26 January 2017

World Health Organization

Proposal for a WHO Model List of Essential In Vitro Diagnostics *(or the EDL)*

Proposed by the Department of Essential Medicines and Health Products, World Health Organization and submitted to the WHO Expert Committee on the Selection and Use of Essential Medicines for comments and recommendations.
Focal point: Dr Francis Moussy (moussyf@who.int)

Action by the Expert Committee on the Selection and Use of Essential Medicines: The committee is invited to consider the proposal to develop a List of Essential In Vitro Diagnostics and to formulate recommendations on the best course of action. A set of questions is provided on page 6 to facilitate the formulation of recommendations.

Goal: To develop a List of Essential In Vitro Diagnostics to improve access to in vitro diagnostics (IVDs)[1] and to guide safe and rationale use of medicines from the WHO Model lists of Essential Medicines.

Rationale: The WHO Model lists of Essential Medicines, often referred to as the Essential Medicines Lists (EMLs) (for adults and children) are some of the most recognised and valued products from WHO. These lists play an important role in health care policy, in particular in

[1] (note: the term In Vitro Diagnostics (IVDs) is used to avoid confusion with other diagnostic tools such as MRI, X-rays,…..). 'In Vitro Diagnostic (IVD) medical device' refers to a medical device, whether used alone or in combination, intended by the manufacturer for the in-vitro examination of specimens derived from the human body solely or principally to provide information for diagnostic, monitoring or compatibility purposes. http://www.imdrf.org/docs/ghtf/final/sg1/technical-docs/ghtf-sg1-n071-2012-definition-of-terms-120516.pdf

LMICs as the lists guide medicines procurement decisions from ministries of health and NGOs. A similar list of Essential IVDs has the potential to do the same for diagnostic tests.

The crucial role of in vitro diagnostics (IVDs) has become widely acknowledged in a diverse range of areas including case finding, treatment, test of cure, outbreak response, surveillance, diseases elimination certification and vaccine efficacy evaluation. Of particular importance is the need for IVDs to ensure the safe and rational use of medicines in the EMLs. For example, Schroeder et al [2] identified 147 laboratory tests that are used for diagnosis, monitoring the medication efficacy, or monitoring the medication toxicity of medicines included in the EMLs.

In addition to improved patient care, the following benefits can be expected from a list of Essential IVDs.

- Help countries to better prioritize their laboratory testing and infrastructure needs.
- Increase affordability by facilitating bulk (e.g. through the WHO Bulk procurement scheme) and advanced purchasing.
- Contribute to reducing antimicrobial resistance by facilitating the appropriate use of antimicrobials through increased access and use of relevant IVDs.
- Help countries to be better prepared for future outbreaks by increasing laboratory capacity at all levels of the health care system.
- Facilitate the development of new IVDs through identification of IVD priority needs to inform IVD developers, industry and funders.
- Help regulatory agencies to prioritize IVDs for review (in particular in countries with limited resources).

A list of Essential In Vitro Diagnostics has already been discussed internally within WHO for some time. For example, such a list was discussed in April 2016 among WHO diagnostics experts and there was strong support for this work across the organization. Other organizations are also acknowledging and expressing this need. A recent example is the June 2016 paper entitled "Time for a model list of essential diagnostics" published in the New England Journal of Medicine[2]. This paper spurred numerous articles also supporting the creation of an EDL. External organizations[3,4] have expressed interest in supporting the creation of the EDL. It appears that now is the ideal time to take action and develop an EDL.

Discussion on the scope of the EDL: At this conceptual stage, it is important to define the scope of the EDL. As a guiding principle, the list should aim for a high public health impact and be implementable by countries, in particular for those with less resources.

[2] Schroeder LF, Guarner J, Elbireer A, Castle PE, Amukele TK. Time for a Model List of Essential Diagnostics. New England Journal of Medicine. June 30;374(26):2511-4

[3] http://www.ghtcoalition.org/pdf/Global-Health-Advocacy-Coalition-Calls-for-WHO-to-Develop-Essential-Diagnostics-List.pdf

[4] http://www.treatmentactiongroup.org/sites/default/files/Essential%20Diagnostic%20List_WHO_Final.pdf

- **IVDs vs diagnostic tools**: An EDL could potentially include both IVDs and other diagnostic tools such as imaging technologies (e.g. X-Rays). In order to facilitate the timely development of the EDL and in consideration of the large number of IVDs for potential inclusion, we propose to focus the list on IVDs.
- **Starting with an EDL to complement the EMLs**: In order to further focus the EDL, we propose to include IVDs in the list that are required for the safe and appropriate use of EML medicines. Further prioritization could potentially reduce the number of IVDs for inclusion in the list by initially focusing on specific conditions where diagnostics have a clear impact on disease diagnosis and management, and where strong support programmes are already in place (e.g. TB, malaria, HIV). This approach may help reduce the risk of countries procuring IVDs without adequate systems in place to ensure proper use. WHO recommendations for the use of IVDs for such diseases are already made according to a rigorous assessment process (see example in Annex 2). Conversely, it could also be argued that including IVDs for priority conditions with weaker support programmes could stimulate the development of stronger support programmes. Further discussions will be needed to determine the best course of action.
- **IVDs for outbreak identification and response**: New IVDs are already available and others are soon to be available for diagnosis of infectious threats, e.g. Ebola and Zika, that do not currently have medicine treatments. There is evidently a need for countries to know which IVDs for such conditions should be procured and included in national lists. A discussion is needed to determine if, and when, such IVDs ought to be included in the EDL and if the same process for inclusion in the list can be applied in such instances. Provisions for fast track mechanisms may be needed to enable an urgent response to an outbreak. As WHO has already listed a number of IVDs for Ebola and Zika in its Emergency and Assessment Use Listing (EUAL), IVDs for these two diseases could be considered for the EDL.
- **Targeted level of the health care system and implications for the IVDs in the list**. IVDs are needed at all levels of the health care system. It is therefore likely that IVDs in the EDL will cover all levels. Different IVDs for the same condition may be recommended depending on the context where they will be used. This will need to be reflected in the list. In contrast to medicines, additional details on the characteristics of the IVDs will be needed to take into account the target users (e.g. trained laboratory staff in a reference laboratory versus minimally trained health worker in a health post). This additional information may include intended use, infrastructure level requirement, target user, sample type and volume, sample handling, performance, and other user requirements. The level of detail will need to be balanced and put in perspective with the objectives and intended use of the list.
- **Presentation of the EDL:** The list should be presented in a format that is both informative and practical for countries. There are several possible presentations/formats. For example: IVDs can be presented by disease, by technology (commercial culture, molecular testing, immunoassays clinical biochemistry), by health care system level (primary, secondary, tertiary), or most likely a combination

of the above, with potential inclusion of necessary specifications. As previously indicated, health care level will have a very strong influence on suitable IVDs and their particular characteristics. This should be reflected in the list. Presentation of the EDL based on health care level may best meet the needs of countries and also facilitate the procurement of IVD technologies that can be used for multiple diseases (e.g. PCR for HIV, hepatitis C, etc.). It may be useful for the EDL to be web based with in-built filters that can be applied to obtain printable lists that meet local needs in countries. It is anticipated that the EDL will have numerous web links to existing WHO disease programmes as well as the WHO Diagnostic prequalification programme in order to take advantage of the extensive relevant information already available on the WHO web site.

Proposed general approach for developing the EDL: The process for establishing the List of Essential In Vitro Diagnostics initially builds on current EML processes and the WHO Expert Committee on the Selection and Use of Essential Medicines. The formulation of recommendations at the next meeting of the expert committee, in March 2017, will help guide next steps. It is anticipated that the development of the EDL will then become the responsibility of another committee (focused on IVDs), which would work in close collaboration with the WHO Expert Committee on the Selection and Use of Essential Medicines, and, where appropriate, with other relevant committees advising disease programmes. After reaching consensus on the scope of the EDL and the process for its development, a stepwise implementation approach is envisioned.

We propose to proceed by first introducing a limited number of needed diagnostics on the EDL and then expanding it. The methodology used for this initial list will be reviewed and adapted as needed. We will initially focus on a few conditions where diagnostics have a clear impact on the diagnosis and management of a disease. We plan to include diseases that already have strong programmes in place for diagnosis and management and where recommendations of diagnostics followed the formal process of the WHO Guidelines Review Committee including the use of the GRADE process. A description of the various steps used to recommend IVDs for tuberculosis is given in Annex 2 as an example. In this case, it is expected that these IVDs may be ready to be included in the list without requiring additional significant review of evidence. Based on Annex 1, part A, the diseases with IVDs that are likely ready to be included in the initial EDL without additional significant review of evidence include TB, malaria/G6PD, HIV and hepatitis B & C.

Once we have demonstrated feasibility with IVDs for TB, malaria/G6PD, HIV and hepatitis B & C, the List of Essential In Vitro Diagnostics will be later expanded to include IVDs for other diseases from the Essential Medicines Lists. It will also be expanded for priority diseases that do not currently have medicines but where diagnostics still play a critical role for case finding, disease management and/or surveillance. Annex 1, Part B shows other priority conditions that have weaker support programmes and/or are lacking GRC compliant recommendations for IVD use. Conditions included in this list are syphilis, diabetes, breast

cancer (HER2), chronic myelogenous leukemia, Ebola and Zika. We anticipate that IVDs for these conditions will be considered for the second round of the EDL with the expectation that the inclusion of these conditions in the EDL will spur development of the required support programmes. Application submissions for inclusion of IVDs for these diseases will need to include substantial information on validation studies and cost effectiveness. This information will need to be collected and assessed according to the WHO GRC procedure in place. An example of what a typical submission may look like is shown in Annex 3. Box 1 below lists the proposed information to be included in an application for inclusion or deletion of an IVD in the WHO EDL.

Box 1. Information to be included with an application for inclusion or deletion of An IVD in the WHO List of Essential IVDs

1. Summary statement of the proposal for inclusion, change or deletion
2. Name of the focal point in WHO submitting the application (where relevant)
3. Name of the organization(s) consulted and/or supporting the application
4. Generic name (and brand name when only one IVD is suitable/available) of the IVD as appropriate
5. Whether listing is requested as a specific IVD from a single manufacturer or as a type of IVD technology. If listing is requested for a specific IVD from a single manufacturer, justification will be required.
6. Information supporting the public health relevance (epidemiological information on disease burden, assessment of current use, target population)
7. Characteristics of the IVD (intended use, product presentation, infrastructure level requirement, target user, sample type and volume, sample handling, performance, time to results, storage conditions, operation conditions, shipping requirements, training requirements, associated equipment, throughput, need for maintenance, connectivity,…)
8. Use of the IVD details (as part of a diagnostic testing algorithm; reference to existing WHO and other clinical guidelines; need for special diagnostic or treatment facilities and skills)
9. Summary of laboratory evaluation studies
10. Summary of validation studies in a variety of clinical settings:
    - Identification of clinical evidence (search strategy, systematic reviews identified, reasons for selection/exclusion of particular data)
    - Summary of available data (appraisal of quality, performance, ease of use, summary of results)
11. Summary of evidence on safety:
12. Summary of available data on comparative cost[1] and cost-effectiveness:
    - range of costs of the proposed IVD
    - comparative cost-effectiveness presented as range of cost per routine outcome (e.g. cost per case, cost per cure, cost per month of treatment, cost per case prevented, cost per clinical event prevented, or, if possible and relevant, cost per quality-adjusted life year gained)
13. Summary of regulatory status of the IVD (in country of origin, and preferably in other countries as well)
14. Proposed (new/adapted) text for the WHO List of Essential IVDs.

[1] All cost analyses should specify the source of the price information.

Timeline:

Key Activity	2017	2018	2019
1.Develop approach to create the EDL.	X		
2. Present to the WHO Expert Committee on the Selection and Use of Essential Medicines.	X		
3. Develop the EDL.	X X	X X X	X
4. The first EDL is developed and endorsed.			X
5. Publish and advocate the EDL and start working on the EDL Version 2.			X X
6. Impact assessment			X

Questions to the Expert Committee on the Selection and Use of Essential Medicines:

1. Can you comment on the need to develop a WHO List of Essential In Vitro Diagnostics?
2. Is the EML a good template to follow or should other models be considered?
3. Can you comment on the initial proposed priority areas (TB, malaria, HIV and Hepatitis B & C) for the first iteration of the EDL?
4. Do you have any suggestions about the proposed process for developing the list?
5. Do you have any other suggestions to facilitate the initiation of the EDL?

For comments or further explanations, please contact:

Dr Francis Moussy
Responsible Officer, List of Essential In Vitro Diagnostics
Innovation, Access and Use
Department of Essential Medicines and Health Products
World Health Organization
(moussyf@who.int)

ANNEX 1: List of diagnostics that are recommended (or suggested) by WHO for selected diseases and the process used for these recommendations.

Diseases	Diagnostics
PART A	**IVDs that have been recommended after a formal assessment following GRC approved guidelines**
Tuberculosis	- LED Microscopy (4 recommendations). 2011 - Commercial culture and DST (to be updated) - Xpert MTB/RIF assay (7 recommendations ranging from strong to conditional recommendations with quality of evidence ranging from high to very low). 2013 - lateral flow urine lipoarabinomannan assay (LF-LAM) for the diagnosis and screening of active tuberculosis in people living with HIV (3 recommendations ranging from strong to conditional recommendations with low quality of evidence for all three). 2015 - molecular line probe assays for the detection of resistance to second-line anti-tuberculosis drugs (2 conditional recommendations, with moderate to low certainty in the evidence for test accuracy). 2016 - loop-mediated isothermal amplification (TB-LAMP) for the diagnosis of pulmonary tuberculosis (2 conditional recommendation, very low quality of evidence for both). 2016 - molecular line probe assays for the detection of resistance to isoniazid and rifampicin. (1 conditional recommendation, moderate certainty in the evidence for the test's accuracy). 2016 - tests for latent TB infection (TST or IGRA) details can be found at: http://apps.who.int/iris/bitstream/10665/162712/1/9789241508612_eng.pdf?ua=1&ua=1 and http://www.who.int/tb/areas-of-work/laboratory/policy_statements/en/
Malaria	- Microscopy - Rapid diagnostic tests (RDTs) Details can be found at: http://www.who.int/malaria/publications/diagnostic_testing/en/ http://apps.who.int/iris/bitstream/10665/162441/1/9789241549127_eng.pdf?ua=1&ua=1
G6PD	- fluorescent spot test (FST) - Rapid diagnostic tests (RDTs) http://apps.who.int/iris/bitstream/10665/162441/1/9789241549127_eng.pdf?ua=1&ua=1 http://www.who.int/malaria/mpac/mpac-march2015-erg-g6pd.pdf

HIV	- Antibody (anti-HIV). Screening (for HIV infection): infants over 18 months of age, children, adolescents, adults. Immunoassay. Consolidated guidelines on HIV Testing Services, July 2015. - Combined antibody/core antigen (anti-HIV/cAg). Screening (for HIV infection). Immunoassay. Consolidated guidelines on HIV Testing Services, July 2015. - Qualitative virological (HIV RNA, DNA, or US p24 Ag). Screening infants under 18 months of age (for HIV infection). Nucleic acid testing. WHO recommendations on the diagnosis of HIV infection in infants and children, 2010. Consolidated guidelines on the use of antiretroviral drugs for treating and preventing HIV infection, 2016. - Quantitive virological. (HIV RNA). Monitoring (of response to antiviral treatment). Nucleic acid testing. Consolidated guidelines on the use of antiretroviral drugs for treating and preventing HIV infection, 2013. Consolidated guidelines on the use of antiretroviral drugs for treating and preventing HIV infection, 2016. - Quantitative immunological. Monitoring (of opportunistic infections). CD4 testing. Consolidated guidelines on the use of antiretroviral drugs for treating and preventing HIV infection, 2013. Consolidated guidelines on the use of antiretroviral drugs for treating and preventing HIV infection, 2016. Immunoassay formats include: rapid diagnostic tests (RDTs), other simple assays (particle agglutination), enzyme immunoassays (EIAs), chemiluminescence immunoassays (CLIA), electrochemiluminescence immunoassay (ECL) **WHO list of prequalified *in vitro* diagnostic products:** http://www.who.int/diagnostics_laboratory/evaluations/PQ_list/en/ **Consolidated guidelines on HIV Testing Services, July 2015.** http://www.who.int/hiv/pub/guidelines/hiv-testing-services/en/ **WHO recommendations on the diagnosis of HIV infection in infants and children, 2010.** http://www.who.int/hiv/pub/paediatric/diagnosis/en/ **Consolidated guidelines on the use of antiretroviral drugs for treating and preventing HIV infection, 2013.** http://www.who.int/hiv/pub/guidelines/arv2013/download/en/ **Consolidated guidelines on the use of antiretroviral drugs for treating and preventing HIV infection, 2016.** http://www.who.int/hiv/pub/arv/arv-2016/en/
Hepatitis B & C	- serological assays (in either RDT or lab-based immunoassays format) (HBsAg, Anti-HCV antibody) - Nucleic Acid Tests (HBV DNA, HCV RNA)

	Guidelines on hepatitis B and C testing: http://apps.who.int/iris/bitstream/10665/251330/1/WHO-HIV-2016.23-eng.pdf?ua=1
PART B	**IVDs that that still need to go through a formal assessment following GRC approved guidelines and/or that do not yet benefit from a strong support programme for implementation.**
Syphilis	- Dark-field microscopy - Direct fluorescent antibody (DFA) test - Nucleic acid amplification tests for T. pallidum - Non-treponemal serological tests - Treponemal serological tests In *Laboratory diagnosis of sexually transmitted infections, including human immunodeficiency virus (2013).* http://who.int/reproductivehealth/publications/rtis/9789241505840/en/
diabetes	- Glucose test - HbA1c can be used as a diagnostic test for diabetes providing that stringent quality assurance tests are in place and assays are standardised to criteria aligned to the international reference values, and there are no conditions present which preclude its accurate measurement. Details can be found at: http://apps.who.int/iris/bitstream/10665/204871/1/9789241565257_eng.pdf?ua=1
breast cancer	- microscopy - Immunohistochemistry (IHC) staining (for hormone receptor testing and HER2) - Fluorescent in situ hybridization (FISH)
chronic myelogenous leukemia	- microscopy - Karyotyping - Fluorescent in situ hybridization (FISH) - Quantitative PCR
Ebola	- RDT - RT-PCR
Zika	- ELISA - RT-PCR

In addition, more than 60 IVD products are listed in the WHO list of prequalified in vitro diagnostic products (HIV serology, HIV virological assays, CD4 assays, hepatitis B and C, malaria RDTs) (http://www.who.int/diagnostics_laboratory/evaluations/PQ_list/en/)

ANNEX 2

WHO's process for developing policies on TB diagnostics

http://apps.who.int/iris/bitstream/10665/162712/1/9789241508612_eng.pdf?ua=1&ua=1

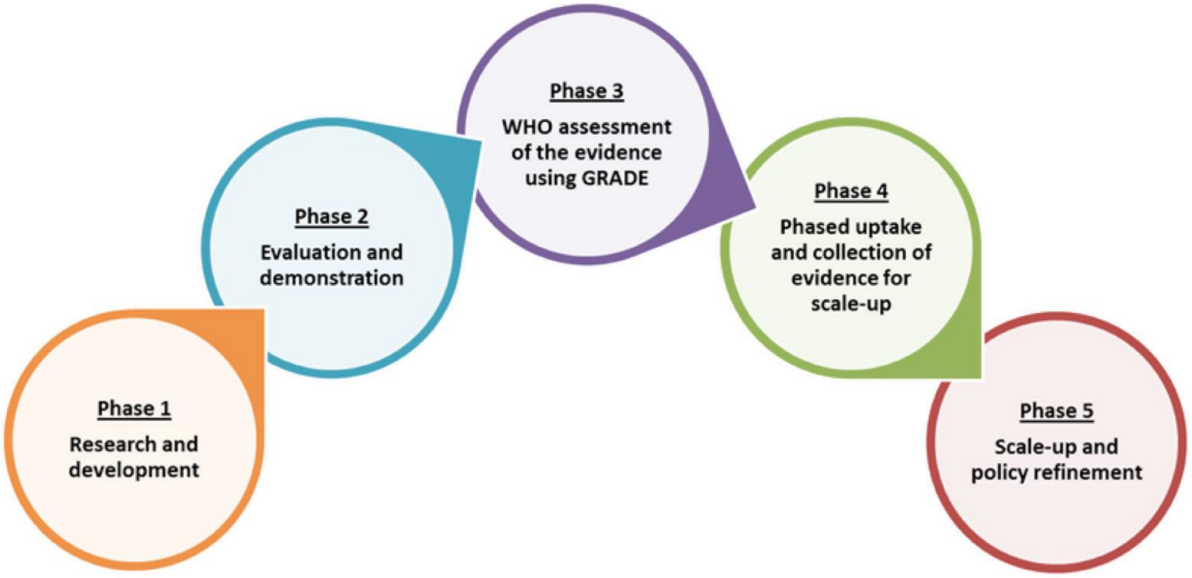


Figure 1. The phases of TB diagnostics development and assessment for WHO recommendation using the GRADE (Grading of Recommendations Assessment, Development and Evaluation) process

Phase 1: Research and development

This is the discovery phase for new diagnostic technologies. It includes a feasibility assessment aimed at developing a final version of the technology (the design-locked product) that can be used in subsequent evaluations. The developers of diagnostic tests are encouraged to engage in early discussions with WHO to ensure that the new technology will be appropriate for the end-users. Priority target product profiles (TPP) for new diagnostics, developed following a consensus building process, are described in the TPP meeting report.5 http://www.who.int/tb/publications/tpp_report/en/

Phase 2: Evaluation and demonstration

Controlled laboratory trials, or evaluation studies, are often conducted at the level of reference laboratories and should be performed in three to five sites in different countries that have a high burden of TB and varying epidemiology in terms of TB, HIV infection and MDR-TB. Data generated from these initial trials are often used for product registration with global or national regulatory authorities, or both. It should be noted that WHO is not a regulatory authority. Following the initial validation of a technology in evaluation studies, field demonstration studies are required in 5–10 sites to validate the specifications and performance characteristics in the intended settings of use. These studies also should be undertaken in high-burden countries in which the burden of TB, HIV infection and MDRTB varies epidemiologically.

Phase 3: Evidence assessment by WHO

During this phase, WHO evaluates a dossier about new technologies or new indications for an existing technology, provided the technology is not intended for use only in a specific country. The dossier contains data from phase 1 and phase 2 (*Figure 1*).

Data on new technologies (or new indications for the use of technologies already recommended by WHO) from controlled evaluations and field demonstration studies, as well as from operational research studies and cost–effectiveness analyses, are systematically assessed by an independent group of experts convened by WHO. The expert group uses the GRADE process as a systematic, structured framework to evaluate the diagnostic accuracy of new tools and their effect on patients and public health. The expert group synthesizes the evidence using systematic reviews and meta-analyses (where possible), in accordance with WHO's standards. For generic versions of technologies already recommended by WHO, comparative data are needed to determine their non-inferiority (that is, their equivalence) in performance. Non-inferiority studies must be multicentre, blinded and conducted independently by at least three members of WHO's TB Supranational Reference Laboratory Network. The end result of the process is the development of policy guidance based on the outcome of the expert group's consensus meeting together with input from the Strategic and Technical Advisory Group for Tuberculosis (STAG-TB) for dissemination to Member States and other stakeholders.

Phase 4: Phased uptake and evidence for scale-up

The new technology is implemented in routine TB services including in high burden TB and HIV settings. WHO subsequently evaluates operational issues associated with implementation, as well as the cost effectiveness of a new technology, by engaging with early implementers in different countries and settings.

Phase 5: Scale-up and policy refinement

WHO's process for policy development is a dynamic mechanism, and diagnostic policies are regularly reviewed (every 3 to 5 years); during these reviews, additional evidence is evaluated, allowing initial guidance to be updated and refined for further country-level scale-up.

See http://apps.who.int/iris/bitstream/10665/250586/1/9789241511261-eng.pdf or http://apps.who.int/iris/bitstream/10665/112472/1/9789241506335_eng.pdf?ua=1 for detailed examples.

ANNEX 3: Example of a possible format and requirements for submission of an IVD for inclusion in the WHO List of Essential In Vitro Diagnostics

PROPOSAL FOR THE INCLUSION OF

THE URINE LIPOARABINOMANNAN (LAM) STRIP-TEST (DETERMINE®-TB ALERE, USA)

IN THE WHO LIST OF ESSENTIAL IN VITRO DIAGNOSTICS

SUBMITTED BY THE WHO Global TB Programme

January 2017

Proposal for the inclusion of lateral flow lipoarabinomannan assay (LF-LAM) for the diagnosis and screening of active tuberculosis in people living with HIV in the WHO list of essential IVDs.

1. Summary statement of the proposal for inclusion, change or deletion

Tests based on the detection of mycobacterial lipoarabinomannan (LAM) antigen in urine have emerged as potential point-of-care tests for tuberculosis (TB). Owing to suboptimal sensitivity, the urinary LAM assays are unsuitable as general screening tests for TB. However, unlike traditional diagnostic methods, they demonstrate improved sensitivity in HIV-TB co-infection which further increases with low CD4 counts. The urine lipoarabinomannan (LAM) strip-test (Determine®-TB Alere, USA) potentially can be used as a rule-in test for TB in patients with advanced HIV-induced immunosuppression, and facilitate the early initiation of antituberculous treatment in them.

The full WHO policy guidance on this IVD with details on the review of evidence is available at http://apps.who.int/iris/bitstream/10665/193633/1/9789241509633_eng.pdf?ua=1&ua=1

2. Name of the focal point in WHO submitting the application (where relevant)

Dr Chris Gilpin,
WHO Global TB Programme

A WHO steering group was established for the guideline development process and comprises Chris Gilpin, Karin Weyer, Alexei Korobitsyn, Haileyesus Getahun, Alberto Matteelli, Fuad Mirzayev, Wayne Van Gemert (WHO Global TB Programme) and Meg Doherty (WHO HIV/AIDS Department). The WHO steering group was setup for scoping the guidelines and to oversee the evidence retrieval. The WHO steering group was responsible for selecting members for the Guideline Development group and External Review group, managing declarations of interest, and for organizing the Guideline Development meeting.

3. Name of the organization(s) consulted and/or supporting the application

WHO Global TB Programme
WHO HIV/AIDS Department
Guideline Review Committee, WHO

4. Generic name or brand name of the IVD as appropriate.

Urine lipoarabinomannan (LAM) strip-test (Determine®-TB Alere, USA)

***5*. Whether listing is requested as a specific IVD from a single manufacturer or as a type of IVD technology. If listing is requested for a specific IVD from a single manufacturer, please justify.**

Listing required as a specific IVD from a single manufacturer. This IVD from Alere is the only Independently validated and commercially available rapid urine lipoarabinomannan (LAM) strip-test.

6. Information supporting the public health relevance (epidemiological information on disease burden, assessment of current use, target population)

Key global priorities for tuberculosis (TB) care and control include improving case-detection and detecting cases earlier, including cases of smear-negative disease which are often associated with co-infection with the human immunodeficiency virus (HIV) and young age. In 2015, an estimated 1.1 million (11%) of the 10.4 million people who developed TB worldwide were HIV-positive. The African Region accounted for 78% of the estimated number of HIV-positive incident TB cases. Globally,

people living with HIV are 29 times more likely to develop TB disease than those who are HIV-negative. Beginning in the 1980s, the HIV epidemic led to a major upsurge in TB cases and TB mortality in many countries, especially in southern and eastern Africa. TB occurs early in the course of HIV infection and shortens survival. Many people infected with HIV in developing countries develop TB as the first manifestation of AIDS.

If accurate, LF-LAM may benefit people with HIV-TB coinfection who have difficulty producing sputum, making it difficult to detect TB using traditional diagnostic methods. HIV-positive patients with TB disease may be missed for the following reasons: they may not be able to provide sufficient and high quality sputum specimens; sputum bacillary load is typically low in these patients; and a substantial proportion of these patients cannot produce sputum at all or have extrapulmonary TB without pulmonary TB.

7. Characteristics of the IVD (intended use, product presentation, infrastructure level requirement, target user, sample type and volume, sample handling, performance, time to results, storage conditions, operation conditions, shipping requirements, training requirements, associated equipment, throughput, need for maintenance, connectivity,…)

The urine lateral flow lipoarabinomannan (LF-LAM) assay is a commercially available strip test for active TB (Alere Determine™ TB LAM Ag, Alere Inc, Waltham, MA, USA). The test is performed manually by applying 60 µL of urine to the Determine™ TB LAM Ag test strip and incubating at room temperature for 25 minutes. The strip is then inspected by eye, and the intensity of any visible bands on the test strip are graded by comparing with those in a series of bands on a manufacturer-supplied reference card. Prior to January 2014, this reference card included five bands (grade 1 representing a very low intensity band to grade 5 representing a high/dark intensity band). After January 2014, the manufacturer revised the reference bands to contain only 4 grades, such that the band intensity for the new grade 1 corresponded to the previous reference card band intensity for grade 2.

Figure 1: Alere Determine™ TB LAM Ag test

(A) Alere Determine™ TB LAM Ag tests. To the sample pad (white pad marked by the arrow symbols) 60 µL of urine is applied and visualized bands are read 25 minutes later. (B) Reference card accompanying test strips to 'grade' the test result and determine positivity (33). Copyright © [2014] [Alere Inc]: reproduced with permission.

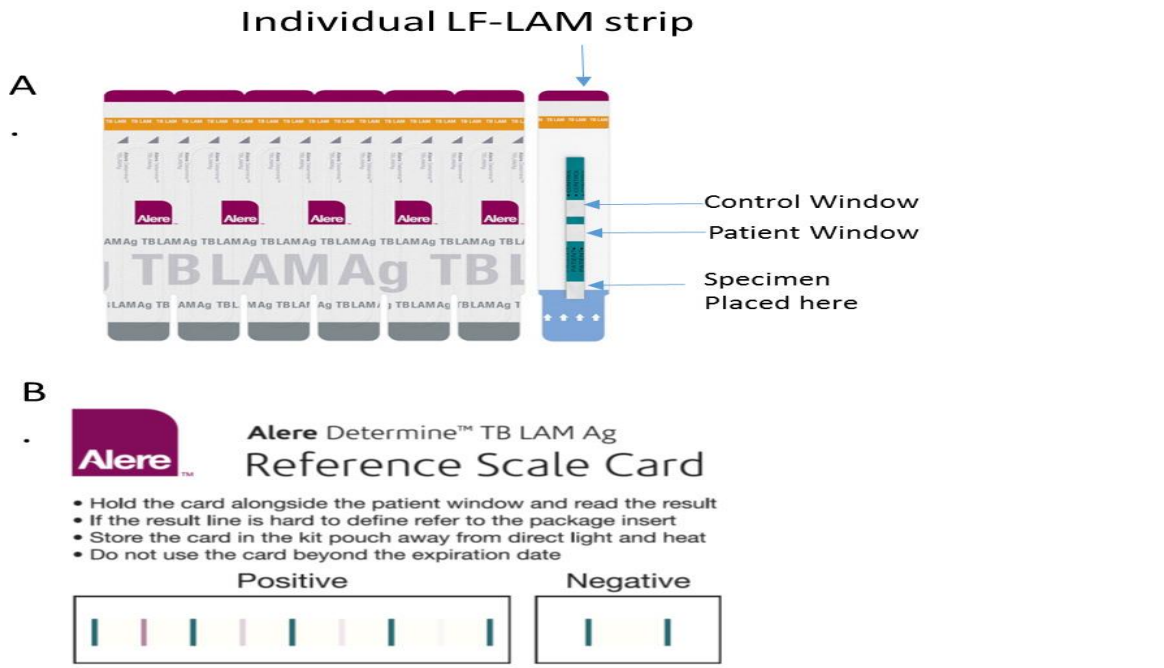
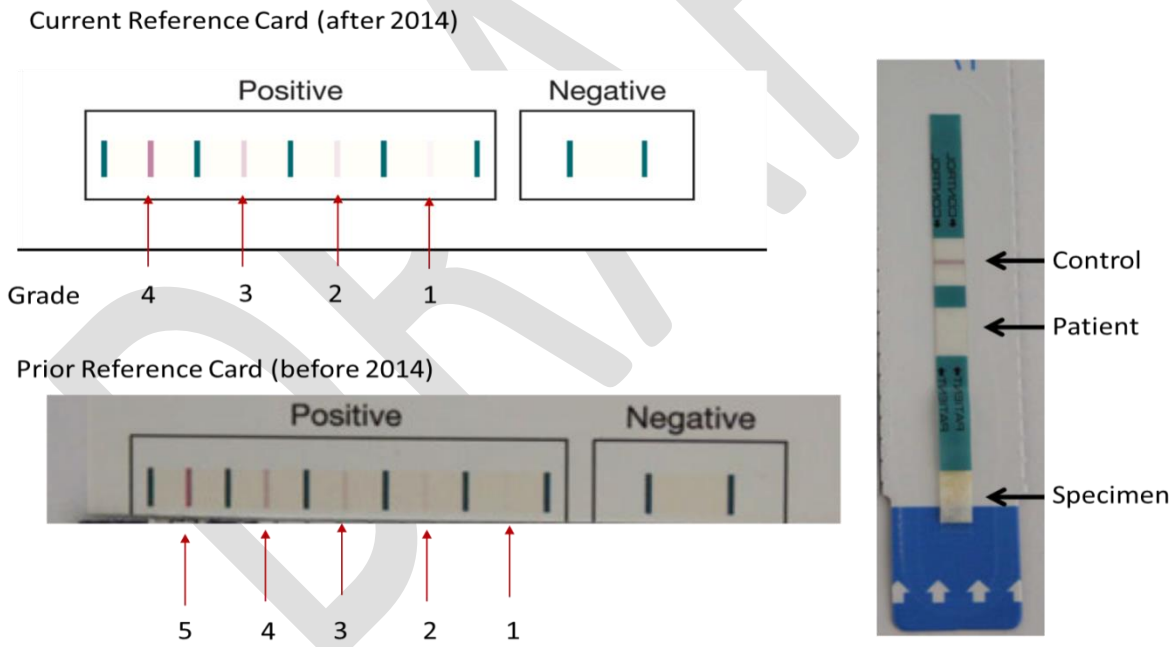


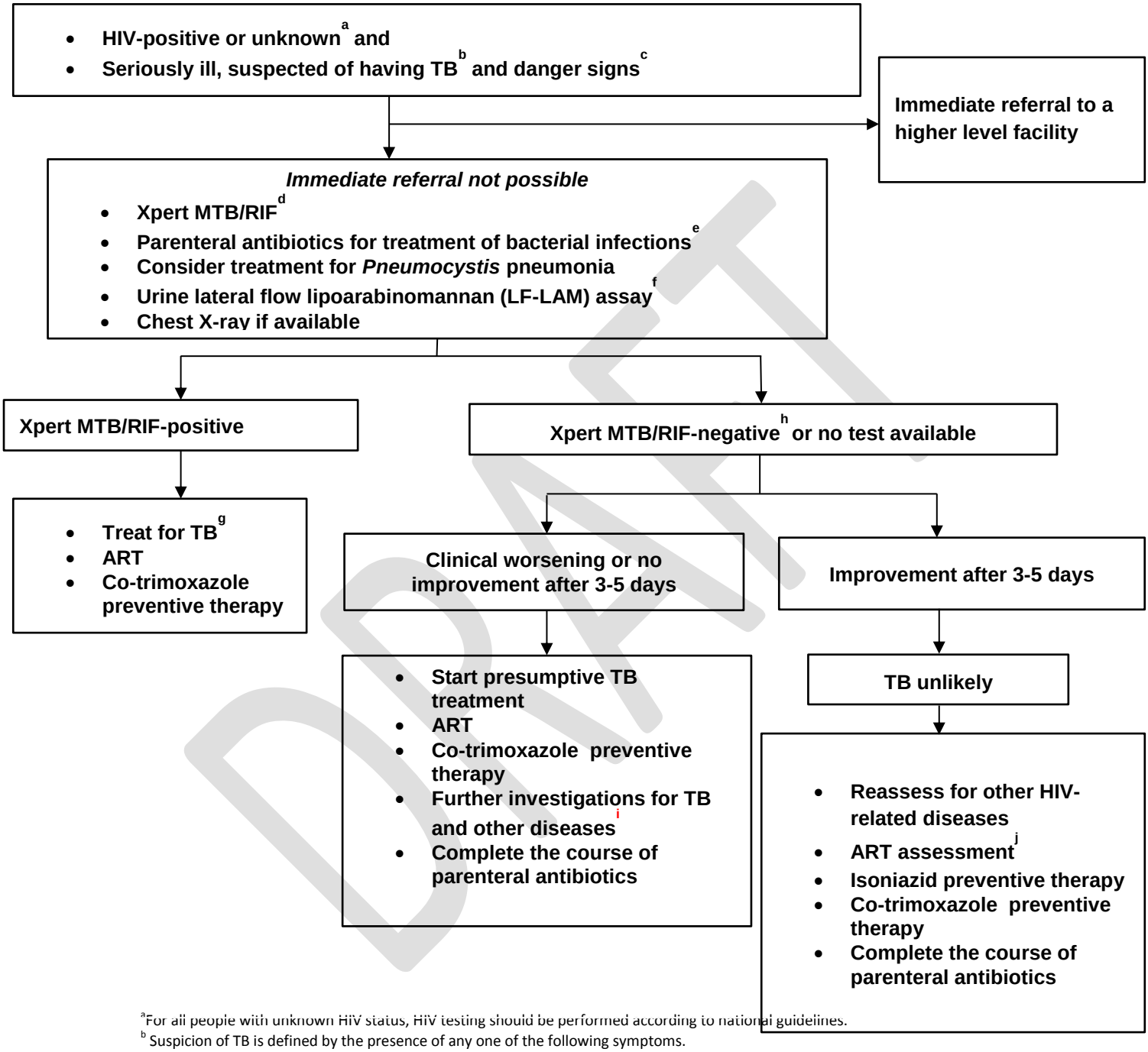
Figure 2. Grading of Determine TB Ag assay



LF-LAM is simple in use, adds only limited workload to the health workers, and does require minimum training for them. LF-LAM test cards must be stored at 2-30°C until the expiration date. Kit components are stable until expiration date when handled and stored as directed. Devices that have become wet or if the packaging has become damaged should not be used.

8. Use of the IVD details (as part of a diagnostic testing algorithm; reference to existing WHO and other clinical guidelines; need for special diagnostic or treatment facilities and skills)

Figure 3: Algorithm for managing people living with HIV and suspected of having TB (seriously ill) http://apps.who.int/iris/bitstream/10665/208825/1/9789241549684_eng.pdf?ua=1 (Annex 15)



[a] For all people with unknown HIV status, HIV testing should be performed according to national guidelines.

[b] Suspicion of TB is defined by the presence of any one of the following symptoms.
– For adults and adolescents living with HIV: current cough, fever, weight loss or night sweats.
– For children living with HIV: poor weight gain, fever, current cough or history of contact with a TB case.

[c] Danger signs include any one of the following: respiratory rate >30 per minute, temperature >39°C, heart rate >120 beats per minute and unable to walk unaided.

[d] For people suspected of having extrapulmonary TB, extrapulmonary specimens should be obtained for Xpert MTB/RIF (cerebrospinal fluid, lymph nodes and other tissues: Xpert MTB/RIF has low sensitivity for pleural fluid and data are limited for stool, urine or blood). If Xpert MTB/RIF is not available, conduct AFB microscopy. AFB-positive is defined as at least one positive smear and AFB-negative as two or more negative smears. Refer the specimen for TB culture where feasible.

[e] Antibiotics with broad-spectrum antibacterial activity (except fluoroquinolones) should be used.

[f] The LF-LAM assay may be used to assist in diagnosing active TB in peripheral settings among both in-and out-patients who are seriously ill, regardless of CD4 count. Whenever possible, a positive LF-LAM should be followed up with a confirmation test such as Xpert MTB/RIF.

While awaiting results of other confirmatory tests, clinicians could consider initiating TB treatment immediately based on the positive LF-LAM and their clinical judgment.

[g] If Xpert MTB/RIF shows rifampicin resistance, treatment for multidrug-reistant TB should be initiated. If the person is considered at low risk for rifampicin resistance, a second Xpert MTB/RIF test should be performed on a fresh specimen. Collect and refer a sample for culture and additional drug sensitivity testing.

[h] If Xpert MTB/RIF shows negative results, the test can be repeated using a fresh specimen.

i Further investigations for TB include chest X-ray, clinical assessment, a repeat Xpert MTB/RIF using a fresh specimen and culture. If extrapulmonary TB is suspected, extrapulmonary specimens should be obtained and sent for culture and abdominal ultrasound may be performed.

[j] ART should be recommended for all adults, regardless of CD4 cell count or clinical stage.

9. Summary of validation studies in a variety of clinical settings:

Identification of clinical evidence (search strategy, systematic reviews identified, reasons for selection/exclusion of particular data)

In June 2015, a Guideline Development Group was convened by WHO's Global TB Programme to assess the data on for the use of LF-LAM. WHO commissioned a systematic review on the use of LF-LAM for the diagnosis and screening for TB among person with HIV as well as a review of the affordability and cost effectiveness of LF-LAM.

In accordance with WHO's standards for assessing evidence when formulating policy recommendations, the Grading of Recommendations Assessment, Development and Evaluation (GRADE) system[5] was used for the evidence synthesis process to provide a systematic, structured framework for evaluating both the accuracy of the test and the test's impact on patients and public health. The evaluations used the GRADE system to determine the quality of the evidence and provide information on the strength of the recommendations using a priori questions (that is, PICO questions) agreed by the Guideline Development Group. 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; the Intervention; the Comparator; and the Outcome.

A systematic review was conducted according to the standards outlined by the Cochrane Collaboration in the Cochrane handbook. A comprehensive search of the following databases was performed on 2 February 2015, without date or language restrictions: Cochrane Infectious Diseases Group Specialized Register; PubMed; EMBASE; lSI Web of Knowledge; MEDION; LILACS; BIOSIS; and SCOPUS. Searches of the metaRegister of Controlled Trials (mRCT) and the search portal of the WHO International Clinical Trials Registry Platform (www.who.int/trialsearch) were also performed to identify ongoing trials and ProQuest Dissertations & Theses A&I to identify relevant dissertations.

Where feasible, meta-analysis was used to summarize the results of independent studies, and these results have been displayed as forest plots. Where meta-analysis was not feasible due to heterogeneity, the evidence has been presented in a narrative synthesis.

The systematic reviewers prepared GRADE evidence profiles for each PICO question. GRADE evidence profiles were prepared to assess the accuracy of LF-LAM for the diagnosis of active TB disease in adults with HIV with signs or symptoms of TB and to assess the accuracy of LF-LAM used as a screening test for active TB disease in adults with HIV irrespective of signs or symptoms of TB. The choice of an optimal reference standard is critical to assess the accuracy of diagnostic tests, since the reference standard is used to determine the presence or absence of the target condition.

[5] Schünemann HJ et al. Grading quality of evidence and strength of recommendations for diagnostic tests and strategies. *BMJ*, 2008, 336:1106–1110.

The target condition was active TB disease, which included both pulmonary and extrapulmonary TB. The studies were required to diagnose TB using at least one of the following two reference standards:

> **A microbiological reference standard** for TB was defined as a positive *M. tuberculosis* culture or NAAT and not TB was defined as a negative M. tuberculosis culture and NAAT[6] (if performed).
>
> **A Composite reference standard** that includes at least one of the following components, *M. tuberculosis* culture, NAAT, smear, or clinical findings. TB was defined as a positive culture or NAAT or positive smear or clinical decision to start TB treatment, and, after at least one month of follow-up, the patient was diagnosed as having TB. Not TB was defined as a negative culture and NAAT (if performed), no TB treatment given, and resolution of signs and symptoms at follow-up.

Using the GRADE framework, results for sensitivity and specificity were used as proxy measures for outcomes seen as important to patients; these outcomes were based on the relative importance or impact of false-positive and false-negative results. Poor sensitivity would result in *false-negative* results so that patients with TB would be missed, and this could have negative consequences in terms of time to treatment initiation, morbidity, mortality and transmission of disease. Poor specificity would result in *false-positive* results so that patients without TB would be prescribed unnecessary treatment.

Rates for true positives, true negatives, false positives and false negatives were calculated based on pretest probabilities – that is, an assumed prevalence of TB of 1% among HIV infected persons irrespective of signs and symptoms for TB, an assumed prevalence of TB of 10% among HIV infected persons with symptoms suggested of TB and an assumed prevalence of TB of 30% among seriously ill, hospitalised HIV infected persons with signs and symptoms of TB.

The evaluation of the impact on patients was based on a balance among the following values:

- *true positives* – the benefit to patients from rapid diagnosis and treatment;
- *true negatives* – the benefit to patients who would be spared unnecessary treatment; the benefit of reassurance and alternative diagnosis;
- *false positives* – the likelihood of anxiety and morbidity caused by additional testing, unnecessary treatment and possible adverse effects; possible stigma associated with a diagnosis of TB; and the chance that a false positive may halt further diagnostic evaluation;
- *false negatives* – the increased risk of morbidity and mortality, delayed treatment initiation and the continued risk of transmission of TB.

For each outcome, the quality of evidence according to GRADE was initially regarded as high since all studies were cross-sectional or cohort studies, prospectively enrolling patients suspected of having pulmonary or extrapulmonary TB. The quality of the evidence and the limitations of the studies were assessed using six GRADE criteria: (1) study design, (2) risk of bias, (3) directness, (4) inconsistency, (5) imprecision, and (6) publication or reporting bias.

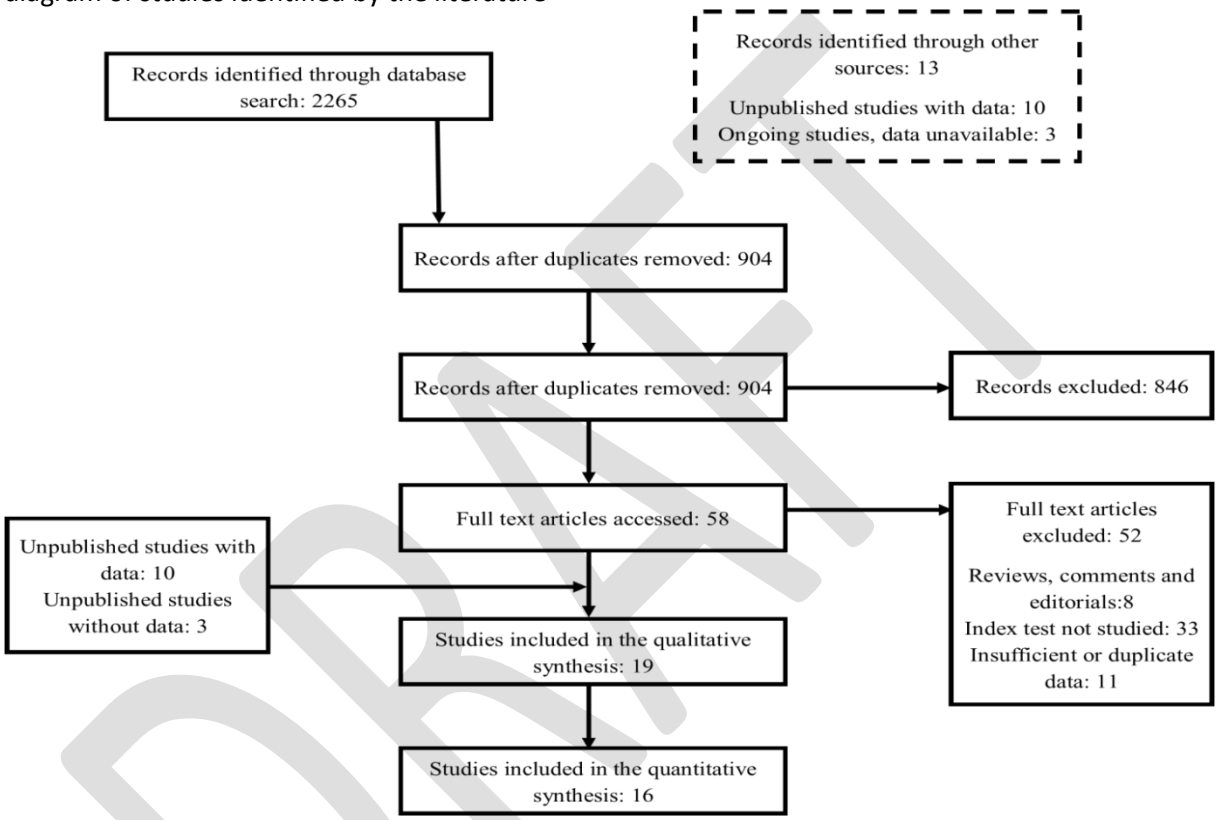
---

[6] NAATs included Enhanced Amplified Mycobacterium Tuberculosis Direct Test (E-MTD, Gen-Probe, San Diego, USA); Amplicor Mycobacterium tuberculosis Test (Amplicor, Roche Diagnostics, Basel, Switzerland); COBAS® TaqMan® MTB Test (Roche Diagnostics); GenoType MTBDRplus (HAIN Lifesciences, Nehren, Germany); and Xpert® MTB/RIF assay (Cepheid, Sunnyvale, USA).

Each review used the Quality Assessment of Diagnostic Accuracy Studies (QUADAS -2) tool to appraise the studies.[7] This tool consists of four domains: patient selection, index test, reference standard, and flow and timing.

Sixteen unique studies evaluating LF-LAM among adults were identified through the search. The 16 studies involved 6588 HIV-infected patients including 1789 (27%) with a microbiological diagnosis of TB. All studies were performed in low- or middle-income countries.

Figure 4: Selection of studies evaluating the accuracy of LF-LAM for the diagnosis of active TB: flow diagram of studies identified by the literature



Summary of available data (appraisal of quality, performance, ease of use, summary of results) Using LF-LAM to diagnose active TB in adults with HIV

Details on the review of evidence is available at:
http://apps.who.int/iris/bitstream/10665/193633/1/9789241509633_eng.pdf?ua=1&ua=1

Of the total 16 studies, seven studies (44%) were identified that used LF-LAM in HIV-positive patients thought to have TB because of signs and symptoms. All seven studies involving 3126 patients (1220 [39%] with TB disease) evaluated LF-LAM using grade 1, while six studies involving 3037 patients (1163 [38%] with TB) additionally reported results using grade 2. The median CD4 cell count in these studies ranged from 71 to 210. Six (86%) of the seven studies were conducted either exclusively or largely in an inpatient setting. Four studies (57%) included participants from outpatient settings.

[7] Whiting PF et al. QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies. *Annals of Internal Medicine*, 2011, 155:529–536.

One of the seven studies applied an approach of testing all patients in an inpatient setting. Given that all included patients had some symptoms, the purpose of testing in this study was considered to be "diagnosis", rather than screening, even though patients were not enrolled in the study on the basis of specific TB symptoms. Four of the seven studies were judged to be at high risk of bias for patient selection, which may have contributed to a decrease in the pooled sensitivity. Three of the seven studies (43%) were judged to be at high risk of bias for the reference standard, which may have contributed to a decrease in the pooled specificity.

Overall accuracy

Six studies evaluated LF-LAM (grade 2) for TB diagnosis. With respect to a microbiological reference standard, pooled sensitivity and specificity of LF-LAM for detecting TB was 44% (95% credible interval [CrI], 31- 60%) and 92% (95% CrI, 83-96%). With respect to a composite reference standard, pooled sensitivity decreased to 28% (95%CrI, 13-51%) and specificity increased to 97% (95% CrI, 93-99%). The increased specificity and reduced sensitivity with respect to a composite reference standard was explained by the broad definition of the reference standard that, in addition to the microbiological criterion, included clinical criteria, i.e., decision to start treatment, and, after at least one month follow-up, the patient was diagnosed as having TB.

By health care setting

The pooled sensitivity of LF-LAM (grade 2) from the six studies was higher among inpatients (54% [95% CrI, 43- 67%]) than among outpatients (21% [95% CrI, 12- 34%], three studies), and the pooled specificity was lower, 90% (95% CrI, 79- 95%) versus 97% (95%CrI, 87-99%), respectively.

By CD4 threshold

Five studies evaluated LF-LAM (grade 2) for TB diagnosis in patients stratified by CD4 count. The pooled sensitivity of LF-LAM increased as the degree of immunodeficiency increased, from 15% (95% CrI, 8-27%) in patients with CD4 cell count >200 cells/μL to 49% (95% CrI, 34- 66%) in patients with CD4 count ≤200, to 56% (95%CrI, 41- 70%) in patients with CD4≤100. Pooled specificity also varied less across CD4 strata but a dose-response relationship was less apparent: 96% (95% CrI, 89- 99%) for CD4>200, 92% (95% CrI, 78- 97%) for CD4>100, and 90% (95% CrI, 81- 95%) for CD4≤100.

10. Summary of evidence on safety:

LF-LAM is suitable for all levels of the health system, does not require additional laboratory equipment, biosafety and security measures and electric supply. All these make it particularly valuable at the lowest levels of health care system, i.e. point-of-care. Implementation of LF-LAM does not eliminate the need in other diagnostic tests, such as sputum smear microscopy, Xpert MTB/RIF, culture, as these tests exceed LF-LAM in diagnostic accuracy. A positive LF-LAM should be followed up with confirmation test such as Xpert MTB/RIF, LPA, and/or bacterial culture with a drug susceptibility test, where possible. LF-LAM is designed to detect Mycobacterial LAM antigen in human urine. Other samples (e.g. sputum, serum, plasma, CSF or other body fluids) or pooled urine specimens should not be used.

11. Summary of available data on comparative cost[1] and cost-effectiveness:

(range of costs of the proposed IVD, comparative cost-effectiveness presented as range of cost per routine outcome e.g. cost per case, cost per cure, cost per month of treatment, cost per case

prevented, cost per clinical event prevented, or, if possible and relevant, cost per quality-adjusted life year gained)

A review of economic evaluations of lateral flow detection of lipoarabinomannan in urine (LF-LAM) for diagnosis of active tuberculosis (TB) in HIV-infected individuals was performed. Two studies were identified, both evaluating populations in sub-Saharan Africa and with a focus on inpatient populations with CD4+ T-cell counts less than 100 cells/µL. Both studies found the addition of LF-LAM to existing diagnostic strategies based on sputum smear microscopy or Xpert MTB/RIF to be highly cost-effective across a wide array of sensitivity analyses, although addition of costs related to future HIV care in one study caused cost-effectiveness ratios to become substantially less favorable.

Incremental cost-effectiveness ratios (without inclusion of HIV care costs) ranged from $21 to $265 per disability-adjusted life year (DALY) averted for the addition of LF-LAM to existing diagnostic algorithms based on sputum smear microscopy in Uganda, from $10 to $3,162 per DALY averted when added to Xpert MTB/RIF-based algorithms in Uganda, and from $135 to $8,707 for addition of LF-LAM to existing diagnosis in South Africa. Excluding HIV care costs, the probability that addition of LF-LAM would cost less than the gross domestic product per capita of the two countries was estimated at >99.8% in one study; including HIV care costs, this probability fell to 72-77% in the second study, with the difference almost entirely reflecting the cost of HIV care for TB survivors. The most important drivers of cost-effectiveness were the specificity of LF-LAM, the prevalence of active TB in the target population, the life expectancy of TB survivors, and the costs of TB and HIV treatment.

While the ability of a positive LF-LAM result to avert mortality was varied in both studies, it was not an important driver of cost-effectiveness, as TB and HIV treatment costs dwarfed the cost of LF-LAM itself, such that patients who would die regardless of test result did not add substantially to the incremental costs nor effectiveness of LF-LAM. Ultimately, the cost-effectiveness of LF-LAM in both analyses closely followed the cost-effectiveness of TB (and, when considered, HIV) treatment, with additional consideration for the excess cost of treating false positives.

12. Summary of regulatory status of the IVD (in country of origin, and preferably in other countries as well)

Not available

13. Proposed (new/adapted) text for the WHO List of Essential IVDs.
WHO Recommendations

Urine lipoarabinomannan (LAM) strip-test (Determine™-TB Alere, USA)

1. LF-LAM **should not** be used for the diagnosis of TB, ***except as specifically described below for persons with HIV with low CD4 counts or who are seriously ill***.[1] (**strong recommendation, low quality of evidence**).

2. LF-LAM **may be used to assist** in the diagnosis of TB in **HIV positive** adult *in-patients* with signs and symptoms of TB (pulmonary and/or extrapulmonary) who have a CD4 cell count less than or equal to 100, or HIV positive patients who are seriously ill[8] regardless of CD4 count or with unknown CD4 count. (**conditional recommendation; low quality of evidence**).

[1] "seriously ill" is defined based on 4 danger signs: respiratory rate> 30/min, temperature>39°C, heart rate>120/min and unable to walk unaided

Remarks

- This recommendation also applies to HIV positive adult *out-patients* with signs and symptoms of TB (pulmonary and/or extrapulmonary) who have a CD4 cell count less than or equal to 100, or HIV positive patients who are seriously ill[1] regardless of CD4 count or with unknown CD4 count, based on the generalisation of data from in-patients.

- This recommendation also applies to HIV positive children with signs and symptoms of TB (pulmonary and/or extrapulmonary) who have a CD4 cell count less than or equal to 100, or HIV positive children who are seriously ill[1] regardless of CD4 count or with unknown CD4 count, based on the generalisation of data from adults while acknowledging very limited data and concern regarding low specificity of the LF-LAM assay in children.