THE USE OF TESTS ON SLAUGHTERHOUSE CATTLE FOR ESTIMATING RELATIVE POTENCIES OF TUBERCULINS AND FOR THE CALCULATION OF DISCRIMINATION TESTS

BY A. B. PATERSON, P. STUART AND I. W. LESSLIE,

Central Veterinary Laboratory, Weybridge

AND F. B. LEECH

Rothamsted Experimental Station, Harpenden

(With 4 Figures in the Text)

Veterinary research work on mammalian tuberculin is principally directed towards the discovery of types or modifications of purified protein derivatives (PPDs), which differentiate between tuberculous and non-tuberculous cattle more accurately than the existing tuberculins, the tuberculous animal being defined as one infected with the bovine type of *Mycobacterium tuberculosis*. Cattle may be sensitized to mammalian tuberculin not only by *M. tuberculosis* var. *bovis*, but also by a variety of other types and species of *Mycobacterium*, those most frequently incriminated being *M. johnei*, the avian and human types of *M. tuberculosis* and the organism found in the condition known as 'skin tuberculosis'. There is also some evidence that *Brucella abortus* (Buxton & Glover, 1939) and *Actinobacillus lignieresi* (Feldman & Moses, 1942) may occasionally cause sensitization to mammalian tuberculin. Sensitizations also occur in the apparent absence of any of these agents. The fact that tuberculin made from an avian strain gives a reaction in some of these sensitizations, as well as in the homologous avian sensitization (M.A.F. statement 1942) has assisted in differentiating them from true tuberculosis.

New types of mammalian tuberculin can be produced in the laboratory with relative ease, but proof of biological superiority is more difficult to obtain. The potency of a new product, relative to a standard, must be determined, and also its specificity, i.e. its value in discriminating between infected and uninfected cattle.

Assays of tuberculins are most easily carried out on guinea-pigs, but, as previously shown (Fisher, 1949), materials can show quite different relative potencies when tested in guinea-pigs and cattle. It is therefore necessary that some method is devised for the assay of new materials on adequate numbers of cattle. The assay designed by Fisher was carried out on herds of cattle containing unknown numbers of infected animals. It can be taken that the distribution of responses in a population of cattle of which some are sensitized and some unsensitized will have two peaks, one near zero being the responses of unsensitized animals, the other being at about 5–6 mm. (Kerr, Lamont & McGirr, 1946). Fisher's method of analysis of this type of data was elaborate though highly effective; however, nothing could be said about the relative specificities of tuberculins from that type of experiment,

because there is no independent criterion for determining whether or not any animal was infected.

Since arrangements could be made to detain animals at slaughterhouses for the few days necessary for the tuberculin test, and since carcases could subsequently be examined with care for the presence of tuberculosis it was decided that a 'test-and-slaughter' experiment would provide the most practical approach to combined estimates of relative potency and specificity. As the analysis of relative potencies would be restricted to the responses of animals known to be infected, the statistical analysis would not require the elaborate methods devised by Fisher.

The present paper deals first with the estimation of relative potencies of tuberculins with their fiducial limits. These were obtained in two 'test-and-slaughter' experiments in which the estimates were calculated from reactions on cows subsequently proved to have been infected. The next section deals with the difference between responses to duplicated injections. The rest of the paper is concerned with discrimination between tuberculous and non-tuberculous cows. The results of the same two experiments are used to demonstrate a statistical analysis that is appropriate to this type of investigation, and may have wider applications.

EXPERIMENTAL

The experiment was carried out in two parts, one at Islington abattoir in 1951 and the other at the Central Veterinary Laboratory, Weybridge (1952–3), the animals used at each place being eighty culled dairy cows obtained from collecting centres in the Midlands. At Islington, they arrived in four groups of twenty; at Weybridge, in sixteen groups of five.

Tuberculins

The preparations and concentrations used are summarized in Tables 1 and 2. BAI culture medium is the Dorset–Henley synthetic medium used by the U.S. Bureau of Animal Industry (Dorset, 1934). The Watson–Reid medium (WR) (Watson, 1935) contains less glycerol (6 %) than BAI (10 %) and has the additional property that when growth of the culture is completed, the pH of the medium is acid, whereas the BAI medium finishes up at alkaline pH (Paterson, 1948).

Table 1. *Details of the ten tuberculin preparations used on all animals at Islington*

Type	Preparation no.	Strains used	Culture medium	Concentrations prepared (mg./ml.)		
Human	8	PN, DT, C	BAI	0·5	1·0	2·0*
Bovine	41	AN 5	BAI	0·5	1·0	2·0
Avian	18	D 4	BAI	0·125	0·25	0·5

* Duplicated.

The tuberculins used were all purified protein derivatives produced as described by Green (1946). Preparation no. 8 (Table 1) is equivalent in potency to the international standard (PPD-S) in both cattle and guinea-pigs. The early growth of

Table 2. *Details of the tuberculin preparations used at Weybridge*

Type	Preparation no.	Strains used	Culture medium	Concentrations prepared (mg./ml.)			Animals inoculated
Human	8	PN, DT, C	BAI	—	2·0*	4·0	All animals
Bovine	41	AN 5	BAI	—	2·0	4·0	101–140
Bovine	82	AN 5	BAI	—	2·0	4·0	141–195
Bovine	61	AN 5	WR	—	2·0	4·0	All animals
Avian	18	D 4	BAI	0·5	2·0	4·0	All animals

* Duplicated.

AN 5 in the production of preparation no. 41 was vigorous, but the pellicle sank at 4–5 weeks and the PPD obtained was brown in colour. Otherwise the preparations followed the routine procedure. The strongest PPD solutions were prepared by direct weighing of the powders which were dissolved in M/30 phosphate buffer (pH 7·0) with 0·5% phenol and 10% glycerol added as preservatives. Weaker solutions were prepared by dilution of the stronger. Protein content was checked from the trichloracetic acid precipitable nitrogen in the concentrates.

Injections and readings

Tuberculins were injected intradermally on the day a batch of cows arrived at the abattoir or the laboratory, using a set of ten McLintock (Arnolds) automatic syringes with short needles (1·5 mm.). The interdermal placement of each injection was checked by palpation and the injection repeated if no bleb could be felt. The volume of the injection was 0·1 ml.

Five sites were injected on each side of the middle third of the neck; four of them formed a diamond pattern, the fifth site being in the centre of the diamond. The distance between adjacent sites was approximately 10 cm. The randomized block experimental design was used, each cow being a block and the ten sites being allocated at random to the ten tuberculin treatments. The practical arrangements ensured absence of bias on the part of the operator, who was unaware of the particular tuberculin in any site, either at time of injection or of measuring reactions.

Measurements of the double fold of skin at each injection site were made with calipers before injection. At Islington, the sites were similarly measured 48 and 72 hr. later; at Weybridge the post-injection measurements were at 72, 96 and 144 hr. Readings were taken to 0·5 mm.

Post-mortem examinations

All animals were slaughtered after the reactions had been measured for the last time. The organs were then examined for lesions of tuberculosis and the lymph nodes removed for detailed examination at the laboratory. The nodes from each animal were distributed in five cartons containing, respectively, the grouped glands of the head, lung, mesentery, other abdominal organs, and the carcase.

https://doi.org/10.1017/S0022172400037505 Published online by Cambridge University Press

Each gland was stripped of fat, sliced into thin strips and examined for macroscopic lesions of tuberculosis. Tuberculous infection was confirmed by guinea-pig inoculation. Material from all lesions of doubtful origin was also subjected to biological test in guinea-pigs.

The experiments provide information on three major points. The first is the relative potencies of the tuberculins used. The second is what may be called the experimental error of tuberculin injection; this was measured from the differences in response to duplicate injections on the same animal. The third is the relative efficiency of different pairs of tuberculins in discriminating between infected and uninfected animals.

(1) *Estimates of relative potencies*

The sums of two successive daily measurements were used for the estimation of relative potencies because this improves the precision of the estimates. The distribution of these data being somewhat skew, they were transformed to square roots for statistical analysis. The effect of this transformation was to reduce the ratio between the variances of responses to high and low doses to a value near enough to unity to ensure the validity of the fiducial limits subsequently calculated.

(*a*) *Islington experiment.* The two mammalian tuberculins had each been injected in concentrations of 0·5, 1·0 and 2·0 mg./ml. This arrangement allowed the responses to human and bovine tuberculins to be treated as a six-point assay in randomized blocks, the duplicate response to the human tuberculin giving a second estimate of relative potency.

Table 3. *Mean responses to human and bovine tuberculins (Islington experiment). Each 'response' was the square root of the sum of the increases in skin thickness measured at 48 and 72 hr. after injection*

	Concentration (mg./ml.)		
Tuberculin	0·5	1·0	2·0
Human (*a*)	—	—	3·32
(*b*)	2·96	3·19	3·41
Bovine	2·64	2·92	3·29
Human minus bovine	0·32	0·27	0·07

The mean responses are shown in Table 3. The differences in the last row of the table show that the potencies of the two preparations tended to converge as the dose was increased, and analysis of variance confirmed that the lines were significantly non-parallel. Thus, even over the limited range of dosages studied, it was not possible to calculate a single estimate of relative potency that would be valid at all points within that range. The mean relative potency of the bovine tuberculin was 0·56. Non-parallel dose-response curves imply qualitative differences in the two tuberculins and they are therefore a hopeful sign when one is looking for differences in specificity.

(*b*) *Weybridge experiment.* The experimental design allowed the results to be treated as a four-point assay in randomized blocks, again with a duplicated dose of

human PPD (2·0 mg./ml.). The human PPD solutions were fresh dilutions made from the sample of PPD that had already been used for the Islington experiment. The two bovine PPDs were fresh preparations, except that no. 41 (Tables 1, 2) was used on the first forty animals. No appreciable difference could be observed between nos. 41 and 82.

Table 4. *Mean responses to human and bovine tuberculins (Weybridge experiment). Each 'response' was the square root of the sum of the measurements of increase in skin thickness taken at 72 and 96 hr. after injection*

	Concentration (mg./ml.)	
Tuberculin	2·0	4·0
Human (*a*)	3·04	—
(*b*)	3·12	3·28
Bovine BAI	2·72	3·02
Bovine WR	3·25	3·44

Table 5. *Potencies of bovine tuberculins in terms of human standard, with 5 % fiducial limits*

		5 % fiducial limits		% limits	
Tuberculin	Relative potency	Lower	Upper	Lower	Upper
Bovine BAI	0·396	0·176	0·605	44	153
Bovine WR	1·616	1·116	2·906	69	180

In this experiment the dose-response lines were tolerably parallel and the relative potencies of both bovine tuberculins could be stated in terms of the human standard tuberculin. The estimates with their 5 % fiducial limits are shown in Table 5, and the graph (Fig. 1) shows the results of both experiments on a comparable basis, i.e. the 72 hr. measurements. The fiducial limits are wider for the relative potency estimate of the tuberculin grown on WR medium than for that grown on BAI, in spite of the fact that the same numbers of observations were made in both cases and a pooled error used in the calculations. This difference results because the square of the log relative potency appears in the fiducial limit formula and would virtually disappear if doses were chosen which produced the same average response to all three tuberculins.

Fig. 1 shows that in the Islington experiment, concentrations of 0·5 mg./ml. of human and 1·2 mg./ml. of bovine (BAI) should give the same level of response, namely, 6·2 mm., in infected cattle. The graphs of the responses to the latter tuberculin in the Weybridge experiment would, however, have to be extrapolated well outside the range of concentration studied to get an equivalent estimate. For example, it would be expected to produce a response of 6·2 mm. in a concentration of about 7·7 mg./ml. and extrapolation of the human tuberculin response line suggests that 6·2 mm. would be produced with a concentration of 4·9 mg./ml. Such calculations are unreliable, first because there is no means of telling from the Weybridge experiment whether the dose-response lines were straight or curved.

An argument of much greater consequence is that the response to any one injection of tuberculin may be partly inhibited when other injections of tuberculin are given to the same animal at the same time; a marked inhibition of the response in guinea-pigs has been demonstrated in these conditions (Paterson & Leech, 1954), the experimental results suggesting that the inhibition may be proportional to the total amount of inflammatory reaction summed over all sites. The two experiments are not incompatible with this hypothesis, for the mean response of the infected

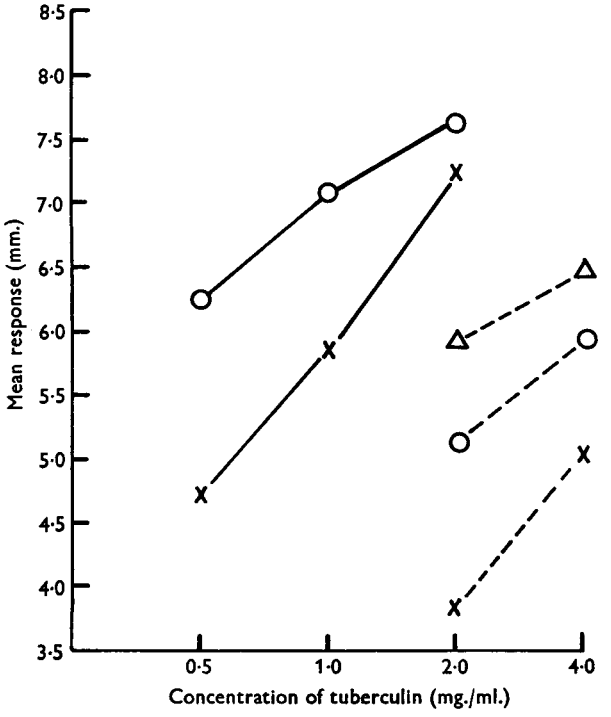

Fig. 1. Mean increase in skin thickness of infected animals, measured at 72 hr. after injection of tuberculin. ——, Islington experiment; – – –, Weybridge experiment; ○, human PPD; ×, bovine PPD ex BAI medium; △, bovine PPD ex WR medium.

animals when measured at 72 hr. after injection was 7·6 mm. to a dose containing 2·0 mg./ml. in the Islington experiment, where this was the highest dose, but 5·2 at Weybridge, where 4 mg./ml. was the highest dose. This is a drop of 32 %. The response to the 2·0 mg./ml. concentration of bovine BAI tuberculin dropped from 7·2 to 3·8 mm. or 47 %.

There is also the possibility that the mean sensitivities of the Islington and Weybridge cows were markedly different, i.e. that if they had been given identical treatment the mean responses would have differed to the extent of 30–40 %.

The experimental errors at Islington and Weybridge are very similar; the standard deviations expressed as percentages of the mean response (s/m) are shown in Table 6 together with values of s/b, which give estimates of the efficiencies of the assays, the smaller value indicating the higher efficiency.

The general conclusions to be drawn about the relative potencies of the bovine

tuberculins are that when grown on BAI medium the tuberculin is about half the potency of the human standard and when grown on WR medium the potency is about one and a half times that of the standard.

Table 6. *Errors and efficiencies of the two assays*

Experiment	S.D.	*s/b*	*s/m*
Islington	± 0·369	0·420	12 %
Weybridge	± 0·367	0·485	12 %

(2) *The repeatability of responses*

The responses to duplicate injections of human tuberculin at the 2 mg./ml. concentration provide a useful set of data for demonstrating one source of error in tuberculin testing. Table 7 shows the frequency distributions of the differences between these duplicates, with their standard deviations.

Table 7. *Frequency distributions of differences between responses to duplicate injections*

Difference between responses to duplicate injections (mm.)	Tuberculous animals		Tubercle-free animals	
	Islington	Weybridge	Islington	Weybridge
0	2	5	10	12
0·5	9	5	14	20
1·0	6	7	4	7
1·5	3	7	4	6
2·0	7	4	—	—
2·5	6	3	—	—
3·0	2	2	1	—
3·5	1	1	—	1
4·0	5	1	1	—
4·5	3	1	—	—
5·0	1	—	—	1
5·5	—	—	—	—
6·0	1	—	—	—
Total	46	36	34	47
S.D.	2·594	1·840	1·107	1·127

In animals apparently tubercle-free, differences greater than 1·5 mm. were relatively uncommon. If duplicate injections of mammalian tuberculin were used in the field it would, therefore, be reasonable to be suspicious of any animal showing a difference greater than 1·5 mm.

(3) *Specificity and discrimination*

In this section it will be necessary to use some common terms in a restricted sense if ambiguity and confusion are to be avoided. These terms with the precise meaning ascribed to them in the rest of the paper, are as follows.

Tuberculous: infected by the bovine type of *Mycobacterium tuberculosis*.

https://doi.org/10.1017/S0022172400037505 Published online by Cambridge University Press

Non-tuberculous or tubercle-free: not infected with the above type, though possibly infected or sensitized by other mycobacteria.

Mammalian tuberculin: tuberculin made from a human or bovine type.

Avian tuberculin: tuberculin made from the Weybridge strain D4.

Specificity of a tuberculin: Green (1946) defined the specificity of a heterologous tuberculin as 'the number of units of the heterologous protein derivative required to elicit the same intradermal reaction in infected guinea-pigs as one unit of the homologous protein derivative'; the 'specificity numbers' were calculated essentially in the same way as relative potencies and were assumed to be independent of the range of doses used. A specificity number as defined above was intended to be used in screening tests for selecting tuberculins likely to be worth field trials, tuberculins with high numbers in non-mammalian sensitization tests, but low numbers in mammalian sensitization tests being selected. If useful results are to be obtained, tuberculins must first be compared in relation to their homologous sensitization; concentrations that produce the same intradermal reactions to the homologous sensitization may then be tested on guinea-pigs with heterologous sensitizations. If this restriction is not imposed, differences in specificity numbers may merely reflect differences in relative potency, because so far as is at present known, the slope of the dose-response curve of all tuberculins in all sensitizations is positive. The 'specificity number' is helpful in the laboratory, but the laboratory results must always be treated with reserve, because they derive generally from guinea-pigs sensitized with one or a few strains of the most important mycobacteria. Far the greatest output of tuberculin is used for testing cattle which may give widely different results after natural infection by these organisms and which may be sensitized by others. Furthermore, it is well known that the characteristics of a strain may alter in the course of time when it is kept under laboratory conditions.

Some indication of specificity can be obtained from cattle tests by plotting, as in Fig. 2, the relationship between the mean responses of tuberculous and tubercle-free groups at each level of dosage. If tuberculins and sensitizations were completely specific, the graph would be a series of points on the x-axis, indicating zero response in tubercle-free animals at all levels of response on tuberculous animals. The actual results show that doses inducing a high level of response on tuberculous animals tend also to induce a relatively high level on tubercle-free animals. This diagram eliminates differences in the relative potencies of tuberculins. It shows also, that although the linear association between the two types of response was very similar in both Islington and Weybridge experiments, there was a higher general level of response in tubercle-free animals in the Weybridge experiment. It is suggested that this may have been due to what is known in the field as a higher level of non-specific sensitization.

We have been dealing so far with features of tuberculins and of the general level of sensitivity induced by an infection. For field use one must develop a 'test' which consists of a technique for using the tuberculin, together with rules for the interpretation of the response of any individual animal. In the present paper we consider single and comparative intradermal techniques and the rational development of rules of interpretation.

The existing advice for the interpretation of test results (M.A.F. statement, 1942, 1947) has developed largely on empirical lines, having been modified from time to time as experience in its use has accumulated in the field. The results of rigid application of these recommendations to the Weybridge and Islington experiments are shown in Table 8, which indicate a total error of between 10 and 20 %. There was evidence of recent tuberculin testing in some animals; this and

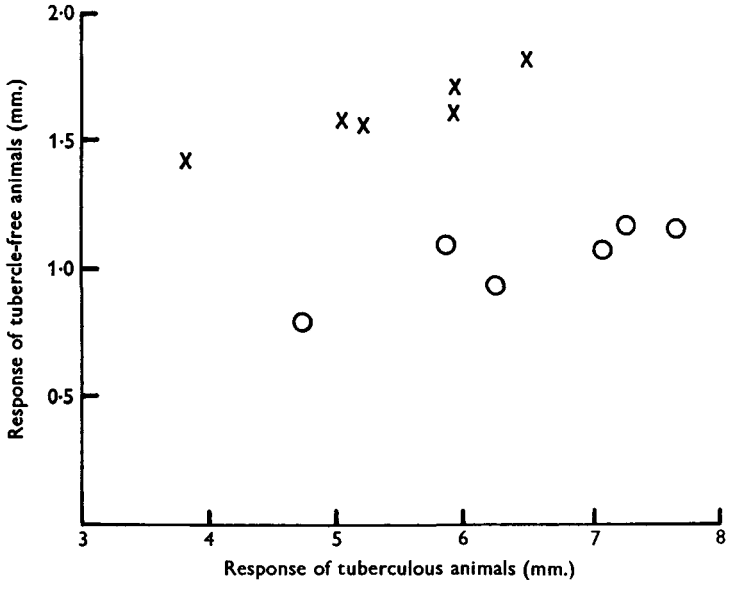

Fig. 2. The association, for the mammalian tuberculins, between the level of response on tuberculous and the level of response on tubercle-free animals, at each of the dosages used. ○, Islington experiment; ×, Weybridge experiment.

multiple injections probably contributed to test errors. On the farm, considerably greater accuracy is undoubtedly obtained by the common-sense use of supplementary information such as the 'test history' of the herd, the overall proportion of reactors in the herds, and the previous 'test history' of an individual animal.

The value of any technique as the basis for a test must be assessed from a comparison of the responses of tuberculous and non-tuberculous animals. Straight differences between means, as in the fourth column of Table 9, take no account of the 'experimental error' involved. A more sophisticated estimate of specificity takes into account the variation within groups of responses to the same material. The differences noted above may be divided by the standard deviation per response, giving values which will be known as d/s (Yates & Healy, 1951) and termed 'discriminating power'. In Fig. 3, the practical significance of this ratio is explained, d being the distance between the means of the two normal distributions each with standard deviation s. The left-hand distribution would generally represent the responses of uninfected and the right-hand the responses of infected animals. Division by s gives the distance between the two means expressed as multiples of the standard deviation and the significance of any value of d/s can therefore be

tested (ignoring errors of estimation) by regarding it as a deviate in the normal distribution with unit standard deviation using, for example, Table 1 in the Statistical Tables of Fisher & Yates (1953). Values of d/s of 1·96 and 2·58 correspond respectively to probabilities of 1/20 and 1/100, that as great or greater

Table 8. *Comparison of official interpretation with autopsy findings. 'Single intradermal'* (1) *and* (2) *are the responses to each of the injections of* 2·0 *mg./ml. human tuberculin considered as a unit. For* 'S.I. *Comparative', the response to* 0·5 *mg./ml. of avian tuberculin is also taken into account*

Test	Interpretation	Autopsy: Uninfected	Autopsy: Infected	Test errors
	Islington, 72 hr. readings			
Single intradermal (1)	32−	28	4	4
	7±	5	2	3·5
	41+	1	40	1 (10·6%)
Single intradermal (2)	32−	27	5	5
	11±	4	7	5·5
	37+	3	34	3 (16·9%)
S.I. Comparative (1)	32−	28	4	4
	15±	6	9	7·5
	33+	0	33	0 (14·4%)
S.I. Comparative (2)	32−	27	5	5
	18±	5	13	9
	30+	2	28	2 (20·0%)
	Weybridge, 72 hr. readings			
Single intradermal (1)	43−	38	5	5
	12±	4	8	6
	28+	5	23	5 (19·3%)
Single intradermal (2)	46−	39	7	7
	4±	1	3	2
	33+	7	26	7 (19·3%)
S.I. Comparative (1)	*44−*	*38*	*6*	*6*
	19±	8	11	9·5
	20+	1	19	1 (19·9%)
S.I. Comparative (2)	46−	39	7	7
	14±	5	9	7
	23+	3	20	3 (20·5%)

Interpretation: + = positive; ± = doubtful; − = negative.

distances should occur by chance. In the more complex case, where two or more variates have been measured, a discriminant function is calculated which when applied to the observations, reduces them to a single variate (X); the ratio d/s calculated from the values of X is exactly analogous to the situation shown diagrammatically in Fig. 3, d being the distance apart of the means of X for the two populations being compared and s being the standard deviation of X. The variate X has the property that the two distributions are further apart than the distri-

https://doi.org/10.1017/S0022172400037505 Published online by Cambridge University Press

Table 9. *Mean responses of tuberculous and tubercle-free animals: readings at 72 hr.*

Tuberculin type	Concentration (mg./ml.)	Mean increase in skin thickness (mm.)		Difference (mm.) (infected minus uninfected)	Values of d/s	Mean d/s
		Tuberculous animals	Tubercle-free animals			
		(*a*) Islington experiment				
Human	0·5	6·24	0·93	5·31	1·53	
	1·0	7·09	1·07	6·02	1·48	1·54
	2·0	7·65	1·15	6·50	1·61	
Bovine	0·5	4·72	0·79	3·93	1·27	
	1·0	5·87	1·09	4·78	1·25	1·31
	2·0	7·27	1·16	6·11	1·42	
		(*b*) Weybridge experiment				
Human	2·0	5·14	1·55	3·59	1·40	1·54
	4·0	5·96	1·70	4·26	1·67	
Bovine (BAI)	2·0	3·85	1·41	2·44	1·05	1·06
	4·0	5·08	1·59	3·49	1·07	
Bovine (WR)	2·0	5·94	1·60	4·34	1·45	1·50
	4·0	6·49	1·83	4·66	1·56	

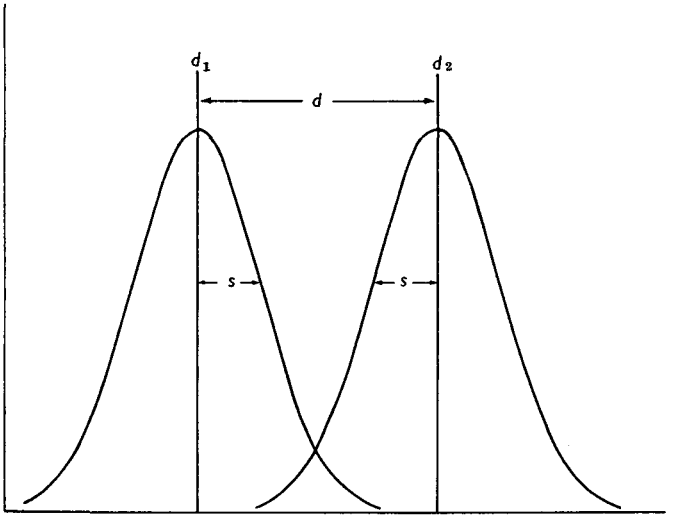

Fig. 3. Diagrammatic representation of normally distributed measurements on two populations with means differing by an amount d, but the same standard deviation s.

butions of any other variate that may be obtained from a linear function of the observation.

The value of d/s calculated from the responses to mammalian tuberculin are shown in the last column of Table 9. Although there was a general tendency for values of d/s to increase with increasing response, this was not entirely consistent. In both experiments the value of d/s was lower for bovine tuberculin grown on

https://doi.org/10.1017/S0022172400037505 Published online by Cambridge University Press

BAI medium than for either of the other tuberculins. This was because the smaller differences were not associated with reduced standard deviations.

The two experiments also provided an opportunity for testing the discriminating powers of different 'comparative' tests, each test being based on one injection each of mammalian and avian tuberculins. From the ten injections, twenty-one comparative tests were available in each experiment; the results for the duplicated injection of human tuberculin were pooled, thus reducing the total number of different tests to eighteen. The ordinary method of discriminant analysis was used

Table 10. *Values of d/s calculated from readings taken at 72 hr.*

Mammalian type	Dose (mg./ml.)	Avian tuberculin (mg./ml.) 0·125	0·25	0·5	Mean
		(*a*) Islington experiment			
Human	0·5	2·18	2·32	2·29	2·26
	1·0	2·01	2·38	2·06	2·15
	2·0	2·39	2·57	2·44	2·47
Bovine	0·5	1·68	1·86	1·68	1·74
	1·0	1·71	1·90	1·77	1·79
	2·0	2·09	2·33	2·17	2·20
Mean		2·01	2·23	2·07	
		(*b*) Weybridge experiment			
		0·5	2·0	4·0	Mean
Human	2·0	1·75	1·62	1·40	1·59
	4·0	1·98	1·94	1·72	1·88
Bovine ex BAI	2·0	1·32	1·22	1·05	1·20
	4·0	1·35	1·18	1·10	1·21
Bovine ex WR	2·0	1·86	1·71	1·57	1·71
	4·0	1·61	1·58	1·44	1·54
Mean		1·64	1·54	1·38	

to calculate a linear function of mammalian and avian responses for each test (equivalent to the variate X explained above). From the means of this function for infected and uninfected groups, and the within-groups standard deviation a value of d/s was obtained for each test. Results for the 72 hr. readings are shown in Table 10. It was found necessary to transform the Islington data to logarithms because the variances of infected and uninfected groups were very different. The full set of calculations was very extensive and was made possible by the availability of an electronic computer.

There are several important features of Table 10. First, the values of d/s are *consistently higher than those in Table 9, which were calculated without the extra* information provided by the responses to avian tuberculin. This shows that the avian tuberculin does help in the discrimination between tuberculous and non-tuberculous animals. Secondly, the values of d/s changed consistently with the concentration of avian tuberculin used and reached a maximum when the concentration was 0·25 mg./ml. (This concentration was not included in the Weybridge

https://doi.org/10.1017/S0022172400037505 Published online by Cambridge University Press

experiment, but the greatest value of d/s in that experiment was given by the smallest concentration used—0·5 mg./ml.) This (0·25 mg./ml.) is half the concentration used in cattle testing in this country. Thirdly, there are marked differences between mammalian tuberculins; the effect on d/s of increase in the concentration of mammalian tuberculin is, however, somewhat irregular.

The mean value of d/s calculated for readings taken at intervals after the injections were made are shown in Table 11; much better results were obtained at 72 hr. than at 48 hr. in the Islington experiment; in the Weybridge experiment a further, though slight, improvement was obtained with 96 hr. readings, but 2 days later, at 144 hr—6 days after injection—the discriminating power of the test had fallen off considerably. It is of interest that two animals which failed to react at 72 hr. became positive reactors at 96 hr.

Table 11. *Means of d/s taken over all mammalian responses*

Time after injection (hr.)	Avian (mg./ml.) 0·125	0·25	0·5	Mean
	(*a*) Islington experiment			
48	1·40	1·53	1·38	1·44
72	2·08	2·28	2·13	2·16
	(*b*) Weybridge experiment			
	Avian (mg./ml.) 0·5	2·0	4·0	
72	1·66	1·55	1·38	1·53
96	1·65	1·66	1·53	1·61
144	1·30	1·38	1·27	1·32

Having determined which test gives the highest discriminating power, the question remains of determining the appropriate rule by which to discriminate between infected and uninfected animals. The regression coefficients of the discriminating function were approximately in the ratio 2 (for human tuberculin) to -1 for avian tuberculin. This ratio did not vary greatly between experiments or between materials or dosages, the range being from $2:-1$ to $2:-1\frac{1}{2}$.

Taking the ratio $2:-1$ for the sake of simplicity (little improvement in discrimination would be obtained by a more exact ratio and one must also remember that the estimates are subject to error), an appropriate test would be to subtract the avian response from twice the mammalian. Application of this function to the Weybridge means yielded the results in Table 12. The values of 1·65 and 8·88 are equivalent to the x-co-ordinates of d_1 and d_2 in Fig. 3 and $\frac{1}{2}(d_1+d_2)$ is the co-ordinate of the point half-way between the two distributions and gives the point which, if used for discrimination in a population of which 50 % of animals are infected, will lead, on the average, to the fewest errors. In these conditions one may specify that if d is greater than 5 mm. (taking a round figure) the animal may be taken to be infected and if conversely d is less than 5 mm. the animals may be taken to be uninfected.

https://doi.org/10.1017/S0022172400037505 Published online by Cambridge University Press

Table 12. *Calculation of rule for discrimination between infected and uninfected animals (Weybridge experiment)*

	Response to human tuberculin at 2·0 mg./ml. (mm.)	Response to avian at 0·5 mg./ml. (mm.)	2 × human − avian = d (mm.)
Tuberculous	5·14	1·40	8·88
Tubercle-free	1·55	1·43	1·65
		Total	10·55
		$\frac{1}{2}(d_1+d_2)$	5·275

The application to the Islington results is slightly less simple, because of the necessity to use a logarithmic transformation. If h and a are the responses in terms of the logarithmic transform, the responses to human at 2·0 mg./ml. and avian at 0·25 mg./ml. yield the discriminant function $2h-a=1{\cdot}848$, at the critical value. The anti-logarithm of this function is $(10H+20)^2/(10A+20)$ with a critical value of 70·5, since the transformation used was $x=\log(10x+20)$. This function is a parabola, but the part of this parabola within the range of responses encountered in the experiment is nearly straight and may be replaced with virtually no loss of discriminating power by the line $2H-A$ with a critical value of 5. Thus the same function was derived from two very different sets of data.

This calculation applies to a sample containing roughly 50 % of infected cattle given ten injections and in such conditions, discrimination midway between the two distributions (Fig. 3) makes fewest errors. Where the sample is believed to be free of tuberculosis the point of discrimination may be shifted to the right, thus reducing the misclassification of uninfected animals while increasing the risk of missing an infected one. The problem is not, however, merely to make the minimum number of errors in given conditions because that might involve too great a risk of missing the occasional infected animal; the safest approach is to specify the minimum admissible risk of missing an infected animal and to estimate from that the appropriate point of discrimination, and the percentage errors to be expected in testing uninfected animals. Fig 4 shows the results on which such a decision might be made.

This figure is based on the results of the Islington experiment using the 2·0 mg./ml. concentration of human tuberculin and the 0·5 mg./ml. of avian. It shows, for example, that if a 2 % error in identifying uninfected animals is allowed for, the chance of identifying infected animals is 65 % and the point of discrimination would be 7·7 mm. It therefore seems probable that a discrimination point at 7 or 7½ mm. would be useful in herds where tuberculosis is believed to be absent.

DISCUSSION

The potency of bovine tuberculin

The strain AN 5 is a strain of bovine type isolated originally from a human. It was shown by McTaggart (1943) and Paterson (1948) that if cultured on BAI

https://doi.org/10.1017/S0022172400037505 Published online by Cambridge University Press

medium instead of WR synthetic medium, the yield of protein obtained was comparable with that obtained from human strains; in guinea-pig assays it was of the same potency as human tuberculin and apparently more specific. The surprising result at Islington of considerably diminished potency of the high-yielding material led to the use of the WR type protein medium in the Weybridge experiment, and to the introduction of a second batch of BAI bovine tuberculin midway through this experiment. The Islington results suggested that the human

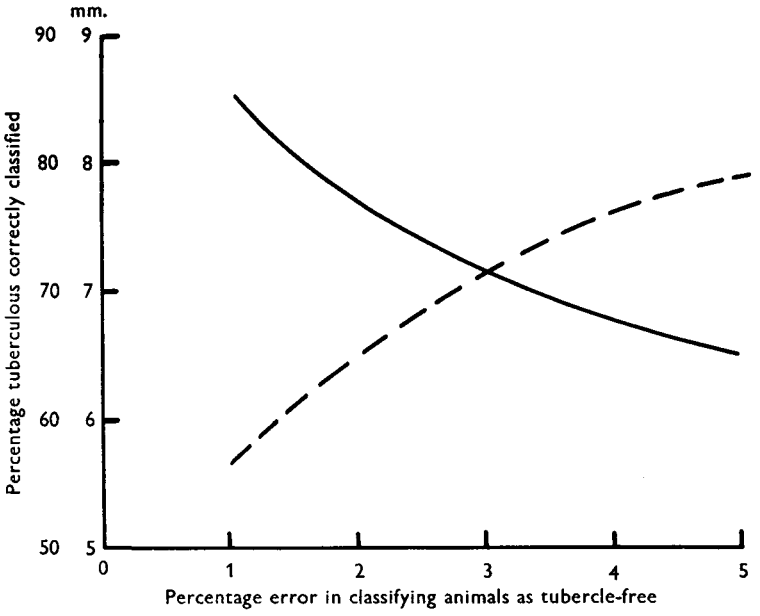

Fig. 4. The relationship between the point of discrimination, using the function 'twice response to human tuberculin minus response to avian tuberculin', and the percentage error in the classification of tuberculous and tubercle-free cows. ——, point of discrimination; – – –, percentage infected animals correctly classified.

and bovine dose-response curves might meet at a strength of about 4·0 mg./ml.; this dosage was therefore included in the Weybridge experiment. The Weybridge results clearly showed that the low potency of AN 5 tuberculin on BAI was real and not a property of the particular sample used, the second sample made in the same way at a much later date exhibiting the same potency. The expected intersection of dose-response lines was not obtained, probably because of the marked difference in average sensitivity in the two groups of cows. The samples of low-potency bovine tuberculin had the common characteristics that they had been prepared from cultures which sank at an early stage (3–4 weeks). While it has long been known that tuberculin of low potency resulted under these conditions, earlier statements have been based on preparations of OT in which the low potency may have been attributable either to low quality or to diminished quantity of protein. Our results indicate clearly that altered protein quality was the major cause, the material obtained possibly having a high proportion of the C-type protein described by Seibert (1949). In the Netherlands, PPD tuberculin is currently produced from

strain AN 5 on BAI medium and has been shown in cattle experiments to have equivalent potency to that of human PPD. It appears now that if AN 5 is maintained for long periods by subculturing on liquid media, the pellicle will sink after 3–4 weeks, but if maintained on glycerol-egg solid medium it will float for the full growth period on synthetic medium. A sample prepared in this way (no. 134) has been shown to resemble the Netherlands PPD in having the same potency as human PPD in infected cattle.

(2) *Differences between duplicate injections*

The discrepancies between the duplicate injections of the same tuberculin might be attributed to a number of causes, acting either singly or in combination, viz. injection error, errors in measurement, faulty syringes, or variations in site reactivity. The actual error involved was very similar at Islington and Weybridge, although injections at the former were made by P.S. and at the latter by A.B.P. At Islington the discrepancies at 48 hr. were in the same direction as those at 72 hr. although measurements were made by different operators at these times (A.B.P. and P.S.). Such a finding tends to eliminate recording and measurement as significant sources of error. The syringes used were all calibrated before use and mechanically it was easy to ensure that only a single 0·1 ml. dose was administered. The average sensitivities of the ten sites of injection were investigated and no significant difference in sensitivity could be found.

It therefore seems probable that the error could be described as mainly 'injection' error and, since it was of very similar size for both operators, the authors believe that it arises mainly from escape of tuberculin at the site of injection either subcutaneously or by back pressure through the needle puncture.

(3) *Discrimination test*

The animals in the present tests might be regarded as a random sample of single animals culled from randomly selected herds. The mean response of the tubercle-free animals to avian tuberculin at the 0·5 mg./ml. strength was 1·78 mm. in the Islington experiment and 1·43 mm. in the Weybridge experiment. Paterson & Hebert (1955) found an average response to the same strength of avian tuberculin in attested herds in Berkshire to be 2·08 mm. in 1953 and 2·32 mm. in 1954. The higher level of response in the field may to some extent be due to the use of only two injections in the official test, whereas each experimental animal had ten (cf. Paterson & Leech, 1954).

A more important point than the somewhat low average reactivity of tubercle-free animals in these two experiments is the well-known fact that in certain tuberculosis-free herds the level of sensitivity to avian and mammalian tuberculins is very high. The application of the $2H - A = 5$ rule in such herds would result in the removal of many more non-infected animals than would the existing interpretation. In these herds, rules of interpretation derived from random samples such as were used in our experiments, need modification—for instance, it might be advisable to discriminate on the square roots of the responses rather than the

responses themselves. The principle that some function of the avian response should be *subtracted* from the mammalian response should be unaltered. The actual function and the most efficient point of discrimination for herds exhibiting particular types of non-tuberculous tuberculin sensitivity could be a matter of further study.

There was no consistent evidence of superiority of bovine over human tuberculin, bovine produced from BAI medium being in fact markedly inferior to human. Furthermore, changes in the discriminating power of the test did not alter consistently with changes in the concentration of mammalian tuberculin. There was, however, a completely consistent relationship between discriminating power and the concentration of avian tuberculin, the maximum discrimination being obtained with 0·25 mg./ml. It would be unwise to assume without further experimentation that this result would apply exactly in the field, where only two injections are given, but it is safe to say that the concentration at present used (0·5 mg./ml.) should not be increased.

SUMMARY

1. The potency of PPD tuberculin made from AN 5, a bovine strain, was high when WR medium was used and low when BAI medium was used. PPD tuberculin made from strains of the human type grown on BAI medium was intermediate in potency between the two bovine PPDs.

2. The standard deviation of the difference between responses to duplicate injections of human type PPD at 2·0 mg./ml. concentration was 2·59 in the Islington and 1·84 in the Weybridge experiment on tuberculous animals, and 1·11 at Islington and 1·13 at Weybridge on tubercle-free animals. Differences greater than 1·5 mm. were relatively infrequent on tubercle-free animals.

3. The Elliott '401' electronic computer at Rothamsted was used to calculate discriminant functions and to estimate the discriminating power of different 'tests', each test involving an injection of a mammalian and an injection of avian tuberculin.

4. There were some differences between the mammalian preparations in their power to discriminate between tuberculous and tubercle-free animals, but these differences were closely allied to differences in potency. It seems probable that if concentrations of equivalent potency were used, there would be little if any difference in their discriminating power.

5. The contribution of avian tuberculin to the discrimination test was demonstrated; the greatest contribution came from a concentration of 0·25 mg./ml. used in the Islington experiment. The Weybridge results were quite consistent with the assumption that a maximum contribution is made by a concentration less than 0·5 mg./ml.

6. Methods for deducing rules for the interpretation of comparative tuberculin tests, using a mammalian and an avian tuberculin are illustrated and results given for the type of population sampled in these experiments.

7. The principle on which this rule might be modified where the population is believed to be free of infection, is outlined and discussed.

We should like to acknowledge the advice and constructive criticisms during the planning of the experiments of Dr A. W. Stableforth and Dr H. H. Green of Weybridge, and also the care with which Miss C. N. Hebert carried out the preliminary arrangement of the large quantity of data. Thanks are also due to the staff of the Islington abattoir for their willing co-operation.

REFERENCES

BUXTON, J. B. & GLOVER, R. E. (1939). *A.R.C. Rep. Series*, no. 4. *Tuberculin test in cattle*. London, H.M.S.O.
DORSET, M. (1934). *J. Amer. vet. med. Ass.* **84**, 439.
FELDMAN, W. H. & MOSES, H. E. (1942). *Amer. J. vet. Res.* **3**, 3.
FISHER, R. A. (1949). *Biometrics*, **5**, 300.
FISHER, R. A. & YATES, F. (1953). *Statistical tables for biological, agricultural and medical research*, 4th ed. London: Oliver and Boyd.
GREEN, H. H. (1946). *Vet. J.* **102**, 267.
KERR, W. R., LAMONT, M. G. & MCGIRR, J. L. (1946). *Vet. Rec.* **58**, 451.
M.A.F. STATEMENT (1942). *Vet. Rec.* **54**, 191.
M.A.F. STATEMENT (1947). *Vet. Rec.* **59**, 261.
MCTAGGART, H. S. (1943). Report to A.R.C. on *ad hoc* research (unpublished).
PATERSON, A. B. (1948). *J. comp. Path.* **58**, 302.
PATERSON, A. B. & HEBERT, C. N. (1955). *Vet. Rec.* **67**, 1143.
PATERSON, A. B. & LEECH, F. B. (1954). *Amer. Rev. Tuberc.* **69**, 806.
SEIBERT, F. B. (1949). *Amer. Rev. Tuberc.* **59**, 86.
WATSON, E. A. (1935). *Canad. publ. Hlth. J.* **26**, 268.
YATES, F. & HEALY, M. J. R. (1951). *Nature, Lond.*, **168**, 1116.

(*MS. received for publication* 27. II. 57)