Comparison between a Whole Blood Interferon-γ Release Assay and Tuberculin Skin Testing for the Detection of Tuberculosis Infection among Patients at Risk for Tuberculosis Exposure

Moriyo Kimura, Paul J. Converse,[1] Jacqueline Astemborski, James S. Rothel, David Vlahov, George W. Comstock, Neil M. H. Graham,[1] Richard E. Chaisson, and William R. Bishai

Center for Tuberculosis Research, and Departments of International Health, Epidemiology, and Molecular Microbiology and Immunology, Johns Hopkins University School of Hygiene and Public Health, and the Department of Medicine, Johns Hopkins University School of Medicine, Baltimore, Maryland; CSL, Ltd., Parkville, Australia

A new test that measures interferon-γ (IFN-γ) release in whole blood following stimulation with tuberculin has the potential to detect tuberculosis infection using a single blood draw. The IFN-γ release assay was compared with the standard tuberculin skin test (TST) among 467 intravenous drug users at risk for tuberculosis in urban Baltimore. Among 300 human immunodeficiency virus (HIV)–seronegative patients, the IFN-γ release assay was positive in 177 (59%), whereas the TST was positive in 71 (24%), for a percent agreement of 59% (κ = 26%). Among 167 HIV-seropositive subjects, the IFN-γ release assay identified 32 reactors (19%); the TST identified 16 reactors (9.6%), for a percent agreement of 82% (κ = 28%). The IFN-γ release assay detected more reactors than did the TST, but its agreement with TST was weak. As the TST is an imperfect standard, further evaluation of the IFN-γ release assay among uninfected persons and persons with culture-confirmed tuberculosis will be useful.

Poor detection of tuberculosis infection and inefficient prevention of reactivation are major impediments to improved tuberculosis control. The tuberculin skin test (TST) has been widely used as a screening test to identify individuals with latent tuberculosis. Despite standardization, the TST has significant limitations [1]. Nonspecific reactivity stemming from exposure to environmental mycobacterial species or bacille Calmette-Guérin vaccination are well-known sources of false-positive TST results. Improper placement or measurement of cutaneous reactions can cause false-negative results. Anergy associated with disseminated tuberculosis, critical illnesses, some acute infectious diseases, and host immunosuppression from human immunodeficiency virus (HIV) infection or cancer chemotherapy can also cause false-negative reactions [2–4]. Finally, an additional drawback to the TST is the need for patients to return for a reading on the second or third day after testing.

Received 26 June 1998; revised 4 December 1998.

Presented in part: American Lung Association/American Thoracic Society International Conference, Chicago, 24–29 April 1998 [poster].

Institutional review board approval was obtained prior to initiating this study, and informed consent was obtained from all patients participating.

J.S.R. is an employee and stockholder of CSL, Ltd.

Financial support: CDC (cooperative agreement U52-CCU300466-17); NIH (AI-36973, AI-37856, and HL-21670); CSL, Ltd.

[1] Present addresses: Glaxo Wellcome Research and Development, Five Moore Dr., P.O. Box 13398, Research Triangle Park, NC, 27709 (N.M.H.G.); IGEN International, Inc., 16020 Industrial Dr., Gaithersburg, MD 20877 (P.J.C.).

Reprints or correspondence: Dr. William Bishai, Johns Hopkins School of Public Health, 615 N. Wolfe St., Baltimore, MD 21205 (wbishai@jhsph.edu).

The Journal of Infectious Diseases 1999; 179:1297–1300

0022-1899/99/7905-0038$02.00

Improved screening tests with greater sensitivity and specificity for tuberculosis infection that would not require multiple patient encounters would significantly enhance tuberculosis control efforts. A promising candidate is the interferon-γ (IFN-γ) release assay developed for the diagnosis of tuberculosis infection in cattle [5]. The test measures IFN-γ release from tuberculin- and control-stimulated whole blood cells. Studies in cattle indicate that the assay is sensitive and specific for the detection of *Mycobacterium bovis* infection [6, 7].

Recently, preliminary evaluations of the IFN-γ release assay have been reported in humans. One study found modest agreement between the IFN-γ release assay and the TST among 67 intravenous drug users and suggested that the new test may be more sensitive than the TST [8]. A second study conducted among 952 Australian pulmonary clinic patients and military recruits, all of whom were HIV-seronegative, demonstrated strong concordance between the 2 tests for this relatively healthy population at low risk for tuberculosis [9]. To date, however, the IFN-γ release assay has not been evaluated in a large population of US subjects at high risk of both tuberculosis and HIV infection. In this report, we describe the agreement between the IFN-γ release assay and the TST among 467 urban intravenous drug users, 36% of whom were HIV-seropositive.

Materials and Methods

Volunteers in the ALIVE (AIDS Linked to Intravenous Experiences) longitudinal study of the natural history of intravenous drug use participated in the study [10]. TST reactions with 5 TU of Tubersol (Pasteur Meriéux-Connaught Laboratories, Toronto, Canada) were measured at 48–72 h by trained nurses, with posi-

tivity defined as a reaction ⩾10 mm for HIV-seronegative and ⩾5 mm for HIV-seropositive subjects [1].

The commercially available IFN-γ release assay (*Quanti*-FERON-TB; CSL) with four test stimuli was used as previously described [8, 9, 11]. The test uses stimulation with *Mycobacterium tuberculosis* (human type) purified protein derivative (PPD), *Mycobacterium avium* (avian type) PPD, *M. bovis* (bovine type) PPD, phytohemagglutinin as a positive control, and PBS as a negative control. After subtracting the background release due to PBS stimulation, the percent response levels were those of IFN-γ release induced by human-type PPD divided by that from phytohemagglutinin induction (defined as maximal). Unless otherwise indicated, a positive test result by IFN-γ release assay was defined as a percent response level of ⩾15% [8].

Results

Of 1008 ALIVE study patients, 467 (46.3%) completed the study ("full participants") by returning in 2 or 3 days to have their TST reactions measured. There were only slight differences in the demographic characteristics (age, sex, race, smoking, alcohol use, and HIV infection status) of full participants and those who withdrew. None of the variables differed significantly between the 2 groups; 300 (64.2%) of full participants were HIV-seronegative, and 167 (35.8%) were HIV-seropositive. Among HIV-seropositive participants, 19 (11.4%) had AIDS-defining illnesses; the mean CD4 cell count was 318/μL (median, 287/μL; interquartile range, 291). Fifty-seven patients had CD4 cell counts <199, 80 had 200–499/μL, and 28 had >500/μL. For HIV-seronegative participants, the mean CD4 cell count was 957/μL. One patient had a history of tuberculosis prior to entry; no subjects developed tuberculosis disease in the 1 year following entry.

Table 1 summarizes the results of the TST and the IFN-γ release assays for the 300 HIV-seronegative and 167 HIV-seropositive subjects when an IFN-γ release assay value of ⩾15% of maximal was defined as positive (as recommended by the manufacturer). Among HIV-seronegative persons, there were more than twice as many reactors by the IFN-γ release assay than by the TST (59% vs. 24%, respectively). The 2 tests showed 59% agreement, and the κ statistic was 26%. κ values <40% indicate weak correlation, while those >75% suggest excellent agreement. Thus, this κ value suggested that the 2 tests correlated poorly in spite of the 59% agreement. Among the HIV-seropositive subjects, the IFN-γ release assay identified twice as many positive reactors as the TST. While the observed agreement was 82% in this group, the κ statistic remained low at 28%, again indicating weak correlation beyond that expected to occur by chance.

Table 1. Tuberculin skin test (TST) and IFN-γ release assay results in HIV-seronegative and -seropositive subjects.

HIV status, TST reactions (in mm)	IFN-γ release (% of maximal response)		
	0–14%	15+%	Total
Negative			
0–9	115 (38.3)	114 (38.0)	229 (76.3)
10+	8 (2.7)	63 (21.0)	71 (23.7)
Total	123 (41.0)	177 (59.0)	300 (100)
Positive			
0–4	128 (76.6)	23 (13.8)	151 (90.4)
5+	7 (4.2)	9 (5.4)	16 (9.6)
Total	135 (80.8)	32 (19.2)	167 (100)

NOTE. Data are no. of subjects (%).

To allow for the fact that the IFN-γ release assay and the TST use scales with markedly different ranges, we divided the range of results for each test into sixths and compared the distributions. Among HIV-seronegative persons, there were large modal values at 0- to 4-mm for the TST, and 0%–14% for the IFN-γ release assay (figure 1A). There were relatively low frequencies beyond the first category, with the TST showing a slight second mode in the 10- to 14-mm range. The major difference between the 2 distributions was that the IFN-γ release assay had a smaller proportion of values in the first category (0%–14% of maximal release) than the TST distribution had in its first group (0- to 4-mm). Using the simplified 6-category system of reaction outcomes, the percent agreement between IFN-γ release assay and TST was 44.3%, and the κ statistic was only 17.0%.

When we applied the 6-category system to HIV-seropositive persons, we again found high frequencies in the first category for both the TST (0- to 4-mm) and the IFN-γ release assay (0% to 14%; figure 1B). The overall tendencies of the 2 distributions were more similar to one another than those for HIV-seronegative persons. Whereas the mean percent agreement between the 2 tests using these 6 categories was high for HIV-seropositive subjects (79.0%), the weighted κ test value was only 28.9%, indicating that the bulk of the observed agreement resulted from the large proportion of participants with near 0 values in both tests.

We considered the possibility that adopting a different cutoff point for positivity in the IFN-γ release assay might produce higher levels of agreement between the 2 tests. To evaluate this hypothesis, we arbitrarily increased the stringency of the IFN-γ release assay by testing percent responses of 15%, 35%, and 70% as cutoff points for positivity. For HIV-seropositive participants, the percent agreement/κ values were as follows: 82%/28%, 85%/24%, and 83%/27% using these 3 different cutoff values, respectively. Moreover, stratifying the subjects by CD4 cell count did not improve the percent agreement or κ values. Therefore, for HIV-seropositive persons, changing the IFN-γ release assay cutoff points made little difference in percent agreement or κ value. On the other hand, for HIV-seronegative persons, percent agreement/κ scores increased with higher IFN-γ release assay cutoff points, with values of 59%/26%, 72%/37%, and 79%/41% for cutoff levels of 15%, 35%, and 70%, respectively. An analysis of all cutoff points for the IFN-γ release assay from 0% to 200% showed that for HIV-seronegative subjects, a maximal κ value of 41% was achieved when 70% of maximal release was used as the cutoff point for the IFN-γ

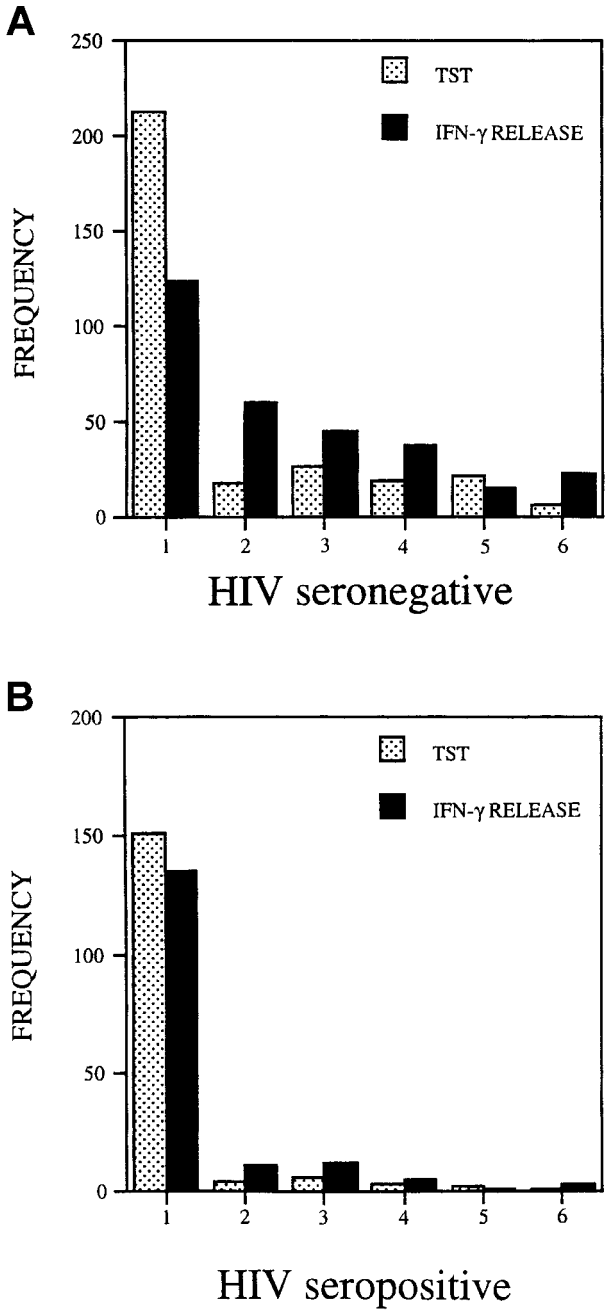

Figure 1. Percent distributions for tuberculin skin test (TST) and IFN-γ release assay test result for 300 human immunodeficiency virus (HIV)-seronegative *(A)* and 167 HIV-seropositive study participants *(B)*. TST groups were defined as follows: group 1 = 0- to 4-mm reaction size, group 2 = 5- to 9-mm, group 3 = 10- to 14-mm, group 4 = 15- to 19-mm, group 5 = 20-to 24-mm, and group 6 = 25- to 29-mm. IFN-γ release assay groups were defined as follows: group 1 = 0%–14% of maximal release, group 2 = 15%–34%, group 3 = 35%–69%, group 4 = 71%–139%, group 5 = 140%–209%, and group 6 = ≥210%.

release assay, given that the cutoff point for the TST was held at 10 mm (data not shown).

Discussion

The whole blood IFN-γ release assay after stimulation with PPD is an in vitro test that evaluates a segment of the complex cell-mediated immunologic events required to produce cutaneous induration in tuberculin-positive individuals. Its simplicity, reproducibility, and rapidity suggest that it might offer advantages over the TST [11]. Significantly, the IFN-γ release assay permits health care providers temporal flexibility in communicating and acting on the test result, unlike the situation with the TST, in which the patient must be seen 48–72 h after the injection.

We compared the commercially produced whole blood IFN-γ release assay for tuberculosis with the TST in 467 injection drug users from urban Baltimore. While other studies of the IFN-γ release assay have been reported, this is the largest evaluation of the test in a US urban population with a high prevalence of HIV infection and a significant risk for tuberculosis exposure.

Among HIV-seronegative patients, the IFN-γ release assay identified 88% (63/71) of tuberculin-positive patients at the 15% cutoff level. In this group, very few TST reactors were missed by the new assay, a result in close agreement with an earlier smaller study of intravenous drug users [8]. Likewise, among HIV- seropositive subjects, the IFN-γ release assay identified 56% (9/16) of those who were tuberculin-positive, similar to the results of the earlier study. The main areas of discrepancy between the IFN-γ release assay and the TST stemmed from a large proportion of tuberculin-negative patients found to be positive by the new test at the 15% cutoff level. Because the TST may be an imperfect measure of latent tuberculosis infection and is prone to significant fluctuations over time, particularly in study populations such as ours [12], it is difficult to ascertain which test is correct among individuals with discordant results.

Overall, our study showed modest agreement between the 2 tests: 59% agreement for HIV-seronegative patients ($\kappa = 26\%$) and 82% for HIV-seropositive patients ($\kappa = 28\%$). Of note, among HIV-seronegative patients, we observed that using a higher cutoff point for the IFN-γ release assay than that applied in earlier evaluations (70% of maximal release rather than 15%) improved the agreement between the 2 tests. In view of the well-documented deficiencies of the TST as a diagnostic tool for tuberculosis infection, it is important to recognize that a new test with perfect sensitivity and specificity would undoubtedly show poor agreement with the TST. Therefore, our results do not exclude the possibility that the IFN-γ release assay could be superior to the TST. Future sensitivity evaluations of the

IFN-γ release assay should be directed towards reference standard populations, which have essentially 100% or 0% rates of tuberculosis infection.

A blood test for which there is an automated assay has 2 important advantages over a skin test. First, there is no necessity for the tested person to return within a narrow window of time. Equally important, automated tests do not have the problem of terminal digit preference—the tendency of human observers to record results in multiples of 5 except when they are using measuring gauges on which the results cannot be seen until the measurement is completed. Terminal digit preference at multiples of 5 leads to serious misclassification [13], especially when the cutoff points are also in multiples of 5, as they are in the current recommendations [1]. These 2 advantages alone would make a test for tuberculosis infection that is based on assays of blood a decided improvement over a skin test.

Acknowledgments

We thank Laura Sheely, Margaret Bonds, Azalia Madison, Sheila Cross, and Caryn Good for technical assistance and Jennifer Doetsch for secretarial assistance.

References

1. American Thoracic Society. Diagnostic standards and classification of tuberculosis. Am Rev Respir Dis **1990**; 142:725–36.
2. Holden M, Dubin MR, Diamond PH. Frequency of negative intermediate-strength tuberculin sensitivity in patients with active tuberculosis. N Engl J Med **1971**; 285:1506–9.
3. Comstock GW. False tuberculin test results. Chest **1975**; 68(Suppl):465–9.
4. Chin DP, Osmond D, Page-Shafer K, et al. Reliability of anergy skin testing in persons with HIV infection. Am J Respir Crit Care Med **1996**; 153: 1982–4.
5. Wood PR, Rothel JS. In vitro immunodiagnostic assays for bovine tuberculosis. Vet Microbiol **1994**; 40:125–35.
6. Whipple DL, Bolin CA, Davis AJ, et al. Comparison of sensitivity of the caudal fold skin test and a commercial γ-interferon assay for diagnosis of bovine tuberculosis. Am J Vet Res **1995**; 56:415–9.
7. Wood PR, Corner LA, Rothel JS, et al. Field comparison of the interferon-gamma assay and the intradermal tuberculin test for the diagnosis of bovine tuberculosis. Aust Vet J **1991**; 68:286–90.
8. Converse PJ, Jones SL, Astemborski J, Vlahov D, Graham NMH. Comparison of a tuberculin interferon-γ assay with the tuberculin skin test in high-risk adults: effect of human immunodeficiency virus infection. J Infect Dis **1997**; 176:144–50.
9. Streeton J, Desem N, Jones S. Sensitivity and specificity of a gamma interferon blood test for tuberculosis infection. Int J Tuberc Lung Dis **1998**; 2:443–50.
10. Vlahov D, Anthony JC, Muñoz A. The ALIVE study, a longitudinal study of HIV-1 infection in intravenous drug users: description of methods and characteristics of participants. NIDA Res Monogr **1991**; 109:75–100.
11. Desem N, Jones SL. Development of a human gamma interferon enzyme immunoassay and comparison with tuberculin skin testing for detection of *Mycobacterium tuberculosis* infection. Clin Diagn Lab Immunol **1998**; 5:531–6.
12. Caiaffa WT, Graham NMH, Galai N, Rizzo RT, Nelson KE, Vlahov D. Instability of delayed-type hypersensitivity skin test anergy in human immunodeficiency virus infection. Arch Intern Med **1995**; 155:2111–7.
13. Bearman JE, Kleinman H, Glyer VV, LaCroix OM. A study of variability in tuberculin test reading. Am Rev Respir Dis **1964**; 90:913–9.