of college, did a far better job in interpreting and communicating between patients and staff. The need for this type of personnel was identified among other groups in which foreign language was not the barrier, such as the blacks, the elderly, and the skid row alcoholics.

As a result, a special classification, community health worker, was developed so that this support staff could be hired for the various medical programs and clinics. Thus, the community health workers in the senior housing projects are well-motivated, alert, and active individuals 60 years of age and older; while those in the skid row area are “dried-out” alcoholics who live in the neighborhood.

The use of community health workers has greatly improved patient acceptability of clinic services. However, it has been found that these workers must be carefully selected and then well trained regarding the resources of the community and the department. Furthermore, there must be an ongoing dialogue between the community health workers and the professional staff if their effectiveness is to continue. In the opinion of both the patients and the staff, the community health workers have been the most important recent addition to the medical programs.

CONCLUSION

In conclusion, it should be emphasized that a well-planned medical program will satisfactorily serve from 65 to 75 percent of the target population, with modifications necessary for the remaining 25 to 35 percent; 15 to 25 percent of patients will require program modifications to meet social, cultural, or ethnic needs, and approximately 10 percent of patients will require program modifications to meet their specific problems. Among the many problems encountered in the development, implementation, and administration of community medical programs, the major difficulties arose from cultural barriers, rather than the generally designated language barriers. Thus, the use of bicultural personnel improved program utilization and acceptability.

When a new major medical program was planned and implemented, or where a marked revision or change was made in an existing program, three separate categories of staff reactions were found to occur. First, some staff welcomed the changes as a means to better serve patients and the community. Second, certain staff members resisted major revisions through fear of change and personal insecurity. Third, a small number of staff were openly hostile and antagonistic to any change because of deep-rooted prejudices against the recipients. These prejudices were frequently ethnic, racial, cultural, or religious; but could also be for lifestyle (“hippie” or skid row drunk), political affiliation (“commie,” “liberal,” or “Bircher”), social strata (“bum” or “chiseler”), or any other of a dozen or more reasons. This group felt that “the patients were already receiving far more than they deserved, and to do more or to improve services was a waste of time, effort, and money.” The hostility and antagonism of this group was so intense that it was found necessary to remove them from the program if the new medical services were to have a chance to succeed.

A successful medical program insures job satisfaction and job security for staff, while at the same time providing better services for patients. A successful medical program depends upon patient cooperation, which, in turn, depends upon program acceptability to patients. Staff must be repeatedly reminded that program availability and accessibility is not the same as program acceptability to patients. Furthermore, anything that will improve program acceptability will favorably influence program success. In San Francisco, it was found that friendly and cooperative staff attitudes combined with the use of carefully selected community health workers were the two most important factors favorably influencing program acceptability by patients.

False Tuberculin Test Results*

George W. Comstock, M.D.

For a meaningful discussion of false tuberculin test results, two prerequisites are needed. The first is to specify the purpose of a tuberculin test.

*From the Department of Epidemiology, School of Hygiene and Public Health, the Johns Hopkins University, Baltimore. Supported in part by research career award HL 21,670 from the National Heart and Lung Institute. Reprint requests: Dr. Comstock, P.O. Box 2067, Hagerstown, Maryland 21740.

Most of us would undoubtedly agree that its fundamental purpose is the identification of individuals who have been infected with mammalian tubercle bacilli. If this definition is accepted, the specification of the four types of false reactions is straightforward. A false-negative tuberculin test is one which fails to identify an individual who has been infected with
tubercle bacilli, and a false-positive tuberculin test erroneously classifies an uninfected individual as infected. If serial tuberculin testing is being carried out, a false-negative test on the first occasion or a false-positive test on the second can lead to a false conversion from negative to positive status and, thereby, suggest a recent infection, when in fact no infection has occurred. False reversions from positive to negative can also occur, but because positive reactors are rarely retested, reversions have little practical importance and will not be discussed further.

The second prerequisite for a meaningful discussion is the recognition that the tuberculin test is a quantitative test. A tremendous range of sensitivity to tuberculin can be demonstrated by the use of different doses of tuberculin and by measuring the resultant induration. In most of the United States, there is little more sense in labelling a tuberculin reaction "positive" or "negative" than in referring to a "positive" or "negative" result of a blood pressure or serum cholesterol determination. Only in a few places are the tuberculin reactions among infected persons distinctly larger than those among uninfected persons; and only in these few regions, mostly in the far Northwest, are the terms, "positive" and "negative," easily defensible. Nevertheless, tradition has long sanctioned the use of the terms, "positive" and "negative," with respect to tuberculin reactions, and I shall continue to use them in this discussion.

CAUSES OF FALSE RESULTS

There are a number of potential causes of false tuberculin reactions that I shall mention only in passing. I shall not consider any dose of tuberculin except the 5 tuberculin-unit (TU) dose, for it has long been known that weaker doses increase the likelihood of false-negative results and that stronger doses increase the likelihood of false-positive results. Methods other than the Mantoux technique of intracutaneous injection by needle and syringe deliver inaccurate doses of tuberculin and, thereby cause obvious problems with false reactions. A further limitation is that, unless otherwise specified, tuberculin in this discussion will refer only to purified protein derivative (PPD) from large batches of well-standardized material. As far as I know, only the preparations distributed by the Center for Disease Control for research purposes and the commercial products manufactured by Connaught Laboratories and by Parke-Davis and Co. meet this fundamental criterion of an acceptable test material.

With these provisos and limitations out of the way, let us return to a consideration of false tuberculin test results. These can arise from four sources: faulty techniques in administering the test; careless measurement of the ensuing reaction; incorrect interpretation of the results; or failure of the tested subject to react in the usual fashion.

Faulty Techniques of Administration

Most of the ordinary mistakes in the technique of administering a tuberculin test lead to smaller reactions and, thus, increase the proportion of false-negative results. The tendency of tuberculin to adhere to glass or plastic surfaces caused serious problems for many years. For example, using the ordinary phosphate-buffered diluent, it was shown that a 5-TU dose was reduced to 3 TU after standing in a syringe for one hour. After eight hours, the dose was reduced to 14 TU and, after 24 hours, to less than one-half TU. At the present time, both the Connaught and Parke-Davis products contain 5 ppm of a detergent, polysorbate 80 (Tween 80), which greatly reduces adsorption. Nevertheless, even with these products, tuberculin should never be transferred from one container to another. Filled syringes should be used as quickly as possible, although with polysorbate 80 in the diluent, the loss in the first hour is probably not important.

Tuberculin is much more sensitive to the effects of light than heat. In fact, the rationale for storing it in the refrigerator is based more on our faith that the light really does go out when the door is closed than on fears concerning the effects of ordinary room temperatures. In testing programs the stock bottles should be protected from light, and filled syringes should be kept covered or shaded until they are used.

Leaky syringes are an obvious cause of reduced dosage and should be discarded at once. Fortunately, plastic disposable syringes practically never leak. If leakage does occur, the injected dose should not be estimated by the size of the wheal. Wheal size depends on a number of factors and is only a very crude estimate of the amount injected. Discard the leaky syringe, and repeat the test in another site several centimeters away from the first attempt.

Injection of tuberculin into the deeper layers of the skin may cause a reaction even if no wheal is produced. However, because the borders of any resulting reaction will be invisible and difficult to palpate, gross errors in measurement can occur. If deep injection has occurred, it is wise to repeat the injection in another area.

Careless Measurement of Reaction

Carelessness in measuring a tuberculin reaction is another obvious cause of misclassification. I have
known physicians to classify tuberculin reactions merely by glancing at them, and even a tuberculosis specialist who measured reactions to the nearest centimeter. Many persons do not realize that induration can usually be seen if lighting and vision are adequate. Obviously, visible induration can be more accurately measured than induration that is identified only by palpation.

Appropriate instruments for measuring induration are hard to come by. An ordinary stiff ruler cannot be used on vesicular reactions, and even a thin flexible one covers half the lesion being measured. Some sort of a caliper is much more convenient to use, an ideal one being the plastic sliding type distributed by the Oregon Thoracic Society, especially if a small piece of sand paper is glued to the sliding component to act as a thumb grip. If you cannot obtain one of these calipers, an ordinary sewing gauge may be adapted for this purpose by gluing a millimeter rule to it, an invention of Dr. R. B. Turnbull of Memphis.

Unfortunately, the ability of a reader to measure tuberculin tests cannot be accurately assessed, because there is no standard against which to judge results. However, a clue may be obtained by looking at the distribution of terminal digits of an individual's readings. Given the decimal orientation of our counting system, most of us tend to record readings that end in 0 or in 5 whenever we are uncertain of the true measurements. Some individuals may prefer other terminal digits. The important point is that an unusually high proportion of readings ending in any one or two digits is a very good indicator that the reader is having difficulties, either because of poor vision, poor working conditions, or poor understanding of the need for careful work.

Misinterpretation of Results

Having obtained accurate measurements of tuberculin reactions in subjects whose ability to react is unimpaired, it is still possible to misinterpret these findings. For practical purposes, errors of interpretation are most likely to increase the numbers of false-positive test results. Among persons known to be infected with tubercle bacilli, the most common reactions to the 5-TU dose of PPD will be between 15 and 20 mm in diameter, with very few being under 5 mm or over 30 mm.4 Cross-reactions to tuberculin resulting from infections with other mycobacteria tend to be smaller than the homologous reactions caused by tuberculous infection, but the overlap between cross-reactions and homologous reactions is considerable. As a rough rule of thumb, it may be considered that reactions larger than 15 mm in diameter are caused by tuberculosis. Below this size, the probability that the reactions are false-positives increases as the reactions become smaller.5

If cross-reactions from other mycobacterial infections are present in the population, as they are in every state with the possible exception of Alaska, there will be no dividing line which will clearly separate tuberculous from nontuberculous infections. If the lower limit for a positive reaction is set very high, the positive reactors may all be infected with tubercle bacilli, but a very high proportion of persons with tuberculous infections will be found among the negative reactions. If the lower limit for a positive reaction is set very low, most truly infected persons will be classified as positive reactors, but so will large numbers of persons infected with other mycobacteria. Setting the lower limit too high increases false-negative reactions; setting it too low increases false-positive reactions.

Because any reasonable definition of positive will falsely classify some persons as negative and others as positive, the problem is to minimize the overall consequences of misclassification in both directions. To do this involves not only balancing false-negative and false-positive reactions, but also deciding on the proper weight for each type of error, since it may sometimes be worse to be a false-negative than a false-positive, or vice versa.6 And because such decisions require considerable detailed information about tuberculin test results, most of us are content to accept the dictum of a committee which sat briefly in one of our holy places and pronounced that induration with a diameter of 10 mm or more to 5 TU of PPD should be called positive.5

Although this arbitrary definition has worked surprisingly well for most of us, an ideal application of principles would require some deviations from this standard. In populations with a high frequency of tuberculous infections relative to other mycobacterial infections, the criterion of positivity should be shifted downward. Such populations are found geographically in the northwestern United States, demographically among the old and among immigrants from high prevalence countries, and epidemiologically among suspected cases and household contacts. The criterion for positivity should also be shifted to minimize the cost-benefit ratios for any procedures to which positive reactors may be subjected. The committee's recommendations took account of only one of these possible deviations by recommending that the criterion for a positive tuberculin reaction among suspected cases and household contacts be reduced from 10 to 5 mm of induration.5

CHEST, 68: 3, SEPTEMBER, 1975 SUPPLEMENT

FALSE TUBERCULIN TEST RESULTS 467
Unusual Test Reactions

Finally, a tuberculin test may not reflect the correct infection status, because the subject fails to react in the expected fashion. For practical purposes, all such failures result in infected persons being falsely classified as negative reactors. The first condition to be considered is early infection. It is generally accepted that tuberculin sensitivity does not develop until two to ten weeks after infection has occurred. During this time, of course, a tuberculin test will not reveal the presence of infection; and one must wait, therefore, for approximately ten weeks after contact is broken before it can reasonably be asserted that a household associate has not been infected.

Anergy is an important and largely uncontrollable cause of false-negative tuberculin reactions. From the clinical point of view, loss of tuberculin sensitivity from anergy is most troublesome in persons with severe life-threatening forms of tuberculosis. This is not so important in patients with overwhelming pulmonary disease, for in these persons a direct smear of the sputum will almost always establish the diagnosis quickly. More distressing is the fact that approximately 10 percent of children with miliary or meningeal tuberculosis do not react to tuberculin at the time they come to medical attention.

Measles and possibly other acute infectious diseases have also been reported to suppress tuberculin sensitivity. My own experience with measles in the occupants of an Alaskan boarding school suggests that this suppression persists for about three weeks after onset.

Administration of immunosuppressant drugs, notably the corticosteroids, may also cause false-negative tuberculin reactions, an important consideration if isoniazid prophylaxis is being considered for persons receiving antitumor chemotherapy. In passing, I might mention that, in spite of widespread general belief, the local application of hydrocortisone ointment does no more to relieve the pain and swelling of a severe tuberculin reaction than any ointment base.

Booster Effect

The final cause of false tuberculin test results to be considered is the so-called booster effect, an analog of the anamnestic reaction in serology. It has long been known that tuberculin sensitivity wanes in some individuals. Every tuberculin testing survey of a general population of which I am aware shows a falling off of the prevalence of positive reactors in the older age groups, and this falling off is greater than can be accounted for by any increased mortality among positive reactors. It has also been known for a long time that, although tuberculin testing with 5 TU of PPD is not sensitizing, repeated tests with this dosage will prevent waning of tuberculin sensitivity after BCG vaccination.

One of the most illuminating studies of the booster effect was done by Ferebee (now Woolpert) and Mount during the course of a controlled trial of isoniazid prophylaxis in mental institutions. Unfortunately, the results of the study by Ferebee and Mount were only published in abstract form, but I have been given permission to review the highlights of their study.

The protocol for the study of isoniazid prophylaxis is required that all participants be tested with tuberculin prior to entry into the trial and again one year later. Early analysis of the year-end results were disconcerting, to say the least. In the three groups of mental hospitals, the apparent conversion rates ranged from 16 to 25 percent (Table 1). These rates seemed too high to be believable in view of the facts that all of these institutions had good tuberculosis control programs and that not a single case of tuberculosis had developed among negative reactors during the year. Among bits of evidence that indicated that nothing was wrong with the tuberculin testing procedures was the fact that the schools for the mentally retarded in Massachusetts showed a low conversion rate, one entirely consistent with the known slight deviations seen in repeated tuberculin tests.

To elucidate this unexpected phenomenon, patients in the next institution to participate in the trial of isoniazid prophylaxis were randomly allocated to two groups, and a double-blind study was set up. Patients in group A were initially tested with 5 TU of PPD; 44 percent of them were classified as reactors (induration larger than 8 mm). Those in group B were initially tested with diluent, and none was classified as a reactor. One month later, both groups were tested with 5 TU of PPD. At this time, 71 percent of group A were reactors, an apparent conversion rate of 47 percent in one month! However,
the control group B showed essentially the same frequency of reactors (45 percent) on their first tuberculin test as group A had on theirs. The only reasonable conclusion was that the stimulus of the first tuberculin test had somehow caused a number of persons in group A to become positive reactors.

A second study was then set up in a prison population in Illinois. Prisoners who volunteered to be tested were randomly divided into five groups. All were initially tested with 5 TU of PPD. Group 1 was retested with the same dose one week later; group 2, two weeks later; and groups 3, 4, and 5 at three, four, and five weeks later, respectively. All groups showed virtually identical increases in tuberculin sensitivity on the second test, showing that the booster phenomenon could be manifested as early as one week after the initial test. Some groups were also tested a third time, with results that were almost the same as those found on the second test. This finding was interpreted as showing that the increase in sensitivity was caused only by the first test and was not increased by subsequent testing.

The major characteristics of the booster phenomenon may be summarized as follows. It has also been observed in studies of histoplasmin sensitivity and presumably can occur in any form of delayed-type hypersensitivity. It is most marked in the age group of 50 to 75 years and among persons without demonstrable disease. It can become manifest within one week and lasts for at least one year. Persons with small reactions are most likely to be affected, but increases in reaction size that anyone would call conversions are often seen. Only the initial test causes an appreciable increase in sensitivity.

How do we know that the booster phenomenon is not caused by the sensitizing effects of skin tests? The evidence against this explanation is extensive, but major pieces of it are the facts that the booster effect is infrequently seen in the very young or very old, and that it is not increased by repeated testing.

The current concept of the booster phenomenon is based on the waning of delayed-type sensitivity and its recall by the minor stimulus of a skin test. In some people, perhaps those infected with few bacilli many years ago and not subsequently exposed to tuberculin, waning of tuberculin sensitivity occurs, usually after a rather long period of time. Not only do their reactions become smaller, but the sensitivity of some will drop below the threshold at which it can be detected by the standard tuberculin test. If sensitivity has not dropped too far below this threshold, the very slight stimulus of a tuberculin test can cause an anamnestic response, so that a second test causes a reaction. In some other persons, particularly the very old, sensitivity may drop so low that a tuberculin test is an insufficient stimulus for recall. While not all aspects of this concept can be considered to be proven, it does have the virtue of fitting the currently known facts about the booster phenomenon.

Conclusion

Finally, let me reemphasize that the tuberculin test is a quantitative test and that it should be administered, measured, and interpreted with care if the maximum amount of information is to be derived from it. Compared to other diagnostic tests, it can accomplish its purpose remarkably well. However, at no point from its administration to its interpretation is it foolproof. The good clinician and the good epidemiologist will keep in mind both its virtues and its faults.

References

1 Bushnell GE: A Study in the Epidemiology of Tuberculosis, with Special Reference to Tuberculosis of the Tropics and of the Negro Race. New York, William Wood and Co, 1920, pp 96-98
7 American Thoracic Society Ad Hoc Committee to Revise Diagnostic Standards: Diagnostic Standards and Classification of Tuberculosis and Other Mycobacterial Diseases. New York, American Lung Association, 1974
9 Bentzon JW: The effect of certain infectious diseases on tuberculin allergy. Tubercle 34:34-41, 1953