Tubercle and Lung Disease (1995) **76**, 114–121

Tubercle and Lung Disease

Methodological issues in the estimation of the tuberculosis problem from tuberculin surveys

H. L. Rieder

Tuberculosis Section of the International Union Against Tuberculosis and Lung Disease, Paris, France

S U M M A R Y. ***Setting:*** **National tuberculin skin test surveys.**
Objectives: **To review the operating characteristics of the tuberculin skin test, to ascertain the validity of estimating prevalence and risk of infection from tuberculin skin test surveys under various conditions, and to review constraints in the estimation of the magnitude of the tuberculosis problem in the community from such surveys.**
Methods: **This report utilizes hypothetical and selected real data obtained in regional and national surveys at various points in time to exemplify methodological issues.**
Results: **Risk of infection, the essence to be abstracted from tuberculin skin test surveys, theoretically allows for a comparison of the extent of transmission of tubercle bacilli in various populations. However, the conduct of tuberculin skin test surveys and the analysis and interpretation of their results are not free from important technical problems. Accurate estimation of infection prevalence is particularly vulnerable to the great variability of the test's specificity under various circumstances. Furthermore, the annual risk of infection has averaging characteristics that preclude a rapid assessment of changes in transmission patterns. Finally, estimates of infection risk do not necessarily provide a standardized parameter to derive incidence of infectious cases, because of variations in the quality of intervention and varying risks of progression from latent infection to overt tuberculosis.**
Conclusions: **While tuberculin skin test surveys provide the currently most widely used means of assessing tuberculosis transmission patterns over prolonged periods of time in a community, results from such surveys must be interpreted with caution when accurate estimates of the tuberculosis problem are sought.**

R É S U M É. ***Cadre:*** **Enquêtes tuberculiniques nationales.**
Objet: **Examen des caractéristiques opératoires du test tuberculinique cutané, évaluation de la validité du calcul de la prévalence et du risque d'infection à partir des enquêtes tuberculiniques sous diverses conditions, et finalement revue des contraintes rencontrées dans le calcul de l'ampleur du problème tuberculeux dans la communauté à partir de telles enquêtes.**
Schéma: **Ce rapport utilise des données hypothétiques ainsi que des données réelles sélectionnées à partir des résultats des enquêtes régionales et nationales effectuées à différentes époques, afin d'illustrer certains problèmes méthodologiques.**
Résultats: **Le risque d'infection, résultat essentiel des enquêtes tuberculiniques, permet en théorie de comparer l'importance de la transmission des bacilles tuberculeux dans différentes populations. Cependant la réalisation des enquêtes tuberculiniques ainsi que l'analyse et l'interprétation des résultats n'échappent pas à des problèmes techniques importants. Le calcul précis de la prévalence de l'infection est particulièrement soumis aux variations importantes de la spécificité du test dans différentes situations. De plus le risque annuel d'infection reflète une 'moyenne' des évènements passés qui empêche toute évaluation rapide de modifications dans les données de transmission. Finalement, les calculs du risque d'infection ne fournissent pas toujours un paramètre standardisé dont on puisse en extraire l'incidence des cas infectieux, en raison des variations liées à la qualité de l'intervention et à la variabilité du risque de progression de l'infection latente vers une tuberculose maladie.**
Conclusion: **Bien que les enquêtes tuberculiniques fournissent la méthode la plus utilisée à l'heure actuelle pour évaluer les caractéristiques de transmission de la tuberculose à travers de longues périodes dans une**

Correspondence to: Dr Hans L. Rieder, Chief, Tuberculosis Section of the IUATLD, Reichenbachstr. 15, 3004 Bern, Switzerland.

Paper received 11 January 1994. Final version accepted 9 August 1994.

communauté, on devrait interpréter les résultats de telles enquêtes avec prudence quand il s'agit d'obtenir une estimation précise du problème tuberculeux.

R E S U M E N. Marco de referencia: **Encuestas tuberculínicas nacionales.**
Objetivos: **Revisar las características operativas del test cutáneo de tuberculina, determinar la validez de la estimación de la prevalencia y del riesgo de infección a partir de las encuestas tuberculínicas bajo diversas condiciones, y revisar las dificultades en la estimación de la magnitud del problema de la tuberculosis en la comunidad a partir de tales encuestas.**
Método: **Este informe utiliza datos hipotéticos y una selección de datos reales obtenidos de encuestas regionales y nacionales en diversos lugares y períodos a fin de ejemplarizar los problemas metodológicos.**
Resultados: **El riesgo de infección, resultado esencial de las encuestas tuberculínicas, teóricamente permite la comparación de la extensión de la transmisión del bacilo tuberculoso en las diversas poblaciones. Sin embargo, la conducción de las encuestas tuberculínicas y el análisis e interpretación de sus resultados no están exentos de importantes problemas técnicos.**

La estimación precisa de la prevalencia de la infección es particularmente vulnerable a la especificidad del test bajo diversas condiciones. Además, el riesgo anual de infección representa un promedio a un momento dado, lo que impide una rápida evaluación de los cambios en los patrones de transmisión. Finalmente, la estimación del riesgo de infección no proporciona necesariamente un parámetro estándar para calcular la incidencia de los casos infecciosos, a causa de las variaciones de la calidad de la intervención y de la variabilidad del riesgo de progresión de la infección latente a la enfermedad tuberculosa.
Conclusiones: **A pesar que las encuestas tuberculínicas constituyen el medio más ampliamente utilizado para evaluar los patrones de transmisión de la tuberculosis en períodos prolongados en una comunidad, los resultados de cada encuesta deben ser analizados con precaución cuando se requiere una estimación precisa del problema de la tuberculosis en una comunidad.**

INTRODUCTION

The most desirable method of ascertaining the current extent of transmission of tubercle bacilli in a society would be to measure the incidence of infection with *Mycobacterium tuberculosis* in susceptible persons. Measuring the incidence of infection is, however, a Herculean task. It requires repeat testing of a large enough number of persons to identify with reasonable precision the few who become newly infected over a specified period of time. Furthermore, the incidence of infection varies across various subpopulations, e.g. various age groups in the same calendar year,[1] and when estimates are based on repeat testing of the same individual both boosting reactions[2] and reversion of initially positive tuberculin skin test reactions[3] may greatly affect the accuracy of the estimates. Some of these problems can be partially overcome through approximation by calculating the average probability that a person has become infected over a specified period of time from a tuberculin skin test prevalence survey in a population which has not been subjected to previous tuberculin skin testing or vaccination with BCG. Detailed technical guidelines on how to conduct such a survey have been published.[4,5]

The purpose of this paper is to highlight methodological issues relevant to the analysis and interpretation of tuberculin skin test survey data and their relation to the tuberculosis problem in the community. Many of the problems have been recognized for quite some time,[6] while other aspects challenge existing concepts and are addressed here to further the discussion on how to critically appraise the meaning of data generated in tuberculin skin test surveys.

TEST CHARACTERISTICS

Each test has its own intrinsic operating characteristics. The sensitivity of the test is the proportion accurately identified by a positive test result among persons with a characteristic in question, while the specificity represents the proportion with a negative test result among persons without that characteristic.[7] A multitude of tuberculin skin test techniques has been developed in this century, but the Mantoux technique is the most quantifiable and, in conjunction with other information, theoretically allows the determination of the sensitivity and specificity of the test at different cut-off points.

In the early 1950s the World Health Organization (WHO) collected information on tuberculin sensitivity in over 3600 hospitalized tuberculosis patients in 10 different countries.[8] The combined results closely fit a normal distribution with a mode at 16–17 mm. The sensitivity in identifying infection with *M. tuberculosis* calculated from this survey is 93% for an induration $\geq$ 10 mm and 78% for an induration $\geq$ 14 mm. Similar normal distributions were found in healthy persons in areas with a very low frequency of small reactions[6,8,9] and among healthy United States Navy recruits who had a history of exposure to tuberculosis.[10] In the latter survey, a mode was identified at 18–19 mm and the calculated sensitivities for $\geq$ 10 mm and $\geq$ 14 mm were 94% and 75% respectively (Table 1).

Thus, the sensitivity of the tuberculin skin test appears to vary relatively little when different populations are compared, except when cellular immunity is seriously compromised. However, as the tuberculin surveys sponsored by the WHO demonstrate,[9] sensitization to envi-

Table 1. Sensitivity and specificity of the tuberculin skin test in United States Navy recruits, using two commonly used cut-off points as indicating the presence of infection (calculated from Tables 1 and 3)[10]

Test criterion	Sensitivity (%)	Specificity (%)
≥ 10 mm	94.2	98.6
≥ 14 mm	74.9	99.5

ronmental mycobacteria apparently varies considerably in different countries, making the specificity of the test unpredictable, and thus the underlying normal distribution attributable to tuberculous infection difficult to ascertain.

Figure 1 shows a hypothetical distribution of tuberculin skin test reaction sizes. Three distributions influence the composite picture that can be ascertained; (1) the distribution among persons without any mycobacterial infection or with skin test anergy, usually not exceeding a few millimeters (thick dotted line); (2) the distribution among persons with tuberculous infection (thin dotted line, shown here with a mode at 17 mm); and (3) the distribution among persons sensitized to mycobacteria other than tubercle bacilli who cross-react with tuberculin PPD (dashed line shown here with a mode at 7 mm). It is apparent that in this example at a cut-off point of 14 mm some true infections will be missed (*c*) and some will be falsely counted (*b*), the latter because they are attributable to non-specific sensitization. In fact, as shown in this example, this observation pertains, albeit to a different degree, to any cut-off point to indicate presence or absence of infection with *M. tuberculosis* between 4 mm and 21 mm. It is thus apparent that cut-off points are compromises to balance false positive and false negative reactions to still obtain the right prevalence, although individuals themselves are misdiagnosed. Tuberculin prevalence surveys in numerous countries have clearly shown that sensitization to environmental mycobacteria is virtually absent in some, but abundant in other countries, and that the relative magnitude of the distribution due to environmental mycobacteria (and/or sensitization due to BCG vaccination) and the magnitude of the distribution due to tubercle bacilli may vary considerably, even within the same country.[6,9–18] Vaccination with BCG itself appears to induce cross-reactivity with tuberculin PPD to a variable extent.[19–23]

Rust and Thomas have developed a model based on tuberculin skin test data in 700 000 United States Navy recruits.[10] By asking about known history of exposure to a tuberculosis case in the same household they separated people into contacts and non-contacts. This allowed a separation of the influences of reactions caused by environmental mycobacteria from those resulting from infection with tubercle bacilli and the development of a model predicting probabilities of tuberculous infection at various cut-off points. These data thus allow the calculation of the specificity of the tuberculin skin test using various cut-off points in the setting of the United States (Table 1).

Even if it is assumed that the frequency of cross-sensitivity reactions is constant (which is clearly not the case),[6,9,12,13] and therefore that the specificity of the tuberculin skin test is unchanged and the information from Rust and Thomas applicable, the predictive value of a significant tuberculin test reaction (*a/(a+b)* in Fig 1) would be extremely sensitive to variations in prevalence of infection with tubercle bacilli (Table 2). If the inci-

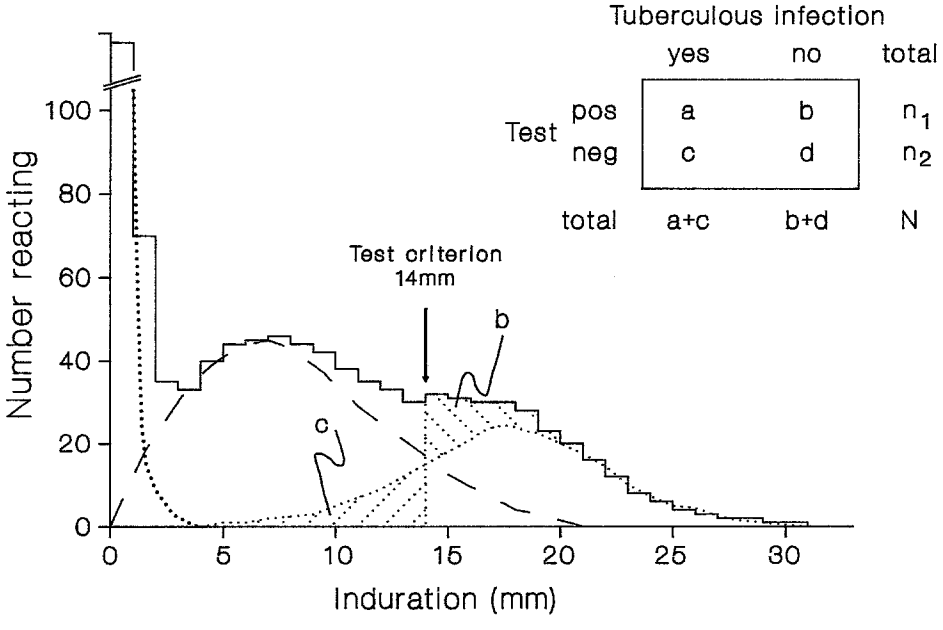

Fig. 1 Example of the distribution of diameters in a tuberculin skin test survey. The distribution of non-specific reactions is shown under the dashed curve, the specific distribution under the dotted curve. The hatched area, labelled *b*, indicates the number of reactions falsely counted as infection using a criterion of ≥ 14 mm to indicate infection, the hatched area, labelled *c*, indicates the infections missed by that criterion. The number of positive reactions (n_1) is found to the left of the test criterion; the number of negative reactions (n_2) is found to its right.

Table 2. Predictive value of a positive test using ≥ 14 mm as indicating infection, assuming (from Table 1) a specificity of 98.6% and a sensitivity of 94.2% with a cut-off of ≥ 10 mm, and specificity and sensitivity of 99.5% and 74.9% respectively using a cut-off of ≥ 14 mm in a population with a prevalence of infection of 0.28% and 10.0% respectively

Prevalence (%)	Predictive value of a positive test result (%) Criterion ≥ 10 mm	Criterion ≥ 14 mm
0.28	15.9	29.6
10.0	88.2	94.3

dence of infection is assumed to be constant at 1% over calendar time and infection risk independent of age, then the expected true prevalence of infection with *M. tuberculosis* in children aged 10.5 years will be 10%. In an industrialized country, where the current risk of infection might be 0.015%, and the decline in risk of infection has been in the order of 10% per year, the expected prevalence of infection in children of the same age is 0.28%. Using these examples, in a country with a low prevalence of infection, a cut-off point of 10 mm indicating infection with *M. tuberculosis* would falsely classify 84.1% of children thus identified as 'infected', with a reduction in mis-classifications to only 70.4% using a cut-off point of 14 mm and more. In a country with a high infection prevalence, the test would falsely classify only 11.8% or 5.7%, using 10 mm or 14 mm induration respectively to indicate infection. It is thus apparent that tuberculin skin test surveys in low infection prevalence countries will almost always preclude meaningful interpretation, unless environmental mycobacteria are of such little importance that the specificity of the test approximates 100% even at low cut-off points. It seems nevertheless that information may be useful in countries with an elevated prevalence of infection.

The commonly used cut-off points of 10 mm and 14 mm, the latter customarily corrected by dividing by 0.82,[11,24] take the sensitivity into account, assuming that with a 10 mm cut-off point virtually all true infections are included, but only about 82% with a cut-off point of 14 mm. The sensitivity of 82% for a cut-off of ≥ 14 mm to indicate infection is a slight overestimation, because in the tuberculin survey in Tanganyika that produced this figure it was assumed that all reactions ≥ 10 mm were 100% specific in two areas with relatively few cross-reactions, which was nevertheless not quite the case.[11] Applied to areas with any non-specific cross-reactions, these criteria have the disadvantage that they do not account for the loss of specificity, a loss that increases the further one moves from the mode of the distribution from tuberculous infection towards the left into increasing contamination by non-specific reactions. Thus, the correction for sensitivity alone will invariably overestimate the prevalence. This is shown in Figure 1, where, to enumerate the infected among those with ≥ 14 mm induration, instead of using $(a+c)$, $(n_1)/[a/(a+c)]$ or $(n_1)/(0.82)$ is erroneously calculated by this technique. Unfortunately, without *a priori* knowledge of both sensitivity and specificity, the proportion of mis-classifications can not be known.

Another technique to estimate prevalence assumes an underlying normal distribution of reactions due to infection with *M. tuberculosis*. This mirror technique attempts to identify the mode, multiplies the number of reactors above the mode by two and adds the number of reactors at the mode to arrive at the number of infected persons. Although the mirror technique partially circumvents the problem of test specificity, the precision of the estimate is by necessity poorer and requires the testing of a much larger number of persons to improve the estimate. Furthermore, using the mirror technique approach, the calculation of infection prevalence is very sensitive to the selection of the location of the mode. If, for example, in Figure 1 the mode is selected to be at 17 mm then the calculated prevalence of infection is 13.7%; if the mode is selected to be at 18 mm, the estimated prevalence of infection is 10.9%, or 20% less.

If the sensitivity (denoted as x) and the specificity (denoted as y) are both known for a certain cut-off point then the calculation of the prevalence is easily done. There are four unknowns (a, b, c and d, see Fig. 1) that can be solved with the four following equations, because n_1 (denoted as the number with a positive test result, i.e. $a+b$) and n_2 (denoted as the number with a negative result, i.e., $c+d$) are defined:

$$a = n_1 - b \qquad \text{eq. (1)}$$

$$b = d/y - d \qquad \text{eq. (2)}$$

$$c = a/x - a \qquad \text{eq. (3)}$$

$$d = n_2 - c \qquad \text{eq. (4)}$$

These equations can be used to solve, e.g., for a:

$$a = \frac{(n_1 + n_2)\,xy - n_2 x}{(x + y - 1)}$$

or any other cell.

EXAMPLES OF TUBERCULIN SKIN TEST SURVEYS

A large tuberculin skin test survey is being carried out in Tanzania under the auspices of the Tuberculosis Surveillance Research Unit of the International Union Against Tuberculosis and Lung Disease. This survey is conducted in 5-year cycles; it encompasses the entire country and is carried out by professional staff, trained by the International Tuberculosis Surveillance Centre. Figures 2 and 3 show the survey data from 1991 from 3 regions (chosen for convenience) combined (Dodoma, Mbeya, and Morogoro) of children with and without BCG scar respectively (reproduced with the permission of the Tanzania National Tuberculosis/Leprosy Programme).[25] Clearly, neither distribution allows rapid identification of the proportion of the proportion infected with tubercle bacilli.

In Figure 3, a mode is proposed (arbitrarily) at 18 mm, allowing for construction of a mirror image of the suspected underlying distribution. Although the mirror image technique accounts for the loss in sensitivity at this

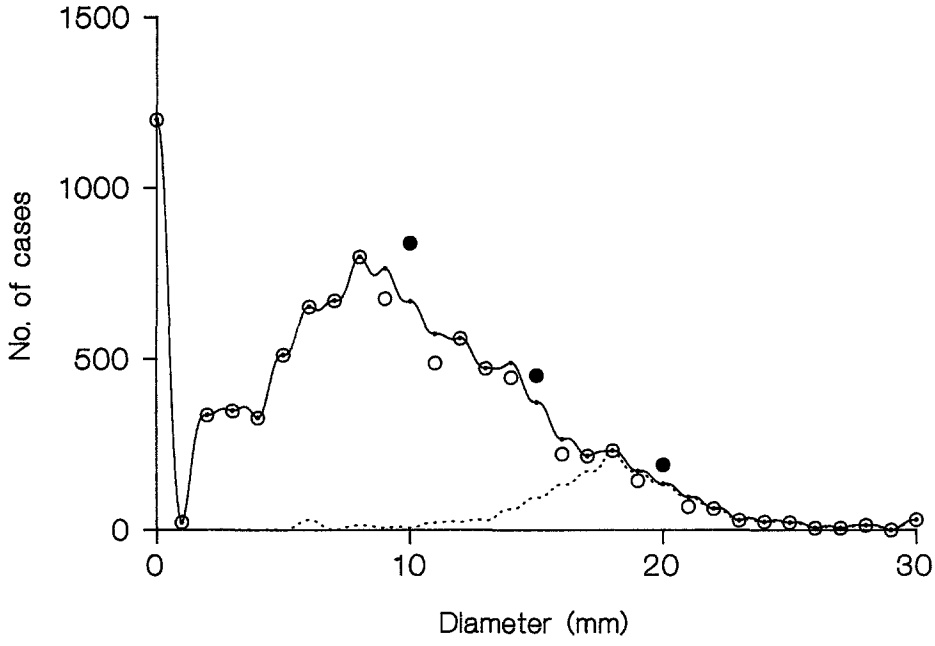

Fig. 2 Measured distribution of tuberculin skin test reaction size diameters in Tanzania 1991 (Dodoma, Mbeya, and Morogoro regions) among children with BCG scar. Large circles are recorded data, filled circles emphasize digit preferences at 10 mm, 15 mm and 20 mm. Points are values linearly adjusted for digit preference (one neighbour only to each side). Line is fitted through adjusted values by inverse squared distance smoothing. Area under dotted line indicates probable true infection with *Mycobacterium tuberculosis*. For better display, the number with 0 mm reaction (10 240 children) has been cut off. (reproduced with the permission of the Tanzania National Tuberculosis/Leprosy Programme).[25]

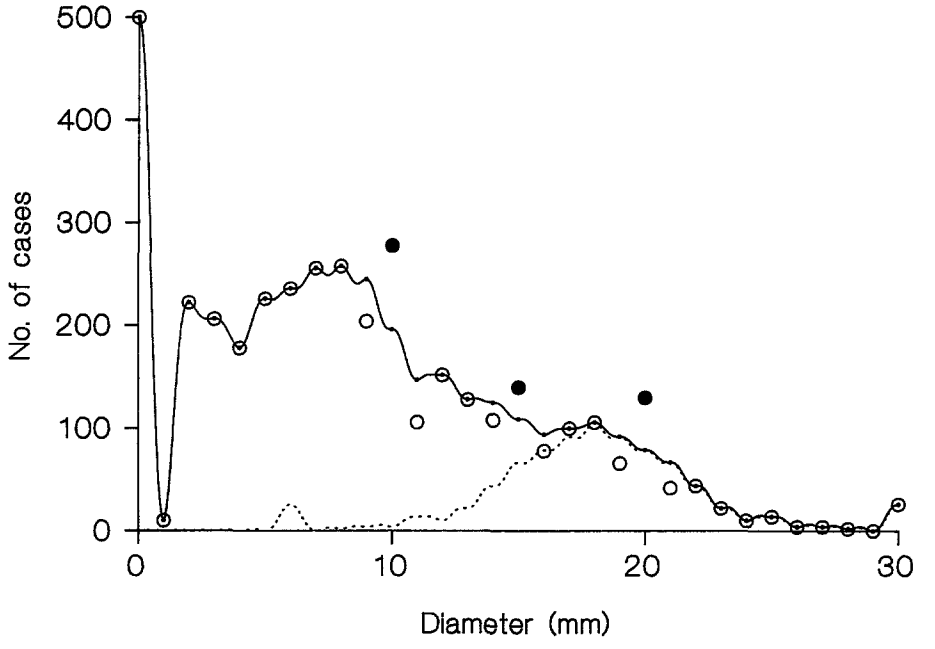

Fig. 3 Measured distribution of tuberculin skin test reaction size diameters in Tanzania 1991 (Dodoma, Mbeya, and Morogoro regions) among children with no apparent BCG scar. Large circles are recorded data, filled circles emphasize digit preferences at 10 mm, 15 mm and 20 mm. Points are values linearly adjusted for digit preference (one neighbour only to each side). Line is fitted through adjusted values by inverse squared distance smoothing. Area under dotted line indicates probable true infection with *M. tuberculosis*. For better display, the number with 0 mm reaction (7210 children) has been cut off. (reproduced with the permission of the Tanzania National Tuberculosis Leprosy Programme).[25]

diameter, the figure suggests that even at this diameter some reactors are still likely to be falsely classified as being infected. Here, the mode was assumed to be at 18 mm and the calculated infection prevalence was 7.9%. Had a mode been selected at 17 mm, the prevalence calculated with the same technique would be 9.8%, 24% higher than with a mode at 18 mm.

Figures 2 and 3 also demonstrate that even experienced readers clearly have a preference for certain digits (shown as full circles at 10 mm, 15 mm, and 20 mm). It would thus be difficult to use a cut-off point of 20 mm or an immediately neighboring value and account algebraically for the loss of sensitivity for the sake of gaining specificity.

Figure 4 shows the results of two tuberculin surveys in Korea conducted 25 years apart with the same technique utilizing 1 TU PPD RT23 in children aged 0–9 years.[26] It demonstrates that a large decrease in infection prevalence fundamentally changes the interpretability of a tuberculin survey in a country even if it has a relatively small contribution from sensitization with environmental mycobacteria. In 1965, the relative contribution of non-specific sensitization was negligible and did not preclude a clear dichotomization between those infected with tubercle bacilli and those not. While the absolute magnitude of non-specific reactions in 1990 was apparently similar to that in 1965, their relative contribution to the overall distribution had become very important by 1990, because true prevalence had declined to very low levels, making it exceedingly difficult to separate the infected from the non-infected. The ratio of reactions attributable to environmental mycobacteria to those resulting from infection with *M. tuberculosis* had inverted to an extent that the mode had shifted to the left (Fig. 4).

CALCULATING THE RISK OF INFECTION FROM PREVALENCE DATA

Assuming that the prevalence of infection with *M. tuberculosis* has been satisfactorily estimated, the essence to be extracted from the data is the estimation of the average annual risk of infection. The annual risk of infection refers to a risk at a specified calendar time $b+x$, where b indicates the calendar time at which the cohort in the survey was born and x is a number between 0 and a, where a is the age of the cohort at calendar time $b+a$, the time when the survey was conducted. It cannot be known at exactly what calendar time this risk existed without inferences from serial surveys.[1,27] Because the risk may change over calendar time, x has been approximated to lie at the midpoint between the year the cohort was born and the year the survey was conducted, if data from a single survey only are available:[1,28]

$$R_{b+a/2} \approx 1 - (1 - P_{b+a})^{1/a},$$

where $R_{b+a/2}$ denotes the annual risk of infection at the midpoint in calendar time between the year the cohort was born and the year of the survey, and P_{b+a} the prevalence of infection at the time of the survey, where both risk and prevalence are expressed as fractions. Thus, if the prevalence of infection among 10.5-year-old children is found to be 10% at the midpoint of the survey (assumed to be in, for example, the end of June 1993, i.e. 1993.5) then the risk of infection is:

$$R_{1988.25} \approx 1 - (1 - 0.1)^{1/10.5} = 0.010,$$

i.e. 1% at the approximated calendar time the end of March 1988. The estimate of $b+x$ can be improved from the first approximation of $b+a/2$ only if serial surveys are available.[27] Serial estimates, only if closely fitting a calculated regression, allow extrapolation to current infection risk.

Two sequential surveys alone a few years apart will not necessarily provide information on the change in infection risk because of the problems associated with the comparison of cross-sectional data across time. If it is assumed that an earlier survey, e.g. conducted in 1988, had also provided an estimated average annual risk of infection of 1% (approximated at calendar time 1983), then the conclusion is not necessarily warranted that the risk of infection has remained unchanged over calendar time up to the time of the second survey in 1993. It may well be that the risk decreased in the first years after the birth of the second cohort, perhaps because an efficient programme for identifying and curing cases spreading infection was implemented, but subsequently the number of infectious cases began to increase because

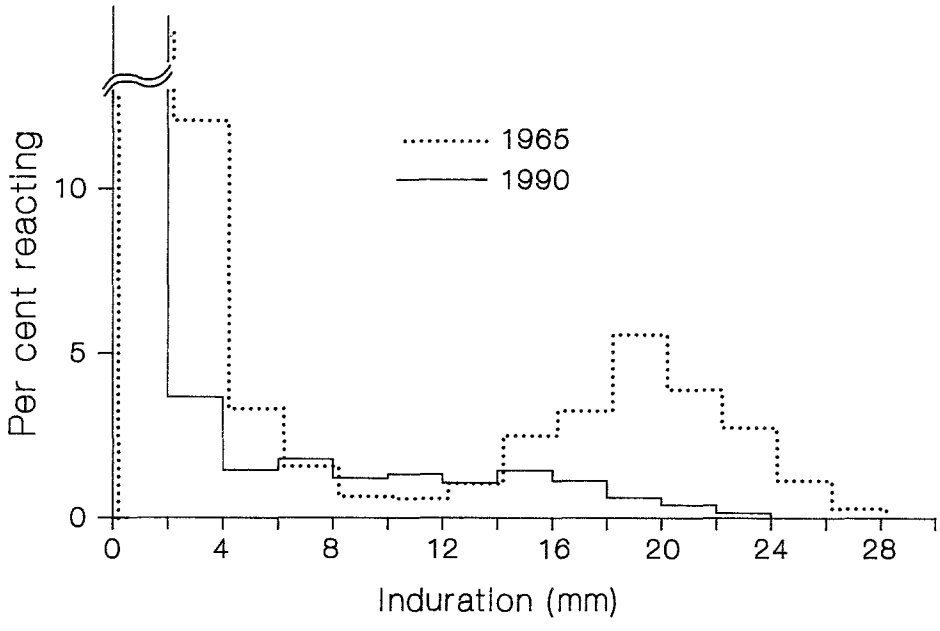

Fig. 4 Frequency distribution of tuberculin skin test reaction sizes among children aged under 10 years, Korea 1965 and 1990. Reproduced with permission.[26]

of increasing prevalence of infection with the human immunodeficiency virus (HIV) among persons with tuberculous infection, leading to an increasing number of transmitters and thus an increased risk of infection in the community. It may just be that the net effect of an initial decline, followed by an increase in infection risk, resulted in no change in infection prevalence.

RISK OF INFECTION AND INFECTIOUS CASES

It is apparent that the rate of transmission of tubercle bacilli is related to the number of sources of infection in a society. The number of successful transmissions by infectious cases to a susceptible population over a defined period of time (usually one year) at a certain calendar time defines the risk of infection in the community during that period. Utilizing the results from a series of WHO sponsored surveys in low income countries and the results of surveys in the Netherlands before chemotherapy, Murray et al correlated annual risk of infection to incidence of tuberculosis by linear regression and obtained an estimate of approximately 50 incident sputum smear-positive cases for each 1% of annual risk of infection.[29] This correlation must be interpreted with caution. The authors performed least squared regression and were thus unable to account for changes in the variance with sample size. More importantly, the authors assume that for areas where information on the incidence was lacking, the incidence was half the prevalence.[29] This is, however, precisely the hypothesis that has to be proven, if a relation between infection risk and incidence rather than person-time of infectiousness is to be demonstrated.

Infection risk is intrinsically coupled to duration of undiagnosed, untreated transmissible tuberculosis, thus with person-time of infectiousness in the community. Intervention with chemotherapy has as its epidemiological aim to reduce the rate of transmission, and where this form of intervention encompasses effectively and efficiently a large proportion of the population, the average duration of infectiousness connecting prevalence and incidence becomes fundamentally changed. Prevalence of infectious tuberculosis thus might correlate better with infection risk than incidence. Nevertheless, it has been pointed out that in countries without a structured programme, the number of infectious (sputum smear-positive) patients remains essentially the same after 2 years with or without such intervention, because the main gain with such intervention lies with a reduction of case fatality at the expense of keeping infectious cases alive.[30] Conversely, in countries where intervention effectively cuts the chain of transmission, the number of transmissions caused by one case will be reduced. Thus, to produce a 1% risk of infection, a larger number of incident cases is required, because the person-time of infectiousness is reduced. This has been shown, for example, for the United States before HIV noticeably affected tuberculosis. In that country, extrapolation would have required some 400 incident cases per 100 000 population to result in 1% risk of infection in the early 1980s,[31] some 8 times the number predicted by the model outlined above.[29]

Furthermore, the risk of tuberculosis following infection with *M. tuberculosis* may vary in different populations. It is certainly increased in persons with HIV infection compared to immunocompetent hosts. Thus, the epidemiological balance usually observed between host and bacillus is no longer present under these circumstances where each case of tuberculosis may produce more than one new infectious case in the HIV-infected segment of the population.

SUMMARY AND CONCLUSIONS

Theoretically, the incidence of infection is epidemiologically the most informative parameter, because it identifies the extent of current transmission in the community. It is usually not feasible to measure infection incidence, and the derivation of the average annual risk of infection from a tuberculin prevalence survey as a proxy has become one of the most cherished tools in tuberculosis epidemiology. Unfortunately, tuberculin skin testing is fraught with problems of a technical nature, including selection of standardized tuberculin, the technique of administration, and reading of the test result. Even if all of these barriers are overcome, it is in many circumstances exceedingly difficult to arrive at an estimate of the prevalence of infection. Sensitization to environmental mycobacteria and *M. bovis* BCG results in cross-reactions with the standard tuberculin. The higher this sensitization is and the lower the prevalence of true infection with tubercle bacilli, the more difficult it becomes to disentangle the truth from confounding factors. It is clear that techniques must be developed to address the problems of interpretation that arise in so many countries which have completed a tuberculin skin test survey. It remains to be determined whether simultaneous testing with different antigens, algebraic manipulation or other approaches can help to overcome some of the apparent shortcomings of tuberculin skin test surveys. Because sensitivity is already largely known for various cut-off points, it appears that the most promising approach would lie with an attempt to determine the specificity of the tuberculin test at a given cut-off point in a country planning a tuberculin skin test survey using different antigens.[14–18] This would help greatly in improving the validity of the results of a tuberculin skin test survey.

The calculation of the risk of infection (should the determination of infection prevalence be successful) from a single or even two sequential surveys provides only information on the extent of transmission at some point in the past, determined by the age of the children that have been tested. The tool is not sensitive to short-term changes, because of its 'averaging' characteristic.

The knowledge of risk of infection cannot precisely

provide information on the number of expected incident cases of tuberculosis; it can only state to what extent such cases are capable of transmitting tubercle bacilli within the community, which is a function of the number of infectious cases, the number of case-contact interactions, the duration of infectivity, and characteristics of exposure.

The determination of the risk of infection has nevertheless been regularly used to compare the extent of the tuberculosis problem in various populations. It is, if technically interpretable, the only available means of measuring the extent of transmission that has occurred, on average, over specified periods of time in the past. The common underlying technical problems will often undermine the precision in estimating the size of the tuberculosis problem in a community from a single survey. It can be useful for global estimates of the level of the tuberculosis problem in a community and of trends over relatively long periods of time. The observation of trends in prevalence or risk of infection over time is by far more informative than a single survey, because the change in slope might be freer from bias than the level of the intercept. A consistent recession of age prevalence curves, as observed for example in the 6 five-yearly Korean prevalence surveys,[26] is enough to convince that the risk of infection has been declining. However, any formal estimation of a change in the risk of infection within a certain precision may be impossible to assure.

To move forward in gaining a better understanding of the intricacies in estimating the tuberculosis problem in a community from tuberculin skin test survey data requires a concerted effort on the part of researchers in the field to address the various issues and problems encountered in the conduct of such surveys and the interpretation of their results.

Acknowledgments

Numerous people have provided input into this article. Particularly acknowledged are the comments made by Nancy Binkin, George M. Cauthen, Donald A. Enarson, Lawrence J. Geiter, and Dixie E. Snider, Jr.

References

1. Sutherland I. Recent studies in the epidemiology of tuberculosis, based on the risk of being infected with tubercle bacilli. Adv Tuber Res 1976; 19: 1–63.
2. Narain R. Interpretation of the repeat tuberculin test. Tubercle 1968; 49: 92–103.
3. Aronson J D. The fluctuation of the tuberculin reaction in different geographic areas and its relationship to resistance. Am Rev Tuberc 1951; 63: 121–139.
4. Deck F, Guld J. Committee on Epidemiology and Statistics. The WHO tuberculin test. Bull Int Union Tuberc 1964; 34: 53–70.
5. Bleiker M A, Sutherland I, Styblo K, ten Dam H G, Misljenovic O. Guidelines for estimating the risks of tuberculous infection from tuberculin test results in a representative sample of children. Bull Int Union Tuberc Lung Dis 1989; 64: 7–12.
6. Edwards P Q, Edwards L B. Story of the tuberculin test. From an epidemiologic viewpoint. Am Rev Respir Dis 1960; 81: 1–47.
7. Last J M. A dictionary of epidemiology, 2nd edn. New York: Oxford University Press, 1988.
8. WHO Tuberculosis Research Office. Further studies of geographic variation in naturally acquired tuberculin sensitivity. Bull World Health Organ 1955; 22: 63–83.
9. Roelsgaard E, Iversen E, Bløcher C. Tuberculosis in tropical Africa. An epidemiological study. Bull World Health Organ 1964; 30: 459–518.
10. Rust P, Thomas J. A method for estimating the prevalence of tuberculous infection. Am J Epidemiol 1975; 101: 311–322.
11. World Health Organization. Tuberculosis survey in Tanganyika. Copenhagen: WHO Research Office, 1958.
12. ten Dam H G. Surveillance of tuberculosis by means of tuberculin surveys. Geneva: World Health Organization, WHO/TB/85.145, 1985.
13. Edwards L B, Acquaviva F A, Livesay V T, Cross F W, Palmer E E. An atlas of sensitivity to tuberculin, PPD-B, and histoplasmin in the United States. Am Rev Respir Dis 1969; 99: 1–132.
14. Lind A, Larsson L O, Bentzon M W et al. Sensitivity to sensitins and tuberculin in Swedish children. 1. A study of schoolchildren in an urban area. Tubercle 1991; 72: 29–36.
15. Larsson L O, Skoogh B E, Bentzon M W et al. Sensitivity to sensitins and tuberculin in Swedish children. 2. A study of preschool children. Tubercle 1991; 72: 37–42.
16. Larsson L O, Skoogh B E, Bentzon M W, Magnusson M, Olofson J, Lind A. Sensitivity to sensitins and tuberculin in Swedish children. 2. Sequential versus simultaneous skin testing. Tubercle 1991; 72: 187–189.
17. Larsson L O, Magnusson M, Skoogh B E, Lind A. Sensitivity to sensitins and tuberculin in Swedish children. 4. The influence of BCG vaccination. Eur Respir J 1992; 5: 584–586.
18. Larsson L O, Bentzon M W, Lind et al. Sensitivity to sensitins and tuberculin in Swedish children. 5. A study of school children in an inland rural area. Tubercle Lung Dis 1993; 74: 371–376.
19. Abrahams E W. Tuberculin hypersensitivity following BCG vaccination in Brisbane school children. Tubercle 1979; 109–113.
20. Tuberculosis Prevention Trial, Madras. Trial of BCG vaccines in south India for tuberculosis prevention. Indian J Med Res 1980; 72 (suppl): 1–74.
21. Bahr G M, Chugh T D, Behbehani K et al. Unexpected findings amongst the skin test responses to mycobacteria of BCG vaccinated Kuwaiti school children. Tubercle 1987; 68: 105–112.
22. Menzies R, Vissandjee B. Effect of Bacille Calmette–Guérin vaccination on tuberculin reactivity. Am Rev Respir Dis 1992; 145: 621–625.
23. Menzies R, Vissandjee B, Amyot D. Factors associated with tuberculin reactivity among the foreign-born in Montreal. Am Rev Respir Dis 1992; 146: 752–756.
24. Styblo K. The first round of the National Tuberculin Survey in Tanzania, 1983–1987. Tuberculosis Surveillance Research Unit, Progress Report 1989, Volume 2. Paris: International Union Against Tuberculosis and Lung Disease, 1989: pp 101–116.
25. International Union Against Tuberculosis and Lung Disease. Tanzania National Tuberculosis/Leprosy Programme. Progress Report No. 27, Paris, April 1992.
26. Hong Y P, Kim S J, Kwon D W, Chang S C, Lew W J, Han Y C. The sixth nationwide tuberculosis survey in Korea, 1990. Tubercle Lung Dis 1993; 74: 323–331.
27. Styblo K, Meijer J, Sutherland I. The transmission of tubercle bacilli. Its trend in a human population. Tuberculosis Surveillance Research Unit Report No. 1. Bull Int Union Tuberc 1969; 42: 1–104.
28. Cauthen G M, Pio A, ten Dam H G. Annual risk of infection. World Health Organization, WHO/TB/88.154, 1988.
29. Murray C J L, Styblo K, Rouillon A. Tuberculosis in developing countries: burden, intervention and cost. Bull Int Union Tuberc Lung Dis 1990; 65: 6–24.
30. Grzybowski S, Enarson D A. The fate of cases of pulmonary tuberculosis under various treatment programmes. Bull Int Union Tuberc 1978; 53: 70–75.
31. Cauthen G M, Rieder H L, Geiter L J. A model of the relation between age-specific prevalence of tuberculous infection and incidence of infectious tuberculosis: implications for screening policies. Tuberculosis Surveillance Research Unit of the IUATLD. Progress Report 1991, Volume 1. The Hague: Royal Netherlands Tuberculosis Association, 1991: pp 1–20.