

Serological tests in leprosy. The sensitivity, specificity and predictive value of ELISA tests based on phenolic glycolipid antigens, and the implications for their use in epidemiological studies

BY P. J. BURGESS*, P. E. M. FINE*, J. M. PONNIGHAUS†
AND C. DRAPER*

* *Department of Tropical Hygiene, London School of Hygiene and Tropical Medicine, Keppel Street, London WC1E 7HT* and † *Lepra Evaluation Project, P.O. Box 46, Chilumba, Karonga District, Malawi*

(Accepted 2 February 1988)

SUMMARY

This paper examines the sensitivity and specificity of two ELISA assays for IgM antibodies to *Mycobacterium leprae*, one employing natural phenolic glycolipid and the other employing a synthetic disaccharide glycoconjugate as antigen. Estimates of sensitivity and specificity are derived, based on a panel of sera from leprosy cases in Malawi and various non-leprosy controls from the UK. Though both assays were able to identify a high proportion of multibacillary patients, neither was able to detect a high proportion of paucibacillary patients without considerable loss of specificity. The implications of the inverse relationship between sensitivity and specificity are discussed with reference to the predictive value of such tests in such areas as Malawi, where the large majority of cases are paucibacillary.

INTRODUCTION

A sensitive and specific test for infection with *Mycobacterium leprae* would be of great potential benefit. Diagnosis of infection prior to the onset of clinical disease might enable more efficient chemotherapy, and so reduce the risk of transmission of infection. Such a test might usefully be incorporated into population-based studies in leprosy-endemic areas, in order to study the natural history of the infection and the disease. It could also be of particular value in assessing whether control programmes reduce the transmission of infection.

In recent years, much of the research on tests for infection with the leprosy bacillus has focused on enzyme-linked immunosorbent assays (ELISA) employing the *M. leprae*-specific phenolic glycolipid (PGL1) antigen or synthetic preparations of its disaccharide epitope (Anonymous, 1986; Bach *et al.* 1986; Brett *et al.* 1983, 1986; Cho *et al.* 1983, 1984, 1986). The literature on these serological tests contains many references to their sensitivity and specificity for detecting 'leprosy'. This literature is confusing, however, as different panels of case and control sera have been employed, different criteria for seropositivity have been used, and the work

has in general not been carried out with reference to any particular epidemiological situation.

All published assessments of test sensitivity have been based upon selected sera from established clinical cases. This is so despite the fact that the greatest interest in such tests for field use is their potential for identifying *M. leprae* infection either prior to, or in the very early phases of, clinical manifestation. Specificity estimates have been based upon the performance of tests on sera from non-leprosy cases living in leprosy-endemic areas or else on sera for non-endemic areas. Neither of these is ideal in so far as control individuals from leprosy-endemic areas might be infected with *M. leprae*, and individuals from non-endemic areas are liable to lack exposure to the other potentially cross-reacting infections which will in fact determine specificity when the tests are applied in endemic areas.

The criterion for seropositivity has generally been set at either two (Bach *et al.* 1986; Brett *et al.* 1986) or three (Cho *et al.* 1983, 1984; Izumi *et al.* 1985) standard deviations above the mean of the non-leprosy controls. The criterion is thus dependent upon the choice of control sera as well as the confidence limit. The fact that the sensitivity and specificity implicit in different test criteria are generally inversely related makes the selection of an optimum criterion particularly complicated.

There are three reasons why serological tests for leprosy should be discussed in the context of particular epidemiological situations. First, there is evidence that the proportion of cases classified as multibacillary may vary between different populations (Fine, 1982). Given that the serological response is a function of the clinical spectrum, this means that different criteria for positivity may be applicable, and that the sensitivity of tests may differ, between areas or populations. Secondly, it is possible that the specificity of these tests will vary, depending upon the exposure of the local population to other potentially cross-reacting – e.g. other mycobacterial-antigens. Thirdly, the usefulness of a test, in terms of its predictive value, will also be a function of the prevalence of *M. leprae* infection or of clinical leprosy in the study population (Daniel & Debanne, 1987; Kronvall, 1981). The predictive value for a positive result is in this context defined as the percentage of all serological positives who have clinical leprosy (if one is interested in identifying diseased individuals), or as the percentage of all serological positives who are infected with *M. leprae* (if one's interest is in identifying infection). In contrast, the predictive value for a negative result is defined as the percentage of all serological negatives who do *not* have clinical leprosy or who are *not* infected with *M. leprae*. These predictive values are direct functions of test sensitivity, test specificity, and the prevalence rate of the condition in the population, and hence will vary between different epidemiological situations.

This paper examines the relationship between the sensitivity and specificity of the ELISA using *M. leprae*-specific PGL1 and synthetic glycoconjugate antigens, and determines the predictive values of these tests with particular reference to a leprosy-endemic area in Northern Malawi.

Table 1. Distribution of study sera in eight groups according to clinical characteristics. Range, mean and median ELISA values are presented in right-hand columns

Group	No. in group	Description	PGL absorbance (range, \bar{x} , median)	Glycoconjugate absorbance (range, \bar{x} , median)
A	7	BI > 0	0.28–0.75 \bar{x} = 0.46 med = 0.38	0.17–0.69 \bar{x} = 0.41 med = 0.32
B	15	BI = 0; activity = 2	0.07–0.50 \bar{x} = 0.27 med = 0.28	0.04–0.25 \bar{x} = 0.14 med = 0.13
C	55	BI = 0; activity = 1 Biopsy classification = BT, BT/BB or not given	0.07–0.43 \bar{x} = 0.24 med = 0.23	0.01–0.34 \bar{x} = 0.12 med = 0.12
D	25	BI = 0; activity = 1 Biopsy classification = TT or TT/BB	0.10–0.45 \bar{x} = 0.24 med = 0.22	0.04–0.25 \bar{x} = 0.12 med = 0.12
E	40	'Non-infected' Europeans with no history of exposure to <i>M. leprae</i> . Age \geq 18 years	0.05–0.29 \bar{x} = 0.13 med = 0.11	0.02–0.13 \bar{x} = 0.06 med = 0.06
F	16	'Non-infected' Europeans with no history of exposure to <i>M. leprae</i> . Age \geq 4 years	0.02–0.19 \bar{x} = 0.09 med = 0.10	0.02–0.14 \bar{x} = 0.05 med = 0.05
G	10	'Non-infected' Europeans with no history of exposure to <i>M. leprae</i> . Serologically positive for rheumatoid factor	0.04–0.25 \bar{x} = 0.14 med = 0.13	0.02–0.11 \bar{x} = 0.07 med = 0.07
H	19	Europeans with no history of exposure to <i>M. leprae</i> but infected with <i>M. tuberculosis</i> or <i>M. avium-intracellulare</i>	0.00–0.27 \bar{x} = 0.14 med = 0.14	0.01–0.15 \bar{x} = 0.08 med = 0.09

MATERIALS AND METHODS

Sera

Leprosy sera: a panel of 102 venepuncture-obtained sera was available from well-documented leprosy cases. All were from previously untreated cases ascertained in the Leprosy Evaluation Project in Northern Malawi (Ponnighaus *et al.* 1987b). Diagnoses were based on rigorous criteria described in detail elsewhere (Ponnighaus *et al.* 1987a). Complete clinical records were available for all cases (including biopsy results on 83 of them) and used for their classification according to the Ridley Jopling scale. Classification was carried out prior to and independent of the serological studies. The cases were allocated into four groups on the basis of clinical classification, bacteriological index (BI), and clinical activity at the time serum was taken. In this context the clinical activity was graded as 2, if the lesions were well-raised plaques or else had aggressively raised edges in combination with a healing centre; or as 1, if the lesions could just be felt as raised above the surrounding skin. Otherwise the activity was coded as 0. The criteria for and numbers in each group are shown in Table 1.

Control sera: the non-leprosy control sera included 66 serum samples taken from Europeans who were not known to have an infectious disease at the time the serum was taken, and who had no history of exposure to *M. leprae*. Of these, 10 were from persons with rheumatoid factor (as measured by indirect haemagglutination), and 16 were from children less than 4 years of age. A further 19 serum samples were from Europeans with other mycobacterial infections – either pulmonary tuberculosis cases who were sputum- and/or culture-positive (17) or individuals with *Mycobacterium avium-intracellulare* infections (2). These patients had no history of exposure to *M. leprae*. The criteria for and numbers in each of these groups are shown in Table 1.

A reference positive serum pool (prepared from serologically positive multi-bacillary leprosy cases) and a reference negative serum pool (from Europeans) were used for standardizing the results.

Malawi sera were stored at below 0 °C for varying periods before shipment to London, where they were aliquotted and stored at –20 °C before use. In order to minimize IgM degradation, no aliquot was used after more than two freeze–thaw cycles.

Antigens

Phenolic glycolipid 1 (PGL1), purified from armadillo-grown *M. leprae*, and the synthetic disaccharide glycoconjugate 8 (Brett *et al.* 1986) were kindly supplied by Dr R. J. W. Rees.

Serology

ELISA using PGL1 antigen

A sonicated suspension of 5 µg/ml PGL1 in carbonate buffer was coated on to wells of alternate rows of a polystyrene microtitre plate (Linbro) and incubated overnight at 37 °C. The following day the plate was washed once in phosphate-buffered saline (PBS) and blocked with 5% bovine serum albumen (BSA) in PBS for 2 h at 37 °C. Following another wash, serum samples diluted to 1 in 100 in 5% BSA in PBS were each added to one antigen-coated and one (adjacent) uncoated well. Five replicate reference positive and two replicate negative sera were included on each plate. Sera were incubated for 2 h at 37 °C and the plate washed six times. A rabbit anti-human – IgM antiserum conjugated to peroxidase (DAKO Laboratories, Denmark), diluted to 1 in 1000 in 5% BSA in PBS was incubated on the plate for 2 h at 37 °C. After four washes, the peroxidase substrate (ortho-phenylene diamine, H₂O₂ in citrate–phosphate buffer) was added and the plate incubated at 37 °C. The reference positive wells were monitored until they reached an absorbance value of 0.85 at 492 nm after stopping with 4M-H₂SO₄, at which time the reaction in all wells was stopped with acid. Absorbance values were read on a Titertek Multiskan (Flow Laboratories). A correction factor was applied to compensate for plate-to-plate variation in results recorded for the reference sera. Results were expressed as the difference between the absorbance values of coated and uncoated wells.

ELISA using synthetic glycoconjugate

Lyophilized synthetic glycoconjugate, diluted to 2 µg/ml in coating buffer, was used to coat wells of a polystyrene microtitre plate (Linbro) overnight at room

temperature (RT). The following day the plate was washed in 0.9% NaCl containing 0.05% Tween 20, and blocked with 5% BSA in PBS for 1 h at RT. The plate was then washed as previously. Serum samples and positive and negative reference sera were diluted to 1 in 20 in incubation buffer and incubated on the plate for 2 h at RT. After six washes of the plate, rabbit anti-human IgM conjugated to peroxidase (DAKO Laboratories) was diluted to 1 in 850 and incubated for 3 h at RT. After six washes, the peroxidase substrate was added and the reaction monitored at RT in the same way as the PGL1 ELISA, except that the end-point absorbance value for the reference positive sera was 0.82.

Analysis

Sensitivity, specificity and predictive value statistics were calculated according to conventional definitions (Daniel & Debanne, 1987; Kronvall, 1981; Last, 1983). Thus *sensitivity* is the percentage of actual leprosy case sera considered ELISA-‘positive’ according to a specified criterion; and *specificity* is the percentage of sera from non-leprosy controls which were considered ELISA-‘negative’ according to that criterion. *Predictive value* is defined as the proportion of test positives representing actual disease (or actual infection, if that is the condition of interest). Predictive values (*PV*) were calculated for different levels of sensitivity (*x*) and specificity (*y*) based on different assumed true prevalences of the condition being studied (*z*):

$$PV = \frac{xz}{xz + (1-y)(1-z)}$$

RESULTS

Figs 1 and 2 show the distribution of IgM activity to the PGL1 and the synthetic glycoconjugate antigens in the various groups of leprosy case and control sera. Summary statistics are presented in Table 1. A wider scatter was consistently observed with the PGL1 than with the glycoconjugate antigen assay.

Among cases, the multibacillary group A, defined as having BI > 0, had the highest antibody levels, but there was little difference between the three groups (B, C, D) with negative BI. Among the latter three groups, there is a suggestion of higher antibody levels in the group (B) with higher clinical activity, but this difference was not statistically significant. There was no significant correlation within group A between the BI at the time the serum was taken and the ELISA result with either antigen.

The median ELISA absorbance values were lower for each of the non-leprosy control groups (E–H) than for any of the leprosy case groups (A–D), for both antigens. There may be some evidence of an age effect in so far as absorbance values were marginally higher in the adult ‘healthy’ non-leprosy control group (E) than in the ‘healthy’ non-leprosy controls under 4 years of age (F), for both antigens. For both antigens the median values for the non-leprosy groups were highest in group H (persons with non-leprosy mycobacterioses) and second highest in G (persons serologically positive for rheumatoid factor), but these differences are not significant statistically.

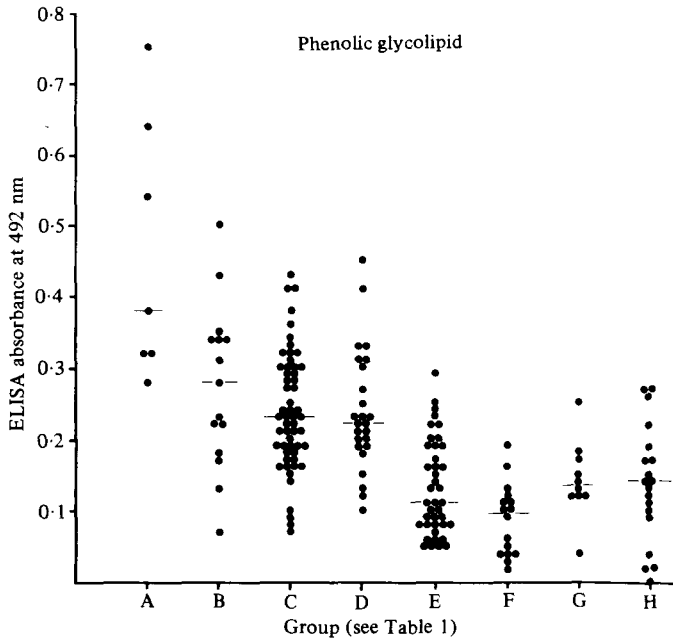


Fig. 1. ELISA results based on natural phenolic glycolipid (PGL1) antigen.

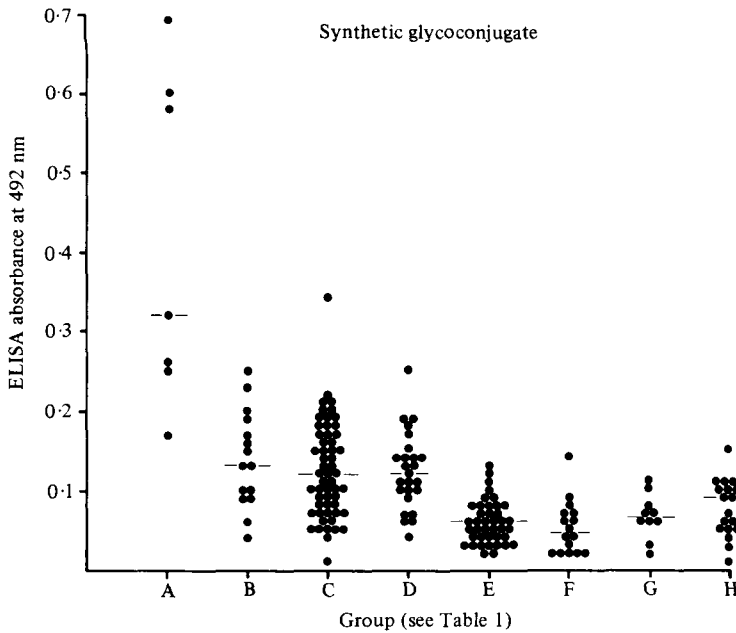


Fig. 2. ELISA results based on synthetic disaccharide glycoconjugate antigen.

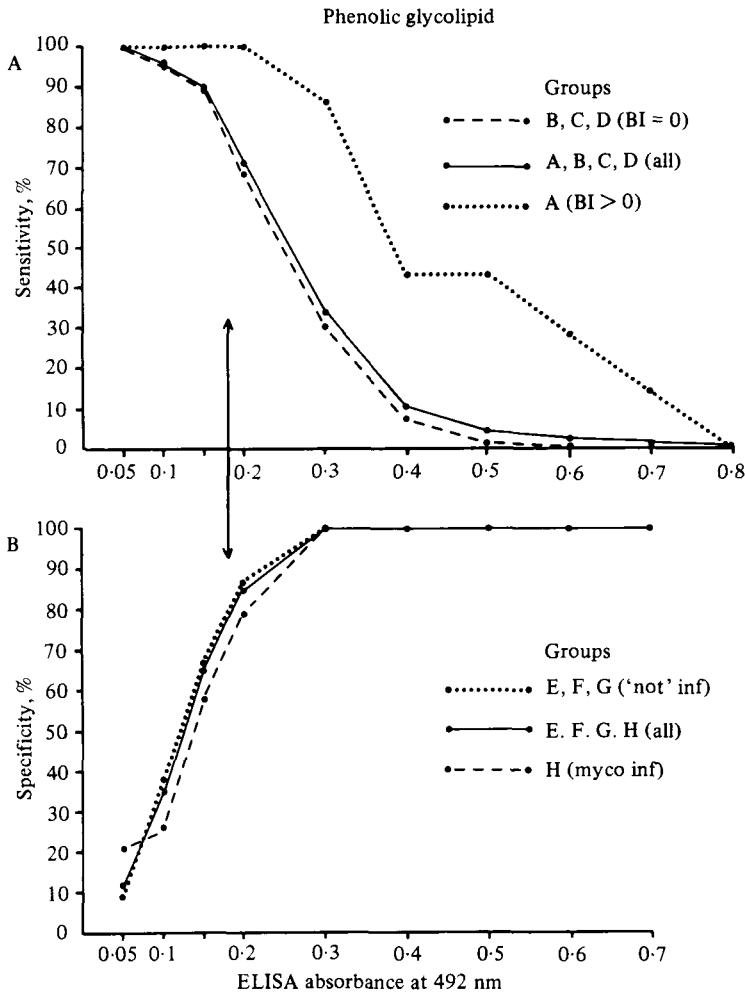


Fig. 3. Sensitivity (A) and specificity (B) of PGL1-based ELISA, as assessed using different case and control groups. Raw data shown in Fig. 1.

Figs 3 and 4 present the sensitivity and specificity values implicit in these results, for different case and control groups and using different criteria for test positivity. Fig. 3 refers to the results based on the PGL1 antigen, and Fig. 4 refers to results based on the glycoconjugate antigen. Figs 3A and 4A show sensitivity values, expressed separately for all cases, for group A (BI > 0) cases and for groups B-C-D (BI = 0) cases. Figs 3B and 4B show specificities, expressed separately with reference to all controls, to groups E-F-G (apparently not infected with mycobacteria) controls and to group H (infected with mycobacteria other than *M. leprae*) controls. The inverse relationship between sensitivity and specificity is shown clearly in these diagrams. For the PGL1-based assay (Fig. 3), an absorbance value of 0.18 is seen to give 78% sensitivity against all leprosy cases together, and 78% specificity against all the non-leprosy groups together. Higher positivity criteria decrease sensitivity, and lower criteria decrease specificity away

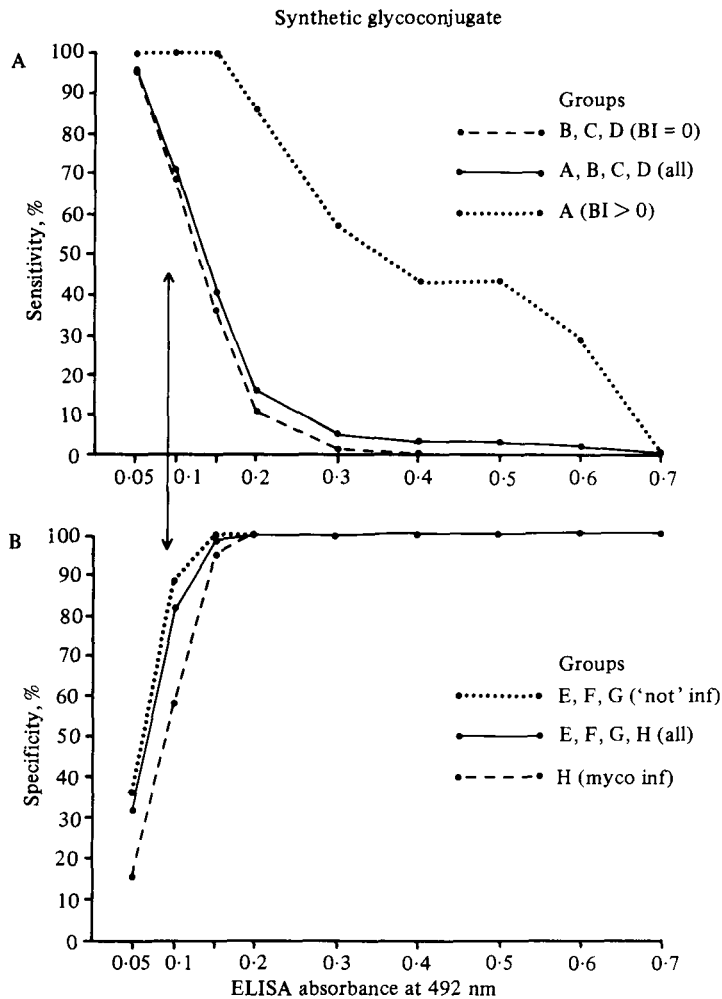


Fig. 4. Sensitivity (A) and specificity (B) of glycoconjugate-based ELISA, as assessed using different case and control groups. Raw data shown in Fig. 2.

from that value. For the glycoconjugate assay (Fig. 4), it appears that the analogous combined maximum is reached with a criterion of 0.09, giving sensitivity and specificity values of approximately 74%. Higher combined sensitivity and specificity are possible against the group A (BI > 0) cases alone, i.e. 93% sensitivity and specificity with a criterion of 0.25 for the PGL1 assay; and 98% sensitivity and specificity with a criterion of 0.16 for the glycoconjugate assay. For both assays the specificity is decreased, and hence the combined maximum sensitivity and specificity are decreased, if mycobacteria-infected individuals are included among the controls.

Figs 5 and 6 present the predictive values of a positive test based on different values of sensitivity (x) and specificity (y), assuming that the true prevalence rate of the condition under study (z) is either 0.02 (Fig. 5) or 0.05 (Fig. 6). These diagrams illustrate the well-known fact that specificity is more important than

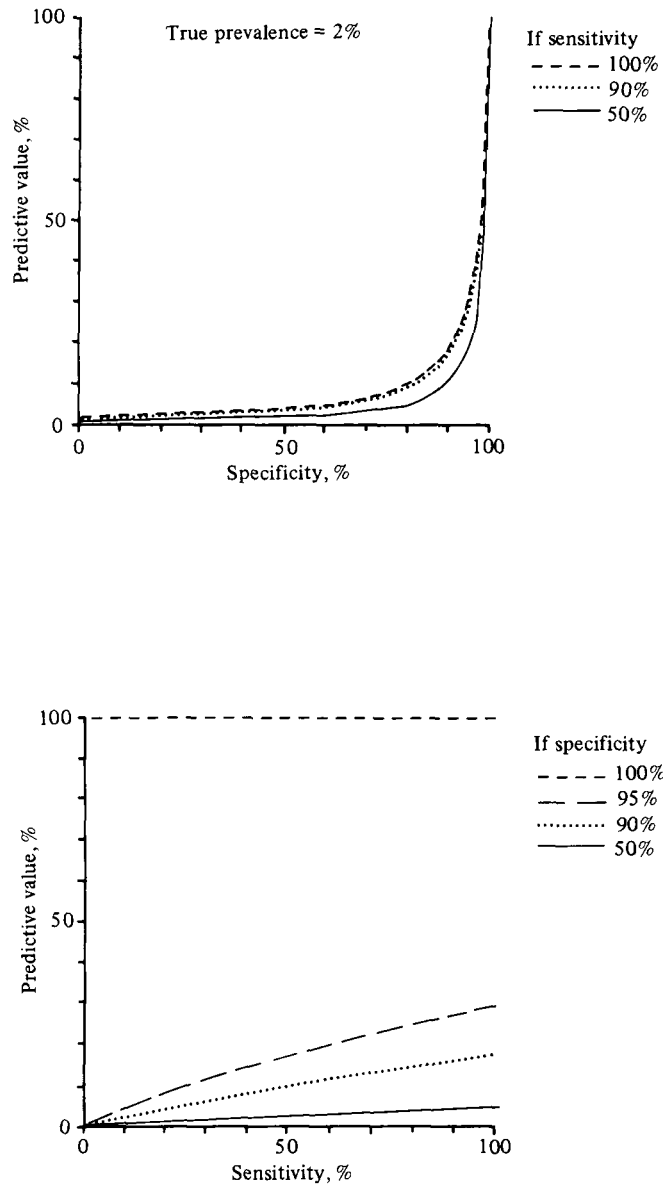


Fig. 5. Predictive value of diagnostic tests of different sensitivity and specificity, assuming a true prevalence of 0.02.

sensitivity, for low prevalence conditions. For example, if the true prevalence is only 2%, then the specificity must exceed 98% for the majority of observed positives to be true positives. If both sensitivity and specificity were 75%, then the predictive value (proportion correctly identified among test positives) would be 14% at a true prevalence of 0.05, and less than 6% at a true prevalence of 0.02.

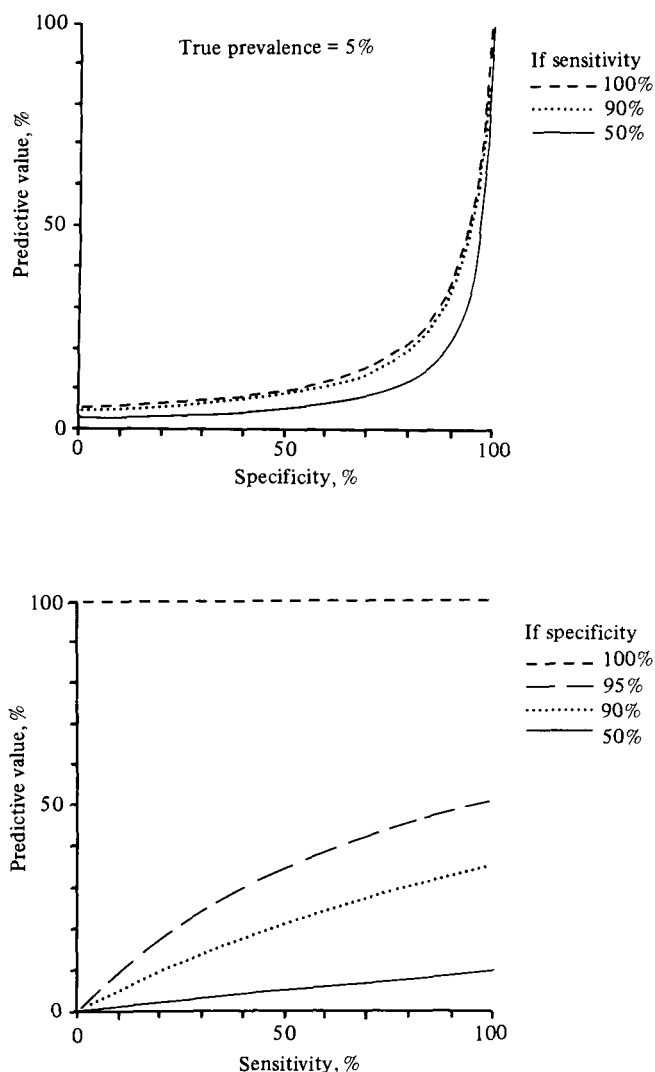


Fig. 6. Predictive value of diagnostic tests of different sensitivity and specificity, assuming a true prevalence of 0.05.

DISCUSSION

This paper explores the sensitivity and specificity of ELISA tests using two antigens thought to be specific for *M. leprae*: natural phenolic glycolipid and synthetic disaccharide glycoconjugate. It should be noted that the tests discussed in this paper, and this laboratory, were included in the recent IMMLEP-sponsored workshop comparing the results of eight different laboratories on a panel of blind-coded sera (Anonymous, 1986). The high degree of comparability observed between laboratories in that workshop suggests that the results presented here may be widely applicable.

Case sera employed in this study were from untreated leprosy cases ascertained in the Lepra Evaluation Project in Northern Malawi. Only 7 out of the 102 case sera were from individuals with positive slit skin smears, reflecting the low proportion of multibacillary cases in this population (McDougall, Ponnighaus & Fine, 1987; Ponnighaus *et al.* 1988). The non-leprosy sera were from individuals living in England, with no history of exposure to leprosy. Though individuals from Northern Malawi without leprosy might have provided a more appropriate control group in terms of background infections, we decided against this approach, in so far as it is impossible to exclude infection with the leprosy bacillus in such persons. On the other hand, it should be noted that our control group may have led to overestimates of test specificity, because of the rarity of potentially cross-reacting mycobacterial infections in England as compared to Malawi.

As expected, the sensitivity of both assays was found to be appreciably higher with reference to multibacillary than to paucibacillary cases. The specificity was found marginally higher when individuals with other mycobacterial infections were excluded from the non-leprosy group. The inverse relationship between sensitivity and specificity was such that with neither ELISA assay did it appear possible to achieve 100% sensitivity and 100% specificity at any criterion for positivity. We now turn to discuss the implications of these results for the use of such assays in studies of leprosy.

Given results such as those shown in Figs 3 and 4, it is in theory possible to select a positivity criterion to fit any desired level of sensitivity *or* specificity, but not both. The question thus arises as to what is the optimum level to use in any given context, recognizing that high diagnostic sensitivity can only be achieved at the expense of a considerable loss in specificity, and vice versa. In general, it will probably be preferable to opt for high specificity in field applications of such tests, for at least two reasons. First, in epidemiological studies low specificity generally introduces a far more severe bias in the assessment of relative risk than does low diagnostic sensitivity (Copeland *et al.* 1977; Kronvall, 1981); and thus a study may be invalidated by the presence of false positives – but only suffer a loss of statistical power by the presence of false negatives. Secondly, in a control context, if such a test were to be used to identify individuals for special attention or treatment, the low prevalence rate of leprosy means that a very high specificity is required to avoid flooding a project with false positives (Daniel & Debanne, 1987).

Reference to the 'prevalence rate of leprosy' in the last sentence raises a most important issue. This evaluation of sensitivity and specificity, and indeed *all* such studies in the leprosy literature, have used leprosy 'cases' as reference 'true' positives. The diagnostic sensitivity thus refers to defined clinical cases, and not to the total population infected with *M. leprae*. This is important, in so far as many authors have hoped that these tests might be useful in identifying incubating or subclinically infected individuals, in order to allow studies of the natural history of the infection and to permit targeted chemo- or immunoprophylaxis or therapy. However, given that we have as yet no valid independent test for *M. leprae* infection, there is no way to assess the sensitivity and specificity of a test for this state. Predictive value statistics for infection can in theory be derived through prospective studies involving the blind follow-up of large

numbers of individuals from whom serum specimens have been collected in the past, though not even this method allows evaluation of the potential of these tests to identify latent or self-healing infections with *M. leprae* (Douglas *et al.* 1987). In this context our only recourse is to make indirect inferences on the basis of population patterns of seropositivity in endemic areas (Fine *et al.* 1988).

The authors wish to thank the British Medical Research Council, the British Leprosy Relief Association and the IMMLEP component of the UNDP/World Bank/WHO Special Programme for Research and Training in Tropical Diseases for funding this work, and Mona Majithia for helping with the manuscript.

REFERENCES

- ANONYMOUS (1986). Serological tests for leprosy. *Lancet* **i**, 533-535.
- BACH, M. A., WALLACH, D., FLAGUEL, B., HOFFENBACH, A. & COTTENOT, F. (1986). Antibodies to phenolic glycolipid I and to whole *Mycobacterium leprae* in leprosy patients: evolution during therapy. *International Journal of Leprosy* **54**, 256-267.
- BRETT, S. J., DRAPER, P., PAYNE, S. N. & REES, R. J. W. (1983). Serological activity of a characteristic phenolic glycolipid from *Mycobacterium leprae* in sera from patients with leprosy and tuberculosis. *Clinical and Experimental Immunology* **52**, 271-279.
- BRETT, S. J., PAYNE, S. N., GIGG, J., BURGESS, P. & GIGG, R. (1986). Use of synthetic glycoconjugates containing the *Mycobacterium leprae* specific and immunodominant epitope of phenolic glycolipid I in the serology of leprosy. *Clinical and Experimental Immunology* **64**, 476-483.
- CHO, S. N., YANAGIHARA, D. L., HUNTER, S. W., GELBER, R. H. & BRENNAN, P. J. (1983). Serological specificity of phenolic glycolipid I from *Mycobacterium leprae* and use in serodiagnosis of leprosy. *Infection and Immunity* **41**, 1077-1083.
- CHO, S. N., FUJIWARA, T., HUNTER, S. W., REA, T. H., GELBER, R. H. & BRENNAN, P. J. (1984). Use of an artificial antigen containing the 3,6-di-*o*-methyl-B-D-glucopyranosyl epitope for the serodiagnosis of leprosy. *Journal of Infectious Diseases* **150**, 311-322.
- CHO, S. N., HUNTER, S. W., GELBER, R. N., REA, T. H. & BRENNAN, P. J. (1986). Quantitation of the phenolic glycolipid of *Mycobacterium leprae* relevance to glycolipid antigenemia in leprosy. *Journal of Infectious Diseases* **153**, 560-569.
- DANIEL, T. M. & DEBANNE, S. M. (1987). The serodiagnosis of tuberculosis and other mycobacterial diseases by enzyme-linked immunosorbent assays. *American Review of Respiratory Diseases* **135**, 1137-1151.
- DOUGLAS, J. T., CELLONA, R. V., ABALOS, R. M., MADARANG, M. G. & FAJARDO, T. (1987). The serological reactivity and early detection of leprosy among the contacts of lepromatous patients in Cebu, Philippines. *International Journal of Leprosy* **55**, 718-721.
- COPELAND, K. T., CHECKOWAY, A., MCMICHAEL, A. J. & HOLBROOK, R. H. (1977). Bias due to misclassification in the estimate of relative risk. *American Journal of Epidemiology* **105**, 488-495.
- FINE, P. E. M. (1982). Leprosy - the epidemiology of a slow bacterium. *Epidemiologic Reviews* **4**, 161-188.
- FINE, P. E. M., PONNIGHAUS, J. M., BURGESS, P. & CLARKSON, J. A. (1988). Seroepidemiological studies of leprosy in Northern Malawi based on an enzyme-linked immunosorbent assay using synthetic glycoconjugate antigen. *International Journal of Leprosy*. In press.
- IZUMI, S., SUGIYAMA, K., FUJIWARA, T., HUNTER, S. W. & BRENNAN, P. J. (1985). Isolation of the *Mycobacterium leprae* specific glycolipid antigen, phenolic glycolipid-I, from formalin-fixed human lepromatous liver. *Journal of Clinical Microbiology* **22**, 680-682.
- KRONVALL, G. (1981). The potential of immunological tests as tools in the epidemiology of leprosy. *Leprosy Review* **52** (Suppl.), 207-219.
- LAST, J. M. (editor) (1983). *A Dictionary of Epidemiology*. Oxford: Oxford University Press.
- MCDUGALL, A. C., PONNIGHAUS, J. M. & FINE, P. E. M. (1987). Histopathological examination

- of skin biopsies from an epidemiological study of leprosy in Northern Malawi. *International Journal of Leprosy* **55**, 88–98.
- PONNIGHAUS, J. M., FINE, P. E. M. & BLISS, L. (1987*a*). Certainty levels in the diagnosis of leprosy. *International Journal of Leprosy* **55**, 454–462.
- PONNIGHAUS, J. M., FINE, P. E. M., BLISS, L., SLINNEY, I. J., BRADLEY, D. J. & REES, R. J. W. (1987*b*). The Lepra Evaluation Project (LEP): an epidemiological study of leprosy in Northern Malawi. I. Methods. *Leprosy Review* **58**, 359–375.
- PONNIGHAUS, J. M., FINE, P. E. M., MAINE, N., BLISS, L., KALAMBO, M. & PONNIGHAUS, I. (1988). The Lepra Evaluation Project (LEP), an epidemiological study of leprosy in Northern Malawi. II. Prevalence rates. *Leprosy Review*. In press.