WHO consolidated guidelines on tuberculosis

Module 3: Diagnosis

Rapid diagnostics for tuberculosis detection

World Health Organization

WHO consolidated guidelines on tuberculosis. Module 3: diagnosis – rapid diagnostics for tuberculosis detection

ISBN 978–92–4-000730–7 (electronic version)

ISBN 978–92–4-000731–4 (print version)

Suggested citation. WHO consolidated guidelines on tuberculosis. Module 3: diagnosis – rapid diagnostics for tuberculosis detection. Geneva: World Health Organization; 2020. Licence: CC BY-NC-SA 3.0 IGO.

Cataloguing-in-Publication (CIP) data. CIP data are available at http://apps.who.int/iris.

Technical editing by Cadman and design by Inis Communication.

WHO consolidated guidelines on tuberculosis

Module 3: Diagnosis

Rapid diagnostics for tuberculosis detection

World Health Organization

Contents

Acknowledgements ... iv

Abbreviations and acronyms ... xv

Definitions ... xvii

Executive summary ... xviii

Introduction ... 1

- Scope of the document ... 2
- Target audience ... 2

Recommendations ... 3

- Section 1. Molecular assays intended as initial tests for TB ... 3
- Section 2. Loop-mediated isothermal amplification ... 43
- Section 3. First-line LPAs ... 49
- Section 4. Second-line LPAs ... 55
- Section 5. Lateral flow urine lipoarabinomannan assay ... 61

Research gaps ... 74

References ... 77

Annex 1: Guideline development methods ... 79

Web Annex 1. List of studies included in systematic review ... 83

Web Annex 2. GRADE profiles ... 83

Web Annex 3. Evidence to decision tables ... 83

Web Annex 4. Evidence synthesis and analysis ... 83

Acknowledgements

The recommendations and remarks in this policy guideline on tuberculosis (TB) are the result of the collaborative effort of professionals from a range of specialties. The World Health Organization (WHO) is grateful for their time and support. There were separate Guideline Development Groups (GDGs) for each of the guidelines that have been included in these consolidated guidelines; hence, the acknowledgements provided below are specific to each WHO guideline.

Molecular assays intended as initial tests

Guideline Development Group

Viet Nhung Nguyen (Co-Chair), National TB Control Programme, Ministry of Health, Hanoi, Viet Nam; Holger Schünemann (Co-Chair), McMaster University, Hamilton, Canada; Denise Arakaki-Sanchez, Ministry of Health, Brasilia, Brazil; David Branigan, Treatment Action Group, New York, United States of America (USA); Petra de Haas, KNCV Tuberculosis Foundation, Hague, Netherlands; Patricia Hall, Centers for Disease Control and Prevention (CDC), Atlanta, USA; Rumina Hasan, Department of Pathology and Microbiology, Aga Khan University, Karachi, Pakistan; Nagalineswaran Kumarasamy, Chief Medical Officer, YRG Centre for AIDS Research and Education, Voluntary Health Services, Chennai, India; Leen Rigouts, Prince Leopold Institute of Tropical Medicine, Brussels, Belgium; Thomas Shinnick, independent consultant, Atlanta, USA; Sabira Tahseen, National Tuberculosis Control Programme, Ministry of National Health Services Regulations and Coordination, Government of Pakistan, Islamabad, Pakistan; Ezio Tavora Dos Santos Filho, Civil Society Task Force Member, Rio de Janeiro, Brazil; Mercy Annapoorani Thiruthuvadoss, Blossom Trust, Tamil Nadu, India; Carrie Tudor, TB Project Director, International Council of Nurses, Durban, South Africa; Diana Vakhrusheva, National Medical Research Center of Phthisiopulmonology and Infectious Diseases, Yekaterinburg, Russian Federation; Anna Vassall, Reader in Health Economics, London School of Hygiene & Tropical Medicine (LSHTM), London, United Kingdom of Great Britain and Northern Ireland (United Kingdom); and Zhao Yanlin, National Center for TB Control, China CDC, Beijing, China.

External Review Group

Martina Casenghi, Elizabeth Glaser Pediatric AIDS Foundation, Washington, USA; Jeremiah Chakaya Muhwa, International Union Against TB and Lung Diseases, Nairobi, Kenya; Moses Joloba, Uganda Supranational Reference Laboratory (SRL), Kampala, Uganda; Katharina Kranzer, LSHTM, London, United Kingdom; Lindiwe Mvusi, NTP, Pretoria, South Africa; Norbert Ndjeka, NTP, Pretoria, South Africa; Marieke van der Werf, ECDC, Stockholm, Sweden; Francis Varaine, Médecins Sans Frontières, Paris, France.

Systematic review team

Flor Lucia Gonzalez Fernandez, International AIDS Society, Geneva, Switzerland; Frederick Haraka, Ifakara Health Institute, Bagamoyo, United Republic of Tanzania; David J Horne, University of Washington, Seattle, USA; Alexander Kay, Baylor College of Medicine, Houston, USA; Mikashmi Kohli, McGill University, Montreal, Canada; Anna M Mandalakas, Baylor College of Medicine, Houston, USA; Eleanor Ochodo, Stellenbosch University, Stellenbosch, South Africa; Klaus Reither, Ifakara Health

Institute, Bagamoyo, United Republic of Tanzania; Karen Steingart, Cochrane Infectious Diseases Group, Liverpool School of Tropical Medicine, Portland, USA; Yemisi Takwoingi, Institute of Applied Health Research, University of Birmingham, Birmingham, United Kingdom; Jerry Zifodya, Tulane University, New Orleans, USA; and Alice Zwerling, School of Epidemiology and Public Health, University of Ottawa, Canada.

Consultants with additional technical expertise

Daniela Cirillo, San Raffaele SRL, Milan, Italy; Luis Cuevas, Liverpool School of Tropical Medicine, Liverpool, United Kingdom; Nora Engel, Maastricht University, Netherlands; Anisa Hajizadeh, McMaster University, Hamilton, Canada; Nazir Ismail, Centre for Tuberculosis, National Institute for Communicable Diseases, Johannesburg, South Africa; Tamara Lotfi, Faculty of Medicine, American University of Beirut, Beirut, Lebanon; Adam Penn-Nicholson, Foundation for Innovative New Diagnostics (FIND), Geneva, Switzerland; Samuel Schumacher, FIND, Geneva, Switzerland; Rosa Stalteri, McMaster University, Hamilton, Canada; and Elisabetta Walters, Desmond Tutu TB Centre, Department of Paediatrics and Child Health, Stellenbosch University, Cape Town, South Africa.

Observers

Karen Heichman, Innovative Technology Solutions, Global Health, Bill & Melinda Gates Foundation, Seattle, USA; Tamara Kredo, South African Medical Research Council, Cape Town, South Africa; Amy Piatek, United States Agency for International Development (USAID), Washington, USA; Morten Ruhwald, FIND, Geneva, Switzerland; Raynal Squires, Public Health Laboratories, WHO Regional Office for the Eastern Mediterranean, Cairo, Egypt; Wayne van Gemert, TB Diagnostics Market Strategies, Stop TB Partnership, Geneva, Switzerland; and Mohammed Yassin, The Global Fund to Fight AIDS, Tuberculosis and Malaria, Geneva, Switzerland.

WHO steering committee

Work on these guidelines was overseen by Alexei Korobitsyn with input from Dennis Falzon, Cecily Miller, Charalampos (Babis) Sismanidis, Irwin Law, Philippe Glaziou, Katherine Floyd, Fuad Mirzaev, and Anna Stukalova (all WHO Global TB Programme), Lara Vojnov and Sathvinder Singh (both WHO Global HIV Programme), Sadia Siddiqui (EDL secretariat), Jean de Dieu Iragena (WHO/AFRO) under the overall coordination of Matteo Zignol and Karin Weyer (WHO Global TB Programme) and the direction of Tereza Kasaeva (Director of the WHO Global TB Programme).

Funding

Funding from USAID, through the USAID–WHO consolidated grant no. US-2016–0961, is gratefully acknowledged. The views of the funding agency have not influenced the development and content of these guidelines.

Conflict of interest assessment for GDG members and ERG members

GDG member	Interests declared	Conclusion
Holger Schünemann	None declared	No conflict of interest
Viet Nhung Nguyen	None declared	No conflict of interest
Rumina Hasan	None declared	No conflict of interest
David Branigan	None declared	No conflict of interest
Nagalineswaran Kumarasamy	None declared	No conflict of interest
Petra de Haas	None declared	No conflict of interest
Denise Arakaki-Sanchez	None declared	No conflict of interest
Zhao Yanlin	None declared	No conflict of interest
Diana Vakhrusheva	None declared	No conflict of interest
Sabira Tahseen	None declared	No conflict of interest
Ezio Tavora Dos Santos Filho	Coordinated CABs to the PROVE-IT study (TREAT-TB grant, Union/ USAID) in Brazil (REDE-TB) from 2010 to 2015. Currently following up, as interested party (not member of study team, but as CAB coordinator of other studies) in the implementation of the Truenat validation study in Brazil, among other BRICS cooperation studies. Intend to follow up the study, settling CABs oversight and protocol analysis in Brazil and other partner countries.	Significant conflict of interest perceived for Molbio Truenat evaluation. Exclude from a discussion on Molbio Truenat.
Leen Rigouts	Her research unit received Genoscholar kits for free for evaluation purposes. Data on PZA drug testing have been published and extension is planned. Data on second-line drug testing are in preparation for publication. Ongoing evaluation PZA testing through FIND, in amount of approx. US$ 10 000.	Non-significant conflict of interest
Carrie Tudor	Employment (starting January 2015) in the International Council of Nurses. ICN TB project received funding from Eli Lilly Foundation – Lilly MDR-TB Partnership. Funding received was approx. US$ 1 000 000 from 2013 to 2019. Current funding period for 2019 is approx. US$ 100 000.	Non-significant conflict of interest

GDG member	Interests declared	Conclusion
Anna Vassall	Research funding for study into hepatitis C diagnostic (unrelated to TB) in FIND, in amount of US$ 30 000.	Non-significant conflict of interest
Thomas Shinnick	Former employee of the US CDC, which has a similar mission to WHO to improve TB control globally. CDC supported travel and research related to the work on laboratory services needed for TB control in January 2016. As an independent consultant, received contracts and travel support from WHO, FIND and USAID for work related to laboratory strengthening and developing global guidance documents.	Non-significant conflict of interest
Patricia Hall	CDC to pay for all travel costs associated with attendance of GDG meeting.	Non-significant conflict of interest
Mercy Annapoorani Thiruthuvadoss	Unitaid has provided travel and accommodation costs to attend Unitaid board meetings as a delegate member of the communities' delegation of Unitaid. The payments for accommodation and travel are made directly by Unitaid, not received in bank account. The amount mentioned here is for one board meeting, attended once a year, and includes a small per diem as well. The amount is around US$ 1700.	Non-significant conflict of interest

BRICS: Brazil, Russian Federation, India, China and South Africa; CAB: Community Advisory Board; CDC: Centers for Disease Control and Prevention; ERG: External Review Group; FIND: Foundation for Innovative New Diagnostics; GDG: Guideline Development Group; MDR-TB: multidrug-resistant TB; PZA: pyrazinamide; TB: tuberculosis; US: United States; USAID: United States Agency for International Development; WHO: World Health Organization.

The External Review Group (ERG) members – Martina Casenghi, Marieke van der Werf, Francis Varaine, Norbert Ndjeka, Lindiwe Mvusi, Jeremiah Chakaya Muhwa, Moses Joloba – did not declare any interests.

Conflict of interest statement for FIND

FIND is a WHO collaborating centre that works with more than 200 academic, industry, governmental and civil society partners worldwide on projects that cross six priority disease areas. All industry partnerships (including that with Molbio) are subject to review by FIND's Scientific Advisory Committee, or another independent review body; selection criteria for technologies and partnerships include due diligence, target product profiles and public sector requirements.

FIND supports the evaluation of publicly prioritized TB assays, and the implementation of WHO-approved assays. To carry out these evaluations, FIND has product evaluation agreements with several private sector companies, which strictly define its independence and neutrality vis-à-vis the companies whose products are evaluated, and which clearly define roles and responsibilities.[1]

Loop-mediated isothermal amplification

The relevant section of this document was prepared by Christopher Gilpin and Alexei Korobitsyn, with input from Wayne van Gemert and Karin Weyer (all at WHO Global TB Programme), on the basis of consensus agreed at a GDG meeting convened by WHO via an online webinar on 16 January 2016.

WHO Steering Group

Christopher Gilpin, Alexei Korobitsyn, Fuad Mirzayev, Wayne van Gemert and Karin Weyer (all at WHO Global TB Programme).

Members of the WHO GDG

Jan Brozek (McMaster University, Hamilton, Canada; Chair/GRADE methodologist), Jeremiah Chakaya Muhwa (Kenya Medical Research Institute, Kenya), Gavin Churchyard (Aurum Institute for Health Research [Aurum Institute], Johannesburg, South Africa), Daniela Maria Cirillo (San Raffaele Hospital [HSR] San Raffaele Scientific Institute/TB SRL, Italy), Paul Klatser (Royal Tropical Institute, Netherlands), Arata Kochi (independent consultant, Switzerland), Satoshi Mitarai (Japan Anti-Tuberculosis Association, Japan), Beatrice Mutayoba (Ministry of Health and Social Welfare, United Republic of Tanzania), Ingrid Oxley Oxland (Nelson Mandela Metropolitan University, South Africa), Thomas M. Shinnick (independent consultant, USA), Karen Steingart (Liverpool School of Tropical Medicine, Liverpool, United Kingdom), Wendy Stevens (University of the Witwatersrand, Johannesburg, South Africa), Francis Varaine (Médecins Sans Frontières, Paris, France), Anna Vassall (LSHTM, London, United Kingdom) and Yasuhiro Yasutomi (National Institutes of Biomedical Innovation, Health and Nutrition, Japan).

Systematic review authors

Adithya Cattamanchi (lead author, systematic review; San Francisco General Hospital, University of California San Francisco [UCSF], San Francisco, USA), Katherine Farr (San Francisco General Hospital, UCSF, San Francisco, USA), Priya B. Shete (San Francisco General Hospital, UCSF, San Francisco, USA), Hojoon Sohn (lead author, economic evaluation; Department of Epidemiology, Biostatistics and Occupational Health, McGill University, Montreal, Canada) and Luke Strnad (San Francisco General Hospital, UCSF, San Francisco, USA).

External Review Group

Kathleen England (KNCV Tuberculosis Foundation, Hague, Netherlands), Levan Gagnidze (International Organization for Migration, Bangkok, Thailand), Rumina Hassan (TB SRL Pakistan, Aga Khan University, Karachi, Pakistan), Nazir Ismail (TB SRL, National Institute of Communicable Diseases, South Africa), Richard Lumb (Adelaide TB SRL, Australia), Enos Masini (National TB Programme, Kenya), Alaine Nyaruhirira (Management Sciences for Health, South Africa), Somsak Rienthong (Bangkok TB SRL, Thailand), Leen Rigouts (Institute of Tropical Medicine, Brussels, Belgium) and Maria Alice Telles (Management Sciences for Health, Brazil).

1 For more information on FIND's policy and guidelines for working with private sector partners, see https://www.finddx.org/wp-content/uploads/2019/03/Private-Sector-Partners-Policy_PL-02-08-01_V1.1_Nov2018.pdf (accessed 12 June 2020).

Acknowledgement of financial support

Funding from USAID is gratefully acknowledged through USAID–WHO Consolidated Grant No. GHA-G-00–09–00003/US-2014–741.

First-line line probe assay

The relevant section of this document was prepared by Christopher Gilpin and Alexei Korobitsyn, with input from Karin Weyer (all at the WHO Global TB Programme), on the basis of consensus agreed at a GDG meeting convened by WHO in Montreux, Switzerland, on 2 March 2016.

WHO Steering Group

Christopher Gilpin, Alexei Korobitsyn, Fuad Mirzayev and Wayne van Gemert and Karin Weyer (all at WHO Global TB Programme).

Members of the WHO GDG

Holger Schünemann (Chair/methodologist; McMaster University, Canada), Gavin Churchyard (Aurum Institute, Johannesburg, South Africa), Daniela Maria Cirillo (HSR San Raffaele Scientific Institute/TB SRL, Italy), Chris Coulter (Queensland Mycobacterium Reference Laboratory, Australia), Greg Fox (University of Sydney, Australia), Moses Joloba (National Reference Laboratory of the National TB and Leprosy Programme, Uganda), James Posey (CDC, USA), Michael Rich (Partners in Health, USA), Leen Rigouts (Institute of Tropical Medicine, Brussels, Belgium), Thomas M. Shinnick (independent consultant, USA), Rebecca Tadokera (Human Sciences Research Council, South Africa) (unable to attend), Marie Alice Telles (independent TB laboratory consultant for the Pan American Health Organization, Brazil) and Francis Varaine (Médecins Sans Frontières, Paris, France).

Systematic review authors

Patrick Cudahy (Yale Medical Center, USA), Claudia M. Denkinger (FIND, Switzerland), Ruvandhi R. Nathavitharana (Beth Israel Deaconess Medical Center, USA), Madhukar Pai (McGill University, Montreal, Canada), Samuel G. Schumacher (FIND, Switzerland) and Karen Steingart (Liverpool School of Tropical Medicine, Liverpool, United Kingdom).

Observers

Sevim Ahmedov (USAID, USA), Emmanuelle Cambau (Groupe Hospitalier Lariboisière-Fernand Widal, France), David Dolinger (FIND, Switzerland), Miranda Langendam (University of Amsterdam, Netherlands), Thomas Schön (Kalmar County Hospital and Linköping University, Sweden) and Belay Tessema (FIND, Switzerland).

External Review Group

Heather Alexander (CDC, USA), Martina Casenghi (Médecins Sans Frontières Access Campaign, Switzerland), Kathleen England (KNCV Tuberculosis Foundation, Hague, Netherlands), Rumina Hasan (TB SRL Pakistan, Aga Khan University, Karachi, Pakistan), Nazir Ismail (TB SRL, National Institute of Communicable Diseases, South Africa), Beatrice Lopez (Buenos Aires TB SRL, Argentina), Richard Lumb (Adelaide TB SRL, Australia), Satoshi Mitarai (Japan Anti-Tuberculosis Association, Japan), Alaine Umubyeyi Nyaruhirira (Management Sciences for Health, South Africa), Rohit Sarin (LRS Institute of TB and Respiratory Diseases, India) and Alena Shrahina (National TB Programme, Belarus).

Acknowledgement of financial support

Funding from the Bill & Melinda Gates Foundation as well as USAID through USAID–WHO consolidated grant no. GHA-G-00–09–00003/US-2014–741 is gratefully acknowledged.

Declaration and management of conflicts of interest

All contributors completed a WHO declaration of interests (DOI) form. All declarations were evaluated by members of the Steering Group to determine whether a possible financial conflict of interest might warrant exclusion from membership of the GDG or the ERG, or from the discussion part of the guideline development process. Contributors with intellectual conflicts of interest were not excluded from membership of the GDG, because broader expertise on drug-susceptibility testing was considered to be part of the criteria for selection. In addition, the diversity and representation in the groups was large enough to balance and overcome any potential intellectual conflicts of interest. During the guideline development process and meeting, the emergence of intellectual conflicts of interest was monitored by the Chair, and any perceived intellectual conflict of interest was discussed with members of the GDG.

The following interests were declared.

None declared

Holger Schünemann (Chair), Patrick Cudahy, Claudia M. Denkinger, Moses Joloba, Miranda Langendam, Ruvandhi R. Nathavitharana, James Posey, Thomas Schön, Samuel Schumacher, Karen Steingart, Belay Tessema, Grant Theron and Francis Varaine declared no conflicts of interest.

Declared, determined to be insignificant

Sevim Ahmedov declared that the costs of his participation in the meeting were covered by USAID.

Emmanuelle Cambau was reimbursed by the European Society of Clinical Microbiology and Infectious Diseases for participating in the European Congress on Clinical Microbiology and Infectious Diseases from 2012 to 2015.

Gavin Churchyard received a research grant to evaluate the national rollout of the GeneXpert® MTB/RIF assay in South Africa (the Bill & Melinda Gates Foundation provided US$ 11 million to the Aurum Institute, Johannesburg, South Africa).

Daniela Maria Cirillo had received research grants from FIND and the Italian government (€ 17 000) to evaluate a new TB test.

Chris Coulter declared taking part in short-term consultancies for WHO (<A$ 5000), receiving a research grant to study TB transmission in Australia, taking part in a whole-genome sequencing study (National Health and Medical Research Council research collaboration grant of <A$ 18 000), and providing laboratory support services to Papua New Guinea (A$ 240 000 for funding provided by the Australian government to the TB SRL).

David Dolinger is employed by a commercial entity and receives US$ 190 000 per year; he is also working with FIND to assess new TB diagnostics.

Gregory Fox received the Otsuka/Union Young Innovators' Award at the 2015 Union World Conference on Lung Health (US$ 10 000 in airfares, accommodation and per diem for the meeting).

Michael Rich is employed by Partners in Health to work on clinical care guidelines and on the programmatic management of multidrug-resistant TB; he has also consulted on behalf of WHO; and he is conducting research on bedaquiline and delamanid as a recipient of Unitaid's endTB grant.

Leen Rigouts has been involved in evaluating Nipro line probe assays for pyrazinamide and second-line agents.

Thomas M. Shinnick was a former employee of the US CDC. The CDC supported his travel and research related to his work on the laboratory services needed to control TB. He had represented the CDC's positions on laboratory services needed for TB diagnosis, treatment and control, and served on the Data and Safety Monitoring Board organized by Otsuka Pharmaceutical for the clinical trial of delamanid. He declared that no remuneration had been received.

Maria Alice Telles had worked for FIND as a consultant providing training on the GeneXpert MTB/RIF assay (US$ 4000) and participated in a meeting on the BACTEC™ mycobacterial growth indicator tube (MGIT; Becton Dickinson, Franklin Lakes, USA), with Beckton Dickinson funding travel and per diem expenses.

Declared, determined to be significant

None.

Second-line line probe assay

The relevant section of this document was prepared by Christopher Gilpin and Alexei Korobitsyn, with input from Karin Weyer (all at the WHO Global TB Programme), on the basis of consensus agreed at a GDG meeting convened by WHO in Montreux, Switzerland, on 2 March 2016.

WHO Steering Group

Christopher Gilpin, Alexei Korobitsyn, Fuad Mirzayev, Dennis Falzon, Matteo Zignol and Karin Weyer (all at WHO Global TB Programme).

Members of the WHO GDG

Holger Schünemann (Chair/methodologist; McMaster University, Canada), Gavin Churchyard (Aurum Institute, Johannesburg, South Africa), Daniela Maria Cirillo (HSR San Raffaele Scientific Institute/TB SRL, Italy), Chris Coulter (Queensland Mycobacterium Reference Laboratory, Australia), Greg Fox (University of Sydney, Australia), Moses Joloba (National Reference Laboratory of the National TB and Leprosy Programme, Uganda), James Posey (CDC, USA), Michael Rich (Partners in Health, USA), Leen Rigouts (Institute of Tropical Medicine, Brussels, Belgium), Thomas M. Shinnick (independent consultant, USA), Rebecca Tadokera (Human Sciences Research Council, South Africa) (unable to attend), Marie Alice Telles (independent TB laboratory consultant for the Pan American Health Organization, Brazil) and Francis Varaine (Médecins Sans Frontières, Paris, France).

Systematic review authors

Karen Steingart, Cochrane Infectious Diseases Group, Liverpool School of Tropical Medicine, Portland, USA; Grant Theron, Stellenbosch University, Department of Biomedical Sciences.

Observers

Sevim Ahmedov (USAID, USA), Emmanuelle Cambau (Groupe Hospitalier Lariboisière-Fernand Widal, France), Miranda Langendam (University of Amsterdam, Netherlands), Thomas Schön (Kalmar County Hospital and Linköping University, Sweden) and Belay Tessema (FIND, Switzerland).

External Review Group

Heather Alexander (CDC, USA), Martina Casenghi (Médecins Sans Frontières Access Campaign, Switzerland), Kathleen England (KNCV Tuberculosis Foundation, Hague, Netherlands), Rumina Hasan (TB SRL Pakistan, Aga Khan University, Karachi, Pakistan), Nazir Ismail (TB SRL, National Institute of Communicable Diseases, South Africa), Beatrice Lopez (Buenos Aires TB SRL, Argentina), Richard Lumb (Adelaide TB SRL, Australia), Satoshi Mitarai (Japan Anti-Tuberculosis Association, Japan), Alaine Umubyeyi Nyaruhirira (Management Sciences for Health, South Africa), Rohit Sarin (LRS Institute of TB and Respiratory Diseases, India) and Alena Shrahina (National TB Programme, Belarus).

Declaration and management of conflicts of interest

All contributors completed a WHO declaration of interests (DOI) form. All declarations were evaluated by members of the Steering Group to determine whether a possible financial conflict of interest might warrant exclusion from membership of the GDG or the ERG, or from the discussion part of the guideline development process. Contributors with intellectual conflicts of interest were not excluded from membership of the GDG, because broader expertise on drug-susceptibility testing was considered to be part of the criteria for selection. In addition, the diversity and representation in the groups was large enough to balance and overcome any potential intellectual conflicts of interest. During the guideline development process and meeting, the emergence of intellectual conflicts of interest was monitored by the Chair, and any perceived intellectual conflict of interest was discussed with members of the GDG.

The following interests were declared.

None declared

Holger Schünemann (Chair), Patrick Cudahy, Claudia M. Denkinger, Moses Joloba, Miranda Langendam, Ruvandhi R. Nathavitharana, James Posey, Thomas Schön, Samuel Schumacher, Karen Steingart, Belay Tessema, Grant Theron and Francis Varaine declared no conflicts of interest.

Declared, determined to be insignificant

Sevim Ahmedov declared that the costs of his participation in the meeting were covered by USAID.

Emmanuelle Cambau was reimbursed by the European Society of Clinical Microbiology and Infectious Diseases for participating in the European Congress on Clinical Microbiology and Infectious Diseases from 2012 to 2015.

Gavin Churchyard received a research grant to evaluate the national rollout of the GeneXpert® MTB/RIF assay in South Africa (the Bill & Melinda Gates Foundation provided US$ 11 million to the Aurum Institute, Johannesburg, South Africa).

Daniela Maria Cirillo had received research grants from FIND and the Italian government (€ 17 000) to evaluate a new TB test.

Chris Coulter declared taking part in short-term consultancies for WHO (<A$ 5000), receiving a research grant to study TB transmission in Australia, taking part in a whole-genome sequencing study (National Health and Medical Research Council research collaboration grant of <A$ 18 000), and providing laboratory support services to Papua New Guinea (A$ 240 000 for funding provided by the Australian government to the TB SRL).

Gregory Fox received the Otsuka/Union Young Innovators' Award at the 2015 Union World Conference on Lung Health (US$ 10 000 in airfares, accommodation and per diem for the meeting).

Michael Rich is employed by Partners in Health to work on clinical care guidelines and on the programmatic management of multidrug-resistant TB; he has also consulted on behalf of WHO; and he is conducting research on bedaquiline and delamanid as a recipient of Unitaid's endTB grant.

Leen Rigouts has been involved in evaluating Nipro line probe assays for pyrazinamide and second-line agents.

Thomas M. Shinnick was a former employee of the US CDC. The CDC supported his travel and research related to his work on the laboratory services needed to control TB. He had represented the CDC's positions on laboratory services needed for TB diagnosis, treatment and control, and served on the Data and Safety Monitoring Board organized by Otsuka Pharmaceutical for the clinical trial of delamanid. He declared that no remuneration had been received.

Maria Alice Telles had worked for FIND as a consultant providing training on the GeneXpert MTB/RIF assay (US$ 4000) and participated in a meeting on the BACTEC™ mycobacterial growth indicator tube (MGIT; Becton Dickinson, Franklin Lakes, USA), with Beckton Dickinson funding travel and per diem expenses.

Declared, determined to be significant

None.

Acknowledgement of financial support

Funding from the Bill & Melinda Gates Foundation as well as USAID through USAID–WHO consolidated grant no. GHA-G-00–09–00003/US-2014–741 is gratefully acknowledged.

Lateral flow urine lipoarabinomannan assay

Guideline Development Group

Holger Schünemann, McMaster University, Hamilton, Canada (Chair); Heather Alexander, CDC, Atlanta, USA; Gavin Churchyard, Aurum Institute, Johannesburg, South Africa; Kathleen England, Médecins Sans Frontières, Geneva, Switzerland; Rumina Hasan, Department of Pathology and Microbiology, Aga Khan University, Karachi, Pakistan; Diane Havlir, UCSF, San Francisco, USA; Nagalineswaran Kumarasamy, Chief Medical Officer, YRG Centre for AIDS Research and Education, Voluntary Health Services, Chennai, India; Gracia Violeta Ross Quiroga, TB civil society representative, La Paz, Bolivia (Plurinational State of); Kenly Sikwese, Coordinator, African Community Advisory Board (AfroCAB), Lusaka, Zambia; Wendy Stevens, National Health Laboratory Services and Medical School – University of the Witwatersrand, Johannesburg, South Africa; Carrie Tudor, TB Project Director, International Council of Nurses, Durban, South Africa; and Anna Vassall, Reader in Health Economics, LSHTM, London, United Kingdom.

External Review Group

Maria Alice da Silva Telles, Management Sciences for Health, São Paulo, Brazil; Levan Gagnidze, International Organization for Migration, Bangkok, Thailand; Jamilya Ismailova, Project HOPE, Tajikistan; Andrei Maryandyshev, Northern University, Arkhangelsk, Russian Federation; Alaine Nyaruhirira, Management Sciences for Health, Pretoria, South Africa; Rohit Sarin, Institute of Tuberculosis and Respiratory Diseases, New Delhi, India; and Francis Varaine, Médecins Sans Frontières, Paris, France.

Systematic review team

Stephanie Bjerrum, Department of Clinical Research, Research Unit of Infectious Diseases / Department of Infectious Diseases, University of Southern Denmark / Odense University Hospital (co-hosted by the Aga Khan University Hospital), Nairobi, Kenya; Maunankh Shah, Division of Infectious Diseases, Center for TB Research and Center for Clinical Global Health Education, Johns Hopkins University, Baltimore, USA; Karen Steingart, Cochrane Infectious Diseases Group, Liverpool School of Tropical Medicine, Portland, USA; and Alice Zwerling, School of Epidemiology and Public Health, University of Ottawa, Canada.

Consultants with additional technical expertise

Nim Arinaminpathy, Faculty of Medicine, School of Public Health, Imperial College London, United Kingdom; Claudia Denkinger, Division of Tropical Medicine, University of Heidelberg, Heidelberg, Germany; Nora Engel, Maastricht University, Netherlands; Helena Huerga, Epicentre / Médecins Sans Frontières, Brussels, Belgium; Emmanuel Moreau, FIND, Geneva, Switzerland; Krishna Reddy, Tobacco Research and Treatment Center, Massachusetts General Hospital, Medical Practice Evaluation Center, Cambridge, USA; Saskia Ricks, Imperial College London, United Kingdom; Samuel Schumacher, FIND, Geneva, Switzerland; and Rita Szekely, FIND, Geneva, Switzerland.

Observers

Patricia Hall, TB and Clinical Monitoring, CDC, Atlanta, USA; Karen Heichman, Innovative Technology Solutions, Global Health, Bill & Melinda Gates Foundation, Seattle, USA; and Wayne van Gemert, TB Diagnostics Market Strategies, Stop TB Partnership, Geneva, Switzerland.

WHO steering committee

Work on these guidelines was overseen by Christopher Gilpin and Alexei Korobitsyn with input from Annabel Baddeley, Licé González-Angulo and Fuad Mirzayev (all at WHO Global TB Programme), WHO, Geneva, Switzerland, and Meg Doherty, Satvinder Singh and Lara Vojnov (all at WHO HIV Department), WHO, Geneva, Switzerland, under the overall coordination of Karin Weyer (WHO Global TB Programme) and the direction of Tereza Kasaeva (Director of the WHO Global TB Programme).

The guidelines were drafted by Alexei Korobitsyn with input from Annabel Baddeley, Christopher Gilpin and Lara Vojnov on the basis of consensus achieved at the GDG meeting, 14–16 May 2019.

Technical editing was provided by Hilary Cadman and her team at Cadman Editing Services, Australia.

Funding

Funding from USAID, through the USAID–WHO consolidated grant no. US-2016–0961, is gratefully acknowledged. The views of the funding agency have not influenced the development and content of these guidelines.

Abbreviations and acronyms

Ag	antigen
AlereLAM	Alere Determine™ TB LAM Ag
CI	confidence interval
COI	conflict of interest
CRS	composite reference standard
CSF	cerebrospinal fluid
DNA	deoxyribonucleic acid
DOI	declaration of interests
DST	drug-susceptibility testing
ERG	External Review Group
EtD	evidence to decision
FIND	Foundation for Innovative New Diagnostics
FL-LPA	first-line line probe assay
FujiLAM	Fujifilm SILVAMP TB LAM
GDG	Guideline Development Group
GRADE	Grading of Recommendations Assessment, Development and Evaluation
HIV	human immunodeficiency virus
LAM	lipoarabinomannan
LAMP	loop-mediated isothermal amplification
LF-LAM	lateral flow urine lipoarabinomannan assay
LPA	line probe assay
MDR-TB	multidrug-resistant tuberculosis
MDR/RR-TB	multidrug- or rifampicin-resistant tuberculosis
MRS	microbiological reference standard
MTBC	*Mycobacterium tuberculosis* complex
NAAT	nucleic acid amplification test
NTP	national TB programme
PCR	polymerase chain reaction
PICO	population, intervention, comparator and outcomes
PLHIV	people living with HIV
QUADAS	quality assessment of diagnostic accuracy studies
RR-TB	rifampicin-resistant tuberculosis
SL-LPA	second-line line probe assay
SLID	second-line injectable drug

STARD	Standards for Reporting Diagnostic Accuracy Studies
TB	tuberculosis
UN	United Nations
USA	United States of America
USAID	United States Agency for International Development
UV	ultraviolet
WHO	World Health Organization
XDR-TB	extensively drug-resistant tuberculosis

Definitions

Advanced HIV disease: for adults, adolescents, and children aged 5 years or more, "advanced HIV disease" is defined as a CD4 cell count of less than 200 cells/mm^3 or a WHO clinical stage 3 or 4 event at presentation for care. All children with HIV aged under 5 years should be considered as having advanced disease at presentation.

Age groups: the following definitions for adults and children are used in these guidelines for the purpose of implementing recommendations (countries may have other definitions under their national regulations)[2]:

- an adult is a person aged 15 years and older;
- a child is a person aged under 15 years.

Grading of Recommendations Assessment, Development and Evaluation (GRADE): a system for rating quality of evidence and strength of recommendations; the GRADE approach is explicit, comprehensive, transparent and pragmatic, and is increasingly being adopted by organizations worldwide.

Inpatient health care setting: a health care facility where patients are admitted and assigned a bed while undergoing diagnosis and receiving treatment and care, for at least one overnight stay.

Outpatient health care setting: a health care facility where patients are undergoing diagnosis and receiving treatment and care but are not admitted for an overnight stay (e.g. an ambulatory clinic or a dispensary).

[2] In Section 5. Lateral flow urine lipoarabinomannan assay the following definitions for adults, adolescents and children were used: an adult is a person older than 19 years of age; an adolescent is a person 10–19 years of age inclusive; and a child is a person under 10 years of age.

Executive summary

Background

The political declaration at the first United Nations (UN) high-level meeting on tuberculosis (TB) held on 26 September 2018 included commitments by Member States to four new global targets.[3] One of these targets is to diagnose and treat 40 million people with TB in the 5-year period 2018–2022. The approximate breakdown of the target is about 7 million in 2018 and about 8 million in subsequent years. The traditional method for diagnosing TB using a light microscope, developed more than 100 years ago, has in recent years been challenged by several new methods and tools. These methods are based on either the detection of mycobacterial antigens or on the detection of mycobacterial DNA.

The novel tools to detect presence of *Mycobacterium tuberculosis* and resistance to anti-TB drugs call for evidence-based policy recommendations. The World Health Organization (WHO) has published a number of guidelines developed by WHO-convened Guideline Development Groups (GDGs), using the Grading of Recommendations Assessment, Development and Evaluation (GRADE) approach to summarize the evidence and to formulate policy recommendations and accompanying remarks. However, the growing number of published guidelines complicates the overview of recommendations for the intended audience (which includes health care personnel, national TB programmes and policy-makers), and WHO recognized the need to consolidate the recommendations into one document. The recommendations in this document have been presented in five guidelines published by WHO between 2016 and 2020, as shown in the box below. Earlier guidelines on diagnostics that were not developed according to the GRADE approach have not been included in this consolidated document.

WHO diagnostic guidelines included in these consolidated guidelines

- Molecular assays intended as initial tests for the diagnosis of pulmonary and extrapulmonary TB and rifampicin resistance in adults and children: Policy update 2020. Issued for the first time as a part of the present document, and corresponds to section 1.
- The use of loop-mediated isothermal amplification (TB-LAMP) for the diagnosis of pulmonary tuberculosis: policy guidance (WHO/HTM/TB/2016.11). Geneva: World Health Organization; 2016.
- The use of molecular line probe assays for the detection of resistance to isoniazid and rifampicin (WHO/HTM/TB/2016.12). Geneva: World Health Organization; 2016.
- The use of molecular line probe assays for the detection of resistance to second-line anti-tuberculosis drugs: policy guidance (WHO/HTM/TB/2016.07). Geneva: World Health Organization; 2016.
- Lateral flow urine lipoarabinomannan assay (LF-LAM) for the diagnosis of active tuberculosis in people living with HIV. Policy update 2019 (WHO/CDS/TB/2019.16). Geneva: World Health Organization; 2019.

[3] Global tuberculosis report 2019 (WHO/CDS/TB/2019.15). Geneva: World Health Organization; 2019 (https://www.who.int/tb/publications/global_report/en/, accessed 26 May 2020).

Introduction

Background

The political declaration at the first United Nations (UN) high-level meeting on tuberculosis (TB) held on 26 September 2018 included commitments by Member States to four new global targets *(1)*. One of these targets is to diagnose and treat 40 million people with TB in the 5-year period 2018–2022. The approximate breakdown of the target is about 7 million in 2018 and about 8 million in subsequent years.

Globally in 2018, 7.0 million people with a new episode of TB (i.e. new and relapse cases) were notified to national TB programmes (NTPs) and reported to the World Health Organization (WHO), a 9% increase from 6.4 million in 2017. Based on these data, the 2018 target of 7 million new and relapse cases to achieve the cumulative target set at the UN high-level meeting on TB – of 40 million in the period 2018–2022 – was achieved *(1)*.

WHO's End TB Strategy calls for the early diagnosis of TB and for universal drug-susceptibility testing (DST), highlighting the critical role of laboratories in the post-2015 era in rapidly and accurately detecting TB and drug resistance *(2)*. Of the 7.0 million new and relapse cases notified in 2018, 5.9 million (85%) had pulmonary TB. Of these, 55% were bacteriologically confirmed, a slight decrease from 56% in 2017 and 58% in 2013.[4] The remaining patients were diagnosed clinically; that is, based on symptoms, abnormalities on chest radiography or suggestive histology.

Activities to strengthen TB diagnosis must be viewed in the context of recent global initiatives to "find the missing cases", and the new global target set at the UN high-level meeting on TB in September 2018. In this context, the proportion of notified cases that are bacteriologically confirmed needs to be monitored. However, the microbiological detection of TB is critical because it allows people to be correctly diagnosed and started on the most effective treatment regimen as early as possible. Most clinical features of TB have low specificity, which may lead to false diagnoses of TB, and hence to people being enrolled in TB treatment unnecessarily. The aim should be to increase the percentage of TB cases confirmed bacteriologically (based on scaling up the use of recommended diagnostics that are more sensitive than smear microscopy).

A range of new diagnostic technologies have been endorsed by WHO during the past 10 years. The amplification and detection of *M. tuberculosis* complex (MTBC) nucleic acids is a technology that has proven to be highly sensitive and specific. Some amplification technologies have the great advantage of also being able to detect resistance to selected anti-TB drugs. The lateral flow technology detecting MTBC antigen in a point-of-care test format has also been endorsed for use in certain groups of presumptive TB patients. In total, four groups of technologies can be identified:

- real-time polymerase chain reaction (PCR) assays – for example, Xpert MTB/RIF® (Ultra) (cartridge-based) and Truenat™ (chip-based);
- line probe assays (LPAs) – for example, GenoType® MTBDR*plus* v1 and v2, Genoscholar™ NTM+MDRTB II and GenoType® MTBDR*sl*;
- loop-mediated isothermal amplification (LAMP) – for example, TB-LAMP; and

[4] A bacteriologically confirmed case is one for whom a biological specimen is positive by smear microscopy, culture or WHO-recommended rapid diagnostic test.

- antigen detection in a lateral flow format (biomarker-based detection) – for example, Alere Determine™ TB LAM Ag.

The real-time PCR applied in some tools is the most widely used technology today. These tools detect MTBC DNA and can distinguish mutations in the gene linked to drug resistance to rifampicin. The available tools use software and hardware (computers) to report results, and they require well-established laboratory networks and trained personnel.

LPAs are a family of DNA strip-based tests that can detect the MTBC strain and determine its drug resistance profile. The tests do this through the pattern of binding of amplicons (DNA amplification products) to probes that target the specific parts of the MTBC genome, common resistance-associated mutations to anti-TB agents or the corresponding wild-type DNA sequence *(3)*. LPAs are more technically complex to perform than the Xpert MTB/RIF assay; however, they can detect resistance to a broader range of first-line and second-line agents (e.g. isoniazid, fluoroquinolones and injectable agents). Testing platforms have been designed for a reference laboratory setting and are most applicable to high TB burden countries. Results can be obtained in 5 hours *(4)*. There are two large groups of assays:

- those detecting MTBC and resistance to first-line anti-TB agents (known as first-line LPAs [FL-LPAs]) – for example, GenoType MTBDR*plus* v1 and v2, Genoscholar NTM+MDRTB II; and
- those detecting resistance to second-line anti-TB agents (known as second-line LPAs [SL-LPAs]) – for example, GenoType MTBDR*sl*.

A third technology is based on LAMP reaction, in which target DNA is amplified at a fixed temperature (in contrast to the PCR, which requires a thermocycler). Detection of amplified product is done visually, using an ultraviolet (UV) lamp, directly in the reaction tubes. The method requires only basic equipment and can be implemented at the lowest levels of the laboratory network. However, detection of mutations in resistance-associated genes is not possible with this technology.

The search for a point-of-care test (i.e. a lateral flow test detecting either MTBC antigen or antibodies to MTBC) has proven complicated. However, the mycobacterial lipoarabinomannan (LAM) antigen in urine has emerged as a potential candidate. The currently available urinary LAM assays have suboptimal sensitivity and specificity, and are therefore not suitable as diagnostic tests for TB in all populations. However, unlike traditional diagnostic methods, urinary LAM assays demonstrate improved sensitivity for the diagnosis of TB among individuals coinfected with HIV.

Scope of the document

This document provides background, justification and recommendations on novel diagnostic tools for the detection of MTBC and for the presence or absence of mutations in target genes proven to be associated with anti-TB drug resistance.

Target audience

The target audience for these guidelines includes laboratory managers, clinicians and other health care staff, HIV and TB programme managers, policy-makers, technical agencies, donors and implementing partners supporting the use of TB diagnostics in resource-limited settings.

Individuals responsible for programme planning, budgeting, mobilizing resources and implementing training activities for the programmatic management of drug-resistant TB may also find this document useful.

Recommendations

Section 1. Molecular assays intended as initial tests for TB

The development of the Xpert MTB/RIF assay (Cepheid, Sunnyvale, United States of America [USA]) was a major step forward for improving the diagnosis of TB and the detection of rifampicin resistance globally. However, Xpert MTB/RIF sensitivity is suboptimal, particularly in smear-negative and HIV-associated TB patients. The Xpert MTB/RIF Ultra (Cepheid, Sunnyvale, USA), hereafter referred to as Xpert Ultra, was developed by Cepheid as the next-generation assay to overcome these limitations. It uses the same GeneXpert® platform as the Xpert MTB/RIF.

New molecular assays – the Truenat MTB, MTB Plus and MTB-RIF Dx assays (Molbio Diagnostics, Goa, India), hereafter referred to as Truenat – were developed in India, and may potentially be used at the same health system level as Xpert MTB/RIF. Of the above-mentioned assays, MTB and MTB Plus are used as initial diagnostic tests for TB, whereas MTB-RIF Dx is used as a reflex test to detect rifampicin resistance for those with positive results on the initial Truenat tests. Multisite international evaluations in settings of intended use are being implemented by the Foundation for Innovative New Diagnostics (FIND), a WHO collaborating centre for the evaluation of new diagnostic technologies. Given the similarity of the operational characteristics for Xpert MTB/RIF and Truenat, the results of the latter study were reviewed within the same Guideline Development Group (GDG) meeting.

1.1 Recommendations

This section contains five sets of recommendations, with each set being specific for a particular type of testing (initial or repeated) and type of TB (pulmonary or extrapulmonary).

1.1.1 Recommendations on Xpert MTB/RIF and Xpert Ultra as initial tests in adults and children with signs and symptoms of pulmonary TB

1.1 In adults with signs and symptoms of pulmonary TB, Xpert MTB/RIF should be used as an initial diagnostic test for TB and rifampicin-resistance detection in sputum rather than smear microscopy/culture and phenotypic DST. *(Strong recommendation, high certainty of evidence for test accuracy; moderate certainty of evidence for patient-important outcomes[5])*

1.2 In children with signs and symptoms of pulmonary TB, Xpert MTB/RIF should be used as an initial diagnostic test for TB and rifampicin-resistance detection in sputum, gastric aspirate, nasopharyngeal aspirate and stool rather than smear microscopy/culture and phenotypic DST. *(Strong recommendation, moderate certainty for accuracy in sputum; low certainty of evidence for test accuracy in gastric aspirate, nasopharyngeal aspirate and stool)*

1.3 In adults with signs and symptoms of pulmonary TB and without a prior history of TB (≤5 years) or with a remote history of TB treatment (>5 years since end of treatment), Xpert Ultra should be used as an initial diagnostic test for TB and for rifampicin-resistance detection in sputum, rather than smear microscopy/culture and phenotypic DST. *(Strong recommendation, high certainty of evidence for test accuracy)*

1.4 In adults with signs and symptoms of pulmonary TB and with a prior history of TB and an end of treatment within the last 5 years, Xpert Ultra may be used as an initial diagnostic test for TB and for rifampicin-resistance detection in sputum, rather than smear microscopy/culture and phenotypic DST. *(Conditional recommendation, low certainty of evidence for test accuracy)*

1.5 In children with signs and symptoms of pulmonary TB, Xpert Ultra should be used as the initial diagnostic test for TB and detection of rifampicin resistance in sputum or nasopharyngeal aspirate, rather than smear microscopy/culture and phenotypic DST. *(Strong recommendation, low certainty of evidence for test accuracy in sputum; very low certainty of evidence for test accuracy in nasopharyngeal aspirate)*

Remarks

For recommendation 1.2: Sputum includes expectorated and induced sputum. Studies assessing the impact of Xpert MTB/RIF on patient-important outcomes in children are lacking. The choice of the specimen will depend on the acceptability (for children, parents, health care workers and other stakeholders) and the feasibility of collecting and preparing specimens in the local context. Regarding Xpert MTB/RIF, the certainty of evidence is higher for sputum and nasopharyngeal aspirates than for other specimen types. The recommendation can be extrapolated for children living with HIV. The direct benefit from testing for rifampicin resistance in sputum (very low certainty of evidence for test accuracy) can be extrapolated to other specimens.

For recommendation 1.4: The justification for a conditional recommendation is based on:

- low certainty of evidence for test accuracy;
- uncertainty about the interpretation of Xpert Ultra trace results in patients with a prior history of disease and the associated high false-positivity rate; and
- uncertainty about the required resources.

[5] Mortality, cure, pretreatment loss to follow-up, time to diagnosis, treatment, and mortality in PLHIV.

For patients with Xpert Ultra trace results, decisions regarding treatment initiation should include considerations of the clinical presentation and the patient context (including prior treatment history, probability of relapse and other test results).

1.1.2 Recommendations on Xpert MTB/RIF and Xpert Ultra as initial tests in adults and children with signs and symptoms of extrapulmonary TB

1.6 In adults and children with signs and symptoms of TB meningitis, Xpert MTB/RIF or Xpert Ultra should be used in cerebrospinal fluid (CSF) as an initial diagnostic test for TB meningitis rather than smear microscopy/culture. *(Strong recommendation, moderate certainty of evidence for test accuracy for Xpert MTB/RIF; low certainty of evidence for test accuracy for Xpert Ultra)*

1.7 In adults and children with signs and symptoms of extrapulmonary TB, Xpert MTB/RIF may be used in lymph node aspirate, lymph node biopsy, pleural fluid, peritoneal fluid, pericardial fluid, synovial fluid or urine specimens as the initial diagnostic test for respective form of extrapulmonary TB rather than smear microscopy/culture. *(Conditional recommendation, moderate certainty of evidence for test accuracy for pleural fluid; low certainty for lymph node aspirate, peritoneal fluid, synovial fluid, urine; very low certainty for pericardial fluid, lymph nodes biopsy)*

1.8 In adults and children with signs and symptoms of extrapulmonary TB, Xpert Ultra may be used in lymph node aspirate and lymph node biopsy as the initial diagnostic test for lymph nodes TB rather than smear microscopy/culture. *(Conditional recommendation, low certainty of evidence)*

1.9 In adults and children with signs and symptoms of extrapulmonary TB, Xpert MTB/RIF or Xpert Ultra should be used for rifampicin-resistance detection rather than culture and phenotypic DST. *(Strong recommendation, high certainty of evidence for test accuracy for Xpert MTB/RIF; low certainty of evidence for Xpert Ultra)*

1.10 In HIV-positive adults and children with signs and symptoms of disseminated TB, Xpert MTB/RIF may be used in blood, as an initial diagnostic test for disseminated TB. *(Conditional recommendation, very low certainty of evidence for test accuracy)*

Remarks

For recommendation 1.6: This recommendation applies to all patients with signs and symptoms of TB meningitis. The recommendation in children with signs and symptoms of TB meningitis is based on very low certainty of evidence for test accuracy for Xpert MTB/RIF. No data were available on the accuracy of Xpert Ultra for TB meningitis in children.

For recommendation 1.7: Clinical judgement and pretest probability should guide treatment. In a high pretest probability setting (>5%), a negative test result will not rule out the condition. Available data on Xpert MTB/RIF for children have included lymph node aspirate and lymph node biopsy specimens; given the similarity of the effects, the recommendation for adults is extrapolated for children.

For recommendation 1.8: The composite reference standard for Xpert Ultra gave similar results when lymph nodes aspirate was compared to lymph nodes biopsy.

For recommendation 1.9: Clinical judgement and pretest probability should guide treatment. In a high pretest probability setting, a negative test result will not rule out the condition.

For recommendation 1.10: Blood was only evaluated in people living with HIV (PLHIV) and under particular processing specifications *(5)*, using third-generation Xpert MTB/RIF cartridges, based on one study with a small number of participants. The recommendation applies only to a particular population (HIV-positive adults with signs and symptoms of disseminated TB). This recommendation cannot be extrapolated to other patient populations.

1.1.3 Recommendations on Xpert MTB/RIF and Xpert Ultra repeated testing in adults and children with signs and symptoms of pulmonary TB[6]

1.11 In adults with signs and symptoms of pulmonary TB who have an Xpert Ultra trace positive result on the initial test, repeated testing with Xpert Ultra may not be used. *(Conditional recommendation, very low certainty of evidence for test accuracy)*

1.12 In children with signs and symptoms of pulmonary TB in settings with pretest probability below 5% and an Xpert MTB/RIF negative result on the initial test, repeated testing with Xpert MTB/RIF in sputum, gastric fluid, nasopharyngeal aspirate or stool specimens may not be used.[7] *(Conditional recommendation, low certainty of evidence for test accuracy for sputum and very low for other specimen types)*

1.13 In children with signs and symptoms of pulmonary TB in settings with pretest probability 5% or more and an Xpert MTB/RIF negative result on the initial test, repeated testing with Xpert MTB/RIF (for total of two tests) in sputum, gastric fluid, nasopharyngeal aspirate and stool specimens may be used. *(Conditional recommendation, low certainty of evidence for test accuracy for sputum and very low for other specimen types)*

1.14 In children with signs and symptoms of pulmonary TB in settings with pretest probability below 5% and an Xpert Ultra negative result on the initial test, repeated testing with Xpert Ultra in sputum or nasopharyngeal aspirate specimens may not be used. *(Conditional recommendation, very low certainty of evidence for test accuracy)*

1.15 In children with signs and symptoms of pulmonary TB in settings with pretest probability 5% or more and an Xpert Ultra negative result on the first initial test, repeated one Xpert Ultra test (for a total of two tests) in sputum and nasopharyngeal aspirate specimens may be used. *(Conditional recommendation, very low certainty of evidence for test accuracy)*

Remarks

For recommendation 1.11: Xpert Ultra trace results will require follow-up, including reassessing clinical symptoms and information on prior history of TB. In the case of suspected rifampicin resistance, repeated testing may provide additional benefit for detection as well as an initial attempt to assess rifampicin resistance.

For recommendation 1.13: The GDG felt that the implementation of the recommendation depends on the acceptability (for children, parents or caregivers, health care workers and other stakeholders) and the feasibility of conducting repeated testing in the local context. The evidence reviewed evaluated repeating the same test on the same type of specimen. However, from the data reviewed on comparing single tests on different specimen types, there appears to be no difference, regardless of which second specimen is obtained. The recommendation can be extrapolated for children living with HIV (for Xpert MTB/RIF). This includes consideration of the direct benefit from detecting rifampicin resistance

6 Based on PICO questions 3 and 4.

7 In low prevalence settings the effect of the second test was less pronounced.

in sputum samples (very low certainty of evidence for test accuracy), which can be extrapolated to other samples. The recommendation applies to a moderate or high pretest setting (>5%). If the first test result is positive, the test should not be repeated. In settings with moderate to high pretest probability, the incremental yield of more than two tests is unknown.

For recommendation 1.15: Desirable and undesirable effects were judged to be moderate, but testing twice in the moderate and high pretest probability (>5%) settings on balance may provide more benefits than harms. The recommendation is applicable for sputum and nasopharyngeal aspirates. No evidence was identified for stool and gastric aspirates.

1.1.4 Recommendations on Xpert MTB/RIF and Xpert Ultra as initial tests for pulmonary TB in adults in the general population either with signs and symptoms of TB or chest radiograph with lung abnormalities or both[8]

1.16 In adults in the general population who had either signs or symptoms of TB or chest radiograph with lung abnormalities or both, the Xpert MTB/RIF or Xpert Ultra may replace culture as the initial test for pulmonary TB. *(Conditional recommendation, low certainty of the evidence in test accuracy for Xpert MTB/RIF and moderate certainty for Xpert Ultra)*

1.17 In adults in the general population who had either a positive TB symptom screen or chest radiograph with lung abnormalities or both, one Xpert Ultra test may be used rather than two Xpert Ultra tests as the initial test for pulmonary TB. *(Conditional recommendation, very low certainty of evidence for test accuracy)*

Remarks

For recommendation 1.16: This recommendation was informed by evidence from recent national surveys of TB disease prevalence in four high TB burden countries. Indirectness of the evidence was classified as serious, given that the methods applied in TB prevalence surveys differ from usual programmatic conditions (e.g. symptom screen limited to cough for 14 days or more, and a requirement in surveys to have the results of both symptom screen and chest radiography available). In addition, inconsistency of the evidence was also classified as serious, owing to variability of the data from different countries. As a result, certainty in the estimates of effect was downgraded to low for sensitivity and moderate for specificity. **The recommendation applies only to the use of Xpert MTB/RIF or Xpert Ultra for clinical case management in situations where an immediate decision on patient treatment needs to be made and recourse to supplementary tests is not available or would incur delays.** It does not apply to scientific studies with other objectives, such as the reliable estimation of the prevalence of TB disease in the community, for which alternative testing algorithms are required (in particular, to address the issue of false-positive results, as illustrated in Table 1.17). Recommendations about the screening and diagnostic algorithms to be used in such studies are beyond the scope of this GDG. Recommendations for the diagnostic algorithm(s) to recommend in national TB prevalence surveys specifically are being developed by WHO and are scheduled for release in 2020.

For recommendation 1.17: There are concerns about losing global and national capacity for culture testing – the current reference standard for identifying active TB disease. An Xpert Ultra trace result was considered as negative in these studies. More false-positive results are expected for Xpert Ultra for pulmonary TB. **The recommendation applies only to the use of Xpert Ultra for clinical case management.** When Xpert Ultra gives a positive result, clinical management should

8 Based on PICO question 5.

be followed according to national guidelines. When Xpert Ultra gives a negative result, the patient should be re-evaluated clinically. In the case of a culture-positive result, clinical management should be followed according to national guidelines. In the case of a culture-negative result, the patient should be re-evaluated clinically. The recommendation does not apply to scientific studies with other objectives, such as the reliable estimation of the prevalence of TB disease in the community, in which alternative testing algorithms (e.g. using more than one test) may be required. Recommendations for the diagnostic algorithm(s) to be used in such studies are beyond the scope of this GDG. Recommendations for the diagnostic algorithm(s) to recommend in national TB prevalence surveys specifically are being developed by WHO and are scheduled for release in 2020.

1.1.5 Recommendations on Truenat MTB, MTB Plus and Truenat MTB-RIF Dx in adults and children with signs and symptoms of pulmonary TB[9]

1.18 In adults and children with signs and symptoms of pulmonary TB, the Truenat MTB or MTB Plus may be used as an initial diagnostic test for TB rather than smear microscopy/culture. *(Conditional recommendation, moderate certainty of evidence for test accuracy)*

1.19 In adults and children with signs and symptoms of pulmonary TB and a Truenat MTB or MTB Plus positive result, Truenat MTB-RIF Dx may be used as an initial test for rifampicin resistance rather than culture and phenotypic DST. *(Conditional recommendation, very low certainty of evidence for test accuracy)*

Remarks

For recommendation 1.18: The recommendation includes patients who are smear negative. There is uncertainty about the use of these assays in PLHIV. In smear-negative patients, the sensitivity is lower than in all patients. The indirect data on test accuracy in smear-negative patients (given that there are no data on PLHIV for this version of Truenat) made it possible to extrapolate this recommendation to PLHIV. However, the certainty of evidence for test accuracy would need to be lowered to account for additional indirectness. In the case of children, there were no data available to assess the accuracy of the test in different specimens, and not enough indirect evidence to extrapolate for specimens other than sputum. This recommendation is extrapolated to children for sputum, although the tests are expected to be less sensitive in children.

For recommendation 1.19: The Truenat is a reflex (two-step) test for rifampicin resistance. Hence, the recommendation for Truenat MTB-RIF Dx is only applicable for those patients with positive Truenat MTB or MTB Plus results.

9 Based on PICO question 7.

1.2 Test descriptions

Xpert MTB/RIF is an automated PCR test (molecular test) using the GeneXpert platform (Fig. 1.1). Xpert MTB/RIF is a single test that can detect both MTBC bacteria and rifampicin resistance within 2 hours of starting the test, with minimal hands-on technical time *(6)*.

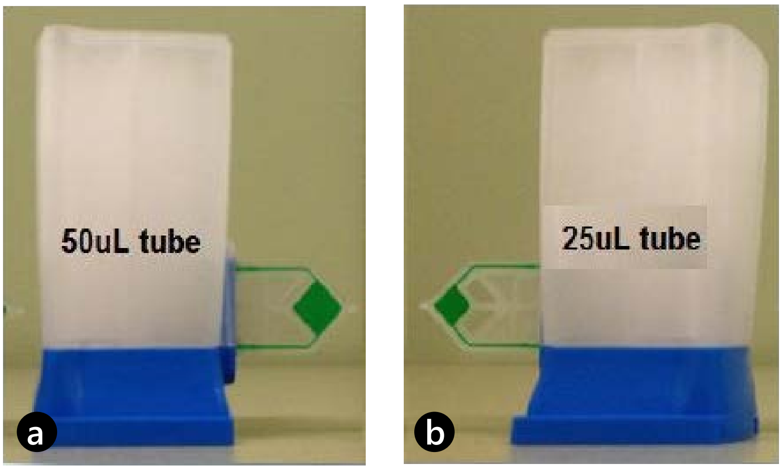
Fig. 1.1. The GeneXpert four-module instrument and the Xpert MTB/RIF test cartridge

Source: Courtesy of Cepheid.

In Xpert MTB/RIF sample processing – in contrast to conventional nucleic acid amplification tests (NAATs) – PCR amplification and detection are integrated into a single self-enclosed test unit; that is, the Xpert MTB/RIF cartridge. Following sample loading, all steps in the assay are automated and contained within the cartridge. In addition, the assay's sample reagent, used to liquefy sputum, is tuberculocidal (i.e. it has the ability to kill TB bacteria), which largely eliminates concerns about biosafety during the test procedure. These features allow the technology to be taken out of a central laboratory or reference laboratory, and to be used nearer to patients. However, Xpert MTB/RIF requires an uninterrupted and stable electrical power supply, temperature control and yearly calibration of the instrument's modules *(7)*.

Xpert Ultra uses the same GeneXpert platform as Xpert MTB/RIF; it was developed by Cepheid as the next-generation assay to overcome limitations in sensitivity for TB diagnosis. To improve assay sensitivity for the detection of MTBC, the Xpert Ultra assay incorporates two different multicopy amplification targets (IS6110 and IS1081) and has a larger DNA reaction chamber than Xpert MTB/RIF (50 µL PCR in Xpert Ultra versus 25 µL in Xpert MTB/RIF, Fig. 1.2). Xpert Ultra also incorporates fully nested nucleic acid amplification, more rapid thermal cycling, and improved fluidics and enzymes. This has resulted in Xpert Ultra having a limit of detection of 16 bacterial colony forming units (cfu) per millilitre (compared with 114 cfu/mL for Xpert MTB/RIF). To improve the accuracy of rifampicin-resistance detection, the Xpert Ultra test incorporates melting-temperature-based analysis rather than real-time PCR. Specifically, four probes identify rifampicin-resistance mutations in the rifampicin-resistance determining region of the *rpoB* gene by detecting shifts in the melting temperature away from the wild-type reference value *(8)*.

Fig. 1.2. (a) The Xpert MTB/RIF Ultra cartridge with its 50 µL reaction tube (green) and (b) the Xpert MTB/RIF cartridge with its 25 µL reaction tube (green)



Source: Courtesy of Cepheid.

The new molecular assays – the Truenat MTB, MTB Plus and MTB-RIF Dx assays – developed in India, may potentially be used at the same health system level as Xpert MTB/RIF. This policy focuses on the following Molbio devices and diagnostic tests:[10]

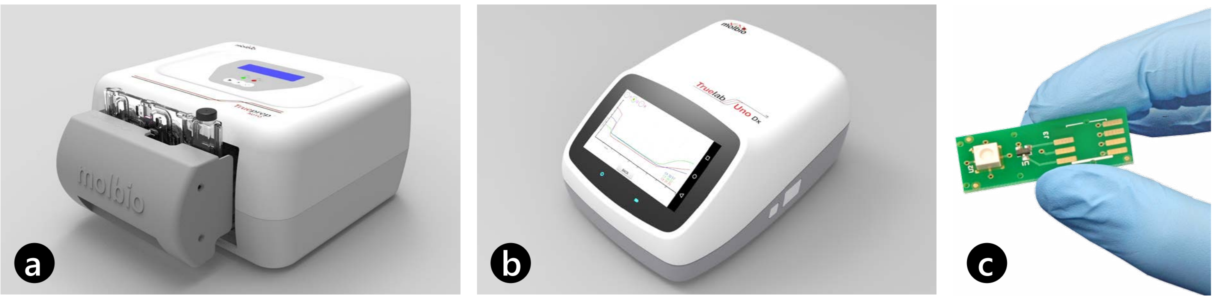
- Trueprep Auto DNA extraction system;
- Truelab DuoDx and Truelab QuattroDx micro-PCR machines;
- Truelab MTB chip;
- Truelab MTB Plus chip; and
- Truelab MTB-RIF Dx chip.

The Truenat MTB and MTB Plus assays and the rifampicin-resistance detection reflex assay (Truenat MTB-RIF Dx) (Molbio Diagnostics, India) use real-time micro-PCR for detection of *M. tuberculosis* and selected rifampicin resistance in DNA extracted from a patient's sputum specimen (Fig. 1.3). The assays use automated, battery-operated devices to extract, amplify and confirm the presence of specific genomic DNA loci, allowing for the rapid diagnosis of TB infections with minimal user input. These products are intended to be operated in peripheral laboratories with minimal infrastructure, and technicians with only minimal training can easily perform these tests routinely in their facilities and report results in under 1 hour. Moreover, with these devices, PCR testing can also be initiated at the field level, on-site.

If the Truenat MTB assay result is positive, the user may then take another aliquot of extracted DNA and run the MTB-RIF Dx assay, to detect the presence of selected rifampicin-resistance-associated mutations. The diagnostic performance of these assays has been evaluated in multicentre prospective clinical evaluation study implemented by FIND, a WHO collaborating centre for the evaluation of new diagnostic technologies, in settings of intended use in four countries (India, Peru, Ethiopia, Papua New Guinea).

[10] See http://www.molbiodiagnostics.com/products-listing.php.

Fig. 1.3. Molbio equipment to run the Truenat MTB, MTB Plus and MTB-RIF Dx assays: (a) Trueprep instrument for sample preparation, (b) Truelab Uno Dx real-time PCR instrument for running the tests, and (c) chip for real-time PCR



PCR: polymerase chain reaction.
Source: Courtesy of Molbio Diagnostics

1.3 Justification and evidence

The WHO Global TB Programme has initiated an update of the current guidelines and commissioned a systematic review on the use of Xpert MTB/RIF and Xpert Ultra for the diagnosis of TB in people with signs and symptoms of TB. The evidence on the use of the Truenat MTB, MTB Plus and MTB-RIF Dx system was generated by multisite international evaluations in settings of intended use, implemented by FIND.

The population, intervention, comparator and outcome (PICO) questions were designed to form the basis for the evidence search, retrieval and analysis.

Box 1.1. PICO questions and subquestions

PICO 1: Among adults with signs and symptoms of pulmonary TB, seeking care at health care facilities, should Xpert MTB/RIF / Xpert Ultra be used as an initial test for diagnosis of pulmonary and rifampicin resistance?

1.1. What is the impact of Xpert MTB/RIF on patient-important outcomes (cure, mortality, time to diagnosis and time to start treatment)?

1.2 What is the diagnostic accuracy of Xpert MTB/RIF for pulmonary TB and rifampicin resistance, as compared with microbiological reference standard (MRS)?[11]

1.3 What is the diagnostic accuracy of Xpert Ultra for pulmonary TB and rifampicin resistance, as compared with MRS?

PICO 2: Among children with signs and symptoms of pulmonary TB, seeking care at health care facilities, should Xpert MTB/RIF / Xpert Ultra be used as an initial test for diagnosis of pulmonary TB and rifampicin resistance?

2.1 What is the diagnostic accuracy of Xpert MTB/RIF for pulmonary TB and rifampicin resistance in children, as compared with MRS and composite reference standard (CRS)?[12]

2.2 What is the diagnostic accuracy of Xpert Ultra for pulmonary TB and rifampicin resistance in children, as compared with MRS and CRS?

[11] Culture.

[12] Positive culture or a clinical decision to initiate treatment for TB.

PICO 3: Among adults with signs and symptoms of extrapulmonary TB, seeking care at health care facilities, should Xpert MTB/RIF / Xpert Ultra be used as an initial test for diagnosis of extrapulmonary TB and rifampicin resistance?

3.1 What is the diagnostic accuracy of Xpert MTB/RIF for extrapulmonary TB and rifampicin resistance in adults, as compared with MRS and CRS?

3.2 What is the diagnostic accuracy of Xpert Ultra for extrapulmonary TB and rifampicin resistance in adults, as compared with MRS and CRS?

PICO 4: Among children with signs and symptoms of extrapulmonary TB and rifampicin resistance, seeking care at health care facilities, should Xpert MTB/RIF / Xpert Ultra be used as an initial test for diagnosis of extrapulmonary TB and rifampicin resistance?

4.1 What is the diagnostic accuracy of Xpert MTB/RIF for extrapulmonary TB and rifampicin resistance in children, as compared with MRS and CRS?

4.2 What is the diagnostic accuracy of Xpert Ultra for extrapulmonary TB and rifampicin resistance in children, as compared with MRS and CRS?

PICO 5: Among people with signs and symptoms of pulmonary TB, seeking care at health care facilities, do repeated Xpert (Ultra) tests on subsequent samples as an initial test for diagnosis of pulmonary TB and rifampicin resistance increase sensitivity/specificity compared with a single initial test?

5.1 Xpert Ultra repeated test for the diagnosis of pulmonary TB in adults with signs and symptoms of pulmonary TB who have an initial Xpert Ultra trace result, as compared with MRS?

5.2 More than one Xpert MTB/RIF versus one Xpert MTB/RIF to diagnose pulmonary TB in children with signs and symptoms of pulmonary TB, as compared with MRS?

5.3 More than one Xpert Ultra versus one Xpert Ultra to diagnose pulmonary TB in children with signs and symptoms of pulmonary TB, as compared with MRS?

PICO 6: Among adults either with signs and symptoms of TB or chest radiograph with lung abnormalities suggestive of pulmonary TB or both, should Xpert MTB/RIF or Xpert Ultra alone be used to define a case of active TB disease *(10)*?

6.1 Xpert MTB/RIF to diagnose pulmonary TB in adults in the general population with signs and symptoms of pulmonary TB or chest radiograph with lung abnormalities or both, as compared with MRS.

6.2 Xpert Ultra to diagnose pulmonary TB in adults in the general population with signs and symptoms of pulmonary TB or chest radiograph with lung abnormalities or both, as compared with MRS.

6.3 Two Xpert Ultra versus one Xpert Ultra to diagnose pulmonary TB in adults in the general population with signs and symptoms of TB or chest radiograph with lung abnormalities or both, as compared with MRS.

PICO 7: Among people with signs and symptoms of pulmonary TB, seeking care at health care facilities, should Molbio Truenat MTB, MTB Plus and MTB-RIF Dx be used as an initial test for diagnosis of pulmonary TB and rifampicin resistance?

7.1 What is the diagnostic accuracy of Truenat MTB to diagnose pulmonary TB in adults with signs and symptoms of pulmonary TB, as compared with MRS?

7.2 What is the diagnostic accuracy of Truenat MTB Plus to diagnose pulmonary TB in adults with signs and symptoms of pulmonary TB, as compared with MRS?

7.3 What is the diagnostic accuracy of Truenat MTB-RIF Dx to diagnose rifampicin resistance in adults with signs and symptoms of pulmonary TB, as compared with MRS?

Additional questions

1. What are the comparative cost, affordability and cost–effectiveness of implementation of Xpert MTB/RIF, Xpert Ultra, and Truenat MTB, MTB Plus and MTB-RIF Dx systems?

2. Are there implications for feasibility, accessibility, patient equity and human rights from the implementation of Xpert MTB/RIF, Xpert Ultra, and Truenat MTB, MTB Plus and MTB-RIF Dx systems?

The systematic reviews were conducted to summarize the current literature on the diagnostic accuracy of Xpert MTB/RIF and Xpert Ultra for the diagnosis of TB and rifampicin resistance. This was done as part of the WHO process to develop updated guidelines for use of molecular assays intended as initial tests for the diagnosis of pulmonary and extrapulmonary TB in adults and children. The data on children, where possible, were reported separately from adults.

The evaluation study of Truenat was carried out in 19 clinical sites (each with a microscopy centre attached) and seven reference laboratories in four countries. The diagnostic accuracy of the assays was evaluated when performed in the intended settings of use (i.e. microscopy centres), against microbiological confirmation (culture) as the reference standard. As part of this assessment, the performance of the Truenat assays was also compared to Xpert MTB/RIF or Xpert Ultra, on the same specimens, in reference laboratories.

The certainty of the evidence was assessed consistently through PICO questions, using the Grading of Recommendations Assessment, Development and Evaluation (GRADE) approach,[13] which produces an overall quality assessment (or certainty) of evidence and a framework for translating evidence into recommendations. The certainty of the evidence is rated as high, moderate, low or very low. These four categories "imply a gradient of confidence in the estimates" *(11)*. In the GRADE approach, even if diagnostic accuracy studies are of observational design, they start as high-quality evidence.[1]

At least two review authors independently completed the quality assessment of diagnostic accuracy studies (QUADAS)-2 assessments. Any disagreements were resolved through discussion or consultation with a third review author.

Finally, where applicable, meta-analyses were performed to estimate pooled sensitivity and specificity separately for Xpert MTB/RIF and Xpert Ultra, and separately for TB (either pulmonary or extrapulmonary) and rifampicin resistance.

Data synthesis was structured around the pre-set PICO questions list below. Details of studies included in the current analysis are given in **Web Annex 1.1** "Molecular assays as initial tests". Summary of

[13] See see https://www.gradeworkinggroup.org/.

the results and details of the evidence quality assessment are available in **Web Annex 2.1** "GRADE profiles molecular assays".

PICO 1: Among adults with signs and symptoms of pulmonary TB, seeking care at health care facilities, should Xpert MTB/RIF / Xpert Ultra be used as an initial test for diagnosis of pulmonary TB and rifampicin resistance?

1.1. What is the impact of Xpert MTB/RIF on patient-important outcomes (cure, mortality, time to diagnosis and time to start treatment)?

The aim of the review was to assess the impact on patient-important outcomes of diagnostic strategies using Xpert MTB/RIF compared with strategies using smear microscopy. The following outcomes were considered: **all-cause mortality, pretreatment loss to follow-up, cure, time to diagnosis and time to treatment initiation**.

For the impact of Xpert MTB/RIF on patient-important outcomes for TB, seven studies were included (16 421 participants): two individually randomized trials (Mupfumi 2014; Theron 2014), four cluster randomized trials (Churchyard 2015; Cox 2014; Ngwira LG 2017; Durovni 2014), and one individual patient data (IPD) meta-analysis (Di Tanna 2019) (see **Web Annex 1** for details of these and other studies). All studies were conducted in high TB burden and high TB/HIV burden countries. There were two trials in South Africa (Churchyard 2015; Cox 2014), one in Zimbabwe (Mupfumi 2014), one in Malawi (Ngwira LG 2017), one in Brazil (Durovni 2014) and two multi-country studies with sites in South Africa, United Republic of Tanzania, Zambia and Zimbabwe (Theron 2014, Di Tanna 2019). All studies were conducted in outpatient settings and enrolled participants aged 18 years or older.

Web Annex 4.1: Impact of diagnostic test Xpert MTB/RIF on patient-important outcomes for tuberculosis: a systematic review.

1.2. What is the diagnostic accuracy of Xpert MTB/RIF for pulmonary TB and rifampicin resistance, as compared with MRS?

The aim of the review was to assess the diagnostic accuracy of Xpert MTB/RIF for pulmonary TB and rifampicin resistance in adults. Randomized trials, cross-sectional studies and cohort studies were included, using respiratory specimens that evaluated Xpert MTB/RIF alone or together with Xpert Ultra against the reference standards of culture for TB detection and culture-based DST or MTBDR*plus* for rifampicin resistance. Only studies that enrolled adults (aged >15 years) were eligible. For the evaluation of TB detection, studies were included that evaluated the index tests in people with signs and symptoms of pulmonary TB, except for studies in PLHIV, where studies were eligible for inclusion irrespective of signs and symptoms of pulmonary TB (e.g. studies that performed TB screening in PLHIV as part of intensified case finding or before TB preventive therapy).

For detection of pulmonary TB, a total of 94 studies were identified. Of these, 85 studies (40 652 participants) evaluated Xpert MTB/RIF and nine studies (3881 participants) evaluated both Xpert Ultra and Xpert MTB/RIF. Of the 94 studies, 50 (53%) took place in high TB burden and 54 (57%) in high TB/HIV burden countries. Most studies had low risk of bias. Also, most studies had low concern about applicability because participants in these studies were evaluated in primary care facilities, local hospitals or both settings.

For detection of rifampicin resistance, 57 studies (8287 participants) evaluated Xpert MTB/RIF. Of the 57 studies, 27 took place in high multidrug-resistant TB (MDR-TB) burden countries. Most studies were judged as having low risk of bias.

Web Annex 4.2: Xpert MTB/RIF and Xpert Ultra for detecting active tuberculosis in adults with signs and symptoms of pulmonary TB: an updated systematic review.

1.3. What is the diagnostic accuracy of Xpert Ultra for pulmonary TB and rifampicin resistance, as compared with MRS?

For detection of pulmonary TB, a total of nine studies (3881 participants) evaluated both Xpert Ultra and Xpert MTB/RIF. For Xpert Ultra, a composite reference standard was also used that included clinical components as defined by the primary study authors. For detection of rifampicin resistance, eight studies (1039 participants) evaluated Xpert Ultra. The total number of Xpert Ultra studies includes one study that provided data for two cohorts; therefore, we classified these as two distinct studies, Mishra 2019a and Mishra 2019b. Most studies were judged as having high certainty of evidence.

Web Annex 4.2: Xpert MTB/RIF and Xpert Ultra for detecting active tuberculosis in adults with signs and symptoms of pulmonary TB: an updated systematic review.

PICO 2: Among children with signs and symptoms of pulmonary TB, seeking care at health care facilities, should Xpert MTB/RIF / Xpert Ultra be used as an initial test for diagnosis of pulmonary TB and rifampicin resistance?

2.1. What is the diagnostic accuracy of Xpert MTB/RIF for pulmonary TB and rifampicin resistance in children, as compared with MRS and CRS?

The initial search resulted in 835 individual records, with one additional reference identified through other sources, giving a total of 836 records, from which 707 were excluded. Initially, the remaining 129 articles were retrieved. After full-text review, 50 studies were included in the quantitative meta-analysis; of these, 40 (80%) took place in high TB burden countries and 10 in high TB/HIV burden countries. For pulmonary TB detection, 43 studies were included that evaluated the diagnostic accuracy of Xpert MTB/RIF in children, and three that evaluated both Xpert Ultra and Xpert MTB/RIF. Forty-two studies evaluated pulmonary TB using a reference standard of culture, and one study evaluated pulmonary TB using smear microscopy only.

In terms of methodological quality, in the patient selection domain, most studies (83%) evaluating pulmonary TB were judged to have low risk of bias. In the index test domain, all studies were judged to have low risk of bias. In the flow and timing domain, most studies (88%) were judged to have low risk of bias. In the reference standard domain, with respect to the MRS, 47% of studies were judged to have unclear risk of bias because only one culture was used to exclude TB. With respect to the composite reference standard, all studies were judged to have unclear risk of bias because of imperfect accuracy of the composite reference standard and differing definitions of this standard used by the primary study authors. Regarding applicability, in the patient selection domain, 50% of studies were judged as having high or unclear risk of bias, because participants were evaluated exclusively as inpatients at tertiary care centres, or the clinical setting was unclear. With respect to applicability of the index test, most studies (72%) were judged as having low concern owing to standardized application of the index tests. Eleven studies evaluating stool as a specimen for Xpert MTB/RIF or Xpert Ultra were judged to have unclear risk of bias because of the absence of a standardized protocol for stool preparation. Applicability of the reference standard was considered as a low concern for most studies (93%).

To generate evidence about the detection of rifampicin resistance, six studies were included. All of the six studies (223 participants) evaluated only Xpert MTB/RIF and were conducted in high TB burden countries and in high MDR-TB burden countries. Among the studies, 50% had a low risk of bias with respect to patient selection, while all studies had a low risk of bias with respect to the reference standard. Risk of bias was considered low for the reference standard if an automated process was used or it was clear that the reference standard results were interpreted without knowledge of the index tests. For all six studies, there were applicability concerns regarding patient selection because of enrolment exclusively from inpatient or tertiary centres.

For the meta-analysis, a total of 23 studies (6612 participants) evaluated sputum specimens; 14 studies (3468 participants) evaluated gastric specimens; four studies (1125 participants) evaluated nasopharyngeal specimens; and 11 studies (1592 participants) evaluated stool specimens – all of these studies evaluated Xpert MTB/RIF alone. Three studies (753 participants) evaluated both Xpert MTB/RIF and Xpert Ultra on frozen sputum specimens. One study (195 participants) evaluated both Xpert MTB/RIF and Xpert Ultra on nasopharyngeal specimens.

2.2. What is the diagnostic accuracy of Xpert Ultra for pulmonary TB and rifampicin resistance in children, as compared with MRS and CRS?

No studies evaluated Xpert Ultra alone. Three studies (753 participants) evaluated both Xpert MTB/RIF and Xpert Ultra on frozen sputum specimens. One study (195 participants) evaluated both Xpert MTB/RIF and Xpert Ultra on nasopharyngeal specimens.

Web Annex 4.4: Xpert MTB/RIF and Xpert Ultra for detecting active tuberculosis in children: an updated systematic review.

PICO 3: Among adults with signs and symptoms of extrapulmonary TB, seeking care at health care facilities, should Xpert MTB/RIF / Xpert Ultra be used as an initial test for diagnosis of extrapulmonary TB and rifampicin resistance?

3.1. What is the diagnostic accuracy of Xpert MTB/RIF for extrapulmonary TB and rifampicin resistance in adults, as compared with MRS and CRS?

There are difficulties in obtaining extrapulmonary specimens both from children and adults, and technical limitations of conventional bacteriological methods to aid diagnosis. Thus, various non-pulmonary specimens and composite reference standards are often used in evaluating the performance of new diagnostic technologies in extrapulmonary TB.

For detection of extrapulmonary TB, 65 studies were included. A total of 63 studies (13 144 participants) evaluated Xpert MTB/RIF, including five that evaluated both Xpert MTB/RIF and Xpert Ultra. The included studies evaluated Xpert MTB/RIF in cerebrospinal fluid (CSF) specimens comprising lymph node aspirate, lymph node biopsy, pleural fluid, urine, synovial fluid, peritoneal fluid, pericardial fluid and blood.

Of the total of 65 studies, 39 (60%) took place in high TB burden and 41 (63%) in high TB/HIV burden countries. Risk of bias was judged to be low in the domains of patient selection, index test, and flow and timing; and high or unclear in the reference standard domain because many studies decontaminated sterile specimens before culture inoculation. Regarding applicability, in the patient selection domain, high or unclear concern was expressed for most studies because either the participants were evaluated exclusively as inpatients at tertiary care centres, or the clinical settings were unclear.

Annex 4.3: Xpert MTB/RIF and Xpert Ultra for detecting active tuberculosis in adults with signs and symptoms of extrapulmonary TB: an updated systematic review.

3.2. What is the diagnostic accuracy of Xpert Ultra for extrapulmonary TB and rifampicin resistance in adults, as compared with MRS?

Six studies (507 participants) evaluated Xpert Ultra for the detection of extrapulmonary TB. The included studies evaluated the test in CSF specimens comprising lymph node biopsy, pleural fluid, urine and synovial fluid. Serious concerns were expressed regarding the indirectness of the evidence; these concerns related to applicability (i.e. evidence was generated in tertiary referral medical centres), and imprecision of the evidence, related mostly to low numbers of participants included in studies. Certainty of evidence was generally judged as being between low and very low.

Web Annex 4.3: Xpert MTB/RIF and Xpert Ultra for detecting active tuberculosis in adults with signs and symptoms of extrapulmonary TB: an updated systematic review.

PICO 4: Among children with signs and symptoms of extrapulmonary TB and rifampicin resistance, seeking care at health care facilities, should Xpert MTB/RIF / Xpert Ultra be used as an initial test for diagnosis of extrapulmonary TB and rifampicin resistance?

4.1. What is the diagnostic accuracy of Xpert MTB/RIF for extrapulmonary TB and rifampicin resistance in children, as compared with MRS and CRS?

4.2. What is the diagnostic accuracy of Xpert Ultra for extrapulmonary TB and rifampicin resistance in children, as compared with MRS and CRS?

To evaluate detection of extrapulmonary TB, studies that evaluated the diagnostic accuracy of Xpert MTB/RIF in children with signs or symptoms of lymph node TB or TB meningitis were included.

For diagnosis of lymph node TB, six studies (210 participants) evaluated Xpert MTB/RIF against an MRS of smear or culture on lymph node specimens. Two studies (105 participants) evaluated Xpert MTB/RIF against a composite reference standard for lymph node TB. For TB meningitis, six studies (241 participants) evaluated Xpert MTB/RIF against culture on CSF. In addition, two studies (155 participants) assessed Xpert MTB/RIF against a composite reference standard that included a clinical diagnosis of TB meningitis. The certainty of evidence was judged to be very low for sensitivity, and low for specificity of detection of both TB meningitis and lymph node TB.

No studies evaluating the accuracy of Xpert Ultra for detecting lymph node TB or TB meningitis were identified.

Web Annex 4.4: Xpert MTB/RIF and Xpert Ultra for detecting active tuberculosis in children: an updated systematic review.

PICO 5: Among people with signs and symptoms of pulmonary TB, seeking care at health care facilities, do repeated Xpert (Ultra) tests on subsequent samples as an initial test for diagnosis of pulmonary TB and rifampicin resistance increase sensitivity/specificity compared with a single initial test?

5.1. Xpert Ultra repeated test for the diagnosis of pulmonary TB in adults with signs and symptoms of pulmonary TB who have an initial Xpert Ultra trace result, as compared with MRS?

For adults, with initial Xpert Ultra trace results, three studies were identified: Mishra 2019a (4 participants), Piersimoni 2019 (4 participants), and Dorman 2018 (42 participants) (see **Web Annex 1** for details of included studies). Piersimoni 2019 retested the same initial sample, whereas Dorman 2018 retested a separately collected sputum sample. Mishra 2019a retested only those participants with discrepant results (i.e. Ultra trace positive/culture negative), and retested new specimens obtained a median of 444 days (range 245–526 days) after initial testing. Owing to limited data, a meta-analysis was not performed. The evidence was downgraded one level for inconsistency and two levels for imprecision. Serious concerns were expressed for inconsistency, and very serious concerns for imprecision. Certainty of evidence was judged to be very low for both sensitivity and specificity.

Web Annex 4.2: Xpert MTB/RIF and Xpert Ultra for detecting active tuberculosis in adults with signs and symptoms of pulmonary TB: an updated systematic review.

5.2. More than one Xpert MTB/RIF versus one Xpert MTB/RIF to diagnose pulmonary TB in children with signs and symptoms of pulmonary TB, as compared with MRS?

For children, five studies (2119 participants) were included that have evaluated the diagnostic accuracy of multiple Xpert MTB/RIF tests compared with a single test. Serious concerns were expressed for indirectness, because patients were enrolled from inpatient tertiary care settings, which could lead to the enrolment of children with more advanced disease. Also, serious concerns were expressed for imprecision, related to the low number of children with pulmonary TB contributing to this analysis for the observed sensitivity. Overall, the certainty of evidence was judged to be very low for sensitivity and moderate for specificity.

Web Annex 4.4: Xpert MTB/RIF and Xpert Ultra for detecting active tuberculosis in children: an updated systematic review.

5.3. More than one Xpert Ultra versus one Xpert Ultra to diagnose pulmonary TB in children with signs and symptoms of pulmonary TB, as compared with MRS?

For children, one study (163 participants) was included that evaluated the diagnostic accuracy of multiple Xpert Ultra tests in sputum compared with a single test. The certainty of evidence was judged to be very low for sensitivity and low for specificity owing to serious concerns for indirectness and imprecision. In addition, one study (130 participants) was included that evaluated the diagnostic accuracy of multiple Xpert Ultra tests in nasopharyngeal aspirates compared with a single test. Overall, the certainty of evidence was judged to be very low both for sensitivity and specificity, owing to very serious concerns for indirectness and imprecision.

Web Annex 4.4: Xpert MTB/RIF and Xpert Ultra for detecting active tuberculosis in children: an updated systematic review.

PICO 6: Among adults either with signs and symptoms of TB or chest radiograph with lung abnormalities suggestive of pulmonary TB or both, should Xpert MTB/RIF or Xpert Ultra alone be used to define a case of active TB disease *(10)*?

The aim of the review was to assess the diagnostic accuracy of Xpert MTB/RIF and Xpert Ultra for pulmonary TB in adults (aged ≥15 years) among the general population. Data from four nationally representative and two subnational prevalence surveys for active TB disease, cross-sectional in design, were included. These surveys used sputum samples that evaluated Xpert MTB/RIF or Xpert Ultra against the reference standard of culture for TB. For the evaluation of TB detection, the surveys evaluated the index tests in adults (aged ≥15 years) with chest X-ray abnormalities or symptoms suggestive of pulmonary TB (or both). For detection of pulmonary TB, a total of six surveys were identified.

6.1. Xpert MTB/RIF to diagnose pulmonary TB in adults in the general population with signs and symptoms of pulmonary TB or chest radiograph with lung abnormalities or both, as compared with MRS?

The analysis reported on the results of four surveys, including 49 556 participants. Assessment of the quality of the evidence revealed serious deficiencies in the evidence quality.

Indirectness: the populations in these prevalence surveys differed from the general population with respect to prior testing (e.g. symptom screen was limited to cough for 14 days or more) and the availability of results of both symptom screen and chest radiography in most participants included in the studies. The evidence was downgraded one level for indirectness.

Inconsistency: the sensitivity estimate for Bangladesh was 84%, which was higher than the sensitivity estimates for the other three countries (range, 68–69%). Lower HIV prevalence in Bangladesh could only partly explain the inconsistency. The evidence was downgraded one level for inconsistency. Overall, the certainty of evidence was judged to be low for sensitivity and moderate for specificity.

6.2. Xpert Ultra to diagnose pulmonary TB in adults in the general population with signs and symptoms of pulmonary TB or chest radiograph with lung abnormalities or both, as compared with MRS.

The analysis reported on the results of four surveys, including 11 488 participants. The included countries were Myanmar, South Africa (TREATS project) and Zambia (TREATS project). The average prevalence of TB in these countries was 2.8% (range 1.6–6.7%).

Indirectness: the populations in these prevalence surveys differed from the general population with respect to prior testing (e.g. symptom screen was limited to cough for 14 days or more) and the availability of results of both symptom screen and chest radiography in most participants included in the studies. The evidence was downgraded one level for indirectness.

Imprecision: there were relatively few participants contributing to this analysis, and a wide 95% confidence interval (CI). The 95% CI around true positives and false negatives may lead to different decisions, depending on which limits are assumed. The evidence was downgraded one level for imprecision. Overall, the certainty of evidence was judged to be low for sensitivity and moderate for specificity.

6.3. Two Xpert Ultra versus one Xpert Ultra to diagnose pulmonary TB in adults in the general population with signs and symptoms of TB or chest radiograph with lung abnormalities or both, as compared with MRS.

The analysis reported on the results of three surveys, including 5080 participants. Serious concerns were expressed about the indirectness of the available evidence. This was because most of the data were from Myanmar, and the results may not be applicable to other settings. In addition, very serious concerns were expressed about imprecision because the analysis was based on data for only a small number of individuals. The 95% CIs for two Xpert Ultra assays and one Xpert Ultra assay were wide. Overall, the certainty of evidence was judged to be very low for sensitivity and moderate for specificity.

PICO 7: Among people with signs and symptoms of pulmonary TB, seeking care at health care facilities, should Molbio Truenat MTB, MTB Plus and MTB-RIF Dx be used as an initial test for diagnosis of pulmonary TB and rifampicin resistance?

7.1. What is the diagnostic accuracy of Truenat MTB to diagnose pulmonary TB in adults with signs and symptoms of pulmonary TB, as compared with MRS?

Evidence for the use of Truenat MTB, MTB Plus and MTB-RIF Dx assays to diagnose pulmonary TB and rifampicin resistance in adults was generated through a multicentre prospective clinical evaluation study implemented by FIND. The study was conducted at 19 clinical sites (each with a microscopy centre attached) and seven reference laboratories in four countries. The aim was to determine the diagnostic accuracy of the Truenat assays when performed in the intended settings of use (i.e. microscopy centres), relative to microbiological confirmation (culture) as the reference standard. The performance of the Truenat assays was also compared head-to-head (on the same specimens) to Xpert or Ultra in reference laboratories, as part of this assessment. All sites performed Xpert except for sites in Peru, which performed Ultra. The analysis for Truenat MTB reported on the results for 1336 participants. Serious concerns were expressed for imprecision and inconsistency of evidence related to sensitivity. Overall, the certainty of evidence was judged to be low for sensitivity but high for specificity.

7.2. What is the diagnostic accuracy of Truenat MTB Plus to diagnose pulmonary TB in adults with signs and symptoms of pulmonary TB, as compared with MRS?

The analysis for Truenat MTB Plus reported on the results for 1336 participants. Serious concerns were expressed for imprecision for sensitivity, related to the few participants contributing to the analysis. Overall, the certainty of evidence was judged to be low for sensitivity and high for specificity.

7.3. What is the diagnostic accuracy of Truenat MTB-RIF Dx to diagnose rifampicin resistance in adults with signs and symptoms of pulmonary TB, as compared with MRS?

The analysis for Truenat MTB-RIF Dx reported on the results for 186 participants. For sensitivity there were serious concerns about indirectness (India and Peru contributed most of the data to the determination of rifampicin resistance) and inconsistency (variable sensitivity estimates: 100% for Peru, based on seven rifampicin-resistant specimens; 100% for Ethiopia, based on one rifampicin-resistant specimen; 100% for Papua New Guinea, based on one rifampicin-resistant specimen; and 81% for India, based on 42 rifampicin-resistant specimens). These results may not be applicable to other settings. In addition, very serious concerns were expressed for imprecision, owing to the small number of participants contributing to this analysis. Overall, the certainty of evidence was judged to be very low for sensitivity. Serious concerns were expressed for indirectness for specificity, related to the low numbers of rifampicin-resistant cases and the fact that most of them were from India and Peru.

Web Annex 4.5: Report on the diagnostic accuracy of the Molbio Truenat tuberculosis and rifampicin-resistance assays in the intended setting of use.

1.4 Performance of the molecular assays

Table 1.1. PICO 1.1: What is the impact of Xpert MTB/RIF on patient-important outcomes (e.g. cure, mortality, time to diagnosis and time to start treatment)?

Patient-important outcome	Studies/design	Certainty of evidence	Patients with outcome of interest/all patients		Effect	
			Xpert MTB/RIF	Smear microscopy	Relative	Absolute
Mortality	5/RT	Moderate	248/5265 (4.7%)	292/5144 (5.7%)	RR 0.88	7 fewer per 1000
Cure	2/RT	High	1786/2500 (71.4%)	1443/2080 (69.4%)	OR 1.09	18 more per 1000
Pretreatment loss to follow-up	3/RT	Moderate	81/642 (12.6%)	95/523 (18.2%)	RR 0.59	74 fewer per 1000
Time to diagnosis	2/RT	High	956	968 (10%)	HR 1.05	5 more per 1000
Treatment	4/RT	Moderate	4055	4153 (10%)	HR 1.00	0 fewer per 1000
Mortality in people with HIV	2/RT	Moderate	66/1211 (5.5%)	75/1055 (7.1%)	RR 0.76	17 fewer per 1000

HIV: human immunodeficiency virus; HR: hazard ratio; OR: odds ratio; PICO: population, intervention, comparator and outcomes; RR: relative risk; RT: randomized trial.

Table 1.2. PICO 1.2: What is the diagnostic accuracy of Xpert MTB/RIF for pulmonary TB in adults, as compared with MRS?

Patient population	Test accuracy	Studies (participants)	Certainty of evidence	2.5% prevalence	10% prevalence	30% prevalence
Adults PTB, MRS	Se: 0.85	70 (10 409)	High	TP: 21 / FN: 4	TP: 85 / FN: 15	TP: 255 / FN: 45
	Sp: 0.98	70 (26 828)	High	TN: 965 / FP: 10	TN: 891 / FP: 9	TN: 693 / FP: 7
Adults PTB, SS–, MRS	Se: 0.67	45 (2315)	High	TP: 17 / FN: 8	TP: 67 / FN: 33	TP: 201 / FN: 99
	Sp: 0.98	45 (16 647)	High	TN: 956 / FP: 19	TN: 882 / FP: 18	TN: 686 / FP: 14
Adults PTB, HIV+, MRS	Se: 0.81	14 (1159)	High	TP: 20 / FN: 5	TP: 81 / FN: 19	TP: 243 / FN: 57
	Sp: 0.98	14 (3505)	High	TN: 956 / FP: 19	TN: 882 / FP: 18	TN: 686 / FP: 14
Adults PTB, previous TB, MRS	Se: 0.86	14 (2197)	Low	TP: 22 / FN: 3	TP: 86 / FN: 14	TP: 258 / FN: 42
	Sp: 0.95	14 (2998)	Moderate	TN: 924 / FP: 51	TN: 853 / FP: 47	TN: 664 / FP: 36

FN: false negative; FP: false positive; HIV+: human immunodeficiency virus positive; MRS: microbiological reference standard; PICO: population, intervention, comparator and outcomes; PTB: pulmonary tuberculosis; Se: sensitivity; Sp: specificity; SS–: sputum smear negative; TB: tuberculosis; TN: true negative; TP: true positive.

Table 1.3. PICO 1.2: What is the diagnostic accuracy of Xpert MTB/RIF for rifampicin resistance in adults with pulmonary TB, as compared with MRS?

Patient population	Test accuracy	Studies (participants)	Certainty of evidence	2% prevalence	10% prevalence	15% prevalence
Adults PTB, RR-TB	Se: 0.96	48 (1775)	High	TP: 19 / FN: 1	TP: 96 / FN: 4	TP: 144 / FN: 6
	Sp: 0.98	48 (6245)	High	TN: 960 / FP: 20	TN: 882 / FP: 18	TN: 833 / FP: 17

FN: false negative; FP: false positive; MRS: microbiological reference standard; PICO: population, intervention, comparator and outcomes; PTB: pulmonary tuberculosis; RR-TB: rifampicin-resistant tuberculosis; Se: sensitivity; Sp: specificity; TB: tuberculosis; TN: true negative; TP: true positive.

Table 1.4. PICO 1.3: What is the diagnostic accuracy of Xpert Ultra for pulmonary TB, as compared with MRS?

Patient population	Test accuracy	Studies (participants)	Certainty of evidence	2.5% prevalence	10% prevalence	30% prevalence
Adults PTB, MRS	Se: 0.90	6 (960)	High	TP: 22 / FN: 3	TP: 90 / FN: 10	TP: 269 / FN: 31
	Sp: 0.96	6 (1694)	High	TN: 932 / FP: 43	TN: 860 / FP: 40	TN: 669 / FP: 31
Adults PTB, SS–, MRS	Se: 0.77	6 (378)	High	TP: 19 / FN: 6	TP: 77 / FN: 23	TP: 231 / FN: 69
	Sp: 0.96	6 (1671)	High	TN: 932 / FP: 43	TN: 860 / FP: 40	TN: 669 / FP: 31
Adults PTB, HIV+, MRS	Se: 0.88	2 (149)	Low	TP: 22 / FN: 3	TP: 88 / FN: 12	TP: 265 / FN: 35
	Sp: 0.95	2 (430)	High	TN: 923 / FP: 52	TN: 852 / FP: 48	TN: 663 / FP: 37
Adults PTB, prior TB, MRS	Se: 0.84	4 (127)	Low	TP: 21 / FN: 4	TP: 84 / FN: 16	TP: 251 / FN: 49
	Sp: 0.86	4 (475)	Low	TN: 842 / FP: 133	TN: 778 / FP: 122	TN: 605 / FP: 95

FN: false negative; FP: false positive; HIV+: human immunodeficiency virus positive; MRS: microbiological reference standard; PICO: population, intervention, comparator and outcomes; PTB: pulmonary tuberculosis; Se: sensitivity; Sp: specificity; SS–: sputum smear negative; TB: tuberculosis; TN: true negative; TP: true positive.

Table 1.5. PICO 1.3: What is the diagnostic accuracy of Xpert Ultra for rifampicin resistance in adults with pulmonary TB, as compared with MRS?

Patient population	Test accuracy	Studies (persons)	Certainty of evidence	2% prevalence	10% prevalence	15% prevalence
Adults PTB, RR-TB	Se: 0.94	5 (240)	High	TP: 19 / FN: 1	TP: 94 / FN: 6	TP: 141 / FN: 9
	Sp: 0.99	5 (690)	High	TN: 970 / FP: 10	TN: 891 / FP: 9	TN: 842 / FP: 8

FN: false negative; FP: false positive; MRS: microbiological reference standard; PICO: population, intervention, comparator and outcomes; PTB: pulmonary tuberculosis; RR-TB: rifampicin-resistant tuberculosis; Se: sensitivity; Sp: specificity; TB: tuberculosis; TN: true negative; TP: true positive.

Table 1.6. PICO 2.1: What is the diagnostic accuracy of Xpert MTB/RIF for pulmonary TB in children, as compared with MRS and CRS?

Patient population	Test accuracy	Studies (persons)	Certainty of evidence	1% prevalence	10% prevalence	20% prevalence
Children sputum, MRS	Se: 0.65	23 (493)	Moderate	TP: 6 / FN: 4	TP: 65 / FN: 35	TP: 129 / FN: 71
	Sp: 0.99	23 (6119)	Moderate	TN: 980 / FP: 10	TN: 891 / FP: 9	TN: 792 / FP: 8
Children sputum, CRS	Se: 0.20	16 (1541)	Low	TP: 2 / FN:8	TP: 20 / FN: 80	TP: 40 / FN: 160
	Sp: 1.00	16 (2838)	Moderate	TN: 990 / FP: 0	TN: 900 / FP: 0	TN: 800 / FP: 0
Children SS–, sputum, MRS	Se: 0.59	12 (184)	Low	TP: 6 / FN: 4	TP: 59 / FN: 41	TP: 118 / FN: 82
	Sp: 0.99	12 (2934)	Moderate	TN: 980 / FP: 10	TN: 891 / FP: 9	TN: 792 / FP: 8
Children HIV+, sputum, MRS	Se: 0.72	10 (88)	Low	TP: 7 / FN: 3	TP: 72 / FN: 28	TP: 144 / FN: 56
	Sp: 0.99	10 (554)	Moderate	TN: 980 / FP: 10	TN: 891 / FP: 9	TN: 792 / FP: 8
Children GA, MRS	Se: 0.73	14 (272)	Very Low	TP: 7 / FN: 3	TP: 73 / FN:27	TP: 146 / FN: 54
	Sp: 0.98	14 (3311)	Low	TN: 971 / FP: 19	TN: 883 / FP: 17	TN: 785 / FP: 15
Children GA, CRS	Se: 0.32	6 (461)	Very Low	TP: 3 / FN: 7	TP: 32 / FN: 68	TP: 64 / FN: 136
	Sp: 0.99	6 (472)	Moderate	TN: 980 / FP: 10	TN: 891 / FP: 9	TN: 792 / FP: 8
Children HIV+, GA, MRS	Se: 0.73	3 (50)	Low	TP: 7 / FN: 3	TP: 73 / FN: 27	TP: 146 / FN: 54
	Sp: 0.99	3 (584)	Moderate	TN: 980 / FP: 10	TN: 891 / FP: 9	TN: 792 / FP: 8
Children NFA, MRS	Se: 0.46	4 (144)	Moderate	TP: 5 / FN: 5	TP: 46 / FN: 54	TP: 92 / FN: 108
	Sp: 1.00	4 (981)	High	TN: 990 / FP: 0	TN: 900 / FP: 0	TN: 800 / FP: 0

Patient population	Test accuracy	Studies (persons)	Certainty of evidence	1% prevalence	10% prevalence	20% prevalence
Children stool, MRS	Se: 0.61	11 (174)	Low	TP: 6 / FN: 4	TP: 62 / FN: 38	TP: 123 / FN: 77
	Sp: 0.98	11 (1418)	Moderate	TN: 975 / FP: 15	TN: 887 / FP: 13	TN: 788 / FP: 12
Children stool, CRS	Se: 0.16	10 (879)	Low	TP: 2 / FN: 8	TP: 16 / FN: 84	TP: 32 / FN: 168
	Sp: 0.99	10 (860)	Moderate	TN: 980 / FP: 10	TN: 891 / FP: 9	TN: 792 / FP: 8
Children HIV+, stool, MRS	Se: 0.70	4 (53)	Low	TP: 7 / FN: 3	TP: 70 / FN: 30	TP: 140 / FN: 60
	Sp: 0.98	4 (473)	High	TN: 970 / FP: 20	TN: 882 / FP: 18	TN: 784 / FP: 16

CRS: composite reference standard; FN: false negative; FP: false positive; GA: gastric aspirate; HIV+: human immunodeficiency virus positive; MRS: microbiological reference standard; NFA: nasopharyngeal aspirate; PICO: population, intervention, comparator and outcomes; Se: sensitivity; Sp: specificity; TB: tuberculosis; TN: true negative; TP: true positive.

Table 1.7. PICO 2.1: What is the diagnostic accuracy of Xpert MTB/RIF for rifampicin resistance in children, as compared with MRS?

Patient population	Test accuracy	Studies (persons)	Certainty of evidence	2% prevalence	10% prevalence	15% prevalence
Children sputum, RR, MRS	Se: 0.90	6 (20)	Very low	TP: 18 / FN: 2	TP: 90 / FN: 10	TP: 135 / FN: 15
	Sp: 0.98	6 (203)	Moderate	TN: 960 / FP: 20	TN: 882 / FP: 18	TN: 833 / FP: 17

FN: false negative; FP: false positive; MRS: microbiological reference standard; PICO: population, intervention, comparator and outcomes; RR: rifampicin resistance; Se: sensitivity; Sp: specificity; TN: true negative; TP: true positive.

Table 1.8. PICO 2.2: What is the diagnostic accuracy of Xpert Ultra for pulmonary TB in children, as compared with MRS and CRS?

Patient population	Test accuracy	Studies (persons)	Certainty of evidence	1% prevalence	10% prevalence	20% prevalence
Children sputum, MRS	Se: 0.73	3 (136)	Low	TP: 7 / FN: 3	TP: 73 / FN: 27	TP: 146 / FN: 54
	Sp: 0.97	3 (551)	High	TN: 960 / FP: 30	TN: 873 / FP: 27	TN: 776 / FP: 24
Children sputum, CRS	Se: 0.24	3 (498)	Low	TP: 2 / FN: 8	TP: 24 / FN: 76	TP: 48 / FN: 152
	Sp: 0.97	3 (255)	Low	TN: 965 / FP: 25	TN: 878 / FP: 22	TN: 780 / FP: 20
Children NFA, MRS	Se: 0.46	1 (35)	Very low	TP: 5 / FN: 5	TP: 46 / FN: 54	TP: 92 / FN: 108
	Sp: 0.98	1 (160)	Low	TN: 970 / FP: 20	TN: 882 / FP: 18	TN: 784 / FP: 16

CRS: composite reference standard; FN: false negative; FP: false positive; MRS: microbiological reference standard; NFA: nasopharyngeal aspirate; PICO: population, intervention, comparator and outcomes; Se: sensitivity; Sp: specificity; TB: tuberculosis; TN: true negative; TP: true positive.

Table 1.9. PICO 3.1: What is the diagnostic accuracy of Xpert MTB/RIF for extrapulmonary TB in adults, as compared with MRS and CRS?

Patient population	Test accuracy	Studies (persons)	Certainty of evidence	2.5% prevalence	10% prevalence	20% prevalence
Adults CSF, MRS	Se: 0.70	28 (521)	Moderate	TP: 18 / FN: 7	TP: 70 / FN: 30	TP: 141 / FN: 59
	Sp: 0.97	28 (2582)	High	TN: 944 / FP: 31	TN: 871 / FP: 29	TN: 774 / FP: 26
Adults CSF, CRS	Se: 0.41	12 (774)	Low	TP: 10 / FN: 15	TP: 41 / FN:59	TP: 81 / FN: 119
	Sp: 0.99	12 (1123)	Moderate	TN: 970 / FP: 5	TN: 896 / FP: 4	TN: 796 / FP: 4
Adults LNA, MRS	Se: 0.89	14 (627)	Moderate	TP: 22 / FN: 3	TP: 89 / FN:11	TP: 177 / FN: 23
	Sp: 0.86	14 (961)	Very low	TN: 839 / FP: 136	TN: 774 / FP:126	TN: 688 / FP:112

Patient population	Test accuracy	Studies (persons)	Certainty of evidence	2.5% prevalence	10% prevalence	20% prevalence
Adults LNA, CRS	Se: 0.81	4 (377)	Low	TP: 20 / FN: 5	TP: 81 / FN: 19	TP: 162 / FN: 38
	Sp: 0.96	4 (302)	Low	TN: 935 / FP: 40	TN: 863 / FP: 37	TN: 767 / FP:33
Adults LNB, MRS	Se: 0.82	11 (220)	Low	TP: 21 / FN: 4	TP: 82 / FN: 18	TP: 164 / FN: 36
	Sp: 0.79	11 (566)	Very low	TN: 773 / FP: 202	TN: 714 / FP:186	TN: 634 / FP:166
Adults, pleural fluid, MRS	Se: 0.50	24 (589)	Very low	TP: 12 / FN: 13	TP: 50 / FN: 50	TP: 99 / FN: 101
	Sp: 0.99	24 (2337)	High	TN: 962 / FP: 13	TN: 888 / FP: 12	TN: 790 / FP: 10
Adults, pleural fluid, CRS	Se: 0.19	10 (616)	Moderate	TP: 5 / FN: 20	TP: 19 / FN: 81	TP: 39 / FN: 161
	Sp: 0.99	10 (408)	High	TN: 964 / FP: 11	TN: 890 / FP: 10	TN: 791 / FP: 9
Adults, peritoneal fluid, MRS	Se: 0.59	13 (94)	Low	TP: 15 / FN: 10	TP: 59 / FN: 41	TP: 118 / FN: 82
	Sp: 0.97	13 (486)	High	TN: 949 / FP: 26	TN: 876 / FP: 24	TN: 778 / FP: 22
Adults, pericardial fluid, MRS	Se: 0.60	5 (57)	Very low	TP: 15 / FN: 10	TP: 60 / FN:40	TP: 121 / FN: 79
	Sp: 0.88	5 (124)	Low	TN: 856 / FP: 119	TN: 790 / FP:110	TN: 702 / FP: 98
Adults, pericardial fluid, CRS	Se: 0.66	2 (60)	Very low	TP: 16 / FN: 9	TP: 66 / FN: 34	TP: 132 / FN: 68
	Sp: 0.96	2 (17)	Very low	TN: 936 / FP:39	TN: 864 / FP: 36	TN: 768 / FP: 32
Adults, urine, MRS	Se: 0.85	9 (72)	Low	TP: 21 / FN: 4	TP: 85 / FN: 15	TP: 169 / FN: 31
	Sp: 0.97	9 (871)	Moderate	TN: 949 / FP: 26	TN: 876 / FP: 24	TN: 778 / FP: 22
Adults, synovial fluid, MRS	Se: 0.97	6 (110)	Moderate	TP: 24 / FN: 1	TP: 97 / FN: 3	TP: 194 / FN: 6
	Sp: 0.94	6 (361)	Very low	TN: 914 / FP: 61	TN: 843 / FP: 57	TN: 750 / FP: 50
Adults, synovial fluid, CRS	Se: 0.88	2 (161)	Low	TP: 22 / FN: 3	TP: 88 / FN: 12	TP: 177 / FN: 23
	Sp: 0.98	2 (44)	Very low	TN: 955 / FP: 20	TN: 881 / FP: 19	TN: 783 / FP: 17
Adults HIV+, blood, MRS	Se: 0.56	1 (9)	Very low	TP: 14 / FN: 11	TP: 56 / FN: 44	TP: 112 / FN: 88
	Sp: 0.94	1 (65)	Very low	TN: 917 / FP: 58	TN: 846 / FP: 54	TN: 752 / FP: 48

CRS: composite reference standard; CSF: cerebrospinal fluid; FN: false negative; FP: false positive; HIV+: human immunodeficiency virus positive; LNA: lymph node aspirate; LNB: lymph node biopsy; MRS: microbiological reference standard; PICO: population, intervention, comparator and outcomes; Se: sensitivity; Sp: specificity; TB: tuberculosis; TN: true negative; TP: true positive.

Table 1.10. PICO 3.1: What is the diagnostic accuracy of Xpert MTB/RIF for rifampicin resistance in adults with extrapulmonary TB, as compared with MRS?

Patient population	Test accuracy	Studies (persons)	Certainty of evidence	2% prevalence	10% prevalence	15% prevalence
Adults, RR, MRS	Se: 0.96	23 (165)	High	TP: 19 / FN: 1	TP: 96 / FN: 4	TP: 144 / FN: 6
	Sp: 0.99	23 (919)	High	TN: 969 / FP: 11	TN: 890 / FP: 10	TN: 841 / FP: 9

FN: false negative; FP: false positive; MRS: microbiological reference standard; PICO: population, intervention, comparator and outcomes; RR: rifampicin resistance; Se: sensitivity; Sp: specificity; TB: tuberculosis; TN: true negative; TP: true positive.

Table 1.11. PICO 3.2: What is the diagnostic accuracy of Xpert Ultra for extrapulmonary TB in adults, as compared with MRS and CRS?

Patient population	Test accuracy	Studies (persons)	Certainty of evidence	2.5% prevalence	10% prevalence	20% prevalence
Adults CSF, MRS	Se: 0.87	4 (40)	Low	TP: 22 / FN: 3	TP: 87 / FN: 13	TP: 174 / FN: 26
	Sp: 0.88	4 (143)	Low	TN: 855 / FP: 120	TN: 789 / FP:111	TN: 702 / FP: 98
Adults LNA, MRS	Se: 0.78	1 (9)	Very low	TP: 20 / FN: 5	TP: 78 / FN: 22	TP: 156 / FN: 44
	Sp: 0.78	1 (64)	Very low	TN: 761 / FP: 214	TN: 702 / FP:198	TN: 624 / FP: 176
Adults LNA, CRS	Se: 0.70	1 (30)	Very low	TP: 17 / FN: 8	TP: 70 / FN: 22	TP: 156 / FN: 44
	Sp: 1.00	1 (43)	Low	TN: 975 / FP: 0	TN: 702 / FP:198	TN: 624 / FP: 176
Adults LNB, MRS	Se: 0.90–1.00	2 (23)	Very low	TP: 23–25 / FN: 0–2	TP: 90–100 / FN: 0–10	TP: 180–200 / FN: 0–20
	Sp: 0.38–0.87	2 (108)	Very low	TN: 371–848 / FP: 127–604	TN: 342–783 / FP: 117–558	TN: 304–696 / FP: 104–496

Patient population	Test accuracy	Studies (persons)	Certainty of evidence	2.5% prevalence	10% prevalence	20% prevalence
Adults LNB, CRS	Se: 0.67	1 (22)	Very low	TP: 18 / FN: 7	TP: 73 / FN: 27	TP: 146 / FN: 54
	Sp: 0.96	1 (57)	Very low	TN: 936 / FP: 39	TN: 864 / FP:36	TN: 768 / FP: 32
Adults pleural fluid, MRS	Se: 0.71	3 (101)	Very low	TP: 18 / FN: 7	TP: 71 / FN: 29	TP: 142 / FN: 58
	Sp: 0.71	3 (156)	Very low	TN: 694 / FP: 281	TN: 641 / FP:259	TN: 570 / FP: 230
Adults pleural fluid, CRS	Se: 0.38–0.61	2 (156)	Very low	TP: 10–15 / FN: 10–15	TP: 38–61 / FN: 39–62	TP: 76–122 / FN: 78–122
	Sp: 0.96–0.99	2 (107)	Moderate	TN: 936–965 / FP: 10–39	TN: 864–891 / FP:9–36	TN: 768–792 / FP: 8–32
Adults synovial fluid, MRS	Se: 0.96	1 (52)	Very low	TP: 24 / FN: 1	TP: 96 / FN: 4	TP: 192 / FN: 8
	Sp: 0.97	1 (34)	Very low	TN: 946 / FP: 29	TN: 873 / FP: 27	TN: 776 / FP: 24
Adults synovial fluid, CRS	Se: 0.96	1 (111)	Low	TP: 24 / FN: 1	TP: 96 / FN: 4	TP: 192 / FN: 8
	Sp: 0.97	1 (34)	Very low	TN: 946 / FP: 29	TN: 873 / FP: 27	TN: 776 / FP: 24
Adults urine, MRS	Se: 1.00	1 (12)	Very low	TP: 25 / FN: 0	TP: 100 / FN: 0	TP: 200 / FN: 0
	Sp: 1.00	1 (12)	Very low	TN: 975 / FP: 0	TN: 900 / FP: 0	TN: 800 / FP: 0

CRS: composite reference standard; CSF: cerebrospinal fluid; FN: false negative; FP: false positive; LNA: lymph node aspirate; LNB: lymph node biopsy; MRS: microbiological reference standard; PICO: population, intervention, comparator and outcomes; Se: sensitivity; Sp: specificity; TB: tuberculosis; TN: true negative; TP: true positive.

Table 1.12. PICO 3.2: What is the diagnostic accuracy of Xpert Ultra for rifampicin resistance in adults with extrapulmonary TB, as compared with MRS and CRS?

Patient population	Test accuracy	Studies (persons)	Certainty of evidence	2% prevalence	10% prevalence	15% prevalence
Adults, RR, MRS	Se: 0.97	3 (19)	Low	TP: 19 / FN: 1	TP: 97 / FN: 3	TP: 145 / FN: 5
	Sp: 0.99	3 (84)	Moderate	TN: 968 / FP: 12	TN: 889 / FP: 11	TN: 840 / FP: 10

CRS: composite reference standard; FN: false negative; FP: false positive; MRS: microbiological reference standard; PICO: population, intervention, comparator and outcomes; RR: rifampicin resistance; Se: sensitivity; Sp: specificity; TB: tuberculosis; TN: true negative; TP: true positive.

Table 1.13. PICO 4.1: What is the diagnostic accuracy of Xpert MTB/RIF for extrapulmonary TB in children, as compared with MRS?

Patient population	Test accuracy	Studies (persons)	Certainty of evidence	1% prevalence	5% prevalence	10% prevalence
Children, CSF, MRS	Se: 0.54	6 (28)	Very low	TP: 5 / FN: 5	TP: 27 / FN: 23	TP: 54 / FN: 46
	Sp: 0.94	6 (213)	Low	TN: 929 / FP: 61	TN: 891 / FP: 59	TN: 844 / FP: 56

CSF: cerebrospinal fluid; FN: false negative; FP: false positive; MRS: microbiological reference standard; PICO: population, intervention, comparator and outcomes; Se: sensitivity; Sp: specificity; TB: tuberculosis; TN: true negative; TP: true positive.

Table 1.14. PICO 5.1: Xpert Ultra repeated test for the diagnosis of pulmonary TB in adults with signs and symptoms of pulmonary TB who have an initial Ultra trace result, as compared with MRS?

Patient population	Test accuracy	Studies (persons)	Certainty of evidence	2.5% prevalence	10% prevalence	30% prevalence
Repeated Ultra for PTB in adults with initial trace result, MRS	Se: 0.69–1.00	3 (15)	Very low	TP: 17–25 / FN: 0–8	TP: 69–100 / FN: 0–31	TP: 207–300 / FN: 0–93
	Sp: 0.47–1.00	3 (25)	Very low	TN: 458–975 / FP: 0–571	TN: 423–900 / FP: 0–477	TN: 329–700 / FP: 0–371

FN: false negative; FP: false positive; MRS: microbiological reference standard; PICO: population, intervention, comparator and outcomes; PTB: pulmonary tuberculosis; Se: sensitivity; Sp: specificity; TB: tuberculosis; TN: true negative; TP: true positive.

Table 1.15. PICO 5.2: More than one Xpert MTB/RIF versus one Xpert MTB/RIF to diagnose pulmonary TB in children with signs and symptoms of pulmonary TB, as compared with MRS?

Patient population	Studies (persons)	Test accuracy (1 MTB/RIF)	Test accuracy (>1 MTB/RIF)	Certainty in evidence	1% prevalence		10% prevalence		20% prevalence	
					1 MTB/RIF	>1 MTB/RIF	1 MTB/RIF	>1 MTB/RIF	1 MTB/RIF	>1 MTB/RIF
1 versus 1+ MTB/RIF for PTB in sputum children, MRS	5 (180)	Se: 0.46	Se: 0.59	Low	TP: 5 FN: 5	TP: 6 FN: 4	TP: 46 FN: 54	TP: 59 FN: 41	TP: 92 FN: 108	TP: 118 FN: 82
	5 (1939)	Sp: 1.00	Sp: 0.99	High	TN: 989 FP: 1	TN: 980 FP: 10	TN: 899 FP: 1	TN: 891 FP: 9	TN: 799 FP: 1	TN: 792 FP: 8
1 versus 1+ MBT/RIF for PTB in GA in children, MRS	1 (32)	Se: 0.09	Se: 0.23	Very low	TP: 1 FN: 9	TP: 2 FN: 8	TP: 9 FN: 91	TP: 23 FN: 77	TP: 19 FN: 181	TP: 46 FN: 154
	1 (903)	Sp: 0.99	Sp: 0.99	Low	TN: 980 FP: 10	TN: 980 FP: 10	TN: 891 FP: 9	TN: 891 FP: 9	TN: 792 FP: 8	TN: 792 FP: 8
1 versus 1+ MBT/RIF for PTB in NPA in children, MRS	2 (91)	Se: 0.41	Se: 0.54	Very low	TP: 4 FN: 6	TP: 5 FN: 5	TP: 41 FN: 59	TP: 54 FN: 46	TP: 82 FN: 118	TP: 108 FN: 92
	2 (614)	Sp: 0.99	Sp: 0.98	Moderate	TN: 980 FP: 10	TN: 970 FP: 20	TN:891 FP: 9	TN: 882 FP: 18	TN:792 FP: 8	TN: 784 FP: 16
1 versus 1+ MTB/RIF for PTB in stool in children, MRS	1 (17)	Se: 0.25	Se: 0.35	Low	TP: 3 FN: 7	TP: 3 FN: 7	TP: 25 FN: 75	TP: 35 FN: 65	TP: 50 FN: 150	TP: 70 FN: 130
	1 (230)	Sp: 0.99	Sp: 0.99	Low	TN: 980 FP: 10	TN: 980 FP: 10	TN: 891 FP: 9	TN: 891 FP: 9	TN: 792 FP: 8	TN: 792 FP: 8

FN: false negative; FP: false positive; GA: gastric aspirate; MRS: microbiological reference standard; NPA: nasopharyngeal aspirate; PICO: population, intervention, comparator and outcomes; Se: sensitivity; Sp: specificity; TB: tuberculosis; TN: true negative; TP: true positive.

Table 1.16. PICO 5.3: More than one Xpert Ultra versus one Xpert Ultra to diagnose pulmonary TB in children with signs and symptoms of pulmonary TB, as compared with MRS?

					1% prevalence		10% prevalence		20% prevalence	
Patient population	**Studies (persons)**	**Test accuracy (1 Ultra)**	**Test accuracy (>1 Ultra)**	**Certainty in evidence**	**1 Ultra**	**>1 Ultra**	**1 Ultra**	**>1 Ultra**	**1 Ultra**	**>1 Ultra**
1 versus 1+ Ultra for PTB in children, MRS	1 (28)	Se: 0.64	Se: 0.75	Very low	TP: 6 FN: 4	TP: 8 FN: 2	TP: 64 FN: 36	TP: 75 FN: 25	TP: 128 FN: 72	TP: 150 FN: 50
	1 (135)	Sp: 1.0	Sp: 0.98	Very low	TN: 990 FP: 0	TN: 970 FP: 20	TN: 900 FP: 0	TN: 882 FP: 18	TN: 800 FP: 0	TN: 784 FP: 16
1 versus 1+ Ultra for PTB in NPA in children, MRS	1 (24)	Se: 0.38	Se: 0.54	Very low	TP: 4 FN: 6	TP: 5 FN: 5	TP: 38 FN: 62	TP: 54 FN: 46	TP: 76 FN: 124	TP: 108 FN: 92
	1 (106)	Sp: 0.98	Sp: 0.96	Low	TN: 970 FP: 20	TN: 950 FP: 40	TN: 882 FP: 18	TN: 864 FP: 36	TN: 784 FP: 16	TN: 768 FP: 32

FN: false negative; FP: false positive; MRS: microbiological reference standard; NPA: nasopharyngeal aspirate; PICO: population, intervention, comparator and outcomes; PTB: pulmonary tuberculosis; Se: sensitivity; Sp: specificity; TB: tuberculosis; TN: true negative; TP: true positive.

Table 1.17. PICO 6.1–6.2: Among adults in the general population with signs and symptoms of pulmonary TB or chest radiograph with lung abnormalities or both, should Xpert MTB/RIF or Xpert Ultra alone be used to define a case of active TB disease, as compared with MRS?

Patient population	Test accuracy	Studies (persons)	Certainty in evidence	1% prevalence	3% prevalence	7% prevalence
Xpert MTB/RIF in adults for PTB, MRS	Se: 0.73	4 (867)	Low	TP: 7 / FN: 3	TP: 22 / FN: 8	TP: 51 / FN: 19
	Sp: 0.99	4 (48 689)	Moderate	TN: 980 / FP: 10	TN: 960 / FP: 10	TN: 921 / FP: 9
Xpert Ultra in adults for PTB, MRS	Se: 0.68	4 (345)	Low	TP: 7 / FN: 3	TP: 20 / FN: 10	TP: 48 / FN: 22
	Sp: 0.98	4 (12 025)	Moderate	TN: 970 / FP: 20	TN: 951 / FP: 19	TN: 911 / FP: 19

FN: false negative; FP: false positive; MRS: microbiological reference standard; PICO: population, intervention, comparator and outcomes; PTB: pulmonary tuberculosis; Se: sensitivity; Sp: specificity; TB: tuberculosis; TN: true negative; TP: true positive.

Table 1.18. PICO 6.3: Two Xpert Ultra versus one Xpert Ultra to diagnose pulmonary TB in adults in the general population with signs and symptoms of TB or chest radiograph with lung abnormalities or both, as compared with MRS.

					1% prevalence		3% prevalence		7% prevalence	
Patient population	**Studies (persons)**	**Test accuracy (>1 Ultra)**	**Test accuracy (1 Ultra)**	**Certainty in evidence**	**>1 Ultra**	**1 Ultra**	**>1 Ultra**	**1 Ultra**	**>1 Ultra**	**1 Ultra**
1 versus 1+ Ultra for PTB in adults for PTB, MRS	3 (187)	Se: 0.75	Se: 0.64	Very Low	TP: 8 FN: 2	TP: 6 FN: 4	TP: 23 FN: 7	TP: 19 FN: 11	TP: 53 FN: 17	TP: 45 FN: 25
	3 (4893)	Sp: 0.97	Sp: 0.98	Moderate	TN: 960 FP: 30	TN: 970 FP: 20	TN: 941 FP: 29	TN: 951 FP: 19	TN: 902 FP: 28	TN: 911 FP: 19

FN: false negative; FP: false positive; MRS: microbiological reference standard; PICO: population, intervention, comparator and outcomes; PTB: pulmonary tuberculosis; Se: sensitivity; Sp: specificity; TB: tuberculosis; TN: true negative; TP: true positive.

Table 1.19. PICO 7.1: What is the diagnostic accuracy of Molbio Truenat MTB for pulmonary TB in adults, as compared with MRS?

Patient population	Test accuracy	Studies (persons)	Certainty in evidence	2.5% prevalence	10% prevalence	30% prevalence
Truenat MTB for PTB, MRS	Se: 0.73	1 (258)	Moderate	TP: 18 / FN: 7	TP: 73 / FN: 27	TP: 220 / FN: 80
	Sp: 0.98	1 (1078)	High	TN: 957 / FP: 18	TN: 884 / FP: 16	TN: 687 / FP: 13
Truenat MTB for PTB, in SS+, MRS[a]	Se: 0.92	1 (174)	Moderate	TP: 23 / FN: 2	TP: 92 / FN: 8	TP: 276 / FN: 24
Truenat MTB for PTB in SS–, MRS	Se: 0.39	1 (84)	Low	TP: 10 / FN: 15	TP: 39 / FN: 61	TP: 117 / FN: 183
	Sp: 0.98	1 (1078)	High	TN: 955 / FP: 20	TN: 881 / FP: 19	TN: 685 / FP: 15

FN: false negative; FP: false positive; MRS: microbiological reference standard; PICO: population, intervention, comparator and outcomes; PTB: pulmonary tuberculosis; Se: sensitivity; Sp: specificity; SS–: sputum smear negative; SS+: sputum smear positive; TB: tuberculosis; TN: true negative; TP: true positive.

[a] Meta-analysis for specificity was not possible because of variability of the data.

Table 1.20. PICO 7.2: What is the diagnostic accuracy of Molbio Truenat MTB Plus for pulmonary TB in adults, as compared with MRS?

Patient population	Test accuracy	Studies (persons)	Certainty in evidence	2.5% prevalence	10% prevalence	30% prevalence
Truenat MTB Plus for PTB, MRS	Se: 0.80	1 (258)	Moderate	TP: 20 / FN: 5	TP: 80 / FN: 20	TP: 239 / FN: 61
	Sp: 0.96	1 (1078)	High	TN: 940 / FP: 25	TN: 868 / FP: 32	TN: 675 / FP: 25
Truenat MTB Plus for PTB, in SS+ MRS	Se: 0.96	1 (176)	Moderate	TP: 24 / FN: 1	TP: 96 / FN: 4	TP: 288 / FN: 12
Truenat MTB Plus for PTB in SS–, MRS	Se: 0.46	1 (84)	Low	TP: 12 / FN: 13	TP: 47 / FN: 53	TP: 142 / FN: 158
	Sp: 0.97	1 (1078)	High	TN: 940 / FP: 35	TN: 868 / FP: 32	TN: 675 / FP: 25

FN: false negative; FP: false positive; MRS: microbiological reference standard; PICO: population, intervention, comparator and outcomes; PTB: pulmonary tuberculosis; Se: sensitivity; Sp: specificity; SS–: sputum smear negative; SS+: sputum smear positive; TB: tuberculosis; TN: true negative; TP: true positive.

Table 1.21. PICO 7.3: What is the diagnostic accuracy of Molbio Truenat MTB-RIF Dx for rifampicin resistance in adults, as compared with MRS?

Patient population	Test accuracy	Studies (persons)	Certainty in evidence	2% prevalence	10% prevalence	15% prevalence
Truenat MTB-RIF Dx for RR	Se: 0.84	1 (51)	Very low	TP: 17 / FN: 3	TP: 84 / FN: 16	TP: 126 / FN: 24
	Sp: 0.97	1 (258)	Moderate	TN: 954 / FP: 26	TN: 876 / FP: 24	TN: 827 / FP: 23

FN: false negative; FP: false positive; MRS: microbiological reference standard; PICO: population, intervention, comparator and outcomes; RR: rifampicin resistance; Se: sensitivity; Sp: specificity; TN: true negative; TP: true positive.

1.5 Cost–effectiveness analysis

This section deals with the following additional question:

What are the comparative cost, affordability and cost–effectiveness of implementation of Xpert MTB/RIF, Xpert Ultra, and Truenat MTB, MTB Plus and MTB-RIF Dx systems?

A systematic review was carried out, focusing on economic evaluations of molecular-based tests for the diagnosis of active TB. The tests included GeneXpert MTB/RIF (referred to as Xpert MTB/RIF), the novel Xpert Ultra and the novel Molbio Truenat MTB test. The objective of the review was to summarize current economic evidence and further understand the costs, cost–effectiveness and affordability of these molecular tests for TB diagnosis. Twenty-eight studies were identified that met the inclusion criteria and addressed one of the PICO questions of interest. Only one study assessing the cost–effectiveness of Truenat was identified, but no studies assessing the cost–effectiveness of Xpert Ultra. Most of the studies assessed Xpert MTB/RIF in outpatient settings in countries in Africa; however, also included were studies among outpatients and hospitalized patients in other countries, such as Brazil, China, Germany, Hong Kong Special Administrative Region (SAR), India, South Africa and the USA.

Studies employed a variety of different modelling approaches, populations and settings. The included studies varied in their costing, effectiveness and epidemiological parameters, making direct comparisons across studies challenging. Furthermore, there were variations in what costing elements, implementation costs and downstream costs were included in the different studies.

Although many studies demonstrated that Xpert MTB/RIF may be cost effective in diagnosing pulmonary TB, key implementation conditions and settings had a strong effect on cost–effectiveness and must be considered when implementing this test. The cost–effectiveness of Xpert MTB/RIF was shown to be improved among certain populations: those with higher TB prevalence, in PLHIV and those where rates of empirical treatment were low. Cost–effectiveness of Xpert MTB/RIF is strongly affected by factors such as location of GeneXpert machines (i.e. centralized versus decentralized facilities), test volume, underlying TB prevalence, level of empirical treatment and pretreatment loss to follow-up.

Only one study assessing the cost–effectiveness of Molbio's Truenat MTB was identified. This study suggests that Truenat MTB is likely to be cost effective if implemented at the point of care in India. However, the study relies on several important modelling assumptions, including improved linkage to care and increased treatment initiation; these assumptions should be evaluated in pragmatic trials (as has been done for Xpert MTB/RIF implementation in South Africa).

Caution should be used when generalizing cost–effectiveness and economic evaluations across settings. Local implementation conditions and settings should be taken into account, and local implementation studies may be helpful to assess the likely impact on case finding, long-term outcomes and cost–effectiveness.

There is a substantial amount of economic evidence around implementation and scale-up of Xpert MTB/RIF in different settings, most notably among outpatients presenting with signs and symptoms of TB. Most of these studies found that Xpert MTB/RIF would probably be cost effective, but there were some exceptions, and it was clear that differences in implementation approaches and settings could have an important impact on cost–effectiveness. Studies employed a wide variety of modelling and analysis approaches, assumptions, diagnostic algorithms and comparators, and they also assessed different study settings, making comparisons across studies and generalizations to other settings challenging.

Studies highlighted that implementation factors and settings need to be taken into account when generalizing cost–effectiveness results to different settings. Important factors in determining whether

Xpert MTB/RIF may be cost effective in any given setting include current standard of care, level of empirical treatment, existing testing facilities, location of Xpert MTB/RIF (centralized or decentralized facilities), TB prevalence, patient volume, pretreatment loss to follow-up and existing linkage to care. Other important cost components include whether implementation costs associated with Xpert MTB/RIF scale-up are considered, and whether downstream costs (e.g. for TB and MDR-TB treatment, and antiretroviral therapy and HIV care) were included.

Web Annex 4.6: Systematic literature review of economic evidence for molecular assays intended as initial tests for the diagnosis of pulmonary and extrapulmonary TB in adults and children.

1.6 User perspective

This section deals with the following question:

Are there implications for feasibility, accessibility, patient equity and human rights from the implementation of Xpert MTB/RIF, Xpert Ultra, and Truenat MTB, MTB Plus and MTB-RIF Dx systems?

The results of the qualitative research show that participants place great value on the ability of Xpert[14] to improve the diagnosis of drug-resistant TB; they also show the impact on patients if they cannot access testing for drug resistance through this technology. The impact on case notification and the value of Xpert for finding more TB cases was less clear, owing to widespread clinical treatment, prolonged turnaround time for results, and the challenges with feasibility and use of Xpert.

Although access has improved, not everybody who needs it can access Xpert testing. Simple laboratory procedures do not automatically translate into feasibility to implement. Rather, the feasibility of Xpert testing depends on government commitment to ensure functioning infrastructure and stable power, supply of cartridges and functioning laboratory services, investment in expertise for handling (discordant) results, effective repair services, staff with monitoring capacities, functioning sample transport, sustainable funding models and transparent donor agreements, and simple diagnostic algorithms.

With regard to acceptability, although Xpert has eased laboratory work through convenience and automation, the preference for Xpert in the laboratory can have undesired consequences for treatment monitoring with microscopy, and for reverting back to microscopy if GeneXpert instruments become non-functional. Clinicians' confidence in Xpert results is fairly high, but the challenges with feasibility and use mean that clinicians are at times deterred from ordering Xpert tests.

1.6.1 Summary of the results

1. **Xpert is unable to bridge disconnects or lack of capacity in general laboratory services**. Participants valued the option to use a specimen other than sputum, but having GeneXpert machines available in the public sector does not necessarily mean that facilities and capacities are available to extract and make use of those specimens. For example, services for histopathology and bacteriology in one country may be disconnected, and sending a specimen to histopathology in the private sector, for instance, may mean that the sample will not return to a public sector GeneXpert machine.
2. **Xpert Ultra trace results complicate decision-making**. Laboratory and clinical management of trace results was rarely straightforward. Study participants reported challenges with obtaining a second fresh sample when patients had left the facility or had been put on treatment and could not easily produce sputum. If repeat tests are conducted after trace, they cause confusion if the second test has a different result (e.g. is negative). Some laboratory managers are unsure which

[14] When not specified, this term applies to both Xpert MTB/RIF and Xpert Ultra.

result to report, and clinicians need expertise and experience to conduct more extensive evaluation for trace patients. This presents challenges in peripheral settings and where turnaround times of confirmatory tests (e.g. phenotypic DST and LPA) slow down clinical decision-making.

3. **Discordant results of repeat tests and confirmatory tests can cause confusion around what should be considered gold standard.** This is particularly the case when specimen quality might be poor. Understanding and contextualizing discordant results requires continuous training, experience and expertise.
4. **Establishing a thorough TB history of patients is uncommon, and "previously treated" is defined differently.** This has implications for potential false-positive results through Xpert testing. Clear guidance is needed on how to define previously treated patients, how to handle their Xpert results, and how to accurately capture outcomes in national databases.
5. **The lack of trained counsellors and of information provided to patients on diagnostics have negative implications.** Patients may be unwilling to accept a diagnosis and invest time and money for clinic visits, follow-up tests and treatment. Patients need better quality counselling by health workers to continue with diagnostic journeys and treatment; such counselling should include information about diagnostic technology and considerations for follow-up testing.
6. **Persistent underuse of GeneXpert machines is compounded by the challenges of delays due to sample transport, module breakdown, stock-out of cartridges or complicated diagnostic algorithms**. The presence of local Cepheid agents is key for repair. However, high workload and staff turnover, combined with infrastructure and environmental conditions, still cause frequent module breakdown, and repair work can be slow or services deemed insufficient. The challenges of cartridge stock-out lead to important delays and disruption of workflows, which in turn lead to underuse.
7. **Diagnostic algorithms that are simple to follow in a specific facility (e.g. test all those with presumptive TB) are more feasible and enhance use, but this simplicity depends on cost and supplies**. Cartridge stock-outs or prohibitive costs can complicate diagnostic algorithms, making them less feasible to follow and thus further compounding underuse. In Uganda, Xpert testing eligibility criteria had to be temporarily restricted to certain patient groups because of cartridge shortages that complicated the algorithm.
8. **Current donor agreements with governments regarding introduction of new diagnostic technologies are not transparent enough for civil society to be able to hold accountable and follow up**. Involving civil society in negotiating agreements and social contracts at national level and local facility levels can enhance accountability and the responsiveness of governments, leading to improved implementation processes and access to diagnostics.

Web Annex 4.7: Report on user perspectives on Xpert testing: results from qualitative research.

1.7 Summary of changes between the 2013 guidance and the 2020 update

Xpert MTB/RIF assay for the diagnosis of pulmonary and extrapulmonary TB in adults and children. Policy update (2013) ***(12)***	Molecular assays intended as initial tests for the diagnosis of pulmonary and extrapulmonary TB and rifampicin resistance in adults and children: rapid communication. Policy update (2020) ***(13)***	Changes
Using Xpert MTB/RIF to diagnose pulmonary TB and rifampicin resistance in adults and children 1. Xpert MTB/RIF should be used rather than conventional microscopy, culture and DST as the initial diagnostic test in adults suspected of having MDR-TB or HIV-associated TB (strong recommendation, high-quality evidence). 2. Xpert MTB/RIF should be used rather than conventional microscopy, culture and DST as the initial diagnostic test in children suspected of having MDR-TB or HIV-associated TB (strong recommendation, very low quality evidence). 3. Xpert MTB/RIF may be used rather than conventional microscopy and culture as the initial diagnostic test in all adults suspected of having TB (conditional recommendation acknowledging resource implications, high-quality evidence). 4. Xpert MTB/RIF may be used rather than conventional microscopy and culture as the initial diagnostic test in all children suspected of having TB (conditional recommendation acknowledging resource implications, very low quality evidence). 5. Xpert MTB/RIF may be used as a follow-on test to microscopy in adults suspected of having TB but not at risk of MDR-TB or HIV-associated TB, especially when further testing of smear-negative specimens is necessary (conditional recommendation acknowledging resource implications, high-quality evidence).	Xpert MTB/RIF and Xpert Ultra as initial tests in adults and children with signs and symptoms of pulmonary TB 1. In adults with signs and symptoms of pulmonary TB, Xpert MTB/RIF should be used as an initial diagnostic test for TB and for rifampicin-resistance detection rather than smear microscopy/culture and DST (strong recommendation, high certainty of evidence for test accuracy and moderate certainty of evidence for patient-important outcomes). 2. In adults with signs and symptoms of pulmonary TB without a prior history of TB (<5 years since end of treatment) or with a remote history of TB treatment (>5 years since end of treatment), Xpert Ultra should be used as the initial diagnostic test for TB and for rifampicin-resistance detection rather than smear microscopy/culture (strong recommendation, high certainty of evidence for test accuracy). 3. In adults with signs and symptoms of pulmonary TB and a prior history of TB with an end of treatment within the past 5 years, Xpert Ultra may be used as the initial diagnostic test for TB and for rifampicin-resistance detection rather than smear microscopy/culture (conditional recommendation, low certainty of evidence for test accuracy). 4. In children with signs and symptoms of pulmonary TB, Xpert MTB/RIF should be used as the initial diagnostic test for TB rather than smear microscopy/culture in sputum (moderate certainty of evidence in test accuracy), gastric aspirate (low certainty of evidence for test accuracy), nasopharyngeal aspirate (moderate certainty of evidence for test accuracy), or stool (low certainty of evidence for test accuracy) specimens (strong recommendation). 5. In children with signs and symptoms of pulmonary TB, Xpert Ultra should be used as the initial diagnostic test for TB rather than smear microscopy/culture in sputum (low certainty of evidence in test accuracy) and nasopharyngeal aspirate (very low certainty of evidence for test accuracy) specimens (strong recommendation).	1. Strong recommendation for use of Xpert MTB/RIF as an initial test for TB and rifampicin resistance in all adults and children with signs and symptoms of pulmonary TB. 2. Xpert Ultra is now recommended as an initial test for TB and rifampicin resistance in all adults and children with signs and symptoms of pulmonary TB. 3. In children, recommended use of Xpert MTB/RIF is expanded to gastric aspirate, nasopharyngeal aspirate, nasopharyngeal aspirate and stool. Use of Xpert Ultra is expanded to nasopharyngeal aspirate.

Xpert MTB/RIF assay for the diagnosis of pulmonary and extrapulmonary TB in adults and children. Policy update (2013) *(12)*	Molecular assays intended as initial tests for the diagnosis of pulmonary and extrapulmonary TB and rifampicin resistance in adults and children: rapid communication. Policy update (2020) *(13)*	Changes
Using Xpert MTB/RIF to diagnose extrapulmonary TB and rifampicin resistance in adults and children 1. Xpert MTB/RIF should be used in preference to conventional microscopy and culture as the initial diagnostic test for CSF specimens from patients suspected of having TB meningitis (strong recommendation given the urgency for rapid diagnosis, very low quality evidence). 2. Xpert MTB/RIF may be used as a replacement test for usual practice (including conventional microscopy, culture or histopathology) for testing specific non-respiratory specimens (lymph nodes and other tissues) from patients suspected of having extrapulmonary TB (conditional recommendation, very low quality evidence).	Xpert MTB/RIF and Xpert Ultra as initial tests in adults and children with signs and symptoms of extrapulmonary TB 1. In adults and children with signs and symptoms of TB meningitis, Xpert MTB/RIF or Xpert Ultra should be used in CSF as an initial diagnostic test for TB meningitis (strong recommendation, moderate certainty of evidence for test accuracy for Xpert MTB/RIF, low certainty of evidence for Xpert Ultra). 2. In adults and children with signs and symptoms of extrapulmonary TB, Xpert MTB/RIF may be used in lymph node aspirate, lymph node biopsy, pleural fluid, peritoneal fluid, pericardial fluid, synovial fluid or urine specimens as the initial diagnostic test for the corresponding form of extrapulmonary TB (conditional recommendation, moderate certainty of evidence for test accuracy for pleural fluid; low for lymph node aspirate, peritoneal fluid, synovial fluid, urine; very low for pericardial fluid, lymph nodes biopsy). 3. In adults and children with signs and symptoms of extra-pulmonary TB an Xpert Ultra may be used in lymph node aspirate and lymph node biopsy as the initial diagnostic test (conditional recommendation, low certainty of evidence). 4. In adults and children with signs and symptoms of extrapulmonary TB, Xpert MTB/RIF or Xpert Ultra should be used for rifampicin-resistance detection rather than culture and DST (strong recommendation, high certainty of evidence for test accuracy for Xpert MTB/RIF; low certainty of evidence for Xpert Ultra). 5. In HIV-positive adults and children with signs and symptoms of disseminated TB, Xpert MTB/RIF may be used in blood, as a diagnostic test for disseminated TB (conditional recommendation, very low certainty of evidence for test accuracy).	1. Improved certainty of evidence for test accuracy for Xpert MTB/RIF when used in CSF as an initial diagnostic test for TB meningitis. 2. High certainty of evidence for Xpert Ultra when used in CSF as an initial diagnostic test for TB meningitis. 3. Use of Xpert MTB/RIF in lymph node aspirate, lymph node biopsy, pleural fluid, peritoneal fluid, pericardial fluid, synovial fluid or urine specimens as the initial diagnostic test for the corresponding form of extrapulmonary TB. 4. Use of Xpert Ultra in lymph node aspirate, lymph node biopsy specimens as the initial diagnostic test for the corresponding form of extrapulmonary TB. 5. Use of Xpert Ultra for rifampicin-resistance detection in adults and children with signs and symptoms of extrapulmonary TB. 6. Use of Xpert MTB/RIF in blood for diagnosis of disseminated TB.

Xpert MTB/RIF assay for the diagnosis of pulmonary and extrapulmonary TB in adults and children. Policy update (2013) ***(12)***	Molecular assays intended as initial tests for the diagnosis of pulmonary and extrapulmonary TB and rifampicin resistance in adults and children: rapid communication. Policy update (2020) ***(13)***	Changes
	Xpert MTB/RIF and Xpert Ultra repeated testing in adults and children with signs and symptoms of pulmonary TB 1. In adults with signs and symptoms of pulmonary TB who have an Xpert Ultra trace positive result on the initial test, repeated testing with Ultra may not be used (conditional recommendation, very low certainty of evidence for test accuracy). 2. In children with signs and symptoms of pulmonary TB in settings with pretest probability below 5% and an Xpert MTB/RIF negative result on the initial test, repeated testing with Xpert MTB/RIF in sputum, gastric fluid, nasopharyngeal aspirate or stool specimens may not be used (conditional recommendation, low certainty of evidence for test accuracy for sputum and very low for other specimen types). 3. In children with signs and symptoms of pulmonary TB in settings with pretest probability 5% or more and an Xpert MTB/RIF negative result on the initial test, repeated testing with Xpert MTB/RIF (for a total of two tests) in sputum, gastric fluid, nasopharyngeal aspirate and stool specimens may be used (conditional recommendation, low certainty of evidence for test accuracy for sputum and very low for other specimen types). 4. In children with signs and symptoms of pulmonary TB in settings with pretest probability below 5% and an Xpert Ultra negative result on the initial test, repeated testing with Xpert Ultra in sputum or nasopharyngeal aspirate specimens may not be used (conditional recommendation, very low certainty of evidence for test accuracy). 5. In children with signs and symptoms of pulmonary TB in settings with pretest probability 5% or more and an Xpert Ultra negative result on the first initial test, repeated one Xpert Ultra test (for a total of two tests) in sputum and nasopharyngeal aspirate specimens may be used (conditional recommendation, very low certainty of evidence for test accuracy).	1. Not recommended repeated Xpert Ultra in adults who have an Xpert Ultra trace positive result on the initial test. 2. Not recommended repeated Xpert MTB/RIF in children in low prevalence settings. 3. Recommended repeated Xpert MTB/RIF in children in high prevalence settings in sputum, gastric fluid, nasopharyngeal aspirate and stool specimens. 4. Recommended repeated Xpert Ultra in children in both low and high prevalence settings in sputum and nasopharyngeal specimens.

Xpert MTB/RIF assay for the diagnosis of pulmonary and extrapulmonary TB in adults and children. Policy update (2013) *(12)*	Molecular assays intended as initial tests for the diagnosis of pulmonary and extrapulmonary TB and rifampicin resistance in adults and children: rapid communication. Policy update (2020) *(13)*	Changes
	Xpert MTB/RIF and Xpert Ultra as initial tests for pulmonary TB in adults in the general population either with signs and symptoms of TB or chest radiograph with lung abnormalities or both 1. In adults in the general population who had either signs or symptoms of TB or chest radiograph with lung abnormalities or both, the Xpert MTB/RIF or Xpert Ultra may replace culture as the initial test for pulmonary TB (conditional recommendation, low certainty of the evidence in test accuracy for Xpert MTB/RIF and moderate certainty for Xpert Ultra). 2. In adults in the general population who had either a positive TB symptom screen or chest radiograph with lung abnormalities or both, one Xpert Ultra test may be used rather than two Xpert Ultra tests as the initial test for pulmonary TB (conditional recommendation, very low certainty of evidence for test accuracy).	Conditional recommendation on use of Xpert MTB/RIF or Xpert Ultra for individual case management in individuals with radiographic abnormalities (but not in surveys estimating burden of disease).
	Truenat MTB, MTB Plus and Truenat MTB-RIF Dx in adults and children with signs and symptoms of pulmonary TB 1. In adults and children with signs and symptoms of pulmonary TB, the Truenat MTB or MTB Plus may be used as an initial diagnostic test for TB (conditional recommendation, low certainty of evidence for test accuracy). 2. In adults and children with signs and symptoms of pulmonary TB and a Truenat MTB or MTB Plus positive result, Truenat MTB-RIF Dx may be used as an initial test for rifampicin resistance (conditional recommendation, very low certainty of evidence for test accuracy).	1. Novel molecular tests Truenat MTB and MTB Plus are recommended as an initial test for TB. 2. Novel molecular assay Truenat MTB-RIF Dx is recommended as an initial test for rifampicin resistance in those with a Truenat MTB or MTB Plus positive result.

CSF: cerebrospinal fluid; DST: drug-susceptibility testing; HIV: human immunodeficiency virus; MDR-TB: multidrug-resistant tuberculosis; TB: tuberculosis.

Section 2. Loop-mediated isothermal amplification

A commercial molecular assay, the Loopamp™ *Mycobacterium tuberculosis* complex (MTBC) detection kit (Eiken Chemical Company, Tokyo, Japan), is based on the loop-mediated isothermal amplification (LAMP) reaction. Referred to as TB-LAMP, this is a manual assay that requires less than 1 hour to perform and can be read with the naked eye under UV light. Because it requires little infrastructure and is relatively easy to use, TB-LAMP is being explored for use as a rapid diagnostic test that would be an alternative to smear microscopy in resource-limited settings. LAMP methods have been used to detect malaria and several neglected tropical diseases.

In 2012, WHO convened a GDG that recognized TB-LAMP as offering a manual molecular approach to TB detection that could feasibly be implemented in peripheral-level microscopy laboratories once laboratory technicians had been adequately trained. The advantages of TB-LAMP are that it has a relatively high throughput, does not require sophisticated instruments, and has biosafety requirements similar to those of sputum-smear microscopy. Since 2012, some 20 additional studies in 17 countries have been conducted. WHO convened a GDG meeting in January 2016 to review evidence from a systematic review and meta-analysis of data from individual participants in these studies.

2.1 Recommendations

2.2 TB-LAMP may be used as a replacement test for sputum-smear microscopy for diagnosing pulmonary TB in adults with signs and symptoms consistent with TB. *(Conditional recommendation, very low quality evidence)*

2.3 TB-LAMP may be used as a follow-on test to smear microscopy in adults with signs and symptoms consistent with pulmonary TB, especially when further testing of sputum smear-negative specimens is necessary. *(Conditional recommendation, very low quality evidence)*

2.2 Remarks

1. These recommendations apply to settings where conventional sputum-smear microscopy can be performed.
2. TB-LAMP should not replace the use of rapid molecular tests that detect TB and resistance to rifampicin, especially among populations at risk of MDR-TB.
3. The test has limited additional diagnostic value over sputum-smear microscopy for testing PLHIV who have signs and symptoms consistent with TB.
4. These recommendations apply only to the use of TB-LAMP in testing sputum specimens from patients with signs and symptoms consistent with pulmonary TB.
5. These recommendations are extrapolated to using TB-LAMP in children, based on the generalization of data from adults, while acknowledging the difficulties of collecting sputum specimens from children.

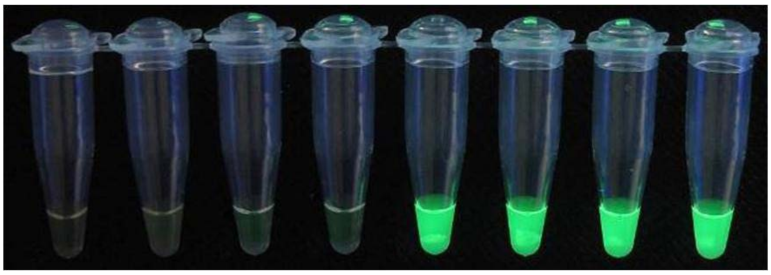
2.3 Test description

The fundamental amplification reaction requires four types of primers, which are complementary to six regions of the target gene. At about 65 °C, double-stranded DNA is in a condition of dynamic equilibrium and one of the LAMP primers can anneal to the complementary sequence of double-stranded target DNA, initiating DNA synthesis with the DNA polymerase; strand displacement activity then displaces and releases a single-stranded DNA. Owing to the complementarity of the 5′-end of the forward inner primer (known as FIP) and the backward inner primer (BIP) in nearby regions of the target amplicon, loop structures are formed. This allows variously sized structures, consisting of alternately inverted repeats of the target sequence on the same strand, to be formed in rapid succession.

The addition of loop primers, which contain sequences complementary to the single-stranded loop region on the 5′-end of the hairpin structure, speeds the reaction by providing a greater number of starting points for DNA synthesis. Using loop primers, amplification by 109–1010 times can be achieved within 15–30 minutes. The version of TB-LAMP that was evaluated includes loop primers for a total of six primers binding to eight locations. This requirement for homogeneous sequences at multiple binding sites preserves the specificity of the assay, even in the absence of a probe.

The LAMP method is relatively insensitive to the accumulation of DNA and DNA byproducts (pyrophosphate salts), so the reaction proceeds until large amounts of amplicon are generated. This feature makes it possible to visually detect successful amplification using double-stranded DNA-binding dyes, such as SYBR green, by detecting the turbidity caused by precipitating magnesium pyrophosphate or by using a non-inhibitory fluorescing reagent that is quenched in the presence of divalent cations. Fig. 2.1 shows calcein, unquenched by pyrophosphate consumption of divalent cations, fluorescing under UV light. The turbid, fluorescent product is easily seen with the naked eye.

Fig. 2.1. Visual display of TB-LAMP results under UV light



- - - - + + + +

LAMP: loop-mediated isothermal amplification; TB: tuberculosis; UV: ultraviolet.

The test procedure has three main steps (Fig. 2.2):

1. Sample preparation – bacteria are heat treated for inactivation and lysis. This step also includes the extraction of DNA.
2. Amplification – the sample is placed in a heating block at 67 °C. At this temperature, the polymerase enzyme amplifies the target DNA.
3. Visualization – the test-tube contains a double-stranded DNA-binding molecule that will fluoresce under UV light, meaning that detection can easily be performed with the naked eye.

Fig. 2.2. Description of the workflow for TB-LAMP

LAMP: loop-mediated isothermal amplification; TB: tuberculosis.

Source: Courtesy of Human Gesellschaft für Biochemica und Diagnostica mbH

2.4 Justification and evidence

The evidence reviewed and this policy guidance apply only to the use of the commercial TB-LAMP manual assay. In accordance with WHO's standards for assessing evidence when formulating policy recommendations, the GRADE approach was used. GRADE provides a structured framework to determine the quality of the evidence and to provide information on the strength of the recommendations, using PICO questions agreed by the GDG. PICO refers to the following four elements that should be included in questions that govern a systematic search of the evidence: the *population* targeted by the action or intervention (in the case of systematic reviews of the accuracy of diagnostic tests, P is the population of interest), the *intervention* (I is the index test), the *comparator* (C is the comparator test or tests) and the *outcomes* (O is usually sensitivity and specificity). The PICO questions for the review are given in Box 2.1.

Box 2.1. PICO questions addressed by the GDG

1. What is the diagnostic accuracy of TB-LAMP for detecting pulmonary TB in adults when TB-LAMP is used as a replacement test for sputum-smear microscopy compared with culture as a reference standard? (Results were stratified by HIV status.)
2. What is the diagnostic accuracy of TB-LAMP for detecting pulmonary TB in adults when TB-LAMP is used as an add-on test following negative sputum-smear microscopy compared with culture as a reference standard?
3. What is the difference in diagnostic accuracy between TB-LAMP and the Xpert MTB/RIF assay (Cepheid, Sunnyvale, USA) for detecting pulmonary TB in reference to mycobacterial culture among all adults?
4. What is the proportion of indeterminate or invalid results when TB-LAMP is used to detect pulmonary TB among all adults and among HIV-positive adults?

The review included all prospective studies that evaluated the use of TB-LAMP on sputum samples from adults with signs and symptoms consistent with pulmonary TB that were conducted in settings with an intermediate or high burden of TB. Twenty studies were identified, including all studies that were directly conducted by FIND or funded through FIND following a request for applications. Study participants who could not be classified as TB-positive or TB-negative based on the reference standard definitions described below were excluded.

The mycobacterial culture reference standards listed below were used to classify TB status. Eligible studies performed one or more sputum cultures on solid media (Löwenstein–Jensen) or on liquid media using the BACTEC™ mycobacterial growth indicator tube (MGIT; Becton Dickinson, Franklin Lakes, USA), or on both liquid and solid media. To account for the different number of cultures performed by studies and the different number of culture results available for participants, three hierarchical culture-based reference standards were used to assess diagnostic accuracy.

Standard 1 comprised:

- TB: at least one positive culture confirmed to be MTBC by speciation testing.
- Not TB: no positive and at least two negative cultures performed on two different sputum samples.

Standard 2 comprised:

- TB: at least one positive culture confirmed to be MTBC by speciation testing.
- Not TB: No positive and at least two negative cultures performed on at least one sputum sample.

Standard 3 comprised:

- TB: at least one positive culture confirmed to be MTBC by speciation testing.
- Not TB: No positive and at least one negative culture.

Across the three standards, there is an expected trade-off between the yield of a confirmed TB diagnosis (highest with Standard 1 and lowest with Standard 3) and the number of studies or participants included in the analysis (lowest with Standard 1 and highest with Standard 3). Thus, using Standard 1, the potential for false-negative index test results is highest and for false-positive index test results is lowest. Also, using Standard 1, the number of studies and study participants included is expected to be lowest because it excludes studies that performed only one culture, and study participants for whom only one negative culture result was available due to culture contamination; in contrast, using Standard 3, the number of studies and study participants is highest.

Of the 4760 adults eligible for inclusion in the analysis, 1810 participants (38%) across seven studies qualified for Standard 1 status, 3110 participants (65%) across 10 studies qualified for Standard 2 and 4596 participants (97%) across 13 qualified for Standard 3 (Table 2.1).

The performance of the test was calculated using the three different reference standards for the following scenarios:

1. TB-LAMP as a replacement for sputum-smear microscopy;
2. TB-LAMP as a replacement for sputum-smear microscopy among PLHIV;
3. TB-LAMP as an add-on test for sputum-smear microscopy negative individuals; and
4. TB-LAMP in head-to-head comparison with Xpert MTB/RIF.

Table 2.1. TB-LAMP as a replacement test for smear microscopy: estimates of pooled sensitivity and specificity

	Reference standard[a]	Pooled sensitivity[b]	Pooled specificity[b]
Replacement for SSM	Standard 1	77.7 (71.2–83.0)	98.1 (95.7–99.2)
	Standard 2	76.0 (69.9–81.2)	98.0 (96.0–99.0)
	Standard 3	80.3 (70.3–87.5)	97.7 (96.1–98.7)
Replacement for SSM for PLHIV	Standard 1	NA	NA
	Standard 2	63.8 (49.0–76.4)	98.8 (85.1–99.9)
	Standard 3	73.4 (51.9–87.6)	95.0 (64.0–99.5)
Add-on for SSM negative individuals	Standard 1	42.1 (30.0–55.3)	98.4 (95.9–99.4)
	Standard 2	42.2 (27.9–57.9)	98.0 (96.0–99.0)
	Standard 3	40.3 (27.9–54.0)	97.7 (96.1–98.6)
Compared to Xpert MTB/RIF	Standard 1	81.1 (70.6–88.5)	98.2 (95.9–99.2)
	Standard 2	80.4 (73.4–85.9)	97.4 (94.9–98.7)
	Standard 3	84.0 (75.6–90.0)	97.2 (94.4–98.6)

LAMP: loop-mediated isothermal amplification; NA: not applicable; PLHIV: people living with human immunodeficiency virus; SSM: sputum-smear microscopy; TB: tuberculosis.

a All reference standards classify patients as having TB if ≥1 positive culture was confirmed as *M. tuberculosis* by speciation testing. To be classified as not having TB, patients had to have no positive and (i) at least two negative cultures on two different sputum specimens (Standard 1), (ii) at least two negative cultures on the same or different sputum specimens (Standard 2), or (iii) at least one negative culture (Standard 3).

b Values are percentages (95% confidence intervals).

2.5 Cost–effectiveness analysis

For the cost analysis, a bottom-up micro-costing analysis was conducted – the aim being to identify, measure and value all resources relevant to providing TB-LAMP and the Xpert MTB/RIF assay as routine diagnostic tests in peripheral laboratories in Malawi and Viet Nam. The two TB-LAMP strategies (used as a replacement test for sputum-smear microscopy and as an add-on test to sputum-smear microscopy for further testing in smear-negative patients) were compared with the base case algorithm, with sputum-smear microscopy followed by clinical diagnosis in those patients with a negative microscopy result.

The weighted average per-test cost of TB-LAMP was US$ 13.78–16.22, and for the Xpert MTB/RIF assay it was US$ 19.17–28.34 when these tests were used as routine diagnostic tests at all peripheral-level laboratories in both countries. The first-year expenditure required for implementation at peripheral laboratories with a medium workload (10–15 sputum-smear microscopy tests per day) in Viet Nam was US$ 26 917 for TB-LAMP and US$ 43 325 for the Xpert MTB/RIF assay. These costs were about US$ 3000 lower in Malawi, because of lower operating and staff costs. Likewise, TB-LAMP was a considerably cheaper test to implement, accounting for 9.33% of the reported TB control budget for 2014 in Malawi and 17.2% in Viet Nam; in comparison, implementing the Xpert MTB/RIF assay accounted for 18% of the reported TB control budget in Malawi and 37% in Viet Nam. In the cost–effectiveness analyses, both of the TB-LAMP scenarios improved case-detection rates, and both strategies were cost effective when compared with WHO's willingness-to-pay threshold levels.

The findings of the cost–effectiveness analysis demonstrate that TB-LAMP is potentially a cost-effective alternative to the base case of sputum-smear microscopy plus clinical diagnosis in settings where the Xpert MTB/RIF assay cannot be implemented because of the infrastructure requirements, including a continuous power supply. However, given the inability of TB-LAMP to detect rifampicin-resistant TB (RR-TB), and its suboptimal sensitivity for detecting TB among PLHIV, national policy-makers must cautiously evaluate the operational feasibility and cost considerations before introducing this technology.

2.6 Implementation considerations

The systematic review supports the use of TB-LAMP as a replacement test for smear microscopy, for diagnosing pulmonary TB in countries with an intermediate or high burden of TB. However, the Xpert MTB/RIF assay should remain the preferred diagnostic test for anyone suspected of having TB, provided that there are sufficient resources and infrastructure to support its use, given the evidence, its ability to simultaneously identify rifampicin resistance and the fact that it is automated.

- Several operational issues accompany the implementation of TB-LAMP; for example, the need for electricity, adequate storage and waste disposal, stock monitoring and temperature control in storage settings where temperatures exceed the manufacturer's recommendation (currently 30 °C for TB-LAMP).
- TB-LAMP is designed and has been evaluated to detect *M. tuberculosis* in sputum specimens. Its use with other samples (e.g. urine, serum, plasma, CSF or other body fluids) has not been adequately evaluated.
- Adoption of TB-LAMP does not eliminate the need for smear microscopy, which should be used for monitoring the treatment of patients with drug-susceptible TB. However, the demand for conventional sputum microscopy may decrease in settings where TB-LAMP fully or partially replaces conventional sputum microscopy.
- TB-LAMP should not replace the Xpert MTB/RIF assay because the latter simultaneously detects *M. tuberculosis* and rifampicin resistance, is automated and is relatively simple to perform.
- In settings where the Xpert MTB/RIF assay cannot be implemented (e.g. because of an inadequate electric supply, or excessive temperatures, humidity or dust), TB-LAMP may be a plausible alternative.

Section 3. First-line LPAs

In 2008, WHO approved the use of commercial LPAs for detecting MTBC in combination with resistance to rifampicin and isoniazid in sputum smear-positive specimens (direct testing) and in cultured isolates of MTBC (indirect testing). A systematic review at that time evaluated the diagnostic accuracy of two commercially available LPAs – the INNO-LiPA Rif.TB assay (Innogenetics, Ghent, Belgium), and the GenoType® MTBDR*plus* (version 1), hereafter referred to as Hain version 1 – and provided evidence for WHO's endorsement *(14, 15)*. Excellent accuracy was reported for both tests in detecting rifampicin resistance, but their diagnostic accuracy for isoniazid resistance had lower sensitivity, despite the high specificity. Because there were inadequate data to allow stratification by smear status, WHO's recommendation for using LPAs was limited to culture isolates or smear-positive sputum specimens. Further data have since been published on the use of LPAs; newer versions of LPA technology have now been developed, such as the Hain GenoType MTBDR*plus* version 2, hereafter referred to as Hain version 2; and other manufacturers have entered the market, including Nipro (Tokyo, Japan), which developed the Genoscholar™ NTM+MDRTB II, hereafter referred to as Nipro.

In 2015, FIND evaluated the Nipro and the Hain version 2 LPAs, and compared them with Hain version 1. The study demonstrated equivalence among the three commercially available LPAs for detecting TB and resistance to rifampicin and isoniazid *(4)*.

3.1 Recommendation

3.1 For persons with a sputum smear-positive specimen or a cultured isolate of MTBC, commercial molecular LPAs may be used as the initial test instead of phenotypic culture-based DST to detect resistance to rifampicin and isoniazid. *(Conditional recommendation, moderate certainty in the evidence for the test's accuracy)*

3.2 Remarks

1. These recommendations apply to the use of LPAs for testing sputum smear-positive specimens (direct testing) and cultured isolates of MTBC (indirect testing) from both pulmonary and extrapulmonary sites.
2. LPAs are not recommended for the direct testing of sputum smear-negative specimens.
3. These recommendations apply to the detection of MTBC and the diagnosis of MDR-TB, but acknowledge that the accuracy of detecting resistance to rifampicin and isoniazid differs and, hence, that the accuracy of a diagnosis of MDR-TB is reduced overall.
4. These recommendations do not eliminate the need for conventional culture-based DST, which will be necessary to determine resistance to other anti-TB agents and to monitor the emergence of additional drug resistance.
5. Conventional culture-based DST for isoniazid may still be used to evaluate patients when the LPA result does not detect isoniazid resistance. This is particularly important for populations with a high pretest probability of resistance to isoniazid.
6. These recommendations apply to the use of LPA in children based on the generalization of data from adults.

3.3 Test description

LPAs are a family of DNA strip-based tests that can detect the MTBC strain and determine its drug resistance profile through the pattern of binding of amplicons (DNA amplification products) to probes targeting the following: specific parts of the MTBC genome (for MTBC detection), the most common resistance-associated mutations to first-line and second-line agents, or the corresponding wild-type DNA sequence (for detection of resistance to anti-TB drugs) *(3)*.

LPAs are based on reverse-hybridization DNA strip technology and involve three steps: DNA extraction from *M. tuberculosis* culture isolates or directly from patient specimens, followed by multiplex PCR amplification and then reverse hybridization with visualization of amplicon binding (or lack thereof) to wild-type and mutation probes *(4)*.

Although LPAs are more technically complex to perform than the Xpert MTB/RIF assay, they can detect isoniazid resistance. Testing platforms have been designed for a reference laboratory setting and are thus most applicable to high TB burden countries. Results can be obtained in 5 hours.

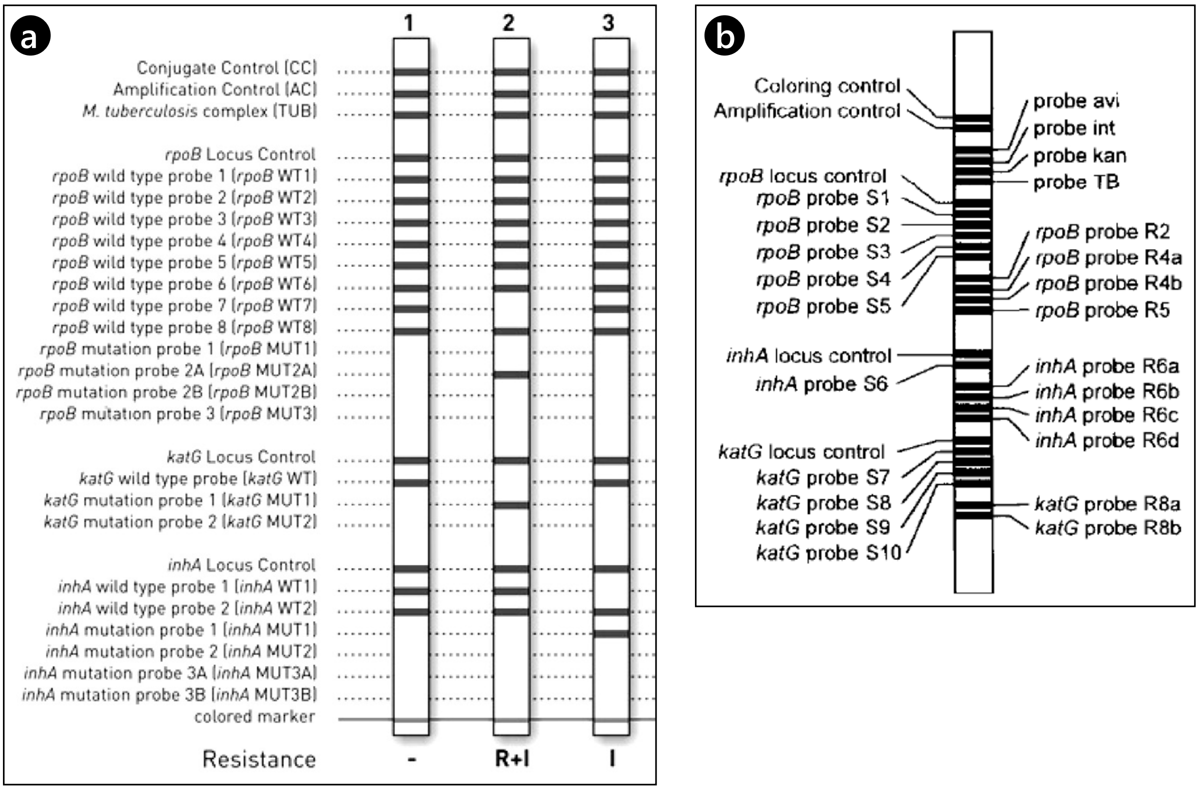
Some of these steps can be automated, making the method quicker and more robust, and reducing the risk of contamination.

The Hain version 1 and version 2 assays include *rpoB* probes to detect rifampicin resistance, *katG* probes to detect mutations associated with high-level isoniazid resistance, and *inhA* promoter probes to detect mutations usually associated with low-level isoniazid resistance. The probes used to detect wild-type and specific mutations are the same for both versions of the Hain LPA (Fig. 3.1a).

Similarly, the Nipro assay allows for the identification of MTBC, and resistance to rifampicin and isoniazid. The Nipro assay also differentiates *M. avium*, *M. intracellulare* and *M. kansasii* from other non-tuberculous mycobacteria (Fig. 3.1b).

The *rpoB*, *katG* and *inhA* promoter mutation probes are the same for the three assays, with the exception of the *katG* S315N mutation, which is included in the Nipro assay but not in Hain version 1 or version 2. There are some minor variations in the codon regions covered for the wild type among Hain version 1 and version 2, and the Nipro.

Fig. 3.1. Examples of different line probe assay strip readouts: (a) Hain GenoType MTBDRplus version 1 and version 2 (Hain Lifescience, Nehren, Germany) and (b) Nipro NTM+MDRTB Detection Kit 2 (Nipro, Tokyo, Japan)



Source: Courtesy of the Foundation for Innovative New Diagnostics (FIND).

3.4 Justification and evidence

In 2015, WHO commissioned an updated systematic review of the accuracy of commercial LPAs for detecting MTBC, and resistance to rifampicin and isoniazid. A total of 74 studies were identified, comprising 94 unique datasets (see **Annex 1.3**). Of these 94 datasets, 83 evaluated Hain version 1, five evaluated Hain version 2, and six evaluated the Nipro assay. Only one of the studies performed head-to-head testing of all three target LPAs on directly tested clinical specimens and indirectly tested isolates, and these data were included as six separate datasets *(16)*. No studies performed LPA testing on specimens and culture isolates from the same patients, precluding direct within-study comparisons.

Following the 2015 systematic review, WHO's Global TB Programme convened a GDG in March 2016 to assess the data and update the 2008 policy recommendations on using commercial LPAs to detect MTBC, and resistance to isoniazid and rifampicin. The PICO questions are given in Box 3.1.

LPAs were compared with a phenotypic culture-based DST reference standard, and a composite reference standard that combined the results from genetic sequencing with results from phenotypic culture-based DST. Phenotypic DST was the primary reference standard applied to all participants for all analyses. These analyses were stratified – first, by susceptibility or resistance to rifampicin or isoniazid (or both) and second, by type of LPA testing (indirect testing or direct testing).

Box 3.1. PICO questions

1. Should LPAs be used to guide clinical decisions to use rifampicin in the direct testing of specimens and the indirect testing of culture isolates from patients with signs and symptoms consistent with TB?
2. Should LPAs be used to guide clinical decisions to use isoniazid in the direct testing of specimens and the indirect testing of culture isolates from patients with signs and symptoms consistent with TB?
3. Should LPAs be used to diagnose MDR-TB in patients with signs and symptoms consistent with TB?
4. Should LPAs be used to diagnose TB in patients with signs and symptoms consistent with TB but for whom sputum-smear results are negative?

Several studies contributed to either sensitivity (no true negatives and no false positives) or specificity (no true positives and no false negatives) but not to both. For these studies, a univariate, random effects meta-analysis of the estimates of sensitivity or specificity was performed separately, to make optimal use of the data. The results from the univariate analysis (using all studies) were compared with the results from the bivariate analysis of the subset of studies that contributed to estimates of both sensitivity and specificity.

If there were at least four studies for index tests with data that contributed only to sensitivity or specificity, a univariate, random effects meta-analysis was performed to assess one summary estimate, assuming no correlation between sensitivity and specificity. In cases in which there were fewer than four studies, or where substantial heterogeneity was evident on forest plots that precluded a meta-analysis, a descriptive analysis was performed for these index tests. Forest plots were visually assessed for heterogeneity among the studies within each index test and in the summary plots, for variability in estimates and the width of the prediction region (a wider prediction region suggests more heterogeneity).

The performance of the tests is summarized in Table 3.1. The results are based on various numbers of studies and specimens tested. In some cases, too few studies were available for meta-analysis. The results from the only head-to-head comparison of the three tests are presented in the right-hand columns for comparison. The data presented are all comparisons with phenotypic culture-based DST as the reference standard.

Table 3.1. Performance of the three LPA tests for detection of rifampicin and isoniazid resistance with phenotypic culture-based DST as the reference standard

		Meta-analysis pooled performance		Nathavitharana et al. 2017 (16)a	
	Line probe assay	**Sensitivity (%)b**	**Specificity (%)**	**Sensitivity (%)**	**Specificity (%)**
Rifampicin sputum specimens	Hain version 1	96.8 (94.7–98.1)	98.1 (96.9–98.8)	97.1 (93.3–99.0)	97.1 (94.3–98.7)
	Hain version 2	95.8 (92.6–97.6)	98.4 (96.9–99.2)	98.2 (95.0–99.6)	97.8 (95.3–99.2)
	Nipro	75–100[c]	96.5–100[c]	96.5 (92.5–98.7)	97.5 (94.8–99.0)
Isoniazid sputum specimens	Hain version 1	88.4 (84.4–91.6)	98.3 (97.4–98.9)	94.4 (90.2–97.2)	96.4 (93.2–98.3)
	Hain version 2	94.5 (91.4–96.5)	99.3 (92.6–100.0)	95.4 (91.5–97.9)	98.8 (96.5–99.8)
	Nipro	50–94.9[c]	96.5–97.8[c]	94.9 (90.9–97.5)	97.6 (94.8–99.1)
Rifampicin culture isolates	Hain version 1	97.3 (95.7–98.3)	99.5 (98.8–98.8)	91.3 (86.0–95.0)	97.1 (94.3–98.7)
	Hain version 2	91.3[d]	98.0[d]	91.3 (86.0–95.0)	97.1 (94.3–98.7)
	Nipro	92.8–98.9 [c]	97.3–98.2[c]	92.4 (87.4–95.9)	97.5 (94.3–99.2)
Isoniazid culture isolates	Hain version 1	91.5 (89.0–93.5)	99.8 (99.3–100)	89.4 (84.3–93.3)	98.9 (96.0–99.9)
	Hain version 2	89.4[d]	98.9[d]	89.4 (84.3–93.3)	98.9 (96.0–99.9)
	Nipro	61.6–91.6[C]	99.4–100[c]	89.9 (84.9–93.8)	99.4 (96.9–100)

DST: drug-susceptibility testing; LPA: line probe assay.

[a] Results of the head-to-head comparison of the three LPA tests by Nathavitharana et al. 2017. *(16)*

[b] Sensitivity and specificity values are shown with 95% confidence interval in parenthesis.

[c] Less than four studies – meta-analysis not possible.

[d] One study.

3.5 Implementation considerations

Adopting LPAs for detecting rifampicin and isoniazid resistance does not eliminate the need for capacity for conventional culture and DST. Culture and phenotypic culture-based DST have critical roles in monitoring patients' responses to treatment and detecting additional resistance to second-line agents.

- The adoption of LPA should be phased in, starting at national or central reference laboratories, or those with proven capability to conduct molecular testing. Expansion could be considered, within the context of a country's plans for laboratory strengthening, the availability of suitable personnel in peripheral centres and the quality of specimen transport systems.
- Adequate and appropriate laboratory infrastructure and equipment should be provided, to ensure that the required precautions for biosafety and the prevention of contamination are met – specimen processing for culture and procedures for manipulating cultures must be performed in biological safety cabinets in TB-containment laboratories.
- Laboratory facilities for LPAs require at least three separate rooms, one each for DNA extraction, pre-amplification procedures, and amplification and post-amplification procedures. To avoid contamination, access to molecular facilities must be restricted, a unidirectional workflow must be implemented and stringent cleaning protocols must be established.
- Appropriate laboratory staff should be trained to conduct LPA procedures. Staff should be supervised by a senior staff member with adequate training and experience in molecular assays. A programme for the external quality assessment of laboratories using LPAs should be developed as a priority.
- Mechanisms for rapidly reporting LPA results to clinicians must be established, to provide patients with the benefit of early diagnosis. The same infrastructure used for performing LPAs can be used also to perform second-line LPAs.
- LPAs are designed to detect TB and resistance to rifampicin and isoniazid in the direct testing of processed sputum samples, and in the indirect testing of culture isolates of MTBC. The use of LPAs with other respiratory samples (e.g. from bronchoalveolar lavage or gastric aspiration) or extrapulmonary samples (e.g. tissue samples, CSF or other body fluids) have not been adequately evaluated.
- The availability of second-line agents is critical in the event that resistance to rifampicin or isoniazid, or both, is detected.
- For patients with confirmed MDR- or rifampicin-resistant TB (MDR/RR-TB), second-line LPAs are recommended to detect additional resistance to second-line anti-TB agents.

Section 4. Second-line LPAs

Genotypic (molecular) methods have considerable advantages for scaling up programmatic management and surveillance of drug-resistant TB, offering rapid diagnosis, standardized testing, potential for high throughput and fewer requirements for laboratory biosafety. Molecular tests for detecting drug resistance – for example, the GenoType MTBDR*sl* assay (Hain Lifescience, Nehren, Germany), hereafter referred to as MTBDR*sl (17)* – have shown promise for the diagnosis of drug-resistant TB. These tests are rapid (can be performed in a single working day) and detect the presence of mutations associated with drug resistance. MTBDR*sl* belongs to a category of molecular genetic tests called second-line LPAs (SL-LPAs).

MTBDR*sl* (version 1.0) was the first commercial SL-LPA for detection of resistance to second-line TB drugs. In 2015, the manufacturer developed and made commercially available version 2.0 of the MTBDR*sl* assay. Version 2.0 detects the mutations associated with fluoroquinolones and second-line injectable drug (SLID) resistance detected by version 1.0, and additional mutations. Once a diagnosis of MDR/RR-TB has been established, an SL-LPA can be used to detect additional resistance to second-line drugs.

The MTBDR*sl* assay incorporates probes to detect mutations within genes that are associated with resistance to either fluoroquinolones or SLIDs (*gyr*A and *rrs* for version 1.0 and those genes plus *gyrB* and the *eis* promoter for version 2.0). The presence of mutations in these regions does not necessarily imply resistance to all the drugs within a particular class. Although specific mutations within these regions may be associated with different levels of resistance (i.e. different minimum inhibitory concentrations) to each drug within these classes, the extent of cross-resistance is not completely understood.

4.1 Recommendations

4.1 For patients with confirmed MDR/RR-TB, SL-LPA may be used as the initial test, instead of phenotypic culture-based DST, to detect resistance to fluoroquinolones.

4.2 For patients with confirmed MDR/RR-TB, SL-LPA may be used as the initial test, instead of phenotypic culture-based DST, to detect resistance to the SLIDs.

4.2 Remarks

- These recommendations apply to the use of SL-LPA for testing sputum specimens (direct testing) and cultured isolates of *M. tuberculosis* (indirect testing) from both pulmonary and extrapulmonary sites. Direct testing on sputum specimens allows for the earlier initiation of appropriate treatment.
- These recommendations apply to the direct testing of sputum specimens from MDR/RR-TB, irrespective of the smear status, while acknowledging that the indeterminate rate is higher when testing smear-negative sputum specimens than with smear-positive sputum specimens.
- These recommendations apply to the diagnosis of extensively drug-resistant TB (XDR-TB), while acknowledging that the accuracy for detecting resistance to the fluoroquinolones and to the SLIDs differs, and hence the accuracy of a diagnosis of XDR-TB overall is reduced.
- These recommendations do not eliminate the need for conventional phenotypic DST capacity, which will be necessary to confirm resistance to other drugs and to monitor the emergence of additional drug resistance.
- Conventional phenotypic DST can still be used in the evaluation of patients with negative SL-LPA results, particularly in populations with a high pretest probability for resistance to fluoroquinolones or SLID (or both).

- These recommendations apply to the use of SL-LPA in children with confirmed MDR/RR-TB, based on the generalization of data from adults.
- Resistance-conferring mutations detected by SL-LPA are highly correlated with phenotypic resistance to ofloxacin and levofloxacin. However, the correlation of these mutations with phenotypic resistance to moxifloxacin and gatifloxacin is unclear, and the inclusion of moxifloxacin or gatifloxacin in an MDR-TB regimen is best guided by phenotypic DST results.
- Resistance-conferring mutations detected by SL-LPA are highly correlated with phenotypic resistance to SLID and are an indication to use an MDR-TB regimen that is appropriately strengthened.
- Given the high specificity for detecting resistance to fluoroquinolones and SLID, the positive results of SL-LPA could be used to guide the implementation of appropriate infection control precautions.

4.3 Test description

The SL-LPA is based on the same principle as the first-line LPA. The assay procedure can be performed **directly** using a processed sputum sample or **indirectly** using DNA isolated and amplified from a culture of *M. tuberculosis*. Direct testing involves the following steps:

1. Decontamination (e.g. with sodium hydroxide) and concentration of a sputum specimen by centrifugation.
2. Isolation and amplification of DNA.
3. Detection of the amplification products by reverse hybridization.
4. Visualization using a streptavidin-conjugated alkaline phosphatase colour reaction.

Indirect testing includes only Steps 2–4. The observed bands, each corresponding to a wild-type or resistance-genotype probe, can be used to determine the drug susceptibility profile of the analysed specimen. The assay can be performed and completed within a single working day.

The index test used was MTBDR*sl*, and the different characteristics of versions 1.0 and 2.0 are presented in Table 4.1. SL-LPAs detect specific mutations associated with resistance to the class of fluoroquinolones (including ofloxacin, levofloxacin, moxifloxacin and gatifloxacin) and SLIDs (including kanamycin, amikacin and capreomycin) in the MTBC. Version 1.0 detects mutations in the *gyrA* quinolone resistance-determining region (codons 85–97) and *rrs* (codons 1401, 1402 and 1484). Version 2.0 additionally detects mutations in the *gyrB* quinolone resistance-determining region (codons 536–541) and the *eis* promoter region (codons –10 to –14) *(17)*. Mutations in these regions may cause additional resistance to the fluoroquinolones or SLIDs, respectively; thus, version 2.0 is expected to have improved sensitivity for resistance to these drug classes. Mutations in some regions (e.g. the *eis* promoter region) may be responsible for causing resistance to one drug in a class more than other drugs within that class. For example, the *eis* C14T mutation is associated with kanamycin resistance in strains from Eastern Europe *(18)*. Version 1.0 also detects mutations in *embB* that may encode for resistance to ethambutol. Because ethambutol is a first-line drug and was omitted from version 2.0, this review did not determine the accuracy for ethambutol resistance.

Table 4.1. Characteristics of GenoType MTBDRsl versions 1.0 and 2.0, as per manufacturer

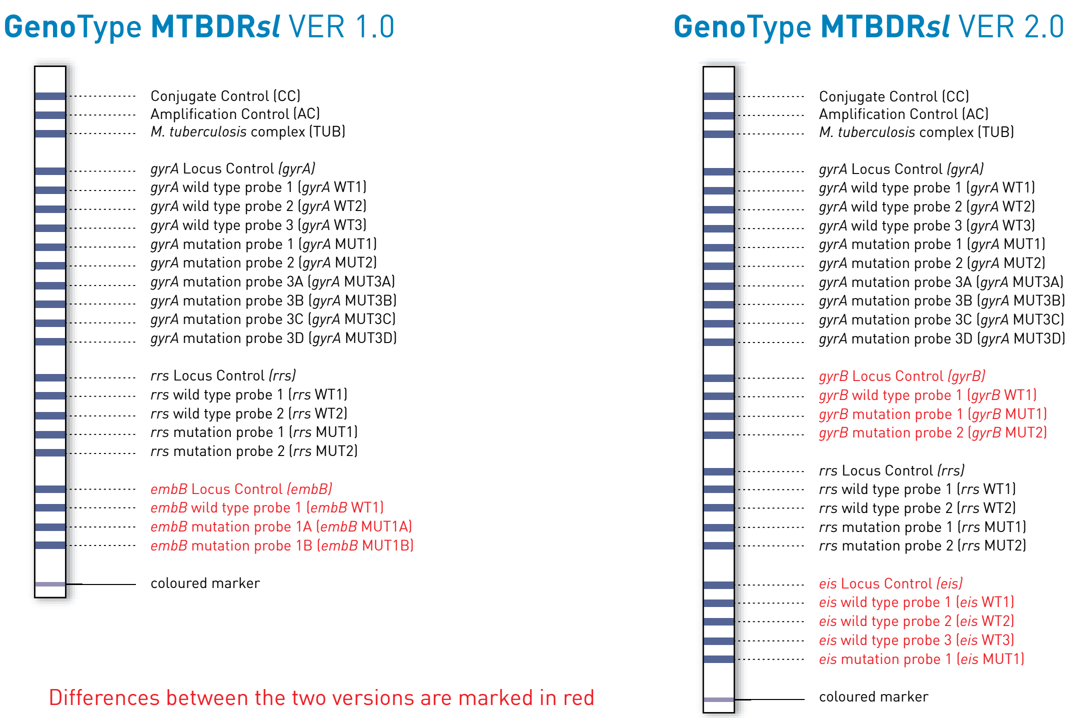
Detection	Version 1.0 MTBC and resistance to fluoroquinolones, SLIDs and ethambutol	Version 2.0 MTBC and resistance to fluoroquinolones and SLIDs
Samples	Smear-positive specimens and culture isolates	Smear-positive and smear-negative specimens and culture isolates
Fluoroquinolone resistance	Mutations in the resistance-determining region of the *gyrA* gene	Mutations in the resistance-determining regions of the *gyrA* and *gyrB* genes
SLID resistance	Mutations in the resistance-determining region of the *rrs* gene	Mutations in the resistance-determining region *rrs* gene and the *eis* promoter region
Ethambutol resistance	Mutations in the *embB* gene	Not included

MTBC: *Mycobacterium tuberculosis* complex; SLID: second-line injectable drug.

More data are needed to better understand the correlation of the presence of certain fluoroquinolone resistance-conferring mutations with phenotypic DST resistance and with patient outcomes.

Fig. 4.1 shows an example of MTBDRs*l* results for version 1.0 and 2.0. A band for the detection of the MTBC (the "TUB" band) is included, as well as two internal controls (conjugate and amplification controls), and a control for each gene locus (version 2.0: *gyr*A, *gyr*B, *rrs*, *eis*). The two internal controls plus each gene locus control should be positive, otherwise the assay cannot be evaluated for that particular drug. A result can be indeterminate for one locus but valid for another (on the basis of a gene-specific locus control failing).

Fig. 4.1. Examples of different GenoType MTBDRs*l* strip readouts



Source: Courtesy of the Foundation for Innovative New Diagnostics (FIND).

A template is supplied by the manufacturer to help the user to read the strips where the banding patterns are scored by eye, transcribed and reported. In high-volume settings, the GenoScan®, an automated reader, can be incorporated to interpret the banding patterns automatically and give a suggested interpretation. If the operator agrees with the interpretation, the results are automatically uploaded, thereby reducing possible transcription errors.

4.4 Justification and evidence

In March 2016, WHO's Global TB Programme convened a GDG to assess available data on the use of the MTBDR*sl* assay. WHO commissioned a systematic review on the accuracy and clinical use of assays for the detection of mutations associated with resistance to fluoroquinolones and SLID in people with MDR/RR-TB.

The PICO questions in Box 4.1 were designed to form the basis for the evidence search, retrieval and analysis.

Box 4.1. PICO questions

1. Should the MTBDR*sl* test be used to guide clinical decisions to use fluoroquinolones in patients with confirmed MDR/RR-TB?

- Direct testing (stratified by smear grade: smear negative; scanty; 1+; ≥2+).
- Indirect testing.

2. Should the MTBDR*sl* test be used to guide clinical decisions to use SLIDs in patients diagnosed with MDR/RR-TB?

- Direct testing (stratified by smear grade: smear negative; scanty; 1+; ≥2+).
- Indirect testing.

Twenty-nine unique studies were identified; of these, 26 evaluated the MTBDR*sl* version 1.0 assay (including 21 studies from the original Cochrane review). Three studies (one published and two unpublished) evaluated version 2.0. Data for version 1.0 and version 2.0 of the MTBDR*sl* assay were analysed separately. A phenotypic culture-based DST reference standard was used for the primary analyses. These analyses were stratified first by susceptibility or resistance to a particular drug, and second by type of SL-LPA testing (indirect testing or direct testing).

4.4 Performance of SL-LPA on sputum specimens and culture isolates

In patients with MDR/RR-TB, a positive SL-LPA result for fluoroquinolone resistance (as a class) or SLID resistance (as a group) can be treated with confidence. The diagnostic accuracy of SL-LPA is similar when performed directly on sputum specimens or indirectly on cultured isolates of *M. tuberculosis*.

Given the confidence in a positive result and the ability of the test to provide rapid results, the GDG felt that SL-LPA may be considered for use as an initial test for resistance to the fluoroquinolones and SLIDs. However, when the test shows a negative result, phenotypic culture-based DST may be necessary, especially in settings with a high pretest probability for resistance to either fluoroquinolones or SLIDs (or both). The use of SL-LPA in routine care should improve the time to the diagnosis of fluoroquinolone and SLIDs, especially when used for the direct testing of sputum specimens of patients with confirmed MDR/RR-TB. Early detection of drug resistance should allow for the earlier initiation of appropriate patient therapy and improved patient health outcomes. Overall, the test performs well in the direct testing of sputum specimens from patients with confirmed MDR/RR-TB, although the indeterminate rate is higher when testing smear-negative sputum specimens compared with smear-positive sputum specimens.

When the MTBDR*sl* assay is used in the direct testing of smear-negative sputum specimens from a population of patients with confirmed drug-resistant TB, up to 44% of the results may be indeterminate (less with version 2.0, although very limited data) and hence require repeat or additional testing. However, if the same test were to be applied to the testing of smear-negative sputum specimens from patients without confirmed TB or drug-resistant TB (i.e. patients suspected of having drug-resistant TB), the indeterminate rate for the test would be significantly higher. Given the test's sensitivity and specificity when an SL-LPA is done directly on sputum, the GDG felt that SL-LPAs can be used for the testing of all sputum specimens from patients with confirmed MDR/RR-TB, irrespective of whether the microscopy result is positive or negative.

Table 4.2. Accuracy of GenoType MTBDRsl (version 1.0) for fluoroquinolone and SLID resistance and XDR-TB, indirect and direct testing (smear-positive specimens), culture-based DST reference standard

Pooled sensitivity (95% CI)	Pooled specificity (95% CI)	Pooled sensitivity (95% CI)	Pooled specificity (95% CI)	Pooled sensitivity *P* value[a]	Pooled specificity *P* value[a]
Fluoroquinolones, indirect testing (19 studies, 2223 participants)		Fluoroquinolones, direct testing (9 studies, 1771 participants)			
85.6% (79.2–90.4%)	98.5% (95.7–99.5%)	86.2% (74.6–93.0%)	98.6% (96.9–99.4%)	0.932	0.333
SLID, indirect testing (16 studies, 1921 participants)		SLID, direct testing (8 studies, 1639 participants)			
76.5% (63.3–86.0%)	99.1% (97.3–99.7%)	87.0% (38.1–98.6%)	99.5% (93.6–100.0%)	0.547	0.664
XDR-TB, indirect testing (8 studies, 880 participants)		XDR-TB, direct testing (6 studies, 1420 participants)			
70.9% (42.9–88.8%)	98.8% (96.1–99.6%)	69.4% (38.8–89.0%)	99.4% (95.0–99.3%)	0.888	0.855

CI: confidence interval: DST: drug-susceptibility testing; SLID: second-line injectable drugs; XDR-TB: extremely drug-resistant tuberculosis.

[a] Likelihood ratio test for evidence of a significant difference between accuracy estimates.

For the reasons mentioned above (inadequate data owing to too few studies on version 2.0), results are not presented here for version 2.0. For MTBDR*sl* version 2.0, the data were either too sparse or too heterogeneous to combine in a meta-analysis or to compare indirect and direct testing.

Three studies evaluated the MTBDR*sl* version 2.0 in 562 individuals, including 111 confirmed cases of TB with fluoroquinolone resistance by indirect testing on a culture of *M. tuberculosis* compared with a phenotypic culture-based DST reference standard. Estimates of sensitivity ranged from 84% to 100% and specificity from 99% to 100%.

See **Web Annex 4.8** for details of the drug concentrations used in culture-based DST to evaluate the performance of SL-LPAs in each included study.

4.5 Implementation considerations

The SL-LPA should only be used to test specimens from patients with confirmed MDR/RR-TB. Adoption of SL-LPAs does not eliminate the need for conventional culture and DST capability. Despite good specificity of SL-LPAs for the detection of resistance to fluoroquinolones and the SLIDs, culture and phenotypic DST is required to completely exclude resistance to these drug classes as well as to other second-line drugs. The following implementation considerations apply:

- SL-LPAs cannot determine resistance to individual drugs in the class of fluoroquinolones. Resistance-conferring mutations detected by SL-LPAs are highly correlated with phenotypic resistance to ofloxacin and levofloxacin. However, the correlation of these mutations with phenotypic resistance to moxifloxacin and gatifloxacin is unclear, and the inclusion of moxifloxacin or gatifloxacin in an MDR-TB regimen is best guided by phenotypic DST results.
- Mutations in some regions (e.g. the *eis* promoter region) may be responsible for causing resistance to one drug in a class more than other drugs within that class. For example, the *eis* C14T mutation is associated with kanamycin resistance in strains from Eastern Europe.
- SL-LPAs should be used in the direct testing of sputum specimens, irrespective of whether samples are smear negative or smear positive.
- SL-LPAs are designed to detect TB and resistance to fluroquinolones and SLIDs from sputum samples. Other respiratory samples (e.g. bronchoalveolar lavage and gastric aspirates) or extrapulmonary samples (tissue samples, CSF or other body fluids) have not been adequately evaluated.
- Culture and phenotypic DST plays a critical role in the monitoring of a patient's response to treatment, and in detecting additional resistance to second-line drugs during treatment.
- SL-LPAs are suitable for use at the central or national reference laboratory level; they can also be used at the regional level if the appropriate infrastructure can be ensured (three separate rooms are required).
- All patients identified by SL-LPAs should have access to appropriate treatment and ancillary medications.

Section 5. Lateral flow urine lipoarabinomannan assay

Tests based on the detection of the lipoarabinomannan (LAM) antigen in urine have emerged as potential point-of-care tests for TB. The currently available urinary LAM assays have suboptimal sensitivity, and are therefore not suitable as general diagnostic tests for TB. However, unlike traditional diagnostic methods, they demonstrate improved sensitivity for the diagnosis of TB among individuals coinfected with HIV. The estimated sensitivity is even greater in patients with low CD4 cell counts. The lateral flow urine LAM assay (LF-LAM) strip-test – the Alere Determine TB LAM Ag (USA), hereafter referred to as AlereLAM – is currently the only commercially available urinary LAM test that potentially could be used as a rule-in test for TB in patients with advanced HIV-induced immunosuppression, and facilitate the early initiation of anti-TB treatment.

5.1 Recommendations

In inpatient settings

5.1 WHO strongly recommends using LF-LAM to assist in the diagnosis of active TB in HIV-positive adults, adolescents and children:

1. with signs and symptoms of TB (pulmonary and/or extrapulmonary)
   (***strong** recommendation, moderate certainty in the evidence about the intervention effects*); or
2. with advanced HIV disease[15] or who are seriously ill[16]
   (***strong** recommendation, moderate certainty in the evidence about the intervention effects*) [1];[17] or
3. irrespective of signs and symptoms of TB and with a CD4 cell count of less than 200 cells/mm^3
   (***strong** recommendation, moderate certainty in the evidence about the intervention effects*) [2].

In outpatient settings

5.2 WHO suggests using LF-LAM to assist in the diagnosis of active TB in HIV-positive adults, adolescents and children:

1. with signs and symptoms of TB (pulmonary and/or extrapulmonary) or seriously ill
   (***conditional** recommendation, low certainty in the evidence about test accuracy*) [3]; and
2. irrespective of signs and symptoms of TB and with a CD4 cell count of less than 100 cells/mm^3
   (***conditional** recommendation, very low certainty in the evidence about test accuracy*) [4].

In outpatient settings

5.3 WHO recommends against using LF-LAM to assist in the diagnosis of active TB in HIV-positive adults, adolescents and children:

1. without assessing TB symptoms
   (***strong** recommendation, very low certainty in the evidence about test accuracy*) [5];
2. without TB symptoms and unknown CD4 cell count or without TB symptoms and CD4 cell count greater than or equal to 200 cells/mm^3
   (***strong** recommendation, very low certainty in the evidence about test accuracy*) [6]; and
3. without TB symptoms and with a CD4 cell count of 100–200 cells/mm^3
   (***conditional** recommendation, very low certainty in the evidence about test accuracy*) [7].

---

[15] For adults, adolescents, and children aged 5 years or more, "advanced HIV disease" is defined as a CD4 cell count of less than 200 cells/mm^3 or a WHO clinical stage 3 or 4 event at presentation for care. All children with HIV aged under 5 years should be considered as having advanced disease at presentation.

[16] "Seriously ill" is defined based on four danger signs: respiratory rate of more than 30/minute, temperature of more than 39 °C, heart rate of more than 120/minute and unable to walk unaided.

[17] Numbers in square brackets indicate the number of the relevant "evidence to decision" (EtD) table in Web Annex 3.

5.2 Remarks

1. The reviewed evidence and recommendations apply to the use of AlereLAM only, because other in-house LAM-based assays have not been adequately validated or used outside limited research settings. Any new or generic LAM-based assay should be subject to adequate validation in the settings of intended use.
2. All patients with signs and symptoms of pulmonary TB who are capable of producing sputum should submit at least one sputum specimen for Xpert MTB/RIF (Ultra) assay, as their initial diagnostic test. This also includes children and adolescents living with HIV who are able to provide a sputum sample.
3. These recommendations also apply to adolescents and children living with HIV, based on generalization of data from adults, while acknowledging that there are very limited data for these population groups.
4. LF-LAM should be used as an add-on to clinical judgement in combination with other tests; it should not be used as a replacement or triage test.

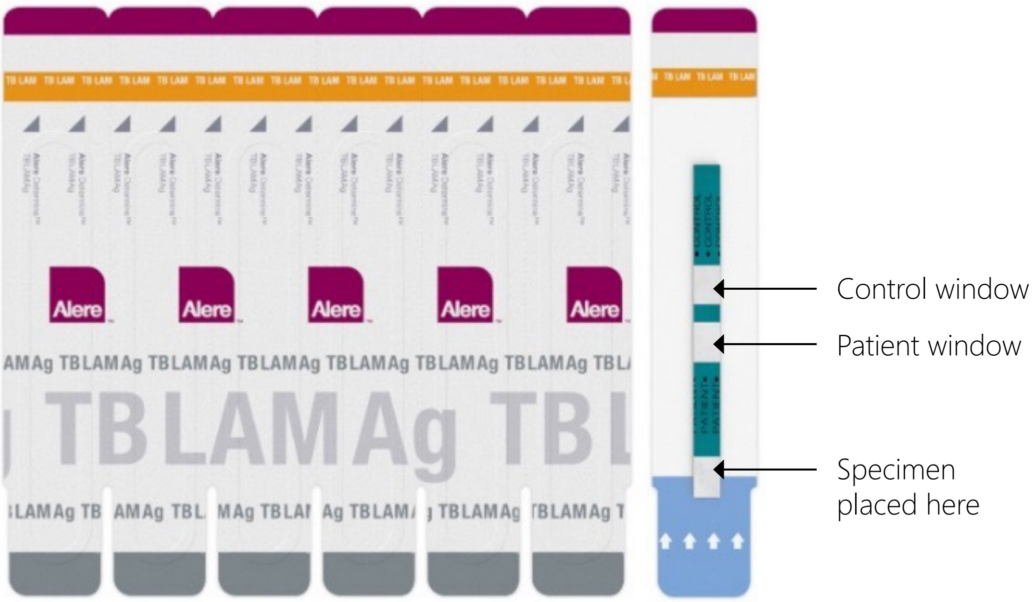
5.3 Test description

The urine-based LF-LAM AlereLAM is a commercially available point-of-care test for active TB *(19)*. AlereLAM is an immunocapture assay that detects LAM antigen in urine, LAM being a lipopolysaccharide present in mycobacterial cell walls that is released from metabolically active or degenerating bacterial cells during TB disease *(19, 20)*.

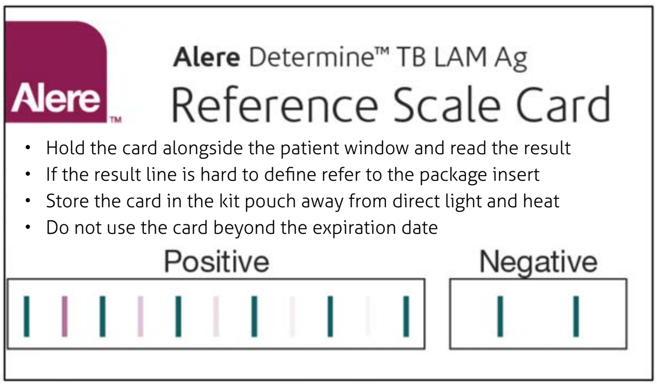
AlereLAM is performed manually by applying 60 µL of urine to the test strip (the white pad marked by the arrow symbols in Fig. 5.1A) and incubating at room temperature for 25 minutes. The strip is then inspected by eye for visible bands. The intensity of any visible band on the test strip is graded by comparing it with the intensities of the bands on a manufacturer-supplied reference scale card (as shown in the example in Fig. 5.1).

Fig. 5.1. Alere Determine TB LAM Ag tests (AlereLAM): (A) individual test strip, and (B) reference card accompanying test strips to "grade" the test result and determine positivity

a



b



AlereLAM is being considered as a diagnostic test that may be used in combination with existing tests for the diagnosis of HIV-associated TB.

5.4 Justification and evidence

WHO commissioned a systematic review to summarize the current scientific literature on the accuracy of AlereLAM for the diagnosis of TB in PLHIV as part of a WHO process to develop updated guidelines for use of the AlereLAM assay.

The PICO questions shown in Box 5.1 were designed to form the basis for the evidence search, retrieval and analysis.

Box 5.1. PICO questions

1. What is the diagnostic accuracy of LF-LAM for the diagnosis of TB in all HIV-positive adults and children with signs and symptoms of TB?

- in inpatient settings (adults, adolescents and older children)
- in outpatient settings (adults, adolescents and older children)
- in all settings (adults, adolescents and older children)
- in inpatient settings (children aged ≤5 years)
- in outpatient settings (children aged ≤5 years)
- in all settings (children aged ≤5 years)

2. What is the diagnostic accuracy of LF-LAM for the diagnosis of TB in all HIV-positive adults and children irrespective of signs and symptoms of TB?

- in inpatient settings (adults, adolescents and older children)
- in outpatient settings (adults, adolescents and older children)
- in all settings (adults, adolescents and older children)
- in inpatient settings (children aged ≤5 years)
- in outpatient settings (children aged ≤5 years)
- in all settings (children aged ≤5 years)

3. What is the diagnostic accuracy of LF-LAM for the diagnosis of TB in adults with advanced HIV disease irrespective of signs and symptoms of TB?

- in inpatient settings, CD4 cell count ≤200
- in outpatient settings, CD4 cell count ≤200
- in all settings, CD4 cell count ≤200
- in inpatient settings, CD4 cell count ≤100
- in outpatient settings, CD4 cell count ≤100
- in all settings, CD4 cell count ≤100

4. Can the use of LF-LAM in HIV-positive adults reduce mortality associated with advanced HIV disease?

- in all settings
- in inpatient settings
- in outpatient settings
- in individuals with CD4 cell count ≤200
- in inpatient settings, CD4 cell count ≤200
- in outpatient settings, CD4 cell count ≤200
- in individuals with CD4 cell count ≤100
- in inpatient settings, CD4 cell count ≤100
- in outpatient settings, CD4 cell count ≤100

5. Additional questions:

- What are the comparative cost, affordability and cost–effectiveness of implementation of LF-LAM (AlereLAM versus FujiLAM) – based on review of the published literature and estimations?
- Are there possible implications for patient equity from the implementation of LF-LAM (AlereLAM versus FujiLAM) – based on review of the published literature and estimations?
- What are the human rights implications from the implementation of LF-LAM – based on review of the published literature and comparative analysis of the two available LF-LAM (AlereLAM versus FujiLAM)?

The review identified 15 unique published studies that assessed the accuracy of AlereLAM in adults, and integrated nine new studies identified since the original WHO and Cochrane reviews in 2015 and 2016, respectively *(21, 22)*. All studies included in the systematic review were performed in high TB/HIV burden countries. The positive AlereLAM results were reported in accordance with the manufacturer's updated recommendations for test interpretation (graded on a scale of 1 to 4, based on band intensity). All analyses were performed with respect to an MRS.

The 15 included studies involved 6814 participants, of whom 1761 (26%) had TB. Eight of the studies evaluated the accuracy of AlereLAM for TB diagnosis in participants with signs and symptoms suggestive of TB; these studies involved 3449 participants, of whom 1277 (37%) had TB. Seven studies evaluated the accuracy of AlereLAM for diagnosis of unselected participants who may or may not have had TB signs and symptoms at enrolment; these studies involved 3365 participants, of whom 439 (13%) had TB.

All studies were performed in high TB/HIV burden countries that were classified as low-income or middle-income countries. The studies had substantial differences in the following characteristics: study population ("studies with symptomatic participants" and "studies with unselected participants"), setting (inpatients versus outpatients), median CD4 cell count, TB prevalence, inclusion and exclusion of participants based on whether or not they could produce sputum, and whether patients were evaluated for pulmonary TB or extrapulmonary TB, or both.

Most studies reported that a valid AlereLAM result was obtained on the first attempt for all tests. Uninterpretable test results (<1%) were reported in only three studies *(23–25)*.

5.4 Summary of the results

For TB diagnosis in HIV-positive adults presenting with signs and symptoms of TB, the diagnostic accuracy of AlereLAM is as follows:

- in *inpatient* settings, sensitivity 52% (40–64%)[18] and specificity 87% (78–93%);
- in *outpatient* settings, sensitivity 29% (17–47%) and specificity 96% (91–99%); and
- in *all* settings, sensitivity 42% (31–55%) and specificity 91% (85–95%).

For TB diagnosis in HIV-positive adults, irrespective of signs and symptoms of TB, the diagnostic accuracy of AlereLAM is as follows:

- in *inpatient* settings, sensitivity 62% (41–83%) and specificity 84% (48–96%);
- in *outpatient* settings, sensitivity 31% (18–47%) and specificity 95% (87–99%); and
- in *all* settings, sensitivity 35% (22–50%) and specificity 95% (89–98%).

[18] The numbers in brackets show the 95% credible interval (CrI).

For diagnosis of TB in adults with advanced HIV disease, irrespective of signs and symptoms of TB, the diagnostic accuracy of AlereLAM (limited data available) is as follows:

- in *inpatient* settings, CD4 cell count ≤200, sensitivity 64% (35–87%) and specificity 82% (67–93%) (one study);
- in *outpatient* settings, CD4 cell count ≤200, sensitivity 21% (8–48%) and specificity 96% (89–99%);
- in *all* settings, CD4 cell count ≤200, sensitivity 26% (9–56%) and specificity 96% (87–98%);
- in *inpatient* settings, CD4 cell count ≤100, sensitivity 57% (33–79%) and specificity 90% (69–97%);
- in *outpatient* settings, CD4 cell count ≤100, sensitivity 40% (20–64%) and specificity 87% (68–94%); and
- in *all* settings, CD4 cell count ≤100, sensitivity 47% (30–64%) and specificity 90% (77–96%).

For diagnosis of TB in HIV-positive children, the diagnostic accuracy of AlereLAM (limited data available) is as follows:

- in *all* settings, including all children, for individual studies, sensitivity and specificity were:
  - 42% (15–72%) and 94% (73–100%) (one study conducted in an outpatient setting);
  - 56% (21–86%) and 95% (90–98%) (one study conducted in an inpatient setting); and
  - 43% (23–66%) and 80% (69–88%) (one study conducted in both inpatient and outpatient settings).

For use of AlereLAM to reduce mortality associated with advanced HIV disease (two randomized trials):

- the pooled risk ratio for mortality was 0.85 (0.76–0.94); and
- the absolute effect was 35 fewer deaths per 1000 (from 14 fewer to 55 fewer) (PICO 4).

Table 5.1 presents pooled sensitivity and specificity results for AlereLAM against an MRS grouped by the study population, TB diagnosis among "symptomatic participants" and TB diagnosis among "unselected participants".

Table 5.1. AlereLAM pooled sensitivity and specificity for TB diagnosis, by study population

Type of analysis	Symptomatic participants				Unselected participants			
	Studies (total participants)	Participants with TB	Pooled sensitivity (95% CrI)	Pooled specificity (95% CrI)	Studies (total participants)	Participants with TB	Pooled sensitivity (95% CrI)	Pooled specificity (95% CrI)
Overall accuracy	8 studies (3449)	1277 (37%)	42% (31–55%)	91% (85–95%)	7 studies (3365)	432 (13%)	35% (22–50%)	95% (89–98%)
By setting								
Inpatient	6 studies (2253)	868 (39%)	52% (40–64%)	87% (78–93%)	3 studies (537)	159 (30%)	62% (41–83%)	84% (48–96%)
Outpatient	4 studies (1196)	409 (34%)	29% (17–47%)	96% (91–99%)	6 studies (2828)	273 (10%)	31% (18–47%)	95% (87–99%)
By CD4 cell count								
CD4 >200	3 studies (738)	163 (22%)	16% (8–31%)	94% (81–97%)	1 study[a] (156)	11 (7%)	Not applicable	Not applicable
CD4 ≤200	4 studies (1825)	722 (40%)	45% (31–61%)	89% (77–94%)	2 studies (706)	82 (12%)	26% (9–56%)	96% (87–98%)
CD4 >100	4 studies (1519)	425 (28%)	17% (10–27%)	95% (89–98%)	4 studies (952)	115 (12%)	20% (10–35%)	98% (95–99%)
CD4 ≤100	4 studies (1239)	512 (41%)	54% (38–69%)	88% (77–94%)	3 studies (417)	130 (31%)	47% (40–64%)	90% (77–96%)
CD4 101–200	4 studies (586)	210 (36%)	24% (14–38%)	90% (77–96%)	1 study[b] (103)	13 (13%)	Not applicable	Not applicable

Type of analysis	Symptomatic participants				Unselected participants			
	Studies (total participants)	Participants with TB	Pooled sensitivity (95% CrI)	Pooled specificity (95% CrI)	Studies (total participants)	Participants with TB	Pooled sensitivity (95% CrI)	Pooled specificity (95% CrI)
By CD4 and setting								
CD4 ≤200 inpatient	2 studies (1009)	348 (34%)	54% (34–73%)	80% (58–91%)	1 study[c] (54)	14 (26%)	Not applicable	Not applicable
CD4 ≤100 inpatient	2 studies (734)	270 (37%)	61% (40–78%)	81% (61–91%)	2 studies (200)	84 (42%)	57% (33–79%)	90% (69–97%)
CD4 101–200 inpatient	2 studies (275)	78 (28%)	32% (16–57%)	81% (55–92%)	1 study[d] (9)	4 (44%)	Not applicable	Not applicable
CD4 ≤200 outpatient	1 study[e] (249)	97 (39%)	Not applicable	Not applicable	2 studies (652)	68 (10%)	21% (8–48%)	96% (89–99%)
CD4 ≤100 outpatient	1 study[f] (121)	48 (40%)	Not applicable	Not applicable	2 studies (217)	46 (21%)	40% (20–64%)	87% (68–94%)
CD4 101–200 outpatient	1 study[g] (128)	51 (40%)	Not applicable	Not applicable	1 study[h] (94)	9 (10%)	Not applicable	Not applicable

AlereLAM: Alere Determine™ TB lipoarabinomannan assay; CrI: credible interval; TB: tuberculosis.

[a] *(7, 26)*, sensitivity 27% (6–61%); specificity 99% (96–100%).

[b] *(7, 26)*, sensitivity 38% (14–68%); specificity 99% (94–100%).

[c] *(7, 26)*, sensitivity 64% (35–87%); specificity 82% (67–93%).

[d] *(7, 26)*, sensitivity 75% (19–99%); specificity 100% (48–100%).

[e] *(4, 23)*, sensitivity 24% (16–33%); specificity 94% (89–97%).

[f] *(4, 23)*, sensitivity 30% (18–46%); specificity 93% (85–98%).

[g] *(4, 23)*, sensitivity 18% (8–31%); specificity 95% (87–99%).

[h] *(7, 26)*, sensitivity 22% (3–60%); specificity 99% (94–100%).

More details are given in **Web Annex 4.9** LF-LAM for detecting active tuberculosis in people living with HIV: an updated systematic review.

5.5 Cost–effectiveness analysis

Economic evidence for the implementation and scale-up of LF-LAM is limited. The studies that have been done show a consistent trend, suggesting that LF-LAM could be cost effective in a population of African adults living with HIV (particularly among hospitalized patients).

More details are given in **Web Annex 4.10** Economic evaluations of LF-LAM for the diagnosis of active tuberculosis in HIV-positive individuals: an updated systematic review.

5.6 User perspective

For a qualitative study on user perspectives, 15 semi-structured interviews were conducted during February and March 2019 with clinicians, nurses, programme officers, laboratory staff and patient advocates in Kenya, South Africa and Uganda. The results showed that LF-LAM clearly addresses a need and makes an important difference in a population in which TB is hard to diagnose. In line with the global discourse on LF-LAM, the participants in this study generally saw LF-LAM as an easy-to-use, rapid test that requires little maintenance and equipment, and crucially does not rely on sputum but on urine, a specimen that is easier to obtain and safer to work with. However, the perceived benefits of specimen, turnaround time, user friendliness, cost and maintenance requirements can also pose a challenge, depending on the particular situation and the capacities in which the test is used. Similarly, the infrastructure requirements are minimal but there can still be challenges with stock-outs, lack of urine containers and shelf life. Finally, even though the turnaround time is in theory only 25 minutes, in many settings, treatment is not initiated until the next day.

Overall, the results from the qualitative study suggest that the benefits outweigh the challenges, especially given the absence of viable diagnostic alternatives for this particular patient group. These results also show that it is essential to pay attention to how diagnostics are operationalized. Just because a technology is quicker, easier to conduct and cheaper than existing diagnostics, this does not mean it is necessarily more successful in being implemented.

More details are given in **Web Annex 4.11** User perspectives on TB-LAM for the diagnosis of active tuberculosis: results from qualitative research.

5.7 Summary of changes between the 2015 guidance and the 2019 update

The use of lateral flow urine lipoarabinomannan assay (LF-LAM) for the diagnosis and screening of active tuberculosis in people living with HIV. Policy guidance (2015) *(22)*	Lateral flow urine lipoarabinomannan assay (LF-LAM) for the diagnosis of active tuberculosis in people living with HIV. Policy update (2019) *(27)*	Changes
LF-LAM may be used to assist in the diagnosis of TB in HIV-positive adults in patients with signs and symptoms of TB (pulmonary and/or extrapulmonary) who have a CD4 cell count ≤100 cells/µL, or HIV-positive patients who are seriously ill[a] regardless of CD4 cell count or with unknown CD4 cell count (conditional recommendation, low quality of evidence).	**In inpatient settings**, WHO strongly recommends using LF-LAM to assist in the diagnosis of active TB in HIV-positive adults, adolescents and children: • with signs and symptoms of TB (pulmonary and/or extrapulmonary) (strong recommendation, moderate certainty in the evidence about the intervention effects); *or* • with advanced HIV disease;[b] or • who are seriously ill (strong recommendation, moderate certainty in the evidence about the intervention effects); or • irrespective of signs and symptoms of TB and with a CD4 cell count <200 (strong recommendation, moderate certainty in the evidence about the intervention effects).	Increased strength of the recommendation. Improved quality of evidence. Increased scope of the recommendation: – all symptomatic or seriously ill inpatients, irrespective of CD4 cell count; – all inpatients with advanced HIV disease; and – inpatients with or without signs and symptoms of TB who have a CD4 cell count <200.
This recommendation also applies to HIV-positive adult outpatients with signs and symptoms of TB (pulmonary and/or extrapulmonary) who have a CD4 cell count ≤100 cells/µL, or HIV-positive patients who are seriously ill regardless of CD4 cell count or with unknown CD4 cell count, based on the generalization of data from inpatients.	**In outpatient settings**, WHO suggests using LF-LAM to assist in the diagnosis of active TB in HIV-positive adults, adolescents and children: • with signs and symptoms of TB (pulmonary and/or extrapulmonary) or seriously ill (conditional recommendation, low certainty in the evidence about test accuracy); and • irrespective of signs and symptoms of TB and with a CD4 cell count <100 (conditional recommendation, very low certainty in the evidence about test accuracy).	Increased scope of the recommendation: – all outpatients with signs and symptoms of TB or seriously ill; and – outpatients with a CD4 cell count <100, irrespective of signs and symptoms of TB.

The use of lateral flow urine lipoarabinomannan assay (LF-LAM) for the diagnosis and screening of active tuberculosis in people living with HIV. Policy guidance (2015) *(22)*	**Lateral flow urine lipoarabinomannan assay (LF-LAM) for the diagnosis of active tuberculosis in people living with HIV. Policy update (2019) *(27)***	**Changes**
Except as specifically described below for persons with HIV infection with low CD4 cell counts or who are seriously ill, LF-LAM should not be used for the diagnosis of TB (strong recommendation, low quality of evidence).	**In outpatient settings**, WHO recommends against using LF-LAM to assist in the diagnosis of active TB in HIV-positive adults, adolescents and children: • without assessing TB symptoms (strong recommendation, very low certainty in the evidence about test accuracy); • without TB symptoms and unknown CD4 cell count, or without TB symptoms and CD4 cell count ≥200 (strong recommendation, very low certainty in the evidence about test accuracy); or • without TB symptoms and with a CD4 cell count of 100–200 (conditional recommendation, very low certainty in the evidence about test accuracy).	Better definition of patient populations for negative recommendation against use of LF-LAM.
LF-LAM **should not** be used as a screening test for TB (strong recommendation, low quality of evidence).	See inpatient and outpatient recommendations above for situations in which LF-LAM is suggested for use among individuals, irrespective of signs and symptoms of TB. See outpatient recommendations above for situations in which WHO recommends against LF-LAM use.	Clarification of recommendation for usage among individuals with and without TB signs and symptoms (i.e. irrespective of signs and symptoms): – LF-LAM is strongly recommended for inpatients with advanced HIV disease, and individuals with a CD4 cell count <200, irrespective of symptoms; and – LF-LAM is suggested for outpatients with a CD4 cell count <100, irrespective of symptoms. See above for patient populations with a recommendation against usage.

The use of lateral flow urine lipoarabinomannan assay (LF-LAM) for the diagnosis and screening of active tuberculosis in people living with HIV. Policy guidance (2015) *(22)*	Lateral flow urine lipoarabinomannan assay (LF-LAM) for the diagnosis of active tuberculosis in people living with HIV. Policy update (2019) *(27)*	Changes
This recommendation also applies to HIV-positive children with signs and symptoms of TB (pulmonary and/or extrapulmonary) based on the generalization of data from adults while acknowledging very limited data and concern regarding low specificity of the LF-LAM assay in children.	These recommendations also apply to adolescents and children living with HIV, based on generalization of data from adults, while acknowledging that data for these population groups are limited.	

HIV: human immunodeficiency virus; LF-LAM: lateral flow urine lipoarabinomannan assay; TB: tuberculosis; WHO: World Health Organization.

[a] "Seriously ill" is defined based on four danger signs: respiratory rate of more than 30/minute, temperature of more than 39 °C, heart rate of more than 120/minute and unable to walk unaided.

[b] For adults, adolescents, and children aged 5 years or more, "advanced HIV disease" is defined as a CD4 cell count of less than 200 cells/mm^3 or a WHO clinical stage 3 or 4 event at presentation for care. All children with HIV who are aged under 5 years should be considered as having advanced disease at presentation.

Research gaps

Current recommendations on the various methods and tools should not prevent or restrict further research on new, rapid molecular drug-susceptibility tests, especially for assays that can be used as close as possible to where patients with a presumptive diagnosis of TB are identified and where treatment can be initiated. Priorities for further operational research on diagnostics are listed below, grouped for each technology.

Molecular assays intended as initial tests

- Evaluation of the impact of Xpert Ultra testing on patient-important outcomes (cure, mortality, time to diagnosis and time to start treatment).
- Evaluation of the diagnostic accuracy of Xpert Ultra in gastric or stool specimens for pulmonary TB and extrapulmonary TB in children.
- Evaluation of the combinatorial benefit of multiple specimen types. There were limited data suggesting that the combination of non-invasive specimens performs comparably with traditional gastric specimens or induced sputum specimens.
- Additional operational and qualitative research to determine the best approach to less-invasive specimen collection.
- Implementation studies on a method of suction for nasopharyngeal aspiration that is appropriate for low-skill or low-resource environments.
- Extensive operational research into the use of stool as a diagnostic specimen in terms of integration into normal diagnostic clinical pathways, definition of laboratory protocols that successfully balance ease of implementation and diagnostic performance, and the impact of stool testing on patient-important outcomes. There is a dearth of qualitative research identifying child and family preferences for and acceptability of comparative diagnostic approaches.
- Identification of an improved reference standard that accurately defines TB disease in children and in paucibacillary specimens because sensitivity of all available diagnostics is suboptimal.
- Development of new tools that correctly diagnose a higher proportion of child TB cases. Ideally, the new tools will be rapid, affordable, feasible, and acceptable to children and their parents.
- Comparison of different tests, including Xpert MTB/RIF and Xpert Ultra, to determine which tests (or strategies) yield superior diagnostic accuracy. The preferred study design is one in which all participants receive all available diagnostic tests or are randomly assigned to receive a particular test. Studies should include children and HIV-positive people. Future research should acknowledge the concern associated with culture as a reference standard, and should consider ways to address this limitation.
- Development of rapid point-of-care diagnostic tests for extrapulmonary TB. Research groups should focus on developing diagnostic tests and strategies that use readily available clinical specimens such as urine, rather than specimens that require invasive procedures for collection.
- Operational research to ensure that tests are used optimally in settings of intended use.
- Evaluation of the diagnostic accuracy of Truenat (MTB, MTB Plus and MTB-RIF) in specific patient populations such as PLHIV, former TB patients for pulmonary TB and extrapulmonary TB in adults and children.

TB-LAMP

- Evaluation of diagnostic algorithms in different epidemiological and geographical settings and patient populations.
- Conducting of more rigorous studies with higher quality reference standards (including multiple specimen types and extrapulmonary specimens) to improve confidence in specificity estimates.
- Determination of training needs, and assessments of competency and quality.
- Gathering of more evidence on the impact on TB treatment initiation, morbidity and mortality.
- Performance of country-specific cost–effectiveness and cost–benefit analyses of targeted TB-LAMP use in different programmatic settings.
- Meeting the Standards for Reporting Diagnostic Accuracy Studies (STARD) for future studies.[19]

First-line LPA

- Development of improved understanding of the correlation between the detection of resistance-conferring mutations using culture-based DST and patient outcomes.
- Review of evidence to confirm or revise different critical concentrations used in culture-based DST methods.
- Determination of the limit of detection for LPA in detecting heteroresistance.
- Determination of needs for training, assessing competency and ensuring quality assurance.
- Gathering of more evidence on the impact on mortality of initiating appropriate treatment for MDR-TB.
- Meeting the STARD for future diagnostic studies.
- Performance of country-specific cost–effectiveness and cost–benefit analyses of LPA use in different programmatic settings.

Second-line LPA

- Development of improved understanding of the correlation between the detection of resistance-conferring mutations with phenotypic DST results and with patient outcomes.
- Development of improved knowledge of the presence of specific mutations detected with SL-LPA correlated with minimum inhibitory concentrations for individual drugs within the classes of fluoroquinolones and SLIDs.
- Determination of the limit of detection of SL-LPA for the detection of heteroresistance.
- Gathering of more evidence on the impact of MTBDR*sl* on appropriate MDR-TB treatment initiation and mortality.
- Strongly encourage that future studies follow the recommendations in the STARD *(28)* statement to improve the quality of reporting.
- Performance of country-specific cost–effectiveness and cost–benefit analyses of the use of SL-LPA in different programmatic settings.

LF-LAM

- Development of simple, more accurate tests based on LAM detection, with the potential to be used for HIV-negative populations.
- Evaluation of the use of LF-LAM in PLHIV without signs and symptoms of TB.
- Evaluation of the use of LF-LAM in children and adolescents with HIV.
- Evaluation of the combination of parallel use of LF-LAM and rapid qualitative CD4 cell count systems.
- Undertaking of implementation research into the acceptance, scale-up and impact of LF-LAM in routine clinical settings.
- Undertaking of qualitative research on user perspectives of LF-LAM for feasibility, accessibility and equity issues.

---

19 See http://www.equator-network.org/reporting-guidelines/stard/.

- Undertaking of implementation research on LF-LAM integrated into HIV care packages.
- Evaluation of the performance of LF-LAM as the HIV epidemic evolves and more people on treatment with viral load suppression are hospitalized.
- Evaluation of the cost–effectiveness of LF-LAM.
- Evaluation of other rapid LAM-based tests such as FujiLAM.

References

1 Global tuberculosis report 2019 (WHO/CDS/TB/2019.15). Geneva: World Health Organization; 2019 (https://www.who.int/tb/publications/global_report/en/, accessed 26 May 2020).

2 Implementing the End TB strategy: the essentials. Geneva: World Health Organization; 2015 (https://www.who.int/tb/publications/2015/end_tb_essential.pdf, accessed 26 May 2020).

3 Line probe assays for drug-resistant tuberculosis detection: interpretation and reporting guide for laboratory staff and clinicians. Geneva: Global Laboratory Initiative; 2018 (http://www.stoptb.org/wg/gli/assets/documents/LPA_test_web_ready.pdf, accessed 26 May 2020).

4 Report for WHO: non-inferiority evaluation of Nipro NTM+MDRTB and Hain GenoType MTBDRplus V2 line probe assays. Geneva: Foundation for Innovative New Diagnostics; 2015 (http://www.finddx.org/wp-content/uploads/2016/04/LPA-report_noninferiority-study_oct2015.pdf, accessed 26 May 2020).

5 Feasey NA, Banada PP, Howson W, Sloan DJ, Mdolo A, Boehme C et al. Evaluation of Xpert MTB/RIF for detection of tuberculosis from blood samples of HIV-infected adults confirms Mycobacterium tuberculosis bacteremia as an indicator of poor prognosis. J Clin Microbiol. 2013;51(7):2311–6 (https://pubmed.ncbi.nlm.nih.gov/23678061/, accessed 26 May 2020).

6 Automated real-time nucleic acid amplification technology for rapid and simultaneous detection of tuberculosis and rifampicin resistance: Xpert MTB. Geneva: World Health Organization; 2013 (https://apps.who.int/iris/handle/10665/112472, accessed 1 June 2020).

7 Rapid implementation of the Xpert MTB/RIF diagnostic test: technical and operational 'How-to'; practical considerations. Geneva: World Health Organization; 2011 (https://apps.who.int/iris/bitstream/handle/10665/44593/9789241501569_eng.pdf?sequence=1, accessed 1 June 2020).

8 WHO meeting report of a technical expert consultation: non-inferiority analysis of Xpert MTB/RIF Ultra compared to Xpert MTB/RIF (WHO/HTM/TB/2017.04). Geneva: World Health Organization; 2017 (https://apps.who.int/iris/bitstream/handle/10665/254792/WHO-HTM-TB-20;jsessionid=52D5C956DADE369AE677BE443C4DF574?sequence=1, accessed 15 December 2019).

9 Molbio: Our products [website]. (http://www.molbiodiagnostics.com/products-listing.php, accessed 11 June 2020).

10 Tuberculosis prevalence surveys: a handbook. Geneva: World Health Organization; 2011 (https://www.who.int/tb/advisory_bodies/impact_measurement_taskforce/resources_documents/thelimebook/en/, accessed 1 February 2020).

11 Schünemann HJ, Oxman AD, Brozek J, Glasziou P, Jaeschke R, Vist GE et al. Grading quality of evidence and strength of recommendations for diagnostic tests and strategies. Bmj. 2008;336(7653):1106–10 (https://pubmed.ncbi.nlm.nih.gov/18483053/, accessed 1 June 2020).

12 Xpert MTB/RIF assay for the diagnosis of pulmonary and extrapulmonary TB in adults and children. Policy update. Geneva, World Health Organization: 2013 (https://www.who.int/tb/publications/xpert-mtb-rif-assay-diagnosis-policy-update/en/, accessed 5 June 2020).

13 Molecular assays intended as initial tests for the diagnosis of pulmonary and extrapulmonary TB and rifampicin resistance in adults and children: rapid communication. Geneva: World Health Organization; 2020 (https://apps.who.int/iris/bitstream/handle/10665/330395/9789240000339-eng.pdf, accessed 5 June 2020).

14 Ling DI, Zwerling AA, Pai M. GenoType MTBDR assays for the diagnosis of multidrug-resistant tuberculosis: a meta-analysis. Eur Respir J. 2008;32(5):1165–74 (https://erj.ersjournals.com/content/32/5/1165, accessed 1 June 2020).

15 Molecular line probe assays for rapid screening of patients at risk of multidrug-resistant tuberculosis (MDR-TB): policy statement. Geneva: World Health Organization; 2011 (https://www.who.int/tb/laboratory/line_probe_assays/en/, accessed 1 June 2020).

16 Nathavitharana RR, Cudahy PG, Schumacher SG, Steingart KR, Pai M, Denkinger CM. Accuracy of line probe assays for the diagnosis of pulmonary and multidrug-resistant tuberculosis: a systematic review and meta-analysis. Eur Respir J. 2017;49(1):1601075.

17 Rapid diagnosis of tuberculosis brochure. Nehren, Germany: Hain Lifescience; 2015 (http://www.hain-lifescience.de/uploadfiles/file/produkte/mikrobiologie/mykobakterien/tb_eng.pdf, accessed 1 June 2020).

18 Gikalo MB, Nosova EY, Krylova LY, Moroz AM. The role of eis mutations in the development of kanamycin resistance in Mycobacterium tuberculosis isolates from the Moscow region. J Antimicrob Chemother. 2012;67(9):2107–9.

19 Alere Determine™ TB LAM Ag: Rapid rule-in TB-HIV co-infection [website]. Abbott; 2019 (https://www.alere.com/en/home/product-details/determine-tb-lam.html, accessed 26 May 2020).

20 Brennan PJ. Structure, function, and biogenesis of the cell wall of *Mycobacterium tuberculosis*. Tuberculosis. 2003;83(1–3):91–7 (https://pubmed.ncbi.nlm.nih.gov/12758196/, accessed 26 May 2020).

21 Shah M, Hanrahan C, Wang ZY, Dendukuri N, Lawn SD, Denkinger CM et al. Lateral flow urine lipoarabinomannan assay for detecting active tuberculosis in HIV-positive adults. Cochrane Database of Syst Rev. 2016;(5)(https://pubmed.ncbi.nlm.nih.gov/27163343/, accessed 26 May 2020).

22 The use of lateral flow urine lipoarabinomannan assay (LF-LAM) for the diagnosis and screening of active tuberculosis in people living with HIV: policy guidance. Geneva: World Health Organization; 2015 (https://www.who.int/tb/publications/use-of-lf-lam-tb-hiv/en/, accessed 26 May 2020).

23 Peter J, Theron G, Chanda D, Clowes P, Rachow A, Lesosky M et al. Test characteristics and potential impact of the urine LAM lateral flow assay in HIV-infected outpatients under investigation for TB and able to self-expectorate sputum for diagnostic testing. BMC Infect Dis. 2015;15(1)(https://bmcinfectdis.biomedcentral.com/articles/10.1186/s12879-015-0967-z, accessed 26 May 2020).

24 Peter JG, Theron G, van Zyl-Smit R, Haripersad A, Mottay L, Kraus S et al. Diagnostic accuracy of a urine lipoarabinomannan strip-test for TB detection in HIV-infected hospitalised patients. Eur Respir J. 2012;40(5):1211–20 (https://erj.ersjournals.com/content/40/5/1211, accessed 26 May 2020).

25 Peter JG, Zijenah LS, Chanda D, Clowes P, Lesosky M, Gina P et al. Effect on mortality of point-of-care, urine-based lipoarabinomannan testing to guide tuberculosis treatment initiation in HIV-positive hospital inpatients: a pragmatic, parallel-group, multicountry, open-label, randomised controlled trial. Lancet. 2016;387(10024):1187–97 (http://dx.doi.org/10.1016/s0140-6736(15)01092-2, accessed 26 May 2020).

26 Bjerrum S, Kenu E, Lartey M, Newman MJ, Addo KK, Andersen AB et al. Diagnostic accuracy of the rapid urine lipoarabinomannan test for pulmonary tuberculosis among HIV-infected adults in Ghana–findings from the DETECT HIV-TB study. BMC Infect Dis. 2015;15(1)(http://dx.doi.org/10.1186/s12879-015-1151-1, accessed 26 May 2020).

27 Lateral flow urine lipoarabinomannan assay (LF-LAM) for the diagnosis of active tuberculosis in people living with HIV. Policy update. Geneva: World Health Organization; 2019 (https://www.who.int/tb/publications/2019/diagnose_tb_hiv/en/, accessed 5 June 2020).

28 Bossuyt P, Reitsma J, Bruns D, Gatsonis C, Glasziou P, Irwig L et al. STARD 2015: an updated list of essential items for reporting diagnostic accuracy studies. BMJ 351: h5527. 2015.

Annex 1: Guideline development methods

Methods used to develop World Health Organization guidelines

To develop new or update existing guidelines for methods and tools to diagnose tuberculosis (TB), the Global TB Programme commissions systematic reviews on the performance or use of the tool or method in question. A systematic review provides a summary of the current literature on diagnostic accuracy or user aspects, for the diagnosis of TB or the detection of anti-TB drug resistance in adults or children (or both) with signs and symptoms of TB.

The certainty of the evidence is assessed consistently for documented evidence using the Grading of Recommendations Assessment, Development and Evaluation (GRADE) approach. GRADE produces an overall quality assessment (or certainty) of evidence and a framework for translating evidence into recommendations. The certainty of the evidence is rated as high, moderate, low or very low. These four categories imply a gradient of confidence in the estimates. Even if a diagnostic accuracy study is of observational design, it would initially be considered high-quality evidence in the GRADE approach.[20]

In addition, the Global TB Programme commissions systematic reviews to collect evidence in the field of resource use (i.e. cost and cost–effectiveness), as well as end-user perspectives on particular diagnostic tests or interventions. This evidence-to-recommendation process will inform domains such as feasibility, accessibility, equity and end-user values.

If systematic review evidence is unavailable or is scarce, the potential subsequent effects can be modelled for both diagnostic accuracy as well as cost and cost–effectiveness. For instance, the prevalence of the disease in question, combined with the sensitivity and specificity of a certain test, can be used to estimate the number of false positives and false negatives in a population. Similarly, data on expenditures and cost–effectiveness ratios can be estimated and modelled, based on economical and epidemiological data. Finally, qualitative evidence on the end-user perspective of using a particular test may be generated through end-user interviews if data are scarce in the public domain.

Following a systematic review, the Global TB Programme convenes a Guideline Development Group (GDG) meeting to review the collected evidence. The GDG is made up of external experts whose central task is to develop evidence-based recommendations. The GDG also performs the important task of finalizing the scope and key questions of the guideline in PICO (i.e. population, intervention, comparator and outcomes) format.

This group should be established early in the guideline development process, once the Steering Group has defined the guideline's general scope and target audience, and has begun drafting the key questions. The GDG should be composed of relevant technical experts; end-users, such as programme managers and health professionals, who will adopt, adapt and implement the guideline; representatives of groups most affected by the guideline's recommendations, such as service users and representatives of disadvantaged groups; experts in assessing evidence and developing guidelines informed by evidence; and other technical experts as required (e.g. a health economist or an expert on equity, human rights and gender).[21]

---

20 Schünemann HJ, Oxman AD, Brozek J, Glasziou P, Jaeschke R, Vist GE et al. Grading quality of evidence and strength of recommendations for diagnostic tests and strategies. Bmj. 2008;336(7653):1106–10 (https://pubmed.ncbi.nlm.nih.gov/18483053/, accessed 1 June 2020).

21 Handbook for guideline development. Geneva: World Health Organization; 2014 (https://www.who.int/publications/guidelines/handbook_2nd_ed.pdf?ua=1, accessed 12 June 2020).

Recommendations are developed based on consensus among GDG members, where possible. When it is not possible to reach consensus, a vote is taken. When a draft guideline is developed by a WHO steering committee, it is reviewed initially by GDG members and subsequently by an External Review Group (ERG). The ERG is made up of individuals interested in the subject, and may include the same categories of specialists as the GDG. When the ERG reviews the final guideline, its role is to identify any errors or missing data, and to comment on clarity, setting, specific issues and implications for implementation – not to change the recommendations formulated by the GDG.[21]

Formulation of the recommendations

Evidence is synthesized and presented in GRADE evidence tables. The evidence to decision (EtD) framework is used subsequently to facilitate consideration of the evidence and development of recommendations in a structured and transparent manner. Finally, recommendations are developed based on consensus among GDG members where possible. If it is not possible to reach consensus, then voting takes place. Decisions on the direction and strength of the recommendations are also made using the EtD framework.

Factors that influenced the direction and strength of a recommendation in this guideline were:

- priority of a problem;
- test accuracy;
- balance between desirable and undesirable effects;
- certainty of:
  - evidence of test accuracy;
  - evidence on direct benefits and harms from the test;
  - management guided by the test results;
  - link between test results and management;
- confidence in values and preferences and their variability;
- resource requirements;
- cost–effectiveness;
- equity;
- acceptability; and
- feasibility.

These factors are discussed below.

Priority of a problem

The GDG considers whether the overall consequences of a problem (e.g. increased morbidity, mortality and economic effects) are serious and urgent. The global situation is considered and available data reviewed. In most cases, the problem must be serious and urgent to be considered by a GDG.

Test accuracy

The pooled sensitivity and specificity presented in the GRADE evidence profile is assessed. Preferably and if available the review includes studies with both microbiological reference standards (culture) as well as composite reference standards (e.g. in children and in patients with extrapulmonary TB).

Balance between desirable and undesirable effects

Under this component, GDG members are asked to judge the anticipated benefits and harms from the test in question, including direct effects of the test (e.g. benefits such as faster diagnosis, and harms such as adverse effects from administration of the test). In addition, the possible subsequent effects of the test must be included; for instance, effects of treatment after a positive diagnosis (cure or decrease in mortality), and the effect of no treatment or further testing after a negative test result.

Evidence, ideally retrieved from systematic reviews of randomized controlled trials (RCTs) of the test, should inform the GDG of these downstream effects. If evidence from RCTs is not available, diagnostic accuracy studies can be used. In the latter, true positive and true negative diagnosed cases are taken as benefits, whereas false positive and false negative cases are taken as harms.

Certainty of the evidence

Certainty of the evidence of test accuracy is judged scored on a scale from very low, via low and moderate, to high. Certainty of the evidence on direct benefits and harms from the test are assessed and scored in a similar way.

Certainty of management

For certainty of patient management being guided by the test results, the GDG focuses on whether the management would be any different, should it be guided by the test results.

For certainty of the link between test results and management, the panel assesses how quickly and effectively test results can transfer to management decisions.

Confidence in values and preferences and their variability

The value of the test to improve diagnosis and its impact on patient care is evaluated and scored with the help of evidence from qualitative research. The impact on notification and, moreover, the ability of the test to increase case notification is also evaluated and scored, taking into account the entire diagnostic cascade, including, for example, issues related to feasibility of implementation, rate of use, staff's confidence in test results and turnaround time of results.

Resource requirements

In relation to resource requirements, the following questions are answered:

- How large are the resource requirements for test implementation?
- What is the certainty of the evidence about resource requirements?
- Does the cost–effectiveness of the intervention favour the intervention or the comparison?

Cost–effectiveness

Available evidence on cost–effectiveness is evaluated and scored.

Equity

GDG members consider whether implementing the tool or method will positively or negatively affect access to health care (e.g. will it be possible to implement the test in distinct levels of health care or through self-administration, or are there other ways to make the tools or method available to all levels of the health care system).

Acceptability

In terms of acceptability, the panel considers whether the tool or method will be acceptable by all relevant stakeholders, such as health workers, health managers and patients.

Feasibility

The GDG considers how feasible it is to implement a tool or method in various settings. Aspects such as training and refresher training needs, hands-on time, biosafety requirements, time to results, service and maintenance, calibration, and effect on diagnostic algorithms are all taken into account in the final score.

For more details on the transition from evidence to recommendations, see **Web Annex 3. Evidence to decision tables**.

Management of conflict of interest

Before being invited to be a GDG member, each potential GDG member is asked to submit a completed declaration of interests (DOI) form and provide a curriculum vitae (CV). In addition, an abbreviated and focused internet search is performed "to identify any obvious public controversies or interests that may lead to compromising situations for WHO and the expert concerned". Members of the steering committee evaluate a potential member's CV, DOI and information retrieved from the internet to determine whether there are, or may be, conflicts of interest (COI) and, if so, whether these require a management plan. COI management is based on the WHO guidelines for DOI for experts,[22] one-on-one consultation with a member of the Ethics Team from the WHO Office of Compliance, Risk Management and Ethics, and the WHO *Handbook for guideline development*.[23]

Both financial and non-financial interests are considered. A "significant" COI would include:

- "intellectual bias", when an individual may have repeatedly taken a public position on an issue under review, which may affect the individual's objectivity and independence in the global policy development process;
- involvement in research or the publication of materials related to the issue under review; and
- financial interest above US$ 5000.

For obvious reasons, developers of any assay are never involved in the process of policy development.

[22] Declaration of interests for WHO experts – forms for submission. Geneva: World Health Organization; 2019 (https://www.who.int/about/declaration-of-interests/en/, accessed 12 June 2020).

[23] Handbook for guideline development. Geneva: World Health Organization; 2014 (https://www.who.int/publications/guidelines/handbook_2nd_ed.pdf?ua=1, accessed 12 June 2020).

Web annexes

Web Annex 1. List of studies included in systematic review

Web Annex 1.1 Molecular assays as initial tests

Web Annex 1.2 TB-LAMP

Web Annex 1.3 FL-LPA

Web Annex 1.4 SL-LPA

Web Annex 1.5 LF-LAM

Web Annex 2. GRADE profiles

Web Annex 2.1 GRADE profiles molecular assays

Web Annex 2.2 GRADE profiles FL-LPA

Web Annex 2.3 GRADE profiles SL-LPA

Web Annex 2.4 GRADE profiles LF-LAM

Web Annex 3. Evidence to decision tables

Web Annex 3.1 Evidence to decision tables molecular assays

Web Annex 3.2 Evidence to decision tables TB-LAMP

Web Annex 3.3 Evidence to decision tables FL-LPA

Web Annex 3.4 Evidence to decision tables SL-LPA

Web Annex 3.5 Evidence to decision tables LF-LAM

Web Annex 4. Evidence synthesis and analysis

Web Annex 4.1 Impact of diagnostic test Xpert MTB/RIF on patient-important outcomes for tuberculosis: a systematic review

Web Annex 4.2 Xpert MTB/RIF and Xpert Ultra for detecting active tuberculosis in adults with signs and symptoms of pulmonary TB: an updated systematic review

Web Annex 4.3 Xpert MTB/RIF and Xpert Ultra for detecting active tuberculosis in adults with signs and symptoms of extrapulmonary TB: an updated systematic review

Web Annex 4.4 Xpert MTB/RIF and Xpert Ultra for detecting active tuberculosis in children: an updated systematic review

Web Annex 4.5 Report on the diagnostic accuracy of the Molbio Truenat Tuberculosis and Rifampicin-Resistance assays in the intended setting of use

Web Annex 4.6 Systematic literature review of economic evidence for molecular assays intended as initial tests for the diagnosis of pulmonary and extrapulmonary TB in adults and children

Web Annex 4.7 Report on user perspectives on Xpert testing: results from qualitative research

Web Annex 4.8 Drug concentrations used in culture-based DST SL-LPA

Web Annex 4.9 LF-LAM for detecting active tuberculosis in people living with HIV: an updated systematic review

Web Annex 4.10 Economic evaluations of LF-LAM for the diagnosis of active tuberculosis in HIV-positive individuals: an updated systematic review

Web Annex 4.11 User perspectives on TB-LAM for the diagnosis of active tuberculosis: results from qualitative research

For further information, please contact:

World Health Organization
20, Avenue Appia CH-1211 Geneva 27 Switzerland
Global TB Programme
Web site: www.who.int/tb

World Health Organization

9789240007307
9 789240 007307