Self-Assessment of Tuberculin Skin Test Reactions by New York City Firefighters: Reliability and Cost-Effectiveness in an Occupational Health Care Setting

David J. Prezant, MD; Kerry J. Kelly, MD; Manoj L. Karwa, MD; and Kathryn Kavanagh, BS

Objective: To determine whether self-assessment of purified protein derivative of tuberculin (PPD) skin test reactions, done using a simple two-choice approach, is an effective screening method for tuberculosis.

Design: Double-blind comparison between self-assessments and trained professional readings of PPD skin test reactions, done 72 hours after test administration.

Setting: The New York City Fire Department's Bureau of Health Services

Participants: 2011 New York City firefighters and fire officers were given PPD skin tests during a mandatory retraining course. Thirty-seven persons were excluded because of a history of a positive PPD skin test result or a bacille Calmette-Guérin vaccination. All others agreed to participate in testing and self-assessment done using simple written instructions. Self-assessment results were submitted just before trained professional readings were done.

Measurements: Self-assessments and trained professional readings of PPD skin test reactions.

Results: 1833 participants (91%) interpreted their test reactions as flat. Of these interpretations, 1824 (99.5%) matched the professional reading and 9 (0.5%) did not. One hundred seventy-eight participants (9%) interpreted their test reactions as not flat; 136 of these interpretations (76.4%) matched the professional reading and 42 (23.6%) did not ($\kappa = 0.828$; lower 95% confidence limit = 0.790). The predictive value of a negative self-assessment reading was 99.5%, and the specificity was 97.7%.

Conclusion: In this occupational health care setting, we follow (and recommend to others with similar populations) a tuberculin screening program based on self-assessment. Repeated tests with follow-up are required for all persons who do not report their results. All persons with self-assessments of "not flat" should return for readings by trained professionals, counseling, and treatment.

Ann Intern Med. 1996;125:280-283.

From the New York City Fire Department, Bureau of Health Services, New York, New York; Albert Einstein College of Medicine and Montefiore Medical Center, Bronx, New York; and Staten Island University Hospital, Staten Island, New York. For current author addresses, see end of text.

Health care and regulatory forces have united to place greater emphasis on mandatory tuberculin skin test screening, done at least annually, for health care providers, school children, teachers, and other high-risk groups. Although some clinicians ask patients or parents to read their own or their children's tuberculin skin tests, it remains unclear whether this practice is reliable. In the largest study of adults (737 volunteers) done to date (1), the authors found that "patient's readings of their own results were inaccurate and should never be relied on to make clinical decisions." However, other investigators (2–4) report consistent agreement between self-assessments and professional readings. Differences between studies may have resulted from different demographics or inadequate or complicated instructions.

We did a prospective study of 2011 New York City firefighters and fire officers to compare the effectiveness of self-assessment of purified protein derivative of tuberculin (PPD) skin test reactions with independently measured readings by trained professionals. We designed simple written instructions explaining that erythema should be ignored, self-assessment should be limited to the area of induration, and only two descriptors of the lesion should be used: "flat" or "not flat." Because all self-assessed readings of "not flat" require follow-up by a clinician, our aim was to eliminate false-negative results even at the expense of increasing the number of false-positive results. Elaborate instructions confuse the person doing the self-assessment and impede the goals of such assessment which are to improve compliance with PPD reading and reporting; maintain reliability; and reduce costs by substantially decreasing the number of persons requiring follow-up visits (from everyone tested to only persons with self-assessed positive results). Cost savings are discussed to highlight the effect of a successful self-assessment program.

Methods

This study was done at the New York City Fire Department's Training Academy under the supervision of the New York City Fire Department's Bureau of Health Services. Firefighters and fire officers assigned to retraining were given a Mantoux PPD

skin test. All persons who reported for retraining ($n = 2048$) gave written consent for PPD testing and agreed to participate in self-assessment. Thirty-seven persons were excluded because of a history of a positive result on a PPD skin test or a bacille Calmette-Guérin vaccination. All participants spoke English, were at least high school graduates, and had passed a competitive employment examination. Demographic characteristics of the study sample are shown in Table 1.

Tuberculin skin testing was done using the Mantoux technique by injecting PPD (5 tuberculin units) intradermally in the vola of the forearm. A wheal at least 6 mm in diameter was produced. After 72 hours, each person assessed their own PPD reaction using the following written instructions:

> Please run your finger over the skin test placed on your arm several days ago. If you feel absolutely no swelling in this area (redness is not important) or in other words if you feel the area to be absolutely flat when you run your finger over it, check the box marked FLAT. If you feel any swelling whatsoever in this area (redness is not important) or in other words if the area feels raised or not flat when you run your finger over it, check the box marked NOT FLAT.

No other instructions were provided. Within 15 minutes after the self-assessment reports were submitted, each person's PPD skin test reaction was measured independently by a trained registered nurse who had no knowledge of the result of the self-assessment. Induration was assessed by palpation and was measured perpendicular to the long axis of the forearm. Erythema was ignored. Induration of at least 10 mm (rather than 15 mm) was considered positive, because our participants live in an area where tuberculosis is prevalent and where contamination with nontuberculous mycobacterium is rare. Five millimeters was considered negative, because this was a screening rather than a contact investigation and because no participants were immunosuppressed. All positive readings by a trained nurse were confirmed the same day by a trained physician (Table 2). Chest radiographs (posteroanterior, lateral, and apical oblique) were negative for disease in all persons with PPD skin test reactions that were confirmed to be positive.

Overall agreement between the participants and trained professionals was calculated for all 2011 paired PPD readings (Table 2), and κ scores with 95% CIs were computed. When a nurse determined a test to have a positive result, the test was subsequently read by a physician; both readings are reported (Table 2). However, data analysis was based on the physician's interpretation because the physicians were considered to be more experienced. The effect of demographics on differences between self-assessed and professional readings was analyzed for significance by analysis of variance (Statgraphics software, version 6.1, 1993, STSC, Inc., Rockville, Maryland). Demographic characteristics analyzed were rank (firefighter or fire officer), age, marital status, and educational status. Sex and race or ethnicity were not analyzed because numbers were inadequate within subgroups. Statistical significance was defined using an overall type I error of 0.05.

Table 1. Demographic Characteristics Related to Self-Assessment of PPD Skin Test Results*

Demographic Characteristic	Persons with Self-Assessed Flat (Negative) PPD Skin Test Reaction		Persons with Self-Assessed Raised (Positive) PPD Skin Test Reaction	
	True-Negative Result	False-Negative Result	True-Positive Result	False-Positive Result
	n (%)			
Total	1824 (99.5)	9 (0.5)	136 (76.4)	42 (23.6)
Rank				
Firefighters	1097 (99.5)	5 (0.05)	103 (83.1)	21 (36.9)
Fire officers	727 (99.4)	4 (0.06)	33 (61.1)	21 (38.9)
Sex				
Male	1816 (99.5)	9 (0.05)	136 (76.4)	42 (33.6)
Female	8 (100)	0 (0)	0 (0)	0 (0)
Age				
<35 y	522 (99.8)	1 (0.02)	16 (64)	9 (36)
≥35 y	1302 (99.3)	8 (0.07)	120 (78.4)	33 (21.6)
Race or ethnicity				
Black	35 (97.2)	1 (0.08)	8 (80)	2 (20)
Asian	6 (100)	0 (0)	0 (0)	2 (100)
White	1751 (99.5)	8 (0.05)	122 (76.7)	37 (33.3)
Hispanic	32 (100)	0 (0)	6 (85.7)	1 (14.3)
Marital status				
Single	355 (99.4)	2 (0.06)	17 (73.9)	6 (36.1)
Married	1469 (99.5)	7 (0.05)	119 (76.8)	36 (33.2)
Level of education				
High school diploma	564 (99.2)	4 (0.08)	88 (84.6)	16 (15.6)
College courses	684 (99.4)	4 (0.06)	28 (57.1)	21 (42.9)
College degree	576 (99.8)	1 (0.02)	20 (80)	5 (20)

* PPD = purified protein derivative of tuberculin.

Results

Because this study protocol was concurrent with mandatory retraining at the New York City Fire Academy, compliance with PPD skin testing was 100% for self-assessment reporting and follow-up reading by a trained nurse. Test reactions were interpreted as flat by 1833 (91%) persons; 1824 (99.5%) of these interpretations matched the final professional reading and 9 (0.5%) did not (Table 2). Test reactions were interpreted as not flat by 178 (9%) of persons; 136 (76.4%) of these interpertations matched the final professional reading and 42 (33.6%) did not ($\kappa = 0.828$; lower 95% confidence limit = 0.790). Using the final professional reading as the gold standard (5), the predictive value of a negative self-assessment reading was 99.5% with a specificity of 97.7%. Rank, age, marital status, and level of education were not significant factors in

Table 2. Self-Assessment Compared with Trained Professional Readings of PPD Skin Test Reactions*

Interpretation of PPD Skin Test Reaction	Interpretation by Trained Nurse		Interpretation by Trained Physician†		Final Professional Interpretation‡	
	Negative	Positive§	Negative	Positive§	Negative	Positive§
Result of self-assessment						
Negative (n = 1833)	1824	9			1824	9
Positive (n = 178)	39	139			42	136
Test found positive by trained nurse (n = 148)			3	136		

* Values are numbers of participants. PPD = purified protein derivative of tuberculin.
† Trained physicians only evaluated tests that were found positive by trained nurses.
‡ Combination of readings by trained nurses and readings by trained physicians.
§ Induration ≥10 mm.

predicting agreement between self-assessment and trained professional readings (Table 1).

Discussion

Medical authorities recommend tuberculin screening for high-risk patients, and regulatory agencies mandate tuberculin screening for health care workers, school children, and teachers. Mantoux PPD is the preferred screening test in most settings (5). With a captive population (such as inpatients or prisoners), variables inherent in a successful screening program (such as participant enrollment and training in PPD placement technique and reading) can be easily controlled (5, 6). In the community, however, the greatest obstacle to screening is compliance with follow-up appointments for skin test readings 48 to 72 hours after placement. Compliance with follow-up appointments ranges from 18% (7) to 84% (8). This can be improved if, as in our study, testing and reading is scheduled around mandated activities that last 48 to 72 hours. In the community, annual follow-up costs for tuberculin skin test readings by professionals are high, especially if salary and lost work time are taken into account. This cost increases as compliance increases, and it doubles in persons who require "booster" or biannual PPD skin tests.

Effects of self-assessment on compliance, reliability, and cost have received surprisingly little attention. There have been few pediatric or adult studies (1–4, 7–9), and those that have been done have included only a small number of participants and have been confounded by family and social issues, comorbid illness, language barriers, and educational abilities. In our experience, follow-up phone calls improve reporting compliance to nearly 100% (Unpublished data). However, improved compliance by return mail, phone, or fax is only valuable if self-assessments and professional readings are shown to agree. For Tine tests done in children, there was 100% agreement between parental and professional readings (2). In college students (4) or veterans (3) trained in self-assessment of PPD skin test reactions, excellent agreement with professional readings was seen (when positivity was defined as ≥ 10 mm induration). Unfortunately, nearly half of the veterans (mean age, 60 years) could not be properly trained in self-assessment (3). In health workers of varying educational status, poor agreement between self-assessments and professional readings was found: Five-hundred twenty of 525 participants correctly read their reactions as negative (<10 mm induration), but only 79 of 212 correctly read them as positive (1).

Because self-assessment requires persons with induration to have follow-up professional readings, we believe that instructions for self-assessment should be as simple as possible, with the aim of eliminating false-negative results even at the expense of increasing the number of false-positive results. Instead of teaching persons how to measure PPD induration, a skill that requires good vision, absence of organic disease, and training (3), we asked persons to ignore erythema and simply state whether the reaction was flat or not flat. We found that this two-choice approach required no teaching and was highly successful. This method is consistent with our self-assessment goal: to provide optimal care by evaluating only test results that are suspected to be positive. This improves compliance and decreases cost by eliminating professional readings of tests for which no action is required.

In our occupational setting, where language skills and education were assured, we found that the ability to assess a negative (flat) PPD skin test reaction was so simple that inability often reflected comorbid illness or a dysfunctional employee–employer relationship. We admit that the high rate of agreement between self-assessment and professional readings observed in our study may not occur in other settings. Our participants were English-speaking, adult volunteers of working age; they were at least high school graduates; they had passed a competitive employment examination; and they lacked comorbid illness. They understood that a positive result on a PPD skin test required counseling, indicated a potential and genuine risk for a contagious disease

that could be almost entirely reduced with treatment, and presented no risk for employment termination or suspension without pay.

Given the proven effectiveness of this self-assessment program, substantial cost savings are possible. Because all fire department personnel (approximately 12 000 persons) and emergency medical service uniformed employees (approximately 3000 persons) in New York City visit a centralized employee's health service annually for medical examinations, a PPD skin test can be administered without substantial cost. However, arranging follow-up readings can be complicated because, unlike employees of hospitals or clinics, our employees are assigned to many locations (> 350) where professional readers are not available. If a professional follow-up reading is required, the cost is substantial, regardless of whether the reading is done by mandated return visit to the employee's health service while on or off duty (both contractually require salary compensation), in private medical offices, or in field-based mobile units. The estimated cost (including lost time at work) of providing annual professional readings of PPD skin tests for 15 000 New York City fire and emergency medical service uniformed employees would be approximately $2.5 million. Projecting costs on a national basis is difficult because of varying contractual agreements and numbers of volunteers. Approximately 1 000 000 firefighters (10) and 700 000 emergency service workers (11) were employed in the United States in 1993. Given the least expensive scenario—no salary reimbursement, no time lost from work, and a professional charge per follow-up visit of $20.00 per worker—the estimated annual cost in the United States would be $34 million. Biannual testing would cost twice that amount. In contrast, self-assessment programs require mandatory follow-up visits only for reactions self-assessed as positive. Using our observed rate of 9% (false-positive results, positive reactions of unknown duration, and new converters), self-assessment would reduce follow-up costs by almost 90% in the first year and by at least 90% thereafter. The cost associated with the low false-negative rate in this study was not factored into that amount because, as a result of known variability (6), similar false-negative rates would be expected from professional readers.

In conclusion, we follow (and recommend to others with similar populations) a tuberculin screening program that uses two-choice self-assessment ("flat" or "not flat") and reporting done by pre-addressed postcard, fax, or phone. Repeated tests with follow-up are required for all persons who do not report results. All employees with reactions found to be positive by self-assessment must return in person for professional reading, counseling, and treatment. In persons who are incapable of self-assessment because of language barriers, lack of education, or comorbid illness and in symptomatic patients being tested for diagnostic purposes, we recommend that all PPD skin test reactions be read professionally.

Acknowledgments: The authors thank the administration and uniformed service personnel of the New York City Fire Department for their cooperation and encouragement; the physicians and nurses of the New York City Fire Department's Bureau of Health Services for advice and help with data collection; Drs. C. Jones, J. Scheuer, T.K. Aldrich, J. Casey, and M. Catalano for encouragement and advice in study design; Ms. B. Einstein for database instruction; and Mr. Colin Jost for data entry.

Grant Support: By a research fellowship grant from the New York Lung Association (Dr. Karwa).

Requests for Reprints: David J. Prezant, MD, Montefiore Medical Center, Pulmonary Division, Centennial 423, Bronx, NY 10467.

Current Author Addresses: Drs. Prezant and Karwa: Montefiore Medical Center, Pulmonary Division, Centennial 423, Bronx, NY 10467.
Drs. Kelly and Prezant: New York City Fire Department, Bureau of Health Services, 251 Lafayette Street, New York, NY 10012.
Ms. Kavanagh: Mount Sinai Medical School, New York, NY 10029.

References

1. **Howard TP, Solomon DA.** Reading the tuberculin skin test. Who, when, and how? Arch Intern Med. 1988;148:2457-9.
2. **Maqbool S, Asnes RS, Grebin B.** Tine test compliance in a clinic setting. Pediatrics. 1975;55:388-91.
3. **Risser NL, Belcher DW, Bushyhead JB, Sullivan BM.** The accuracy of tuberculin skin tests: self-assessment by adult outpatients. Public Health Rep. 1985;100:439-45.
4. **Navin JA, Kaplan JE, Desilvio EL.** Self-reading of PPD skin tests. J Am Coll Health. 1994;43:37-8.
5. **Snider DE Jr.** The tuberculin skin test. Am Rev Respir Dis. 1982;125(3 Pt 2):108-18.
6. **Bearman JE, Kleinman H, Glyer VV, et al.** A study of variability in tuberculin skin test reading. Am Rev Respir Dis. 1985;132:177-8.
7. **Weinberger HL, Terry C.** Tuberculin testing in a pediatric outpatient clinic. J Pediatr. 1969;75:111-15.
8. **Asnes RS, Maqbool S.** Parent reading and reporting of children's tuberculin skin test results. Chest. 1975;68(3 Suppl):459-62.
9. **Wiecha JM, Lim M.** Tine test cards for TB screening: rates of return and associated factors. Fam Pract Res J. 1994;14:51-7.
10. **National Fire Protection Association.** United States Fire Department Profile through 1993 Report. Quincy, MA: National Fire Protection Assoc; 1994:3.
11. **Emergency Medical Services Communications.** United States Emergency Medical Services Market Report. Carlsbad, CA: Emergency Medical Services Communications; 1995:28.