CLINICAL AND VACCINE IMMUNOLOGY, Apr. 2009, p. 541–543 Vol. 16, No. 4
1556-6811/09/$08.00+0 doi:10.1128/CVI.00375-08

Differences in Tuberculin Reactivity as Determined in a Veterans Administration Employee Health Screening Program▽

Sanjay R. Mehta,[1]* Cathy MacGruder,[2] David Looney,[1,2] Scott Johns,[2] and Davey M. Smith[1,2]

University of California, Division of Infectious Diseases, 9500 Gilman Drive, La Jolla, California 92093,[1] and San Diego Veterans Administration Medical Center, 3350 La Jolla Village Dr., San Diego, California 92161[2]

Received 5 September 2008/Returned for modification 10 December 2008/Accepted 9 February 2009

In response to a difference in pricing, the San Diego Veterans Administration Medical Center changed its tuberculin preparation from Tubersol to Aplisol in the fall of 2006. Following the change, an increased number of employee skin test conversions was noted. Employee tuberculin skin test converters from 2006 were screened with the QuantiFERON Gold (QFT-G) gamma interferon release assay. Those employees who tested negative by QFT-G were asked to repeat their skin test with both Tubersol and Aplisol tuberculin preparations. Of the new purified protein derivative converters, 12 of 14 returned for repeat testing with QFT-G, and the assay was negative for 83% (10/12), positive for 8% (1/12), and indeterminate for 8% (1/12) of the individuals. Nine of the individuals who were QFT-G negative agreed to repeat skin testing with both tuberculin preparations, and 7/8 (87.5%) demonstrated reactivity with the Aplisol preparation, while 0/8 (0%) reacted to the Tubersol preparation. A change from Tubersol to Aplisol resulted in elevated tuberculin skin test conversion rates that may be due to false-positive reactions. The differences in skin test reactivity between preparations support CDC guidelines that recommend that institutions should not change tuberculin preparations, as doing so may falsely increase the number of positive reactions.

Despite the introduction of gamma interferon release assays (IGRA), the standard tuberculin skin test (TST) remains the dominant method of screening for latent *Mycobacterium tuberculosis* infection (LTBI) in the United States, likely due to both cost and ease of use (8). However, since its inception, the TST has notoriously been subject to observer bias and other pitfalls that may lead to both false-negative and false-positive interpretations of reactions. False-negative reactions often occur in immunocompromised individuals and in some individuals with active tuberculosis, due to the intrinsic mechanism of the test, which requires the individual being tested to have the ability to mount a type IV immune response in the skin. False-positive reactions of the TST usually depend upon the antigen or antigens used for the skin testing preparation. This can be due to the cross-reactivity of the antigens themselves or the purity of the preparations.

Currently, the purified protein derivatives (PPD) that are used in the TST for LTBI consist of the purified protein from *M. tuberculosis* culture filtrates. These preparations have been standardized to PPD-S, the standardized PPD prepared by F. B. Seibert in 1939 (7). Although previous exposures to nontuberculous mycobacteria may lead to cross-reactivity to these preparations, they were a marked improvement from the original Koch preparation (7). Currently in the United States, two tuberculin preparations are marketed, Tubersol (Aventis-Pasteur) and Aplisol (Parkdale). Each of these preparations is prepared from a master lot of tuberculin, to prevent lot-to-lot variability, and both companies have supplies expected to last many years (7). Previous reports have noted differences in reactivity between preparations when tested on the same individual (1, 4, 10, 11); however, without a gold standard assay, it is unclear when an individual reacts to only one of the preparations whether it reflects differences in sensitivity or problems with specificity. The introduction of the IGRAs, in particular, QuantiFERON Gold (QFT-G), which uses peptide antigens derived from two genes (ESAT-6 and CFP-10) that are relatively specific for *M. tuberculosis* and pathogenic strains of *M. bovis*, has provided an alternative testing method with improved specificity; however, even this test has some cross-reactivity to mycobacteria other than *M. tuberculosis*, including *M. kansasii*, *M. marinum*, and *M. szulgai* (8).

At our institution, the San Diego Veterans Administration Medical Center (SDVAMC), we changed from the Tubersol tuberculin preparation to the Aplisol tuberculin preparation in the fall of 2006 in order to reduce costs, and soon afterward we noticed an increase in skin test conversion rates. We then performed the following investigation.

MATERIALS AND METHODS

We investigated the increase in positive TSTs in the SDVAMC employee tuberculosis screening program after the change in our tuberculin preparation from Tubersol to Aplisol in the fall of 2006. Tuberculin tests are mandatory for employees (~2,000) not known to be PPD positive on a yearly basis. Tests are performed using the Mantoux method with intermediate-strength (5 tuberculin units) PPD-S, and reading is performed by trained nurses in the tuberculosis screening program. A two-step method, as recommended by CDC guidelines (5), is used for initial screening at our institution, whereby an individual receiving his or her first annual TST is brought back for a second TST 2 to 3 weeks later if the initial test is negative. This is done to augment the response in individuals who have a diminished response after the first exposure to tuberculin antigen in many years, thereby avoiding false-negative TSTs. Individuals may elect to have QFT-G testing in lieu of a two-step or single TST.

A positive skin test conversion was defined as an induration of ≥10 mm in diameter in an individual without a previous reaction or an increase in size by ≥10 mm of the reaction from the previous year. All patients with positive skin tests were evaluated for active tuberculosis by assessment of symptoms and chest

* Corresponding author. Mailing address: Stein Clinical Research Building, MC 0711, 9500 Gilman Drive, La Jolla, CA 92093. Phone: (858) 822-4092. Fax: (858) 822-5362. E-mail: srmehta@ucsd.edu.

▽ Published ahead of print on 18 February 2009.

radiography. Testing with QTF-G is routinely offered for converters if there is concern over the accuracy of the interpretation of the TST.

A positive QFT-G test was defined according to CDC guidelines (8). If the normalized level of gamma interferon produced in response to either of the antigens ESAT-6 and CFP-10 was greater than 0.35 IU and greater than or equal to 50% of that of the negative control, the test was considered positive. If the normalized level of gamma interferon produced in response to ESAT-6 and CFP-10 was less than 0.35 IU, the negative control had a response less than 0.7 IU, and the mitogen-normalized response was greater than or equal to 0.5 IU, the test was considered negative. Otherwise, the results were read as indeterminate (8).

Each year, our institution screens roughly 2,000 employees with tuberculin skin testing. In 2006, our institution had 14 positive skin tests, while the average annual rate for the previous 5 years had been 2.2. Due to the concern of false-positive reactions, the administrators of the employee health tuberculosis control program asked these individuals to return for QFT-G testing. Twelve of these individuals were retested with QFT-G. Based upon these results, the administrators had concerns about the tuberculin preparations in use. In order to guide policy decisions, those individuals that initially tested positive by the Aplisol TST but negative by QFT-G were asked to return for further testing. A subset of nine of these individuals were retested using both the previous Tubersol stock and the Aplisol formulation simultaneously, with one on each arm.

RESULTS

The SDVAMC averaged 2.2 skin test conversions per year from 2001 to 2005, while the Tubersol tuberculin preparation was being used for the TST. During 2006, when the Aplisol tuberculin preparation was used, we had 14 employee conversions (new TST reaction of >10 mm), most without any identifiable source of exposure. Of these conversions, 12 were retested with QFT-G. Ten of these individuals tested negative (83%), one of them tested positive (8%), and one individual had an indeterminate test (8%). Of our Aplisol-positive/QFT-G-negative population, nine agreed to be retested with both tuberculin preparations, although one individual did not return for skin test reading. Of the eight who tested and returned for results, seven (87.5%) tested positive with Aplisol but none tested positive with the Tubersol preparation ($P = 0.0014$, two-tailed Fisher's exact test). Measurements of erythema and induration are shown in Table 1. The difference in induration produced by Aplisol and Tubersol was highly significant at 48 and 72 h ($P = 0.0006$ and $P = 0.0002$, respectively; two-tailed Mann-Whitney U test), while the difference in erythema was nonsignificant at 48 h and marginal at 72 h ($P = 0.25$ and $P = 0.04$, respectively; two-tailed Mann-Whitney U test).

Since returning to the Tubersol preparation for our screening, we have noticed no further skin test conversions as of February 2008.

TABLE 1. Sizes of induration and erythema in subjects tested concurrently with both Aplisol and Tubersol

Preparation	Patient	Size[a] (mm)			
		48 h		72 h	
		Induration	Erythema	Induration	Erythema
Aplisol	1			14	14
	2	15	18	10	16
	3	14	20	12	12
	4	18	18	15	20
	5	12	12	10	11
	6	20	20	14	15
	7	5	5	5	5
	8	14	0	14	0
Tubersol	1			0	0
	2	0	15	0	14
	3	0	0	0	0
	4	0	13	0	12
	5	0	12	0	0
	6	0	15	0	10
	7	5	15	0	0
	8	0	0	0	0

[a] Measurements were taken at 24, 48, and 72 h. Subject 1 did not have measurements taken at 48 h.

DISCUSSION

The differences in skin test reactivity between the tuberculin preparations Aplisol and Tubersol have been documented in several publications. Most reports comparing the two preparations found a higher number of positive reactions when the Aplisol preparation was used (1, 4, 12). The population found to be Aplisol positive but Tubersol negative ranged from 1.0% (1) to 1.4% (4) of the population tested, much lower than that observed here. Information on discordance in which the individuals are Tubersol positive but Aplisol negative is less well documented. Some have suggested that Aplisol may have a tendency to cause more erythema than Tubersol and that this erythema may be mistaken for a positive reaction (4). Others have suggested that Aplisol produces larger reactions than Tubersol (4, 12). With our small subset of eight employees that received both preparations and returned for results, our experienced tuberculosis coordinator found that Aplisol produced induration and increased erythema but that for the individuals with discordant reactions between preparations, the Tubersol created no induration, suggesting either that the Aplisol results represented false-positive reactions or that the Aplisol preparation was more sensitive than the Tubersol preparation.

Although the reported increase in the number of positive reactions after the change from Tubersol to Aplisol is on the order of 1% of the population tested, for a large institution this represents a major expense and inconvenience to a substantial number of individuals. Positive reactions require clinical evaluation, chest radiography, and possibly repeat QFT-G testing, and false-positive reactions could lead to 9 months of unnecessary and potentially toxic therapy for LTBI. At our institution, in addition to the costs of a chest X ray (approximately $120) and an IGRA (approximately $65), there is the cost of visits to the employee health physician, the tuberculosis control nurse, and the epidemiologic investigation that accompanies most of our employee conversions; the cost for each false-positive reaction can add up very quickly, even when the costs and potential morbidity of treatment are not considered.

At our institution, we attempted to obtain QFT-G assays for all of our TST converters receiving Aplisol since November 2006. We found that the results of the QFT-G test were more consistent with the skin test results derived from the Tubersol preparation than with those derived from the Aplisol preparation. As the QFT-G test uses only the ESAT-6 and CFP-10 antigens, it is considered a more specific test for the diagnosis of LTBI than standard tuberculin skin testing. The current tuberculin preparations in use in the United States are PPD of heat-inactivated, stationary-phase *M. tuberculosis* culture, with numerous proteins and epitopes present (7). Ultimately, how-

ever, as there is no gold standard test for the diagnosis of LTBI and since we do not have long-term data to know the true specificity and sensitivity of IGRAs, such as QFT-G, it is difficult to assess the utility of QFT-G in determining which of the two tuberculin preparations is more accurate.

By using populations thought to be free from tuberculosis and populations with active tuberculosis, the CDC has estimated the sensitivity and specificity of QFT-G to be 67 to 81% and 96 to 98%, respectively (3, 6, 9). Cellestis (Melbourne, Australia), the manufacturer of QFT-G, followed 41 individuals who tested positive with QFT-G but were untreated for 2 years and found that 6 (14.6%) progressed to active disease. Of 219 individuals within the same cohort who tested positive by a TST (Tuberculin-10-GT [Chiron Behring, Marburg, Germany] or RT23 [Statens Serum Institute, Copenhagen, Denmark]) and were untreated for 2 years, only 5 (2.3%) progressed to active disease (2). These studies suggest that the QFT-G assay is indeed more specific than the TST; however, without knowledge of the sensitivity of the QFT-G assay for LTBI, this test cannot be used either to usurp the traditional skin test or to confirm it, although the CDC recommends QFT-G in all situations where the TST may be used (8). We are also precluded from using QFT-G to determine whether the Aplisol preparation is more sensitive or less specific than the Tubersol preparation. Positive reactions require clinical evaluation, radiography, and 9 months of potentially toxic therapy. This makes it extremely important to reduce the rates of false-positive reactions.

The CDC has put forth a statement recommending that tuberculosis screening programs use a single tuberculin product, as changing preparations may make serial changes in skin tests difficult to interpret. The Advisory Council for the Elimination of Tuberculosis (ACET) has recommended that when switching from Tubersol to Aplisol, (i) the appropriate users are notified when the change is taking place, (ii) a systematic assessment is performed to exclude the possibility of an outbreak if a cluster of false-positive reactions in a health care setting is seen after the change, and (iii) Tubersol is used to retest if ongoing transmission has been ruled out and the QFT-G test is considered for use in ruling out or in the positive reactions (5).

These recommendations bring up several questions. Once again, since there is no gold standard test for asymptomatic LTBI, how can we determine whether one tuberculin preparation is more sensitive or less specific? A prospective study in which individuals are randomized to receive one preparation or the other could answer this question but would require a large population (with long-term follow-up), for whom isoniazid therapy would need to be withheld, which may not be ethically justifiable. However, until such data are available, institutions screening for tuberculosis around the country have to decide whether to use the less expensive Aplisol tuberculin preparation, which may be less specific and in the end cost more money through false-positive reactions, or the more expensive but potentially less sensitive Tubersol preparation. IGRAs may be helpful to rule in positive skin test reactions for therapy; however, a negative IGRA result would remain unhelpful, as data on the sensitivity of IGRAs need to be collected prospectively over many years. Additionally, it is possible that the negative results obtained with Tubersol and QFT-G tests corresponding to positive Aplisol results actually represent false-negative reactions of the Tubersol and QFT-G tests. Without a clear winner, tuberculosis screening departments around the country have to choose which test is appropriate for their population and resources, while awaiting further studies.

ACKNOWLEDGMENTS

This work was conducted under the auspices of the UCSD IRB, HRPP 081018X, and supported by grants 5T32AI007036, AI69432, AI043638, MH62512, MH083552, AI077304, AI36214, AI047745, and AI74621.

We report no conflict of interest.

REFERENCES

1. **Blumberg, H. M., N. White, P. Parrott, W. Gordon, M. Hunter, and S. Ray.** 2000. False-positive tuberculin skin test results among health care workers. JAMA **283:**2793.
2. **Diel, R., R. Loddenkemper, K. Meywald-Walter, S. Niemann, and A. Nienhaus.** 2008. Predictive value of a whole blood IFN-gamma assay for the development of active tuberculosis disease after recent infection with Mycobacterium tuberculosis. Am. J. Respir. Crit. Care Med. **177:**1164–1170.
3. **Ferrara, G., M. Losi, M. Meacci, B. Meccugni, R. Piro, P. Roversi, B. M. Bergamini, R. D'Amico, P. Marchegiano, F. Rumpianesi, L. M. Fabbri, and L. Richeldi.** 2005. Routine hospital use of a new commercial whole blood interferon-gamma assay for the diagnosis of tuberculosis infection. Am. J. Respir. Crit. Care Med. **172:**631–635.
4. **Gillenwater, K. A., S. C. Sapp, K. Pearce, and G. K. Siberry.** 2006. Increase in tuberculin skin test converters among health care workers after a change from Tubersol to Aplisol. Am. J. Infect. Control **34:**651–654.
5. **Jensen, P. A., L. A. Lambert, M. F. Iademarco, and R. Ridzon.** 2005. Guidelines for preventing the transmission of Mycobacterium tuberculosis in health-care settings, 2005. MMWR Recommend. Rep. **54:**1–141.
6. **Kang, Y. A., H. W. Lee, H. I. Yoon, B. Cho, S. K. Han, Y. S. Shim, and J. J. Yim.** 2005. Discrepancy between the tuberculin skin test and the whole-blood interferon gamma assay for the diagnosis of latent tuberculosis infection in an intermediate tuberculosis-burden country. JAMA **293:**2756–2761.
7. **Lee, E., and R. S. Holzman.** 2002. Evolution and current use of the tuberculin test. Clin. Infect. Dis. **34:**365–370.
8. **Mazurek, G. H., J. Jereb, P. Lobue, M. F. Iademarco, B. Metchock, and A. Vernon.** 2005. Guidelines for using the QuantiFERON-TB Gold test for detecting Mycobacterium tuberculosis infection, United States. MMWR Recommend. Rep. **54:**49–55.
9. **Mori, T., M. Sakatani, F. Yamagishi, T. Takashima, Y. Kawabe, K. Nagao, E. Shigeto, N. Harada, S. Mitarai, M. Okada, K. Suzuki, Y. Inoue, K. Tsuyuguchi, Y. Sasaki, G. H. Mazurek, and I. Tsuyuguchi.** 2004. Specific detection of tuberculosis infection: an interferon-gamma-based assay using new antigens. Am. J. Respir. Crit. Care Med. **170:**59–64.
10. **Page, E. H., and R. J. Driscoll.** 2007. Evaluation of results from occupational tuberculin skin tests—Mississippi, 2006. MMWR Morb. Mortal. Wkly. Rep. **56:**1316–1318.
11. **Rupp, M. E., A. W. Schultz, Jr., and J. C. Davis.** 1994. Discordance between tuberculin skin test results with two commercial purified protein derivative preparations. J. Infect. Dis. **169:**1174–1175.
12. **Villarino, M. E., W. Burman, Y. C. Wang, L. Lundergan, A. Catanzaro, N. Bock, C. Jones, and C. Nolan.** 1999. Comparable specificity of 2 commercial tuberculin reagents in persons at low risk for tuberculous infection. JAMA **281:**169–171.