of Ultra on important patient outcomes including time to TB treatment initiation, morbidity and mortality.

Acknowledgements

Annex 1. Members of the Technical Expert Group

Technical Expert Group Members
Martina Casenghi (Médecins San Frontières, Rome, Italy); Chris Coulter (Queensland Mycobacterium Reference Laboratory, Brisbane, Australia); Moses Joloba (National Reference Laboratory of the National TB and Leprosy Programme, Uganda); Chakaya J Muhwa (Kenyatta University School of Medicine, Nairobi, Kenya); VP Myneedu (National Institute of TB and Respiratory Diseases, New Delhi, India); Thomas M Shinnick (Independent consultant, United States of America); Wendy Stevens (National Priority Programmes, National Health Laboratory service, South Africa); Elisa Tagliani (San Raffaele Scientific Institute Milan, Italy); Sabira Tahseen (National TB Control Programme, Islamabad, Pakistan); Maria Alice Telles (Tuberculosis Laboratory Independent Consultant for PAHO, Brazil); Armand van Deun (The Union, Antwerp, Belgium).

Declaration and management of conflict of interest
All the contributors completed a WHO Declaration of Interest form. All stated declarations of interest were evaluated by the Global TB Programme for the existence of any possible financial conflict of interest which might warrant exclusion from membership of the Expert Group. In addition, the diversity and representation in the Expert Group was large enough to balance and overcome any potential intellectual conflict of interest.

The following interests were declared:

None declared
Jeremiah Chakaya, Moses Joloba, VP Myneedu, Sabira Tahseen, Maria Alice Telles and Armand van Deun.

Declared, insignificant
Martina Casenghi declared that as part of the position held at MSF Access Campaign, has presented MSF experience with Xpert MTB/RIF roll-out, highlighted the need for improved tests during conferences and public fora.

Chris Coulter has declared research funding from the Australian National Health and Medical Research Council for a TB WGS study, USD 10,000 per annum, provided for the project from this source, funding ceased in 2016.

Elisa Tagliani has declared research support: 250 Ultra and G4 tests for migrant screening (Cepheid/No income). Re-testing of discordant Xpert Ultra cases (FIND/ EUR 15,000).

Thomas Shinnick declared that he was a former employee of the United States Centres for Disease Control and Prevention (CDC) until January 2016. As an employee, he had often represented CDC's positions on laboratory services needed for tuberculosis diagnosis, treatment and control

Wendy Stevens declared that she had received funding for a number of TB assay validations in the form of reagents from different diagnostic companies (Cepheid, Abbott, Roche, Hain Lifesciences, DNA Genotek and Alere). Family member is employed by Cepheid in South Africa. Was involved in development of EQA/verification product for molecular TB testing. Has received multiple institutional grants for scientific projects related to TB/MDR-TB diagnostic.

Technical Resource Persons
Claudia Denkinger (FIND, Geneva, Switzerland); David Dowdy (John Hopkins); Emily Kendell (John Hopkins); Samuel Schumacher (FIND, Geneva, Switzerland).

WHO Steering Committee
Christopher Gilpin, Alexei Korobitsyn, Karin Weyer, Fuad Mirzayev, Wayne van Gemert

Annex 2: Abstract: The Expected Clinical Impact of Xpert Ultra: A Modeling Analysis

Authors: Emily A. Kendall and David W. Dowdy, Johns Hopkins University

Objective: The Xpert MTB/RIF Ultra cartridge has shown improved sensitivity for TB detection in recent field trials, but at the expense of diminished specificity. We explored the likely clinical implications, in different populations, of replacing the existing Xpert cartridge with the Xpert Ultra cartridge for diagnosis of pulmonary TB.
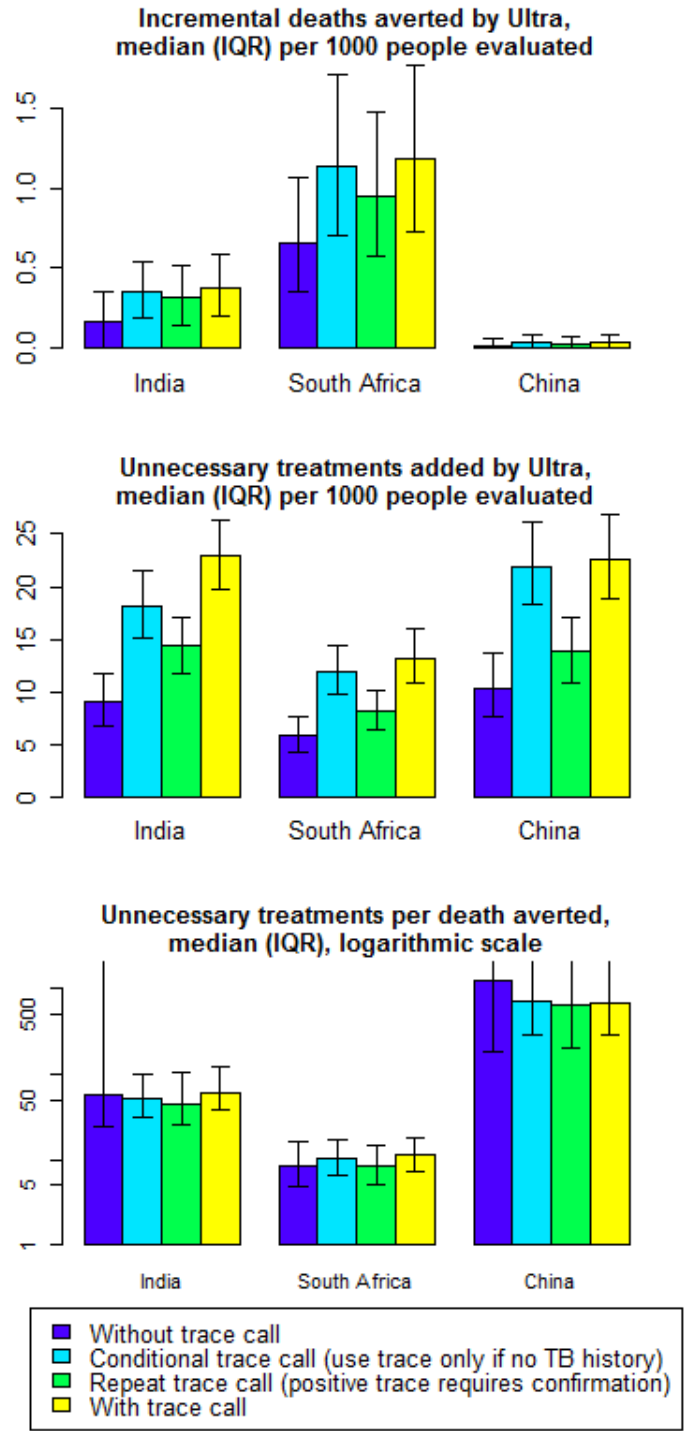
Methods: We developed a Markov microsimulation model of hypothetical cohorts of 100,000 adults undergoing diagnostic sputum evaluation with Xpert MTB/RIF for suspected pulmonary TB, in each of three emblematic settings: an HIV clinic in South Africa, a public TB center in India, and a primary care setting in China. In each setting, we used existing data to project likely diagnostic results, treatment decisions, and ultimate clinical outcomes, assuming use of the standard Xpert versus Xpert Ultra cartridge. Our primary outcome was the projected number of unnecessary treatments generated per TB death averted, if standard Xpert were switched to Xpert Ultra. We also simulated alternative approaches to interpreting positive results of the Ultra cartridge's semi-quantitative trace call.

Results: In the Indian setting, replacing the standard Xpert cartridge with Xpert Ultra was projected to avert 0.5 TB deaths (95% uncertainty range [UR]: -0.1, 1.3) and generate 22 unnecessary treatments (95% UR: 12, 35) per 1000 individuals evaluated – a ratio of 49 incremental unnecessary treatments per incremental death averted (95% UR: 14, indefinite upper bound). In the South African HIV-care setting – where TB mortality rates are higher and Ultra's improved sensitivity has greater absolute benefit – this ratio improved to 10 unnecessary treatments per TB death averted (95% UR: 3, 64). By contrast, in a Chinese primary care setting, this ratio was much less favorable, at 501 unnecessary treatments per TB death averted (95% UR: 93, indefinite upper bound). Alternative interpretations of the trace call had little effect on this ratio (see Figure). Limitations of this analysis include uncertainty in key parameters (including the clinical implications of false-negative results), the exclusion of transmission effects, and restriction of this analysis to adult pulmonary TB.

Conclusion: Modelling of the impact of switching from standard Xpert to the Xpert Ultra cartridge for diagnosis of adult pulmonary TB suggests that the consequences will differ depending on the underlying TB and MDR-TB epidemiology and the clinical setting. Using the Ultra cartridge may therefore require a more nuanced, setting-specific approach to implementation.

Incremental deaths averted by Ultra, median (IQR) per 1000 people evaluated

Unnecessary treatments added by Ultra, median (IQR) per 1000 people evaluated

Unnecessary treatments per death averted, median (IQR), logarithmic scale

Without trace call
Conditional trace call (use trace only if no TB history)
Repeat trace call (positive trace requires confirmation)
With trace call