Influence of the Prevalence of Infection on Tuberculin Skin Testing Programs

JAMES F. JEKEL, M.D., M.P.H., RICHARD A. GREENBERG, Ph.D., M.P.H., and BENJAMIN M. DRAKE, M.D., M.P.H.

THE Surgeon General's Task Force on Tuberculosis Control in 1963 recommended a "child-centered program" to control tuberculosis. This plan calls for skin testing of children at least twice during their school experience, once when they enter school and once in junior high school (*1*, *2*). Such a program, if applied throughout the United States, would require at least 7 million skin tests per year, tests which would have to be performed by community agencies such as health departments or schools.

When yearly screening of this magnitude is to be accomplished, the screening test must have several characteristics in addition to accuracy. It must be quick, inexpensive, and easy to administer and be acceptable to the health professionals and to the public, especially children. Unless a test has these operating characteristics, overworked public health personnel simply will not use it often. The multiple puncture tuberculin tests, such as the tine, Heaf, and Mono-Vac, have advantages over the intradermal Mantoux test in most of these operating characteristics. Largely because of widespread doubts about their accuracy, however, the multiple puncture tests have not gained the popularity that might be expected on the basis of these advantages. The doubts about these tests have arisen chiefly as a result of published studies comparing them with the intradermal Mantoux test (*3–7*) and from certain experiences with them in school tuberculin testing programs.

Dr. Jekel is assistant professor of public health and Dr. Greenberg is associate professor of biometry at the Yale University School of Medicine, New Haven, Conn. Dr. Drake is the health director of Gaston County, N.C.

Comparative Studies

The intradermal Mantoux test, in which 5 tuberculin units of purified protein derivative (PPD) are used, has long been considered the standard of comparison, and the rates of error for multiple puncture tests are usually computed on the basis of disagreement of their results with those of the Mantoux test. In such studies the authors usually, although it is not always so stated, assume zero rates of error for the Mantoux test. Greenberg and Jekel, in addition to questioning the zero error rates that are usually assumed for the Mantoux test in comparative studies, have shown that an incorrect assumption that the rates of error for one test are zero will result in maximum estimates of the error rates for the other test and hence lead to unjustified pessimism about its accuracy (*8*).

The recommended dose for Mantoux tuberculin tests is 5 tuberculin units of PPD in 0.1 ml. of solution, which is relatively dilute compared with the dose of 250 tuberculin units per 0.1

Another View of Tuberculin Testing

The authors very properly point out the statistical basis for estimation of the prevalence of testing error and also very properly reemphasize that the multiple puncture tests are screening tests. They assume, however, that the advantage of the multiple puncture test justifies the principle of using screening tests.

I question the validity of this assumption for the following reasons:

1. Our experience has been that the needle Mantoux test is just about as acceptable to children as a screening test.

2. The actual cost of the Tine and the Mono-Vac is higher than that of the Mantoux, assuming the use of a disposable multiple dose syringe with individual disposable needles.

3. A well-trained team can do a Mantoux test as rapidly as the Heaf or the Stern test. The small amount of time saved by using the Mono-Vac or the Tine test is more than offset by the need for another visit for retesting. We have found that in the long run a considerable amount of time, money, and effort is saved by doing the Mantoux test alone.

4. The authors do not mention that the Tine test is painful as compared with the other tests.

In short, while the statistical analysis is sound and well worth stressing, I do not believe that it is sufficient in itself to justify the use of a screening test in view of the many other factors to be considered.—M. Stuart Lauder, M.D., *Director, Tuberculosis Control Program, Kentucky State Department of Health.*

ml. used in the second strength test, a test which has also been used in the past (*9*, *10*). The relatively low dose of antigen increases the specificity of the test and leads to fewer false positives, but by the same token it would be expected to decrease the sensitivity somewhat, thus increasing the rate of false negative tests. This expected increase in false negative tests from a more specific test may be very small, and in the currently recommended Mantoux test, it probably is. Greenberg and Jekel showed that one kind of disagreement between the two tests would contribute to the false negative error of the Mantoux test or to the false positive error of the multiple puncture test, but not to both. If there had been any appreciable false negative error in the PPD Mantoux test in the comparative studies, even one as small as 0.5 to 1 percent, most of the published studies would have demonstrated false positive errors of much less than 10 percent for the multiple puncture tests. Bearman and others have pointed out that there are many possible sources of error in the intradermal Mantoux test (*11*).

Prevalence of Infection

The main concern of local public health workers about multiple puncture tests probably originates from their observations while using them in the tuberculin testing of school children. In most schools in this country, the prevalence of tuberculosis infection is low, less than 5 percent. In fact, in most first grade classes, the prevalence of infection is less than 2 percent. We have observed or heard of numerous instances in which local health workers have found, when testing school children with multiple puncture tests (particularly with the tuberculin tine test), that most of the positive reactors were negative when retested with PPD Mantoux. The workers have then reached what seemed to them to be a reasonable conclusion, namely, that the tuberculin tine test is not very useful because of the high proportion of false positive tests. In fact, however, neither the accuracy nor the usefulness of a screening test can be determined with certainty from an examination of the proportion of test positives which have been correctly identified. The reason is that not only the false positive error of the screening test, but also the low prevalence of tuberculosis infection, is responsible for the results observed.

False Positive and False Negative Rates

In any discussion of the false positive and false negative errors of a test, the numerator and denominators must be clearly defined. In a true rate, the denominator consists of only those persons who are at risk of being in the numerator. Thus, only those persons who are actually not infected with tuberculosis are at risk of being falsely identified as positive. The proper equation for the false positive error is therefore as follows:

False positive error (alpha) =

$$\frac{\text{number of false positives}}{\text{number of true negatives.}}$$

By similar reasoning, the equation for the false negative error rate becomes:

False negative error (beta) =

$$\frac{\text{number of false negatives}}{\text{number of true positives.}}$$

Therefore, what is most obvious to the public health workers—the ratio of false positives to the total number of tuberculin reactors—is not the rate of false positive error, but rather the effect of that rate combined with the prevalance of tuberculosis in the population being skin tested.

Influence of Prevalence

The prevalence of tuberculosis infection is not always considered when the results of a skin testing program are being evaluated. Yet a clear understanding of what is going on depends on consideration of the effect of prevalence. The practical effect of the prevalence of a condition upon the result of screening can be seen in the table. Here a false positive error of 6 percent and a false negative error of 4 percent are assumed for the screening test. (These error rates, incidentally, seem reasonable for the tuberculin tine test in the light of what has been found in the better comparative studies.) The table shows the results of applying these error rates to several populations of 100 persons each with a different proportion of persons infected with tuberculosis. As the prevalence of infection drops from 50 to 1 percent, the number of reactors correctly identified would, on the average, drop from 48 of 100 tests to one of 100 tests. At the same time the number of false positives found would increase, on the average, from three of 100 tests to the maximum of six of 100. The actual false positive error does not exceed the error rate we assumed in creating the table, but the proportion of reactors who are actually false positives has increased from 6 to 86 percent. It is not surprising that many health officers have become discouraged when using a screening test in populations with low prevalence. Moreover, it should be emphasized that the same principle would apply when performing any screening test (chest X-ray, diabetes screening, and so forth) in populations with a low prevalence of the condition under study.

Implications for Tuberculosis Control

What are the implications of the decreasing prevalence of tuberculosis for tuberculin testing? Do the multiple puncture tests have any value? The answer, in our opinion, is a definite affirmative. In the table, the positives, as a proportion of tests performed, would not on the average exceed the false positive error rate no matter how high the ratio of false positives to test positives. In effect, what the screening test does is to identify that 5 to 10 percent or so of the population among whom all, or nearly all, of the infected persons will be found. This information enables the health officer to apply his more specific but less acceptable PPD Mantoux test to only this small subset of his study population. A screening test—and this is how the multiple puncture tests should be regarded—always requires a followup diagnostic test

Variation in the proportion of test positives who are true positives as the prevalence of a test condition in the population changes [1]

Percent of prevalence	Expected true positives per 100 tests	Expected false positives per 100 tests	Expected percent of true positives among all test positives	Expected percent of false positives among all test positives
50	48	3	94 (48 of 51)	6 (3 of 51)
25	24	5	83 (24 of 29)	17 (5 of 29)
10	10	5	67 (10 of 15)	33 (5 of 15)
5	5	6	45 (5 of 11)	55 (6 of 11)
3	3	6	33 (3 of 9)	67 (6 of 9)
1	1	6	14 (1 of 7)	86 (6 of 7)

[1] Assuming that the false positive error is 6 percent (number of false positives ×100 ÷ number of true negatives) and that the false negative error is 4 percent (number of false negatives ×100 ÷ number of true positives).

before a diagnosis can be made, but the screening test is still invaluable if it makes the case-finding feasible in the first place.

An actual example of the value of the multiple puncture test can be seen in the following tuberculin testing results from Gaston County, N.C., where Drake is health officer.

Tine tests read	4, 963	(100 percent)
Tine reactors	239	(4.8 percent)
PPD reactors	55	(1.1 percent)
PPD reactors among tine reactors	$\frac{55\times 100}{239}$	(23 percent)

$$\frac{\text{Number of false positives} \times 100}{\text{number of PPD nonreactors}}$$

$$=\text{estimated maximum false positive rate}$$

$$=\frac{(239-55)\times 100}{4{,}963-55}=\frac{18{,}400}{4.908}=3.7 \text{ percent.}$$

Only about 25 percent of the persons with positive reactions in the Lederle Company's tuberculin tine test also reacted to 5 tuberculin units of PPD in the intradermal Mantoux test. These results can be satisfactorily explained in a population with a tuberculosis prevalence of 1.1 percent by assuming a rate of false positive error for the tine test of 3.7 percent. The tine test was used because of its operating advantages—speed combined with acceptability to both the public health nurses and the school children. In all, less than 5 percent of the tine reactors needed to be retested with PPD, and only 23 percent of those retested needed further followup with chest X-rays. We believe that this combined approach was quicker and easier than using the Mantoux test on 5,000 children. A simple cost-benefit study showed that the total cost of this tuberculin testing program, including overhead and travel, was approximately 50 cents per child tested, or approximately 8 cents per child in the school system. Use of the tine test followed by Mantoux tests on the tine reactors proved acceptable to nurses, the school system (including administrators), children, and parents.

Thus, it was possible to establish the screening phase of an ongoing, child-centered program of tuberculosis control for approximately 8 cents per school child per year.

REFERENCES

(*1*) U.S. Public Health Service: The future of tuberculosis control. PHS Publication No. 1119. U.S. Government Printing Office, Washington, D.C., 1963.

(*2*) U.S. Public Health Service: A child-centered program to prevent tuberculosis. PHS Publication No. 1280. U.S. Government Printing Office, Washington, D.C., 1965.

(*3*) Affronti, L., Parlett, R. C., Pierson, F., and Anello, C.: An epidemiologic comparative study in Delaware of the tine and Mantoux tests. Amer Rev Resp Dis 95: 81–88, January 1967.

(*4*) Badger, T. L., Breitwieser, E. R., and Muench, H.: Tuberculin tine test. Amer Rev Resp Dis 87: 338–353, March 1963.

(*5*) Capobres, D. B., Tosh, F. E., Yates, J. L., and Langeluttig, H. V.: Experience with the tuberculin tine test in a sanatorium. JAMA 180: 1130–1132, June 30, 1962.

(*6*) Maha, G. E.: Comparative study of tuberculin tine and Mantoux tests in 676 college students. JAMA 182: 304–305, Oct. 20, 1962.

(*7*) Maxwell, K. W., Dietz, T., Marcus, S., and Hill, G. A.: Comparative sensitivity of tuberculin tine test and Mantoux test in guinea pigs. Amer Rev Resp Dis 89: 926–928, June 1964.

(*8*) Greenberg, R. A., and Jekel, J. F.: Some problems in the determination of the false positive and false negative rates of tuberculin tests. Amer Rev Resp Dis. In press.

(*9*) Badger, T. L., et al.: Tuberculin skin-testing techniques: Current status. Statement of the committee on diagnostic skin testing. Amer Rev Resp Dis 87: 607–610, April 1963.

(*10*) Hsu, K. H. K., Carreon, A. T., Jeu, F., and Jenkins, D. E.: Today's concept of the tuberculin test. Dis Chest 46: 648–664, December 1964.

(*11*) Bearman, J. E.: A study of variability in tuberculin test reading. Amer Rev Resp Dis 90: 913–919 (1964).

Tearsheet Requests

Dr. James F. Jekel, Yale University School of Medicine, 60 College St., New Haven, Conn. 06510