# The Efficacy of the Tuberculin Test An Analysis Based on Results from <sup>33</sup> Countries

JØRGEN NYBOE, cand. act.<sup>1</sup>

Is it always justified to take a tuberculin reaction over a certain size limit as indicative of infection with tubercle bacilli and reactions below the limit as indicative of absence of such infection ? Information on this point is here derived from a quantitative analysis of test reactions found in general population groups. The data were collected by specially trained international teams operating simultaneously with or as a preliminary to large-scale tuberculosis control programmes. Because of the international character of the work it has been possible to compile a picture which shows the tuberculin sensitivity in a large part of the world, yet which is based on highly uniform techniques. It is hereby demonstrated that the pattern of tuberculin sensitivity varies widely between different populations, but follows a definite geographical trend. In temperate and subtropical countries almost all test reactions are either clearly " positive " or clearly " negative ", indicating that the test is highly efficient. In tropical regions, on the other hand, a large proportion of the reactions are intermediate in size, and distinction between two kinds of reaction is therefore difficult. The data strongly suggest that the cause of the intermediate reactions is that the population is being massively exposed to certain unidentified agents producing cross-reactions to tuberculin. In tropical regions a clear-cut distinction between tuberculosis infected and uninfected evidently cannot be made by means of the present tuberculin test.

#### INTRODUCTION

In most of the tuberculosis projects currently conducted with international assistance in Africa, Asia and South America, the tuberculin skin test is of critical importance. It is used in BCG vaccination campaigns for selecting subjects for vaccination, in tuberculosis surveys for estimating the prevalence of tuberculous infection, in tuberculosis control projects for diagnosis of disease.<sup>2</sup> In each of these applications, however, the purpose of the test is the same: to distinguish between those persons infected with tubercle bacilli and those uninfected.

While the test seemed to fuffil its purpose without presenting serious problems in the countries in which it was developed, no systematic data were available on its efficacy in other parts of the world. The need for such data was recognized, and the first

studies to obtain them began shortly after the WHO Tuberculosis Research Office (TRO) was established in 1949. During this decade pertinent studies have been included in many different projects carried out under the technical direction of the TRO. Most of the data were collected by the so-called BCG assessment teams that were established to evaluate the results of the mass BCG vaccination campaigns. Within this programme it was natural to include studies of the effectiveness of the test in selecting persons for vaccination. Further, data were collected from a number of studies of various technical aspects of BCG vaccination and from tuberculosis prevalence surveys carried out in several countries in Africa. In all these projects the field personnel were trained in the use of uniform testing techniques, a fact which contributes considerably to the comparability of the material from the various projects and countries.

These data show that the efficacy of the tuberculin test differs widely in different parts of the world. In some regions the reactions to the test fall into two distinct groups, large reactions and small reactions, and the population can thereby be divided into two

<sup>1</sup> Division of Communicable Diseases, WHO, Geneva, Switzerland (formerly Statistician, WHO Tuberculosis Research Office, Copenhagen).

<sup>&#</sup>x27;The test is also used to evaluate tuberculin allergy induced by BCG vaccination, but the present paper is confined to an evaluation of the efficacy of the test in distinguishing tuberculin sensitivity induced by natural infection.

categories, presumably the infected and the uninfected. But in many other regions a large proportion of the test reactions are intermediate in size, and the distinction of two categories becomes difficult. The reasons for the varying efficacy are not yet clear, but a good deal of observation has been made that begins to throw some light on the nature and cause of the perplexing intermediate tuberculin sensitivity. The data available are summarized in the present paper, and their interpretation is discussed.

#### METHODOLOGY

The history of the tuberculin test is a history of varying opinions as to the appropriate dose. When Charles Mantoux intoduced the intradermal test technique, he suggested <sup>a</sup> dose of <sup>1</sup> TU (1/20 ml of <sup>a</sup> 1:5000 dilution of OT) (Mantoux, 1910), but by 1935 a dose of 100 TU—preceded by weaker doses was often used, and some authors even advocated <sup>1000</sup> TU as <sup>a</sup> final dose (Saye, 1936; Paretzky, 1937). This development towards a high test dose was no doubt influenced by the clinical observation that tuberculous patients do not always react to a low dose. But the consequences of using a high test dose in uninfected persons were not recognized.

It was not until the introduction of a new study technique-a quantitative study of tuberculin reactions—that it was shown by Palmer that a lowdose test was the more appropriate. By testing patients in tuberculosis hospitals he showed that a dose of <sup>5</sup> TU (PPD) produces definite reactions in all but a very small proportion; more precisely, that when the reactions are arranged according to size of diameter of induration they will approximate to <sup>a</sup> normal distribution, with <sup>a</sup> mean of 15-20 mm and <sup>a</sup> standard deviation of 3-4 mm (WHO Tuberculosis Research Office, 1955a). And when the same test dose is used in groups of healthy persons, a varying but certain proportion present reactions that are distributed like the patients' reactions, whereas the remainder have only small reactions or none at all. Palmer (1953) also showed that the frequency of reactions to the low dose of tuberculin depends on the degree of contact with cases of tuberculosis, but that the distribution of reactions to a high dose does not; the latter varies with location of residence. On the basis of these findings he concluded that <sup>a</sup> dose around <sup>5</sup> TU is best suited for selecting persons infected with tubercle bacilli, and that reactions to high doses of <sup>100</sup> TU or more are

caused by infection with an organism other than Mycobacterium tuberculosis of the human or bovine type.

The *quantitative* study of tuberculin reactions has also been applied in the present analysis of the efficacy of the test. A Mantoux test was given in various population groups in each country, usually without further medical examination. The Mantoux technique is well suited for quantitative studies of tuberculin sensitivity because the resulting reaction (erythema or induration) forms a confluent area that can be measured relatively easily. Other test techniques such as Moro's or Heaf's may by easier to apply but the reactions cannot be objectively quantified: they require a judgement on the part of the reader for classification. In addition to this essential property the Mantoux test has certain other advantages: with a suitable dose the reaction sizes are distributed over a relatively large range, which is convenient, and the dose of tuberculin can be more precisely determined than with other testing techniques.

<sup>5</sup> TU <sup>1</sup> was used as the standard dose of tuberculin. In addition, persons with small reactions to the Mx <sup>5</sup> TU test were tested with Mx <sup>100</sup> TU in some groups. The results of the latter testing have been included to elucidate certain trends in the pattern of small reactions to 5 TU.

The reactions were read by measuring the diameter of the indurated skin area in millimetres, and these readings were recorded on individual cards or on lists together with information on age and sex. The analysis of the data comprises a study of the distributions of tuberculin reactions according to size. While simple size-distributions of reactions do not allow conclusions to be drawn as to the medical significance of test results, they do provide some reference information about the behaviour of the test in different countries, which appears to be essential before specific studies of interpretation can fruitfully be made.

#### STUDY POPULATION

The study population comprises about 190 000 persons tested in 33 different countries in Africa, America, Asia and Europe during the period 1950-58

<sup>&</sup>lt;sup>1</sup> The dilutions were prepared to contain 5 TU. However, recent studies have shown that PPD is adsorbed to the walls of the glass containers in varying amounts; consequently the doses actually administered are now estimated to have been about 2 TU (Waaler et al., 1958).

(Table 1). About  $44\%$  of the tested were schoolchildren, and the remainder were general population groups cf all ages.

The study population groups were not randomly selected except in some African countries, and they cannot therefore be regarded as truly representative of the populations of the countries in question. This deficiency in the study population is due, in the main, to two circumstances. First, a large part of the data was collected among schoolchildren tested for inclusion in the BCG studies. No attempt was made to obtain a representative selection of schools because schoolchildren are not necessarily representative of the total population of their agegroup. Secondly, <sup>a</sup> mass BCG campaign was often under way in the countries concerned, and the population already covered by the campaign could not be included in the present studies. As the BCG campaigns usually proceeded systematically from district to district, there was no possibility of selecting a sample that would be representative of the whole area.

Possibilities for choice were thus often very limited, but when feasible the team leaders tried to select groups well scattered over the country and representing different terrains-mountainous, lowland, desert, etc. In countries where BCG vaccination campaigns were in progress, assessment team leaders were instructed to examine all persons for BCG vaccination lesions if mass campaign teams had been operating in the neighbourhood. All persons presenting a scar that might have been due to BCG vaccination have been excluded from the data.

The number tested varies widely from country to country, from 450 in Mexico to 48 000 in the Sudan; in most countries the number lies between 3000 and 10 000 (Table 1). In order to give an idea of the scatter of the groups included, the number of localities visited and the number of major administrative divisions visited out of the total number in the country have also been given in Table 1. In some countries the data comprise one or a few communities only and in others the localities visited are concentrated in a limited part of the country. On the other hand, in those countries where the data are relatively complete an analysis by locality has shown that the findings are consistent; and moreover the data for all countries follow a very clear trend. We therefore believe that the non-randomness and other limitations of the data do not seriously impair the validity of the conclusions drawn.

#### TECHNIQUE OF TUBERCULIN TESTING

The tuberculin product used in these studies was PPD prepared from human strains of tubercle bacilli in the Statens Seruminstitut, Copenhagen. One batch, RT 19-21, was used in all countries except Ecuador, where batch RT <sup>22</sup> was used. One unit of RT 19-21 is equivalent in powder weight to one unit of the international standard tuberculin, PPD-S, i.e., 1/50 <sup>000</sup> mg. One unit of RT 22, <sup>a</sup> more potent preparation, is defined as 1/75 000 mg. Dilutions ready for use were prepared at the Statens Seruminstitut and forwarded to the teams by air, except in the two American countries, where stock solution was sent to local laboratories for further dilution.

The tests were given by injecting 0.1 ml of dilution very superficially into the skin on the dorsal aspect of the forearm. The injections were made with all-glass syringes, and the amount injected was judged by the scale on the barrel of the syringe, not by the weal produced on the skin; since 1953 all syringes have been pretested for leakage. The criterion for giving the Mx <sup>100</sup> TU test has varied somewhat, from the early criterion of <sup>4</sup> mm or less induration to the Mx <sup>5</sup> TU test up to <sup>9</sup> mm or less in groups recently tested.

The tests were read three or four days after they had been given, by measuring the transverse diameter of the indurated skin area in millimetres. The nurses reading the reactions had received extensive training in measuring reactions precisely and in maintaining a constant level of measurement. They were specifically trained not to think of the reactions in terms of " positives " or " negatives ", and even when the result had certain consequences for the person tested (for example, whether BCG vaccination or <sup>a</sup> <sup>100</sup> TU test should be given), the nurses were expected to maintain impersonal, objective measurement of the reaction size. It is general experience that nurses who have been given this training tend to read reactions slightly larger than " untrained" nurses, and it should be borne in mind that as far as the absolute size of reactions is concerned the observations reported here might have been slightly lower if made by other persons.

In spite of all efforts to obtain uniform results the data are still influenced by certain variations in the testing technique. Even experienced nurses do not measure reactions alike: a difference between readers of <sup>3</sup> mm may occur. Further, recent studies show that tuberculin dilutions prepared to contain <sup>5</sup> TU

# TABLE <sup>1</sup> EXTENT OF DATA



 $^a$  The figures in parenthesis give the total number of major administrative divisions in the country.<br> $^b$  G = general population groups S = schoolchildren

vary considerably because a high but variable proportion of the active substance is adsorbed to the surface of the container; the dose injected, therefore, has not always been the same.<sup>1</sup> These variations in the measuring instrument must therefore be kept in mind when evaluating the observed differences in reaction size.

#### RESULTS

#### Variations in sensitivity pattern

The tuberculin testing results in each country are given in Appendix Tables <sup>1</sup> and 2 as frequency distributions of reactions to Mx <sup>5</sup> TU and Mx <sup>100</sup> TU according to size (in 2-mm groups) in specified agegroups. In order to present first a picture of the kind of variation in the sensitivity patterns found in various populations the results of testing with <sup>5</sup> TU in one age-group (10-14 years) in <sup>33</sup> countries are shown graphically in Fig. 1.

The shape of the distributions varies considerably in several respects, but the most important difference in the present context is that some distributions can be described as bipartite whereas others can not. Bipartite distributions-those, for example, in Basutoland, Cyprus, Denmark and Egypt-are characterized by an accumulation of reactions around 0-3 mm with rapidly decreasing frequencies in the succeeding size-groups, very few reactions in the range 8-11 mm and <sup>a</sup> second accumulation of reactions between <sup>12</sup> and <sup>25</sup> mm with <sup>a</sup> maximum around 15-20 mm.

This sensitivity pattern indicates that the population comprises persons with two essentially different responses to tuberculin: some respond with large reactions and may be termed reactors, others with small or no reactions, the non-reactors. An observed bipartite distribution with an estimated breakdown into the reactor and non-reactor components is given in Fig. 2. The breakdown is made so that both component distributions are regular in shape and taper off smoothly.

The left-hand component has <sup>a</sup> mode at 2-3 mm and tails off rapidly to the right; only  $1\%$  of the reactions exceeds <sup>7</sup> mm. Its shape resembles the distribution of reactions obtained when buffered diluent alone is used for testing (Table 2). The right-hand component has the shape of a normal distribution with <sup>a</sup> mean of <sup>19</sup> mm and <sup>a</sup> standard deviation of 3-4 mm. There is very little overlapping between the two component distributions, or, in

TABLE <sup>2</sup> DISTRIBUTIONS OF REACTIONS TO BUFFERED DILUENT

ACCORDING TO SIZE IN THREE COUNTRIES



other words, the tail-end of each component that stretches into the range of the other comprises only a small proportion of the reactions. Thus if the limit between positive and negative reactions is set at, say, <sup>10</sup> mm the proportion misclassified in either direction will be small.

But it is clear from Fig. <sup>1</sup> that the sensitivity pattern in many countries is not bipartite. In Bechuanaland, Burma, Cambodia, the Philippines and Viet Nam, for example, no less than  $15\%$ - $20\%$  of all reactions fall into the intermediate range 8-11 mm, and at no place on the scale is there an interval with few reactions separating the distribution into two parts. And the non-bipartite distributions vary considerably in shape. Some distributions show a maximum percentage of reactions around 5-6 mm, with decreasing frequencies both towards smaller and larger reaction-sizes (Burma, Cambodia, Mauritius). In others the frequencies do not vary much between <sup>3</sup> and <sup>18</sup> mm (Bechuanaland, Nigeria, the Philippines). It is significant, however, that in most of the non-bipartite distributions two modes can be seen more or less clearly, the one around 5-6 mm and the other around 15-16 mm, indicating that these populations-like those characterized by bipartite distributions-include two groups of persons with two different responses to tuberculin. Even in those few distributions where no bimodality can be observed (East Pakistan, for example), two component parts merging into a flat or unimodal shape can be supposed: as composite distributions are characteristic of the great majority of countries it seems reasonable to assume

<sup>&</sup>lt;sup>1</sup> See footnote, page 6.



DISTRIBUTIONS OF REACTIONS TO THE Mx <sup>5</sup> TU TEST ACCORDING TO SIZE AMONG PERSONS FIG. <sup>I</sup> 10-14 YEARS OLD IN <sup>33</sup> COUNTRIES





this is true of them all. Further evidence supporting this assumption will be given in the next section.

An observed non-bipartite but bimodal distribution is given in Fig. 3 with three hypothetical breakdowns into the two component parts. Hypothesis C, where the two components split abruptly at one point on the scale, is contrary to general biostatistical experience and quite implausible. Distributions of biological reactions generally assume a smooth shape, tapering off more or less regularly, and do not end abruptly. Hypotheses A and B depict two plausible explanations of how the two components might be placed. In A the righthand (reactor) component has the shape and location of the reactor component in a bipartite distribution, while the left-hand component extends further to the right than in a bipartite distribution; in B the left-hand component is assumed to end at about the same place on the size-scale as in the bipartite distributions, while the reactor group spreads considerably further to the left. Both hypotheses are plausible, and it is clear that from the shape of an observed non-bipartite distribution alone the correct placement of the two components cannot be deduced. All that can be said is that the right-hand component cannot be located much higher up the scale than depicted in A, nor much lower than shown in B, if the distributions for both components are to be unimodal and regular in shape. Some further information about the distribution of the two components of a non-bipartite distribution can be gained from a study of the pattern of sensitivity by age and will be presented in the following section.

The most important implication of the nonbipartite distribution pattern is that it is not possible to distinguish sharply between the two components by means of the test reactions. No matter where the limit between positive and negative reactions is set, a not inconsiderable proportion of persons belonging to one category will inevitably be included in the other.

#### Trends according to age

The foregoing analysis was based on a study of reaction sizes in a single age-group, 10-14 years. In order better to define the components of the size-distributions, an analysis of reaction size will be made in this section for all age-groups.

Distributions of reactions for specified age-groups are given in Fig. 4 for Basutoland, Iraq, Libya and West Pakistan, areas where the separation between the two components in the age-group 10-14 years is distinct. Mx <sup>5</sup> TU reactions are given in the upper half of the figure, Mx <sup>100</sup> TU reactions in the lower. The <sup>5</sup> TU distributions for each age-group contain the two components observed in the age-group 10-14 years, but in varying proportions. In the youngest ages, 0-4 years, the non-reactor component is by far the largest, the reactor component very small. With increasing age the non-reactor component diminishes and the reactor component increases correspondingly until in the age-group 50 years and over it comprises the great majority of reactions. As people get older they pass from the non-reactor into the reactor group, and this increase in tuberculin sensitivity with age is commonly found wherever tuberculin testing surveys are done.

This general trend, however, is not of so much interest here as the observation that the separation



FIG. 3 OBSERVED NON-BIPARTITE DISTRIBUTION WITH THREE HYPOTHETICAL BREAKDOWNS INTO TWO COMPONENT PARTS

between the two components becomes somewhat blurred with increasing age: the proportion of reactions falling between <sup>8</sup> and <sup>11</sup> mm increases. A closer scrutiny of the distributions in the different age-groups reveals that this increase is at least in part due to changes in the left-hand component. Among the youngest persons the majority of reactions fall between 0 and <sup>3</sup> mm, and few fall between 4 and 9 mm; but with increasing age a comparatively larger number measure 4-9 mm, and a smaller number measure 0-3 mm. There seems to be a shift towards slightly larger reactions among the non-reactors with age. Retesting them with <sup>100</sup> TU (cf. lower part of Fig. 4) supports the trends observed in the <sup>5</sup> TU distributions: the youngest non-reactors have essentially no tuberculin sensitivity, even to a high dose, but as age increases the response to <sup>100</sup> TU also increases.

Thus it becomes clear that the persons termed non-reactors are not entirely lacking in sensitivity to tuberculin. Their reactions to Mx <sup>5</sup> TU, though small, increase with age, and their reactions to Mx <sup>100</sup> TU are clearly larger than those to buffered diluent, shown in Table 2. This tuberculin sensitivity among persons belonging to the left-hand component, whether observed in reactions to <sup>5</sup> TU or brought out only by a higher dose, will in the following be termed weak sensitivity or low-grade sensitivity in contrast to the strong sensitivity manifesting itself as large reactions to 5 TU.

There are a few further observations to be made in connexion with the results of testing with 100 TU. First, considerable geographical variation is observed in the level of weak tuberculin sensitivity. Among non-reactors to the low-dose test in the age-group 20 years and over in Libya, for example, most of the <sup>100</sup> TU reactions are relatively small, but in Basutoland most are large, and in Iraq and West Pakistan there is approximately the same frequency of reactions of all sizes between 0 and 22 mm.' Secondly the distributions of reactions to <sup>100</sup> TU are in general not bipartite. In some countries the distribution is bimodal in some age-groups, but it should be clear that the non-reactor component of

<sup>1</sup> The fact that the criterion for giving the <sup>100</sup> TU test was <sup>7</sup> mm or less of induration to <sup>5</sup> TU in Libya and <sup>9</sup> mm or less in the three other countries could hardly change this general picture. There are so few <sup>5</sup> TU reactions in the range 8-9 mm in Libya that the size-distribution of <sup>100</sup> TU reactions would only be slightly affected if these persons had been included in the <sup>100</sup> TU testing.

AGE-GROUP **BASUTOLAND IRAQ LIBYA** PAKISTAN, WEST  $(YEARS)$  $RO$  $5TU$ 60  $O - 4$ 40  $20$  $\Omega$ 40  $5 - 9$ 20 <u>si 1873 v</u> 40  $|0-19|$  $\overline{20}$ <u>Sillisin</u>  $\mathcal{C}$ 40  $20 - 49$ 20 <u> Manan</u>  $\circ$  $40$ 50  $\begin{array}{r} 50 \\ \text{AND OVER} \\ \hline \\ \text{EECLUTAGE} \\ \text{ECEU} \\ \text{$ IOO TU 60  $O - 4$ 40  $20$ mand William  $\Omega$ 40  $5 - 9$  $20<sup>1</sup>$  $\Omega$ 40  $10 - 19$ 20 40 20 AND OVER 20 <u> Millian Milli</u> DIAMETER OF INDURATION (mm)

FIG. 4 DISTRIBUTIONS OF REACTIONS TO THE Mx <sup>5</sup> TU AND Mx <sup>100</sup> TU TESTS ACCORDING TO SIZE IN SPECIFIED tAGE-GROUPS IN BASUTOLAND, IRAQ, LIBYA AND WEST PAKISTAN

Mx <sup>100</sup> TU given to persons with Mx <sup>5</sup> TU reactions measuring less than <sup>10</sup> mm (in Libya <sup>8</sup> mm)

the Mx <sup>5</sup> TU distributions cannot be separated into two distinct groups by testing with Mx <sup>100</sup> TU.

The reactor component in these size-distributions of Mx <sup>5</sup> TU reactions does not differ markedly between the various age-groups except that it grows with increasing age. As pointed out previously, this component approximates to the shape of a normal curve and is therefore characterized by the mean size and the standard deviation. Estimates of these statistics are given in Table 3 for the reactor component in various age-groups in eight countries; the countries included are those in which the reactor component is clearly discernible and for which data are available over a wide age-range.

The means are very stable within each country but a slight decrease in size is observed in most countries with increasing age: on the average the mean measures 1.3 mm less among persons <sup>50</sup> years of age and over than among children 5-9 years old. The standard deviation increases slightly but consistently with age in nearly all countries; on the average it is higher by about one millimetre in the oldest age-group than in the young children.

It is clear that as the spread of reactions increases and the mean size drops slightly with increasing age there will be an increase in the proportion of reactions in the reactor component measuring, say, less than <sup>11</sup> mm and that this increase will of necessity blur the distinction between the two components in older persons. With increasing age there are thus trends in the sensitivity pattern of both the nonreactors (an increase) and the reactors (a decrease) that make the distinction between them a little less clear.

Common to the populations so far discussed is the finding that in general a relatively good distinction between the two components can be made by means of the Mx <sup>5</sup> TU test. If attention is turned to populations where the distinction is poorer, a different picture emerges, and one that is more difficult of interpretation. Data from South India, Nigeria, East Pakistan and Tanganyika are shown in Fig. 5, the Mx <sup>5</sup> TU test results in the upper part, the Mx <sup>100</sup> TU results in the lower part. The general pattern of reactions to <sup>5</sup> TU resembles somewhat the trend observed in Fig. 4 in that tuberculin sensitivity is seen to increase with age. However, the shape of the distributions in the ages between 10 and 49 years is markedly different from those in Fig. 4: none of them are bipartite.

In three of the countries the size-distributions are bimodal in some age-groups, indicating again that the populations comprise two groups of persons with different responses to tuberculin. In East

ESTIMATED MEAN SIZE AND STANDARD DEVIATION OF REACTIONS COMPRISING THE REACTOR COMPONENT TABLE 3 IN SPECIFIC AGE-GROUPS IN EIGHT COUNTRIES \*

Country	Mean size of diameter of induration (mm)							Standard deviation of diameter of induration (mm)							
	age-group (years)							age-group (years)							
	$0 - 4$	5-9	$10 - 14$	$15 - 19$	20-49	$\geqslant 50$	$0 - 4$	$5-9$	10-14	$15-19$	20-49	$\geqslant50$			
<b>Basutoland</b>	17.8	18.6	18.1	18.1	17.3	16.3	3.3	3.1	3.1	3.3	4.0	3.9			
Cyprus	$\overline{\phantom{a}}$	19.3	19.7	18.2	19.0	18.2	$\qquad \qquad \blacksquare$	4.0	3.5	3.8	3.6	4.1			
Iraq	19.0	18.8	18.9	18.7	18.4	17.4	4.6	3.3	3.2	3.1	4.2	6.0			
Jordan	$\overline{\phantom{a}}$	20.3	20.1	20.1	19.8	19.9	$\overline{\phantom{a}}$	3.6	3.6	3.2	3.5	4.2			
Lebanon	$\overline{\phantom{a}}$	19.1	20.0	19.4	19.4	19.4	$\qquad \qquad \longrightarrow$	3.5	3.7	3.3	3.6	3.8			
Libya		19.5	19.2	19.3	19.0	18.8	$\qquad \qquad \qquad$	3.7	3.1	3.9	3.1	4.9			
Pakistan, West	18.5	18.6	18.2	17.8	17.5	17.4	4.3	3.2	3.3	3.3	3.6	4.1			
Turkey	$\overline{\phantom{m}}$	19.9	19.0	16.9	16.7	16.5	$\overline{\phantom{m}}$	3.4	3.6	3.5	4.1	4.3			
Simple average	$\overline{\phantom{m}}$	19.3	19.2	18.6	18.4	18.0	$\overline{\phantom{0}}$	3.5	3.4	3.4	3.7	4.4			

\* For persons under <sup>20</sup> years of age the estimates have been computed from reactions measuring <sup>10</sup> mm or more. For older age-groups, where the separation between the reactor and non-reactor components is less clear, estimates have been computed from reactions measuring 13 mm or more and adjustment for the truncation made as described by Fish



FIG. 5 DISTRIBUTIONS OF REACTIONS TO THE Mx <sup>5</sup> TU AND Mx <sup>100</sup> TU TESTS ACCORDING TO SIZE IN SPECIFIED AGE-GROUPS IN SOUTH INDIA, NIGERIA, EAST PAKISTAN AND TANGANYIKA

Mx <sup>100</sup> TU given to persons with Mx <sup>5</sup> TU reactions measuring less than <sup>10</sup> mm (in South India <sup>8</sup> mm)

Pakistan none of the distributions are bimodal, but the way their shape changes with age suggests a transfer of persons from one group to another. If the distributions reflected a uniform response to tuberculin, the increase in sensitivity would be seen as a gradual increase in the modal point, but this is not what is observed. As the age of the tested increases the mode remains at 2-3 mm up to about 20 years while the right tail swells; after age 20 the mode is located around 14-18 mm, and the left tail seems to shrink. This is exactly the pattern that would be expected if the distributions were composed of two widely overlapping parts.

An attempt has been made to study the variation in shape and location of the two components with age as was done for bipartite distributions, but the extensive overlapping of the two components makes this difficult, and the results are limited.

The left-hand component can be studied rather well in the young ages, where the reactor component is small. Examination of Fig. 5 shows that the tuberculin sensitivity of the left-hand component increases rapidly with age: in the youngest agegroup, 0-4 years, most of the reactions are less than <sup>4</sup> mm and only <sup>a</sup> few measure 4-9 mm (except in South India), but in the next age-groups, 5-9 and 10-19 years, the proportion of very small reactions decreases and the proportion of reactions measuring 4-9 mm increases. In South India the left-hand component is located rather high up on the sizescale even among the youngest children, but a shift to the right is nevertheless discernible with increasing age. These trends are well corroborated by testing with 100 TU: in the youngest age-group a large proportion of the reactions are small, but strong reactions are already in the majority among children 5-9 years old, and after the age of 10 years practically everyone reacts strongly to 100 TU. The tuberculin sensitivity of persons in the left-hand component thus increases much more rapidly in these populations than in the populations characterized by a bipartite pattern of sensitivity. Whether the increase continues beyond the age-group 10-19 years is difficult to ascertain because the two components become thoroughly intermixed beyond this point.

The right-hand component is more difficult to analyse. Among young children 5-9 years old it sometimes stands out rather clearly, and from these few instances the inference can be made that this component again assumes the shape of a normal distribution and that the spread is approximately

the same as that observed for the right-hand component in bipartite distributions. In the next agegroups the distinction becomes very poor because of the extensive overlapping with the left-hand component, but in persons 50 years of age and over the right-hand component is again well defined. Here again the mean size of reactions is 1-2 mm less than among young children.

Even with this limited knowledge at hand, breakdowns of the composite distributions into their component parts can be attempted. In illustration, an estimated breakdown of the data for Nigeria is shown in Fig. 6. The divisions have been made by eye, in such a way that the left-hand component is regular in shape and the right-hand component is symmetrical and approximately " normal " and its mean that indicated by the observed distributions. With these conditions the possibilities of subjective variation in estimation are very limited. In other countries such as East Pakistan and Burma breakdowns of this nature would be much less certain because the observed distributions give no indication of the mean of the right-hand component in any age-group except the oldest.

Analysis of the shape and location of the two components by age thus shows that the sensitivity pattern of a population is primarily determined by variation in the tuberculin sensitivity of the nonreactors. There is, in all populations, an increase in tuberculin sensitivity among the non-reactors with age, but whereas in some populations this sensitization is so faint that persons with weak sensitivity are well distinguished from those with strong sensitivity until late in adulthood, in others the sensitization is so much stronger and more rapid that it obscures the separation between the two components at a very early age. In these latter populations persons belonging to the left-hand component can hardly be termed non-reactors: except among young children, an appreciable proportion of them do in fact react to <sup>5</sup> TU with <sup>10</sup> mm of induration or more, and when tested with <sup>100</sup> TU they all react strongly.

Variations in the location of the right-hand component also occur, and it appears that in those populations where the sensitization of the lefthand component is most marked, the right-hand component is depressed, i.e., the mode is usually located a few millimetres lower on the size-scale. The connexion between a high level of weak sensitivity and a slightly depressed level of strong sensitivity is not found in all countries, but sufficient

## OBSERVED DISTRIBUTIONS OF REACTIONS TO THE Mx <sup>5</sup> TU TEST ACCORDING TO SIZE IN SPECIFIED AGE-GROUPS IN NIGERIA WITH ESTIMATED BREAKDOWNS INTO COMPONENTS



observations are available to establish a definite correlation between the two phenomena. A schematic illustration of the varying shape and position of the two components and the effect of this variation on the composite distribution is shown in Fig. 7.

## Geographical trends

Analysis of size-distributions of reactions has shown that the clarity of distinction between two kinds of reaction varies considerably from country



to country. In order to evaluate the geographical trends in this variation, a quantitative measure of the degree of distinction (or merging) is called for. The simple two-way description of distributions as bipartite and non-bipartite used to illustrate the sensitivity patterns in the foregoing sections is not adequate for this purpose because in fact all degrees of merging occur. Brief reference to Fig. <sup>1</sup> will show the range: in Jordan, for example, the two components are very distinct, they are a little less so in Iraq, much less so in North India, and in Viet Nam they coalesce to <sup>a</sup> large extent. As <sup>a</sup> measure of the degree of merging we shall use in this section the percentage of reactions falling in the size-range 8-11 mm. While this index is not an absolute measure of the merging, a low figure will indicate a good distinction, the higher the percentage the greater the merging.

As has been shown, the frequency of Mx <sup>5</sup> TU reactions in the size-range 8-11 mm varies with the age of the tested. The age-factor must therefore be taken into account in a comparison between countries, and this has been done in the following analysis by using data from one age-group only. The age-group 10-14 years has been chosen because it includes a fair number of observations in most countries; the trends, however, are much the same in the age-groups 5-9 and 15-19 years, as the reader may verify for himself from Appendix Table 1.

The percentages of reactions between 8 and <sup>11</sup> mm are given for the age-group 10-14 years by country in Fig. 8. The percentages vary from  $0.4\%$ in Cyprus to  $25.9\%$  in Burma, but a striking geographical trend is immediately clear from the map: the percentages are low in the northern temperate and subtropical countries and, with two exceptions (Mexico and Ecuador), high in tropical areas. Data are available for only two countries south of the Tropic of Capricorn: Swaziland with an index of  $7.2\%$  and Basutoland with a very low index of 1.9%. It would thus appear that the efficacy of the tuberculin test in distinguishing between reactors and non-reactors is rather sharply delineated by climatic zones: it effects a sharp distinction in temperate and subtropical areas and usually a poor one in populations living in the tropical belt.

In order to study the uniformity of the data within countries, breakdowns by individual localities are given in Table 4. In countries where the index is low—less than  $2\%$ —there is little variation from locality to locality: in no locality does the index exceed  $4\%$ . In those countries where the average index is  $10\%$  or more there is some variation—from  $6\%$  to  $34\%$ —but with a single exception (Viet Nam) the index is never low. For areas with either a very low or a very high index, then, the data are rather uniform.

In some of the countries where the index lies between  $2\%$  and  $10\%$  the variation is more pronounced-for example, North India, Iran and Swaziland. In North India this variation has previously been shown to be related partly to geographical location, partly to altitude above sea level (Bates, Busk & Palmer, 1951; WHO Tuberculosis Research Office, 1955c). In Gauhati, for example, a lowland town in Assam State (bordering East Pakistan),  $27\%$  of reactions fall in the intermediate range and the distinction between the two components is very poor indeed; in the nearby town of Shillong, however, situated 6000 feet (about 1800 m) above sea level, the distinction is much better. A similar difference is observed between villages at an altitude of 7500 feet (about 2300 m) in Uttar Pradesh and localities in the lowlands of



Country ________	% in individual localities												
<b>Basutoland</b>	0.0	0.0	2.2	2.4	2.4	3.1							1.9
Cambodia	8.3	11.5	14.2	15.5	17.4	20.7	21.1	21.8	23.6	24.9	26.5	31.3	20.6
Cyprus	0.0	0.0	0.0	0.0	0.0	0.3	0.4	1.1	1.4	1.4			0.4
Denmark	0.0	0.0	0.0	0.0	0.0	2.5	2.6	2.9	3.9				1.2
India, North	1.3	1.8	3.7	4.2	4.2	4.8	4.8	6.4	7.8	27.0			7.5
Indonesia	5.7	6.2	6.9	8.2	8.7	10.2	12.3	13.3	14.1	21.4	21.8	29.8	10.9
Iran	2.0	2.6	3.2	3.4	4.6	6.2	7.7	9.4	16.0				4.8
Lebanon	0.0	0.0	0.0	0.0	0.0	0.8	0.8	1.3	1.6				0.6
Nigeria	6.7	7.7	11.0	11.2	13.3	14.3	17.9	20.0	22.2	23.6	24.6		13.9
Pakistan, West	0.0	1.5	3.1	3.1	3.1	3.8	4.2	5.2					2.8
Philippines	7.1	13.3	14.4	14.8	15.3	15.5	17.2	20.0	20.8	23.9	24.2	33.5	18.1
Somalia	2.9	6.0	7.6	9.1	10.1	10.1	10.1	10.2	12.4	12.7	13.4	13.6	9.8
Somaliland Protectorate	2.9	3.9	5.4	5.5	6.2	6.7	7.0						5.3
Sudan	8.0	8.0	8.1	11.0	11.3	12.0	12.1	12.4	14.3	14.6	21.2	23.0	15.7
Swaziland	1.2	4.4	4.5	8.1	8.7	10.8	17.3						7.2
Turkey	0.0	0.0	0.0	0.0	1.1	1.3	1.4	1.9	3.2				1.3
Viet Nam	1.7	12.1	15.2	16.7	19.0	23.0							17.5

TABLE <sup>4</sup> PERCENTAGE OF REACTIONS TO THE Mx <sup>5</sup> TU TEST MEASURING BETWEEN <sup>8</sup> AND <sup>11</sup> mm AMONG PERSONS 10-14 YEARS OLD IN INDIVIDUAL LOCALITIES IN <sup>17</sup> COUNTRIES \*

\* Only localities with 50 or more observations in the age-group 10-14 years have been included. Where less than six such loca-<br>lities were available, the country has been excluded from the table; where more than 12 were av is given.

Delhi and Rajasthan. In this connexion it should also be pointed out that the one group tested in Mexico lay more than 6000 feet above sea level, and that a high proportion of the children tested in Ecuador were living in the mountains. These observations suggest that it is the climatic conditions rather than the geographical latitude that determine the sensitivity pattern.

On the whole the data indicate that there exist in tropical surroundings certain factors that impair the efficacy of the test in distinguishing two kinds of reaction.

#### DISCUSSION

The purpose of a biological selective test is to effect in a given population a separation into two categories of individuals who differ from each other in a specific biological aspect. While it is often difficult if not impossible to observe the biological phenomena directly, their occurrence is usually accompanied by physiological changes that can be observed, and the biological selective test consists in determining whether or not the physiological change has occurred. For example, the tuberculin test measures skin sensitivity to tuberculin from which infection with tubercle bacilli is inferred.

In an evaluation of the efficacy of the test, it must be borne in mind that there are two kinds of error involved: persons infected with tubercle bacilli may be classified as uninfected (hereafter referred to as error of the first kind) and uninfected persons may be classified as infected (error of the second kind). In any given population the frequency of both kinds of error may be affected by changing the limit between positive and negative reactions, or by changing

the test dose, but it can be shown that efforts to minimize the one error will often result in an increase in the other. If, in the observed distribution shown in Fig. 2, the right-hand component is assumed to comprise the infected and the left-hand component the uninfected, and the limit for positive reactions is changed from <sup>10</sup> mm to <sup>8</sup> mm, the frequency of errors of the first kind will be reduced, but, simultaneously, the frequency of the second kind of error will be increased. Conversely, if the limit is raised to <sup>12</sup> mm there will be fewer errors of the second kind, but more of the first. By the same token, if the test dose is increased but the limit between positive and negative reactions remains unchanged, there will be fewer errors of the first kind but probably more of the second.

There has in the past been a tendency within the medical profession to attach the greater importance to avoiding the first kind of error, the false classification of infected as uninfected. From the beginning Mantoux (1910) stated that the negative reaction was more useful from a clinical point of view because it could be used to exclude the diagnosis of tuberculosis, whereas a positive reaction had little clinical significance, as the majority of older children and adults reacted so. With that epidemiological background and with that purpose in mind it was of course important that a diagnosis of non-infection was correct. However, it does not follow that errors of the first kind are to be avoided at all costs in every situation. In tuberculin testing surveys, for example, it is important that the two kinds of error be of equal magnitude so that the estimate of the frequency of infection will be correct. Furthermore, the epidemiological situation has changed markedly since the days of Mantoux. In some countries we are now approaching a situation where a positive reaction among children and young adults is the exception rather than the rule. With this development the correctness of a diagnosis of infection may become the more important.

In the present discussion we have taken the view that the tuberculin test is effective only when the frequencies of both errors are low. This requires that two conditions be met: first, the distinction between large skin reactions and small reactions must be sharp, and, secondly, the inference of infection with tubercle bacilli from a large reaction, and of non-infection from small or no reactions, must hold true almost without fail. The two conditions are equally important: the distinction between infected and uninfected will break down both when

there are many exceptions to the interpretation of the skin reactions and when the two kinds of reaction cannot clearly be distinguished.

The present studies, conducted with uniform test procedures throughout a large part of the world, provide extensive data with which to judge the capacity of the tuberculin test to distinguish between two kinds of reaction. Mantoux (1910), in launching the technique named after him, emphasized the ease with which the test distinguished between positive and negative reactions, although he did mention a certain type of reaction (réactions frustes) that presented some difficulty in classification. The present data, compiled with approximately the same test dose as Mantoux's but on a wide geographical basis, indicate that his statement will not hold true for all populations. While reactions to the Mx <sup>5</sup> TU test do generally seem to be of two kinds, the differentiation between them cannot always be made with equal ease, and in some instances is difficult indeed. The separation is usually better among young children than among older persons. But even more important than the age factor is the geographical location of the population: the distinction is usually good where the climate is temperate or subtropical but poor where the climate is tropical. It seems, in fact, fair to postulate that if Mantoux had been working, say, in the Philippines he would not have been so convinced that the positives and the negatives are easily distinguished.

Further, differences in the capacity of the test to distinguish between the two kinds of reaction reflect variations in both kinds. The sensitivity of persons with the smaller reactions to Mx <sup>5</sup> TU varies widely among different population groups, from a very low level characterized by small reactions to both 5 TU and  $100$  TU—reactions only slightly larger than those to buffered diluentto a relatively high level characterized by intermediate-sized reactions to <sup>5</sup> TU and strong reactions to 100 TU. The sensitivity of persons with large reactions to Mx <sup>5</sup> TU also differs slightly among different population groups and seems to be lowest exactly in those populations where the small reactions are intermediate in size, a circumstance that aggravates the merging of the two kinds of reaction.

The validity of the inference of infection or noninfection from the two kinds of tuberculin reaction is the second condition of the efficacy of the test. Although information on the medical significance of the test reactions is not available from the data

given here, present knowledge and theory will be reviewed as a basis for evaluation.

Relatively good evidence exists that the large reactions to Mx 5 TU-the reactions characterized by an approximately normal size-distribution with <sup>a</sup> mean of 15-20 mm and <sup>a</sup> standard deviation of 3-4 mm-are primarily caused by infection with pathogenic tubercle bacilli, i.e., mainly of the human or bovine type. First, the shape and location of the distribution of these reactions closely resembles that obtained by testing patients in tuberculosis hospitals (WHO Tuberculosis Research Office, 1955a).<sup>1</sup> Moreover, the frequency of large reactions has been shown to be related to the degree of contact with cases of tuberculosis (Palmer, 1953). Finally, in follow-up studies in populations with little new infection, clinical cases of tuberculosis most often develop in persons with strong tuberculin sensitivity (Groth-Petersen, Knudsen & Wilbek, 1957; Palmer, Shaw & Comstock, 1958). These findings can be taken to indicate rather conclusively that large reactions to the Mx <sup>5</sup> TU test are caused by pathogenic tubercle bacilli. On the other hand it is not known whether strong sensitivity to tuberculin can be caused by agents other than the tubercle bacilli. It is generally believed not to be the case, but the possibility cannot be excluded. The small Mx <sup>5</sup> TU reactions—the left-hand component of the sizedistribution-present greater problems in interpretation. Generally, small reactions are taken as indicative of non-infection. There are some wellrecognized exceptions, however, which it might be advisable to enumerate. For example, persons tested shortly after infection do not react. This can be a serious impediment if testing is done in a group just after a tuberculosis epidemic, but as the annual infection rate is less than  $10\%$  in practically all populations—and often very much less and as the period between infection and the development of sensitivity is about one month, the proportion of infected missed on this account will on the average be low indeed. Further, persons suffering from certain diseases such as measles and scarlet fever lose their sensitivity temporarily (Weis Bentzon, 1953), but as these diseases are recognizable and occur in epidemics it should be rather easy to avoid the difficulties they create. In those populations where the distinction between large and small reactions is good there is the possibility of directly assessing the importance of these exceptions by

testing the population on two occasions at a short interval. As an illustration, some data are given in Appendix Table 3 for groups of schoolchildren in Cyprus, Iran and Lebanon tested at an interval of 2-3 months. Although there is some variation in the observed reaction sizes, there are only very few instances where the reaction is small at the first test and large at the second and none where the reaction size shifts in the other direction. The data are too limited to draw valid conclusions as to the frequency of such exceptions but do show that in these populations they do not play a very important role. Finally, it is generally recognized that tuberculous patients sometimes lose their tuberculin sensitivity when critically ill, but although this exception is important in clinical work it is hardly a serious obstacle when the test is used as a public health tool.

However, the critical question remains: How are the varying levels of weak sensitivity to be interpreted? And in this connexion naturally the question of what causes the weak tuberculin sensitivity also arises. The most plausible explanation of the latter question is that the weak sensitivity is caused by an agent or agents antigenically related to the human type of tubercle bacilli (that used in the preparation of the present tuberculin). On this hypothesis the variation in level of sensitivity among different populations could be explained partly by the agents being of different kinds in different populations, partly by differences in the degree of exposure to the agents.

The risk of exposure to the weak sensitizing agents is very high in populations living in tropical areas: practically everyone has acquired a weak sensitivity to tuberculin by the age of 10 years, while a similar prevalence of strong sensitivity is usually found only after the age of 50 years. Thus the exposure to the weak sensitizing agents must be reckoned to be several times heavier than exposure to pathogenic tubercle bacilli.

The nature of the weak sensitizing agents is not yet known, but because of their antigenic relation to human tubercle bacilli the suggestion has been made that they belong to the genus Mycobacterium (Palmer, 1953; WHO Tuberculosis Research Office, 1955b). It is also possible that they are tubercle bacilli, if this term is applied in accordance with common terminology, i.e., mycobacteria that may produce progressive tuberculous disease in man or an animal species. Various " atypical" acid-fast organisms have been isolated in recent years from

<sup>&</sup>lt;sup>1</sup> See also Fig. 9 below.

patients with disease resembling tuberculosis, and it has been demonstrated that some of these patients have only small reactions to human tuberculin but strong reactions to the homologous antigen (Edwards & Palmer, 1958).

While the weak sensitizing agents considerably alter the tuberculin sensitivity pattern, there is no indication that they contribute much to the tuberculosis morbidity. If they did, a certain proportion, anyway, of the tuberculosis patients in the tropics might be expected to have only weak sensitivity to human tuberculin, but such a phenomenon has not been reported in the medical literature. Further, in connexion with the studies reported here, tuberculous patients were tested in many countries with the same test technique and tuberculin product as used in the general population groups. Mx <sup>5</sup> TU reactions among these patients are given in Fig. 9

#### FIG. 9

DISTRIBUTIONS OF REACTIONS TO THE Mx <sup>5</sup> TU TEST ACCORDING TO SIZE AMONG PATIENTS IN TUBERCULOSIS HOSPITALS IN (A) COUNTRIES WITH A HIGH LEVEL \* AND (B) COUNTRIES WITH A LOW LEVEL\*\* OF WEAK TUBERCULIN SENSITIVITY



\* Burma, Ghana, Indonesia, Nigeria, East Pakistan, the Philippines, Sudan, Thailand, Viet Nam \*\* Basutoland, Cyprus, Denmark, England, Iraq, Lebanon, Libya, West Pakistan

in two histograms; the upper figure comprises reactions of patients in countries where the level of weak sensitivity is high, the lower figure reactions of patients in countries where the level is low. The two distributions resemble each other (and the reactor component in the size-distributions of reactions among the general populations): there is no appreciable proportion of small reactions in either. From this we can conclude that if the weak sensitizing agents are infecting micro-organisms, either they are seldom pathogenic for man or the disease they cause is seldom mistaken for tuberculosis. If they were pathogenic one would expect to find a considerable proportion of persons with weak sensitivity hospitalized for tuberculosis, at least in those areas where exposure to the weak sensitizing agents is several times heavier than exposure to pathogenic tubercle bacilli. Apart from this essential property nothing is known of the weak sensitizing agents. Whether they provoke a disease different from tuberculosis, or whether they cause a benign " tuberculous " infection with or without consequent immunity to infection with other mycobacteria, cannot be stated. It is not even certain that they are micro-organisms multiplying in the human body.

The existence of agents inducing a weak sensitivity to tuberculin may well explain the connexion observed between a slightly lower level of sensitivity among persons infected with pathogenic tubercle bacilli in those areas where weak sensitivity is widespread. Two explanations of this connexion have been offered by Bates, Busk & Palmer (1951): (a) that an initial weak sensitization (acquired by the majority of the population in tropical areas) may prevent the development of as high a degree of sensitivity as would otherwise appear from subsequent infection with pathogenic tubercle bacilli; or (b) persons infected with pathogenic tubercle bacilli may become partly desensitized through constant exposure to the weak sensitizing agents. A third possible explanation of the lower allergy level might be that part of the tuberculin injected is fixed by the cellular antibodies responsible for the weak sensitivity.

It is thus reasonable to assume that although the levels of sensitivity of the two component parts of the size-distribution of Mx <sup>5</sup> TU reactions vary among different population groups, their interpretation with regard to tuberculous infection is the same: infection with pathogenic tubercle bacilli can be inferred from strong tuberculin sensitivity

and non-infection from weak or no sensitivity. The problem of the efficacy of the test depends therefore essentially on how well the two components can be distinguished, and this, in turn, has been shown to be associated with climatic conditions.

In temperate and subtropical regions the test is effective: when a suitable reaction size is used as the limit between positive and negative reactions, only a few per cent of the infected will be wrongly classified as uninfected, and vice versa. In these areas, in fact, the tuberculin test is perhaps one of the best of biological tests at our disposal today.

In tropical regions, on the other hand, where the two component parts overlap to a large extent, the efficacy of the test is sharply reduced. For diagnostic purposes, its use is of limited value. Only from a very large or a very small reaction can a diagnosis of infection be made with some confidence; for the great bulk of persons with intermediate-sized reactions this is not possible. For epidemiological purposes the test is perhaps more useful. Although a large proportion of the population cannot be classified as infected or uninfected, an estimate of the frequency of infection can nevertheless be made by estimating the size of the right-hand component, as illustrated in Fig. 6. Such estimates are of course more or less accurate depending, again, on the level of weak sensitivity, and it must be admitted that in some countries they will be pretty uncertain.

When the test is used to select persons for BCG vaccination in tropical countries applying the same criterion for vaccination as is applied in temperate and subtropical areas, a higher proportion of error will be made: more uninfected persons will be excluded from vaccination and more infected persons will be vaccinated. The limit for vaccination may be raised for the purpose of giving BCG to more uninfected persons, but if so it will also be given to more infected. Further, there is some question of the value of vaccinating persons who have already acquired a weak sensitivity to tuberculin (Palmer, 1957). Specific studies are needed to determine whether the weak sensitivity is accompanied by an increased resistance to infection with human or bovine tubercle bacilli or whether this sensitivity is incidental and has nothing to do with the resistance mechanism.

Many problems remain, then, in the use of the tuberculin test in tropical regions, which comprise a large proportion of the world's population. Further research is indicated in two directions. Investigations should be made to identify the weak

sensitizing agent or agents and their possible contribution to host resistance in tuberculosis. And study must be continued on finding a means of improving the tuberculin test for use in tropical areas. As the weak sensitizing agents have an antigenic link to the pathogenic tubercle bacilli, a possible method would seem to be to fractionate the present tuberculin product in an attempt to remove those components common to both but retain those specific for the pathogenic tubercle bacilli. The need for an improvement of the test by this or another method is clearly demonstrated by the data presented here. Somehow the tuberculin test must be modified so as to become as effective for the populations in the vast tropical regions as it is for the populations of the temperate regions, where it was developed.

#### SUMMARY

For the tuberculin test to be effective in distinguishing between persons infected with tubercle bacilli and persons not infected, two requirements must be fulfilled: first, the test reactions must be of two distinct kinds and, secondly, the one kind must represent the infected, the other the uninfected. The present paper deals with the problem of the extent to which the first requirement is fulfilled in various populations. The data presented are very extensive, comprising tuberculin test results from about 190 000 persons in 33 countries, and they have been compiled with a high degree of uniformity in testing procedures. The study method consists in a quantitative evaluation of the pattern of tuberculin sensitivity: intradermal (Mantoux) <sup>5</sup> and <sup>100</sup> TU tests have been given to schoolchildren and general population groups, the size of the reactions has been carefully measured and recorded and the resulting distributions of reactions according to size subjected to statistical analysis.

By means of this technique it has been shown that in many countries reactions to <sup>a</sup> dose of <sup>5</sup> TU are of two distinct kinds—small reactions and large reactions. If a suitable reaction size is used for separating the two kinds only a few per cent of reactions of the one kind will be erroneously classified with the other kind and vice versa. But in many other countries there is no natural distinction between the two kinds of reaction, as reactions of intermediate size occur with appreciable frequency. A closer scrutiny of the size distributions obtained in these areas indicates that the reactions are indeed of two kinds, but that they merge considerably. There are two tendencies which bring about this merging a decreased sensitivity among the reactors and an increased sensitivity among the non-reactors. In consequence, the two kinds of reaction cannot be clearly distinguished in such populations: no matter where a point of separation is established a not inconsiderable proportion of one kind will be included among the other.

Analysis of the data by geographical location indicates that the distinction between the two kinds of reaction is usually good among persons living in temperate or subtropical climates and poor among those living in the tropics.

The paper contains no data concerning the medical

significance of the two kinds of tuberculin reaction, but other studies indicate that in general persons with the large reactions to Mx <sup>5</sup> TU are infected with pathogenic tubercle bacilli, those with the small reactions are not. This seems to hold true all over the world, regardless of how well the two kinds of reaction are distinguished from each other. The increase in the sensitivity of non-reactors in some populations is best explained by a sensitization by agents other than pathogenic tubercle bacilli, possibly other types of mycobacteria. The presence of such sensitizing agents might also explain the somewhat lower level of the strong sensitivity. The need for improving the tuberculin test for use in tropical areas is emphasized.

# RÉSUMÉ

L'épreuve à la tuberculine ne peut permettre de distinguer les sujets infectés des sujets non infectés qu'à deux conditions: les reactions qu'elle provoque doivent être de deux types nettement différenciés et chaque type doit correspondre exclusivement aux cas infectés ou aux cas non infectés. Le but de l'article est d'examiner dans quelle mesure la première condition se trouve remplie dans diverses populations. Les données présentées sont très vastes, puisqu'elles intéressent les résultats d'épreuves tuberculiniques pratiquées selon une technique uniforme sur quelque 190 000 personnes reparties dans 33 pays. La méthode adoptée a consisté à évaluer quantitativement les réactions de sensibilité à la tuberculine: à cet effet, on a pratiqué des épreuves intradermiques (Mantoux) à 5 et 100 UT sur des écoliers, ainsi que dans des groupes prélevés parmi la population générale et l'on a soigneusement mesuré et enregistré la dimension des reactions; la distribution des reactions selon leur dimension a ensuite fait <sup>l</sup>'objet d'une analyse statistique.

Cette méthode a permis de se rendre compte que dans beaucoup de pays les réactions à une dose de 5 UT sont de deux ordres - les réactions de petite dimension et les réactions de grande dimension  $-$  et si l'on choisit une dimension adequate pour operer la discrimination entre ces deux types, le risque de confusion sera minime. Dans beaucoup d'autres pays, en revanche, la distinction est beaucoup moins nette car il se produit assez souvent des reactions de dimension intermediaire. Un examen plus attentif de la distribution des dimensions observées dans ces regions montre que les reactions sont, en fait,

de deux genres, mais avec une zone de chevauchement importante. Ce chevauchement résulte de deux facteurs, à savoir une sensibilité diminuée chez les sujets positifs et une sensibilité accrue chez les sujets négatifs. La distinction devient alors difficile et, quelle que soit la ligne de démarcation adoptée, une proportion non négligeable de l'un des genres de réactions sera classée à tort dans l'autre catégorie. En analysant les données par région géographique, on s'aperçoit que, si la distinction est, en général, satisfaisante chez les habitants des zones tempérées et subtropicales, elle est très imparfaite chez ceux des zones tropicales.

L'article ne fournit aucune donnée sur la signification médicale des deux genres de réactions, mais d'autres études montrent qu'en général les personnes qui présentent les fortes réactions au Mantoux à 5 UT sont infectées par des bacilles pathogènes, tandis que celles qui font de petites réactions ne le sont pas. Il semble en etre ainsi dans le monde entier, quelle que soit <sup>l</sup>'exactitude avec laquelle la distinction est etablie entre les deux genres de réactions. Ce qui explique peut-être le mieux l'accroissement de sensibilité chez les sujets négatifs dans certaines populations est un phénomène de sensibilisation par des agents autres que des bacilles tuberculeux pathogènes, peut-être par d'autres types de mycobactéries. La présence de ces agents sensibilisants expliquerait peut-être aussi le niveau quelque peu plus bas de la forte sensibilité. L'article souligne qu'il importe d'améliorer l'épreuve à la tuberculine pour les régions tropicales.

- Bates, L. E., Busk, T. & Palmer, C. E. (1951) Publ. Hlth Rep. (Wash.), 66, 1427
- Edwards, L. B. & Palmer, C. E. (1958) Amer. J. Hyg., 68, 213
- Fisher, R. A. (1931) Truncated normal distribution. In: British Association for the Advancement of Science, Mathematical tables, vol. 1, London
- Groth-Petersen, E., Knudsen, J. & Wilbek, E. (1957) Nord. Med., 58, 1361
- Mantoux, C. (1910) Presse méd., 18, 10
- Palmer, C. E. (1953) Amer. Rev. Tuberc., 68, 678
- Palmer, C. E. (1957) Bull. int. Un. Tuberc., 27, 105
- Palmer, C. E., Shaw, L. W. & Comstock, G. W. (1958) Amer. Rev. Tuberc., 77, 877

Paretzky, M. (1937) Amer. Rev. Tuberc., 36, 313

- Sayé, L. (1936) La tuberculose pulmonaire chez les sujets apparemment sains et la vaccination antituberculeuse, Paris
- Waaler, H., Guld, J., Magnus, K. & Magnusson, M. (1958) Bull. Wld Hlth Org., 19, 783
- Weis Bentzon, J. (1953) Tubercle (Lond.), 34, 34
- WHO Tuberculosis Research Office (1955a) Bull. Wld Hlth Org., 12, 63
- WHO Tuberculosis Research Office (1955b) Bull. Wld Hlth Org., 12, 85
- WHO Tuberculosis Research Office (1955c) Bull. Wld Hlth Org., 12, 101

# APPENDIX TABLE <sup>1</sup> DISTRIBUTIONS OF REACTIONS TO THE Mx <sup>5</sup> TU TEST ACCORDING TO SIZE IN SPECIFIED



## J. NYBOE



#### APPENDIX TABLE <sup>1</sup> (continued) DISTRIBUTIONS OF REACTIONS TO THE Mx <sup>5</sup> TU TEST ACCORDING TO SIZE IN SPECIFIED AGE-GROUPS IN <sup>33</sup> COUNTRIES

## APPENDIX TABLE <sup>1</sup> (continued)

## DISTRIBUTIONS OF REACTIONS TO THE Mx <sup>5</sup> TU TEST ACCORDING TO SIZE IN SPECIFIED AGE-GROUPS IN <sup>33</sup> COUNTRIES



# 30 J. NYBOE

## APPENDIX TABLE <sup>1</sup> (continued)

## DISTRIBUTIONS OF REACTIONS TO THE Mx <sup>5</sup> TU TEST ACCORDING TO SIZE IN SPECIFIED AGE-GROUPS IN <sup>33</sup> COUNTRIES



## APPENDIX TABLE <sup>I</sup> (concluded)

## DISTRIBUTIONS OF REACTIONS TO THE Mx <sup>5</sup> TU TEST ACCORDING TO SIZE IN SPECIFIED AGE-GROUPS IN <sup>33</sup> COUNTRIES



#### J. NYBOE

#### APPENDIX TABLE <sup>2</sup>

DISTRIBUTIONS OF REACTIONS TO THE Mx <sup>100</sup> TU TEST ACCORDING TO SIZE AMONG PERSONS WITH Mx <sup>5</sup> TU REACTIONS MEASURING LESS THAN <sup>10</sup> mm\* IN SPECIFIED AGE-GROUPS IN <sup>32</sup> COUNTRIES



\* In Cambodia and Cyprus <sup>8</sup> mm, in Denmark <sup>5</sup> mm, in Egypt <sup>6</sup> mm

#### APPENDIX TABLE <sup>2</sup> (continued)

DISTRIBUTIONS OF REACTIONS TO THE Mx <sup>100</sup> TU TEST ACCORDING TO SIZE AMONG PERSONS WITH Mx <sup>5</sup> TU REACTIONS MEASURING LESS THAN <sup>10</sup> mm\* IN SPECIFIED AGE-GROUPS IN <sup>32</sup> COUNTRIES



\* In England, India and Iran <sup>8</sup> mm

#### APPENDIX TABLE <sup>2</sup> (continued)

DISTRIBUTIONS OF REACTIONS TO THE Mx <sup>100</sup> TU TEST ACCORDING TO SIZE AMONG PERSONS WITH Mx 5 TU REACTIONS MEASURING LESS THAN 10 mm\* IN SPECIFIED AGE-GROUPS IN 32 COUNTRIES



\* In Lebanon and Libya 8 mm, in Mexico 6 mm

#### APPENDIX TABLE 2 (continued)

DISTRIBUTIONS OF REACTIONS TO THE Mx <sup>100</sup> TU TEST ACCORDING TO SIZE AMONG PERSONS WITH Mx 5 TU REACTIONS MEASURING LESS THAN 10 mm\* IN SPECIFIED AGE-GROUPS IN 32 COUNTRIES



\* In the Philippines and Sudan <sup>8</sup> mm

## APPENDIX TABLE <sup>2</sup> (concluded)

## DISTRIBUTIONS OF REACTIONS TO THE Mx <sup>100</sup> TU TEST ACCORDING TO SIZE AMONG PERSONS WITH Mx <sup>5</sup> TU REACTIONS MEASURING LESS THAN <sup>10</sup> mm\* IN SPECIFIED AGE-GROUPS IN <sup>32</sup> COUNTRIES



\* In Thailand and Viet Nam <sup>8</sup> mm, in Turkey <sup>5</sup> mm

#### APPENDIX TABLE <sup>3</sup>

## CORRELATION BETWEEN SIZES OF REACTIONS TO TWO Mx <sup>5</sup> TU TESTS GIVEN AT AN INTERVAL OF 2-3 MONTHS AMONG SCHOOLCHILDREN IN CYPRUS, IRAN AND LEBANON

