BELYAVIN, G., WESTWOOD, J. C. N., PLEASE, N. W. & SMITH, WILSON (1951). *J. gen. Microbiol.* 5, 546-552.

Influenza Haemagglutination

An Evaluation of the Photoelectric Titration Method

By G. BELYAVIN, J. C. N. WESTWOOD, N. W. PLEASE
AND WILSON SMITH

*Department of Bacteriology, University College Hospital Medical School
and the Department of Statistics, University College, London*

SUMMARY: A modification of the Hirst & Pickels photoelectric densitometer and its application to work on influenza virus haemagglutination are described. The methods used are believed to reduce to a minimum the inherent margin of error of virus and serum antibody assays based upon the haemagglutination reaction. In order to make full use of data thus obtained precise assessment of the significance of differences of titre is essential, and for this purpose two simple statistical tests have been applied and their application illustrated by an example.

Several different methods, based upon the haemagglutination reaction, are available for titrating influenza viruses and serum inhibitory substances. Only those, however, which minimize the subjective element in the assessment of agglutination end-points are suitable for certain types of investigation, notably investigations concerned with relatively small differences between samples. The method originally used in this laboratory was a modification of that devised by Hirst & Pickels (1942) in which a photoelectric densitometer is employed to measure the light transmission through each tube of a titration series, and the 50 % agglutination end-point is calculated from the galvanometer readings (Smith & Westwood, 1949). During the past two years several modifications have been introduced and a new photoelectric densitometer constructed with the object of reducing the inherent margin of error of the test to a minimum. The employment of such a method is, however, only justified if full use is made of the data obtained. This requires a precise assessment of the significance of deviations of titre. It is the purpose of this paper to describe the method now in use, how the error of the test has been determined; and to outline two statistical tests for assessing the significance of titre differences.

Technique of the test

A twin-cell instrument of the type used by Needham for recording bacterial growth curves was used (Needham, 1947). The general layout is shown in Fig. 1. The light from a 6 V., 24 W. bulb is allowed to fall simultaneously on two matched photo-cells. The cells are connected in opposition and bridged by a sensitive, centre-reading galvanometer. So long as the same light intensity is falling on both cells, the system remains balanced. When a tube of fowl erythrocyte suspension is placed in front of one cell, the decrease of the light

intensity falling on that cell causes a potential difference to develop across the galvanometer terminals, which is registered by a deflexion of the needle. This deflexion is brought back to zero by increasing a potentiometer resistance inserted in the circuit of the opposing cell. The potentiometer reading is thus an accurate measure of the density of the red-cell suspension.

As fluctuations in the intensity of the light source affect both photo-cells simultaneously, the instrument is self-compensating. The 'base line' reading of the instrument is checked by measuring the density of a standard red screen.

The sensitivity of such an instrument is determined by the intensity of the light source, the potentiometer resistance and the sensitivity of the galvanometer. For our instrument the specifications were as follows:

Galvanometer. Weston centre-reading, d.c. type, sensitivity $2\mu\text{a.}$ per division.

Potentiometer. Enclosed pattern, 1 K. resistance, 15 W.

Lamp. Car headlight bulb, 6 V., 24 W.

Fig. 1. *A*, galvanometer; *B*, potentiometer dial; *C*, lamp housing; *D*, photoelectric cells; *E*, tube holder; *F* 1, 2, lens holders; *G*, standardizing screen (hinged to lens holder, *F* 2); *H*, holder for infra-red absorbent screen and balancing filters; *I*, 6 V. transformer.

Hard wood blocks containing holes drilled vertically to a depth of $1\frac{1}{2}$ in. were substituted for the more conventional racks. With tubes inserted the clearance is only just sufficient for smooth manipulation. Thus the tubes containing the test mixtures are held in a strictly vertical position so that maximum sedimentation of the agglutinated cells to the bottom is ensured. This avoids spuriously high density readings due to the accumulation of cell clumps on the sides of the tubes, an error which is common when racks of the ordinary type are used. Glass tubes, 7.0×0.8 cm., are carefully selected for internal diameter by checking the column height of a 2 ml. volume of water to within ± 1 mm. of a standard. They are then standardized by filling with a suspension of mammalian red blood corpuscles, and taking density readings; all tubes giving a reading deviating more than ± 1 potentiometer division from a standard are rejected. Mammalian cells are used because of their greater stability in suspension, and to obviate the streaming effect which occurs in fowl cell suspensions.

The test must be precisely timed. At the end of 75 min. standing, sedimentation is still proceeding rapidly in those tubes showing agglutination, and the 50 % end-point is moving towards the higher dilution region. A delay in reading of 1 or 2 min. may produce a considerably enhanced titre.

Where the 50 % agglutination end-point falls between two dilutions it is calculated by logarithmic interpolation as in Hirst's original method (Hirst & Pickels, 1942).

The error of the test

Since the titre is calculated by logarithmic interpolation it may theoretically attain any value between consecutive dilutions, and may thus be regarded as a continuous variable. The standard error of a titre may be estimated by replicating the titration, but it must be noted that the actual observations obtained may be biased by the operator performing the test. The original experiments carried out to measure the standard deviation of both virus and antibody titrations were so designed that the operators' bias could be extracted from the data by analysis of variance. This technique had the further advantage that it was possible to obtain additional information from the experiment, e.g. whether there was a significant difference in the titres between operators, and whether alternative methods of making the twofold dilutions produced different results.

Table 1. *Replicate titrations by three different operators*

Row	Titres of a virus suspension estimated in replicate tubes by three operators					
	M. A. W.		G. B.		J. C. W.	
	Single pipette	Multiple pipette	Single pipette	Multiple pipette	Single pipette	Multiple pipette
1	290	282	285	310	269	261
2	277	255	209	286	331	255
3	282	233	252	282	266	249
4	243	251	265	320	289	255

Table 2. *Partial analysis of variance of titres in Table 1*

Source of variation	Sum of squares	D.F.	Mean square	Variance ratio	P
Operators	592.40	2	296.20	—	—
Methods	15.08	1	15.08	—	—
Operator-Methods	7264.27	2	3632.14	6.95	< 0.01 > 0.001
Residual Error	9403.00	18	522.39	—	—

Each of three operators carried out eight titrations on the same virus preparation, four in which a single pipette was used for each dilution series and four in which a fresh pipette was used for each dilution. The same fowl cells were used throughout. The titres obtained in this experiment are tabulated in Table 1, and the analysis of variance is shown in Table 2. The analysis is incomplete, because the errors due to the operators and operator-methods are not separated, but in the present instance the information obtained by separating these errors is irrelevant. It is clear that there is no significant difference between operators or methods of dilution *per se*, but the operators differed in their use of the dilution methods, and the difference is highly significant. The standard deviation of the whole sample is the square root of the error mean square.

This type of experiment has been repeated a number of times, and the results compared with the standard deviations of samples from individual operators. The coefficient of variation for the virus titrations varied between 5.5 and 9.5 % with a mean of 7.8 %, and for the serum antibody titrations between 5.5 and 9 % with a mean of 6.6 %.

Significance of differences between titres

There are two methods of assessing significant differences. An arbitrary limit may be adopted or the difference may be assessed against the error of the experiment. The assignment of arbitrary limits is unsatisfactory; they may either be set so finely relative to the error of the test as to lead to dangerous conclusions, or so widely as to be wasteful of information. It is clearly of little value to perfect an exacting technique, merely to subject the data it provides, to wide arbitrary limits of significance. Nevertheless, this method is that most widely used in haemagglutination titration work.

Tests of significance between single observations

The determination of the significance of the difference between the titres of two samples is not difficult. The comparison may be made as precise as is necessary, merely by replicating the titrations of each sample a sufficient number of times and comparing the means of such replications, using 'Student's' t test. The error in such a case is directly estimated from the data. This method is statistically ideal, and it must be clear that it requires no previous standardization of the test for error.

It is not, however, possible in haemagglutination experiments always to carry out an adequate number of replications, and the question frequently arises as to whether two single observations are or are not significantly different. Data consisting of two single measurements cannot supply their own estimate of error. This must, therefore, be pre-determined. If the variable were normally distributed and its standard error, σ , were constant for all its values, then we could use the test criterion

$$\frac{x_1 - x_2}{\sqrt{(2S)}} = t,$$

where x_1 and x_2 are the two single observations and S is the previously determined estimate of σ .

But in haemagglutination titrations the standard error (σ) of the titre (x) is not constant for all values of x , although it appears that the ratio of the standard deviation to the mean (coefficient variation) does remain constant. Denoting the coefficient of variation by Vc we have $Vc = \sigma/x = \text{constant}$ and we can, therefore, obtain an estimate of σ for any particular value of x from a previously determined estimate of Vc . It thus appears that the test criterion in this case should be

$$\frac{x_1 - x_2}{Vc\bar{x}} = M_t, \quad (1)$$

where $\bar{x} = \frac{1}{2}(x_1 + x_2)$. In order to calculate significance points for this expression it is necessary to consider the distribution of M_t which depends on the distribution

of x . In haemagglutination titrations it would appear that the titre, x , is not in fact normally distributed, but $\log x$ is so distributed. On this assumption, and assuming that the correct value of Vc is used, it can be shown that the 1% significance point of M_t is approximately 3.6 and the 5% point is 2.8. It is interesting to note that these significance points are almost constant for true values of Vc between 4 and 10 %. It is obvious, however, that if the previously determined value of Vc were larger than the true value, then the ratio $(x_1 - x_2)/Vc\bar{x}$ would always be too small and significant differences between x_1 and x_2 would be missed. Also if the estimate of Vc were too small, then the test would give too large a number of significant differences. Since the estimates of Vc varied between 5.5 and 9.5 % it was decided, before the significance points had been fully examined, to take the value 10 % thus ensuring that differences were not wrongly considered significant, and the 1 % significance point of M_t was taken as 3. It appears, however, that the use of this smaller value was compensated for by the use of the overestimate of 10 % for Vc . It was found that if the true value of Vc were 8.8 % and the value used in calculating $(x_1 - x_2)/Vc\bar{x}$ was the overestimate of 10 % then the 1 % significance point of M_t would be reduced to 3. From the data the average estimate of Vc was 7.8 %, which suggests that 3 was a safe value.

The logarithmic transformation. It requires such a very large number of observations to obtain the exact distribution of a variable that a direct test was impracticable in this case. However, tests of significance are less affected by slight deviations from normality than by changes in the standard error. Though the standard error of x increases as x increases, the standard error of $\log x$ remains constant. It was, therefore, assumed that $y = \log x$ was normally distributed, and this assumption was supported by the fact that the titre was calculated by interpolation of values of $\log x$. It was on this assumption that significance values of M_t were obtained.

However, it seems more convenient to use the normal variable y in testing for differences, since tables of the normal curve and 'Student's' t distribution have been widely published. The significance test of two individual values of x is made in the following way. An estimate of the constant standard error of y (σ_y) is made from previous experiments and it is seen that

$$\frac{y_1 - y_2}{\sqrt{(2\sigma_y^2)}} = t, \quad (2)$$

where $y_1 = \log x_1$, $y_2 = \log x_2$ and the number of degrees of freedom are those on which the estimate of σ_y is based. If the number of degrees of freedom is large, then tables of the normal distribution may be used.

It is one of the advantages of this method that a value of P for any difference can be measured.

An example of the tests of significance

In a cross-immunizing experiment, two sera, A and B, obtained in two successive bleedings from the same animal, were titrated for influenza-virus antibody by the haemagglutination inhibition titration. The titres (x_1 and x_2) were 639 and 755 respectively. The tests of significance, to determine whether

the difference between the two titres represents a true antibody rise, or merely represents random error, are applied as follows.

The 'M_t' test. *V_c*, previously determined for haemagglutination inhibition titrations, = 6.6 % = 6.6/100 = 0.066; and $\bar{x} = (639 + 755)/2$. Substituting these values in equation (1), we have

$$M_t = \frac{755 - 639}{0.066 \times \frac{1}{2}(639 + 755)} = 2.5.$$

The 5 % probability value for M_t is 2.8. A deviation between two antibody titres as great as that observed, would be obtained as a result of random error more than once in twenty trials. It is unlikely, therefore, that the difference is significant.

The logarithmic transformation. Applying equation (2), we have, $\sigma_y = 0.061$, determined experimentally from a series of titrations and, being the standard error of log x , calculated in the usual way. Then

$$t = \frac{2.8779 - 2.8055}{0.061} = 1.187.$$

Since σ_y was determined with 88 degrees of freedom, the value of P corresponding to $t = 1.187$ is $0.3 > P > 0.2$. The difference in titres observed will occur as a result of random error with a mean frequency of between two and three times in every ten trials. It is, therefore, unlikely to be a significant deviation. It is worth noting that the logarithmic transformation gives a more precise estimate of P than the M_t test.

DISCUSSION

The method described is of general applicability and can be used to measure the significance of differences between single observations in any test of which the parameter σ_y^2 can be reliably estimated and y is normally distributed. It is not, however, applicable to techniques where the titre is restricted to one of a series of tube dilutions or to an arbitrarily interpolated value, since in such cases the titre does not form a continuous variable. It is essentially a research method. For routine work the significance of small differences of titres are not likely to be important. Significance in the diagnostic sense is not concerned with the question of whether two titres differ statistically, but whether the difference justifies a positive diagnosis. It has been shown, for instance (Wolff, 1949), that a fourfold increase in antibody level is required to justify such diagnosis in the case of influenza. Therefore, for work of this nature, a simpler method such as the Salk test is preferable, possessing as it does the great advantage that it does not require any special apparatus (Salk, 1944).

Expenses of the work were defrayed partially from a grant by the Medical Research Council. One of us, J. C. N. Westwood, was seconded for virus research from the staff of the Public Health Laboratory Service.

REFERENCES

- HIRST, G. K. & PICKELS, E. G. (1942). A method for the titration of influenza haemagglutinin and influenza antibodies with the aid of a photoelectric densitometer. *J. Immunol.* **45**, 273.
- NEEDHAM, N. V. (1947). A photometric method for the comparative evaluation of disinfectants. *J. Hyg., Camb.*, **45**, 1.
- SALK, J. (1944). A simplified procedure for titrating haemagglutinating capacity of influenza-virus and the corresponding antibody. *J. Immunol.* **49**, 87.
- SMITH, W. & WESTWOOD, M. A. (1949). Factors involved in influenza haemagglutination reactions. *Brit. J. exp. Path.* **30**, 48.
- WOLFF, H. I. (1949). The use of the haemagglutination inhibition test in epidemiological influenza-virus studies. *J. Hyg., Camb.*, **47**, 396.

(Received 13 January 1951)