Evolution of the tuberculin skin test and challenges in its interpretation

A N Shashidhara¹, V K Chadha ²

Introduction

Tuberculin skin test (TST) is routinely carried out as a diagnostic aid for TB by pediatricians' and clinicians all over the country. Compared to India perhaps no other country has carried out tuberculin testing in different communities, epidemiological and clinical sub groups. The number of tests administered may exceed 20-30,000 or more on any given day in India. It is also used extensively in epidemiological studies, the most prominent being the ARTI surveys. Tuberculin surveys are usually conducted among children to estimate prevalence of ARTI, which may vary in different epidemiological settings. Since the population is large, the sample included in ARTI surveys are indeed large. Therefore, it is very necessary to possess an in depth understanding of this test which is wide in use.

The TST has four major components: a) methods of administration b) the antigen c) aspects of measurement and evaluation of the reaction at test site d) interpretation and end uses of the results. Experienced users have understood both the advantages and limitations of these four components, as well as the pitfalls. A broad summary is as under:

Technique of administration

Clement von Pinquet was the first physician to put to use the almost forgotten Old Tuberculin (OT) discovered by Robert Koch in 1890 in a gentler way than Koch did, by devising the scratch test¹, ². During 1907-08, he conducted extensive studies of how OT could be effectively used and discovered that a tiny scratch through a little quantity of tuberculin is enough to elicit evidence of tuberculin sensitivity. He also introduced the terms allergy and anergy. He observed that the tuberculin test was almost always positive in TB patients. He further noted that measles could suppress tuberculin reactivity. Around the same time, Charles Mantoux introduced the concept of administering a measured dose to the desired depth of the skin by using a syringe. In the same year, Alberto Moro introduced the concept of non-invasiveness by developing a patch test. In the later years, other modes of administration (Tine test, Heaf test etc) were developed. Currently the most used method of administration of tuberculin skin test is the Mantoux method. This has advantages over the other methods because a recommended dose of the PPD dilution could be administered to the upper layer of the skin (hence called intradermal test) accurately by trained practitioners with a leak proof syringe fitted with a sharp, short length needle. The results could be read at 42 hours and interpreted to great benefit. This method has been improved from time to time and has currently achieved a high state of respectability.

The antigen

From the days of Koch who announced the discovery of tuberculin in a somewhat mysterious way as a thanapeutic agent in 1890, Old Tuberculin (OT) was being prepared by many. As variations between different preparations came to be noticed, OT was standardized somewhat in 1928 ( the soluble products of the culture and lysis of one or more strains of Mycobacterium tuberculosis grown in a liquid medium which may be glycerolated broth or a synthetic medium). As this product was not stable and contained many impurities, researchers

¹ Research Officer, ARTI surveys; ² Sr. Epidemiologist, National Tuberculosis Institute, Bangalore
began to look for a standard preparation (high specificity, potency and constancy; stability; samples made at different times should be of equal strength; it should not give rise to false positive reactions and it should not sensitize the individual). Florence Seibert and co-workers of the Henry Phipps Institute, USA, finally succeeded in their attempts to standardize a methodology of making purer, less denatured preparations of tuberculin (1934) which was designated purified protein derivative (PPD). They made two large batches for international use. One batch, designated as PPD-S has been recognized as the US standard. Since 1952, it is also referred to as the international standard. The standard first dose (5TU) is defined as the delayed skin test activity elicited by a 0.0001 mg per 0.01 ml dose of PPD-S. Seibert and associates also prepared another large batch of PPD Tuberculin for the WHO for international use. Designated as the PPD RT 23 batch, 1TU or the first or the low dose and contains 0.00002mg of PPD +0.000008mg of buffer salts + Tween 80 (diluents) and is preserved in the Statens Serum Institute, Copenhagen. Preparation, preservation and comparative assays on different PPD batches are difficult and highly complex. Later attempts made by several investigators in preparing better or purer antigen/s or sensitin/s have not yet borne fruit. Therefore, informed investigators have been using the above two preparations in all their studies.

Aspects of measurements and evaluation of the reaction at the test site

Currently, measurement of the transverse diameter in mms of the reaction (induration) that develops at the test site on the 3rd or 4th day is the standard practice. This method ensued from an earlier one, in which Caroll E Palmer had introduced a "4 category quantitative description of reaction density". Category I was used to describe a typical text book reaction with an area of induration which is firm, elevated, clearly defined, and well circumscribed. At the other end of the scale, Category IV was used to describe questionable induration which was very soft, ill defined and not well circumscribed. Category II and III were used to describe reactions which did not entirely fulfill the exacting conditions of either I or IV but fell somewhere between the two: Category II denoted a reaction showing greater similarity to that described as Category I, and Category III showing greater similarity to Category IV. However, training of readers to categorize reactions was difficult and was slowly given up. Pioneering work of C E Palmer and other workers in the 1940s and, by the WHO in the later years showed that the delayed type of responses at the test site following a low dose tuberculin test would be most prominent and most evident in most people between 2-5 days, and its measurement in mms would be the best indicator of tuberculin sensitivity. Studies conducted throughout the world have confirmed that this indicator when grouped together as frequency distributions yielded valuable information especially on the epidemiological classification content of the community. Information generated thus could be used by the clinicians to great benefit and be of assistance in diagnosis. For an example, in a particular epidemiological setting where tuberculin testing has been carried out, a cut off point to signify the infected and the non infected, say at 10, 12 or 15 mm, could be deliberated from the frequency distributions of tuberculin reactions plotted in the form of histogram. Even though such classifications may have some short comings, in the least, it would finger point at the group of persons with similar antigenic status. This is in conformity with the findings of researchers everywhere: the larger the size of the induration at the test site, the probability of infection with M. tuberculosis increases. The reverse probability that the infection is not due to M. tuberculosis
stands on a brittle ground especially in high prevalence countries like India and gets further confounded by the proliferation of Mycobacterium other than Tuberculosis (MOTT), which affect on the specificity on any antigen used for TST. However, neither the time, duration or quantification of infection is measured or implied.

Beginning from the early seventies, a lot of work has been done in the immunology of tuberculosis which has widened the understanding of the nature of the disease itself. Substantial progress has also been made in elucidating the physiological changes that occur in the skin at the test site and in the host in general. In tuberculosis, damages to the tissue are caused more by the host’s reaction or immunologic response to the bacilli than by the actions of the bacilli itself. Before tubercle bacilli are destroyed by macrophages, these cells must be activated by T-lymphocytes and their lymphokines.

As more factors that determine immune mechanisms underlying the pathogenesis of tuberculosis became better understood, it also became increasingly clear that the process of tuberculin sensitization result from a diverse spectrum of T-cell functions. Different functions might be attributed to separated T-Cell populations or to a multifunctional cell at a certain stage of maturation. Activation of T-cells and their lymphocytes are key factors which form the foundation of cell mediated immunity (CMI), of which the delayed type of hypersensitivity (DTH) is one component. Observers have noticed that although tuberculin sensitivity response is classically a DTH response, maximal at 48-72 hours, though some might elicit immediate response maximal at 6-8 hours. Such reactions are common in areas pervaded by MOTT, or as a consequence of frequent contact with patients with disease.

Thus, there is a prophylatic CMI and tissue damaging anaphylactic DTH. The later causes caseous necrosis (Tissue Necrosis Factor) of host tissues whenever the bacillary antigens reach high levels. A low level of antigens in a tuberculin positive host is beneficial because they stimulate the development of CMI, whereas high levels stimulate the development of DTH (Tumor necrosis factor-TNF). Susceptibility of the host is therefore directly related to the increased levels of DTH which causes caseation, tissue destruction and discharging of caseous material into the bronchiactic tree.

Though the interplay between the CMI and DTH are not readily quantifiable, yet, the development of induration at test site is usually taken as a measure of the DTH and TNF components than CMI and is related to CMI in general terms. Accordingly, a significant reaction means the individual is infected with M. tuberculosis. Experience of tuberculosis workers is analogous to the phenomenon that the homologous antigen-antibody reactions are stronger and more enduring than the heterologous ones.

In a single host according to whether he has a single or multiple exposure to the risk of infection, either Listeria or Koch type or both types of responses may be present. And, in contrast to homologous antigen-antibody reactions which are stronger, firmer and more intensely indurated, the heterologous reactions (ex: BCG induced visa-a-visa PPD-B antigen) are less firm and tend to fade sooner.

A comparison of tuberculin sensitivity patterns of TB patients with healthy individuals who were given BCG and later subjected to a tuberculin test will clearly illustrate these factors. Whereas most reactors at the test site of TB patients were denser, harder and long lasting, those of healthy individuals given BCG were less dense and would not last as long. A majority of reactions were softer and some were so soft that their edges were barely palpable. Indeed, a large controlled study has
shown that tuberculin sensitivity following BCG vaccination wanes steadily till it touches a low peak by about 2½ years, whereas the sensitivity induced by *M. tuberculosis* sensitivity does not show such a decline in such a short time\(^2^4\). Further, observers point out that in areas where prevalence of tuberculosis is high, tuberculin sensitivity hardly wanes and most adults continue to show significant reactions. In contrast, in areas where the prevalence of tuberculosis is very low waning is very substantial among adults. Tuberculin sensitivity most probably gets boosted by repeated exposures that bolster up recall memories and keep the re-activation processes of the T. lymphocytes going\(^2^1-^2^3\).

**Interpretation and end uses of tuberculin skin test results**

The aim of the TST is to determine tuberculin sensitivity, i.e., to classify those infected with *M. tuberculosis* as significant reactors and those not infected as not. The basis of the quantitative aspect is that a predetermined dose of tuberculin when injected intradermally to a sensitized host, will cause a reaction indicating DTH at the test site. A standard measurement of this reaction could be taken to represent tuberculin sensitivity and therefore could be used for meaningful interpretations. This, however, gets into difficulties as there are hosts of determining factors that come into play. Not all persons react in the same manner to any dose of tuberculin or antigen. Given any quantified definition, unambiguous classification may not be achieved in a group of persons or community as well. Researchers everywhere are beginning to understand that there are more complexities which are less understood than are dealt with under the present system of interpretations, analysis of data and how the end results are used. It, therefore becomes very important to consider the following factors before a person or a group of persons is/are subjected to the tuberculin test.

a) **Purpose:** Besides classifying an individual as infected or not, a TST, however accurately performed, may not amount to much of anything else. Neither the force, extent, quality or history of previous infection/s can be clearly made out. Except perhaps a recent episode of re-infection established on the basis of a well categorized and accepted definition of it.

   It is important for the user to be clear why he needs to order the test to a particular individual.

b) **Person being tested:** Age (new born, elderly), epidemiological setting (high or low prevalence area), drugs (especially anti-TB drugs, corticosteroids), recent vaccination status (especially BCG vaccination), nutritional status, over whelming stress etc.

c) **Antigen/ product:** dose (resorting to higher doses; BCG Test etc), storage, maintainance of cold chain (2-8°C) combination.

d) **Administration:** Trained/untrained person administering and interpreting the test, less understood problems with repeat test.

e) **Wastage of tuberculin:** In the current practice of one time use and throw disposable syringes, well experienced observers point out that only 8 or 9 tests can be given from a 20 doses vial. No more. All the rest goes as waste, the largest quantity as a filler to the barrel of the needle and sticking to the inside sidewalls of the barrel. More tuberculin is thus wasted than 'used'.

   This is highly noticeable because the quantity of tuberculin in a dose is very little: 0.1 ml.

**Some difficulties in the current methodology of Tuberculin Skin Testing**
1. Training inadequacies in both testing and reading: As has been observed during training of tuberculin testers and readers for research purposes (especially at the NTI, Bangalore and TRC, Chennai), however much may be the effort put in and on the job practice given, not all trainees respond adequately to training and become reasonably standard testers or readers. A mere 30% will turn out to be preferred in research studies and designated and used as standard tester or standard reader. All the rest are "trained" and can practice testing and reading elsewhere. This raises a host of questions on the credibility of testing and reading techniques adopted by the hundreds if not thousands of practitioners practicing testing and reading who are exposed to less rigorous training methodologies.

2. As to the evaluation and end uses of test results, it must be first understood that both epidemiologists and clinicians face ambiguities in deciding who is infected and who is not. Epidemiologists face difficulties in arriving at scientifically valid cut-off points and to estimate infection figures. They are turning out to be exercises in further statistical technicalities. Clinicians use a host of standard and non-standard tuberculin of different strengths, even BCG tuberculin, bringing in increased even less understood complexities.

3. Interpretation of tuberculin reactivity and classification of whether a particular subject is positive or negative is approached judiciously and is intent based. Whereas it could be generally agreed that as the induration levels tend to shift from the lower to the higher ranges (say, 15-20 mms or more) individuals found in higher ranges may be classified as tuberculinised or infected. Indeed, it is in this group the largest number of TB patients are routinely found. However, its reverse reasoning, individuals found in the lower ranges (< 14mms to 10mms or less) may not all be not tuberculinised and classified as non-infected may not turn out to be always true. This is because individuals react differently at different stages to tuberculin antigen and the summed up reaction ranges of individuals presented in frequency distributions are necessarily overlapping and continuous. Therefore, in a country like India, where both specific and nonspecific infections are highly prevalent, classification of infection, re-infection tends to be difficult. There are temporal factors like boosting and waning, low grade and high grade sensitivities to consider.

4. Both researchers in epidemiology and other users like the clinicians and pediatricians tend to play down an important fact that the lone measure representing tuberculin sensitivity, which is the maximum measurable transverse diameter in mms of the induration at the test site in a subject; and/or pooled indices teased out from different studies (one of which is ARTI), have underlying negotiated considerations. There may be pitfalls not yet sufficiently understood; one does well by proceeding with abundant caution and plan further studies on both direct and indirect determinants that affect the engineering of tuberculin sensitivity.

5. Our country also covers infants and young children with BCG vaccination under the UPI. Large proportions of BCG vaccinated children have been later found not to have lasting scars or durable tuberculin sensitivity. Problems arise in interpreting the tuberculin test results among such vaccinated children, especially when paediatricians use stronger doses of tuberculin, or the BCG test because both these confound and cloud the antigenic sensitivity.

6. In children's hospitals, child suspects of TBM who are already on intense medication will in addition be usually administered the tuberculin test. They will sometimes be getting additional shots of medication. The tuberculin reaction may indeed be suppressed in such conditions.

7. Currently, our country is vigorously
implementing RNTCP which impacts both the exposure and force of transmission of infection. This is an area of challenge and astute epidemiological minds may work on and come out with solutions. Among those with pulmonary symptoms of 3 weeks or more who are coming under TB purview. The childhood contacts of TB cases are a high, risk group who should be administered the TST and screened for eligibility to an efficient chemoprophylaxis programme to prevent, or control, or preempt breaking down into overt disease at that future time.

8. There are many other less understood factors influencing the transmission of infection. Some of these factors e.g., poverty, socio-economic status, migration, terrain and habitat have also their effects and call for accounting while defining and estimating the force of infection. It seems strange but as observed in repeat surveys, TB cases may not, and were not seen to emerge from TB case households; and that not all children in TB case households might become tuberculinsed!

9. WHO recommends, 2TU of PPD RT23 with tween 80 as a standard dose compared to 1TU employed earlier. The appropriateness of the increased dose has though not been established beyond doubt. Besides, there may be many more explorable avenues which widen both the scope and reliability of the test. All these call for attention to take a re-look on the TST, and, if possible, do basic work to experiment and innovate.

To conclude, the above issues highlight that utmost caution should be observed in conduct and interpretation of the tuberculin test. A lot depends on the individual being tested, the PPD product used, and the interpretation of the size of the induration /reaction which is sought to be measured. Its use therefore must be cautious and selective.

References
2. The story of the tuberculin test from an epidemiologic view point, Ame Rev Resp Dis 1960, 1, 1-47.
6. Reichman LB. Tuberculin skin testing, State of the art; Chest, 1979, 76S, 764S.
9. Chaparas S D. Why monospecific tuberculin skin test antigens have not been isolated, Bull IUAT 1979, 54, 156.


18. Raj Narain Vallishayee RS and Venkatesh Reddy A. Value of dual testing with PPD-S and PPD-B. In


