Assessing the Validity of Tuberculin Skin Test Readings by Trained Professionals and Patients*

Philip O. Ozuah, MD; William Burton, PhD; Keith A. Lerro, MD, PhD; Jordan Rosenstock, MD; and Michael Mulvihill, DrPH

**Study objectives:** To assess the validity of purified protein derivative (PPD) readings by patients and trained health-care professionals as compared with a calibrated model.

**Design and participants:** Survey of a group of patients, nurses, medical assistants, and physicians at five neighborhood health centers in the Bronx, NY.

**Interventions:** Participants were asked to read a calibrated model with four PPD indurations measuring 0 mm, 3 mm, 7 mm, and 13 mm. Indurations ≥ 5 mm were to be considered “positive” reactions.

**Measurements and results:** Data were obtained from 233 patients and 80 trained professionals. All patients correctly measured the 0-mm induration site and were able to detect the presence of an induration in 99.3% of possible observations. Compared with professionals, patients had more variability in measurements and interpretations of the 3-, 7-, and 13-mm sites. Professionals detected 100% of all indurations. Patients’ specificity for the 0- and 3-mm sites was 97.4% and 62.7%, respectively; whereas sensitivity for the 7- and 13-mm sites was 68.2% and 59.3%, respectively. Professionals’ specificity for the 0- and 3-mm sites was 98.7% and 65.3%, respectively; their sensitivity for the 7- and 13-mm sites was 86.7% and 97.3%, respectively. Seventy percent of professionals agreed that the model was a realistic representation of PPD indurations.

**Conclusions:** Patients can reliably distinguish between the presence and absence of an induration at a PPD injection site. They are not as reliable in the measurement and interpretation of test reactions. Professionals had considerable variability in their assessments of PPDs but were more precise overall in their assessments than patients. (CHEST 1999; 116:104–106)

**Key words:** Mantoux purified protein derivative; self-assessment; tuberculin skin test; tuberculosis

**Abbreviation:** PPD = purified protein derivative

The Mantoux 5 Todd unit of purified protein derivative (PPD) of tuberculin PPD is now the accepted screening modality for tuberculosis.1–3 The recommendation is that PPD test reactions be read and interpreted only by trained health-care workers, thereby necessitating a return visit by patients. This recommendation is based on reports that found test reaction readings by patients to be unreliable4; however, other reports have shown that patients could reliably interpret some PPD test reactions.5,6 All of these studies have measured the ability of patients to read PPD reactions on the basis of determinations of the level of agreement between self-assessments and assessment by health professionals. It has been documented, however, that there can be considerable variability in the measurement of test reactions even among skilled professionals.7,8 Hence, we undertook a survey to assess the validity of PPD readings by professionals and patients as compared with a calibrated model developed in a research laboratory.

**Materials and Methods**

We designed a calibrated model made of acrylic and rubber that had four randomly placed induration sizes measuring 0, 3, 7, and 13 mm. The model was pilot tested among a group of five infectious diseases clinicians with expertise in tuberculosis and was revised until it was judged to be a realistic representation by all the members of the group. Each of these clinicians read tuberculin skin tests several times a week including a large number of positive results. The clinicians unanimously certified the model to be consistent with the feel, texture, depth, and

*From the Departments of Pediatrics (Dr. Ozuah) and Family Medicine (Drs. Burton, Lerro, Rosenstock, and Mulvihill), Albert Einstein College of Medicine, Bronx, NY.

This manuscript was presented at the Annual Meetings of the Pediatric Academic Societies in Washington, D.C., on May 3, 1997.

Manuscript received August 4, 1998; revision accepted February 24, 1999.

Correspondence to: Philip O. Ozuah, MD, Montefiore Medical Center, 3544 Jerome Ave, Bronx, NY 10467; e-mail: Pozuah@pol.net

104 Clinical Investigations

Downloaded From: http://journal.publications.chestnet.org/pdfaccess.ashx?url=/data/journals/chest/21927/ on 06/27/2017
contours of tuberculin skin test indurations. There was no facsimile of erythema on the model. This model was used as the validation criterion in this study. For the purposes of this study, we defined a positive PPD reaction as an induration of 5 mm or greater. Because the diagnostic level in clinical situations varies from 5 to 15 mm depending on the risk category, we chose the most conservative criterion of 5 mm to represent a positive reaction to assess the ability of participants to detect infection even among the highest risk patients, including those with immunosuppressed conditions.

The study was performed at several community health centers located in the inner city of the Bronx, NY, where 233 adult patients voluntarily participated in the study. Data were collected from 80 professionals, of whom 28 (36%) were physicians, 21 (28%) were nurses, and 27 (35%) were medical assistants (three professionals did not indicate their titles). Sixty-five percent of the professionals read PPDs at least once weekly and 32.4% read at least two PPDs per month. On average, the professionals had 10 years of experience and read three PPDs a week.

After obtaining consent, the model was presented to participants along with a plastic ruler calibrated in millimeters. The following written instructions were read by an investigator: “The areas labeled A, B, C, D, are reactions that may form at the site of a PPD injection. Please use your finger to feel for the presence of a bump in each area. If you feel a bump, please use the plastic ruler to measure it from one side to the other in the transverse direction (investigator demonstrates). If you do not feel a bump or if you measure the bump to be less than 5 mm, you should interpret that site as a negative PPD test. Any bump that is at least 5 mm or more is a positive PPD test. You may begin with any of the four sites.” All measurements and interpretations were recorded by the investigators at the time of the reading. In addition a questionnaire was administered to the professionals asking their opinions of the model. Demographic data were collected from the patients.

Specificity measures were obtained for the 0- and 3-mm indurations to assess for interpretation of these sites as true negative reactions. Sensitivity measures were obtained for the 7- and 13-mm induration sites because these were considered positive reactions. We did not combine the two “negative” and two “positive” induration sizes each for the specificity and sensitivity measures, because the differences between induration levels could be important. χ² analyses were used to determine differences between patients’ and professionals’ sensitivity and specificity rates for each induration size. Two-tailed t tests were used to analyze differences between the mean measurements of the four induration sites, and an F ratio was used to evaluate whether there was a significant difference in variability in the readings between the groups.

To assess whether individual patient characteristics were pre-
dicators of ability to measure and interpret PPDs, regression analyses were performed. Two regression models were used. For the first model, using multiple linear regression, the dependent variable was measurement of indurations; and for the other, using logistic regression, the dependent variable was the interpretation of results. Independent variables for both equations were sex, age, educational level, and occupation (health-related field/non–health-related field).

Attributes of professionals were assessed by t test analysis. This study was approved by the Institutional Review Board at the Montefiore Medical Center, Bronx, NY.

**RESULTS**

Data were collected from 233 patients, of whom 84% were women and 94% were black or Hispanic. The mean age was 28.9 years (range, 19 to 61 years). Sixty-four percent of the patients had completed high school, and 96% had at least an eighth-grade education.

The vast majority of patients accurately assessed the 0-mm site, but there was considerable variability in the patients’ assessment of the 3-, 7-, and 13-mm sites, as depicted in Tables 1 and 2. However, of the 699 total observations (233 for each induration size), patients detected the presence of an induration in 694 of the observations (99.3%).

Educational level, age, sex, and employment in a health-related field were neither strong predictors of patients’ ability to measure PPD indurations ($R^2 = 0.025$) nor their ability to interpret the results ($R^2 = 0.036$).

The professionals’ performance on various induration sites is also presented in Tables 1 and 2. Professionals had the most difficulty assessing the 3-mm induration size. Educational qualification and experience with reading PPDs were not significant factors in the professionals’ ability to measure and interpret PPDs except on the 3-mm induration site. Nonphysicians had a mean measurement of this induration size that was 1.1 mm greater than that of physicians ($p = 0.0003$). In addition, professionals who read PPDs less often than once weekly had a

### Table 1—Comparison of Patients’ and Professionals’ Interpretations of PPD Reactions*

<table>
<thead>
<tr>
<th>Induration Size</th>
<th>Participants</th>
<th>+ PPD</th>
<th>PPD</th>
<th>Specificity, %</th>
<th>Sensitivity, %</th>
<th>p Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>0 mm Patients</td>
<td>6</td>
<td>227</td>
<td>97.4</td>
<td>0.523</td>
<td></td>
<td></td>
</tr>
<tr>
<td>0 mm Professionals</td>
<td>1</td>
<td>74</td>
<td>98.7</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3 mm Patients</td>
<td>57</td>
<td>146</td>
<td>62.7</td>
<td>0.676</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3 mm Professionals</td>
<td>26</td>
<td>49</td>
<td>65.3</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7 mm Patients</td>
<td>159</td>
<td>74</td>
<td>68.2</td>
<td>0.002</td>
<td></td>
<td></td>
</tr>
<tr>
<td>7 mm Professionals</td>
<td>65</td>
<td>10</td>
<td>86.7</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>13 mm Patients</td>
<td>208</td>
<td>25</td>
<td>89.3</td>
<td>0.032</td>
<td></td>
<td></td>
</tr>
<tr>
<td>13 mm Professionals</td>
<td>73</td>
<td>2</td>
<td>97.3</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*+ = positive for PPD; = negative for PPD.
mean measurement of this site that was 0.9 mm greater than that of professionals who read PPDs more frequently (p = 0.006).

Seventy percent of professionals agreed that the acrylic/rubber model was a realistic representation of PPD indurations.

We found no differences between patients and professionals in their specificity rates for the 0-mm site, that is, in their ability to detect the absence of an induration.

Professionals generally had significantly better sensitivity rates (Table 1) and less variability in their measurements (Table 2).

**DISCUSSION**

In our study, patients were able to detect the presence of an induration > 99% of the time. Patients performed as well as professionals in detecting the absence of an induration. This suggests that patients and parents can be relied on to assess whether there has been a reaction around the site of a tuberculin injection. There was, however, more variability in the patients’ ability to interpret actual induration sizes.

We also found variability among professionals, particularly with the “borderline” indurations of 3 and 7 mm. Overall, accuracy in interpretation was lower for these induration sites compared with the 13-mm induration. We did not allow professionals to use the ballpoint method to avoid marks on the model that might have biased subsequent observers. Three of the professionals indicated their preference for this method. Excluding the data from these three participants did not affect our results, and hence we decided to include them in the analysis.

Compliance with follow-up appointments for PPD results is a major problem in practice settings. We recommend that patients receiving tuberculin skin tests or their guardians be instructed as in our study. All patients who detect the presence of an induration of any size or who are uncertain must return for assessment of test results by trained professionals. All other patients can relay their test results by telephone at 48 h. Follow-up phone calls can be made to nonresponders between 48 and 72 h. This approach will greatly reduce health costs by eliminating the vast majority of unnecessary follow-up appointments for nonreactive patients, patients’ transportation costs, and lost time from work, as well as the expense of recalling and retesting a large number of noncompliant patients.

We suggest that patients without telephones should make a return visit for test results. We do not recommend the use of preaddressed postcards because of poor compliance with this modality in the past and also because it would effectively eliminate the 48-h “window” during which nonresponders can still be contacted for results. We recognize the possibility that patients may falsely report a negative result to avoid the consequences of a positive result. However, we are also aware that some patients do not comply with therapy even after test results have been interpreted by a trained professional. There are clearly many complex issues and motivations involved in human behavior, but such a discussion is beyond the scope of this study.

ACKNOWLEDGMENTS: The authors thank Jennifer Pierce, MPH, for assistance in data gathering. Part of this project was completed by two medical students (J.R. and K.L.) in partial fulfillment of the requirements for the Family Medicine Clerkship.

### References

2. Centers for Disease Control, the American Thoracic Society. Core curriculum on tuberculosis. 2nd ed. Atlanta, GA: Centers for Disease Control and Prevention, 1991

<table>
<thead>
<tr>
<th>Induration Size</th>
<th>Patients Mean ± SD</th>
<th>Professionals Mean ± SD</th>
<th>p Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>3 mm</td>
<td>4.45 ± 2.75</td>
<td>4.42 ± 1.53</td>
<td>0.88</td>
</tr>
<tr>
<td>7 mm</td>
<td>6.10 ± 2.71</td>
<td>7.12 ± 2.33</td>
<td>0.003</td>
</tr>
<tr>
<td>13 mm</td>
<td>8.30 ± 3.44</td>
<td>9.76 ± 2.58</td>
<td>0.0001</td>
</tr>
</tbody>
</table>

Table 2—Comparison of Patients’ and Professionals’ Measurements of Various Induration Sizes