Official American Thoracic Society/Infectious Diseases Society of America/Centers for Disease Control and Prevention Clinical Practice Guidelines: Diagnosis of Tuberculosis in Adults and Children


¹Oregon Health & Science University, Portland, Oregon; ²Emory University School of Medicine and ³Centers for Disease Control and Prevention, Atlanta, Georgia; ⁴Denver Public Health Department, Denver, Colorado; ⁵National Jewish Health and the University of Colorado Denver, and ⁶California Department of Public Health, Richmond; ⁷St James’s Hospital, Dublin, Ireland; ⁸Francis J. Curry International TB Center, San Francisco, California; ⁹Foundation for Innovative New Diagnostics, Geneva, Switzerland; ¹⁰McGill University and McGill International TB Centre, Montreal, Canada; ¹¹University of Southampton, United Kingdom; ¹²National Jewish Health, Denver, Colorado; ¹³Vanderbilt University School of Medicine, Vanderbilt Institute for Global Health, Nashville, Tennessee, ¹⁴Wisconsin State Laboratory of Hygiene, Madison, and ¹⁵University of Arkansas for Medical Sciences, Little Rock

Background. Individuals infected with Mycobacterium tuberculosis (Mtb) may develop symptoms and signs of disease (tuberculosis disease) or may have no clinical evidence of disease (latent tuberculosis infection [LTBI]). Tuberculosis disease is a leading cause of infectious disease morbidity and mortality worldwide, yet many questions related to its diagnosis remain.

Methods. A task force supported by the American Thoracic Society, Centers for Disease Control and Prevention, and Infectious Diseases Society of America searched, selected, and synthesized relevant evidence. The evidence was then used as the basis for recommendations about the diagnosis of tuberculosis disease and LTBI in adults and children. The recommendations were formulated, written, and graded using the Grading, Recommendations, Assessment, Development and Evaluation (GRADE) approach.

Results. Twenty-three evidence-based recommendations about diagnostic testing for latent tuberculosis infection, pulmonary tuberculosis, and extrapulmonary tuberculosis are provided. Six of the recommendations are strong, whereas the remaining 17 are conditional.

Conclusions. These guidelines are not intended to impose a standard of care. They provide the basis for rational decisions in the diagnosis of tuberculosis in the context of the existing evidence. No guidelines can take into account all of the often compelling unique individual clinical circumstances.

EXECUTIVE SUMMARY

Individuals infected with Mycobacterium tuberculosis (Mtb) may develop symptoms and signs of disease (TB disease) or may have no clinical evidence of disease (latent tuberculosis infection [LTBI]). TB disease is a leading cause of infectious disease morbidity and mortality worldwide, with many diagnostic uncertainties. A task force supported by the American Thoracic Society, Centers for Disease Control and Prevention, and Infectious Diseases Society of America appraised the evidence and derived the following recommendations using the Grading, Recommendations, Assessment, Development, and Evaluation (GRADE) approach (Table 1):

Testing for LTBI

Our recommendations for diagnostic testing for LTBI are based upon the likelihood of infection with Mtb and the likelihood of progression to TB disease if infected, as illustrated in Figure 1.

- We recommend performing an interferon-γ release assay (IGRA) rather than a tuberculin skin test (TST) in individuals 5 years or older who meet the following criteria: (1) are likely to be infected with Mtb, (2) have a low or intermediate risk of disease progression, (3) it has been decided that testing for LTBI is warranted, and (4) either have a history of BCG vaccination or are unlikely to return to have their TST read (strong recommendation, moderate-quality evidence). Remarks: A TST is an acceptable alternative, especially in situations where an IGRA is not available, too costly, or too burdensome.

- We suggest performing an IGRA rather than a TST in all other individuals 5 years or older who are likely to be infected with Mtb, who have a low or intermediate risk of disease progression, and in whom it has been decided that testing for LTBI is warranted (conditional recommendation, moderate-quality evidence). Remarks: A TST is an
Table 1. Interpretation of Strong and Weak (Conditional) Recommendations

<table>
<thead>
<tr>
<th>Patients</th>
<th>Strong Recommendation</th>
<th>Weak (Conditional) Recommendation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Most individuals in this situation would want the recommended course of action, and only a small proportion would not.</td>
<td>The majority of individuals in this situation would want the suggested course of action, but many would not.</td>
<td></td>
</tr>
</tbody>
</table>

| Clinicians | Most individuals should receive the intervention. Adherence to this recommendation according to the guideline could be used as a quality criterion or performance indicator. Formal decision aids are not likely to be needed to help individuals make decisions consistent with their values and preferences. | Recognize that different choices will be appropriate for individual patients and that you must help each patient arrive at a management decision consistent with his or her values and preferences. Decision aids may be useful in helping individuals to make decisions consistent with their values and preferences. |

| Policymakers | The recommendation can be adopted as policy in most situations. | Policymaking will require substantial debate and involvement of various stakeholders. |

Testing for TB Disease

• We recommend that acid-fast bacilli (AFB) smear microscopy be performed, rather than no AFB smear microscopy, in all patients suspected of having pulmonary TB (strong recommendation, moderate-quality evidence). Remarks: False-negative results are sufficiently common that a negative AFB smear result does not exclude pulmonary TB. Similarly, false-positive results are sufficiently common that a positive AFB smear result does not confirm pulmonary TB. Testing of 3 specimens is considered the normative practice in the United States and is strongly recommended by the Centers for Disease Control and Prevention and the National Tuberculosis Controllers Association in order to improve sensitivity given the pervasive issue of poor sample quality. Providers should request a sputum volume of at least 3 mL, but the optimal volume is 5–10 mL. Concentrated respiratory specimens and fluorescence microscopy are preferred.

• We suggest that both liquid and solid mycobacterial cultures be performed, rather than either culture method alone, for every specimen obtained from an individual with suspected TB disease (conditional recommendation, low-quality evidence). Remarks: The conditional qualifier applies to performance of both liquid and solid culture methods on all specimens. At least liquid culture should be done on all specimens as culture is the gold standard microbiologic test for the diagnosis of TB disease. The isolate recovered should be identified according to the Clinical and Laboratory Standards Institute guidelines and the American Society for Microbiology Manual of Clinical Microbiology.

• We suggest performing a diagnostic nucleic acid amplification test (NAAT), rather than not performing a NAAT, on the initial respiratory specimen from patients suspected of having pulmonary TB (conditional recommendation, low-quality evidence). Remarks: In AFB smear-positive
patients, a negative NAAT makes TB disease unlikely. In AFB smear-negative patients with an intermediate to high level of suspicion for disease, a positive NAAT can be used as presumptive evidence of TB disease, but a negative NAAT cannot be used to exclude pulmonary TB. Appropriate NAAT include the Hologic Amplified Mycobacteria Tuberculosis Direct (MTD) test (San Diego, California) and the Cepheid Xpert MTB/Rif test (Sunnyvale, California).

- We recommend performing rapid molecular drug susceptibility testing for rifampin with or without isoniazid using the respiratory specimens of persons who are either AFB smear positive or Hologic Amplified MTD positive and who meet one of the following criteria: (1) have been treated for tuberculosis in the past, (2) were born in or have lived for at least 1 year in a foreign country with at least a moderate tuberculosis incidence (≥20 per 100,000) or a high primary multidrug-resistant tuberculosis prevalence (≥2%), (3) are contacts of patients with multidrug-resistant tuberculosis, or (4) are HIV infected (strong recommendation, moderate-quality evidence). Remarks: This recommendation specifically addresses patients who are Hologic Amplified MTD positive because the Hologic Amplified MTD NAAT only detects TB and not drug resistance; it is not applicable to patients who are positive for types of NAAT that detect drug resistance, including many line probe assays and Cepheid Xpert MTB/RIF.

- We suggest mycobacterial culture of respiratory specimens for all children suspected of having pulmonary TB (conditional recommendation, moderate-quality evidence). Remarks: In a low incidence setting like the United States, it is unlikely that a child identified during a recent contact investigation of a close adult/adolescent contact with contagious TB was, in fact, infected by a different individual with a strain with a different susceptibility pattern. Therefore, under some circumstances, microbiological confirmation may not be necessary for children with uncomplicated pulmonary TB identified through a recent contact investigation if the source case has drug-susceptible TB.
We suggest sputum induction rather than flexible bronchoscopic sampling as the initial respiratory sampling method for adults with suspected pulmonary TB who are either unable to expectorate sputum or whose expectorated sputum is AFB smear microscopy negative (conditional recommendation, low-quality evidence).

We suggest flexible bronchoscopic sampling, rather than no bronchoscopic sampling, in adults with suspected pulmonary TB from whom a respiratory sample cannot be obtained via induced sputum (conditional recommendation, very low-quality evidence). Remarks: In the committee members’ clinical practices, bronchoalveolar lavage (BAL) plus brushings alone are performed for most patients; however, for patients in whom a rapid diagnosis is essential, transbronchial biopsy is also performed.

We suggest that postbronchoscopic sputum specimens be collected from all adults with suspected pulmonary TB who undergo bronchoscopy (conditional recommendation, low-quality evidence). Remarks: Postbronchoscopic sputum specimens are used to perform AFB smear microscopy and mycobacterial cultures.

We suggest flexible bronchoscopic sampling, rather than no bronchoscopic sampling, in adults with suspected miliary TB and no alternative lesions that are accessible for sampling whose induced sputum is AFB smear microscopy negative or from whom a respiratory sample cannot be obtained via induced sputum (conditional recommendation, very low-quality evidence). Remarks: Bronchoscopic sampling in patients with suspected miliary TB should include bronchial brushings and/or transbronchial biopsy, as the yield from washings is substantially less and the yield from BAL unknown. For patients in whom it is important to provide a rapid presumptive diagnosis of tuberculosis (ie, those who are too sick to wait for culture results), transbronchial biopsies are both necessary and appropriate.

We suggest that cell counts and chemistries be performed on amenable fluid specimens collected from sites of suspected extrapulmonary TB (conditional recommendation, very low-quality evidence). Remarks: Specimens that are amenable to cell counts and chemistries include pleural, cerebrospinal, ascitic, and joint fluids.

We suggest that adenosine deaminase levels be measured, rather than not measured, on fluid collected from patients with suspected pleural TB, TB meningitis, peritoneal TB, or pericardial TB (conditional recommendation, low-quality evidence).

We suggest that free IFN-γ levels be measured, rather than not measured, on fluid collected from patients with suspected tuberculosis infection.
suspected pleural TB or peritoneal TB (*conditional recommendation, low-quality evidence*).

- We recommend that mycobacterial cultures be performed, rather than not performed, on specimens collected from sites of suspected extrapulmonary TB (*conditional recommendation, very low-quality evidence*). Remarks: A positive result can be used as evidence of extrapulmonary TB and guide decision making because false-positive results are unlikely. However, a negative result may not be used to exclude TB because false-negative results are exceedingly common.

- We recommend that mycobacterial cultures be performed, rather than not performed, on specimens collected from sites of suspected extrapulmonary TB (*conditional recommendation, very low-quality evidence*). Remarks: A positive result can be used as evidence of extrapulmonary TB and guide decision making because false-positive results are unlikely. However, a negative result may not be used to exclude TB because false-negative results are exceedingly common.

- We recommend that histological examination be performed, rather than not performed, on specimens collected from sites of suspected extrapulmonary TB (*conditional recommendation, very low-quality evidence*). Remarks: Both positive and negative results should be interpreted in the context of the clinical scenario because neither false-positive nor false-negative results are rare.

- We recommend one culture isolate from each mycobacterial culture-positive patient be submitted to a regional genotyping laboratory for genotyping (*strong recommendation, very low-quality evidence*).

Persons infected with *Mycobacterium tuberculosis* (*Mtb*) have a broad array of presentations, ranging from those with clinical, radiographic, and microbiological evidence of tuberculosis (TB disease) to those who are infected with *Mtb* but have no clinical evidence of TB disease (latent tuberculosis infection [LTBI]). Individuals with LTBI who have been recently exposed have an increased risk of developing TB, whereas those with remote exposure have less risk over time unless they develop a condition that impairs immunity. Operationally, recent exposure can be defined either epidemiologically (ie, as might occur in the setting of the household of an infectious case or occupational exposure) or immunologically (ie, conversion of a tuberculin skin test or interferon-γ release assay [IGRA] from negative to positive).

These clinical practice guidelines on the diagnosis and classification of tuberculosis in adults and children were prepared by a task force supported by the American Thoracic Society (ATS), the Centers for Disease Control and Prevention (CDC), and the Infectious Diseases Society of America (IDSA). Additionally, Fellows of the American Academy of Pediatrics participated in the development of these guidelines. The specific objectives of these guidelines are as follows:

- To define high- and low-risk patient populations based upon the results of epidemiological studies.
- To provide diagnostic recommendations that lead to beneficial treatments and favorable clinical outcomes.
- To describe a classification scheme for tuberculosis that is based on pathogenesis.

These guidelines target clinicians in high-resource countries with a low incidence of TB disease and LTBI, such as the United States. The recommendations may be less applicable to medium- and high-tuberculosis incidence countries. For such countries, guidance documents published by the World Health Organization (WHO) may be more suitable.

**HOW TO USE THESE GUIDELINES**

These guidelines are not intended to impose a standard of care. They provide the basis for rational decisions in the diagnostic evaluation of patients with possible LTBI or TB. Clinicians, patients, third-party payers, stakeholders, or the courts should never view the recommendations contained in these guidelines as dictates. Guidelines cannot take into account all of the often compelling unique individual clinical circumstances. Therefore, no one charged with evaluating clinicians’ actions should attempt to apply the recommendations from these guidelines by rote or in a blanket fashion. Qualifying remarks accompanying each recommendation are its integral parts and serve to facilitate more accurate interpretation. They should never be omitted when quoting or translating recommendations from these guidelines.

**METHODS**

**Committee Selection**

The criteria for committee selection were an (1) established track record in the relevant clinical or research area; (2) involvement with the ATS Assembly on Microbiology, Tuberculosis and Pulmonary Infections, the IDSA Tuberculosis Committee, or employment by the United States CDC Division of Tuberculosis Elimination; and (3) absence of disqualifying conflicts of interest. Conflicts of interest were managed according to the policies and procedures agreed upon by the participating organizations [1].

The committee was divided into subcommittees assigned to develop drafts for each of the following areas: (1) LTBI, (2) clinical and radiological aspects of TB diagnosis, (3) microbiological evaluation for TB diagnosis and detection of drug resistance.
Evidence Synthesis
Each subcommittee identified key diagnostic questions and then performed a pragmatic evidence synthesis for each question, to identify and summarize the related evidence. The subcommittees first sought studies comparing one diagnostic intervention with another and measuring clinical outcomes. Such evidence was unavailable, so the subcommittees next sought diagnostic accuracy studies. When published evidence was lacking, the collective clinical experience of the committee was used. The evidence syntheses were used to inform the recommendations. Though comprehensive, the evidence syntheses should not be considered systematic reviews of the evidence.

Developing and Grading Recommendations
Recommendations were formulated and the quality of evidence and strength of each recommendation were rated using the Grading, Recommendations, Assessment, Development and Evaluation (GRADE) approach [2, 3].

The quality of evidence is the extent to which one can be confident that the estimated effects are close to the actual effects and was rated as high, moderate, low, or very low. The quality of evidence rating derived from the quality of the accuracy studies that informed the panel’s judgments, as randomized trials and controlled observational studies were lacking. Well-done accuracy studies that enrolled consecutive patients with legitimate diagnostic uncertainty and used appropriate reference standards represented high-quality evidence; lack of these characteristics constituted reasons to downgrade the quality evidence. Normally, the quality of evidence for first-line therapy would have been factored into such quality of evidence ratings but, in this case, the quality of evidence that treatment of TB disease and LTBI improve outcomes is high quality, so the overall quality of evidence rating was determined entirely by the accuracy study.

The decision to recommend for or against an intervention was based upon consideration of the balance of desirable consequences (ie, benefits) and undesirable consequences (ie, harms, burdens), quality of the evidence, patient values and preferences, cost, resource use, and feasibility. The subcommittees used open discussion to arrive at a consensus for each of the recommendations. An open voting procedure was reserved for situations when the subcommittee could not reach consensus through discussion, but this was not needed for any recommendation.

The strength of a recommendation indicates the committee’s certainty that the desirable consequences of the recommended course of action outweigh the undesirable consequences. A strong recommendation is one for which the subcommittee is certain, whereas a conditional recommendation is one for which the subcommittee is uncertain. Uncertainty may exist if the quality of evidence is poor, there is a fine balance between desirable and undesirable consequences (ie, the benefits may not be worth the costs or burdens), the balance of desirable and undesirable consequences depends upon the clinical context, or there is variation about how individuals value the outcomes. A strong recommendation should be interpreted as the right thing to do for the vast majority of patients; a weak recommendation should be interpreted as being the right thing to do for the majority of patients, but maybe not for a sizeable minority of patients.

TUBERCULOSIS: EPIDEMIOLOGY, TRANSMISSION, AND PATHOGENESIS
A full discussion of these topics can be found in the Supplementary Materials. TB disease remains one of the major causes of morbidity and mortality in the world. The WHO estimates that 8.6 million new cases of tuberculosis occurred in 2014 and approximately 1.5 million persons died from the disease [4]. The emergence of drug-resistant tuberculosis has become apparent over the past 2 decades, and in particular, multidrug-resistant tuberculosis (MDR-TB; resistant to isoniazid and rifampin) and extensively drug-resistant tuberculosis (XDR-TB; resistant to isoniazid and rifampin, plus any fluoroquinolone and at least 1 of 3 injectable second-line drugs [ie, amikacin, kanamycin, or capreomycin]), which are more difficult to treat than drug-susceptible disease [5, 6]. The approximate number of cases of MDR-TB in the world is roughly 500 000 reported from at least 127 countries, and XDR-TB has been reported from 105 countries [4].

In the United States, 9412 cases of TB disease were reported in 2014, with a rate of 3.0 cases per 100 000 persons. Sixty-six percent of cases were in foreign-born persons; the rate of disease was 13.4 times higher in foreign-born persons than in US-born individuals (15.3 vs 1.1 per 100 000, respectively) [7]. An estimated 11 million persons are infected with Mtb [8]. Thus, although the case rate of TB in the United States has declined during the past several years, there remains a large reservoir of individuals who are infected with Mtb. Without the application of improved diagnosis and effective treatment for LTBI, new cases of TB will develop from within this group, which is therefore a major focus for the control and elimination of tuberculosis [9].

Mtb is transmitted from person to person via the airborne route [10]. Several factors determine the probability of Mtb transmission: (1) infectiousness of the source patient—a positive sputum smear for acid-fast bacilli (AFB) or a cavity on chest radiograph being strongly associated with infectiousness; (2) host susceptibility of the contact; (3) duration of exposure of the contact to the source patient; (4) the environment in which the exposure takes place (a small, poorly ventilated space providing the highest risk); and (5) infectiousness of the Mtb strain. In the United States, among contacts of patients with TB disease evaluated during a contact investigation, about 1% have TB disease themselves and 23% have a positive tuberculin skin test (TST) without evidence of tuberculosis disease and are considered to have LTBI [11]. Those who are household contacts and are
exposed to patients who are smear positive have higher rates of both infection and disease [12]. Medical procedures that generate aerosols of respiratory secretions, such as sputum induction and bronchoscopy, entail significant risk for \( Mtb \) transmission unless proper precautions are taken [13].

Initial Infection: Acquisition of Latent \( Mtb \) Infection

After inhalation, the droplet nucleus is carried down the bronchial tree and implants in a respiratory bronchiole or alveolus. Whether or not inhaled tubercle bacilli establish an infection depends on both host and microbial factors [14]. It is hypothesized that, following infection, but before the development of cellular immunity, tubercle bacilli spread via the lymphatics to the hilar lymph nodes and then through the bloodstream to more distant anatomic sites [15]. The majority of pulmonary tuberculosis infections are clinically and radiographically unapparent [16]. A positive TST or IGRA result, most commonly, is the only indication that infection with \( Mtb \) has taken place.

Those who develop a positive TST are considered to have LTBI. It is estimated that, in the absence of treatment, approximately 4%-6% of individuals who acquire LTBI will develop active TB disease during their lifetime. The greatest risk of progression is during the first 2 years following exposure [11, 17]. The ability of the host to contain the organism is reduced in young (<4 years) children and by certain diseases such as silicosis, diabetes mellitus, and diseases associated with immunosuppression (eg, human immunodeficiency virus [HIV] infection), as well as by corticosteroids and other immunosuppressive drugs such as tumor necrosis factor alpha (TNF-\( \alpha \)) inhibitors. In these circumstances, the likelihood of progression to TB disease is greater. For example, individuals who have a prior latent infection with \( Mtb \) (not treated) and then acquire HIV infection will develop TB disease at an approximate rate of 5%-10% per year (in the absence of effective HIV treatment) [18, 19].

DIAGNOSTIC TESTS FOR LTBI

The aim of testing for LTBI is to identify those who will benefit from prophylactic therapy. At present, the likelihood of completing LTBI treatment is relatively modest. In some reports, only 17%-37% of those eligible for LTBI therapy ultimately complete the treatment course, with higher rates of completion associated with shorter courses of therapy [20, 21]. Once therapy has been initiated, completion rates are more favorable [22]. It is hoped that better diagnostic tests, testing strategies, and treatment regimens will allow for resources to be focused on patients who are most deserving of evaluation and treatment of LTBI and, therefore, result in increased completion of therapy rates.

Tuberculin Skin Testing

The tuberculin skin test (TST) detects cell-mediated immunity to \( Mtb \) through a delayed-type hypersensitivity reaction using a protein precipitate of heat-inactivated tubercle bacilli (purified protein derivative [PPD]–tuberculin). The TST has been the standard method of diagnosing LTBI.

The TST is administered by the intradermal injection of 0.1 mL of PPD (5 TU) into the volar surface of the forearm (Mantoux method) to produce a transient wheal. The test is interpreted at 48–72 hours by measuring the transverse diameter of the palpable induration. TST interpretation is risk-stratified [23]. A reaction of 5 mm or greater is considered positive for close contacts of tuberculosis cases; immunosuppressed persons, in particular persons with HIV infection; individuals with clinical or radiographic evidence of current or prior TB; and persons receiving TNF blocking agents. A reaction of ≥10 mm is considered positive for other persons at increased risk of LTBI (eg, persons born in high TB incidence countries and those with at risk of occupational exposure to TB) and for persons with medical risk factors that increase the probability of progression from LTBI to TB (Figure 1). A reaction of 15 mm or greater is considered positive for all other persons. Serious adverse reactions to PPD-tuberculin are rare. However, strong reactions with vesiculation and ulceration may occur.

The sensitivity of the TST, as measured in clinically well persons with previously treated tuberculosis, is high (95%-98%). False-negative reactions occur more frequently in infants and young children, early (<6–8 weeks) after infection, in persons having recently received viral vaccination, in persons with clinical conditions associated with immunosuppression (eg, HIV infection) or overwhelming illness (including extensive or disseminated tuberculosis), after recent viral and bacterial infections, and in association with treatment with immunosuppressive drugs (eg, high-dose corticosteroids, TNF inhibitors).

Test specificity of the TST is decreased among persons with prior BCG vaccination, especially those vaccinated postinfancy and those with repeat vaccination. Similarly, persons living in areas where nontuberculous mycobacteria are common are at increased risk of having false-positive TST reactions. Repeated administration of TSTs cannot induce reactivity; however, a repeat TST can restore reactivity in persons whose TST reactivity has waned over time. Because of this “boosting phenomenon,” initial repeat testing is recommended for persons with a negative TST who are to undergo periodic TST screening and who have not been tested with tuberculin recently (eg, 1 year).” This “2-step” testing, with a repeat TST within 1–3 weeks after an initial negative TST, is intended to avoid misclassification of subsequent positive TSTs as a TST conversion, indicating recent infection, when they are actually a result of boosting.

Benefits and Limitations of the TST

The benefits of the TST include its simplicity to perform (it does not require a laboratory or equipment and can be done by a trained healthcare worker in remote locations), its low
cost, no need for phlebotomy, the observation that it reflects a polyclonal immune response, and the foundation of well-controlled studies that support the use of the TST to detect LTBI and guide the use of prophylactic therapy [24]. In addition, there are well-established definitions of TST conversion, which are particularly helpful when using the TST in the setting of serial testing.

Limitations include the need for trained personnel to both administer the intradermal injection and interpret the test, inter- and intrareader variability in interpretation, the need for a return visit to have the test read, false-positive results due to the cross-reactivity of the antigens within the PPD to both BCG and nontuberculous mycobacteria, false-negative results due to infections and other factors, rare adverse effects, and complicated interpretation due to boosting, conversions, and reversions [24].

Interferon-Gamma Release Assays

Overview of IGRAs

Until recently, the TST has been the only method to test for latent infection with Mtb. Ideally, an improved diagnostic test would specifically identify those with Mtb infection and would delineate those at risk for disease progression. In this regard, the TST has well-known strengths and limitations [23, 25, 26]. The IGRAs are newer tests to diagnose infection with Mtb. IGRAs are in vitro, T cell–based assays that measure interferon gamma (IFN-γ) release by sensitized T cells in response to highly specific Mtb antigens.

Immune Basis of IGRAs

Like the TST, the IGRA is a reflection of the cellular immune response. The discovery of antigens that have elicited robust immune responses and are relatively specific for infection with Mtb has enabled the development of IGRA assays, which are more specific for Mtb infection than the TST [27], particularly in the setting of BCG vaccination. Of particular interest has been the RD-1 gene segment, a 9.5-kb DNA segment absent from all strains of M. bovis but present in wild-type M. bovis and Mtb [28]. This region, containing 11 open reading frames, is responsible for the transcription and translation of a variety of antigenic proteins, including early secretory antigen (ESAT-6) [29–33] and culture filtrate protein (CFP-10) [34–37]. Both antigens are absent from all attenuated strains of M. bovis (BCG strains) and most nontuberculous mycobacteria with the important exceptions of Mycobacterium kansasii, Mycobacterium szulgai, Mycobacterium marinum [32, 38], and Mycobacterium leprae [39, 40].

IGRA assays are primarily a reflection of a CD4+ T-cell immune response to these antigens. Immunologic memory is characterized by the clonal expansion of antigen-specific T cells following exposure to an antigen. Effector memory T cells are defined by their capacity to respond rapidly to subsequent antigenic exposure. This response is characterized by the release of cytokines, as well as further expansion of these cells. Responses measured in current short-term IGRA assays reflect the presence of these cells. Although it has been postulated measurement of these short-term effectors might reflect recent infection and/or ongoing bacterial replication, current evidence does not support this hypothesis [41–43].

Commercially Available IGRAs

Currently, there are 2 commercially available IGRA platforms that measure interferon-γ release in response to Mtb-specific antigens: the QuantiFERON TB Gold in Tube (QFT-GIT; Cellestis Limited, Carnegie, Victoria, Australia) and T-SPOT.TB test (T-SPOT, manufactured by Oxford Immunotec Ltd, Abingdon, United Kingdom). The QFT-GIT measures IFN-γ plasma concentration using an enzyme-linked immunosorbent assay (ELISA), while the T-SPOT assay enumerates T cells releasing IFN-γ using an enzyme-linked immunospot (ELISPOT) assay.

QuantiFERON Assays

The QFT-GIT method has been approved by the US Food and Drug Administration (FDA) and has replaced the QuantiFERON-TB Gold (QFT-G) test. Whole blood (minimum 3 mL) is drawn directly into heparinized tubes coated with lyophilized antigen and agitated. In this case, peptides from ESAT-6, CFP-10, and TB7.7 are found within the same tube. Two additional tubes are drawn as controls (mitogen control and nil control). The mitogen control (phytohemagglutinin [PHA]) stimulates T-cell proliferation and ensures that viable cells are present. After incubation for 16–24 hours at 37°C, plasma is collected from each tube and the concentration of IFN-γ is determined for each by ELISA. The in-tube methodology requires no additional sample handling. Perhaps because of the nearly immediate exposure of T cells to antigen, as well as the addition of the TB7.7 peptide, the QFT-GIT may be more sensitive than the QFT-G test. Studies reporting the sensitivity and specificity of the QFT-GIT test are provided in Supplementary Tables 1 and 2, respectively. The next generation of QFT (QFTPlus) has been introduced in Europe and is pending approval in the United States. QFTPlus contains a tube of short peptides derived from CFP-10, which are designed to elicit an enhanced CD8 T-cell response. There is no TB7.7 peptide. No published information is available to evaluate the performance of this test.

The QFT-GIT assay is considered positive if the difference in the IFN-γ concentration in response to the Mtb antigens and the IFN-γ response to the nil control is ≥0.35 IU. In addition, to control for high background in the nil control, the IFN-γ response to antigen must be 25% greater than the IFN-γ concentration in the NIL control. An indeterminate response is defined as either a lack of response in the PHA control well (IFN-γ concentration ≤0.5 IU) or a nil control that has a very high background (IFN-γ concentration >8 IU).
**T-SPOT.TB Assays**

The T-SPOT.TB assay is currently available in Europe, Canada, and has been approved for use in the United States with revised criteria for test interpretation. For the T-SPOT.TB assay, blood (minimum 2 mL) is drawn into either a heparin or CPT Ficoll tube, and must be processed within 8 hours. More recently, this time has been extended to 32 hours if the "T-cell Xtend" additive is used and the blood kept between 10°C and 25°C. Peripheral blood mononuclear cells (PBMCs) are separated using density gradient centrifugation, enumerated, and then added to micropellet wells at 2.5 × 10^5 viable PBMCs per well that have been coated with monoclonal antibodies to IFN-γ (ELISPOT assay). Peptides derived from ESAT-6 and CFP-10 antigens are then added and the plate is developed following overnight (16–20 hours) incubation at 37°C. Cells are then washed away and "captured" IFN-γ is then detected via a sandwich capture technique by conjugation with secondary antibodies hence revealing a “spot.” These spots are then enumerated as “footprints” [44] of effector T cells [44, 45].

For the T-SPOT.TB assay, a positive response is based on spot-forming units (SFU). Outside of the United States, if the negative control well contains ≤5 SFU and there are >6 SFU above the media nil control in either of the antigen wells, then this is considered positive. If the negative control well has ≥6 SFU, then the antigen wells must be at least 2 times the negative control well for a response to be considered positive. An invalid response is defined as high background in the negative control well (≥10 SFU) or if the positive control well is not responsive to mitogen (PHA, <20 SFU). The FDA has published revised criteria for T-SPOT.TB interpretation in the United States, in which a test is considered negative if there are ≤4 spots. Eight spots or greater is considered positive. Five, 6, and 7 spots are considered “borderline” and would be interpreted in conjunction with the subject's pretest probability of infection with Mtb. Studies reporting the sensitivity and specificity of the T-SPOT test are provided in Supplementary Tables 3 and 4, respectively.

**Indeterminate/Invalid IGRA Responses**

Unlike the TST, in which the results are interpreted categorically based on the size of the reaction [46], the IGRA results currently have a trichotomous outcome yielding a positive, negative, or indeterminate result (T-SPOT may also yield a borderline result as described above). As described above, an indeterminate/invalid IGRA can result from either a high background (nil) response or from a poor response to positive control mitogen. Indeterminate IGRA results are associated with immunosuppression [47–49], although they may occur in healthy individuals (studies reporting the test characteristics of IGRA in individuals with immunosuppression are provided in Supplementary Tables 5 and 6). With regard to those with a poor response to the positive control mitogen, there are at least 2 possibilities. First, the test may not have been correctly performed. For example, errors in specimen collection, long delays in specimen processing, incubator malfunction, or technical errors might result in a poor mitogen response. Here, it is reasonable to simply repeat the assay. Second, a persistently diminished response to mitogen may be a reflection of anergy. Thus, the reproducibility and details regarding the reason for an indeterminate result may provide clinically useful information.

**Reproducibility of IGRA**

Because IGRA is predicated on in vitro release of cytokines from stimulated cells, there is likely to be more variability in these tests than those based on the measurement of a circulating substance such as sodium. There are at least 4 sources of variability which are inherent in the IGRA: (1) the type of measurement itself (ie, ELISA or ELISPOT), (2) reproducibility of a complex biological reaction, (3) the natural variability of immune responses, and (4) variability introduced during the course of test performance or manufacturing variances.

Reproducibility has been evaluated for both the QFT and T-SPOT assays. Although published information regarding currently available tests is limited [50, 51] the QFT-IT result was reported to have an 11% variance (http://www.accessdata.fda.gov/cdrh_docs/pdf/P010033). Studies on within subject variability of the QFT-IT are limited and most were performed in areas of the world where Mtb is endemic and variability over time due to re-infection would be expected [50, 51]. Recently, intrasubject variability of QFT-IT was assessed using available plasma, and a discordance rate of 8% between the first and second tests was observed. While the variations were quantitatively modest, results at or near the cutoff resulted in differing test results [52]. This variability might spuriously change the test result (positive to negative or negative to positive). Consequently, values at or near the test cutoff should be interpreted with caution. Variability of the T-SPOT was dependent on the strength of the response, and varied from 4% in those with robust responses, to 22% in those whose responses were close to the cutoff (http://www.accessdata.fda.gov/scripts/cdrh/cfdocs/cftopic/pma/cfm?num=p070006).

**Boosting of IGRA**

Initial studies found that repeat TST testing did not alter the IGRA response [53, 54]. However, more recent evidence [50, 51] suggests that the prior placement of a TST can boost an IGRA, particularly in those individuals who were already IGRA positive to begin with (ie, previously sensitized to Mtb or possibly other mycobacteria). Additionally, it was found that this could be observed in as little as 3 days post-TST administration, and that the boosting effect may wane after several months [50, 51]. While these data do not detract from the excellent overall agreement that has been reported, they suggest when dual testing is to be considered that the IGRA be collected either concurrently or prior to TST placement.

Diagnosis of TB in Adults and Children • CID 2017:64 (15 January) • e9
Special Considerations
Because IGRA
ded to be or
with TST
and "non-
ry prevent
"Fig-1-3
0.5-
-2

Benefits and Limitations of IGRA
drug abuse); while the relative
risk factors) and intermediate risk for progression (ie, diabetes,
ren
ease is more likely to be disseminated and elude discovery, and
firmities [62] and has public health benefits, as each new
ase is likely to infect others. The consequences of failing to
 progression to active TB disease may be especially severe in
the young or immunocompromised host, in whom the dis-
ease is more likely to be disseminated and elude discovery, and
has a higher mortality rate. Failure to rapidly diagnose TB dis-
 ease also poses a risk of widespread transmission in hospitals,
homeless shelters, and prisons.

Patients with LTBI have a 4%-6% lifetime risk of developing TB
disease, with approximately half of these cases occurring follow-
ing recent exposure [11, 17]. Multiple placebo-controlled trials in
adults and children with LTBI have shown that isoniazid reduces
the subsequent development of TB disease in patients at high risk
of progression. As an example, in a trial of 28,000 individuals with
LTBI and radiographic evidence of healed tuberculosis, isonia-
ized taken for 52 weeks reduced the subsequent development of
TB disease from 14.3% to 3.6% [62]. Other groups in which the
treatment of LTBI has been demonstrated to reduce the incidence
of TB disease include household contacts of active TB patients
[55, 63], native Alaskan communities [64], residents of mental
health facilities [65], persons with HIV infection [66-69], and
individuals treated with TNF inhibitors [70, 71]. These data can
be extrapolated to populations at low risk for progression (ie, no
risk factors) and intermediate risk for progression (ie, diabetes,
chronic renal disease, intravenous drug abuse); while the relative
benefit of treatment is probably similar in these lower risk popu-
lations, the absolute benefit is almost certainly smaller due to the
lower baseline risk of progression to TB disease.

These studies provide high-quality evidence that treatment
of LTBI reduces the incidence of TB disease in populations at
high risk for progression. However, they provide only moder-
ate-quality evidence that treatment of LTBI reduces the inci-
dence of TB disease in populations at low or intermediate risk for
progression because the data are from high-risk populations.

Our recommendations for the diagnosis of LTBI reflect both
the likelihood of infection (either likely or unlikely, based upon
studies that used the TST to detect LTBI) and the risk of pro-
gression if infected (low; intermediate, RR 1.3-3; and high, RR
3-10). This paradigm is summarized in Figure 1.

Question 1: Should an IGRA or a TST be performed in indi-
viduals 5 years or older who are likely to be infected with
Mtb, who have a low or intermediate risk of disease progres-
sion, and in whom it has been decided that testing for LTBI
is warranted?

Evidence
In individuals who are likely to be infected with Mtb but at low
or intermediate risk of disease progression, the sensitivity of
IGRAs in the detection of Mtb infection has been consistently
reported as either equal (QFT; 81%-86%) or superior (T-SPOT;
90%-95%) to the sensitivity of the TST (71%-82%) [47, 72-94]
when a final diagnosis of either microbiologically confirmed or clinical TB is used as the reference standard. Individuals who are likely to be infected with Mtb include household contacts (studies reporting the test characteristics of IGRA in contacts are provided in Supplementary Table 7), recent exposures of an active case, mycobacteriology laboratory personnel, immigrants from high-burden countries, and residents or employees of high-risk congregate settings. Individuals at low risk of progression to TB include those with no risk factors, while those at intermediate risk of progression to TB include those with diabetes, chronic renal failure, or intravenous drug abuse.

In patients who are known to have received vaccination with BCG, the specificity of IGRA has also been consistently superior to TST testing, presumably because IGRA rely on responses to antigens absent in BCG and many nontuberculous mycobacteria. In contrast, among patients who have not received vaccination with BCG, the specificity of IGRA and TST appears similar. Meta-analyses estimate that the specificity of QFT–IT to be >95%, whereas the specificity for TST is roughly 75% in those with no prior exposure to BCG. The specificity is reduced to roughly 60% in those with a history of BCG vaccination [24]. Data for the commercially available T-SPOT are more limited. In German healthcare workers, specificity (using a cutoff of 6 spots) was reported at 97%, whereas in Korean adolescents the specificity was 85% [49]. In Navy recruits, specificity was 99% using the 8-spot cutoff [95].

Our confidence in the estimated test characteristics was moderate because many of the studies did not report whether the subjects were consecutively enrolled.

Recommendation 1a: We recommend performing an IGRA rather than a TST in individuals 5 years or older who meet the following criteria: (1) are likely to be infected with Mtb, (2) have a low or intermediate risk of disease progression, (3) it has been decided that testing for LTBI is warranted, and (4) either have a history of BCG vaccination or are unlikely to return to have their TST read (strong recommendation, moderate-quality evidence). Remarks: A TST is an acceptable alternative, especially in situations where an IGRA is not available, too costly, or too burdensome.

Recommendation 1b: We suggest performing an IGRA rather than a TST in all other individuals 5 years or older who are likely to be infected with Mtb, who have a low or intermediate risk of disease progression, and in whom it has been decided that testing for LTBI is warranted (conditional recommendation, moderate-quality evidence). Remarks: A TST is an acceptable alternative, especially in situations where an IGRA is not available, too costly, or too burdensome.

Rationale

Accuracy studies indicate that IGRA are more specific and equally or more sensitive than TST in individuals who have received the BCG vaccination; therefore, false-positive results are less likely with IGRA than TST. This is important because false-positive results may lead to unnecessary treatment and its accompanying risks (ie, hepatotoxicity) [96–98]. To minimize these risks, the guideline development panel chose to recommend IGRA testing for individuals who received the BCG vaccination.

In contrast, the accuracy of TST and IGRA appears similar in those without a history of BCG vaccination. Despite the similar test characteristics, the guideline development committee chose to suggest IGRA testing over TST testing in such patients because it was concerned about the reliability of a history of having received or not received the BCG vaccination. Because many of the individuals who fall into the likely to be infected with Mtb category are from regions of the world in which the BCG vaccination is routinely administered, the committee concluded that individuals who are likely to be infected with Mtb and provide history of not having received the BCG vaccination should be treated the same as those who provide a history of having received the BCG vaccination, unless there is a reason to choose an alternate approach such as IGRA testing not being available, being too costly, or being too burdensome.

The recommendation to perform IGRA testing rather than TST testing is strong for those who have received the BCG vaccination or who are not likely to return for TST read, reflecting the guideline development committee's certainty that avoiding the serious consequences of false-positive results and obtaining a result to guide therapy outweigh the additional cost and need to perform phlebotomy for IGRA testing. In contrast, the suggestion to perform IGRA testing rather than TST testing on all other patients who are likely to be infected with Mtb and have a low or moderate risk of progressing to TB disease is conditional, reflecting the committee's recognition that the choice should depend upon the clinical context as the test characteristics are similar. While the committee concluded that IGRA testing is preferable in most patients, it recognized that TST testing may be more appropriate in a sizeable minority due to availability, feasibility, cost, or burden.

Justification for Extending the Recommendation Down to 5 Years of Age

Young children are at increased risk of developing TB following infection and more likely to develop severe disease than older children and adults [99, 100]. This risk is highest in the youngest infants, diminishes with increasing age, and becomes equivalent with older children and adults at approximately 5 years of age. Thus, children ≥5 years old have a similar risk of TB as adults and display a similar disease spectrum. With respect to Mtb infection, children aged ≥5 years possess a functional immune response equivalent to that of adults. In addition, the results of existing studies of IGRA performance in children ≥5 years of age, albeit limited, are consistent with results of studies of IGRA performance in adults. The sensitivity of IGRA in
children with TB [56, 59, 101–105] and in older children who are household [106, 107] or school [108] contacts and the specificity of IGRA in children [102, 108] are comparable to those of adults. For these reasons, it seems reasonable to extrapolate the results of studies of IGRA performance in healthy adults to children aged ≥5 years.

Cautions and Limitations

While both IGRA and TST testing provide evidence for infection with Mtb, they cannot distinguish active from latent tuberculosis. Therefore, the diagnosis of active TB must be excluded prior to embarking on treatment for LTBI. This is typically done by determining whether or not symptoms suggestive of TB disease are present, performing a chest radiograph and, if radiographic signs of active tuberculosis (eg, airspace opacities, pleural effusions, cavities, or changes on serial radiographs) are seen, then sampling is performed and the patient managed accordingly.

Quantitative aspects of the tests are poorly understood. With respect to the TST, the result is categorized as positive or negative and quantitative data are of limited utility, with the exception of recognition that a large (>15 mm) skin test reaction is more likely to reflect infection with Mtb [109, 110]. The dichotomous characterization of the result, coupled with the fact that repeat testing is not recommended in the setting of a prior positive test result, has resulted in a paucity of information about the variability of the TST result over time. With respect to IGRA, measurement of IFN-γ over time may reflect inherent variability in the test result (the FDA accepts a variance of 11%) or true immunological variation due to alterations in the abundance of Mtb antigens, exposure to other antigens, and/or the health and nutritional status of the host. As an example, it is possible that a rise in IFN-γ might reflect ongoing exposure and/or growth of the bacteria. Alternatively, a rise in IFN-γ may reflect variability of the test. At present, there are insufficient data upon which to base any recommendations for quantitative interpretation of IGRA beyond those cut-points recommended by the FDA. However, it is important to recognize that the optimal cut-points are controversial and results near the cut-point are less reliable than results far above or below the cut-point. The results of IFN-γ testing should be reported quantitatively such that these immune correlates of the natural history of TB can be prospectively discerned and ultimately applied to clinical practice.

Discordance between TST and IGRA testing is common. Not surprisingly, TST-positive/IGRA-negative discordance is often seen in persons with prior exposure to BCG. However, TST-positive/IGRA-negative discordant results where the TST is well over 15 mm have also been reported. The reasons for this delayed type hypersensitivity are not understood. It could relate to the possibility that discordance may reflect immune responses that have occurred in the remote past (and where the antigen is currently not available to drive an ongoing response that can be measured by IGRA), may reflect immune differences inherent in a delayed-type hypersensitivity versus blood assay, or may reflect exposures to nontuberculous mycobacteria. In low-risk populations, discordant tests are likely to be false positives [61, 111]. Clearly, more information is desirable regarding which test best reflects productive infection and, therefore, best reflects the likelihood of disease progression.

The benefit of targeted testing for LTBI resides not in the test employed, but in its programmatic use. We acknowledge that programmatic considerations such as cost, test availability, prevalence of BCG exposure in the target population, ability to reevaluate the patient 2–3 days after testing, and the training and expertise of program staff might all affect the decision to use IGRA- or TST-based evaluations.

Question 2: Should an IGRA or a TST be performed in individuals 5 years or older who are likely to be infected with Mtb, who have a high risk of progression to disease, and in whom it has been decided that testing for LTBI is warranted?

Evidence

Individuals at high risk of progression to TB include those with HIV infection, an abnormal chest radiograph consistent with prior TB, or silicosis. It also includes those who are receiving immunosuppressive therapy. Most data about the accuracy of the TST and IGRA are from patients who are immunocompromised.

Studies have compared TST and IGRA in the setting of immunocompromise. Both diagnostic tests have diminished sensitivity in this setting. The sensitivity of IGRA (QFT-IT and T-SPOT) for detecting LTBI in individuals with HIV infection has been estimated to be from 65% to 100% [112–114], while the sensitivity of TST is only 43% (25,85) when a final diagnosis of either microbiologically confirmed or clinical TB is used as the reference standard. These limited data suggest that IGRA are at least as sensitive as TST in the setting of HIV infection. Studies have also compared IGRA with TST in populations that were heterogeneous with respect to both the type of underlying immunocompromise and the reasons for testing. These studies demonstrated significant discord between TST and IGRA results, but the source of the discordance has not been elucidated [61]. The panel's confidence in the estimated test characteristics of IGRA and TST testing was moderate because it was not reported whether patients were consecutively enrolled or whether there was true diagnostic uncertainty.

Recommendation 2: There are insufficient data to recommend a preference for either a TST or an IGRA as the first-line diagnostic test in individuals 5 years or older who are likely to be infected with Mtb, who have a high risk of progression to disease, and in whom it has been determined that diagnostic testing for LTBI is warranted.
Rationale

The committee judged the body of evidence insufficient to render a recommendation for either IGRA or TST testing in patients likely to be infected with Mtb who are at high risk for progression to disease because the estimated test characteristics were widely variable and derived from only a small subgroup of such patients (ie, immunocompromised patients).

As part of the discussion about which diagnostic test to perform in patients likely to be infected who are at high risk for progression to disease, many committee members acknowledged that they perform a second test in their clinical practices when such patients test negative; specifically, they perform a TST if an initial IGRA is negative or an IGRA if an initial TST is negative. If the second test is positive, they consider this evidence for infection with Mtb. Their practice is not based upon empirical evidence, but rather, the following clinical rationale. A sensitive diagnostic test is important for individuals who are likely to be infected with Mtb and at high risk of progression, so that such individuals are less likely to receive false-negative results that will result in delayed diagnosis and treatment. Performing a second diagnostic test when the initial test is negative is one strategy to increase sensitivity. While this strategy to increase sensitivity may reduce the specificity of diagnostic testing, this may be an acceptable tradeoff in situations in which it is determined that the consequences of missing LTBI (ie, not treating individuals who may benefit from therapy) exceed the consequences of inappropriate therapy (ie, hepatotoxicity).

Caution and Limitations

While both QFT and T-SPOT rely on the release of IFN-γ in response to RD-1 antigens, limited data have suggested that indeterminate results are more common for QFT-IT when the CD4 count is <200 cells/µL than T-SPOT [115]. This may be the result of T-SPOT using a defined number of PBMCs, which may better, but not completely, normalize for the lack of CD4+ T cells. By incorporating a measure of anergy into the test (Mitogen control), IGRA may more accurately allow the clinician to discriminate a test that is negative from one that is indeterminate (anergic by virtue of inadequate responses to mitogen).

Question 3: Should an IGRA or a TST be performed in individuals 5 years or older who are unlikely to be infected with Mtb, but in whom it has been decided that testing for LTBI is warranted?

Evidence

There is a lack of direct evidence regarding the relative test characteristics of IGRA and TST testing in individuals who are unlikely to be infected with Mtb. Indirect evidence from individuals likely to be infected with Mtb indicates that IGRA testing is more specific than TST testing and equally or more sensitive than TST testing. We have no reason to suspect that these relative test characteristics will be different among individuals who are unlikely to be infected with Mtb. However, it is likely that false-positive results are more common for both IGRA and TST in populations with a lower prevalence of LTBI. This is supported by a study of longitudinal testing of healthcare workers residing in areas of low TB prevalence, which found that most conversions were false-positive results as evidenced by a negative result on repeat testing [116].

The evidence provides low confidence in the estimated test characteristics in our population of interest because many of the estimates are based upon evidence from patients who are likely to be infected with a high risk for progression rather than patients who are unlikely to be infected, and many of the studies did not report whether subjects were consecutively enrolled.

Guidelines recommend that persons at low risk for Mtb infection and disease progression NOT be tested for Mtb infection. We concur with this recommendation. However, we also recognize that such testing may be obliged by law or credentialing bodies. If diagnostic testing for LTBI is performed in individuals who are unlikely to be infected with Mtb despite guidelines to the contrary.

Recommendation 3a: We suggest performing an IGRA instead of a TST in individuals 5 years or older (conditional recommendation, low-quality evidence). Remarks: A TST is an acceptable alternative in settings where an IGRA is unavailable, too costly, or too burdensome.

Recommendation 3b: We suggest a second diagnostic test if the initial test is positive in individuals 5 years or older (conditional recommendation, very low-quality evidence). Remarks: The confirmatory test may be either an IGRA or a TST. When such testing is performed, the person is considered infected only if both tests are positive.

Rationale

Current ATS/CDC and American Academy of Pediatrics guidelines recommend that testing for LTBI not be performed in individuals at low risk for infection with Mtb because the risk of isoniazid chemoprophylaxis may outweigh the potential benefit [117]. Despite this, testing is often performed in conjunction with school enrollment, employee health testing, and other institutional settings. In such patients, many conversions are false results, which may lead to unnecessary therapy and, therefore, unnecessary and age-related risk of hepatotoxicity.

The evidence indicates that false-positive results are frequent (ie, more common than true-positive results) among individuals who are unlikely to be infected with Mtb. Use of a more specific test may result in fewer false-positive results and, therefore, fewer persons receiving unnecessary LTBI treatment and being placed at risk for adverse outcomes. In addition to the risk
associated with isoniazid chemoprophylaxis, those with a positive test for LTBI often undergo additional screening, including a chest radiograph. Avoiding such unnecessary screening has both cost and health benefits. The desire for a more specific test favors IGRA testing over TST, according to evidence described above from patients who are likely to be infected and who have a low or intermediate risk for progression. The notion of performing a second, confirmatory test following an initial positive result is based upon the evidence that false-positive results are common among individuals who are unlikely to be infected with Mtb and the committee's presumption that performing a second test on those whose initial test was positive will improve specificity.

The recommendations are both conditional because the quality of evidence provided the committee with limited confidence in the estimated test characteristics of IGRAs and TST in individuals who are unlikely to be infected; therefore, the committee could not be certain that the desirable consequences of performing IGRAs instead of TST, or of performing a second test following a positive result, outweigh the undesirable consequences in the vast majority of patients.

Cautions and Limitations (Testing for TB in Healthcare Workers)
Traditionally, once an individual has had a positive TST, future use of the TST for screening is not recommended due to the belief that the skin test will remain positive for life. In those who are TST negative, serial testing can be complicated by random variability, boosting (ie, increased reactions upon retesting due to immunological memory), conversions (ie, new reactions due to new infection), and reversions (ie, decreased reactions). Criteria for the placement and reading of the TST, as well as the effect of boosting with PPD, criteria for TST conversion have been established. Receiver operating characteristic (ROC) analysis has been used to establish criteria for positive and negative IGRA results in those thought to be unlikely to be infected or those with TB disease. However, IGRAs have not proven to be the solution to the problem of false-positive results associated with serial testing in low-risk individuals. At present, there is insufficient information available to guide the establishment of definitive criteria for the conversion and possible reversion of IGRAs. The issue of interpreting IGRA conversions and reversions in the context of serial testing has proven especially problematic. For example, in a study of 216 Indian healthcare workers, a QFT conversion rate of 12% and a reversion rate of 24% were observed, with many of these apparent changes occurring near the cutoff values [118]. A longitudinal study involving 2563 in healthcare workers demonstrated an IGRA conversion (6%–8%) in those undergoing serial testing [116]. These rates were 6–9 times higher than that seen for the TST and were thought to have represent false conversions. Such studies have not yielded useful criteria that can be used to distinguish Mtb infection from a false-positive result [119]. As discussed above, there are a number of sources of variability in the IGRA assay related to laboratory technique such as sample agitation, time elapsed prior to incubation, duration of incubation, agitation technique, and blood volume that could result in variability around the cutoff value. In this instance, this variability may reflect the inherent variability of a biologic measurement, and is the rationale behind the committee's recommendation that quantitative values be reported. The optimal cut-points for IGRA testing are controversial. While results close to the cut-point tend to be less reliable than results substantially above or below the cut-point, this is not absolute; in many instances, positive values well above the threshold were not reproduced in subsequent testing [116]. It is for this reason that the committee felt that quantitative guidance regarding the interpretations of conversions and reversions in the context of healthcare worker screening could not be provided. Given the varied sources of IGRA variability [24], the committee thought that a positive test in a low-risk individual was likely to be a false-positive result, and recommended repeat testing.

Question 4: Should an IGRA or a TST be performed in healthy children <5 years of age in whom it has been decided that testing for LTBI is warranted?

Evidence
The body of evidence regarding IGRA performance in young children is limited. Compared with adults, a limited number of children have been enrolled in IGRA studies. Even fewer children from nonendemic countries have been studied, and many reports do not include a separate analysis of young children.

The sensitivity of IGRAs in young children with TB ranges from 52% to 100% when a final diagnosis of either microbiologically confirmed or clinical TB disease is used as the reference standard, which is comparable to adults [56, 59, 76, 102–105, 120]. The sensitivity of the TST has been reported as equivalent or increased compared with IGRAs in children [56, 59, 76, 101–105, 120], with young age associated with decreased IGRA positivity [107]. Important caveats to this comparison, however, are that some studies used earlier, less sensitive versions of the IGRA and results have been inconsistent. As examples of the inconsistencies, a study using an IGRA similar to a currently available IGRA test demonstrated increased sensitivity of the IGRA compared with TST in children aged <3 years, especially among those coinfected with HIV and/or malnourished [121]. None of these studies were performed in nonendemic countries.

The specificity of IGRAs appears to be excellent in children in the range of 90%–100% [122] according to a study conducted in children who had nontuberculous mycobacteria. The study found that IGRAs were more specific than TST in children with nontuberculous mycobacterial disease [102].

Our confidence in the estimated test characteristics of IGRAs and TST in children is very low because most of the studies did
not report whether or not they enrolled consecutive patients, were not performed in nonendemic countries, and have provided inconsistent results.

Recommendation: We suggest performing a TST rather than an IGRA in healthy children <5 years of age for whom it has been decided that diagnostic testing for LTBI is warranted (conditional recommendation, very low-quality evidence). Remarks: In situations in which an IGRA is deemed the preferred diagnostic test, some experts are willing to use IGRA in children over 3 years of age.

Rationale
The limited direct evidence described above suggests that the TST might be more sensitive than IGRA in young children, and IGRA may be more specific than the TST, particularly in those given BCG. Because young children have a high risk for progression to active TB disease, the committee believed that the sensitivity of the diagnostic test (ie, avoiding false-negative results, missed opportunities to treat) is more important than the specificity of the test (ie, avoiding false-positive results, unnecessary therapy). This is supported by the observations that the potential consequences of delayed treatment are high, while the risk of hepatotoxicity is greatly reduced in young children. An additional reason to favor TST testing over IGRA testing in young children is that the management of the most at-risk young children (ie, young household contacts) depends upon the results of serial testing for infection, for which there are no data for IGRA in young children.

While there are theoretical benefits from IGRA testing (eg, improved acceptance of LTBI therapy), these benefits have not been proven. Therefore, there is insufficient evidence that the benefits of IGRA testing exceed the well-known limitations of the TST. For these reasons, it is too early to recommend replacing the TST with IGRA testing. The recommendation is conditional because the quality of evidence provided the committee with limited confidence in the estimated test characteristics of IGRA and TSTs in children; therefore, the committee could not be certain that the desirable consequences of performing IGRA instead of TSTs outweigh the undesirable consequences in the vast majority of patients.

Cautions and Limitations
In studies of young children that report rates of indeterminate IGRA results, the frequency ranges from 0 to 35%, which is generally higher than in studies that reported in adults. Several studies have reported an increased rate of indeterminate IGRA results in children <5 years of age [47, 56–60]. As phlebotomy is more difficult in young children, inability to perform the IGRA due to insufficient blood volumes represents an additional practical limitation to IGRA testing in young children. A relatively high incidence of failed phlebotomy has been documented in some studies [106, 123].

DIAGNOSTIC TESTS FOR TB
The diagnosis and management of TB disease rely on accurate laboratory tests, both for the benefit of individual patients and the control of TB in the community through public health services. Therefore, laboratory services are an essential component of effective TB control at the local, state, national, and global levels.

In the United States, up to 80% of all initial TB-related labora-
tory work (eg, AFB smear and culture inoculation) is performed in hospitals, clinics, and independent laboratories outside the public health system, whereas >50% of species identification and drug susceptibility testing (DST) is performed in public health laboratories [124]. Thus, effective TB control requires a network of public and private laboratories to optimize laboratory testing and the flow of information. Public health laboratory workers, as a component of the public health sector with a mandate for TB control, should take a leadership role in developing laboratory networks and in facilitating communication among laboratory workers, clinicians, and TB controllers.

Seven types of tests for the diagnosis of TB disease and detection of drug resistance are performed within the tuberculosis laboratory system and recommended for optimal TB control services (Table 2). These laboratory tests should be available to every clinician involved in TB diagnosis and management, and to jurisdictional public health agencies charged with TB control.

For suspected cases of pulmonary TB, sputum smears for AFB are correlated with the likelihood of transmission and then, for AFB smear–positive pulmonary cases, a nucleic acid amplification assay provides rapid confirmation that the infecting mycobacteria are from the Mtb complex. Both sputum smears for AFB and nucleic acid amplification tests (NAATs) should be available with rapid turnaround times from specimen collection.

Table 2. Essential Laboratory Tests for the Detection of Mycobacterium tuberculosis

<table>
<thead>
<tr>
<th>Test</th>
<th>Time Required</th>
</tr>
</thead>
<tbody>
<tr>
<td>I. Nucleic acid amplification test, detection (NAAT-TB)</td>
<td>1 d</td>
</tr>
<tr>
<td>II. Nucleic acid amplification test, resistance markers (NAAT-R)</td>
<td>1–2 d</td>
</tr>
<tr>
<td>III. Acid-fast bacilli microscopy</td>
<td>1 d</td>
</tr>
<tr>
<td>IV. Growth detection</td>
<td>Up to 6–8 wk</td>
</tr>
<tr>
<td>Liquid</td>
<td>(average 10–14 d)</td>
</tr>
<tr>
<td>Solid</td>
<td>(average 3–4 wk)</td>
</tr>
<tr>
<td>V. Identification of Mycobacterium tuberculosis complex by DNA probe</td>
<td>1 d*</td>
</tr>
<tr>
<td>or HPLC</td>
<td></td>
</tr>
<tr>
<td>VI. First-line drug susceptibility testing (liquid medium)</td>
<td>1 to 2 wk*</td>
</tr>
<tr>
<td>VII. Second-line and novel compound drug susceptibility testing</td>
<td></td>
</tr>
<tr>
<td>i. Liquid (broth-based) medium</td>
<td>1 to 2 wk*</td>
</tr>
<tr>
<td>ii. Solid (agar- or egg-based) medium</td>
<td>3 to 4 wk*</td>
</tr>
</tbody>
</table>

These tests facilitate decisions about initiating treatment for TB or a non-TB pulmonary infection, infection control measures (eg, patient isolation), and, if TB is diagnosed, for reporting the case and establishing priority for the contact investigation.

**DIAGNOSTIC APPROACH: TESTING FOR SUSPECTED PULMONARY TB**

Pulmonary TB is often first suspected on the basis of chest computed tomographic findings (Supplementary Table 8). Randomized trials and controlled observational studies that directly compared diagnostic tests for pulmonary tuberculosis and measured patient-important outcomes have not been performed. Therefore, the recommendations in this section are based upon data that describe how accurate a diagnostic test is at confirming or excluding pulmonary TB, coupled with the widely accepted knowledge that diagnosing pulmonary TB leads to therapy that dramatically improves patient-important outcomes and reduces disease transmission [125, 126]. Finally, it was the consensus of the committee that testing for LTBI (TST or IGRA) cannot be used to exclude a diagnosis of TB and, hence, should not be used in the evaluation of those with suspected TB.

**Question 5: Should AFB smear microscopy be performed in persons suspected of having pulmonary TB?**

**Evidence**

Performing 3 AFB smears confirms pulmonary TB with a sensitivity of approximately 70% when culture-confirmed TB disease is the reference standard. The reason for performing 3 AFB smears is that each specimen increases sensitivity. The sensitivity of the first specimen is 53.8%, which increases by a mean of 11.1% by obtaining a second specimen. Obtaining a third specimen increases the sensitivity by a mean of only 2%–5% (ie, false-positive results could exceed the additional true-positive results obtained from a third specimen).

The sensitivity of a first morning specimen is 12% greater than a single spot specimen [127]. Concentrated specimens have a mean increase in sensitivity of 18% compared with non-concentrated specimens (using culture as the standard) and fluorescence microscopy is on average 10% more sensitive than conventional microscopy [128, 129]. The specificity of microscopy is relatively high (≥90%), but the positive predictive value (PPV) varies (70%–90%) depending upon the prevalence of tuberculosis versus nontuberculous mycobacterial disease [130, 131]. These accuracy studies provide moderate confidence in the estimated test characteristics because many did not report having enrolled consecutive patients.

Recommendation 5: We recommend that AFB smear microscopy be performed, rather than no AFB smear microscopy, in all patients suspected of having pulmonary TB (strong recommendation, moderate-quality evidence). Remarks: False-negative results are sufficiently common that a negative AFB smear result does not exclude pulmonary TB. Similarly, false-positive results are sufficiently common that a positive AFB smear result does not confirm pulmonary TB. Testing of 3 specimens is considered the normative practice in the United States and is strongly recommended by the CDC and the National Tuberculosis Controllers Association to improve sensitivity given the pervasive issue of poor sample quality. Providers should request a sputum volume of at least 3 mL, but the optimal volume is 5–10 mL. Concentrated respiratory specimens and fluorescence microscopy are preferred.

**Rationale**

AFB smear microscopy can be performed in hours, is inexpensive, and is technically simple. Our committee predetermined that AFB smear microscopy would be recommended if false-negative results occur <30% of the time (ie, sensitivity is ≥70%) and false-positive results occur <10% of the time (ie, specificity is ≥90%). The likely outcome of a false-negative result is additional diagnostic testing and/or delayed diagnosis due to the duration required for cultures to become positive, whereas the likely outcome of a false-positive result is unnecessary therapy with its associated risk for hepatotoxicity. The different thresholds for false results reflect the committee’s recognition that the consequences of false-negative results are generally transient, whereas the consequences of false-positive results may be long lasting. In this case, the sensitivity and specificity of AFB smear microscopy were approximately 70% and ≥90%, respectively, so AFB smear microscopy is recommended.

The recommendation is strong because the quality of evidence provided the committee with moderate confidence in the estimated test characteristics of AFB smear microscopy, and the committee therefore felt certain that the desirable consequences of AFB smear microscopy (ie, an early presumptive diagnosis, initiation of therapy, and possibly less transmission) outweigh the undesirable consequences (ie, cost, burden, effects of false results) in the vast majority of patients.

**Question 6: Should both liquid and solid mycobacterial cultures be performed in persons suspected of having pulmonary TB?**

**Evidence**

A meta-analysis comparing 2 liquid culture methods with solid cultures found that both liquid culture methods were more sensitive (88% and 90%) than the solid culture method (76%) when a combination of conventional solid media with a broth-based method was the reference standard, and also had a shorter time to detection (13.2 and 15.2 days for liquid culture methods versus 25.8 days for the solid culture method) [132]. The specificity of all 3 methods exceeded 99%. Liquid culture medium has a higher contamination rate than solid culture medium due to the growth of bacteria other than...
mycobacteria (4%–9% in the meta-analysis), which interferes with obtaining a valid culture result. This evidence provides low confidence in the estimated test characteristics for 2 reasons. First, there may be selection bias, as many of the studies did not state whether they enrolled consecutive patients. Second, there is indirectness, since the studies address the test characteristics of either test alone but the question is about the tests combined.

Recommendation 6: We suggest that both liquid and solid mycobacterial cultures be performed, rather than either culture method alone, for every specimen obtained from an individual with suspected TB disease (conditional recommendation, low-quality evidence). Remarks: The conditional qualifier applies to performance of both liquid and solid culture methods on all specimens. At least liquid culture should be done on all specimens as culture is the gold standard microbiologic test for the diagnosis of TB disease. The isolate recovered should be identified according to the Clinical and Laboratory Standards Institute guidelines and the American Society for Microbiology Manual of Clinical Microbiology [133, 134].

**Rationale**
Mycobacterial culture is the laboratory gold standard for tuberculosis diagnosis, but the preferred type of cultures is uncertain. Liquid cultures alone are reasonably sensitive and highly specific, but limited by contamination. Solid cultures alone are not sufficiently sensitive to reliably diagnose TB and generally take longer to yield results; however, some Mtb isolates are detected only on solid medium. Performing both liquid and solid cultures likely improves the sensitivity of mycobacterial cultures, while the liquid cultures provide a more rapid answer and the solid cultures serve as a safeguard against contamination. The recommendation is conditional because the quality of evidence provided the committee with limited confidence in the estimated test characteristics of the culture methods; therefore, the committee could not be certain that the desirable consequences of performing both culture methods instead of only one method outweigh the undesirable consequences in the vast majority of patients.

**Question 7: Should NAAT be performed on the initial respiratory specimen in persons suspected of having pulmonary TB?**

**Evidence**
Three meta-analyses were identified and reviewed. The first stratified the performance characteristics of NAAT based upon AFB smear results [135]. When AFB smear microscopy was positive, the sensitivity and specificity of NAAT were 96% and 85%, respectively. Most studies used culture as the reference standard. When AFB smear microscopy was negative, the sensitivity decreased to 66% and the specificity increased to 98.

When further stratified by whether the patient received treatment, the specificity in untreated patients was 97%. The second meta-analysis reported an overall sensitivity of 85% and specificity of 97% and did not stratify according to the results of AFB smear microscopy [136]. There was significant heterogeneity in both meta-analyses. The third meta-analysis stratified the NAAT test characteristics in AFB smear microscopy–negative suspects according to clinical suspicion of tuberculosis [137]. It found that in AFB smear microscopy–negative individuals, a positive NAAT result is beneficial when the clinical suspicion of tuberculosis was intermediate or high (>30%) and a negative NAAT result is of little use in excluding the presence of Mtb. This evidence provides low confidence in the estimated test characteristics because there may be selection bias since many of the studies did not state whether they enrolled consecutive patients with legitimate diagnostic uncertainty and there was significant inconsistency in the meta-analyses.

Recommendation 7: We suggest performing a diagnostic NAAT, rather than not performing a NAAT, on the initial respiratory specimen from patients suspected of having pulmonary TB (conditional recommendation, low-quality evidence). Remarks: In AFB smear–positive patients, a negative NAAT makes TB disease unlikely. In AFB smear–negative patients with an intermediate to high level of suspicion for disease, a positive NAAT can be used as presumptive evidence of TB disease, but a negative NAAT cannot be used to exclude pulmonary TB. Appropriate NAATs include the Hologic Amplified Mycobacteria Tuberculosis Direct (MTD) test (San Diego, California) and the Cepheid Xpert MTB/Rif test (Sunnyvale, California).

**Rationale**
Mycobacterial culture results require at least 1–2 weeks; therefore, rapid diagnostic tests that can be performed within hours are desirable, such as AFB smear microscopy and diagnostic NAAT. Diagnostic NAAT has the added advantage over AFB smear microscopy of being able to distinguish Mtb from nontuberculous mycobacteria. However, NAAT is appropriate only as an adjunct to mycobacterial culture and AFB smear microscopy. It is used as an adjunct to mycobacterial culture because it is not sensitive enough to replace mycobacterial culture for diagnosis and does not produce an isolate, which is needed for phenotypic DST. It is used as an adjunct to AFB smear microscopy because the test characteristics of NAAT are highly variable depending upon the AFB smear results and clinical suspicion.

In AFB smear–positive patients, NAAT yields false-negative results only 4% of the time, indicating that it is reliable for excluding pulmonary TB. In AFB smear–negative patients, clinical suspicion needs to be considered. When there is an intermediate to high level of suspicion for disease, NAAT yields sufficiently few false-positive results that a positive NAAT result can be used as presumptive evidence of TB and guide therapeutic decisions;
however, false-negative results are sufficiently common that NAAT cannot be used to exclude pulmonary TB. When the clinical suspicion for TB is low, NAAT is generally not performed because false-positive results are unacceptably frequent. An algorithm for interpretation and use of NAAT results in conjunction with AFB smear results has been published [138].

The recommendation is conditional because the quality of evidence provided the committee with limited confidence in the estimated test characteristics of NAAT; therefore, the committee could not be certain that the desirable consequences of performing NAAT (ie, promptly diagnosing TB disease and initiating treatment), instead of not performing NAAT, outweigh the undesirable consequences (ie, cost, false-positive results leading to unnecessary treatment, and false-negative results provided false reassurance) in the vast majority of patients.

**Cautions and Limitations**
Laboratory-based diagnostic tests are not a replacement for clinical judgment and experience. A diagnosis of pulmonary tuberculosis can be made in the absence of laboratory confirmation, especially in children [139]. Although there appears to be little increase in accuracy achieved by routinely performing NAAT on multiple specimens rather than on a single specimen, some clinicians may find it beneficial in the diagnosis of individual patients [140, 141]. As an example, the presence of inhibitors can cause false-negative results for some NAATs [142] and, therefore, if a specimen has a positive AFB smear result and a negative NAAT result, evaluation of the sample for the presence of inhibitors should be considered if the NAAT being used is subject to inhibition. If inhibitors are detected, collection of a new specimen for NAAT should be considered. The recommendation for use of NAATs is based on studies of commercial test kits. The data on in-house tests show even greater heterogeneity [143]. If in-house tests are to be used, they should be validated and be shown to have analytical performance accuracy comparable to or better than that of commercial tests.

**Question 8: Should rapid molecular drug susceptibility testing for isoniazid and rifampin be performed as part of the initial diagnostic evaluation for all patients suspected of having pulmonary TB or only in selected subgroups?**

**Evidence**
Rapid molecular DST can be performed via line probe or molecular beacon assays. We evaluated systematic reviews of meta-analyses of 2 line probe assays [144, 145]. Both line probe assays detected rifampin resistance with a sensitivity and specificity of ≥97% and ≥98%, respectively, when conventional, culture-based DST was used as the reference standard. More recently, a molecular-beacon based method for rapid rifampin resistance detection was evaluated in a large international accuracy study [146]. This assay, Xpert MTB/RIF, was >92% sensitive and >99% specific for detection of rifampin resistance when performed on a single specimen; the sensitivity increased to >97% when performed on 3 specimens [146]. Despite its good sensitivity and specificity, the PPV of rapid molecular DST for the detection of rifampin resistance is low in populations with a low prevalence of drug resistance (Supplementary Table 9) [147].

One of the assays also detects isoniazid resistance. It identified isoniazid resistance with a sensitivity and specificity of 84% and 99%, respectively, when culture-based DST is used as the reference standard. However, when the meta-analysis was performed on a subgroup of studies that evaluated a newer version of the assay, the sensitivity increased to approximately 90%. This indicates that in appropriate subgroups of patients, false-positive and false-negative results occur in 1% and 10% of patients, respectively. In contrast to rifampin resistance, the PPV of a test indicating isoniazid resistance is quite high, a reflection of isoniazid resistance being fairly prevalent in the United States (approximately 8%) [144].

This evidence provides moderate confidence in the estimated sensitivities and specificities among patient subgroups with increased rates of drug resistance. The confidence is moderate instead of high because the absence of reporting that patients were enrolled consecutively suggests that there is a risk of bias.

**Recommendation 8:** We recommend performing rapid molecular DST for rifampin with or without isoniazid using the respiratory specimens of persons who are either AFB smear positive or Hologic Amplified MTD positive and who meet one of the following criteria: (1) have been treated for tuberculosis in the past, (2) were born in or have lived for at least 1 year in a foreign country with at least a moderate tuberculosis incidence (≥20 per 100 000) or a high primary MDR-TB prevalence (≥2%), (3) are contacts of patients with MDR-TB, or (4) are HIV infected (strong recommendation, moderate-quality evidence). Remarks: This recommendation specifically addresses patients who are Hologic Amplified MTD positive because the Hologic Amplified MTD NAAT only detects TB and not drug resistance; it is not applicable to patients who are positive for types of NAAT that detect drug resistance, including many line probe assays and Cepheid Xpert MTB/RIF.

**Rationale**
Conventional, culture-based DST is the laboratory gold standard [134, 148, 149]. It is performed routinely any time Mtb complex is isolated in culture. Drug susceptibility testing is essential because treatment success for patients with MDR-TB (can reach 75% or higher [150, 151]) is dependent upon patients being treated with an effective antimicrobial regimen [152]. An important limitation of culture-based DST, however, is that it can take >2 weeks to grow the isolate that is necessary for testing.
Rapid molecular DST addresses this limitation. It can be performed within hours, enabling earlier initiation of an appropriate antimicrobial regimen. Rapid molecular DST is an adjunct and not a replacement for culture-based DST because it only evaluates susceptibility to rifampin and occasionally isoniazid. Nonetheless, detection of rifampin resistance is helpful to clinicians because it is a good surrogate for MDR-TB in locations where rifampin monoresistance is uncommon. However, an important limitation is that the PPV is expected to be lower in the United States than in areas where rifampin resistance is more common [153–155].

The committee recommends rapid molecular DST only for subgroups in which drug resistance is more likely, as the PPV for rifampin resistance testing is low in populations with a low prevalence of drug resistance. Examples of appropriate persons for testing include those who are NAAT or AFB smear positive and meet one of the following criteria: (1) have been treated for tuberculosis in the past, (2) were born in or have lived for at least 1 year in a foreign country with at least a moderate tuberculosis incidence (≥20 per 100 000) or a high primary MDR-TB prevalence (≥2%), (3) are contacts of patients with MDR-TB, or (4) are HIV infected [154, 156–158].

The sensitivity and specificity of rapid molecular DST for detecting rifampin resistance are both >97%, indicating that false-positive and false-negative results occur <3% of the time; thus, rapid molecular DST can be used confirm or exclude rifampin resistance in respiratory specimens. The sensitivity and specificity of rapid molecular DST for detecting isoniazid resistance are estimated to be 90% and 99%, respectively, indicating that false-positive and false-negative results occur roughly 1% and 10% of the time, respectively; thus, rapid molecular DST can be used to confirm isoniazid resistance in respiratory specimens, but not exclude it.

The recommendation is strong because the moderate-quality evidence provided the committee with sufficient confidence in the test characteristics to be certain that the benefits of rapid molecular DST (ie, early identification of possible MDR-TB and initiation of an appropriate antimicrobial regimen) outweigh the costs and burden of testing in the overwhelming majority of patients who have increased risk for drug resistance.

Cautions and Limitations
Line probe and molecular beacon assays have not been sufficiently validated for use on specimens other than respiratory specimens. The recommendation for line probe assays and molecular beacon on respiratory specimens is based upon studies of commercial test kits, only one of which is currently approved by the FDA: the molecular beacon–based method, Xpert MTB/RIF. It is the only FDA-approved assay and it integrates diagnosis of TB and detection of rifampin resistance. If this test is used for the diagnosis of TB, a rifampin resistance result is automatically provided regardless of patient risk.

Other assays for rapid detection of drug resistance using alternative molecular techniques (eg, automated real-time polymerase chain reaction [PCR] with sequencing, loop-mediated isothermal amplification [LAMP]) are being developed. These assays are promising, but are not yet commercially available [146, 159, 160]. The data on in-house tests show substantial heterogeneity [161]. If in-house tests are to be used, they should be validated and shown to have performance accuracy at least comparable to that of commercial tests. The same cautions also apply to new commercial assays that may become available in the near future.

Some clinicians and health departments may opt for broader use of the molecular detection of drug resistance assays than recommended above, especially in regions where MDR-TB is more common. Because the prevalence of rifampin resistance (and therefore MDR-TB) is low in the United States, the PPV of Xpert MTB/RIF and other assays for rifampin resistance will be lower than in settings where Xpert MTB/RIF has been predominantly studied. Therefore, confirmation of a positive test result for rifampin resistance has been recommended [147]. To confirm a positive result, genetic loci associated with rifampin resistance (to include rpoB), as well as isoniazid resistance (to include inhA and katG), should be sequenced to assess for MDR-TB. If mutations associated with rifampin resistance are confirmed, rapid molecular testing for other known mutations associated with drug resistance (to first-line and second-line drugs) is needed for healthcare providers to select an optimally effective treatment regimen. All molecular testing should prompt growth-based DST.

Alternative methods for rapid molecular DST are being developed and other technologies are likely to become available in the near future (eg, automated real-time PCR with sequencing, LAMP) [162]. It is possible that these techniques will be sufficiently sensitive to be used for AFB smear–negative specimens. Laboratories in the United States should only use tests approved by the FDA or tests that have been produced and validated in accord with applicable FDA and Clinical Laboratory Improvement Amendments regulations.

Question 9: Should respiratory specimens be collected from children with suspected pulmonary TB disease?

Evidence
Respiratory specimens that can be collected from children include gastric aspirates; sputum collected by spontaneous expectoration, induction, or nasopharyngeal aspiration; and bronchoalveolar lavage (BAL). Gastric aspirates involve intubating the stomach after an overnight fast to collect swallowed sputum before the stomach empties. Collection of specimens on 3 consecutive mornings from patients with suspected pulmonary TB provides a diagnostic yield of up to 40%–50%, with higher yields for infants (up to 90%), symptomatic children, and children with extensive disease (up to 77%), using a clinical
diagnosis of TB disease in a low prevalence country as criteria for the diagnosis of TB disease [163–165]. Meticulous attention to detail during the collection and processing of the specimen can improve yield (details are provided at http://www.curytb-center.ucsf.edu/pediatric_tb/). Sputum collected from children by nasopharyngeal aspiration or sputum induction with a bronchodilator has a yield of 20%–30% [166], whereas BAL in children with pulmonary TB has a yield of 10%–22% [167]. These estimates of diagnostic yield are based upon moderate-quality evidence—accuracy studies for which it was not documented whether the subjects were enrolled consecutively.

Recommendation 9: We suggest mycobacterial culture of respiratory specimens for all children suspected of having pulmonary TB (conditional recommendation, moderate-quality evidence).

Remarks: In a low incidence setting like the United States, it is unlikely that a child identified during a recent contact investigation of a close adult/adolescent contact with contagious TB was, in fact, infected by a different individual with a strain with a different susceptibility pattern. Therefore, under some circumstances, microbiological confirmation may not be necessary for children with uncomplicated pulmonary TB identified through a recent contact investigation if the source case has drug-susceptible TB.

**Rationale**

Despite the observation that less than half of pediatric specimens yield a positive culture, the committee judged that the desirable consequences of mycobacterial cultures of respiratory specimens outweigh the undesirable consequences of specimen collection in children for several reasons. First, a positive mycobacterial culture is likely to be reassuring to parents and staff that the diagnosis of tuberculosis is correct. Second, cultures are necessary for DST, which is particularly important in situations in which TB drug resistance is prevalent. Third, susceptibility data are not always available from the presumed source case. Finally, after-the-fact culture collection in the face of treatment failure may have even lower yield than sampling a drug-naive child. Specimens that can be used for mycobacterial culture include gastric aspirates, sputum, and BAL; the panel decided that there was insufficient evidence to advocate one collection method over another.

With respect to the need for DST, overtreatment for presumed drug-resistant TB may lead to unnecessary toxicities and cost, while undertreatment due to unidentified drug resistance may lead to treatment failure, risk of dissemination, and even death. While it is tempting to avoid collection from the child contact when a putative source case is identified (especially when susceptibility results are already available), prior case series indicate that 2%–10% of children have susceptibility patterns that differ from the presumed source case [168] and more recent US studies have found up to 15% discordance of molecular fingerprinting between the isolates collected from children with culture-proven TB compared to their presumed source case [169, 170]. In contrast, no discordance was found between pediatric TB cases and their presumed source cases from 2000 to 2004 in Houston [171].

The recommendation is conditional because the moderate quality of evidence provided the committee with insufficient confidence in the estimated diagnostic yield; thus, the committee felt uncertain that a diagnosis was rendered frequently enough that the desirable consequences of collecting respiratory specimens (ie, confirming the diagnosis of TB, obtaining an isolate for DST) outweigh the undesirable consequences (ie, cost, burden, effects of false results) in the vast majority of children with suspected pulmonary TB.

The highest yields for gastric aspirates are in the youngest infants, in children with extensive or symptomatic disease, and for the first gastric aspirate collected. While there are situations where a presumed source case is not the child’s true source case, in the case of a very recent contact investigation of a household-type contact with pan-susceptible disease, performing only one gastric aspirate or relying on the source case susceptibility may be appropriate. For infants, immunocompromised hosts, children with extensive, disseminated, or extrapulmonary disease, exposure to other potential source cases, or risk of drug-resistance, respiratory specimens should be collected. Studies comparing the yield of gastric aspirates to sputum have shown discrepant results. Selection of an appropriate respiratory specimen (ie, gastric aspirates, spontaneous or induced sputa, or rarely bronchoalveolar lavage) should be based upon the expertise of the clinic and provider, the patient’s age and developmental level, and the likelihood of an alternative diagnosis. Most investigators have not found increased yield for bronchoalveolar lavage compared to gastric aspirates. Bronchoscopy should be reserved for situations where an alternative diagnosis is being considered or when the anatomy is unclear.

**Cautions and Limitations**

Gastric aspirates are rarely AFB smear positive and the yield of cultures is suboptimal in children with pulmonary TB; thus, gastric aspirate culture results are helpful only if they are positive. Negative results should not dissuade the provider from empirically treating tuberculosis in children in the appropriate clinical setting. Gastric aspirate, sputum induction, and nasopharyngeal aspiration in children are not comfortable and not without financial cost. The procedures have modest risk (bleeding from the nose, bronchospasm, airway intubation).

**Question 10:** Should sputum induction or flexible bronchoscopic sampling be the initial respiratory sampling method for adults with suspected pulmonary TB who are either unable to expectorate sputum or whose expectorated sputum is AFB smear microscopy negative?
Evidence

We identified 6 studies [172–177] that compared the diagnostic yield of induced sputum with the yield of specimens obtained by flexible bronchoscopy, using a positive mycobacterial culture or evidence of a response to therapy as criteria for the diagnosis of pulmonary TB. Five of the 6 studies demonstrated a higher yield from induced sputum than bronchoscopy, with the remaining study [176] demonstrating a similar yield. The diagnostic yield of induced sputum increases with multiple specimens, with detection rates by AFB smear microscopy of 91%–98% and mycobacterial culture of 99%–100% reported when 3 or more specimens are obtained [178].

Two cost-analysis studies favored sputum induction over bronchoscopy [172, 174]. In the first study, direct costs for bronchoscopy measured in Canadian dollars were $187.60, compared with $22.22 for sputum induction [172]. In the second study, induced sputum was about one-third the cost of flexible bronchoscopy, and the most cost-effective strategy was 3 induced sputa without bronchoscopy [174].

Our confidence in the accuracy of the study results is low because there was a risk of bias and indirectness. With respect to risk of bias, most of the studies did not report whether or not consecutive patients were enrolled. Supporting this concern, the variability of prevalence among studies suggests that the degree of diagnostic uncertainty likely differed among studies. With respect to indirectness, there appeared to be indirectness of the intervention because the studies varied in the number of specimens collected (from 1 to 3 per patient), the concentrations of hypertonic saline, the type of nebulizers, and the culture techniques.

Recommendation 10: We suggest sputum induction rather than flexible bronchoscopic sampling as the initial respiratory sampling method for adults with suspected pulmonary TB who are either unable to expectorate sputum or whose expectorated sputum is AFB smear microscopy negative (conditional recommendation, low-quality evidence).

Rationale

Induced sputum has equal or greater diagnostic yield than bronchoscopic sampling, has fewer risks, and is less expensive. These features all favor induced sputum as the initial respiratory sampling method in patients with suspected pulmonary TB who are either unable to expectorate sputum or whose expectorated sputum is AFB smear microscopy negative. The committee recognizes that a potential advantage of bronchoscopy over sputum induction is the possibility of making a rapid presumptive diagnosis of tuberculosis by performing biopsies and identifying typical histopathologic findings, but felt that the balance of the upsides to downsides of induced sputum outweighed that of bronchoscopic sampling. The recommendation is conditional because the quality of evidence does not provide sufficient confidence in the study results for the committee to be absolutely certain that the balance of desirable to undesirable consequences favors induced sputum over bronchoscopy.

Question 11: Should flexible bronchoscopic sampling be performed in adults with suspected pulmonary TB from whom a respiratory sample cannot be obtained via induced sputum?

Evidence

Numerous studies reported the diagnostic yield of respiratory specimens obtained by flexible bronchoscopy, using a positive mycobacterial culture or evidence of a response the therapy as criteria for the diagnosis of pulmonary TB [172–175, 178–182]. Generally speaking, bronchoscopic sampling appears to have a diagnostic yield of 50%–100% when based on culture in patients suspected of having pulmonary TB. This yield appears unaffected by HIV infection, with bronchoscopy leading to an early presumptive diagnosis of TB in 34%–48% of HIV-infected patients according to 2 studies [183, 184]. In one study, bronchial washings had the same culture yield (9%) as BAL but higher frequency of positive AFB smears (26% vs. 4%) [235]. Bronchoscopic brushings yield AFB smear–positive results in 9%–56% [180, 185].

Transbronchial biopsy (TBB) provides histopathologic findings suggestive of pulmonary TB in 42%–63% of specimens from smear-negative HIV-uninfected patients [183, 186]. HIV-infected patients are less likely (9%–19%) to demonstrate granulomas on TBB [183, 186], although in 2 studies TBB was the exclusive means of diagnosing pulmonary TB in 10%–23% of patients [183, 184].

Our confidence in the accuracy of these estimated diagnostic yields is very low because most of the studies did not report whether or not consecutive patients were enrolled, the range of reported diagnostic yields is wide, and the studies varied in how specimens were collected (bronchial aspirates and/or BAL and/or bronchial brushings and/or TBB) and the culture techniques.

Recommendation 11: We suggest flexible bronchoscopic sampling, rather than no bronchoscopic sampling, in adults with suspected pulmonary TB from whom a respiratory sample cannot be obtained via induced sputum (conditional recommendation, very low-quality evidence). Remarks: In the committee members’ clinical practices, BAL plus brushings alone are performed for most patients; however, for patients in whom a rapid diagnosis is essential, transbronchial biopsy is also performed.

Rationale

The committee judged that the desirable consequences of bronchoscopic sampling outweigh the undesirable consequences among patients with suspected pulmonary TB from whom respiratory samples could not be obtained noninvasively. The most important reason to perform bronchoscopy in a patient with possible pulmonary TB is to differentiate TB disease from alternative diseases. Another reason to perform bronchoscopy is to obtain specimens for cultures that provide isolates for DST.
Empiric treatment for presumed drug-resistant TB may lead to unnecessary toxicities and cost if the patient actually has drug-sensitive TB, while empiric treatment for drug-sensitive TB may lead to treatment failure, risk of dissemination, and even death if the patient actually has drug-resistant TB. Moreover, delayed diagnosis of drug resistance will prolong therapy and increase risk of default. Bronchoscopy also provides the opportunity of obtaining a rapid presumptive diagnosis by identifying histopathologic findings consistent with tuberculosis. These benefits of bronchoscopic sampling were thought to outweigh the risks of bronchoscopy and the accompanying sedation, as well as the costs and burdens.

The recommendation is conditional, reflecting the guideline development committee’s uncertainty that the desirable consequences of bronchoscopy outweigh the undesirable consequences in many situations. Reasons for the committee’s uncertainty included the highly variable estimates of diagnostic yield, the very low quality of evidence, recognition that bronchoscopy is an invasive procedure and the risk of harm varies according to the patient’s clinical condition, and recognition that the feasibility of timely bronchoscopy varies according to the clinical setting. For example, in the context of a public health clinic, the benefits of obtaining a bronchoscopy may not justify the thousands of dollars that it will cost due to professional fees, hospital charges, pathology costs, and laboratory fees, or the days to weeks of delays that will be necessary to refer the patient to a pulmonologist for bronchoscopy. In some situations, the potential harm associated with delayed diagnosis may warrant empiric initiation of therapy based upon a reasonable suspicion of TB disease.

Question 12: Should postbronchoscopy sputum specimens be collected from adults with suspected pulmonary TB?

Evidence
Postbronchoscopy sputum specimens are typically sent for AFB smear microscopy and mycobacterial culture. Postbronchoscopy AFB smears have a diagnostic yield of 9%–73% and postbronchoscopy mycobacterial cultures have a yield of 35%–71% according to multiple studies [182, 185, 187, 188]. In HIV-infected patients, the yield of postbronchoscopy sputum cultures was 80% in a single study [186]. Our confidence in the accuracy of the estimated diagnostic yields is low because most of the studies did not report whether or not consecutive patients were enrolled and the diagnostic yields reported varied greatly as indicated by the wide ranges described above. Recommendation 12: We suggest that postbronchoscopy sputum specimens be collected from all adults with suspected pulmonary TB who undergo bronchoscopy (conditional recommendation, low-quality evidence). Remarks: Postbronchoscopy sputum specimens are used to perform AFB smear microscopy and mycobacterial cultures.

Rationale
The rationale for postbronchoscopy sputum collection is the same as that described above for the bronchoscopic collection of respiratory specimens.

Question 13: Should flexible bronchoscopic sampling be performed in adults with suspected miliary TB and no alternative lesions that are accessible for sampling whose induced sputum is AFB smear microscopy negative or from whom a respiratory sample cannot be obtained via induced sputum?

Evidence
Specimens obtained via bronchoscopy can undergo AFB smear microscopy, mycobacterial culture, NAAT, and histopathological analysis. There is a paucity of evidence regarding the diagnostic characteristics of various types of bronchoscopic specimens (ie, washings, BAL, brushings, TBB) obtained from patients with possible miliary TB. It has been reported that bronchial washings, brushings, and TBB have diagnostic yields of 14% [234], 27%–78% [181, 189, 190], and 32%–75% [181, 189, 190], respectively. The diagnostic yield of BAL has not been reported. Our confidence in the accuracy of the estimated diagnostic yields is very low because most of the studies did not report whether or not consecutive patients were enrolled (the prevalence of miliary TB ranged from 55% to 90% in the studies, suggesting that the degree of diagnostic uncertainty differed among the studies), the ranges of diagnostic yields were wide for both brushings and TBB, reflecting the variable results reported by the individual studies, and the studies varied in the technique used to perform the sampling (particularly TBB) and the number of specimens collected.

Recommendation 13: We suggest flexible bronchoscopic sampling, rather than no bronchoscopic sampling, in adults with suspected miliary TB and no alternative lesions that are accessible for sampling whose induced sputum is AFB smear microscopy negative or from whom a respiratory sample cannot be obtained via induced sputum (conditional recommendation, very low-quality evidence). Remarks: Bronchoscopic sampling in patients with suspected miliary TB should include bronchial brushings and/or TBB, as the yield from washings is substantially less and the yield from BAL unknown. For patients in whom it is important to provide a rapid presumptive diagnosis of tuberculosis (ie, those who are too sick to wait for culture results), TBB is both necessary and appropriate.

Rationale
The rationale for bronchoscopic sampling in individuals with suspected miliary TB and no alternative lesions that are accessible for sampling (eg, enlarged lymph nodes or draining lesions) whose induced sputum is AFB smear negative or from whom a respiratory
sample cannot be obtained via induced sputum is essentially the same as that described above for patients with suspected pulmonary TB from whom a respiratory sample cannot be obtained via induced sputum. That is, it is important to differentiate miliary TB from other diseases and also to obtain specimens for mycobacterial culture because cultures provide isolates for DST, which may prevent unnecessary drug toxicities and cost, mitigate treatment failure, and reduce the risk of dissemination and death. Moreover, bronchoscopy also provides the opportunity of obtaining a rapid presumptive diagnosis by identifying histopathologic findings consistent with tuberculosis. These benefits outweigh the risks of both bronchoscopy and the accompanying sedation.

The recommendation is conditional, reflecting the guideline development committee’s uncertainty that the desirable consequences of bronchoscopy outweigh the undesirable consequences and its recognition that the balance of desirable and undesirable consequences depends upon clinical context. Reasons for the committee’s uncertainty included the variable estimates of the diagnostic yield and very low quality of evidence. Clinical considerations that may affect the balance of desirable and undesirable effects include the patient’s condition (ie, bronchoscopy is an invasive procedure and the risk of harm varies according to the patient’s clinical condition) and the availability, feasibility, and cost of timely bronchoscopy in a particular clinical setting.

**DIAGNOSTIC APPROACH: TESTING FOR SUSPECTED EXTRAPULMONARY TB**

Randomized trials and controlled observational studies that directly compared diagnostic tests or approaches for extrapulmonary TB and measured patient-important outcomes have not been performed. Therefore, the recommendations in this section are based upon data that describe how accurate a diagnostic test is at confirming or excluding extrapulmonary TB, coupled with evidence that the diagnosis of extrapulmonary TB leads to therapy that improves patient-important outcomes. Tests used to diagnose extrapulmonary TB are described in Supplementary Figure 1.

With respect to the evidence that the diagnosis of extrapulmonary TB leads to therapy that improves patient-important outcomes, trials directly comparing treatment with no treatment will never be done. However, indirect evidence from patients with pulmonary TB (described above) and evidence from patients with disseminated TB suggests that extrapulmonary TB is treatable with high cure rates in most drug-susceptible cases and that untreated extrapulmonary TB has significant morbidity and mortality, particularly meningeal TB [191]. As an example, an observational study that excluded patients with tuberculous meningitis found that mortality due to disseminated TB fell from 100% to <5% with the introduction of isoniazid-based antimicrobial regimens [192]. The large mortality reduction in this study would probably have been less dramatic if patients with tuberculous meningitis had been included, as 2 case series that included patients with tuberculous meningitis found that mortality due to disseminated tuberculosis is 20% [193, 194]. Nevertheless, mortality is clearly improved with a large magnitude of effect. This evidence constitutes moderate-quality evidence that treatment of extrapulmonary TB improves mortality because there are observational studies with a very large magnitude of effect, but the increase in confidence in the results due to the large magnitude of effect is mitigated by the indirectness of the population.

**Question 14: Should cell counts and chemistries be performed on amenable (ie, liquid) specimens collected from sites of suspected extrapulmonary TB?**

**Evidence**

We identified no studies that reported the sensitivity and specificity of cell counts and chemistries in the identification of extrapulmonary TB. Therefore, the committee used its collective clinical experience to inform its recommendation. Clinical experience constitutes very low-quality evidence.

**Recommendation 14:** We suggest that cell counts and chemistries be performed on amenable fluid specimens collected from sites of suspected extrapulmonary TB (conditional recommendation, very low-quality evidence). Remarks: Specimens that are amenable to cell counts and chemistries include pleural, cerebrospinal, ascitic, and joint fluids.

**Rationale**

Cell counts and chemistries can be performed in hours, are inexpensive, and are technically simple. Any risks are related to the sampling procedure. Although their sensitivity and specificity for extrapulmonary TB have not been reported, the committee suspects that the sensitivity is moderate to high and the specificity is poor if interpreted alone, but substantially better if interpreted in the context of the clinical setting, radiographic findings, and other laboratory results. Most importantly, it is believed that cell counts and chemistries can provide useful information to guide the clinician toward either confirmatory diagnostic testing for tuberculosis or diagnostic testing for alternative etiologies; this alone provides enough benefit to justify the costs of the additional tests. The strength of the recommendation is conditional because it is believed that the balance of the benefits of the additional information versus the cost of the testing may be finely balanced in some clinical situations, and the quality of evidence provides little confidence in the estimates upon which the committee based its judgments.

**Question 15: Should adenosine deaminase (ADA) and free IFN-γ levels be measured on specimens collected from sites of suspected extrapulmonary TB?**

**Evidence**

Test characteristics of ADA in the diagnostic evaluation of meningeal, pleural, peritoneal, and pericardial tuberculosis
have been reported in meta-analyses of accuracy studies. Two meta-analyses estimated the sensitivity and specificity of an elevated ADA level in the cerebrospinal fluid [195, 196]. The first meta-analysis included 10 studies and found a sensitivity and specificity of 79% and 91%, respectively [195], using final clinical diagnosis, consistent pathology/cytology, or microbiologic confirmation as the reference standard. Most of the studies used a threshold of 9 U/L or 10 U/L to define an elevated ADA. The second meta-analysis included 13 studies and showed that the sensitivity and specificity are excellently sensitive to the threshold used to define an elevated ADA level [196]. If 4 U/L was used as the threshold, the sensitivity and specificity were >93% and <80%, respectively. In contrast, if 8 U/L was used as the threshold, the sensitivity and specificity were <59% and >96%, respectively.

Five meta-analyses that included 9–63 studies estimated that the sensitivity and specificity of an elevated ADA level in the pleural fluid are 89%–99% and 88%–97%, respectively, with all but one of the meta-analyses estimating that the specificity is ≥90% [197–201]. A more recent meta-analysis reported similar sensitivity and specificity [202]. Final clinical diagnosis, consistent pathology/cytology, or microbiologic confirmation were used as the reference standard in most studies. Thresholds used to define an elevated ADA level ranged from 10 U/L to 71 U/L, with most clustering around 40 U/L.

A meta-analysis of 31 studies estimated that the sensitivity and specificity of an elevated ADA level in pericardial fluid are 88% and 83%, respectively [203]. The threshold to define an elevated ADA level was 40 U/L. Finally, a meta-analysis of 4 studies estimated that the sensitivity and specificity of an elevated ADA level in peritoneal fluid are 100% and 97%, respectively [204]. The threshold used to define an elevated ADA level ranged from 36 U/L to 40 U/L.

The test characteristics of free IFN-γ levels have not been as extensively studied. A meta-analysis of 6 studies estimated that the sensitivity and specificity of an elevated free IFN-γ level in peritoneal fluid are 93% and 99%, respectively [205]. The threshold used to define an elevated IFN-γ level ranged from 0.35 U/L to 9 U/L or 20 pg/mL to 112 pg/mL. A meta-analysis of 22 studies estimated that the sensitivity and specificity of an elevated free IFN-γ level in pleural fluid are 89% and 97%, respectively [206]. The threshold used to define an elevated IFN-γ level ranged from 0.3 U/L to 10 U/L or 12 pg/mL to 300 pg/mL. We did not identify any studies that looked at the test characteristics of free IFN-γ levels on pericardial fluid or cerebrospinal fluid.

No studies were identified that reported the test characteristics of using both ADA and free IFN-γ to evaluate specimens from patients with suspected extrapulmonary TB. This evidence provides low confidence in the accuracy of the estimated test characteristics for both ADA and free IFN-γ levels. The studies did not report whether consecutive patients were enrolled (ie, risk of bias) and the studies reported variable results (ie, inconsistency), probably due in large part to the different thresholds used to define an elevated level.

Recommendation 15a: We suggest that ADA levels be measured, rather than not measured, on fluid collected from patients with suspected pleural TB, TB meningitis, peritoneal TB, or pericardial TB (conditional recommendation, low-quality evidence).

Recommendation 15b: We suggest that free IFN-γ levels be measured, rather than not measured, on fluid collected from patients with suspected pleural TB or peritoneal TB (conditional recommendation, low-quality evidence).

Rationale

Neither the ADA level nor the IFN-γ level provide a definitive diagnosis of extrapulmonary TB disease; rather, they provide supportive evidence that must be interpreted in the entire clinical context. In any type of diagnostic testing for extrapulmonary TB, both false-negative results and false-positive results have important consequences. False-negative results delay diagnosis and treatment while diagnostic testing continues, whereas false-positive results may lead to unnecessary therapy and the associated risks of drug toxicity and cost. Therefore, it is desirable for diagnostic tests to be both sensitive and specific.

Our committee made the judgment that measurement of ADA levels and free IFN-γ levels are indicated if false-negative results occur <30% of the time (ie, sensitivity is ≥70%) and false-positive results occur <20% of the time (ie, specificity is ≥80%). The different thresholds for false results reflect the committee’s recognition that the consequences of false-negative results are generally transient, whereas the consequences of false-positive results may be long lasting. In this case, the sensitivity and specificity of ADA were ≥79% and ≥83%, respectively, for detecting TB in cerebrospinal fluid, pleural fluid, peritoneal fluid, and pericardial fluid, so ADA measurements are recommended in these fluids. Similarly, the sensitivity and specificity of free IFN-γ measurements were ≥89% and ≥97%, respectively, for detecting TB in pleural fluid and peritoneal fluid, so free IFN-γ measurements are recommended in these fluids.

The recommendations are conditional because the low quality of evidence does not provide sufficient confidence in the estimated sensitivities and specificities for the committee to be certain that the balance of desirable to undesirable consequences favors testing and obtaining the specimens to test usually requires an invasive procedure and, therefore, the balance of benefits versus risks may vary substantially depending upon the clinical condition of the patient. Furthermore, the committee recognized that these tests often required the services of an off-site laboratory, and that standards were variable across laboratories and across published studies.
Question 16: Should AFB smear microscopy be performed on specimens collected from sites of suspected extrapulmonary TB?

Evidence

The diagnostic yield and sensitivity of AFB smear microscopy tend to be lower in extrapulmonary TB than pulmonary TB because the former is paucibacillary. Accuracy studies indicate that AFB smear microscopy has a sensitivity of 0–10%, 14%–39%, 10%–30%, <5%, and 0–42% in pleural fluid (Supplementary Table 10), pleural tissue (Supplementary Table 10), urine (Supplementary Table 11), cerebrospinal fluid (Supplementary Table 12), peritoneal fluid (Supplementary Table 13), and pericardial fluid (Supplementary Table 14), respectively [207–218], when final clinical diagnosis, consistent pathology/cytology, or microbiologic confirmation is used as the reference standard. In contrast, the specificity of AFB smear microscopy tends to be quite high as described for pulmonary TB (≥90%). This evidence provides very low confidence in the estimated test characteristics because many studies do not report enrolling consecutive patients, most studies were small with few samples, and ranges are wide due to variable results from the individual studies.

Recommendation 16: We suggest that AFB smear microscopy be performed, rather than not performed, on specimens collected from sites of suspected extrapulmonary TB (conditional recommendation, very low-quality evidence). Remarks: A positive result can be used as evidence of extrapulmonary TB and guide decision making because false-positive results are unlikely. However, a negative result may not be used to exclude TB because false-negative results are exceedingly common.

Rationale

AFB smear microscopy provides the opportunity for early diagnosis and treatment. The estimated specificity of ≥90% for AFB smear in the diagnosis of extrapulmonary TB indicates that false-positive results occur only ≤10% of the time; thus, if a positive AFB smear result is obtained, it is reasonable to assume that infection is present and to act accordingly. In contrast, the estimated sensitivity of <50% for all specimen types indicates that false-negative results are more common than true-negative results and, thus, a negative AFB smear result should not be used to exclude extrapulmonary TB; additional diagnostic testing is indicated.

Even though a positive AFB smear result is infrequent, the committee judged the benefits of early diagnosis (early initiation of treatment, potential to reduce transmission) to outweigh the cost and burden of AFB smear microscopy. The recommendation is conditional because the very low quality of evidence does not provide sufficient confidence in the estimated sensitivities and specificities for the committee to be certain that the balance of desirable to undesirable consequences favors testing and obtaining the specimens to test usually requires an invasive procedure and, therefore, the balance of benefits versus risks may vary substantially depending upon the clinical condition of the patient.

Question 17: Should mycobacterial cultures be performed on specimens collected from sites of suspected extrapulmonary TB?

Evidence

Accuracy studies indicate that mycobacterial culture has a sensitivity of 23%–58%, 40%–58%, 80%–90%, 45%–70%, 45%–69%, and 50%–65% in pleural fluid (Supplementary Table 10), pleural tissue (Supplementary Table 10), urine (Supplementary Table 11), cerebrospinal fluid (Supplementary Table 12), peritoneal fluid (Supplementary Table 13), and pericardial fluid (Supplementary Table 14), respectively [207, 208, 211, 213–216, 219–223], when final clinical diagnosis, consistent pathology/cytology, or microbiologic confirmation is used as the reference standard. The specificity of mycobacterial culture tends to be comparatively higher than the sensitivity (>97%). This evidence provides low confidence in the estimated test characteristics because many studies do not report enrolling consecutive patients and most studies were small with few samples.

Recommendation 17: We recommend that mycobacterial cultures be performed, rather than not performed, on specimens collected from sites of suspected extrapulmonary TB and obtaining the specimens to test usually requires an invasive procedure and, therefore, the balance of benefits versus risks may vary substantially depending upon the clinical condition of the patient.

Rationale

The estimated specificity of >97% for mycobacterial cultures in the diagnosis of extrapulmonary TB indicates that false-positive results occur only <3% of the time; thus, a positive mycobacterial culture is a reliable indicator that infection is present. In contrast, the estimated sensitivity of mycobacterial culture is widely variable depending upon the specimen type. However, even the specimen that provides the highest sensitivity (urine has an 80%–90% sensitivity for the diagnosis of urinary TB) provides false-negative results 10%–20% of the time. Such frequent false-negative results suggest that mycobacterial cultures should not be used to exclude extrapulmonary TB; additional diagnostic testing is indicated.

The committee judged the diagnostic yield and benefits of mycobacterial culture sufficient to outweigh the cost and burden. Importantly, positive mycobacterial cultures are the only way to obtain isolates for DST. Empiric treatment for presumed drug-resistant TB may lead to unnecessary toxicities and cost if the patient actually has drug-susceptible TB, whereas empiric treatment for drug-susceptible TB may lead
to treatment failure, risk of dissemination, and even death if the patient actually has drug-resistant TB. Moreover, delayed diagnosis of drug resistance will prolong therapy and increase risk of default.

The recommendation is strong despite the low quality of evidence because the committee is certain that the balance of desirable to undesirable consequences favors mycobacterial culture. This certainty reflects the committee's recognition of the importance of obtaining mycobacterial isolates for DST compared with the costs and burdens of performing the cultures, and the belief that additional data would not alter the balance of desirable and undesirable consequences in the overwhelming majority of patients.

**Question 18: Should NAAT be performed on specimens collected from sites of suspected extrapulmonary TB?**

**Evidence**

Meta-analyses have been published for the use of NAAT in suspected pleural and meningeal tuberculosis [224, 225]. Most studies used final clinical diagnosis, consistent pathology/cytology, or microbiologic confirmation as the reference standard. Nucleic acid amplification performed on pleural fluid and cerebrospinal fluid has a sensitivity of 56% and 62%, respectively, indicating false-negative rates of 44% and 38%, respectively. In contrast, the specificity of NAAT is high for both pleural fluid and cerebrospinal fluid (98% for both), indicating that only about 2% of positive results are false-positives. Individual studies have been published describing the test characteristics of nucleic acid amplification on other body fluids and tissues. The studies showed considerable variability in the sensitivity and specificity based upon the disease site; generally speaking, the sensitivity was usually <90%, while the specificity was >95% (Supplementary Table 15) [191, 226]. This evidence provides very low confidence in the estimated test characteristics because many studies do not report enrolling consecutive patients, findings were inconsistent as exemplified by the wide ranges, and the studies were small with few samples.

Recommendation 18: We suggest that NAAT be performed, rather than not performed, on specimens collected from sites of suspected extrapulmonary TB (conditional recommendation, very low-quality evidence). Remarks: A positive NAAT result can be used as evidence of extrapulmonary TB and guide decision making because false-positive results are unlikely. However, a negative NAAT result may not be used to exclude TB because false-negative results are exceedingly common.

**Rationale**

NAAT cannot replace mycobacterial culture for diagnosis because it is not sensitive enough and it does not produce an isolate, which is needed for DST. However, NAAT is appropriate as an adjunct to mycobacterial culture because mycobacterial culture results require at least 1–2 weeks, but NAAT can be performed within hours, thereby offering the opportunity for early diagnosis and treatment. The committee felt that NAAT gives positive results frequently enough that the potential benefits outweigh the costs and burden of testing. Moreover, the committee felt that if the test results are applied correctly (ie, a positive NAAT result is considered adequate to confirm extrapulmonary TB, but a negative NAAT result is not used to exclude extrapulmonary TB), then the risks associated with false results are minimal.

The recommendation is conditional because the very low quality of evidence does not provide sufficient confidence in the estimated sensitivities and specificities for the committee to be certain that the balance of desirable to undesirable consequences favors testing and obtaining the specimens to test usually requires an invasive procedure and, therefore, the balance of benefits versus risks may vary substantially depending upon the clinical condition of the patient.

**Cautions and Limitations**

At this time there are no FDA-approved NAATs for use with extrapulmonary specimens.

**Question 19: Should histological examination be performed on specimens collected from sites of suspected extrapulmonary TB?**

**Evidence**

Accuracy studies indicate that histological examination has a sensitivity of 69%–97%, 86%–94%, 60%–70%, 79%–100%, and 73%–100% in pleural tissue (Supplementary Table 10), urologic tissue (Supplementary Table 11), endometrial curettage (Supplementary Table 11), peritoneal biopsy (Supplementary Table 13), and pericardial tissue (Supplementary Table 14), respectively, when final clinical diagnosis, consistent pathology/cytology, or microbiologic confirmation is used as the reference standard. The specificity of mycobacterial culture microscopy tends to be low because necrotizing and nonnecrotizing granulomas are seen in other infectious and noninfectious diseases. This evidence provides very low confidence in the estimated test characteristics because many studies do not report enrolling consecutive patients, the wide ranges of sensitivity are due to the variable results of individual studies, and the studies were small with few samples.

Recommendation 19: We suggest that histological examination be performed, rather than not performed, on specimens collected from sites of suspected extrapulmonary TB (conditional recommendation, very low-quality evidence). Remarks: Both positive and negative results should be interpreted in the context of the clinical scenario because neither false-positive nor false-negative results are rare.
**Rationale**

Tissue sampling with histological examination generally occurs after other types of diagnostic testing have failed to identify a definitive diagnosis. Thus, at this stage in the diagnostic process, the committee thought testing was worthwhile if sensitivity and specificity were both >50%, meaning that true results were more likely than false results. Histological examination surpassed these thresholds and, therefore, is recommended. However, the committee emphasizes the importance of interpreting the results within the clinical context, to lessen the impact of false results. The recommendation is conditional because the very low quality of evidence does not provide sufficient confidence in the estimated sensitivities and specificities for the committee to be certain that the balance of desirable to undesirable consequences favors testing and obtaining the specimens to test usually requires an invasive procedure and, therefore, the balance of benefits versus risks may vary substantially depending upon the clinical condition of the patient.

**GENOTYPING OF M. TUBERCULOSIS**

Over the past 2 decades, genotyping of TB strains has been shown to be a valuable tool in TB control. Molecular epidemiology has helped to identify unsuspected transmission, determine likely locations of transmission, measure the extent of transmission, and differentiate reactivation from newly acquired infection [227]. Often traditional contact investigations focus on persons in the household and workplace. Numerous reports describe TB cases linked through genotyping of Mtb isolates, when detection of transmission was initially missed by conventional contact investigation because the setting was nontraditional. This type of transmission occurs frequently among members of a “social network” that is centered around a specific activity, including illicit drug use, excess alcohol use, or gambling, or location such as a homeless shelter, adult entertainment club, or HIV residential care facility [228–231]. When genotyping detects previously unrecognized transmission of TB in a nonconventional setting, public health interventions to contain and subsequently end the outbreak can be redirected to focus on the social network or location associated with transmission.

Genotyping or DNA fingerprinting of Mtb can be used for determining the clonality of bacterial cultures. PCR-based, and sometimes Southern blotting, methods are used. The PCR-based methods are mycobacterial interspersed repetitive units (MIRU) and spacer oligonucleotide typing (spoligotyping) [232, 233]. A standardized protocol has been developed to permit comparison of genotypes from different laboratories [232].

**Question 20: Should genotyping be performed on a culture isolate from culture-positive patients with TB?**

**Evidence**

We identified no empirical evidence that estimated the frequency with which the availability of genotyped isolates changed public health practices or affected patient outcomes. Therefore, the recommendation is based upon the committee's collective clinical experience, which constitutes very low-quality evidence.

**Recommendation 20: We recommend one culture isolate from each mycobacterial culture–positive patient be submitted to a regional genotyping laboratory for genotyping (strong recommendation, very low-quality evidence).**

**Rationale**

Genotyping is useful in detecting false-positive results due to confirming laboratory cross-contamination [234, 235], investigating outbreaks of TB (both detecting unsuspected outbreaks and confirming suspected outbreaks) [236], evaluating contact investigations [237], and determining whether new episodes of TB are due to re-infection or reactivation [238]. In addition, genotyping is useful for elucidating sites and patterns of Mtb transmission within communities [237, 239]. This information is used by state and local tuberculosis control programs to focus interventions to interrupt further TB transmission. Genotyping is used to aid public health departments in the control of TB and poses no risk to individual patients.

Recently, whole-genome sequencing (WGS) has been applied to investigation of tuberculosis outbreaks [240]. This technique may add discriminatory power to strain identification, but the role of WGS in outbreak investigation is still being determined.

In response to nosocomial outbreaks and tuberculosis among HIV-infected patients, the CDC established a national universal tuberculosis genotyping system for the United States. The merger of modern molecular protocols for strain identification at the DNA level and conventional epidemiological methodologies has given birth to an enhanced collaborative strategy to impact tuberculosis control efforts. Regional TB genotyping laboratories can be contacted through the state public health laboratories or TB control programs.

The recommendation is strong because the committee felt certain that the public health benefits of genotyping far outweigh the modest costs and burdens of genotyping. Even though the evidence can provide very little confidence in the magnitude of the benefits, costs, and burdens used by the committee to make its decision, the differences seemed so overwhelming that the committee thought it extraordinarily unlikely that additional data would lead to a judgment that the costs and harms exceed the benefits.

**RESEARCH NEEDS**

As described by Abu-Raddad et al [241], improved detection of those with TB and improved identification of those at risk...
to progress once infected have the potential to substantially decrease the prevalence of TB and its associated mortality.

**Tuberculosis**

The ability to rapidly and accurately identify *Mtb* as well as drug resistance (eg, through NAAT, line probe, molecular beacon, and Xpert MTB/RIF assays) reflects substantial advances. While rapid tests for TB diagnosis still have a sensitivity of 70%–90%, they may fail to detect paucibacillary pulmonary TB. They also remain relatively expensive, making them difficult to implement in high-burden, low-resource settings. Ideally, what is needed is a simple, inexpensive, rapid (ie, hours) test that is highly accurate (>95% sensitivity and specificity). Rapid tests for detection of drug resistance are approaching the desired level of accuracy, at least for rifampin. However, these tests also are relatively expensive and need to be expanded to allow for detection of resistance to other TB medications. Such expansion is currently limited by gaps in knowledge of the molecular basis of resistance to most first- and second-line drugs. In this regard, improved functional tests for resistance may prove useful.

Other significant gaps remain in the diagnosis of pediatric and extrapulmonary TB. First, the yield of AFB smear and culture in children is low compared to that in adults, which leads to excessive morbidity and mortality due to delayed and missed diagnoses, especially in resource-limited settings. Conversely, the inability to exclude TB results in overtreatment when the diagnosis cannot be excluded. In areas of the world where TB is diagnosed entirely based on smear microscopy, children will be almost completely neglected and untreated for TB. In areas with greater resources, low yields of microbiologic specimens in children deter many clinicians from even attempting culture collection. This may result in prolonged treatment with extra TB drugs (in jurisdictions that use 4 drugs for 6 months in patients lacking susceptibility data). Alternatively, drug resistance will not be identified and the child could suffer dire consequences receiving inadequate care. Second, similar challenges exist for the accurate diagnosis of those with extrapulmonary TB. Finally, diagnostic approaches to the identification of those likely to fail TB treatment are needed. These limitations in the diagnosis of paucibacillary TB highlight the need to develop testing strategies based on either host or bacterial markers of infection that can be measured from readily available clinical sources such as plasma or urine.

**Latent Tuberculosis Infection**

Individuals with immunological evidence of exposure to *Mtb* antigens, but without evidence of clinical disease are termed “latently” infected. However, it is clear that there is considerable heterogeneity within this classification. As was described previously, those with recent infection (<2 years) are at increased risk for progression to clinical disease, and functionally might be considered acutely infected. Conversely, those more remotely infected are at substantially lower risk, but remote infection remains difficult to define operationally. Those with negative TSTs or IGRAst are unlikely to progress to TB. However, the PPV of either test is relatively modest, as current estimates would suggest that among household contacts, 20–40 people require treatment to prevent one case. While the number needed to treat based on IGRAst is not known with certainty, early evidence suggests that it is not likely to be dramatically different [24, 55]. Nonetheless, a careful evaluation of which diagnostic is more closely associated with the development of TB remains a research priority. Given the relatively poor PPVs of current diagnostics for the prediction of progression to TB disease, a diagnostic that can accurately identify those at risk is needed. It should be noted that both QFT and T-SPOT are largely measures of CD4 T-cell immunity, but additional markers of inflammation, cellular, or humoral immunity may prove useful. Clearly, the identification of biomarkers associated with the development of tuberculosis following infection will require carefully performed prospective investigation. In this regard, the prospective evaluation of populations at risk for disease progression, and the use of sophisticated imaging such as positron emission tomography–computed tomography (PET-CT) are likely to further delineate markers associated with either disease progression or subclinical infection.

Operationally, the intent of targeted testing is to identify those who would benefit from treatment. While much is now known about the accuracy of both the TST and IGRAst, much less is known about their performance with regard to treatment completion. Additional research on the use of IGRAst with regard to provider and patient perceptions is needed to establish optimal diagnostic and treatment strategies. Finally, the literature addressing the performance of IGRAst in children <5 years old is still limited and studies to inform the appropriate use of these tests to accurately diagnose LTBI in this age group are needed. Studies of young household contacts in low-incidence countries would be especially informative.

**GUIDELINE STATEMENT**

These guidelines are not intended to impose a standard of care. They provide the basis for rational decisions in the diagnostic evaluation of patients with possible latent tuberculosis or tuberculosis. Clinicians, patients, third-party payers, stakeholders, or the courts should never view the recommendations contained in these guidelines as dictates. Guidelines cannot take into account all of the often compelling unique individual clinical circumstances. Therefore, no one charged with evaluating clinicians’ actions should attempt to apply the recommendations from these guidelines by rote or in a blanket fashion. Qualifying remarks accompanying each recommendation are its integral parts and serve to facilitate more accurate interpretation. They should never be omitted when quoting or translating recommendations from these guidelines.

**Supplementary Data**

Supplementary materials are available at Clinical Infectious Diseases online. Consisting of data provided by the author to benefit the reader, the posted materials are not copyedited and are the sole responsibility of the author, so questions or comments should be addressed to the author.
Notes

Acknowledgments. The writing committee thanks Drs Mike Isman and Jeffrey Starke for their critical examination of the manuscript. The committee is particularly indebted to Kevin Wilson for his patience and his editing skills.

Potential conflicts of interest. D. L. C. has received speaking fees from Qiagen. C. L. D. serves on the data and safety monitoring board (DSMB) for Otsuka America Pharmaceutical, Inc, served on a DSMB for Sanofi Pasteur Inc, received research support from Insmed, and received travel support from Qiagen. L. R. serves as a speaker and on an advisory committee for Boehringer Ingelheim and F. Hoffmann-La Roche, serves on an advisory committee and received research support from Biogen Inc, served on an advisory committee for AstraZeneca, GlaxoSmithKline, and Sanofi Pasteur, served as a consultant for Bayer HealthCare, served as a speaker for Cipla. T. M. S. is employed by the US Centers for Disease Control and Prevention (CDC). T. R. S. serves on a DSMB for Otsuka. P. A. L. is employed by the CDC. All other authors report no potential conflicts. All authors have submitted the ICMJE Form for Disclosure of Potential Conflicts of Interest. Conflicts that the editors consider relevant to the content of the manuscript have been disclosed.

References


Diagnosis of TB in Adults and Children • CID 2017;64 (15 January) • e29
recommendations to prevent reactivation of latent tuberculosis infection in Carmona L, Gomez-Reino JJ, Rodriguez-Valverde V, et al. Effectiveness of trials. BMJ


Committee on Prophylaxis. Bull World Health Organ

Efficacy of various durations of isoniazid preventive therapy for tuberculosis: five 9:15. tuberculosis infection among HIV-infected individuals. BMC Infect Dis


169. Schoch OD, Rieder P, Tueller C, et al. Diagnostic yield of sputum, induced spu-
179. Schoch OD, Rieder P, Tueller C, et al. Diagnostic yield of sputum, induced spu-


