Multiple Puncture Skin Test and Mantoux Test in Southeast Asian Refugees*

Antonino Catanzaro, M.D., F.C.C.P.

Skin test reactions to PPD applied by Mantoux techniques were compared with reactions to tuberculin tine test (PPD-tine); tuberculin, old tine test (OT tine); Aplitest; and Mono-Vacc, tuberculin, old (Mono-Vacc) in newly arrived refugees from Cambodia and Laos. The reaction to Mantoux test was accepted as the "true reading" and compared to the reaction size to one of the multiple puncture tests (MPT). A $2 \times 2$ table was constructed and sensitivity, specificity, false negative, and false positive rates computed over a broad range of cut-off points for MPT. The MPTs were very sensitive (100 percent to 78 percent) but lacked specificity (78 percent to 18 percent) when a cut-off of 1 mm was used. Predictably, as the cut-off is moved to larger reactions, sensitivity decreases and specificity increases. The relationship between the two is emphasized in a receiver-operator characteristics analysis. The MPTs are not intended for diagnostic use. Reaction to the MPT should be interpreted with the same careful consideration that clinicians use to interpret reactions to the Mantoux test.

Since 1975, 600,000 refugees have been resettled from Vietnam, Cambodia, and Laos into the United States. Tuberculosis is the most important public health problem for the refugees. The active case rate in refugees during the peak influx was 100-fold greater than among the general population (1,500/100,000) either had tuberculosis at the time of entry or developed it before the end of 1980. The incidence of positive tuberculin skin test varies significantly with age; in general, however, 65 percent of incoming refugees have positive reaction on testing by Mantoux test.

Intradermal skin testing presents a number of technical problems. Multiple puncture skin tests have been designed to simplify skin testing. Unfortunately, the multiple puncture skin tests present a new set of problems, primarily related to interpretations of reaction size. Some investigators have found the new problems to lead to an incidence of false-negative reaction as to sharply limit the usefulness of these tests. Since the refugees from Southeast Asia have a high incidence of tuberculin reactions, they provide an excellent study population.

In this study, a group of refugees were studied shortly after arrival in San Diego, and their reactions to PPD applied by Mantoux technique with tuberculin tine test (PPD-tine); tuberculin, old tine test (OT tine); Aplitest; and Mono-Vacc, tuberculin, old (Mono-Vacc) were compared. An extensive analysis of reaction size is presented.

METHODS

Study subjects were drawn from the Cambodian and Laotian

Methods supported in part by American Cyanamid. Manuscript received July 9, revision accepted September 12.

*From the Department of Medicine, University of California, School of Medicine, San Diego.
†Associate Professor of Medicine.

Skin Testing

All skin tests were applied and read by one of four well-trained skin tests technicians. Mantoux skin tests were administered using plastic tuberculin syringes with 27 gauge needles. Syringes were loaded immediately prior to use with Tubersol. A dose of 0.1 ml was injected intradermally. In addition, each subject received two of the four multiple puncture skin tests. One group received OT tine (tuberculin, old tine test; and Mono-Vacc, tuberculin, old). The other group received PPD tine (tuberculin tine test) and Aplitest (PPD). All reagents were purchased through commercial sources. Readings were performed 48 hours after skin test antigen had been administered. Three skin test technicians each read one skin test site, without knowledge of the other two readings or examining the other skin test reactions. All reactions were measured in millimeters. Reactions were read as follows: the reaction at each tine puncture site was recorded separately as the largest cross-sectional diameter to either a single tine or the coalescent tine; if there was any induration, it was scored as 1 mm. If the reaction to the four tines coalesced into one area of induration, the cross-sectional diameters of that coalescent skin test reaction were recorded.

For the analysis conducted in this report, the largest cross-sectional diameter to either a single tine or the coalescent tine or the Mantoux test was used.

Statistical Analysis

The operating characteristics of the multiple puncture skin tests were determined using a standard $2 \times 2$ table (Fig 1). The Mantoux skin test was defined as the true reading and one of the multiple puncture skin tests as the comparison reading. Sensitivity, specificity, false-negative rate, and false-positive rates were determined. Note that (1 - sensitivity) equals false negative rate and that (1 - specificity) equals false-positive rate. In order to extract the greatest amount of information from this comparison, the procedure was
FIGURE 1. The Mantoux skin test was defined as the true reading and one of the multiple puncture skin tests as the comparison reading. Sensitivity, specificity, false-negative rate, and false-positive rates were determined for various cut-offs.

Repeated for each of several cut-off points for the multiple puncture skin test: 1, 2, and all even numbers to 24 mm. However, the cut-off for the Mantoux test was held fixed at 10 mm.

The receiver-operator characteristics of each of the multiple puncture skin tests were determined by comparing the true positive rate with the false-positive rate.

RESULTS

The frequency of various size reactions to each skin test is illustrated in Figure 2. Examination of the bar graph of reactions to the Mantoux test suggests that the population is nearly unimodal, and reaction sizes are distributed roughly normally with a slight excess in the number of reactors in the 2 to 5 mm range. The distribution of reaction sizes with the multiple puncture technique skin tests is quite different. The PPD tine, OT tine, and Aplitest clearly demonstrate a bimodal distribution with reactors 6 mm and greater with a normal distribution and those below 6 mm highly skewed to the left. Reactions to the Mono-Vacc are also skewed to the left, but lack the subpopulation of large reaction sizes found in the other four tests.

OPERATING CHARACTERISTICS

Using the $2 \times 2$ table presented in Figure 1, sensitivity and specificity of each of the multiple puncture skin tests in comparison with the Mantoux skin test were determined. Using a cut-off of 10 mm for the Mantoux test and 1 mm for the multiple puncture skin tests, sensitivity was 100 percent for PPD tine, 98 percent for Aplitest, 96 percent for OT tine, and 78 percent for Mono-Vacc. Regarding the PPD tine, at the 1 mm level, the false negative rate is extremely low (0), but the false positive rate is extremely high (82 percent). As the cut-off for the multiple puncture skin tests is set at higher levels of induration, the sensitivity decreases.

FIGURE 2. The frequency of various size reactions to each skin test are illustrated.
The drop is particularly precipitous with the Mono-Vacc; less so with the Aplitest and OT tine, and minimal with the PPD tine (Fig 3).

The specificity of the multiple puncture skin test in comparison with Mantoux test is quite poor at the 1 mm cut-off. It is best with Mono-Vacc (75 percent); next with Aplitest (49 percent); next OT tine (40 percent); and last PPD tine (18 percent). In the multiple puncture skin tests, as one defines greater amounts of induration for the cut-off, specificity increases. At the

**Figure 3.** Sensitivity for each multiple puncture skin test is plotted against the test cut-off. Sensitivity is high at low cut-offs and low at high cut-offs.

**Figure 4.** Specificity for each multiple puncture skin test is plotted against the test cut-off.
The sensitivity (true-positive rate) is plotted against 1-specificity (false-positive rate) for each multiple puncture skin test. In each case, the first point on the right is the value for a 1 mm cut-off. Moving to the left, the next value is for 2.4, 5, 8, 10... 24 mm. The most favorable location is the upper left portion of the graph. In this area, one has the highest true-positive rate and the lowest false-positive rate.

**Receiver-Operator Characteristics Analysis**

A graph can be constructed that correlates true positive rates with false-positive rates (sensitivity and 1-specificity, respectively) for a series of cut-off points. Figure 5 illustrates such a receiver-operator characteristics analysis. Four points are plotted comparing each of the multiple puncture skin tests with the Mantoux test. For the PPD tine, sensitivity was 100 percent at 1 mm and drops very little as one raises the cut-off. Concomitantly, the false-positive rate drops from very high levels at 1 mm to 14 percent at 10 mm cut-off. Similar characteristics are displayed by all the multiple puncture skin tests. As one raises the cut-off from 1 mm to 10 mm, sensitivity decreases as does the false-positive rate.

**Discussion**

The fact that reaction sizes by Mantoux test in this population are unimodal with only a slight increase in the number of reactors below 5 mm suggests that reaction is primarily to *Mycobacterium tuberculosis* with little cross-reaction to other mycobacteria, including BCG.

The distribution of reaction sizes to the dry multiple puncture skin tests (OT tine, PPD tine, Aplitest) suggests that the population of reactors actually consists of two subpopulations. The division of the two populations is not 2 mm as suggested by the package insert, but instead is 6 mm. The bar graphs of the 6 mm reactions show a marked skewing to the left (Fig 2). This probably reflects the high concentration of antigen applied to a very small volume of tissue. Reactions larger than 6 mm were normally distributed. This observation suggests that an analysis of larger reactions to the multiple puncture skin tests compared to the Mantoux test might yield different results than the comparisons previously reported. The Southeast Asian population was an excellent study population for the purpose of this study. However, we do not recommend MPT as a screening tool for this group. The Mantoux test remains the preferred skin test, in screening as well as diagnostic situations. However, when an MPT has been performed, very useful data should not be ignored. The most recent ATS statement recommends against using MPT in populations likely to manifest a positive tuberculin skin test. The role for MPT is not yet defined.

As one might expect, when a 1 mm cut-off for the multiple puncture skin tests is used, it is very sensitive but not very specific. As the cut-off is moved higher, sensitivity drops, but specificity increases. It is remarkable to note that the sensitivity of the PPD tine
drops modestly from 100 percent at 1 mm, to 99 percent at 2 mm, to 96 percent at 6 mm, and 85 percent at 10 mm. Specificity, however, increases dramatically from 15 percent at 1 mm to 92 percent at 12 mm. Of course, a screening test should be 100 percent sensitive. The consideration of the actual size of skin test reactions, rather than noting only positive/negative, allows the incorporation of a great deal of specificity into this "screening" test. Examination of the receiver-operator characteristics over a range of cut-offs provides a more complete picture of the relationship between the multiple puncture skin test reactivity and the reactivity to the Mantoux test.

Ideally, reactivity to all these tests should be compared with an independent, 100 percent specific and sensitive assessment of which subjects are infected. Unfortunately, such a "gold-standard" does not exist. For many years, we have assumed that the Mantoux skin tests provided us with a first approximation of who is infected vs who is not infected. Unfortunately, twice in the past ten years, the reliability of the commercially-available tuberculin for intradermal injection has been called into question. Adherence of Aplisol protein to glass caused a decrease in available antigen. Recently, problems with false-positive reactions to Aplisol administered by Mantoux skin test technique have been noted. The MPT requires meticulous attention to detail of antigen application to the tine and to the skin to obtain dependable results. These occurrences remind us that the skin test is a biological test, and as such, subject to a wide variety of confounding circumstances. One must keep open the possibility that one MPT might detect some infected persons who are not identified by the Mantoux test.

CONCLUSIONS

In most situations, an arbitrary cut-off point does a disservice to a quantitative assay. The reaction to multiple puncture skin test should be interpreted with the same careful consideration that clinicians use to interpret the reaction to the Mantoux test. Specifically, results should be recorded in millimeters of induration and interpretation based on the same factors that influence interpretation of the Mantoux test. One can then use a cut-off selected in accordance with the needs of the particular study; one can choose the benefits of a high degree of sensitivity at the low cut-off or the increasing specificity at high cut-offs. For example, if one is selecting individuals for a more definitive test such as PPD Mantoux test or chest x-ray film, one would want the greatest sensitivity and would be quite willing to sacrifice some specificity. If one is selecting individuals who will be considered for chemoprophylaxis, then one needs the greatest degree of specificity and would demand a much larger skin test reaction. In certain situations, the clinician may wish to maximize information when it is presented. The ATS statement regarding the interpretation of MPT may be too restrictive. The techniques described in this report provide a sound basis for such analysis.

ACKNOWLEDGMENTS: I wish to thank Ellen Logue, Sompi Phommay, Rhonda Gustofson, and Barry Peters for the skin testing; Robert Moser for administrative support; Quyen Cao, Andrew Catanzaro, Brian Catanzaro, and Leslie Schumacher for computer programming calculations and graphics; Sharon Batcher for editorial assistance; and Barbra Blair for preparation of the manuscript.

REFERENCES

7 Griner PF, Mayewski RJ, Mushlin AI, Greenberg P. Selection and interpretation of diagnostic tests and procedures. Ann Intern Med 94:571-85