Clinical Application and Limitations of Interferon-γ Release Assays for the Diagnosis of Latent Tuberculosis Infection

Victor Herrera,1 Sharon Perry,1 Julie Parsonnet,1 and Niaz Banaei1,2,3

1Departments of Medicine, Division of Infectious Diseases and Geographic Medicine, 2Pathology, Stanford University School of Medicine, and 3Clinical Microbiology Laboratory, Stanford University Hospital, Palo Alto, California

Interferon-release assays (IGRAs) represent advances in tuberculosis immunology and evolutionary biology. IGRAs were designed to replace tuberculin skin test (TST) for the diagnosis of latent tuberculosis infection because of their logistical advantages and enhanced specificity over TST. Although IGRAs and TST have been useful in epidemiologic studies, they lack the sensitivity and reproducibility normally expected from diagnostic tests in clinical practice. In this review, we present an overview of the current recommendations and knowledge in the field and discuss practical approaches in areas of uncertainty related to discordant IGRA results.

GLOBAL IMPACT OF TB INFECTION

One-third of the worldwide population is infected with Mycobacterium tuberculosis (MTB) [1]. After exposure to MTB, ~30% of individuals are thought to develop latent infection [2]. Because each case of infection carries a 5%–10% lifetime risk of progressing to the active tuberculosis (TB), latent infections comprise a significant reservoir of future epidemics. Risk factors associated with progression include malnutrition, acute infections, and other conditions that compromise immunity [3].

BENEFITS OF SCREENING FOR LATENT TB INFECTION (LTBI)

Treatment of latent infection with isoniazid reduces risk of future disease by 75%–90% [1]. Because each case of active TB may cause 10–20 additional infections [4], identifying individuals with LTBI has substantial societal and individual benefits. In the United States, it is estimated that ~11 million individuals have LTBI [5]; however, prevalence of LTBI varies widely from <1% among younger, native-born individuals to >50% among recent immigrants from countries of endemicity [6, 7].

SCREENING TESTS FOR LTBI

There is no gold standard for diagnosis of LTBI. Indications for prophylactic therapy are based on epidemiologic, clinical, and laboratory criteria and patient acceptance [8].

The Tuberculin Skin Test
First described by Koch in 1890 and developed by Mantoux in 1907, the tuberculin skin test (TST) relies on principles of delayed hypersensitivity to recruit memory T cells to the site of an intradermal injection of purified protein derivative (PPD) of Mycobacterium bovis. In patients with culture-confirmed TB, the sensitivity of the TST ranges from 75%–90%; however, cross-reactivity of PPD with nontuberculous mycobacteria (NTM) and the bacille-Calmette Guérin (BCG) vaccine limits the specificity of the test in populations in which exposure to MTB is common. Other limitations of the TST include the boosting phenomenon, the necessity of scheduling a reading visit, and inter-observer
variability. In addition, false-negative results may be common in immunocompromised patients, young persons, and older persons [8].

Interferon-γ Release Assays
A hallmark of the immune response to MTB infection is the release of interferon (IFN)–γ by CD4 cells [9]. In the past decade, 2 standardized IFN-γ release assays (IGRAs) have been developed to measure IFN-γ responses to infection in blood samples (Table 1). Both are approved by the US Food and Drug Administration and recommended by the US Centers for Disease Control and Prevention (CDC) [10] as an aid in the detection of LTBI. The QuantiFERON-TB GOLD test (Cellestis) uses an enzyme-linked immunosorbent assay to measure the amount of IFN-γ released in response to in vitro stimulation of whole blood with MTB antigens. A more recent version of the QuantiFERON test, known as the QFT-gold in-tube (GIT) [11], enables blood to be collected directly into precoated tubes to expedite incubation and to minimize sample processing. T-SPOT TB assay (Oxford Immunotec) uses an enzyme-linked immunospot assay to count the number of IFN-γ–producing cells on precoated plates [12]. An advantage of this assay is the standardization of the number cells added to each well. Both assays use highly specific MTB antigens, such as early-secreted antigen 6 (ESAT6) and culture filtrate protein 10 (CFP10), which are not present in any strains of BCG or most environmental mycobacteria, thus providing improved specificity over TST [13]. Both assays are also available with a positive control (phytohaemagglutinin). Important logistical advantages of the IGRAs include a single patient visit and objective output [14].

In the past decade, hundreds of papers have appeared in the evidence-based literature regarding the performance characteristics of the new IGRAs in different populations and settings. The present review focuses on the use of the IGRAs as an aid for detection of LTBI. Although they offer important logistical advantages, several challenges confront the physician using these tests. These issues will be discussed and include (1) low or variable sensitivity, including discordance with TST; (2) poor reproducibility; (3) limited interpretive criteria; (4) lower efficacy in children; and (5) unknown prognostic value.

### INDICATIONS FOR TESTING

In the United States and other low-incidence countries, a moderate proportion of active TB cases result from reactivation of a recent infection [15]. With the recognition in the 1950s that isoniazid could prevent reactivation of LTBI, the public health service began to recommend TB screening broadly in its population [16]. Deaths from isoniazid-induced hepatitis, however, combined with the recognition that screening the whole population was impractical, soon led to more targeted recommendations. Because screening policies vary in much of the world, this review will focus on testing indications in the United States.

### Screening Indications and Risk Stratification

The American Thoracic Society, in conjunction with the CDC, has provided guidelines for tuberculosis screening in the United States [10]. Since 2005, the recommendations embrace either TST or IGRA testing for diagnosis of LTBI. With either test, the key recommendation is that testing be targeted to those at highest risk and, conversely, avoid those at low risk. This latter recommendation prevents excessive treatment of false-positive results, which regardless of the test, are always more likely in low-prevalence populations [8].

Currently, individuals considered to be at high risk can be categorized into 2 groups: those likely to have acquired infection recently and those with conditions that increase the risk for LTBI reactivation. In the former category are contacts of patients with active TB; individuals known to have converted their LTBI test within the preceding 2 years; children aged <5 years; recent immigrants (within 2 years) from and frequent travelers to countries with high incidence of TB; individuals living or working in homeless shelters, prisons, nursing homes, or residential facilities for patients with AIDS;

---

Table 1. Characteristics of FDA-Approved IGRAs

<table>
<thead>
<tr>
<th>Variable</th>
<th>QFT-GIT</th>
<th>T-SPOT</th>
</tr>
</thead>
<tbody>
<tr>
<td>Format</td>
<td>ELISA</td>
<td>ELISPOT</td>
</tr>
<tr>
<td>Specimen</td>
<td>Whole Blood (1 mL per tube, 3 tubes)</td>
<td>PBMCs (2.5×10⁵ per well, 4 wells)</td>
</tr>
<tr>
<td>TB antigen peptides</td>
<td>ESAT6, CFP10, TB7.7</td>
<td>ESAT6, CFP10</td>
</tr>
<tr>
<td>Positive control</td>
<td>Phytohemagglutinin</td>
<td>Phytohemagglutinin</td>
</tr>
<tr>
<td>Output</td>
<td>IU/ml</td>
<td>Spot forming units (SFU)</td>
</tr>
<tr>
<td>Positive result</td>
<td>&gt;=0.35 IU/ml, and &gt;25% of nil</td>
<td>&gt;=8 SFU</td>
</tr>
<tr>
<td>Borderline result</td>
<td>None</td>
<td>5.6 SFU</td>
</tr>
<tr>
<td>Indeterminate result</td>
<td>Mitogen &lt;0.5 IU/ml and/or Nil &gt;8.0</td>
<td>&gt;10 Spots in Nil, or &lt;20 SFU mitogen</td>
</tr>
</tbody>
</table>
Figure 1. An algorithm for LTBI screening with IGRAs. (1) Risk stratification is based on criteria established by the CDC. (2) After an exposure, consider waiting at least 4–7 weeks before testing. (3) In certain groups (e.g., close contacts of patients with active TB), consider excluding active TB and treating for LTBI despite negative IGRA results. (4) IGRAs have lower sensitivity in children and immunocompromised patients.

and health care workers likely to care for patients with TB. In the second category are individuals with immunocompromising conditions (e.g., HIV infection; treatment with immunosuppressant therapy, including TNF-α antagonists and high-dose steroids; hematopoietic malignancy; diabetes; and chronic renal disease), those with fibrotic lung disease or gastrectomy, and individuals who use drugs or alcohol or are grossly (>10%) underweight. It should be noted that the decision to be screened should be based on the individual’s risk, not on employment considerations. Thus, although some states mandate screening of teachers, school volunteers, and others, this practice is not cost-effective or beneficial for individuals at low risk.

Several caveats must be considered when using any screening test. For example, recent contacts of patients with active cases may have negative test results if tested too soon after exposure [17]. Thus, it is recommended that, if the result of an initial screening test is negative, the test should be repeated 8–10 weeks after the last known exposure to the patient [18]. In addition, the clinician must always apply clinical judgment to the specific case scenario. For example, a close contact of a patient with active TB or an immunocompromised patient at high risk for TB should be strongly considered for prophylactic treatment regardless of the result of the test. An algorithm on screening for LTBI is presented in Figure 1.

**PERFORMANCE CHARACTERISTICS OF IGRAs AND TST**

**Sensitivity and Specificity**

Evaluating the performance of IGRAs and TST has been complicated by the absence of a reference standard for LTBI [19]. Most studies use surrogate markers of LTBI to estimate performance measures for IGRAs. Specificity has been assessed by testing individuals at very low risk for LTBI, and sensitivity has been estimated using culture-confirmed cases. The latter approach is problematic, because active disease and latent infection are 2 distinct conditions with substantial clinical and immunological differences. More accurate estimates of sensitivity based on progression of LTBI to active TB are starting to emerge and are discussed below [20, 21].

There is significant variation in performance estimates for T-SPOT, QFT-GIT, and TST depending on the study type and the population tested [22]. Comparing studies is challenging.
because of the use of different definitions and cutoff values and the limited number of direct comparisons [23, 24]. In a recent comprehensive review of studies comparing the sensitivity of IGRAs and TST in patients with culture-confirmed TB, the pooled sensitivity for T-SPOT, QFT-GIT, and TST was calculated to be 90%, 83%, and 89%, respectively [10]. In the same study, pooled specificity was 88%, 99%, and 85%, respectively. T-SPOT has a slightly superior sensitivity to QFT-GIT in most studies, whereas QFT-GIT tends to have a higher specificity than T-SPOT [24]. The IGRAs have superiority in specificity over TST in BCG-vaccinated populations because of use of MTB-specific antigens [25].

Longitudinal studies involving contacts of patients with active TB are starting to provide important information on the sensitivity of IGRAs in infected individuals who progressed to active TB. The sensitivities have been variable, ranging from 40% to 100%, even in contacts with strains linked to the source case by strain typing (Table 2). A caveat in these studies is the lack of standardization of time lapse between TB exposure and LTBI testing. Nonetheless, these findings indicate that, in typical contact investigations, IGRAs lack sufficient sensitivity for detection of LTBI after a recent exposure. These findings also question the efficacy of confirming positive TST results with IGRAs. For the clinician and public health officials, these results imply that a negative IGRA result does not rule out the diagnosis of LTBI.

**Prognostic Value for Progression to Active TB**

The risk of developing active TB after a positive TST result has been defined in large longitudinal studies [29]. It is known that only a small proportion of infected individuals develop active TB; therefore, only this subset would theoretically benefit from receiving prophylaxis [29]. IGRAs could provide prognostic insight in identifying cases that are most likely to progress to active disease [30]. Thus far, the quantitative results from IGRAs have not been shown to have prognostic value and, therefore, should not be used for that purpose in clinical practice. It is possible that a unique signature of differential cell surface markers and secreted cytokines in functional T cell assays could identify a protective immunophenotype and be of prognostic value for clinical decision making [31]. Further development of current functional assays and longitudinal studies are necessary to validate this hypothesis.

**Reproducibility of Results After Serial Testing**

Health care workers represent one of the largest TB surveillance groups [32]. Because health care workers are tested annually, reproducibility of IGRAs is crucial. Studies examining the reproducibility of IGRAs after serial testing have found variability in IFN-\(\gamma\) responses. Both conversions in the apparent absence of TB exposure and reversions in the absence of therapy have been observed, with frequencies of 12%–50% [33–35]. Conversions and reversions are thought to be secondary to within-subject fluctuations and/or attributable to variations in laboratory procedures, but biological and environmental causes remain possible [36, 37]. To date, there is no consensus on how to interpret conversions and reversions in terms of their accuracy. The main challenge is to differentiate nonspecific variations from true conversions. For this reason, a grey zone has been proposed for individuals with fluctuating QFT-GIT results close to the cutoff value of .35 IU/mL [38]. Results outside this zone are presumed to be true reversions or conversions. A borderline category already exists for T-SPOT.

**Efficacy in Children**

Evaluation of IGRAs in children remains limited because of small study sizes, resistance to phlebotomy, and difficulty in obtaining culture-confirmed results for reference [10]. In a study of 204 children, 81% of 99 TST-positive (induration, ≥10 mm) children considered to be at low to moderate risk of...
TB were QFT-GIT negative, as were 100% of 5 close contacts of a patient with TB [39]. Outside the United States, sensitivities of QFT and T-SPOT in children are generally reported to be similar to or lower than that of TST [40–42]. Reports of lower mitogen levels in children and higher rates of indeterminate results are consistent with the concern that age-related immunologic factors may affect the sensitivity of these tests [39, 43, 44]. Because young children are at greater risk of disseminated disease [45], a negative IGRA result should not be used to exclude infection.

**Approach to Interpreting IGRA and TST Results**

The decision to evaluate patients for LTBI should be guided by risk assessment (Figure 1). This is an important step, because a positive screening result in an individual at low risk leads to diagnostic and therapeutic dilemmas. In the ensuing section, we present common scenarios that arise in clinical practice and discuss ways to address them.

**Conversions and Reversions**

In the event that a test is performed for an individual at low risk of LTBI, conversion from a negative to a positive result may represent a false conversion. A reasonable approach is to withhold treatment and repeat testing. We and others have shown that, in individuals at high risk, results close to the cutoff for QFT-GIT are more likely to revert and convert during repeat testing [46]. Therefore, a high positive result (ie, >1.0 IU/mL) is more likely to remain positive and, therefore, should be confirmed and treated as if it is still positive. It is not clear why false-positive results occur, but postulated reasons include concomitant illness at the time of testing, laboratory factors, and nonspecific boosting of IFN-γ responses [34, 47].

Reversions from positive to negative come to attention when the positive result is suspected to be a false positive and the test is repeated. Reversions can also occur spontaneously or after therapy and are postulated to represent immune clearing of the infection [30]. Studies evaluating the effect of LTBI treatment on positive IGRA results have not confirmed that therapy increases the rate of reversions, although a decrease in quantitative results has been reported [48, 49].

**Discordant Results Between IGRA and TST**

Longitudinal studies suggest that the TST is more sensitive than IGRA in high-risk populations [50]. However, IGRA may be more sensitive at detecting recent TB exposure than is the TST [51]. In the case of an individual with a history of BCG vaccination and a positive TST and negative IGRA results, if risk for TB infection is otherwise low, it is reasonable to assume that the TST result is false positive and to withhold further evaluations. In the case of a negative TST and positive IGRA results, if the individual is considered to be at high risk for TB infection, the negative TST result alone should not prevent further examination.

Another consideration when interpreting discordant TST and IGRA results is the observation that TST preceding IGRA testing could boost IGRA [47]. This effect appears more pronounced on the days after the TST and could wane with time.

**Interpretation of Quantitative IGRA Results**

According to recent recommendations by the CDC, “both the qualitative results and the quantitative assay measurements for IGRA results should be reported” [10, p. 10]. Reporting IFN-γ measurements does provide useful information in individuals undergoing serial IGRA testing [51]. However, in the absence of interpretive guidelines, this practice could lead to false assumptions and misinterpretation of the results. There are limited data on the significance of changing IFN-γ levels [52]. It is not clear whether higher IFN-γ responses correlate with greater risk of progression to active TB. Pre-analytical factors, such as delays in incubation and sample processing, are also known to negatively affect IFN-γ responses [37, 53]. These factors should be taken into consideration when interpreting quantitative results.

On the basis of the available data, the following approach to interpretation of quantitative results is recommended: (1) the quantitative results should not be used for prognostic or therapeutic monitoring purposes at this time, because evidence is lacking or nonsupportive and (2) the quantitative results are useful for predicting the likelihood for reversion or conversion of test results when the IFN-γ signal is close to the assay cutoff.

**CONCLUDING REMARKS AND FUTURE NEEDS**

IGRAs have the potential to improve health care efficiency and streamline the diagnosis and management of LTBI. However, rigorous investigational studies conducted over the past decade, including longitudinal studies, have shown that IGRAs lack the sensitivity that is routinely demanded from other laboratory diagnostics [20, 21, 26, 27, 54]. Below, we present crucial deficiencies related to IGRA performance and discuss ways to overcome these hurdles.

**Assay Standardization**

It has been shown that pre-analytical factors negatively impact IGRAs. Thus, it is critical that preanalytics of IGRA are standardized and appropriate quality controls included. It will also be important to determine whether biological and/or environmental factors have an impact on IGRA performance. Further investigation is needed to determine whether there is a circadian or seasonal effect on T cell circulation and function, which may impact IGRA results.
Sensitivity
In their current format, IGRA s have limited sensitivity for detection of either latent or active TB. It is plausible that multiplexed assays measuring additional T cell cytokines and biomarkers downstream of IFN-γ signaling could enhance the sensitivity of IGRA s. Another approach, although of limited practicality, includes performing IGRA s on lung-derived mononuclear cells. Whether these modifications can improve the sensitivity of IGRA s remains to be demonstrated in longitudinal studies using bacteriological culture as the reference standard.

Specificity
IGRA s cannot distinguish between either active or latent infection, de novo or treated infection, or recent or remote infection. The latter is crucial for risk stratification, because infections are more likely to progress to active TB within the first few years after exposure. Several studies have shown that the profile of T cell surface markers and secreted cytokines in IGRA s is heterogeneous during the various stages of infection and that it may be possible to use these functional signatures to improve the specificity of IGRA s. These findings require validation in large prospective cohorts.

Reproducibility
Conversions and reversions close to the assay cutoff value pose a major challenge to making an accurate diagnosis of LTBI. Reproducibility studies with large cohorts are needed to better delineate the grey zone, and longitudinal studies are necessary to determine their significance. Guidelines are also needed on how to manage individuals with results in the grey zone.

In summary, IGRA s and the TST have similar performance for the detection of LTBI. In certain contexts, IGRA s offer economic and logistical advantages over TST. Although both tests are the detection of LTBI. In certain contexts, IGRA s offer economic


