THE INTRACEREBRAL MOUSE-PROTECTION TEST FOR PERTUSSIS VACCINES

By J. O. IRWIN

Statistical Research Unit of the Medical Research Council, London School of Hygiene and Tropical Medicine, Keppel Street, W.C.1

AND A. F. B. STANDFAST

The Lister Institute of Preventive Medicine, Elstree, Hertfordshire

CONTENTS

GENERAL CONSIDERATION OF THE TEST

During the last eight years a great many determinations of the relative potencies of pertussis vaccines have been made by the mouse-protection test using the intracerebral route for challenge. Procedures for this test have been described in detail by Kendrick, Eldering, Dixon & Misner (1947) and are given in *Minimal Requirements* (1948) and W.H.O. (1953). Briefly the method is as follows:

Selection of mice. White mice weighing 12–14 g., all of one sex, from the same known stock, are distributed at random into cages of fifteen mice. Three cages of fifteen mice are required for each antigen to be tested and a further four cages of unvaccinated controls are used for the titration of the challenge dose of viable *Bordetella* (*Haemophilus*) *pertussis*.

https://doi.org/10.1017/S002217240006126X Published online by Cambridge University Press

Immunization of mice. A single dose of vaccine is given. Each antigen is usually diluted to contain 2000, 400 and 80 million bacilli in 0·2 ml. saline and a group of fifteen mice injected intraperitoneally with each dose. The unvaccinated control mice are set aside at the same time and all the cages are kept together in the animal house. Period before challenge 10 days.

Challenge of mice. The challenge suspension of *B. pertussis* strain 18-323 is prepared from an 18–20 hr. culture on Bordet–Gengou medium by emulsifying a little of the growth in a 1 % aqueous solution of Difco–Casamino acids (technical grade) so that 0·03 ml. contains 50,000 organisms by opacity, using the National Institute's of Health, Washington, U.S.A. (N.I.H.) ground-glass standard opacity tube. For the titration of the challenge dose three further dilutions to contain 5000, 500 and 50 organisms in 0·03 ml. are usually prepared.

The mice are lightly anaesthetized with ether or ether-chloroform mixture and a dose of 0·03 ml. of a suitable dilution injected. Not more than 3 hr. is allowed to elapse between harvesting the challenge culture and injecting the last mouse.

Test period and calculation of results. The mice are observed for 14 days and a record kept of each death. Deaths occurring in the first 48 hr. after challenge are not included in the calculations. Mice which are paralysed on the 14th day, the last day of the assay, are considered as 'deaths'. Several methods are available for the calculation, in routine tests, of the ImD 50, the amount of vaccine which would protect 50 % of mice, and the LD 50, the amount of challenge suspension that would kill 50 % of the control mice. The methods usually used are the Reed–Muench (1938), the Worcester–Wilson (1943) or the Litchfield–Fertig (1941).

Our object was to discover whether the test was practicable and to determine its accuracy, and our conclusions in this respect are based in the main on four series of tests all of which had been carried out and analysed statistically by February 1950. Later tests provided further information but did not modify the essential conclusions reached. The four series were:

A. Six protocols containing in all twenty-six tests provided by Dr P. L. Kendrick (Michigan Dept. Health, U.S.A.)

B. A series of repeated tests carried out by one of us (A.F.B.S.) on seven vaccines. One of these was an American Standard, the remainder were of British manufacture.

C. Simultaneous tests carried out (A.F.B.S.) on vaccines, V 1, V 2 and V 3 (which were among those tested in B), twice a week over a period of 7 weeks.

D. A further fourteen simultaneous tests carried out by Mr H. Proom (Wellcome Research Laboratories, Beckenham) on vaccines V 1, V 2 and V 3.

The details of the statistical analysis of these four series of tests are given below.

In carrying out any test of this type it is essential to keep the reagents and procedure as constant as possible, using mice obtained from the same closely bred stock. In spite of such 'standardization' the main difficulties of the mouse test are:

(*a*) the day-to-day variation in the infective potency of the challenge suspension;

https://doi.org/10.1017/S002217240006126X Published online by Cambridge University Press

(*b*) the relatively restricted range of dosage of vaccine that can be given;

(*c*) the large day-to-day variation in ImD 50 (50 % immunizing dose) of the same vaccine;

(*d*) the large number of mice necessary to get a statistically significant result.

(*a*) In spite of every effort to standardize the challenge dose, its potency varies widely. Kendrick *et al.* (1947) reported that the LD 50 (50 % lethal dose) of the challenge dose varied from 80 to 1870 organisms in ten laboratories collaborating in parallel assays, so that the challenge doses ranged from 20 to 500 × LD 50. We have found a similar variation from day to day. Our results suggest that when the challenge dose is more than 200 × LD 50, the ImD 50 tends to be greater, but our data are too irregular to determine the precise form of the relationship. In any test, therefore, the challenge suspension may affect the variation described below in (*c*).

(*b*) The range of dosage possible is restricted. The usual scheme is three doses, 2000, 400 and 80 million: sometimes a fourth dose of 1000 million is added. 2000 million is about the maximum number of *B. pertussis* cells that can be given to a mouse without obvious toxic effects. Twice this dose is usually toxic for a proportion of the mice; at the other end of the scale 80 million may protect few if any of the mice so immunized, and doses below this are of little value, except on rare occasions when vaccines have, for reasons unknown, a very low ImD 50 (see under (*c*)). The highest practicable dose is thus only 25 times the lowest, and, moreover, ImD 50's near either end of the range are less likely to be accurate than those in the middle; the demonstration of significant differences in ImD 50 is therefore not easy.

(*c*) Kendrick, Updyke & Eldering (1949) reported that the ImD 50 in forty-five mouse-protection tests with vaccines prepared from culture 10-536, ranged from 50 to 1100 millions with a median of 230, and they consider that these values for any vaccine might be so distributed. We have observed a somewhat similar range which must be regarded as the expected day-to-day variation. Unfortunately, Kendrick did not compare two vaccines at the same time, but in a series of comparative tests we made with two vaccines on fourteen occasions (Table 1), vaccine V 3 was better than vaccine V 1 ten times, and worse than V 1 four times, although when results of all these tests were combined vaccine V 3 was significantly better than V 1. When the results obtained each day are inspected it will be seen that the ImD 50 of the two vaccines do not vary in parallel, and though in the first three tests vaccine V 1 was better than V 3, this trend is not maintained in subsequent results.

This apparently random variation, coupled with the wide range of LD 50 obtained with any vaccine and the restricted practicable dose range, means that any single result is far too dependent on chance to be reliable, even when a simple comparison 'better than' or 'worse than' is all that is required.

(*d*) Large numbers of mice used in multiple tests are therefore necessary to obtain significant results. It was decided to aim at being able to assert a significant difference (at the 5 % level) when the estimated potency ratio was greater than 2 or less than $\frac{1}{2}$. The number of animals necessary for this purpose depends mainly on the slope of the dosage-response curve connecting the proportion of survivors

with the logarithm of the dose. When the probit transformation is used these curves are linear. For series A, the slope was found to be $1{\cdot}46 \pm 0{\cdot}09$, for series B $0{\cdot}79 \pm 0{\cdot}06$, for series C 0·73 and for series D 0·93. At the time when these tests were analysed, it was thought that series A over-estimated the average accuracy obtainable.

Table 1. *Comparison of* ImD 50*'s of two vaccines* V 1 *and* V 3 *tested together on fourteen occasions*

	ImD 50 in millions	
Date of test	Vaccine V 3	Vaccine V 1
6. xii. 48	210	150
7. xii. 48	1300	650
13. xii. 48	1400	850
14. xii. 48	400	2700
20. xii. 48	510	900
21. xii. 48	400	2100
28. xii. 48	630	950
1. i. 49	710	1300
3. i. 49	400	1300
4. i. 49	2000	3300
10. i. 49	1300	1400
11. i. 49	∞	2000
17. i. 49	830	2200
28. i. 49	240	750

MAIN RESULTS AND ANALYSIS

Statistical Analysis of Series A (Kendrick's tests)

(*a*) *The determination of the* ImD 50

Dr Pearl L. Kendrick provided protocols of six experiments which were analysed statistically. The six experiments contained in all twenty-six tests, for each of which she had already estimated the ImD 50, by the Reed–Muench method—essentially a form of Behrens's method. The individual tests were, with one exception, carried out with three immunizing doses and fifteen animals to a dose; the exception was test C 224 in which there were four doses. Consecutive doses were in the ratio of 5 to 1 in three (164, 184, 190) and 4 to 1 in two (186, 187). In experiment 176 the same antigen was used throughout, but the ratio between consecutive doses was changed from one test to another.

The results and their errors. The Reed–Muench method is a rapid method of estimation which in itself provides no adequate estimate of error. The object of the statistical analysis presented here is to make the best possible estimates of the ImD 50, to determine fiducial limits for their error and to examine the constancy of the relation between dose and response. All the results are shown in Table 2.

Each test was first treated as a self-contained test. The probit (or normal equivalent deviation) corresponding to the percentage of survivors on any dose was assumed to be a linear function of the logarithm of the dose and the best fitting straight lines were obtained. From these the values of the ImD 50 were

Table 2. *Results of tests in series A—Kendrick's tests*

Exp. no.	Vaccine	ImD 50 in millions: Kendrick's result	ImD 50 in millions: As single test	ImD 50 in millions: With pooled slope		Fiducial limits: As single test: Actual	Fiducial limits: As single test: %	Fiducial limits: With pooled slope: Actual	Fiducial limits: With pooled slope: %	Slope: b	Slope: s.e.	χ^2	D.F.
164	10536	90	85	68	$P=0{\cdot}95$	0·016–0·143	18–167	0·031–0·147	46–216	2·20	0·80	0·02	1
					$P=0{\cdot}99$	0·000–0·177	0–208	0·024–0·188	36–277				
	C 224	130	138	139	$P=0{\cdot}95$	0·071–0·263	51–190	0·079–0·246	57–177	1·59	0·34	2·48	2
					$P=0{\cdot}99$	0·053–0·344	38–249	0·065–0·293	47–211				
	Comb. 1	240	247	252	$P=0{\cdot}95$	0·094–0·516	38–209	0·134–7·474	53–188	1·54	0·45	0·04	1
					$P=0{\cdot}99$	0·043–0·895	17–362	0·111–0·577	44–229				
	Comb. 2	220	187	182	$P=0{\cdot}95$	0·664–0·363	35–194	0·091–0·364	50–201	1·80	0·53	2·11	1
					$P=0{\cdot}99$	0·026–0·486	14–260	0·073–0·455	40–250				
	Pooled slope	—	—	—	—	—	—	—	—	1·67	0·23	—	—
176	Factor 2	120	149	67	$P=0{\cdot}95$	0–∞	0–∞	0·030–0·149	45–221	1·05	0·79	0·83	1
					$P=0{\cdot}99$	0–∞	0–∞	0·024–0·191	35–284				
	Factor 3	100	79	68	$P=0{\cdot}95$	0–0·192	0–245	0·037–0·124	55–183	1·05	0·52	1·36	1
					$P=0{\cdot}99$	0–∞	0–∞	0·030–0·149	45–221				
	Factor 4	110	98	72	$P=0{\cdot}95$	0·039–0·176	39–179	0·036–0·143	50–199	1·81	0·54	1·79	1
					$P=0{\cdot}99$	0·017–0·226	17–229	0·029–0·179	40–248				
	Factor 5	80	66	60	$P=0{\cdot}95$	0–0·132	0–200	0·029–0·125	48–209	1·54	0·50	2·98	1
					$P=0{\cdot}99$	0–0·167	0–253	0·023–0·158	38–264				
	Factor 6	90	81	69	$P=0{\cdot}95$	0·045–0·257	55–316	0·031–0·154	45–224	1·87	0·58	0·14	1
					$P=0{\cdot}99$	0·015–0·357	18–439	0·024–0·199	35–289				
	Factor 10	140	162	162	$P=0{\cdot}95$	0·072–0·360	44–222	0·080–0·333	49–205	1·51	0·36	0·28	1
					$P=0{\cdot}99$	0·049–0·519	30–319	0·063–0·417	39–257				
	Pooled slope	—	—	—	—	—	—	—	—	1·50	0·21	—	—
184	18533	70	60	63	$P=0{\cdot}95$	0·005–0·127	9–213	0·030–0·132	48–209	1·44	0·51	0·74	1
					$P=0{\cdot}99$	0–0·160	0–269	0·024–0·166	38–263				
	K 1289	460	395	388	$P=0{\cdot}95$	0·185–0·969	47–246	0·210–0·718	54–185	1·39	0·39	2·95	1
					$P=0{\cdot}99$	0·125–1·787	32–453	0·175–0·869	45–224				
	23 A	530	557	556	$P=0{\cdot}95$	0·279–1·653	50–297	0·284–1·090	51–196	1·54	0·45	0·01	1
					$P=0{\cdot}99$	0·208–1·839	37–330	0·228–1·351	41–243				
	10536 Susp. 1	140	139	139	$P=0{\cdot}95$	0·052–0·268	38–192	0·072–0·267	52–192	1·56	0·44	0·16	1
					$P=0{\cdot}99$	0·025–0·350	18–252	0·058–0·328	42–236				
	10536 Susp. 2	80	71	58	$P=0{\cdot}95$	0·025–0·121	34–170	0·026–0·133	44–229	2·43	0·82	0·01	1
					$P=0{\cdot}99$	0·004–0·154	6–216	0·020–0·172	34–297				
	Pooled slope	—	—	—	—	—	—	—	—	1·54	0·21	—	—

186	B 1	230	203	201	$P = 0{\cdot}95$	0·116–0·352	57–173	0·094–0·425	47–212	2·10	0·52	1·05	1
					$P = 0{\cdot}99$	0·087–0·460	43–227	0·071–0·564	35–281				
	L 1	610	732	874	$P = 0{\cdot}95$	0·398–4·193	55–573	0·407–2·36	47–270	1·81	0·62	0·05	1
					$P = 0{\cdot}99$	0·327–206·4	45–28,200	0·320–3·69	36–422				
	M 213 (Ref.)	140	124	129	$P = 0{\cdot}95$	0·029–0·270	23–218	0·060–0·251	47–195	1·30	0·46	0·01	1
					$P = 0{\cdot}99$	0·001–0·421	1–339	0·044–0·316	34–245				
	10536 B	230	119	224	$P = 0{\cdot}95$	0–∞	0–∞	0·106–0·502	47–224	0·89	0·45	0·32	1
					$P = 0{\cdot}99$	0–∞	0–∞	0·081–0·696	36–311				
	Pooled slope	—	—	—	—	—	—	—	—	1·44	0·25	—	—
187	Ech. 10536 B	220	222	250	$P = 0{\cdot}95$	0·067–0·512	30–231	0·133–0·446	53–179	0·99	0·28	1·26	2
					$P = 0{\cdot}99$	0·027–0·737	12–332	0·105–0·542	42–217				
	M 213 (Ref.)	95	95	95	$P = 0{\cdot}95$	0·058–0·154	62–163	0·037–0·245	39–257	3·09	0·85	0·04	1
					$P = 0{\cdot}99$	0·044–0·203	47–215	0·026–0·343	28–361				
	Ro. 10536 B	480	469	470	$P = 0{\cdot}95$	0·245–0·917	52–195	0·247–0·902	53–192	1·46	0·33	1·84	2
					$P = 0{\cdot}99$	0·185–1·244	39–265	0·197–1·139	42–242				
	M 213	450	481	557	$P = 0{\cdot}95$	0·283–1·022	59–212	0·245–1·352	44–243	2·23	0·64	3·26	1
					$P = 0{\cdot}99$	0·222–1·849	46–384	0·185–1·884	33–338				
	Pooled slope	—	—	—	—	—	—	—	—	1·39	0·20	—	—
190	9797	71,000	3,301	2,046	$P = 0{\cdot}95$	0·766–1·651 $\times 10^{329}$	25–500 $\times 10^{329}$	0·749–12·12	37–593	0·92	0·47	0·96	1
					$P = 0{\cdot}99$	0–∞	0–∞	0·570–35·69	28–1740				
	Cult. 18297	270	197	196	$P = 0{\cdot}95$	0·105–0·368	53–187	0·074–0·514	38–262	1·84	0·45	6·59	1
					$P = 0{\cdot}99$	0·077–0·497	39–252	0·049–0·775	25–395				
	Ref. 10536 B	600	1,196	762	$P = 0{\cdot}95$	0·356–7·623 $\times 10^{7}$	30–637 $\times 10^{7}$	0·330–2·80	43–368	0·82	0·38	0·26	1
					$P = 0{\cdot}99$	0–∞	0–∞	0·256–5·85	34–768				
	Pooled slope	—	—	—	—	—	—	—	—	1·16	0·25	—	—

In calculating fiducial limits for the pooled slope, the exact formula was used for experiments 186, 187, 190. In the other three experiments the slope is more than seven times its standard error and the approximate formula is adequate.

estimated and the fiducial limits for $P = 0{\cdot}95$ and $P = 0{\cdot}99$ were calculated. The estimates of slope made from each separate test inevitably have very large sampling errors, and a large part of the errors of the estimates of the ImD50 is due to the uncertainty of the slope. In a few cases (six tests out of twenty-six) the slope did not differ significantly from zero, in which case the fiducial limits become 0–∞. This means that the test, regarded as a self-contained test, provides no evidence of increasing protection with increasing dose and that consequently no valid estimate of the ImD50 can be made from it.

No significant departures from linearity and no significant differences in slope from test to test within the same experiment were found. Accordingly, a pooled estimate of slope was next made for each experiment and the ImD50's and their fiducial limits of error were recalculated, using the pooled estimate.

Columns (3), (4) and (5) of Table 6 give (i) Kendrick's estimate, (ii) the estimate when each test is treated as a single self-contained test, (iii) the estimate when the pooled slope from each experiment is used. Columns (6) and (8) give the fiducial limits of error for the estimates (ii) and (iii), and in columns (7) and (9) these are expressed as percentages of the corresponding estimate. When a pooled slope is used the fiducial limits are very considerably narrower and approximate to the real limits of error of the test when the effect of uncertainty of slope is eliminated. Omitting experiment 190 for which the slope is less than 5 times its standard error, the average fiducial limits are close to 50–200% at $P = 0{\cdot}95$ (49–209% without experiment 190 and 48–232% with experiment 190). Thus the true value can be taken to be between half and double the estimate.

Analysis of the slopes. No significant differences between experiments were found in the pooled slopes. The average overall slope is 1·46 with a standard error of 0·09, and the data as a whole are consistent with this constant value.

Table 3 gives an analysis of variance of the slopes.

It is worth noting that the expected values of the mean squares in Table 3 are unity, on the assumption that response is linearly related to the logarithm of the dose, and that the animals form a homogeneous group in respect to their reactions to the antigen. The mean square between experiments is somewhat below its expected value but not significantly so. The other two mean squares are remarkably close to their expected values. Since the mean square between experiments is not significantly subnormal no special explanation is called for, but the result does suggest that the experiments selected in these protocols are a little better than the usual.

The error when a constant slope is used for interpretation. The data are consistent with the hypothesis that the true slope is constant with a value close to 1·5. The actual estimate obtained is $1{\cdot}46 \pm 0{\cdot}09$. The average value of the sum of the weights Σnw is close to 20 (19·60). These values are used below to calculate the average errors of a single determination of the ImD50 with forty-five animals and of the potency ratio of a test vaccine to a reference vaccine with forty-five animals on each preparation.

For a single determination these errors depend, although only slightly, on the values of $5 - \bar{y}$, the deviation of the average probit response from 5 and, for the

https://doi.org/10.1017/S002217240006126X Published online by Cambridge University Press

determination of a potency ratio, on $\bar{y}_2 - \bar{y}_1$, the difference between the mean-probit responses to the test and reference vaccine. The results are shown in Table 4.

Table 3. *Analysis of variance of slopes—series A*

	Sum of squares	D.F.	Mean square
Between experiments	2·6387	5	0·5278
Within experiments: Between tests	19·1364	20	0·9568
Within tests	31·5252	29	1·0871

Table 4. *Fiducial limits of error—series A*

	Single determination of the ImD 50		Potency ratio—test against standard	
$5-\bar{y}$ or $\bar{y}_2-\bar{y}_1$	$P=0·95$ (%)	$P=0·99$ (%)	$P=0·95$ (%)	$P=0·99$ (%)
0	50–201	40–251	37–269	27–368
0·25	50–202	40–251	37–269	27–369
0·50	49–203	40–253	37–271	27–372
0·75	48–204	39–256	36–275	26–377
1·00	48–207	39–260	36–279	26–385

(*b*) *The determination of the* LD 50 *of the culture and its relation to the* ImD 50 *and the challenge dose*

Method of statistical analysis. One determination was made in each experiment with three or four doses at five-fold intervals. The values of LD 50 were estimated by finding the best-fitting straight line connecting the probit, corresponding to the mortality observed, with the logarithm of the dose. Fiducial limits corresponding to $P=0·95$ and $P=0·99$ were then obtained. In fact, the statistical method used was the same as that for the values of ImD 50.

With the possible exception of experiment 186 no significant differences in slope from experiment to experiment were found. The results could therefore be interpreted with regard to a pooled slope. This average slope was 1·11 including experiment 186 and 1·06 excluding it; there was therefore no point in the exclusion and the seven experiments could be regarded as homogeneous in slope. Table 5 gives the results and their errors. The estimates of the LD 50 by the three methods (i) Kendrick's own estimate by the Reed–Muench method, (ii) interpretation as a single test, (iii) interpretation with respect to a pooled slope, do not differ greatly compared with differences due to the sampling variation of the animals. The latter are of course large. When interpreted with regard to the pooled slope the average fiducial limits for a single determination from forty-five animals are 40–250 %.

The χ^2 values show that there were no significant differences of the probit-log-dose relationship from linearity.

The analysis of variance of slopes is shown in Table 6. The theoretical expectation is unity for each mean square, and neither diverges significantly from unity at the 5 % level of significance. If there is any heterogeneity in slope it is due to the value 2·47 for experiment 186.

https://doi.org/10.1017/S002217240006126X Published online by Cambridge University Press

Table 5. *Statistical analysis of series A—Kendrick's tests*

Exp. no.	LD50 Kendrick's result	LD50 As single test	LD50 With pooled slope		Fiducial limits: As single test Actual	As single test %	With pooled slope Actual	With pooled slope %	Slope b	Slope s.e.	χ^2	D.F.
164	160	150	119	$P=0{\cdot}95$	61–272	41–181	45–315	38–264	1·85	0·52	0·54	2
				$P=0{\cdot}99$	30–348	20–232	33–429	28–359				
176	1,000	1,126	1,314	$P=0{\cdot}95$	591–2,754	52–245	552–3,153	42–240	1·51	0·37	2·64	3
				$P=0{\cdot}99$	465–4,715	41–419	420–4,152	32–316				
184	100	43	75	$P=0{\cdot}95$	0·125–157	0·3–363	31–184	41–246	0·83	0·30	0·89	3
				$P=0{\cdot}99$	0–202	0–467	23–243	31–326				
186	691	659	695	$P=0{\cdot}95$	392–1,135	60–172	243–1,966	35–283	2·47	0·42	0·32	2
				$P=0{\cdot}99$	372–1,235	56–187	181–2,723	26–392				
187 Ro.	1,380	2,908	2,613	$P=0{\cdot}95$	1,129–10,140	39–349	1,202–5,722	46–219	0·92	0·21	3·16	3
				$P=0{\cdot}99$	803–19,816	28–681	941–7,342	36–281				
187 Ech.	2,550	5,146	4,316	$P=0{\cdot}95$	1,896–21,974	37–427	1,899–9,711	44–225	0·89	0·20	3·46	3
				$P=0{\cdot}99$	1,357–51,086	26–993	1,468–12,560	34–291				
190	280	291	259	$P=0{\cdot}95$	129–554	44–191	111–609	43–235	1·50	0·35	0·30	3
				$P=0{\cdot}99$	83–707	29–243	86–796	33–307				

Table 6. *Analysis of variance of slopes*

	Sum of squares	D.F.	Mean square	
Between slopes	12·382	6	2·064	V.R. 2·19
Residual	11·308	12	0·942	

Relation of LD 50, ImD 50 *and challenge dose.* The logarithms of the LD 50 and their standard errors are shown in Table 7.

Table 7. *Logarithms of the* LD 50*'s and their standard errors from series A*

Experiment	LD 50	log LD 50	S.E.	Fiducial limits ($P = 0·95$)
176	1314	3·119	0·194	
164	119	2·077	0·215	
184	75	1·873	0·199	
190	259	2·414	0·189	
186	695	2·842	0·230	
187 Ro.	2613	3·417	0·174	
187 Ech.	4316	3·635	0·180	
Mean				
164, 184, 190	136	2·134	0·116	18–228
186, 187, Ro. and Ech.	2330	3·368	0·110	1418–3828

Experiment 176 was one on antigen 16945 with different dose intervals in the different tests, and we are not further concerned with it here. The remaining values of the LD 50 fall into groups within which there are no significant differences. The mean value for the first group is 136 and for the second 2330, which, taking the error into account, may be called 150 and 2500 with as much accuracy as is justifiable.

It is interesting to compare the values of the ImD 50 with those of the LD 50 and the challenge doses. There are in all seven tests with antigen 10536, for which we may construct Table 8. The estimates of the ImD 50 in the first group do not differ

Table 8. *Comparison of* ImD 50 *and* LD 50 *of challenge dose in series A*

Exp.	Antigen	ImD 50 in millions	Challenge dose	LD 50	$R = \frac{\text{C.D.}}{\text{LD 50}}$
164	Ref.	68	40,000	150	270
184	Susp. 1	139	40,000	150	270
184	Susp. 2	58	40,000	150	270
186	B	224	40,000	2500	16
187 Ro.	B	250	50,000	2500	20
187 Ech.	B	470	50,000	2500	20
190	B	762	50,000	150	330

significantly, nor do those in the second group. Assuming that we can regard the two groups as referring to two different antigens, or two different batches of the

https://doi.org/10.1017/S002217240006126X Published online by Cambridge University Press

same antigen, it is clear that the ImD 50's in the first group are comparable. Both the challenge dose and the ratios R are the same for all three. The value of R for experiment 190 is about the same as for those in the first group; the inference from this is that the second antigen is really weaker than the first. There is no doubt about this however we regard the data. In the second group while the values of the ImD 50 do not differ significantly the R ratio is greater for experiment 190 than for the other three. It is possible, however, that the differences between the ImD 50's may be real. In this case the result supports the view that the ImD 50 increases as the ratio of challenge dose to LD 50 increases. But there is not enough evidence in these data to prove that the ImD 50 is directly proportional to R.

When an antigen is always tested by means of a comparison with a standard, the same challenge dose of the same culture being used for both, the difficulty probably disappears, because one would expect the potency ratio of test to standard to be independent of R. In the data analysed there were no significant changes of slope although the values of R varied, and this is a hopeful sign.

STATISTICAL ANALYSIS OF SERIES B (A.F.B.S. FIRST SERIES)

(*a*) *The determination of the* ImD 50

The data consisted of repeated tests on seven vaccines. Each vaccine was tested at three dose levels at five-fold intervals with fifteen mice on a dose. The LD 50 in the controls was determined from four doses at 10-fold intervals with fifteen mice on a dose. The results have been evaluated (i) by the Reed–Muench method and (ii) by the probit method, so as to determine the maximum-likelihood estimates and fiducial limits for their error. All the results are shown in Table 9.

The results and their errors. Each test was first treated as a self-contained test. The probit (or normal equivalent deviation) corresponding to the percentage of survivors on any dose was assumed to be a linear function of the dose and the best-fitting straight line was obtained. From these values the ImD 50's were estimated and the fiducial limits for $P = 0{\cdot}95$ and $P = 0{\cdot}99$ were calculated. As in Kendrick's results, and as is indeed inevitable, the estimates of slope made from each separate test have very large sampling errors, and a large part of the errors of the estimates of ImD 50 is due to the uncertainty of slope. Three tests (those on V 1 and V 2 on 6 August 1948 and that on V 1 on 21 September 1948) failed because there were no survivors except at the highest dose. In about half the tests (twenty at $P = 0{\cdot}95$ and twenty-six at $P = 0{\cdot}99$) the slope did not differ significantly from zero, in which case the fiducial limits become $0–\infty$. This means that the test regarded as a self-contained test, provides no evidence of increased protection with increasing dose and that consequently no valid estimate of the ImD 50 can be made from it.

With two possible exceptions no significant departures from linearity and no significant differences in slope were found. Accordingly, an overall pooled estimate of slope was made and the ImD 50's and their fiducial limits of error were recalculated using the pooled estimate, which proved to be 0·788 with a standard error of 0·064.

Table 9. *Results of tests in series B*

Date of experiment	Vaccine no.	ImD 50 (millions) Reed–Muench	ImD 50 (millions) As single test	ImD 50 (millions) With overall pooled slope	Fiducial limits	As single test Actual	As single test %	With overall pooled slope Actual	With overall pooled slope %	Slope b	Slope S.E.	χ^2	D.F.
30. vii. 48	V 3	> 2,000	$(3{\cdot}34 \times 10^{19})$	26,400	$P = 0{\cdot}95$	0–∞	—	4,230–161,000	16–610	0·09	0·52	0·003	1
					$P = 0{\cdot}99$	0–∞	—	2,380–284,000	9–1,100				
	V 1	> 2,000	2,500,000	10,000	$P = 0{\cdot}95$	0–∞	—	2,310–43,700	23–435	0·28	0·44	7·01	1
					$P = 0{\cdot}99$	0–∞	—	1,400–69,300	14–690				
	V 2	> 2,000	$(6{\cdot}70 \times 10^{77})$	50,000	$P = 0{\cdot}95$	0–∞	—	6,500–391,000	13–781	0·02	0·57	0·87	1
					$P = 0{\cdot}99$	0–∞	—	3,500–745,000	7–1,490				
	Pooled slope	—	—	—	—	—	—	—	—	0·15	0·29	—	—
6. viii. 48	V 3	732	683	1,060	$P = 0{\cdot}95$	418–1,290	61–189	200–5,570	19–528	2·86	0·80	0·08	1
					$P = 0{\cdot}99$	330–2,050	48–299	116–9,390	11–890				
	V 1	> 2,000	—	39,500	$P = 0{\cdot}95$	—	—	3,550–464,000	9–1,180	∞	—	—	—
					$P = 0{\cdot}99$	—	—	1,580–1,010,000	4–2,550				
	V 2	> 2,000	—	161,000	$P = 0{\cdot}95$	—	—	8,030–3,030,000	5–1,880	∞	—	—	—
					$P = 0{\cdot}99$	—	—	3,210–7,620,000	2–4,740				
	Pooled slope	—	—	—	—	—	—	—	—	—	—	—	—
16. viii. 48	Y 2	230	101	166	$P = 0{\cdot}95$	0–∞	—	53–516	32–310	0·48	0·34	1·11	1
					$P = 0{\cdot}99$	0–∞	—	38–737	23–443				
	V 3	870	1,870	1,330	$P = 0{\cdot}95$	0–∞	—	411–4,240	31–320	0·60	0·36	0·02	1
					$P = 0{\cdot}99$	0–∞	—	292–6,120	22–462				
	V 4	335	401	419	$P = 0{\cdot}95$	154–1,150	39–288	121–1,420	29–339	1·24	0·40	1·05	1
					$P = 0{\cdot}99$	73–3,110	18–774	84–2,080	20–497				
	Pooled slope	—	—	—	—	—	—	—	—	0·73	0·21	—	—
17. viii. 48	Y 2	400	483	491	$P = 0{\cdot}95$	126–2,200	26–456	147–1,610	30–328	0·98	0·37	1·76	1
					$P = 0{\cdot}99$	6–634,000	1–1,310	103–2,340	21–477				
	V 3	360	386	381	$P = 0{\cdot}95$	66–1,800	17–466	118–1,240	31–325	0·92	0·37	0·39	1
					$P = 0{\cdot}99$	0–∞	0–∞	80–1,800	21–471				
	V 4	230	195	179	$P = 0{\cdot}95$	4–614	2–315	54–599	30–335	0·91	0·39	1·02	1
					$P = 0{\cdot}99$	0–∞	0–∞	36–876	20–490				
	Pooled slope	—	—	—	—	—	—	—	—	0·94	0·21	—	—
23. viii. 48	V 3	310	386	386	$P = 0{\cdot}95$	68–2,190	—	123–1,190	32–309	0·87	0·36	2·48	1
					$P = 0{\cdot}99$	0–∞	—	89–1,700	23–441				
	V 1	2,000	$(1{\cdot}92 \times 10^{4})$	4,320	$P = 0{\cdot}95$	0–∞	—	1,210–15,400	28–355	0·47	0·39	0·04	1
					$P = 0{\cdot}99$	0–∞	—	822–22,800	19–528				
	V 2	> 2,000	5,260	11,700	$P = 0{\cdot}95$	1,562–$(3{\cdot}91 \times 10^{14})$	30–74×10^{9}	2,460–56,800	21–484	1·13	0·53	0·89	1
					$P = 0{\cdot}99$	0–∞	—	1,530–93,200	13–794				
	Pooled slope	—	—	—	—	—	—	—	—	0·77	0·23	—	—

Table 9 (*cont.*)

Date of experiment	Vaccine no.	ImD 50 (millions) Reed-Muench	ImD 50 (millions) As single test	ImD 50 (millions) With overall pooled slope	Fiducial limits	As single test Actual	As single test %	With overall pooled slope Actual	With overall pooled slope %	Slope b	Slope s.e.	χ^2	D.F.
24. viii. 48	V 3	160	134	83	$P=0{\cdot}95$	23–245	17–183	22–307	27–371	1·37	0·44	0·25	1
					$P=0{\cdot}99$	2–376	2–281	15–462	18–559				
	V 2	900	1,570	1,300	$P=0{\cdot}95$	0–∞	—	402–4,230	31–326	0·81	0·36	0·01	1
					$P=0{\cdot}99$	0–∞	—	273–6,130	21–472				
	Y 2	200	180	150	$P=0{\cdot}95$	24–466	13–259	43–508	29–339	1·06	0·39	0·21	1
					$P=0{\cdot}99$	0·02–804	0·02–448	30–746	20–498				
	Pooled slope	—	—	—	—	—	—	—	—	0·98	0·23	—	—
30. viii. 48	V 3	250	222	179	$P=0{\cdot}95$	66–490	30–221	54–596	30–333	1·26	0·40	0·96	1
					$P=0{\cdot}99$	13–1,030	6–463	38–868	21–485				
	V 2	> 2,000	($2{\cdot}46\times10^{20}$)	4,560	$P=0{\cdot}95$	0–∞	—	1,280–16,500	28–362	−0·64	0·37	0·00	1
					$P=0{\cdot}99$	0–∞	—	821–24,700	18–542				
	Y 2	190	135	120	$P=0{\cdot}95$	1–386	0·9–286	36–395	30–330	0·89	0·37	1·49	1
					$P=0{\cdot}99$	0–∞	0–∞	25–575	21–480				
	Pooled slope	—	—	—	—	—	—	—	—	0·69	0·21	—	—
31. viii. 48	V 3	290	284	429	$P=0{\cdot}95$	0–∞	—	131–1,370	31–319	0·58	0·36	0·12	1
					$P=0{\cdot}99$	0–∞	—	94–1,970	22–459				
	Y 2	670	884	811	$P=0{\cdot}95$	0–∞	—	227–2,920	28–360	0·69	0·40	0·42	1
					$P=0{\cdot}99$	0–∞	—	154–4,370	19–539				
	NIH 4	730	871	1,110	$P=0{\cdot}95$	350–9,670	40–1,110	322–3,810	29–343	1·08	0·40	0·07	1
					$P=0{\cdot}99$	226–($1{\cdot}14\times10^{10}$)	26–(131×10^{7})	222–5,620	20–506				
	Pooled slope	—	—	—	—	—	—	—	—	0·77	0·23	—	—
6. ix. 48	V 3	680	554	520	$P=0{\cdot}95$	0–∞	—	172–1,560	33–299	0·62	0·34	0·39	1
					$P=0{\cdot}99$	0–∞	—	125–2,200	24–423				
	Y 2	180	157	95	$P=0{\cdot}95$	48–313	30–199	24–374	25–394	1·58	0·47	1·91	1
					$P=0{\cdot}99$	15–410	10–260	15–576	16–607				
	NIH 4	360	475	461	$P=0{\cdot}95$	0–∞	—	138–1,510	30–328	0·71	0·37	0·29	1
					$P=0{\cdot}99$	0–∞	—	97–2,190	21–476				
	Pooled slope	—	—	—	—	—	—	—	—	0·86	0·21	—	—
7. ix. 48	V 3	2,000	6,360	4,320	$P=0{\cdot}95$	0–∞	—	1,210–15,600	28–362	0·66	0·40	0·68	1
					$P=0{\cdot}99$	0–∞	—	778–23,500	18–543				
	NIH 4	860	923	1,300	$P=0{\cdot}95$	431–3,300	47–358	364–4,690	28–361	1·36	0·41	0·04	1
					$P=0{\cdot}99$	312–11,900	34–1,290	234–7,030	18–541				
	WRL	1,570	2,030	4,440	$P=0{\cdot}95$	914–41,100	45–2,023	1,020–19,000	23–428	1·42	0·53	0·70	1
					$P=0{\cdot}99$	718–($2{\cdot}47\times10^{13}$)	35–($1{\cdot}22\times10^{12}$)	666–30,000	15–676				
	Pooled slope		—	—	—	—	—	—	—	1·09	0·26	—	—

13. ix. 48	V3	470	422	422	$P=0{\cdot}95$	0–∞	—	114–1,590	27–377	0·43	0·40	0·14	1
					$P=0{\cdot}99$	0–∞	—	72–2,410	17–572				
	N1H	1,140	2,660	1,990	$P=0{\cdot}95$	0–∞	—	558–7,020	28–352	0·66	0·39	0·02	1
					$P=0{\cdot}99$	0–∞	—	379–10,400	19–522				
	WRL	1,040	3,580	2,010	$P=0{\cdot}95$	0–∞	—	522–7,680	26–383	0·58	0·40	0·42	1
					$P=0{\cdot}99$	0–∞	—	341–11,700	17–584				
	Pooled slope	—	—	—	—	—	—	—	—	0·56	0·23	—	—
14. ix. 48	V3	350	393	414	$P=0{\cdot}95$	153–1,100	39–280	108–1,610	26–389	1·39	0·44	0·12	1
					$P=0{\cdot}99$	78–2,590	20–659	70–2,470	17–597				
	WRL	850	831	1,150	$P=0{\cdot}95$	449–1,900	54–228	287–4,680	25–408	1·92	0·56	3·82	1
					$P=0{\cdot}99$	334–3,570	40–429	183–7,280	16–635				
	V4	180	141	107	$P=0{\cdot}95$	10–355	7–252	31–376	29–350	1·07	0·40	0·67	1
					$P=0{\cdot}99$	$(0{\cdot}53\times10^{-5})$–522	$(0{\cdot}38\times10^{-5})$–371	20–556	19–518				
	Pooled slope	—	—	—	—	—	—	—	—	1·37	0·26	—	—
20. ix. 48	V3	> 2,000	23,500	7,210	$P=0{\cdot}95$	0–∞	—	1,880–28,000	26–389	0·54	0·41	0·15	1
					$P=0{\cdot}99$	0–∞	—	1,230–43,000	17–596				
	WRL	> 2,000	2,910	7,930	$P=0{\cdot}95$	1,270–$(1{\cdot}22\times10^{6})$	44–419	1,430–44,000	18–555	1·61	0·70	0·19	1
					$P=0{\cdot}99$	0–∞	0–∞	873–75,500	11–952				
	V4	1,080	2,850	1,580	$P=0{\cdot}95$	0–∞	—	489–5,440	29–344	0·52	0·37	0·65	1
					$P=0{\cdot}99$	0–∞	—	316–8,000	20–508				
	Pooled slope	—	—	—	—	—	—	—	—	0·68	0·26	—	—
21. ix. 48	V3	> 2,000	3,230	8,780	$P=0{\cdot}95$	1,200–$(1{\cdot}58\times10^{8})$	37–49,100	167–4,510	19–513	1·38	0·62	0·37	1
					$P=0{\cdot}99$	0–∞	0–∞	105–7,530	12–857				
	V1	> 2,000	—	7,070	$P=0{\cdot}95$	—	—	849–61,000	12–862	∞	—	—	—
					$P=0{\cdot}99$	—	—	424–120,000	6–1,700				
	V4	580	—	1,240	$P=0{\cdot}95$	0–∞	—	274–5,640	22–453	0·29	—	4·77	1
					$P=0{\cdot}99$	0–∞	—	174–9,070	14–729				
	Pooled slope	—	—	—	—	—	—	—	—	0·66	0·36	—	—
27. ix. 48	V3	510	767	653	$P=0{\cdot}95$	0–∞	—	209–2,060	32–316	0·58	0·36	0·95	1
					$P=0{\cdot}99$	0–∞	—	144–2,970	22–454				
	V1	> 2,000	106,000	15,600	$P=0{\cdot}95$	0–∞	—	328–7,570	21–484	0·50	0·47	0·27	1
					$P=0{\cdot}99$	0–∞	—	203–12,400	13–794				
	V4	620	774	748	$P=0{\cdot}95$	161–$(1{\cdot}20\times10^{7})$	21–15,500	232–2,380	31–318	0·74	0·36	0·16	1
					$P=0{\cdot}99$	0–∞	0–∞	164–3,420	22–457				
	Pooled slope	—	—	—	—	—	—	—	—	0·62	0·23	—	—

Columns (3), (4) and (5) of Table 13 give (i) the Reed–Muench estimate, (ii) the estimate when each test is treated as a single self-contained test, (iii) the estimate when the pooled slope is used. Columns (6) and (8) give the fiducial limits of error from the estimates (ii) and (iii), and in columns (7) and (9) these are expressed as percentages of the corresponding estimates. When a pooled slope is used the fiducial limits are very considerably narrower, on the average, and approximate to the real limits of error of the test when the effect of uncertainty of slope is eliminated. The average fiducial limits are 26–450 % at $P = 0{\cdot}95$ or 27–365 %, omitting the three tests which 'failed'. The corresponding limits in those of Kendrick's tests which we examined were 49–209 %. The Kendrick tests examined were therefore decidedly more accurate, due to a slope which is almost double the slope obtained here.

The Reed–Muench method is clearly accurate enough for purposes of routine estimation, since errors due to the method of estimation are small compared with those due to the sampling variation of the mice. The Reed–Muench method, however, provided no valid estimate of error.

Analysis of the slopes. No significant differences between tests were found in the slopes. The average overall slope is 0·788 with a standard error of 0·064 and the data as a whole are consistent with this constant value. Table 10 gives the analysis of variance of the slopes.

Table 10. *Analysis of variance of slopes in series B*

	Sum of squares	D.F.	Mean square
Between slopes:			
Between days	14·753	13	1·135
Between different vaccines on same day	23·279	27	0·862
Residual within tests	36·905	41	0·900

The expected value of the mean squares in Table 10 is unity, on the assumption that response is linearly related to the logarithm of the dose and that the animals form a homogeneous group in their response to the antigen. The mean squares are very close to their expected values and there is no significant departure from the values expected.

The error when a constant slope is used for interpretation. The data are consistent with the hypothesis that the true slope is constant with a value close to 0·8. The actual value is $0{\cdot}79 \pm 0{\cdot}06$. This may be compared with Kendrick's value of $1{\cdot}46 \pm 0{\cdot}09$. These tests give a decidedly flatter slope than Kendrick's, which means that the variability of response in the mice used here was considerably greater than in Kendrick's tests. The error of the test is consequently larger.

The average value of the sum of the weights Σnw is 20·1; in Kendrick's tests it was 19·6, so that for practical purposes, the two averages are identical. These values can be used to determine the average errors of a single determination with forty-five animals and of the potency ratio of a test vaccine to a reference vaccine with forty-five animals on each preparation. For a single determination these

https://doi.org/10.1017/S002217240006126X Published online by Cambridge University Press

errors depend, although only slightly, on the values of $5-\bar{y}$, the deviation of the average probit response from 5; for the determination of a potency ratio they depend slightly on $\bar{y}_2-\bar{y}_1$, the difference between the mean-probit responses to the test and reference vaccine. The results which agree with the average values found on p. 64, are shown in Table 11.

Table 11. *Average fiducial limits of error for tests with 45 animals based on the pooled slope and the average value of* Σnw, *series B*

	Single determination on the ImD 50		Potency ratio—test against standard	
$5-\bar{y}$ or $\bar{y}_2-\bar{y}_1$	$P=0{\cdot}95$ (%)	$P=0{\cdot}99$ (%)	$P=0{\cdot}95$ (%)	$P=0{\cdot}99$ (%)
0	27–364	16–547	16–623	9–1106
0·25	27–367	18–551	16–628	9–1118
0·50	27–372	18–563	16–641	9–1150
0·75	26–382	17–582	15–665	8–1206
1·00	25–395	16–608	14–698	8–1291

The different vaccines compared. Table 12 shows the ratio of the ImD 50 of each vaccine to that of V 3 for determinations made on the same day. A weighted mean has been determined for each vaccine with the results shown in Table 13.

Table 12. *Ratio of* ImD 50*'s to corresponding* ImD 50 *of* V 3 *in series B*

Date of experiment	V 3	V 1	V 2	Y 2	NIH 4	WRL	V 4
30. vii. 48	1·00	0·38	1·89	—	—	—	—
6. viii. 48	1·00	37·4	152	—	—	—	—
16. viii. 48	1·00	—	—	0·13	—	—	0·32
17. viii. 48	1·00	—	—	1·29	—	—	0·47
23. viii. 48	1·00	11·2	30·4	—	—	—	—
24. viii. 48	1·00	—	15·7	1·81	—	—	—
30. viii. 48	1·00	—	25·5	0·67	—	—	—
31. viii. 48	1·00	—	—	1·89	2·59	—	—
6. ix. 48	1·00	—	—	0·18	0·87	—	—
7. ix. 48	1·00	—	—	—	0·30	1·03	—
13. ix. 48	1·00	—	—	—	4·73	4·75	—
14. ix. 48	1·00	—	—	—	—	2·77	0·26
20. ix. 48	1·00	—	—	—	—	1·10	0·22
21. ix. 48	1·00	0·81	—	—	—	—	0·14
27. ix. 48	1·00	23·9	—	—	—	—	1·14

Table 13. *Average ratios of* ImD 50*'s—series B*

Vaccine	$\frac{\text{ImD 50}}{\text{ImD 50 of V 3}}$	Fiducial limits ($P=0{\cdot}95$) (%)
V 2	19·6	39–256
V 1	6·05	38–266
WRL	2·05	48–208
NIH 4	1·34	42–238
V 3	1·00	—
Y 2	0·62	50–200
V 4	0·37	48–208

The ratio of the ImD 50 of V 2 to that of V 1 is 3·2 with fiducial limits $P = 0{\cdot}95$ of 31–326 %. Thus V 2 is less potent than V 1 (just significantly), and these two are decidedly less potent than V3. The American Standard (NIH) and vaccine WRL may be a little less potent than V 3, but the difference is not significant for the former and only just so for the latter. V 4 is definitely more potent than V 3 and Y 2 may also be, but in this case the difference is not statistically significant.

(*b*) *The determination of the* LD 50

One determination was made on each day on which tests of the vaccine were carried out. There were four doses at 10-fold intervals in each determination with fifteen mice on each dose. The values of the LD 50 were estimated by finding the best-fitting straight line connecting the probit corresponding to the mortality observed with the logarithm of the dose. Fiducial limits corresponding to $P = 0{\cdot}95$ and 0·99 were then obtained. The statistical method used was, in fact, the same as that for the values of the ImD 50.

There were no significant differences in slope from experiment to experiment. The results could therefore be interpreted with regard to a pooled slope. This average slope was 0·79 with a standard error of 0·06; this may be compared with Kendrick's values of 1·10 with a standard error of 0·11. The difference is significant and the average slope is lower in the present series of experiments. This was also the case with the dosage-response curves for the vaccines—a result which suggests that the mice used in the present series of tests were, in general, somewhat more heterogeneous.

Table 14 gives the results and their errors. The estimates of the LD 50 by the three methods (i) Reed–Muench, (ii) interpretation as a single test, (iii) interpretation with respect to a pooled slope, do not differ greatly compared with differences due to the sampling variation of the animals. When interpreted with regard to the pooled slope the average fiducial limits ($P = 0{\cdot}95$) for a single determination from sixty animals are 27–410 %.

The χ^2 values show that there were no significant departures of the probit-log dose relationship from linearity.

The analysis of variance of slopes is shown in Table 15. The variance ratio indicates that there is no significant difference between the two mean squares, and neither departs significantly from its expected value of unity.

In spite of the large errors there are significant differences between the estimates of the LD 50 obtained on different days; on the average a ratio of about 7 : 1 in two determinations is just significant at the 5 % level. The ratios of the challenge dose to the LD 50 are also given in Table 14, column (5). There is some tendency for the ImD 50 to be higher when the ratio of the challenge dose to the LD 50 is higher. This is illustrated by Table 16, where the corresponding values for the two quantities are shown for vaccine V 3. When the ratio is below 200 the ImD 50 and R are unrelated: for the values above 200 there seems to be some correlation, but the data clearly do not permit of any precise determination of the relation between the two. Proportionality, certainly, cannot be inferred.

Table 14. *The determination of the* LD50, *results of tests in series B*

Date of experiment	LD50 (Actual no. of organisms) Reed–Muench	As single test	With pooled slope	$R = \frac{50{,}000}{LD50}$	Fiducial limits* As single test Actual	%	With pooled slope Actual	%	Slope b	S.E.	χ^2	D.F.
30. vii. 48	82	69	42	1196	20–141 5–197	30–204 7–285	9–200 5–327	21–479 13–784	1·80	0·57	0·005	2
6. viii. 48	145	88	77	653	4–286 0·1–399	5–326 0·1–454	21–275 15–411	28–359 19–536	0·85	0·26	0·29	2
16. viii. 48	203	208	174	288	85–448 52–632	41–216 25–304	43–699 28–1080	25–402 16–623	1·52	0·38	0·55	2
17. viii. 48	1490	1260	1250	40	403–3780 239–6150	32–301 19–490	424–370 300–5220	34–297 24–418	0·86	0·19	3·65	2
23. viii. 48	185	108	154	324	4–430 0·1–648	4–397 0·1–599	51–466 35–660	33–302 23–428	0·63	0·19	1·42	2
24. viii. 48	1300	1310	1320	38	362–4290 189–7280	28–327 14–556	447–3830 329–5340	34–291 25–406	0·77	0·19	1·03	2
30. viii. 48	256	241	242	207	37–751 9–1110	15–312 4–462	77–759 53–1090	32–314 22–450	0·78	0·20	1·85	2
31. viii. 48	281	247	321	156	14–1010 0·3–1700	6–411 0·1–691	106–985 74–1400	33–307 23–437	0·62	0·19	2·96	2
6. ix. 48	1900	1850	1770	28	674–4830 427–7190	36–260 23–388	513–6010 354–8840	29–340 20–500	1·08	0·23	1·17	2
7. ix. 48	50	6	29	1720	0–∞ 0–∞	— —	6–138 4–226	21–477 13–779	0·50	0·27	5·97	2
13. ix. 48	281	285	241	208	80–723 38–1030	28–254 13–362	72–794 51–1160	30–330 21–480	1·02	0·25	0·35	2
14. ix. 48	600	476	452	111	105–1500 41–2370	22–31 87–50	135–1520 90–2230	30–337 20–493	0·86	0·22	1·14	2
20. ix. 48	80	—	118	424	— —	— —	15–896 8–1700	13–760 7–1440	∞	—	—	—
21. ix. 48	260	91	253	198	0·0002–679 0–∞	0·0002–744 —	78–825 53–1200	31–326 21–473	0·46	0·20	3·24	2
27. ix. 48	< 50	—	12	4340	— —	— —	1–94 1–182	12–816 6–1580	∞	—	—	—

* The upper figures of each pair give the limits for ($P = 0{\cdot}95$), the lower for ($P = 0{\cdot}99$).

Table 15. *Analysis of variance of slopes*

	Sum of squares	D.F.	Mean square	
Between slopes	15·005	12	1·250	V.R. 1·38
Residual within tests	23·620	26	0·908	

Table 16. ImD50 *of vaccine* V3 *and ratio* (*R*) *of challenge dose to* LD50

Date of experiment	ImD50 (millions)	*R*
30. vii. 48	26,414	1,196
6. viii. 48	1,055	653
16. viii. 48	1,325	288
17. viii. 48	381	40
23. viii. 48	386	324
24. viii. 48	83	38
30. viii. 48	179	207
31. viii. 48	429	156
6. ix. 48	520	28
7. ix. 48	4,323	1,724
13. ix. 48	422	208
14. ix. 48	414	111
20. ix. 48	7,210	424
21. ix. 48	8,784	198
27. ix. 48	653	4,344

The challenge dose was in each case 50,000 organisms.

STATISTICAL ANALYSIS OF SERIES C (A.F.B.S. SECOND SERIES)

Series A and B showed that the error of a single determination of the ImD50 with forty-five animals was large. As replication was the only practicable method of obtaining more accurate results it was determined to carry out simultaneous tests on vaccines V1, V2 and V3 as far as possible twice a week for 7 weeks.

Table 17 shows the actual results obtained, numbers of survivors at dose levels of vaccine of 80, 400 and 2000×10^6 and numbers of deaths when challenge doses of 50,000, 500, 50 and 5 organisms were given.

Table 18 shows the values of the ImD50 and of the LD50 calculated by the Reed–Muench method. It also shows the ratio of the challenge dose to the LD50. Where no value is given for the ImD50 of V2, its true value is extremely high, but no estimate could be made of it. For example, on 13 December 1948 none of the animals survived at any of the three dose levels of 80, 400 and 2000×10^6. On 21 December 1948 the respective numbers of survivors were 2/15, 3/15, 0/15 respectively.

There is no doubt that V2 is inferior to the other two vaccines. The ImD50 was greater in every case. V1 is less potent than V3 but not quite significantly. The ratio of ImD50 is 1·5 with fiducial limits 0·9–2·6, or 60–175 %, at $P = 0{\cdot}95$. In the first series of tests the order of potency of the three vaccines was the same, but

https://doi.org/10.1017/S002217240006126X Published online by Cambridge University Press

Table 17. *Results of fourteen experiments forming series C*

Date of experiment	Dose of vaccine millions	Survivors/total V 3	Survivors/total V 1	Survivors/total V 2	Dose of challenge culture	Survivors total
6. xii. 48	80	5/14	6/13	6/14	50,000	4/12
	400	10/15	12/15	9/14	500	6/13
	2,000	12/14	13/15	9/14	50	10/14
					5	13/15
7. xii. 48	80	5/15	4/14	1/13	50,000	0/14
	400	5/14	4/14	1/10	500	10/15
	2,000	4/15	9/13	2/11	50	11/14
					5	9/14
13. xii. 48	80	0/13	0/15	0/15	50,000	0/16
	400	3/12	2/14	0/14	500	5/12
	2,000	7/14	10/11	0/15	50	8/14
					5	13/13
14. xii. 48	80	5/14	1/14	3/15	50,000	1/14
	400	9/15	3/15	0/15	500	3/15
	2,000	7/15	4/13	1/15	50	3/14
					5	8/15
20. xii. 48	80	2/13	2 15	2/14	50,000	0/15
	400	8/14	7/16	1/13	500	11/11
	2,000	9/15	6/14	2/13	50	12/14
					5	12/12
21. xii. 48	80	3/14	3/14	2/15	50,000	2/15
	400	9/15	1/15	3/15	500	11/13
	2,000	8/14	5/15	0/15	50	11/13
					5	14/15
28. xii. 48	80	1/14	0/13	2/13	50,000	4/12
	400	6/13	4/15	0/15	500	4/13
	2,000	8/12	8/12	1/11	50	11/15
					5	5/11
1. i. 49	80	2/11	2/15	2/11	50,000	0/14
	400	5/10	1/14	2/12	500	2/11
	2,000	7/12	6/14	1/13	50	5/12
					5	5/10
3. i. 49	80	3/14	2/15	0/12	50,000	0/14
	400	9/15	1/14	2/13	500	2/13
	2,000	9/15	2/5	0/12	50	5/12
					5	2/6
4. i. 49	80	0/15	0/15	1/13	50,000	1/15
	400	3/13	1/13	2/14	500	2/11
	2,000	6/15	3/11	3/15	50	1/15
					5	1/13
10. i. 49	80	1/13	1/15	0/15	50,000	1/14
	400	1/15	4/13	2/15	500	5/13
	2,000	9/15	6/14	0/15	50	8/13
					5	12/14
11. i. 49	80	0/15	1/15	0/12	50,000	1/15
	400	1/15	0/14	2/13	500	4/12
	2,000	2/15	4/9	1/15	50	9/15
					5	10/11
17. i. 49	80	0/13	0/13	1/11	50,000	1/12
	400	6/14	2/13	0/15	500	2/13
	2,000	7/12	5/13	0/11	50	2/15
					5	4/15
28. i. 49	80	1/13	1/15	1/14	50,000	2/13
	400	9/13	4/14	3/13	500	12/15
	2,000	12/12	12/15	0/14	50	12/12
					5	12/13

https://doi.org/10.1017/S002217240006126X Published online by Cambridge University Press

V 1 was significantly less potent than V 3, with a ratio of ImD 50 of 6·1 with fiducial limits 2·3–16·1 or 38–266 %. The two fiducial ranges overlap so that the mean ratio in the two series might be the same, but the result is significantly greater in the first series of tests at the 5 % level.

Table 18. *Values of* ImD 50 *in millions and of* LD 50 *and the ratio of* LD 50 *to challenge dose in the fourteen experiments in series C*

	ImD 50						LD 50		
	V 3		V 1		V 2				
Date of experiment	Value millions	log	Value millions	log	Value millions	log	Value	log	$R = \frac{50{,}000}{\text{LD 50}}$
6. xii. 48	210	2·32	150	2·16	250	2·40	880	2·94	57
7. xii. 48	1,300	3·12	650	2·82	4,900	3·69	440	2·64	110
13. xii. 48	1,400	3·16	850	2·93	—	—	260	2·41	190
14. xii. 48	400	2·60	2,700	3·43	13,000	4·10	19	1·28	2,600
20. xii. 48	510	2·71	900	2·95	5,000	3·69	5,000	3·70	10
21. xii. 48	400	2·60	2,100	3·32	—	—	3,900	3·59	13
28. xii. 48	630	2·80	950	2·98	8,800	3·95	53	1·72	940
1. i. 49	710	2·85	1,300	3·10	—	—	11	1·04	4,500
3. i. 49	400	2·60	1,300	3·11	—	—	< 5	—	> 10,000
4. i. 49	2,000	3·30	3,300	3·52	4,200	3·62	< 5	—	> 10,000
10. i. 49	1,300	3·12	1,400	3·14	—	—	170	2·23	300
11. i. 49	12,000	4·09	2,000	3·30	—	—	150	2·17	330
17. i. 49	830	2·92	2,200	3·34	—	—	< 5	—	> 10,000
28. i. 49	240	2·37	750	2·87	—	—	3,500	3·54	14

Table 19 gives the analysis of variance for the ImD 50 of vaccines V 1 and V 3. The average fiducial limits for a single determination of the ImD 50 are 20–470 %, showing about the same accuracy as the values found in the first series of tests but somewhat less accuracy than the values found from Kendrick's data.

Table 19. *Analysis of variance of log* ImD 50*'s for vaccines* V 1 *and* V 3, *series C*

	Sum of squares	D.F.	Mean square	Variance ratio
Dates	2·9407	13	0·2262	2·38 n.s. ($P = 0·95$)
Vaccines	0·2074	1	0·2074	2·18 n.s. ($P = 0·95$)
Error	1·2364	13	0·0951	
Total	4·3845			

Average fiducial limits ($P = 0·95$) for a single determination 20–470 %.

The LD 50 and their logarithms determined by the Reed–Muench method are given in Table 18. There are significant differences at different dates as judged by the average error obtained from the first series of tests. There is a suggestion of negative correlation between the log ImD 50 and the log LD 50, but the actual values ($r = 0·29$ for V 3 and $r = 0·33$ for V 1) are not statistically significant on eleven observations.

https://doi.org/10.1017/S002217240006126X Published online by Cambridge University Press

Statistical Analysis of Series D
(Proom's tests)

Table 20 shows the numbers of survivors at vaccine dose levels of 80, 400 and 2000×10^6, also numbers of survivors when challenge doses of 5×10^n ($n = 4, 3, 2, 1$) of the culture were given, in a series of tests carried out by Mr H. Proom, Wellcome Research Laboratories (Biological Division).

Table 20. *Results of fourteen experiments forming series D*

Date of experiment	Dose of vaccine millions	Survivors/total V3	Survivors/total V1	Survivors/total V2	Dose of challenge culture	Survivors/total
28. vii. 48	80	0/15	1/15	0/15	50,000	0/5
	400	11/15	2/15	1/15	5,000	—
	2,000	14/15	11/15	0/15	500	2/10
					50	4/10
					5	9/10
6. viii. 48	80	1/15	0/15	0/15	50,000	0/5
	400	3/15	0/15	0/15	5,000	0/10
	2,000	12/15	6/15	0/15	500	0/10
					50	0/10
					5	5/10
23. viii. 48	80	1/13	0/15	0/15	50,000	0/5
	400	4/13	1/15	0/15	5,000	0/5
	2,000	7/9	5/15	0/15	500	1/10
					50	5/10
					5	10/10
13. ix. 48	80	0/15	1/15	0/15	50,000	0/5
	400	6/15	1/15	0/15	5,000	0/10
	2,000	9/15	5/15	1/15	500	7/10
					50	10/10
					5	—
13. xii. 48	80	0/15	0/15	0/15	5,000	0/5
	400	1/15	1/15	0/15	500	0/10
	2,000	5/15	1/15	1/15	50	1/10
					5	8/10
20. xii. 48	80	2/15	0/15	0/15	5,000	0/5
	400	4/15	1/15	0/15	500	6/10
	2,000	9/15	6/15	0/15	50	9/10
					5	10/10
3. i. 49	80	0/15	0/15	0/15	5,000	0/5
	400	4/15	0/15	0/15	500	2/10
	2,000	9/15	6/15	1/15	50	8/10
					5	10/10
10. i. 49	80	0/15	0/15	0/15	5,000	0/5
	400	5/15	0/15	0/15	500	1/10
	2,000	5/15	6/15	1/15	50	9/10
					5	10/10
17. i. 49	80	1/15	0/15	0/15	5,000	0/5
	400	4/15	0/15	0/15	500	1/10
	2,000	12/15	6/15	0/15	50	10/10
					5	10/10

https://doi.org/10.1017/S002217240006126X Published online by Cambridge University Press

Table 20 (*cont.*)

Date of experiment	Dose of vaccine millions	Survivors/total V3	Survivors/total V1	Survivors/total V2	Dose of challenge culture	Survivors/ total
24. i. 49	80	0/15	0/15	0/15	5,000	0/5
	400	2/15	1/15	1/15	500	4/10
	2,000	13/15	6/15	0/15	50	9/10
					5	10/10
31. i. 49	80	2/15	0/15	0/15	5,000	0/5
	400	9/15	4/15	0/15	500	5/10
	2,000	13/15	10/15	4/15	50	9/10
					5	10/10
7. ii. 49	80	5/15	1/15	0/15	5,000	0/5
	400	8/15	2/15	2/15	500	2/6
	2,000	13/15	5/15	3/15	50	9/10
					5	10/10
14. ii. 49	80	3/15	1/15	0/15	5,000	0/5
	400	8/15	3/15	0/15	500	8/10
	2,000	12/15	2/15	1/15	50	10/10
					5	10/10
21. ii. 49	80	1/15	0/15	0/15	5,000	0/5
	400	3/15	0/15	1/15	500	5/10
	2,000	7/15	3/15	0/15	50	9/10
					5	10/10

Challenge dose 28 July to 13 September 1948 inclusive was 50,000 organisms, from 13 December 1948 to 21 February 1949, 5,000 organisms.

There was no doubt that V 2 was inferior to the other two vaccines; and as not more than one animal ever survived the protection tests, there was no point in calculating an ImD 50 for this vaccine.

Table 21 shows the values of the ImD 50 and of the LD 50 calculated by the Reed–Muench method for vaccines V 1 and V 3. V 1 is less potent than V 3. The ratio of the two mean values of the ImD 50 is 4·2 with fiducial limits 2·7–6·6, or 63–157 %, at $P = 0{\cdot}95$. This result may be compared with series C (also fourteen in number) which gave a result of 1·5 with fiducial limits 0·9–2·6, or 60–175 %, at $P = 0{\cdot}95$; also with the series B result, 6·1 with fiducial limits 2·3–16·1, or 38–266 %. The value 1·5 differs significantly from the other two at the 5 % level.

Table 22 gives the analysis of variance for the ImD 50's of vaccines V 1 and V 3. The average fiducial limits for a single determination of the ImD 50 are 33–300 %. This is somewhat more accurate than in series B and C (27–365 and 20–470 %), but less accurate than in series A (49–209 %). The explanation is a slope intermediate between the other two. This slope may be estimated as 0·93 compared with $1{\cdot}46 \pm 0{\cdot}09$ for series A, $0{\cdot}79 \pm 0{\cdot}06$ for series B and 0·73 for series C. We may conclude that the colonies of mice used in different laboratories differ in the variability of tolerance of the individual mice.

The LD 50's and their logarithms are given in Table 21. They differ significantly at different dates even if judged by the average error obtained from series B. There is no significant correlation between log ImD 50 and {log (challenge dose) − log LD 50} as was suggested by Kendrick.

https://doi.org/10.1017/S002217240006126X Published online by Cambridge University Press

Table 21. *Values of* ImD 50 *in millions and of* LD 50 *and the ratio of* LD 50 *to challenge dose in the fourteen experiments in series D*

Date of experiment	ImD 50 V 3 Value millions	ImD 50 V 3 log	ImD 50 V 1 Value millions	ImD 50 V 1 log	LD 50 Value	LD 50 log	R = challenge dose / LD 50
28. vii. 48	240	2·38	980	2·99	42	1·62	1,200
6. viii. 48	840	2·92	2,600	3·42	5	0·70	10,000
23. viii. 48	730	2·86	2,800	3·44	62	1·79	800
13. ix. 48	890	2·95	2,500	3·41	970	2·98	52
13. xii. 48	2,800	3·44	61,000	4·79	14	1·13	370
20. xii. 48	940	2·97	2,400	3·38	620	2·79	8
3. i. 49	1,100	3·04	2,600	3·42	160	2·20	32
10. i. 49	2,000	3·30	2,600	3·42	160	2·20	32
17. i. 49	770	2·88	2,600	3·42	180	2·25	28
24. i. 49	900	2·95	2,400	3·38	290	2·46	17
31. i. 49	310	2·49	980	2·99	400	2·60	12
7. ii. 49	270	2·44	2,300	3·77	240	2·38	21
14. ii. 49	360	2·56	6,800	3·83	1200	3·07	4
21. ii. 49	1,500	3.17	6,700	3·83	400	2·60	12

Table 22. *Analysis of variance of log* ImD 50*'s for vaccines* V 1 *and* V 3 *from series D*

	Sum of squares	D.F.	Mean square	Variance ratio
Dates	3·035	13	0·233	4·0 sign.
Vaccines	2·713	1	2·713	46·8 sign.
Error	0·761	13	0·058	
	6·509	27		

Average fiducial limits for a single determination (30–330 %).

RECOMMENDATIONS FOR ROUTINE TESTING

The routine testing of pertussis vaccine for potency, in terms of the British Standard for Pertussis Vaccine, will be recommended in the report of the Medical Research Council's Trial (1956). The British Standard, itself, will be described in a separate communication by Armitage & Perry (1957).

The actual legal requirements for batches of vaccine to be used in children is a matter for Regulations made under the Therapeutic Substances Act; and similar requirements may be included in the monograph on Pertussis Vaccine of the *British Pharmacopoeia*.

Whatever the final requirements may be (and they will take into account the effect, in children, of the Standard itself as well as the practical difficulties of the assay) they will probably take some such form as the following: 'The estimated potency of the vaccine under test should be at least x % of that of the standard and the lower fiducial limit ($P = 0{\cdot}95$) should be at least y % of the estimated potency. Here, for example, x might be 200 and y might be 50.'

If the manufacturer can, on the average, produce a vaccine k times as potent as the standard, he can estimate how many tests he need carry out to satisfy the specification, with reasonable certainty. If $k=4$, for instance, with $x=200$, $y=50$ about five tests would be necessary to satisfy the specification. This assumes a slope of $b=1{\cdot}0$ which, on existing information, seems a fair estimate. If a slope of 1·5 were regularly obtained the number of tests needed would fall to about three, and it should be possible to use a 2×2 assay with twenty mice per dose for routine purposes (W. L. M. Perry, 1955, private communication).

The results for individual assays can be worked out by the method suggested in Appendix XV of the *British Pharmacopoeia*, 1953. The method given there for pooling the results of a number of assays might often be inapplicable for pertussis vaccine because the quantity 'g'* exceeded 0·1. This difficulty can be overcome by calculating a pooled slope from the set of five assays and reinterpreting each assay with regard to the pooled slope. The modified results so obtained can then be pooled by calculating their weighted mean in the manner described in Appendix XV. The weights would now be the reciprocals of the sampling variances of the estimates obtained by using a pooled slope.

Example

(1) An assay is carried out with twenty animals per dose and two doses of each of the standard and test preparations. The two doses of the test preparation are 50×10^6 and 500×10^6 and of the standard 200×10^6 and 2000×10^6. The numbers of survivors are 6/20 and 12/20 for the standard preparation and 5/20, 14/20 for the test preparation.

These figures are the same as in Example VI of Appendix XV, but the dose interval is now 10-fold, so that

$$I=\log 10=1 \quad \text{and} \quad b=0{\cdot}98/I=0{\cdot}98,$$

$$M=0{\cdot}06/b=0{\cdot}0612. \quad \text{Potency ratio}=1{\cdot}15.$$

The actual potency will be 4 times this or 4·60 times that of the standard.

We now have

$$A=0{\cdot}0865, \quad B=\frac{0{\cdot}0865}{I}=0{\cdot}0865, \quad g=\frac{0{\cdot}0865\,(3{\cdot}84)}{(0{\cdot}98)^2}=0{\cdot}346,$$

log fiducial limits per cent

$$=2+\frac{(0{\cdot}346)\,(0{\cdot}0612)}{0{\cdot}654}\pm\frac{1{\cdot}96}{(0{\cdot}98)\,(0{\cdot}654)}\sqrt{\{(0{\cdot}0865)\,(0{\cdot}654)+(0{\cdot}0865)\,(0{\cdot}00375)\}}$$

$$=2+\frac{(0{\cdot}0212)}{0{\cdot}654}\pm\frac{1{\cdot}96}{0{\cdot}6409}\sqrt{\{0{\cdot}0566+0{\cdot}0003\}}$$

$$=2{\cdot}0324\pm 3{\cdot}058\sqrt{(0{\cdot}0569)}$$

$$=2{\cdot}0324\pm 3{\cdot}058\,(0{\cdot}2385)$$

$$=2{\cdot}0324\pm 0{\cdot}7293$$

$$=1{\cdot}3031 \text{ to } 2{\cdot}7617.$$

Fiducial limits of error $=20$–$578\,\%$.

* $g=t^2\,V(b)/b^2$, where b is the slope and $V(b)$ is its variance, for quantal assays $t=1{\cdot}96$.

https://doi.org/10.1017/S002217240006126X Published online by Cambridge University Press

(2) Suppose this test and four similar tests gave the following values for b and B:

	b	B	$W_b = 1/B$
(1)	0·98	0·0865	11·56
(2)	1·26	0·0829	12·06
(3)	1·02	0·0850	11·76
(4)	0·82	0·0875	11·43
(5)	0·88	0·0835	11·98

The weighted mean is $\dfrac{\Sigma W_b b}{\Sigma W_b} = \dfrac{5843{\cdot}46}{58{\cdot}79} = 0{\cdot}994.$

With standard error $\sqrt{(1/W_{\bar{b}})} = 0{\cdot}130.$

We now have

$$V(\bar{b}) = 1/\Sigma W_b = 0{\cdot}0170, \quad g = t^2\ V(\bar{b})/\bar{b}^2 = \frac{(0{\cdot}0170)(3{\cdot}84)}{0{\cdot}988} = 0{\cdot}066.$$

Thus we are justified in taking the result for any one of the assays as $M = F/\bar{b}$ with a standard error

$$s_M = \frac{1}{\bar{b}}\sqrt{\{A + M^2\ V(\bar{b})\}}.$$

The log limits per cent for any one assay are now

$$2 \pm \frac{t}{\bar{b}}\sqrt{\{A + M^2\ V(\bar{b})\}}.$$

In particular, the modified result for the first assay is now

$$M = \frac{0{\cdot}06}{0{\cdot}994} = 0{\cdot}0604 \quad \text{and} \quad \text{potency ratio} = 1{\cdot}15$$

as before, but the log limits per cent are now:

$$\begin{aligned} 2 \pm \frac{1{\cdot}96}{0{\cdot}994}\sqrt{\{0{\cdot}0865 + 0{\cdot}00368(0{\cdot}0170)\}} \\ = 2 \pm 1{\cdot}972\sqrt{\{0{\cdot}0865 + 0{\cdot}0001\}} \\ = 2 \pm 1{\cdot}972\sqrt{\{0{\cdot}0866\}} \\ = 2 \pm 1{\cdot}972\ (0{\cdot}294) \\ = 2 \pm 0{\cdot}583 \\ = 1{\cdot}420 \text{ to } 2{\cdot}580. \end{aligned}$$

Limits of error are 26–380 %.

When the individual assays have been interpreted with regard to the pooled slope, a weighted mean of the five results may be obtained in the manner described in Appendix XV, p. 790.

The variance of a single estimate of M is $(A + M^2\ V(\bar{b}))/\bar{b}^2$ and the appropriate (W) is $1/V(M)$. Thus

$$\bar{M} = \frac{\Sigma(WM)}{\Sigma W},$$

$$V(\bar{M}) = \text{variance of } \bar{M} = \frac{1}{\Sigma W}.$$

The fiducial limits of error for $\bar{M}$ are $\bar{M} \pm t\sqrt{\{V(M)\}}$, where $t = 1{\cdot}96$ at the 5 % level.

https://doi.org/10.1017/S002217240006126X Published online by Cambridge University Press

In this example a 10-fold dose ratio was employed, so as to give a slope in accordance with pertussis experience, while at the same time retaining the same responses as in the example in Appendix XV of the *British Pharmacopoeia*. In practice it might be more advantageous to use some other dose ratio; say a 5-fold dose ratio with doses of 80×10^6, 400×10^6 on the test vaccine and 320×10^6, 1600×10^6 on the standard, if the test vaccine was in fact expected to be 4 times as strong as the standard.

When groups of tests are carried out as a matter of routine, it would be worth while to keep a control chart for the values of $\bar{b}$.

When say six sets of five tests have been carried out, the weighted mean of the six values of $\bar{b}$, say $\hat{b}$, is obtained and its position marked by a horizontal line on the chart. The average variance of $\bar{b}$ is obtained; it is given by $6/\Sigma W_{\bar{b}}$, the summation being over all the six sets. The standard error of $\bar{b}$, $s_{\bar{b}}$, is the square root of this; the standard error of $\hat{b}$ is $1/\Sigma W_{\bar{b}}$. Inner and outer control limits can now be marked on the chart by horizontal lines. The inner control limits can be taken at $\hat{b} + 1{\cdot}96\ s_{\bar{b}}$ and the outer at $\hat{b} \pm 2{\cdot}58\ s_{\bar{b}}$. The values of $\bar{b}$ can now be plotted on the chart, as they are obtained. As long as the true slope remains unchanged there is a probability of 0·95 that a point lies between the inner control limits, and a probability of 0·025 that it falls above the upper of the two inner limits with an equal probability of 0·025 that if falls below the lower. For the outer control limits the corresponding probabilities are 0·99 and 0·005. If a point falls outside the inner limits this can be taken as a signal that the situation needs watching; if it falls beyond the outer limits it can be taken as an indication that the slope is changing or is not 'under control'.

If the slope is under control the value of $\hat{b}$ can be used with its appropriate weight or sampling variance in calculating the result of any assay, with a consequent gain in precision. From time to time, always assuming the slope remains under control, the value of $\hat{b}$ can be revised by including the results of new tests with those of the old. This will produce a further gain in precision.

SUMMARY AND CONCLUSIONS

Four series of tests, A, B, C and D, have been considered in detail. In the first two series, to obtain the values of the ImD 50, percentages of survivors were transformed into probits, and the best-fitting straight lines connecting the probits with the logarithms of the doses were obtained by the method of maximum likelihood. This is the most accurate method to use. These results showed that the error of the test was large, but there were no significant differences in the slope of the probit-log dose relation in the results of tests from any one laboratory. The slope ($1{\cdot}46 \pm 0{\cdot}09$) obtained from series A was greater than that from series B ($0{\cdot}79 \pm 0{\cdot}06$). with a consequent increase in accuracy.

It was clear from the first two series that many mice would be needed to detect significant differences in potency between vaccines; in series C, therefore, repeated tests on three vaccines were carried out—twice a week for 7 weeks. Later, in series D, fourteen further replicate tests on the same vaccines were carried out. The

values of the ImD 50 in both these series were calculated by the Reed–Muench* (1938) method, and the error and slope were estimated from an analysis of variance of the results. The estimates of slope were 0·73 for series C and 0·93 for series D.

Table 23 shows for each of series A, B, C, D the average fiducial limits of error ($P = 0{\cdot}95$) for a single determination of the ImD 50 from forty-five animals (three doses) and for a determination of the relative potency of two vaccines using, in all, ninety animals. These tables also show the average slopes obtained. The results from series C agree closely with those from series B. The agreement is even closer than Table 23 suggests; the increase in the second series is due to the method of estimation adopted (only 13 D.F. were available for error) and not to an increase in the average logarithmic standard error. The accuracy of series D is intermediate between that of A and B or C.

Table 23. *Average fiducial limits per cent* ($P = 0{\cdot}95$) *and slope in the four series of tests*

	Single ImD 50	Potency ratio of two vaccines	Slope	and S.E.
A. Kendrick (6 experiments with 26 tests)	49–209	35–283	1·46	0·09
B. A.F.B.S. (repeated tests on 7 vaccines)	27–365	16–624	0·79	0·06
C. A.F.B.S. (repeated tests on vaccines V 1, V 2, V 3)	20–470	11–891	0·73	—
D. Proom (repeated tests on vaccines V 1, V 2, V 3)	33–300	21–473	0·93	—

On the basis of series B or C we may take the average fiducial limits for an ImD 50 obtained from forty-five animals to be 25–400 %, on the basis of series D 33–300 %. For a slope of 1·4 the corresponding result is approximately 50–200 %.

Tables 24 and 25 show (i) on the basis of series B or C and (ii) on the basis of series D, the fiducial limits ($P = 0{\cdot}95$) for tests with given numbers of mice. If the slope is 1·4 Table 24 will be approximately correct, if the number of mice in each row of the table is divided by 4.

Table 24 shows for example, that in a comparison between two vaccines, the result obtained would not with any certainty indicate a difference between them, when 180 animals in all are used (not counting those used to determine the LD 50 of the culture), unless the estimated potency ratio is > 4 or $< \frac{1}{4}$. The corresponding figures from Table 25 are > 3 or $< \frac{1}{3}$.

The standard tests employ 150 mice of which ninety are used in the comparison of the two vaccines and sixty for the titration of the challenge dose. Using the slope 0·79 obtained from series B, it may be calculated that, on the average, eight tests are needed (and series C supports this) to enable one to assert a significant

* A number of other approximate methods of estimating median effective doses have been given. Armitage & Allen (1950) and Finney (1952) have discussed their accuracy. Errors of estimation by most of these methods, including of course the Reed–Muench, are usually small compared with the sampling variation of the animals.

https://doi.org/10.1017/S002217240006126X Published online by Cambridge University Press

Table 24. *Average fiducial limits of error ($P = 0{\cdot}95$) for tests with varying numbers of mice. Based on series B*

(It is assumed that the limits for a single determination of the ImD 50 with forty-five animals are 25–400%)

Single determination of the ImD 50			Potency ratio of two vaccines		
Total no. of mice	1·96 σ*	Fiducial limits (%)	Total no. of mice	1·96 σ*	Fiducial limits (%)
10	1·2792	5·3–1900	20	1·8089	1·7–6400
20	0·9045	12– 800	40	1·2792	5·3–1900
40	0·6396	23– 430	80	0·9045	12– 800
45	0·6021	25– 400	90	0·8515	14– 700
50	0·5720	27– 370	100	0·8089	16– 640
90	0·4258	38– 270	180	0·6030	25– 400
100	0·4039	39– 250	200	0·5720	27– 370
135	0·3665	43– 235	270	0·4923	32– 320
200	0·2956	52– 190	400	0·4039	39– 250
360	0·2126	61– 160	720	0·3010	50– 200

* σ = standard error of logarithm of result.

Table 25. *Average fiducial limits of error ($P = 0{\cdot}95$) for tests with varying numbers of mice. Based on series D*

(It is assumed that the limits for a single determination of the ImD 50 with forty-five animals are 33–300%)

Single determination of the ImD 50			Potency ratio of two vaccines		
Total no. of mice	1·96 σ*	Fiducial limits (%)	Total no. of mice	1·96 σ*	Fiducial limits (%)
10	1·0057	10–1000	20	1·4221	37–2700
20	0·7112	19– 510	40	1·0057	10–1000
40	0·5029	32– 310	80	0·7112	19– 510
45	0·4741	33– 300	90	0·6705	21– 470
50	0·4498	35– 280	100	0·6361	23– 430
90	0·3352	46– 220	180	0·4741	33– 300
100	0·3180	48– 210	200	0·4498	35– 280
135	0·2737	53– 190	270	0·3871	41– 240
200	0·2249	60– 170	400	0·3180	48– 210
360	0·1676	68– 150	720	0·2371	58– 170

* σ = standard error of logarithm of result.

difference (at the 5% level) when the estimated potency is greater than 2 or less than $\frac{1}{2}$. For a slope of 1·46 (series A) the corresponding number of tests necessary is 2; for a slope of 0·93 (series D) it is 5.

Later tests (unpublished) gave slopes from $0{\cdot}99 \pm 0{\cdot}13$ to $1{\cdot}42 \pm 0{\cdot}13$, and it may therefore well be that the slopes obtained in the earlier series of assays are underestimating the accuracy now obtainable.

Tables 24 and 25 tell us, for a given number of animals, under what circumstances we can detect any difference at all. There is, however, another question which may need answering. If one vaccine is in fact 4 times as potent as the other, what is the least number of animals needed in order that the chance of failing to get a

https://doi.org/10.1017/S002217240006126X Published online by Cambridge University Press

significant result in the test will be 0·05? The number of animals may be obtained as follows. We look in the last column of Table 24 or Table 25 for limits of 25–400 %. We take the corresponding number of animals from the fourth column and multiply by 3·4. i.e. $(1{\cdot}96+1{\cdot}64)^2/(1{\cdot}96)^2$. The figure obtained is 612 or about 600. Table 25 gives about 400. Other particular cases may be similarly treated. Table 26 gives some further information.

Table 26. *Number of tests necessary in order to be 'reasonably sure' that a real difference is not missed*

	Slope		
True ratio	$b=0{\cdot}73$	$b=0{\cdot}93$	$b=1{\cdot}46$
	No. of tests		
2:1	28	18	7
3:1	12	7	3
4:1	7	5	2
5:1	6	3	1

'Reasonably sure' here means that the chance of failing to detect the difference, when the usual significance test is used, is 0·05. The number of tests necessary depends on the true ratio and on the slope 'b' of the probit-log dose response curve.

In an attempt to reduce the number of mice required a series of assays were carried out, subsequent to the series A–D, in which litter mates were used, one member of each litter being placed on each dose. The procedure did not, however, result in any appreciable increase of accuracy, commensurate with the large amount of work involved (Irwin & Standfast, 1955).

In spite of the large numbers of animals necessary, significant differences between various vaccines have been found. For instance, series B, C and D all agreed in showing that of the three vaccines V 1, V 2 and V 3 used in the pertussis field trials V 2 was much less potent than the other two and V 1 less potent than V 3 as judged by trials on mice.

The LD 50's of the cultures have been examined in detail. There were sometimes significant differences in tests carried out at different times, but no significant changes of slope in the same laboratory. It was not possible from the data examined to establish any precise relation between the ImD 50 of a vaccine and the ratio of the challenge dose to the LD 50.

Recommendations for routine testing have been made. The requirements for batches of pertussis vaccine to be used in children is a matter for Regulations made under the Therapeutic Substances Act. Whatever these may be they must be in terms of a reference vaccine such as the British Standard for Pertussis Vaccine and will probably take such a form as: 'the estimated potency of the vaccine should be at least x % of that of the standard and the lower fiducial limit should be at least y % of the estimated potency.' When these standards are known, then the manufacturer can estimate with reasonable certainty the number of tests, and so the number of mice, he will require for a complete assay as he will know the level of accuracy obtained in his own testing laboratories, and will probably have some idea of the customary or expected potency of his unknown vaccine.

This work was done under the auspices of the M.R.C. Whooping Cough Immunization Committee. We must acknowledge our indebtedness to Dr P. Armitage for many helpful discussions and to Miss Irene Allen for the great amount of arithmetical work which she carried out.

REFERENCES

ARMITAGE, P. & ALLEN, I. (1950). Methods of estimating the LD 50 in quantal response data. *J. Hyg., Camb.*, **48**, 298.

ARMITAGE, P. & PERRY, W. L. M. (1957). The British Standard for pertussis vaccine. To be published.

FINNEY, D. J. (1952). *Statistical Method in Biological Assay*, chapter 20. London: Griffin.

IRWIN, J. O. & STANDFAST, A. F. B. (1955). Litter-mate assays of pertussis vaccine. *J. Hyg., Camb.*, **53**, 106.

KENDRICK, P. L., ELDERING, G., DIXON, M. K. & MISNER, J. (1947). Mouse protection tests in the study of pertussis vaccine. *Amer. J. publ. Hlth*, **37**, 803.

KENDRICK, P. L., UPDYKE, E. L. & ELDERING, G. (1949). Comparison of pertussis cultures by mouse protection and virulence tests. *Amer. J. publ. Hlth*, **39**, 179.

LITCHFIELD, J. T. & FERTIG, J. W. (1941). On a graphical solution of the dosage-effect curve. *Johns Hopk. Hosp. Bull.* **69**, 276.

MEDICAL RESEARCH COUNCIL (1956). Report on vaccination against whooping cough. *Brit. med. J.* **2**, 454.

Minimal Requirements (1948). Pertussis Vaccine. National Institutes of Health, Bethesda, Md., U.S.A.

REED, L. J. & MUENCH, H. A. (1938). A simple method of estimating fifty per cent endpoints. *Amer. J. Hyg.* **27**, 493.

WORCESTER, J. & WILSON, E. B. (1943). A table for determining LD 50 or the fifty per cent endpoint. *Proc. nat. Acad. Sci., Wash.*, **29**, 207.

WORLD HEALTH ORGANIZATION TECHNICAL REPORT SERIES (1953). *Diphtheria and Pertussis Vaccination*, no. 61, p. 66.

(*MS. received for publication* 29. VII. 56)