AN INTERNATIONAL EXPERIMENT ON THE WIDAL-REACTION

REPORT BY A. D. GARDNER, M.D., F.R.C.S.

From the Medical Research Council's Standards Laboratory, School of Pathology, Oxford

Laboratories participating in the Experiment

Bucarest:	Institut des Serums et Vaccins.	Dr M. CIUCA.
Copenhagen:	Statens Seruminstitut.	Prof. TH. MADSEN.
		Dr F. KAUFFMANN.
Oxford:	Standards Laboratory, School of Pathology.	Dr A. D. GARDNER.
Warsaw:	L'Institut d'Hygiène de l'Etat.	Prof. L. HIRSZFELD.

(With 3 Charts in the Text)

THIS experiment had its origin in a suspicion that the Widal methods used in different centres might be of unequal accuracy and reliability. Acting on behalf of the Health Organization of the League of Nations, Prof. Th. Madsen of Copenhagen in July 1935 took steps to arrange a comparative test of methods in a number of European laboratories; and, as a preliminary, consulted various English bacteriologists on the question whether the "Standardized Suspensions" that have been issued for many years by the Medical Research Council's "Standards Laboratory" in Oxford could be accepted as standard reagents with which to compare the suspensions made and employed in the various participating laboratories.

The following Memorandum sets out the results of these preliminary conversations, and describes the method that was recommended as the standard for comparison.

MEMORANDUM

THE AGGLUTINATION REACTION IN THE DIAGNOSIS OF TYPHOID FEVER

Experimental Comparison of Methods. Organized by Dr TH. MADSEN *on behalf of the Health Organization of the League of Nations*

The purpose of the proposed experiment is to ascertain whether greater uniformity can be achieved by different workers using standardized killed suspensions by an agreed method than with the diverse suspensions and methods generally employed in the different laboratories.

The principle of the experiment is that each laboratory should titrate samples of a number of serums, divided into fractions and distributed by Dr Madsen:

(1) With Oxford "Standardized Suspensions" of *Bact. typhosum* H and *Bact. typhosum* O, by a uniform technique; and

(2) With the method and suspensions generally used in the different laboratories.

At Dr Madsen's suggestion a meeting of bacteriologists especially interested in agglutination work[1] was held in London on 26 July 1935, and it was agreed that the standardized suspensions made at Oxford on behalf of the Medical Research Council were suitable for use in the proposed experiment. Further, it was felt that the experiment might fail in its purpose unless a uniform method of employing the suspensions were adopted and practised in all detail.

In spite of considerable differences of opinion on certain technical matters, such as the size and form of the agglutination tubes and the volumes of the reagents, it was found possible to agree and unanimously to recommend the technique described herewith.

In this method certain features, such as the proportion of the volumes of the reagents, are necessitated by the character of the standardized suspensions, which are considerably less dense than those generally employed on the Continent; certain other features, such as the form of the tubes, represent a choice that had to be made between two alternative possibilities, the one chosen being preferred by a majority of the committee.

In thus recommending a standard method for the purposes of this experiment no suggestion is made that it is necessarily superior to the methods employed in the routine work of the laboratories taking part in the experiment.

Technique for use with "standardized suspensions" in the comparative experiment on the Widal reaction

(*a*) The agglutination tubes shall be of an internal diameter of approximately 6 mm.

If tubes of this type are not obtainable locally, they can be supplied free of charge by the Standards Laboratory, School of Pathology, Oxford, England.

(*b*) The total volume of the agglutinating mixture shall be 1 c.c.; consisting of 0·4 c.c. of the appropriate serum dilution and 0·6 c.c. of the standardized suspensions.

(*c*) Each serum is to be completely titrated against the two standardized suspensions of *Bact. typhosum* (H and O) by means of a series of dilutions of the serum and a control tube without serum.

Make preliminary dilutions of the serum with sterile 0·85 per cent NaCl solution: 1 in 10, 1 in 100 (1 in 1000 and so on, if necessary); and with each dilution set up three tubes as follows:

	Serum 1 in 10			Serum 1 in 100			Control
Saline solution	0·0	0·2	0·3	0·0	0·2	0·3	0·4
Diluted serum	0·4	0·2	0·1	0·4	0·2	0·1	0·0
Standardized suspension	0·6	0·6	0·6	0·6	0·6	0·6	0·6
Total serum-dilution (one in)	25	50	100	250	500	1000	Nil

[1] A. Felix (Lister Institute); A. D. Gardner (Oxford Medical Research Council); H. Schütze (Lister Institute); W. M. Scott (Ministry of Health Laboratories); P. Bruce White (National Institute of Medical Research). Also consulted, W. W. C. Topley (London School of Hygiene and Tropical Medicine).

These quantities may be measured either with graduated or with dropping pipettes.[1] The saline is measured first, then the serum and lastly the suspension. The reagents must be mixed either by the inversion of each tube, closed by a finger, or by bubbling air very gently through the fluid with a Pasteur pipette. In either case it is best to start with the control tube and work from right to left, washing the finger or the pipette between each series.

(*d*) *Water bath.* A covered bath at a temperature of between 50 and 53° C. (52° being the best).

The water level must be controlled and constant; so arranged that only a quarter or a third part of the column of fluid in the tube is immersed in the water.

(*e*) *Times of incubation.* For the H suspension 2 hours. For the O suspension 18 to 24 hours. (It is useful to inspect also the O test after 2 hours and the H test after 18–24 hours.)

(*f*) *Readings.* These are to be taken at the end of the stated incubation times, after the tubes have stood for about 20 min. at room temperature. Artificial light and a dark background are essential; the simplest apparatus being a standing electric lamp with a dark shade and a black board propped up directly behind the shade. The tubes are held against the edge of the shade, which is tilted so that the lamp, while illuminating the tubes, is just hidden from the eye.

Examine with the naked eye, noting down the significant readings, and then determine with a low-power lens the last trace of agglutination by direct comparison with the control tube.

(*g*) *Records.* The various degrees of agglutination should be recorded thus: ++++ (complete); +++ (nearly complete); ++ (well-marked partial agglutination); + (trace, visible with the naked eye); (+) (trace visible only with the lens).

If other or additional symbols are used, their significance should be stated.

The clumps formed by the O suspensions are much smaller and more finely graded than those of the H. In judging degrees of O agglutination attention should be paid to the residual opacity of the fluid, compared with the control, as well as to the relative size of the clumps.

(In H suspensions after 24 hours, especially if kept at 50–53° C. all the time, fine traces of O agglutination are sometimes seen—these should be neglected.)

It is desirable to record all significant readings, from "complete" to zero:

Example: 100 = ++++; 250 = ++; 500 = (+); 1000 = 0.

Complete titration down to the zero point should be carried out in all cases.

(*h*) Full records of the tests performed (1) with the standardized suspensions by the agreed technique, and (2) with the suspensions and technique employed

[1] Ten standard drops = 0·4 c.c. If dropped, the serum in each series of 3 tubes should be 10, 5, 2; saline to bring each to 10; suspension 15 drops. The final dilutions are slightly different, viz. 25, 50, 125, 250, 500, 1250. (This note was not included in the circulated directions.)

for routine work in the laboratory, should be sent on completion of the experiment to Dr A. D. Gardner, Standards Laboratory, School of Pathology, Oxford, England.

RESULTS OF THE INTERNATIONAL EXPERIMENT

OBJECTS OF THE EXPERIMENT

To discover whether the Widal methods used by any of the participating laboratories are liable to serious qualitative or quantitative error; and whether the use of a uniform technique with specially standardized suspensions would improve the accuracy of the test.

PRINCIPLE OF EXPERIMENT

Sixty-four human serums from cases of typhoid (or paratyphoid) fever were divided into four parts and distributed to the four participating laboratories, where they were titrated in the following ways:

(1) With the customary routine method and suspensions of each laboratory.

(2) By a prearranged, uniform technique,[1] using killed, standardized suspensions of the H and the O type, prepared at the Standards Laboratory, Oxford.

(3) Two laboratories, in which the routine Widal reactions are done with living cultures, tested the serums also with locally prepared, killed H and O suspensions.

The titrations were carried out in all the laboratories during October, November and December, 1935.

METHODS OF CALCULATION

Titres, ratios and errors

Apart from the question of correctness of actual diagnosis, which is considered elsewhere, it is necessary to assess the accuracy and consistency of the titrations done by the different laboratories by the different methods. The material for calculation consists of a large number of sets of four titrations, arranged according to the type of agglutinin (H or O) titrated and to the method used.

The titres in the protocols were all expressed in the familiar form of $++++$. $+++$, $+$ and $\pm$, or in numbers signifying the number of pluses, according to the degrees of agglutination observed.

Following the method in use at Oxford the values of the titres have been calculated from the highest positive reading; or when this was only $+$ or $\pm$, from the mean of the highest and of a stronger reading next below it. Again,

[1] The term "standard method", used sometimes in this report, implies only that the method is accepted as the standard method for the purposes of this research.

following the Oxford system, which assigns different titre values to different degrees of agglutination, the following correction-factors were employed:

$$+++ = 1{\cdot}3;\ ++ = 1{\cdot}0;\ + = 0{\cdot}7.$$

Readings of ± were disregarded.

Where the last two or more tubes have shown weak agglutination a somewhat lower figure has been adopted than where strong agglutination is recorded up to the last or the penultimate positive tube.

All the figures should be taken as approximations, since there is necessarily a considerable degree of arbitrariness and uncertainty in assigning values to the readings of different observers by different methods.

Standard of comparison

Geometrical mean

The only way of judging the accuracy of a given titre is to calculate a mean of the four tests done on each serum by a given method.

The arithmetical mean, however, which is commonly used in such calculations, is not the right one for this purpose, since agglutination reactions are performed and expressed in geometric series.

It is, for example, desirable to adopt a system whereby a pair of titres such as 1 in 50, and 1 in 200 shall give a mean of 1 in 100, since this is midway between them in the accepted technique and terminology. Moreover, with only four observations from which to calculate a mean (the experiments would have been much better if there had been more), and since in many cases one figure is widely different from the others, it seemed that the most probable true mean would be obtained by calculating only from the two middle values. Accordingly, in each set of four the highest and lowest were omitted, the logarithms of the two middle figures were taken, and the antilogarithm of their arithmetical mean was taken as the *geometrical mean* of the set.

Ratios. Each titre was then divided by the geometrical mean of the set, to give its ratio to that mean, taken as unity.

Errors. From these ratios the differences from the mean were calculated, and expressed as percentages of the mean.

Further, in order to obtain truly comparable figures the ratios below unity were first expressed as reciprocals, for the following reason:

With three titres such as 50, 100 and 200, whose geometric mean is 100, the ratios are 0·5, 1·0 and 2·0. If they are expressed directly as percentages of the mean, we get 50, 100 and 200; i.e. the first differs from the mean (100) by 50 per cent, and the third by 100 per cent. But in reality they both have the same "error", i.e. one "tube" or "dilution" away from the mean. If, however, we now convert 0·50 (being below unity) into its reciprocal, 2·0, the two ratios, and their percentage errors (100) become the same.

In support of the view that the geometrical mean gives the most probable approximation to the true mean, consider, for example, serum 9 in Table V.

Here three of the figures are reasonably near the geometrical mean of the two middle ones, whereas the use of the arithmetical mean, 1186, would throw a probably undeserved burden of error on the two lower titres. Many similar instances are to be seen in the tables. Our method, in fact, accepts the judgement of the majority, and throws the main burden of error on the more widely divergent titres.

CORRECTION FOR "CONSTANT LOCAL DEVIATION"

Several of the tables show that one or other of the laboratories tends constantly to produce high titres, others low. This phenomenon is in some cases so regular that it seems desirable, if possible, to try the effect of eliminating constant local factors before calculating the errors.

In the case of local methods the differences may be due to the use of an unusually sensitive (or insensitive) strain for making the suspensions; or to different sizes and shapes of agglutination tubes, etc. In the prearranged or "standard" method series the laboratories may differ in such matters as the height of the water in the water bath, in the values given to different degrees of agglutination, and so on.

A simple method of introducing a correction for these points in any series of observations is to divide all the *ratios* (titre/mean) of a given laboratory by the mean (median) of those ratios. This shifts the whole series of ratios upwards or downwards, till its mean is 1·0.

Having done this for all the laboratories we can then recalculate the percentage errors, and so determine the degree of irregularity in that series of observations apart from any constant local sources of error.

These calculations are not shown in the tables, but are represented by the graphs with broken lines in Charts 1 and 2. We shall see that in some cases the recalculation transforms a "bad" graph, i.e. one that shows many high errors, into a very good one.

EXAMPLE OF A COMPLETE CALCULATION

In Table II, serum 1 shows titres of

500, 650, 700, 250.

The two middle values are 500 and 650. Their logarithms are 2·75 and 2·81, with a mean of 2·75. Antilog. of 2·75 = 570 = geometrical mean. The ratios, titre ÷ mean, are:

0·88, 1·14, 1·23, 0·44.

A. Taking the reciprocals of those under 1·0, the ratios become

1·14, 1·14, 1·23, 2·27.

The percentage differences of these from 1·0 (i.e. errors) are:

14, 14, 23, 127.

B. Correction for constant local differences. The geometric means, or

medians of all the *ratios* of each laboratory in Table II are (calculation not shown):

0·89, 1·00, 1·47, 0·88.

The ratios of serum 1, given above, are each divided by their appropriate mean ratio; giving

0·99, 1·14, 0·84, 0·50.

Taking the reciprocals of those under 1·0, we have:

	1·01,	1·14,	1·19,	2·00.
Errors per cent	1,	14,	19,	100.

In this particular case the correction does not greatly reduce the errors, but in many cases it does so.

The graphs (Charts 1, 2, 3)

These are all constructed on the same principle. The ordinates give the numbers or percentages of titres that show the "errors" indicated in the abscissae. These errors are the deviations per cent of the titres from the geometric mean of each set of four. In each, the continuous line represents the errors of the uncorrected ratios, the broken line the errors after correction for constant local differences. A "good" graph consists of a very steeply falling line. The higher it rises on the left, and the less it travels to the right, the better it is. Graphs which start low and straggle far to the right are "bad", i.e. contain many high errors.

Discussion of results

A. *H agglutination with the agreed uniform method and standardized H suspension*

Table I and the upper half of Chart 1 show the titres, errors and graphical arrangements of 51 serums. The calculations of ratios and corrected errors are omitted from this and the other tables.

The range of errors is surprisingly great, though considerably less than in the other tables.

It may be said that similar titrations, if carried out by a single laboratory with workers and staff specially trained and practised in agglutination work, would show few errors as great as 50 per cent and practically none of 100 per cent.

The mean (median) errors are shown at the bottom of the table.

To correct for constant local error the mean ratios were worked out, viz. 0·92, 1·27, 1·23 and 0·64. After dividing the ratios by these factors, the recalculated errors show median values of 11, 31, 23 and 28. The graph of Lab. IV is thereby changed from poor to very good, which shows that the chief constituent of its uncorrected errors is a constant. The other graphs are also improved,

Table I. *International Widal reaction experiment. Standard method—H agglutination*

Serum no.	Titres (recalculated) Lab. I	Lab. II	Lab. III	Lab. IV*	Geom. mean	% difference from geom. mean Lab. I	Lab. II	Lab. III	Lab. IV
1	250	250	250	130	250	0	0	0	92
2	65	75	130	65	70	7	7	86	7
3	1350	3200	1500	1300	1430	6	124	5	10
4	140	150	250	100	145	4	3	72	45
5	140	150	130	70	135	4	11	4	92
6	70	150	100	65	85	33	87	18	30
7	310	300	275	100	290	7	4	5	194
8	25–	25–	—	25–	—	—	—	—	—
9	25–	25–	25–	25–	—	—	—	—	—
10	135	130	170	65	130	4	0	31	100
11	145	130	250	100	135	7	4	85	35
12	250	250	250	130	250	0	0	0	92
13	110	130	100	70	105	5	24	5	49
14	900	2500	1500	500	1175	30	113	28	132
15	25	30	70	15	30	20	0	133	100
16	170	50	300	30	95	79	89	216	212
17	250	500	300	130	275	11	82	9	113
18	50	30	250	70	60	20	100	316	17
19	1000	3200	2000	675	1410	41	127	42	109
20	625	5000	600	500	615	2	710	2	23
21	40	30	25	17	30	33	0	20	72
22	65	70	140	30	70	8	0	100	132
23	80	70	150	70	75	7	7	100	7
24	25–	25–	25–	25–	—	—	—	—	—
25	25–	25–	25–	25–	—	—	—	—	—
26	140	250	300	100	190	35	32	57	89
27	100	120	135	70	110	10	9	23	56
28	140	350	300	130	200	43	75	50	54
29	50	50	135	35	50	0	0	170	43
30	500	650	650	325	560	12	16	16	72
31	1200	3500	1500	675	1350	12	159	11	100
32	230	250	330	150	240	4	4	37	61
33	500	700	675	325	580	16	21	16	79
34	25–	25–	25–	25–	—	—	—	—	—
35	310	700	500	325	400	28	75	25	24
36	25–	25	25–	25–	—	—	—	—	—
37	600	2500	700	650	675	12	270	4	4
38	25–	25–	25–	25–	—	—	—	—	—
39	140	350	325	130	215	54	63	51	67
40	30	70	35	25	32	9	115	8	30
41	60	65	100	65	65	9	0	54	0
42	65	100	70	65	67·5	4	48	4	4
43	25–	25–	25–	25–	—	—	—	—	—
44	4400	10000	5000	2500	4680	6	114	7	89
45	25–	25–	25–	25–	—	—	—	—	—
46	25–	25–	—	25–	—	—	—	—	—
47	140	500	350	130	220	56	117	59	70
48	880	1300	1000	650	935	6	39	7	45
49	30	50	70	35	40	33	25	75	14
50	50	65	170	70	67	33	3	154	5
51	330	700	500	170	410	35	71	22	143
52	25–	25–	25–	25–	—	—	—	—	—
53	55	170	170	65	105	92	62	62	61
54	440	2500	650	350	530	21	372	23	51
55	1650	3250	2500	1350	2040	24	59	22	51
56	30	65	35	15	32·5	9	100	8	117
57	30	65	35	30	32·5	9	100	8	9
58	1060	2500	1000	500	1030	3	142	3	104
59	45	70	120	35	55	22	27	118	56
60	105	250	170	130	150	43	67	13	15
61	440	100	700	500	470	6	376	49	6
62	530	250	700	350	425	25	70	65	22
					Median	11	59	23	56

* The order of the laboratories is not the same as given on the first page of the paper.

https://doi.org/10.1017/S0022172400034859 Published online by Cambridge University Press

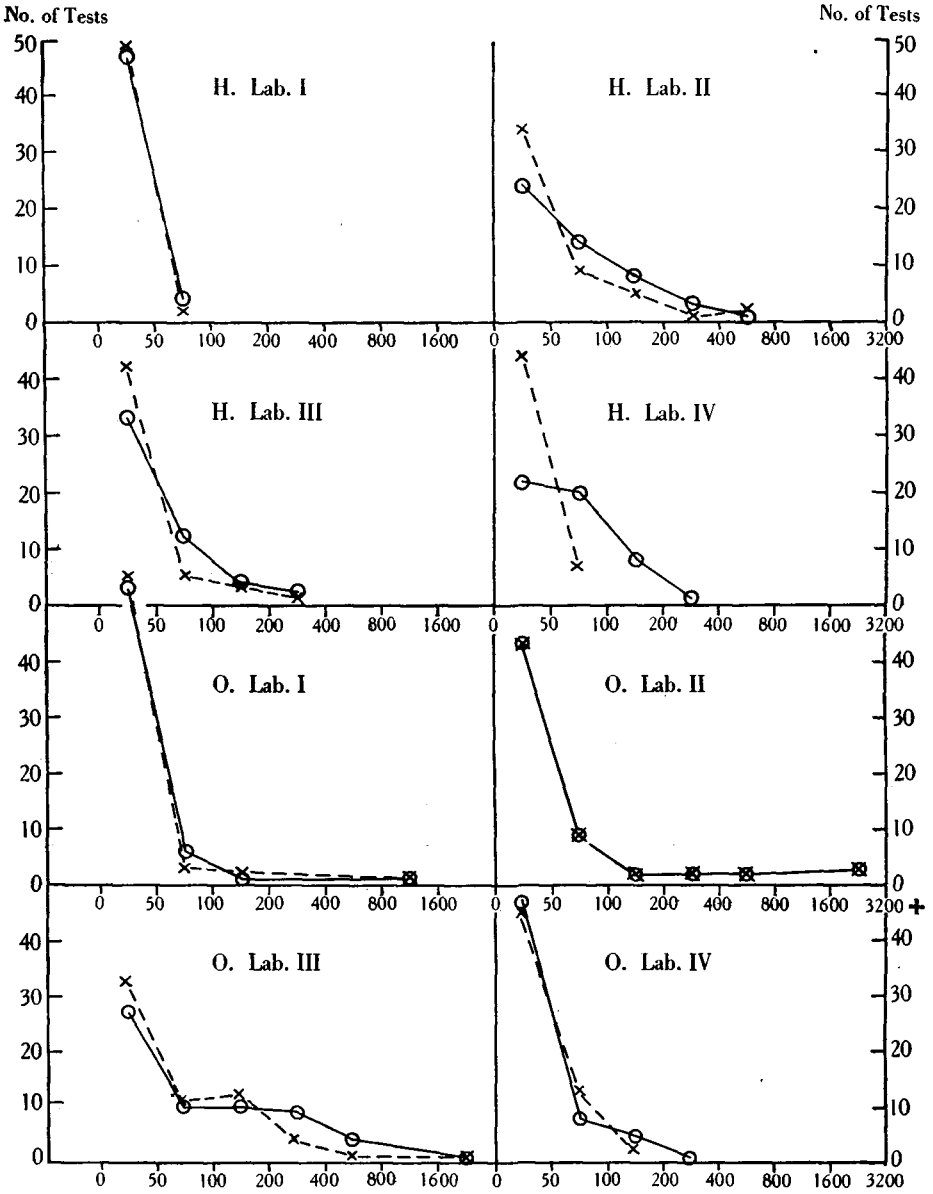

Chart 1. H and O, standard methods. Graphs of errors.

○——○ Uncorrected.
×– – –× Corrected for constant local differences.
Abscissae: deviations per cent from mean (i.e. "errors").
Ordinates: number of tests.

https://doi.org/10.1017/S0022172400034859 Published online by Cambridge University Press

but not to such a striking degree; in Labs. II and III many errors of over 100 per cent fail to be eliminated by the correction.

This set of observations proves that even with a standard technique and a single batch of suspension different laboratories often show "errors" of over 100 per cent; and that in an appreciable number of tests the errors are even grosser. It is moreover to be noted that the titres in a set of four generally differ from each other much more than they do from the mean of the four.

Some of the protocols of Lab. II (not reproduced here) show an anomalous diminution of agglutination between 2 and 24 hours (nos. 3, 51, 60, 61, 62), which can only be due to some artificial factor. The same phenomenon is seen in many of the "local H" tests. Can it be due to shaking of the tubes before observing the reaction?

B. *O agglutination with the agreed uniform method and standardized O suspension.* 61 *serums*

The agreement of the four titres and the appearance of their graphs (Table II and Chart 1, lower half) are on the whole even less satisfactory than those just described; though the mean errors are on an average about the same.

Each laboratory shows errors of over 100 per cent, and there are a number of extremely large ones recorded. Lab. IV comes out best. Lab. I would be good but for one gross error (serum 12). This has been traced to the number 17 on the label of a phial having been mistaken for 12. Since the two serums happen to have practically the same H titre, the error is only apparent in the O series.

This regrettable but illuminating mistake suggests that some of the large errors of other laboratories may be due to similar causes.

A remarkable block of three very large errors is shown by Lab. II in nos. 48, 49 and 50, suggesting some temporary failure of technique.

This laboratory (II) shows a large proportion of good observations (i.e. small errors), but also a sprinkling of gross ones, in addition to the three just mentioned. The graph of Lab. III is distinctly unsatisfactory. The graphs show that the correction for constant local errors does not effect much improvement in any.

C. *H agglutination with local H suspensions and local technique*

A survey of the titres in this series (Table III) shows such enormous differences that it is not thought possible to apply the methods of calculation used in the other series. Lab. III shows numerous high titres that can only be interpreted as due to O agglutination, e.g. nos. 9, 10, 11, 22, etc. The other laboratories generally show a fair agreement with each other and with the means of the standard H tests, which are reproduced here from Table I; but there are many very wide differences between them.

The measurements of H agglutination in living suspensions by Labs. II and

Table II. *Standard method.—O agglutination, 24 hour readings*

Serum no.	Titres (recalculated)				Geom. mean	% difference from geom. mean			
	Lab. I	Lab. II	Lab. III	Lab. IV		Lab. I	Lab. II	Lab. III	Lab. IV
1	500	650	700	250	570	14	14	23	127
2	170	250	170	150	170	0	47	0	14
3	350	500	1000	320	415	19	20	141	30
4	360	1000	1000	340	600	67	66	66	75
5	350	1000	350	320	350	0	185	0	10
6	490	1000	700	340	590	21	69	19	72
7	100	175	170	100	130	30	34	31	30
8	50	60	—	35	—	—	—	—	—
9	330	600	350	550	435	32	38	25	26
10	340	700	700	170	485	39	44	44	186
11	330	700	700	250	480	45	46	46	92
12	675	60	70	70	70	864	16	0	0
13	140	30	250	170	155	11	426	61	9
14	330	250	700	350	340	3	37	106	3
15	150	70	250	100	125	20	79	100	25
16	110	15	350	30	57·5	91	285	508	92
17	570	600	350	700	590	3	1	70	19
18	150	170	700	250	205	37	21	241	22
19	65	70	250	70	70	7	0	257	0
20	125	170	170	70	145	16	17	17	108
21	55	35	70	70	60	9	72	17	17
22	140	100	350	130	135	4	35	159	4
23	1350	1000	1700	1300	1320	2	32	29	2
24	150	250	100	130	140	7	78	41	7
25	25	15	35	15	20	25	33	75	33
26	320	170	500	170	235	36	39	113	39
27	60	15	350	120	85	41	456	312	41
28	1200	2500	5000	1700	2070	72	21	142	22
29	250	170	700	70	210	19	24	233	203
30	140	100	170	100	120	17	108	24	108
31	660	500	700	650	655	1	32	7	1
32	65	15	70	65	65	0	335	8	0
33	150	250	700	170	210	41	19	233	24
34	570	350	0 (10)	700	450	27	28	4900	86
35	130	170	170	170	170	32	0	0	0
36	100	50	130	100	100	0	100	30	0
37	140	170	170	170	170	22	0	0	0
38	140	70	700	70	100	40	43	600	43
39	150	250	350	170	205	37	22	71	21
40	110	350	100	170	185	67	89	85	9
41	35	50	25	70	40	14	25	59	75
42	15	15	70	35	25	67	67	180	40
43	40	70	700	120	90	127	28	680	33
44	500	700	1300	650	675	35	4	92	4
45	0	0	70	35	—	—	—	—	—
46	100	70	—	100	—	—	—	—	—
47	280	350	250	350	315	12	11	25	11
48	250	0 (10)	1000	350	295	18	2900	240	19
49	1750	70	7500	1300	1500	17	1900	400	15
50	300	0 (10)	1700	350	325	9	2900	423	8
51	50	50	170	50	50	0	0	240	0
52	150	350	700	170	245	64	43	186	43
53	145	250	260	130	190	32	24	37	47
54	280	250	350	170	265	5	6	32	56
55	250	500	700	170	355	43	41	97	108
56	15	25	15	25	20	33	25	33	25
57	130	250	170	100	150	15	67	13	49
58	250	250	700	170	250	0	0	180	47
59	170	350	700	170	245	43	43	186	43
60	340	650	700	350	475	39	27	47	35
61	320	500	500	250	430	35	16	16	72
62	320	350	350	375	350	9	0	0	7
63	170	170	350	150	170	0	0	106	14
64	440	500	1700	350	470	6	6	262	35
					Median	20	33	70	25

https://doi.org/10.1017/S0022172400034859 Published online by Cambridge University Press

Table III. *Local methods.—H agglutination*

Serum no.	H titres Killed suspensions Lab. I	Lab. II	Lab. III	Lab. IV	Standard H geometrical mean	H titres Living suspensions Lab. II	Lab. III
1	175	280	700	175	250	300	350
2	70	140	350	110	70	300	170
3	1650	280	1700	1100	1430	200	350
4	125	70	500	70	145	120	70
5	75	50	350	110	135	35	350
6	55	70	500	70	85	70	350
7	320	800	500	110	290	300	350
8	25–	25–	—	25–	—	25–	—
9	25–	25–	350	25–	—	25–	1700
10	80	30	500	55	130	30	350
11	165	70	1000	150	135	70	350
12	230	140	500	150	250	70	350
13	115	100	350	110	105	70	170
14	1000	560	2500	550	1175	70	350
15	30	70	250	15	30	140	100
16	165	35	1000	30	95	15	170
17	175	280	700	170	275	140	350
18	45	140	700	35	60	140	350
19	1000	1100	7000	550	1410	1100	1500
20	1000	2200	700	350	615	2200	600
21	25	35	70	15	30	30	70
22	70	50	700	35	70	70	100
23	96	280	1700	70	75	560	250
24	25–	35	170	25–	—	25–	100
25	25–	25–	35	25–	—	25–	35
26	150	140	350	70	190	280	170
27	85	140	700	50	110	70	170
28	100	560	1000	100	200	560	250
29	35	70	1000	25	50	280	350
30	355	280	700	250	560	560	500
31	1000	560	3500	1000	1350	2200	1700
32	165	280	700	70	240	560	170
33	350	1100	1000	170	580	560	350
34	25–	25–	25–	25–	—	35	170
35	280	560	500	170	400	560	1000
36	15	35	70	0 (10)	—	25–	170
37	675	1120	1000	500	675	560	700
38	25–	25–	250	25–	—	25–	170
39	140	280	500	150	215	280	170
40	30	140	300	35	32·5	140	350
41	55	280	250	50	65	140	170
42	65	140	250	50	67·5	140	70
43	25–	25–	50	25–	—	25–	170
44	2850	3200	12500	3500	4680	1120	3500
45	25–	25–	50	25–	—	25–	25–
46	25–	25–	—	25–	—	70	—
47	130	1120	1500	170	220	1120	1700
48	700	1600	2000	1000	935	1120	1000
49	30	70	3000	35	40	70	3000
50	50	200	700	70	67	35	700
51	300	560	700	100	410	140	500
52	25–	25–	500	25–	—	35	1000
53	55	100	300	70	105	70	70
54	440	1120	1500	350	530	140	1000
55	1440	2240	7000	1300	2040	280	1000
56	30	25–	150	25–	32·5	25–	25–
57	30	50	600	25–	32·5	100	170
58	1130	1120	13000	500	1030	280	2000
59	30	560	600	25	55	280	170
60	90	140	700	100	150	280	250
61	290	280	1000	350	470	140	500
62	500	200	1000	350	425	140	700
63	25–	35	150	25–	—	140	350
64	25–	35	600	25–	—	140	350

https://doi.org/10.1017/S0022172400034859 Published online by Cambridge University Press

III, done by different methods, are given for comparison on the right of the table. In some instances it seems that O agglutination is confusing the issue, especially in Lab. III (e.g. nos. 6, 23, 36, 40, 49, 63, 64).

The protocols of Lab. II show, in many instances, the same remarkable reduction of agglutination between the 2 and the 24 hour readings that we have already mentioned (p. 133). This sometimes amounts to three tubes (no. 59), and since the reaction itself cannot proceed backwards, the effect must be artificial.

D. *O agglutination with local O suspensions and local technique.* 62 *serums*

In Table IV the titres, means and errors are set out, as in Tables I and II. The mean (median) errors are much higher than in the standard O series of Table II. This is to be expected, since the sensitiveness of the suspensions used in the different laboratories is likely to differ and thereby to cause considerable constant deviations from the mean. Technical differences must also be expected to contribute.

The uncorrected graphs (Chart 2, top, continuous lines) are on the whole worse than those of the standard O method. The order of merit of the four laboratories is IV, I, II, III. The two "combined" graphs at the bottom give a direct comparison of the standard and the local O series (Tables II and IV). Here the errors of all four laboratories are shown in a single graph, the ordinates representing the percentage of observations showing the degrees of percentage error indicated in the abscissae, both before and after correction for the constant local errors. The standard method shows a considerably smaller proportion of large errors (100–800 per cent); and a larger proportion of small ones. Extreme errors, however, are rather more common in the standard than in the local graph. These as we have suggested are probably due to gross technical mistakes. As might have been predicted, the correction-process improves the local graph considerably while making little difference to the standard one.

E. *Local routine Widal methods: Highest H or O titres* (Table V and Chart 3)

Since the routine methods of the laboratories differ considerably in principle, Labs. I and IV using killed H and O suspensions, Lab. II living HO suspensions with an orginal system of time and temperature, and Lab. III living HO suspensions with a more usual technique, it has been impossible to compare separately the H and the O factors in this section of the work.

It has, however, been thought useful to compare the full titres of the serums as observed by the various laboratories irrespective of the type of agglutination. Thus in some of the serums it is the H titres that are compared, in others the O, and in a few it is a mixture of both.

Table IV. *Local methods—O agglutination*

Serum no.	Titres (recalculated) Lab. I	Lab. II	Lab. III	Lab. IV	Geom. mean	% difference from geom. mean Lab. I	Lab. II	Lab. III	Lab. IV
1	340	280	700	350	345	1	24	103	1
2	90	140	300	175	155	72	11	93	13
3	340	280	3000	350	345	1	24	770	1
4	360	280	1500	550	445	24	59	237	24
5	360	35	600	175	250	44	614	140	43
6	385	200	600	270	325	18	61	85	21
7	75	70	300	70	73	3	4	311	4
8	30	25–	—	25–	—	—	—	—	—
9	265	140	1200	700	430	61	178	179	63
10	250	140	1500	350	295	18	113	409	19
11	285	280	2500	700	445	56	59	462	57
12	655	15	300	70	145	351	900	107	108
13	145	15	135	70	95	53	525	42	35
14	340	140	600	170	240	42	72	150	41
15	160	35	325	70	105	52	203	196	49
16	95	15	600	35	57·5	65	285	945	64
17	700	200	600	350	455	54	127	32	30
18	180	35	600	70	115	56	233	422	64
19	75	140	600	55	105	41	33	471	92
20	80	140	300	170	155	92	11	94	10
21	30	35	100	35	35	16	0	186	0
22	125	70	300	170	145	16	108	107	17
23	1500	800	1350	1700	1430	5	79	6	19
24	100	140	300	170	155	56	11	93	10
25	25	0 (10)	30	15	20	25	100	50	33
26	190	140	650	170	180	5	28	261	6
27	30	0 (10)	250	70	50	67	400	400	40
28	1470	800	1300	700	1010	45	27	29	45
29	180	200	1000	350	270	50	35	270	30
30	70	100	600	170	130	85	30	362	31
31	655	280	1350	700	685	4	144	97	2
32	40	70	300	170	110	178	56	173	54
33	125	200	1000	170	185	47	8	440	9
34	500	280	600	350	420	19	49	43	21
35	100	140	170	350	155	54	11	10	126
36	100	50	500	70	85	18	70	488	22
37	100	140	600	350	220	122	56	173	59
38	80	140	1000	350	220	178	56	354	59
39	100	140	100	100	100	0	40	0	0
40	125	280	350	150	205	64	36	71	37
41	35	70	350	70	70	100	0	400	0
42	15	10	70	10	12·5	20	25	460	25
43	44	35	300	70	55	25	56	446	27
44	590	560	2500	1000	765	30	37	227	30
45	0 (10)	0 (10)	1000	35	18·5	85	85	4940	89
46	65	70	—	70	—	—	—	—	—
47	290	560	2000	700	630	117	12	218	11
48	260	140	2500	700	425	64	203	488	65
49	1500	1120	6000	1700	1600	6	43	275	6
50	330	70	1000	350	340	3	376	194	3
51	55	35	650	100	75	37	113	767	33
52	150	280	1500	350	315	108	12	376	11
53	125	140	300	170	155	23	11	94	10
54	270	200	600	170	230	18	15	161	35
55	290	280	1500	170	285	2	2	426	67
56	15	70	60	35	45	203	56	33	28
57	70	140	600	350	220	212	56	172	59
58	280	560	1500	70	400	43	40	275	456
59	180	560	600	350	440	144	27	36	25
60	230	280	1500	700	445	92	59	237	57
61	260	280	600	350	315	21	12	90	11
62	300	140	1000	170	225	33	61	345	32
63	150	140	600	170	160	6	14	275	6
64	440	560	1500	1700	925	108	64	62	84
					Median	445	56	195	30

https://doi.org/10.1017/S0022172400034859 Published online by Cambridge University Press

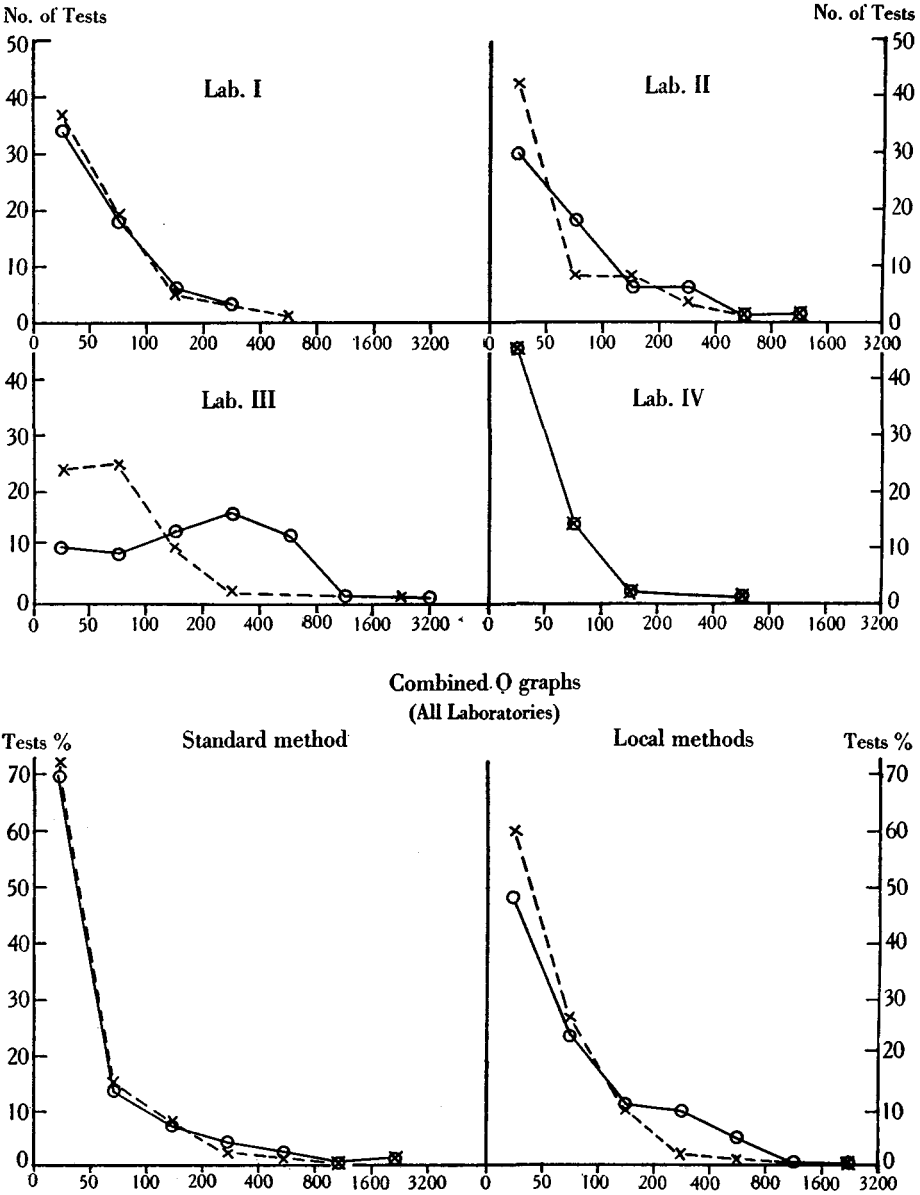

Chart 2. O, local methods. Graphs of errors.

○——○ Uncorrected.
×– – –× Corrected for constant local differences.
Abscissae: deviations per cent from mean (i.e. "errors").
Ordinates: number or percentage of tests.

https://doi.org/10.1017/S0022172400034859 Published online by Cambridge University Press

Table V. *Local routine Widal methods. Highest titre, H or O*

Serum no.	Lab. I (killed)	Lab. II (live)	Lab. III (live)	Lab. IV (killed)	Geom. mean	% difference from geom. mean			
						Lab. I	Lab. II	Lab. III	Lab. IV
1	340	280	700	350	345	1	24	103	1
2	90	560	350	175	245	170	129	43	41
3	1650	560	1000	1100	1050	57	89	5	5
4	360	560	350	550	450	25	24	28	22
5	360	280	500	175	320	12	14	56	19
6	385	560	500	270	440	14	27	14	64
7	320	400	500	110	360	12	11	39	223
8	30	25–	—	25–	—	—	—	—	—
9	265	280	3500	700	445	67	59	686	57
10	250	140	700	350	295	18	113	137	19
11	285	280	500	700	360	27	28	39	94
12	655	140	700	150	315	108	127	122	108
13	145	140	350	110	145	0	3	141	32
14	1000	140	500	550	525	91	270	5	5
15	160	70	170	70	105	53	49	62	49
16	165	15	1000	35	75	120	400	1235	113
17	700	560	1000	350	630	11	12	59	79
18	180	140	350	70	160	13	14	119	127
19	1000	560	3500	550	750	33	33	367	37
20	1000	1120	700	350	840	19	33	21	144
21	30	15	350	35	32·5	9	117	975	8
22	125	140	170	170	155	34	11	10	10
23	1500	800	700	1700	1095	37	37	56	55
24	100	70	350	170	130	30	85	169	31
25	25	25–	70	25–	—	—	—	—	—
26	190	200	350	170	195	3	3	79	15
27	85	70	170	70	77·5	10	11	119	11
28	1470	800	350	700	750	96	7	113	8
29	180	280	1000	350	315	75	12	217	11
30	355	280	700	170	315	13	12	122	85
31	1000	2240	2000	1000	1415	45	55	38	45
32	165	200	350	170	185	12	8	89	9
33	350	560	500	170	415	19	35	20	144
34	500	170	350	350	350	43	104	0	0
35	280	560	1000	350	440	56	27	127	25
36	100	0 (10)	350	70	85	18	733	312	22
37	675	560	1100	500	620	9	11	81	24
38	80	70	700	350	165	108	138	324	112
39	140	280	700	150	210	49	33	233	41
40	125	140	350	150	145	16	3	142	3
41	55	170	350	70	110	100	54	218	56
42	65	170	170	50	105	61	62	62	108
43	44	0 (10)	350	70	55	25	456	536	27
44	2850	2240	8750	3500	3125	9	39	180	12
45	0 (10)	0 (10)	170	35	20	100	100	750	75
46	65	70	—	70	—	—	—	—	—
47	290	1120	1500	700	890	203	26	68	27
48	700	560	1500	1100	880	25	56	71	25
49	1500	1120	3500	1700	1600	6	43	118	6
50	330	25	1750	350	340	3	1330	415	3
51	300	140	1300	110	205	46	47	533	85
52	150	280	2000	350	310	108	11	545	13
53	125	70	170	170	145	16	108	17	17
54	440	280	1700	350	390	13	39	336	11
55	1440	280	2800	1500	1475	2	426	90	2
56	30	0 (10)	170	35	32·5	9	223	424	8
57	70	70	350	350	155	122	122	126	126
58	1130	280	3900	550	785	44	178	397	43
59	180	280	700	350	310	72	11	126	13
60	230	560	700	700	630	178	12	11	11
61	290	280	700	350	315	9	12	122	11
62	500	140	1400	350	415	21	194	237	19
63	150	140	700	170	160	6	14	337	6
64	440	560	350	1700	495	12	13	41	244
					Median	25	37	119	25

https://doi.org/10.1017/S0022172400034859 Published online by Cambridge University Press

For Labs. I, III and IV the H titres given are those observed after 2 hours, and the O after 24 hours. For Lab. II the titres for both H and O are those observed after 20 hours in the cold followed by 2 hours at 50° C. Here the H titres are often lower than they were after 20 hours in the cold, so that it is difficult to judge which is the real titre.

The titres (Table V) show numerous and large differences from their geometrical means. Some of the errors are exceptionally great (nos. 9, 16, 21, 36, 45, 50, etc.).

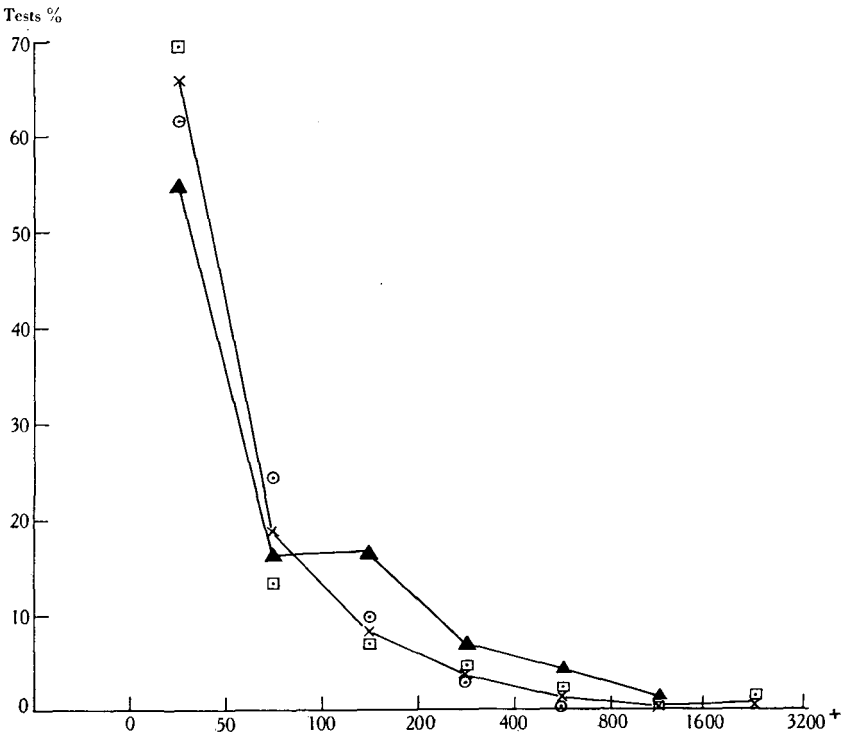

Chart 3. Local Widal-routine methods (▲——▲) compared with combined results of standard H and O methods (×——×).

H = ⊙, O = ⊡, × = mean of H and O.
Abscissae: deviations per cent from mean (i.e. "errors").
Ordinates: percentage of tests.

The titres of Labs. I, II and IV generally approximate fairly well to the mean, but those of Lab. III run far above it.

The median of Lab. III's ratios (not shown in the table) is as high as 2·19, whereas Lab. I gives 0·92, Lab. II 0·88, and Lab. IV 1·0; this last being the ideal mean ratio.

The sum of the median errors is much higher than that of either the standard H or the standard O series.

Chart 3 gives a graphical comparison of the errors of Table V with those of the standard H and O series combined (Tables I and II).

The graph with triangular points gives the percentages of Routine-Widal titres that differ to various degrees from the mean of their set of four. The other graph shows the mean of the combined H and O graphs of all laboratories, standard method, expressed as percentages. None of the figures are corrected for constant local errors.

It is sufficiently clear that the standard graph is a better one than the local, since it has more low errors and fewer high ones—though a small group of extreme errors in the O series (Table II, nos. 48–50) reduces the superiority of the standard graph (see also Table VI).

Table VI. *Percentage of titres showing various differences percentage from the geometrical mean*

Method	0–50	50–100	100–200	200–400	400–800	800–1600	1600–3200+	0–100	100–3200+
Standard technique H	62·3	24·5	9·8	2·9	0·5	0	0	87	13
Standard technique O	69·7	13·5	7·4	4·9	2·5	0·1	1·6	83	17
Local killed suspensions O	48·0	23·8	11·3	10·2	5·6	0·8	0·4	72	28
Local Widal methods H + O	55·0	16·4	16·8	7·0	4·1	1·2	0	71	29

From this we may deduce that with a standard technique and uniform suspensions a greater uniformity can be achieved than with different local methods and suspensions; but that the standardization of method does not remove all the causes of gross error.

STRAINS OF *BACT. TYPHOSUM* USED FOR LOCAL SUSPENSIONS

The strain used by Lab. III has been examined and found to be semi-rough. It is slightly unstable in saline suspensions, and gives a partial but definite Bruce-White's (Millon's) reaction for roughness. This explains its general hypersensitiveness. It has still plenty of O antigen, but is not as rich in flagella as the average strain.

SUMMARY AND CONCLUSIONS

Accuracy of titration may be taken as a measure of the reliability of the methods and work of a laboratory.

With such a small number of titrations as four, a convincing mean, representing the most probable true titre, is hard to calculate. A fairly close approximation, however, is possible in the form of the geometrical mean.

It is also possible to correct for constant local factors which cause the titrations of a given laboratory to run above or below the mean. By all methods of titration a surprisingly large proportion of large errors are recorded.

The most consistent results, with the smallest proportion of large errors, are obtained by the prearranged standard method with a single standardized suspension. The H titres are rather more consistent than the O; probably because H agglutination is easier than O to read correctly.

Local methods give greater differences than the agreed (standard) method. Since the H suspensions used in one laboratory were also sensitive to O agglutination the supposed H titres are unreal, and cannot be treated statistically.

Local O methods give a considerably greater range of error than the standard method, as is shown by a comparison of the mean errors of Tables II and IV, and by the percentages of errors over and under 100 per cent given in Table VI.

The full titres (H or O or both) of the serums measured by local routine Widal methods, with living suspensions in two laboratories and killed ones in the other two, show a considerably greater proportion of large differences than are shown by the titrations with the standard method. They do not, however, differ more widely than the titres obtained with local O suspensions.

The results of the laboratories that use killed H and O suspensions in routine work are more consistent and accurate than those of the laboratories that use living suspensions.

It is suggested that gross technical mistakes are a commoner cause of error than is usually supposed. Errors of dilution of the serums and other technical variations must also be responsible for a part of the very wide range of differences observed.

The use of old strains, tending to roughness, is to be avoided.

(*MS. received for publication* 19. x. 1936.—Ed.)