THE ANTIBODY RESPONSE IN MAN FOLLOWING INFECTION WITH VIRUSES OF THE POX GROUP

I. AN EVALUATION OF THE POCK COUNTING METHOD FOR MEASURING NEUTRALIZING ANTIBODY

By K. McCARTHY, and A. W. DOWNIE

Department of Bacteriology, University of Liverpool

and P. ARMITAGE

Statistical Research Unit of the Medical Research Council, London School of Hygiene and Tropical Medicine

(With 2 Figures in the Text)

Various techniques have been used in studying antibody responses following infection with pox viruses. Of the *in vitro* tests the complement fixation and haemagglutination inhibition techniques have proved most convenient, but the antibodies concerned are apparently different from neutralizing antibody (Chu, 1948). Estimation of this antibody to certain members of the group, such as ectromelia and vaccinia viruses, may be made in susceptible animals. Variola virus does not, however, produce lesions in the common laboratory animals with sufficient regularity to make such neutralization tests practicable. But it does constantly give rise to recognizable focal lesions on the chorio-allantois of developing chick embryos. This tissue, susceptible to most members of the pox group, has therefore been used in previous studies (McCarthy & Downie, 1948; Downie & McCarthy, 1950), employing the technique developed by Burnet and his colleagues for vaccinia and other viruses (Burnet, 1936; Keogh, 1936; Burnet, Keogh & Lush, 1937; Burnet & Lush, 1939).

Our main interest in the present studies has been the time of appearance and the persistence of antibody after vaccination and in smallpox patients. Neutralizing antibody in sera has been estimated by the inhibition of variola virus on the chorio-allantois. This virus produces small lesions on the chorio-allantois, so that accurate counts are possible over a wide range and consequently it is particularly suitable for antibody titrations by this technique. Also the fact that variola is more readily neutralized by immune serum than vaccinia (Downie & McCarthy, 1950) makes its use a more sensitive method for detecting antibody. Some of our findings have already been published (Downie, 1951; Downie & McCarthy, 1954), and the technique has been used for antibody titrations in convalescent sera from alastrim patients (McCarthy & Downie, 1953).

In this paper there are recorded the results of experiments made in 1950, and subsequently, to determine the effect of minor variations in technique on pock counts and the variation to be expected when serum-virus mixtures are tested in replicate using a standard technique. The details are now presented with a

https://doi.org/10.1017/S002217240003758X Published online by Cambridge University Press

statistical analysis by one of us (P.A.). This analysis has provided a simple method of determining the smallest significant reduction in pock count for different control means and has served as a guide in assessing the results in the two subsequent papers.

MATERIALS AND METHODS

Eggs were obtained from one poultry farm from a flock of hens which were Rhode Island Reds crossed with Black Leghorn or with Buff Rocks. The eggs were incubated at 38° C. for 12 days before use. They were carefully candled and a square window, side approximately 0·8 cm., was cut in the shell over a well vascularized part of the chorio-allantoic membrane. (It was found on opening the eggs 3 days later that this objective was not always achieved.) While the shell was being cut care was taken not to damage the underlying shell membrane. A small opening into the air space at the blunt end of the egg was made and paraffin vaseline mixture, melting point 45° C., was then placed along the cuts in the shell and the square of shell was turned back off the shell membrane. A small slit was made in the shell membrane to allow the chorio-allantois to drop away from it, aided if necessary by gentle suction with a rubber teat over the opening at the air space end of the egg. No fluid was placed over the slit in the shell membrane before the dropping of the chorio-allantois, except in experiments made to test the value of this procedure, which showed, as Burnet & Faris (1942) had found, that the use of saline broth as dropping fluid did tend to avoid the occasional scarring of membranes which slightly lowers the mean count. A piece of the shell membrane in the window was torn out to allow the inoculum to be delivered on the dropped chorio-allantois. The square of shell was then replaced over the window and sealed by the paraffin vaseline mixture with the help of a warm metal spatula run along the edges of the window. The opening over the original air space was not sealed. After the eggs had been rocked to distribute the inoculum they were incubated for 3 days at 36° C.

The virus used in most of the tests was the fifth egg-passage of a strain isolated in 1945 from the skin lesions of a case of variola major. The suspension was prepared from chorio-allantoic membranes inoculated 3 days previously with fourth egg-passage virus. These membranes were extracted with McIlvaine's citric acid-phosphate buffer, 0·004M, pH 7·2. Virus in the supernatant from low-speed centrifugation was sedimented and thrice washed by resuspension and high-speed centrifugation in fresh buffer; the virus was finally resuspended in buffer to which nutrient broth and penicillin were added to give concentrations of 20% and 100 units/ml., respectively. Sterile saline containing the same concentration of broth and penicillin was used as diluent for serum and virus before setting up mixtures, unless otherwise stated. The virus suspensions were stored at 4° C.

All sera were inactivated at 58° C. for 15 min. before use to destroy the activity of heat labile inhibitory factors known to be present in fresh human sera (McCarthy & Germer, 1952). Virus-serum mixtures were set up by adding 0·2 ml. of serum or serum dilution to 0·4 ml. of virus suspension. This suspension was obtained by diluting the stock preparation so that it would produce approximately 100 lesions

https://doi.org/10.1017/S002217240003758X Published online by Cambridge University Press

per membrane. The mixtures (0·6 ml.) were kept at room temperature for appropriate periods of time and then distributed among five or six 12-day-old embryos, using a fine pipette delivering about 60–70 drops per ml. With this technique occasional eggs received one drop more or one drop less than the others in the group. After incubation the eggs were opened, the chorio-allantoic membranes were excised, placed in black dishes containing formol saline, and the number of pocks on each membrane counted using a concealed tally counter. The few membranes showing severe scarring or tears due to faulty inoculation technique were excluded from the counts (see footnotes, Tables 3 and 4). Eggs with dead embryos were likewise discarded.

RESULTS

Effect on pock count of minor variation in technique

It is commonly appreciated that membranes with non-specific lesions may have lower pock counts than membranes without such lesions. Burnet & Faris (1942) employed dropping fluid to minimize scarring. Experiments were made to test this point after most of the results recorded in this paper had been obtained. Tuberculin syringes were used to deliver 0·1 ml. of inoculum on each of forty membranes, twenty dropped with and twenty without dropping fluid; the mean pock counts in this experiment were seventy-eight and sixty-three, respectively. The first group of membranes included one membrane with a non-specific lesion, while the second group included seven such membranes. In the second experiment the mean count with dropping fluid was 26·6 and without dropping fluid 27·4. In the first group of nineteen membranes there were three with non-specific lesions, while in the second group of twenty there were four. These results suggest that there may be some slight advantage in the use of dropping fluid, because by its use non-specific lesions may be reduced. Hahon, Louie & Ratner (1957) have introduced a modified inoculation technique to decrease non-specific lesions; but the proportion of their membranes which showed such lesions (70 %), when eggs were prepared by the usual method, was much higher than we have encountered. The Vibro-Tool which they used for shell cutting may have been responsible for these unfavourable results.

It has been suggested that membranes vary in susceptibility to vaccinia according to the time that has elapsed since they were dropped, but with variola virus no great variation in susceptibility has been noted over a period of 4 hr. The means of counts of groups of ten membranes inoculated immediately, 30 min., 1 hr. and 3 hr. after dropping were 43, 54, 49 and 51. Keeping eggs at room temperature for a period up to 1 hr. after dropping the membranes and before inoculation, seemed to have no deleterious effect as shown by the mean count in such eggs and in control eggs kept at 36° C. throughout. In the first group the mean pock count of ten eggs was twenty-eight and in those kept and inoculated in the incubator room it was twenty-five.

It has been possible to count the lesions of variola virus on the chorio-allantois up to 300 or even more, but such high counts may not give an accurate measure of the amount of infective virus because of suppression of lesions through over-

https://doi.org/10.1017/S002217240003758X Published online by Cambridge University Press

crowding. However, when doubling dilutions of virus suspension have been tested the mean pock counts have borne a linear relation to dilutions up to counts in the region of 150–200. Table 1 gives the results of such a titration when the commencing dilution was 10^{-5} of a purified suspension.

Table 1. *Titration of twofold dilutions of variola virus*

Dilution factor	Individual membrane counts	Mean
1	116, 151, 171, 194, 196, 198, 208, 259 (D, D)	187
2	71, 74, 79, 93, 94, 115, 121, 123, 135, 142	105
4	27, 33, 34, 44, 49, 51, 52, 59, 67, 92	51
8	8, 10, 15, 22, 26, 27, 30, 41, 44, 48	28
16	5, 6, 9, 11, 13, 15, 16, 19, 20, 22	14
32	6, 6, 7, 7, 8, 9, 9, 9, 11, 20	9

D = dead embryo.

Effect on neutralization of varying some conditions of the test

Strain of variola virus. Alastrim and variola virus have been equally neutralized by sera from alastrim patients (McCarthy & Downie, 1953) and in most of the present work with human sera the above-mentioned 1945 strain of variola was used. Previous tests with immune fowl-sera showed no immunological differences among strains coming from different parts of the world in different years (McCarthy & Downie, 1948; Downie & McCarthy, 1950).

Effect of egg passage. Virus suspensions were prepared directly from human smallpox crusts and also from the first, second, third and fifth passage on the chorio-allantois. No appreciable difference in the degree of neutralization of these different suspensions was observed with either post-vaccination or smallpox convalescent sera.

Effect of storage of test virus suspension. Most of the neutralization tests considered in this paper were made over a period of several months with the same stock virus preparation. During this time the pock-producing titre decreased from 10^8 to 10^7/ml. As this involved an increase in the amount of dead virus and of soluble antigen in the preparation used for the later tests in the series, the effect of excess of these materials on neutralization of virus by immune serum was examined. Virus for this test was suspended in the following: (*a*) a suspension of heat-killed variola virus from smallpox crusts, (*b*) a similar suspension from chorio-allantoic membranes, (*c*) virus-free soluble antigen from extracts of infected chorio-allantoic membranes, and (*d*) broth saline. The neutralizing potency of the serum was slightly reduced in the presence of heated virus, but not significantly by the soluble antigen (Table 2). As in our stock-virus suspension the amount of dead virus accumulating in later tests was no more than 1·0 % of that added in the experiment shown in Table 2, it was considered that dead virus or soluble antigen were not important interfering agents in the testing of serum by our technique.

Time of contact of virus and serum. It has long been known that increased time of contact of serum and virus prior to inoculation may increase the degree of

https://doi.org/10.1017/S002217240003758X Published online by Cambridge University Press

neutralization whether these tests are carried out in susceptible animals or in chick embryos. While variation in the time of preliminary contact of variola virus and serum had little effect when a high titre immune fowl serum was tested after $\frac{1}{2}$, $1\frac{1}{2}$ and 4 hr. contact, the effect was definite when weaker human post-vaccination serum was used. The results of one such test is shown in Table 3.

Table 2. *Percentage of virus neutralized by immune fowl serum in the presence of dead virus or soluble antigen*

Dilutions of Sera		Virus diluent: Heated crust virus	Heated egg virus	Soluble antigen	Broth saline
Immune fowl	1/10	99	99	98	98
	1/50	98	92	97	100
	1/250	95	85	93	98
	1/1250	84	86	90	96
	1/6250*	46	0	75	93
Normal fowl	1/10	0	0	0	0

* In previous neutralization tests this had been found to be the neutralizing titre of this serum.

Table 3. *Effect on neutralization of prolonging time of contact of serum with virus prior to inoculation*

Serum	Dilution of serum	Time of contact	Pock counts	Mean	% virus neutralized
Post vaccination	1/5	30 min.	30, 47, 64, 64, 67, 78	58	55
	1/25		(2),* 82, 93, 130, 134, 185	125	2
	1/125		38, 108, 118, 132, 170, 190	126	2
Normal	1/5		145, 188, 146, 110, 116, 62	128	0
Post vaccination	1/5	4 hr.	4, 6, 7, 19, 41, D	15	88
	1/25		5, 28, 35, 35, 42, 86	38	70
	1/125		60, 77, 87, 97, 115, 158	99	23
Normal	1/5		45, 56, 119, 140, 196, 211	128	0

* (2) badly scarred membrane—discounted. D = dead embryo.

Experiments with vaccinia virus bearing on this point are to be found in the paper by Boulter (1957). Because of the increase in neutralizing power of weak sera with increased time of contact, an endeavour was made to reduce variation of this period to a minimum.

Variation in pock counts in replicate titrations

The following experiments were made to determine the extent to which the pock count varied when replicate mixtures of serum and virus were inoculated on groups of eggs in the same experiment. A knowledge of this inherent variation was necessary before tests for the detection and measurement of antibody could be evaluated.

https://doi.org/10.1017/S002217240003758X Published online by Cambridge University Press

Expts. 1 *and* 2. For these experiments two pools of sera from unvaccinated healthy adults were used. In each experiment one pipette was used to deliver 0·2 ml. quantities of undiluted pooled serum into a series of sterile tubes and another pipette to add the 0·4 ml. of diluted virus suspension. After the mixtures had stood 30 min. the contents of each tube were inoculated on five eggs; a separate finely drawn capillary pipette was used for each mixture. By this method the variation in the volume of inoculum from egg to egg in one group was never more than 16 %. The results of the pock counts are shown on Tables 4 and 5.

Table 4

Expt. 1. Lesions developing on the chorio-allantois of eighteen groups, each of five eggs. Each group was inoculated with replicate mixtures made separately from the same suspension of variola virus and the same pool of normal serum.

Serum sample no.	Lesion counts on individual eggs	Arithmetic mean
1	65, 91, 96, 114, 121	97
2	25, 33, 61, 76, 95	58
3	43, 61, 111, 148, 151	103
4	67, 70, 74, 109, 118	88
5	(1), 41, 24, 54, 97	54
6	56, 86, 88, 119, 125	95
7	58, 75, 104, 118, 120	95
8	57, 78, 87, 96, D	84
9	65, 75, 80, 104, 144	94
10	(9), 60, 84, 101, D	82
11	77, 83, 93, 112, 140	101
12	44, 57, 75, 89, 110	75
13	48, 57, 68, 83, 135	78
14	(6), 37, 61, 77, 80	64
15	38, 50, 51, 51, 82	54
16	18, 18, 44, 64, 65	42
17	14, 23, 60, 64, 103	65
18	13, 78, 83, 84, 98	71

D = dead embryo.
( ) = counts excluded from the mean because the membranes were avascular or badly scarred.

Expt. 3. Sera from twenty unvaccinated adults were tested separately in a similar manner, a fresh pipette being used to measure out each serum and separate capillary pipettes for inoculating the mixtures into eggs. The results are shown in Table 6.

Expt. 4. Sera obtained from three adults who had been vaccinated for the first time 2–3 weeks previously were pooled, and from this pool seventeen replicate mixtures were prepared and tested (Table 7).

Expt. 5. Sixteen replicate mixtures were prepared using a pool of four sera from revaccinated adults (Expt. 5*a*, Table 8); at the same time four replicate mixtures were made with the same virus suspension and the serum of an unvaccinated adult (Expt. 5*b*, Table 8).

https://doi.org/10.1017/S002217240003758X Published online by Cambridge University Press

Table 5

Expt. 2. This experiment was made as Expt. 1 (Table 4) using a different pool of normal sera and a suspension of variola virus made from the same stock.

Serum sample no.	Lesion counts on individual membranes	Arithmetic mean
1	10, 21, 29, 32, 51	29
2	27, 37, 53, 65, D	45
3	19, 33, 47, 48, D	37
4	17, 22, 22, 30, 82	37
5	22, 38, 46, 54, 66	45
6	26, 29, 39, 70, 83	49
7	26, 28, 41, 49, 61	41
8	7, 21, 69, 84, D	45
9	38, 49, 52, 60, 67	53
10	12, 29, 35, 44, D	30
11	48, 58, 77, 87, 98	74
12	17, 42, 49, 67, 75	50
13	40, 57, 67, 84, 112	72
14	30, 39, 43, 51, D	41
15	21, 36, 43, 46, 47	39
16	14, 34, 39, 48, 80	43
17	26, 32, 43, 54, 64	44

Table 6

Expt. 3. The procedure in this experiment was similar to that used in the experiment shown in Table 4, except that the serum samples were obtained from twenty unvaccinated adults.

No. of patient	Lesion counts on individual eggs (five eggs per mixture)	Arithmetic mean
1	68, 77, 79, 100, D	81
2	6, 90, 240, D, D	112
3	89, 133, 152, D, D	125
4	63, 91, 103, 113, D	93
5	36, 96, 103, 110, 142	97
6	64, 120, 129, 148, D	115
7	85, 106, 109, 135, 198	127
8	34, 48, 51, 96, 108	67
9	35, 98, 104, 145, D	96
10	109, 110, 114, 136, D	117
11	62, 89, 155, 172, 199	135
12	37, 118, 125, 141, 145	113
13	110, 128, 187, 210, D	159
14	103, 108, 122, 140, 159	126
15	30, 43, 97, 105, 126	80
16	49, 87, 103, 112, 118	94
17	62, 72, 93, 104, 150	96
18	55, 92, 96, 111, 184	108
19	57, 68, 110, 130, D	86
20	65, 78, 87, D, D	77

https://doi.org/10.1017/S002217240003758X Published online by Cambridge University Press

Table 7

Expt. 4. Lesions resulting from the inoculation of eighteen mixtures of variola virus and post-vaccination serum. The virus suspension used was added to eighteen samples of the same immune serum.

Serum sample no.	Lesion counts on individual membranes	Arithmetic mean
1	14, 19, 23, 23, 33	22
2	9, 34, 41, D, D	28
3	6, 16, 21, 26, 29	19
4	2, 11, 20, 27, 50	22
5	8, 11, 11, 22, 70	24
6	7, 16, 32, 56, 68	36
7	2, 10, 17, 24, 54	21
8	9, 8, 17, 33, 42	22
9	3, 5, 17, 19, D	11
10	9, 14, 18, 19, D	15
11	6, 11, 14, 28, 49	22
12	8, 9, 16, 43, 52	26
13	11, 15, 17, 25, 37	21
14	16, 17, 21, 30, 38	24
15	4, 17, 28, 32, 45	25
16	3, 28, 37, D, D	23
17	1, 14, 24, 31, D	17

Table 8

Expt. 5. The technique was similar to the experiment shown in Table 7 except that in addition to sixteen mixtures of virus and immune serum (5*a*) four groups of eggs were inoculated with mixtures of virus and normal serum (5*b*).

Serum sample no.	Lesion counts on individual membranes	Arithmetic mean
5*a* 1	5, 13, 15, 21, 28	16
2	17, 26, 26, 34, 41	29
3	22, 34, 35, 57, 98	49
4	18, 20, 20, 36, 55	30
5	19, 34, 41, 48, D	36
6	7, 28, 57, 63, 140	59
7	18, 19, 21, 44, D	26
8	8, 18, 22, 54, 60	32
9	11, 21, 58, 100, D	48
10	52, 55, 74, 81, 108	74
11	26, 29, 61, 78, D	49
12	11, 12, 42, 51, 76	38
13	16, 22, 37, 38, 70	37
14	17, 20, 35, 42, 80	39
15	10, 16, 18, 22, 62	26
16	6, 16, 33, 43, 75	35
5*b* 1	65+, 75, 94+, 107, 140	107
2	66+, 98, 120, 130, 130	120
3	153, 175, 231, 242, D	200
4	109, 137, 156, 186, 219	161

The counts marked + had confluent areas and consequently these counts should have been higher by an unknown amount.

https://doi.org/10.1017/S002217240003758X Published online by Cambridge University Press

Statistical analysis

The object of this analysis is, first, to determine whether consistent results are obtained with different normal sera (Expt. 3), and with different samples of the same mixture of serum and virus (the other experiments); and, secondly, to examine the variability of the replicate counts in each group of three, four or five eggs, in order to discover the smallest difference between the mean counts for test and control sera which could be taken as evidence for the presence of neutralizing antibody in the test serum.

Consistency of results

Tables 4–8 contain ninety-two groups of replicate counts, each group having three, four or five counts. These are divided amongst six experiments (Expts. 5*a* and 5*b* being kept separate).

Table 9. *Analysis of variance of pock counts, between and within groups of replicates*

	Expt. 1				Expt. 2				Expt. 3			
	S.S.	D.F.	M.S.	V.R.	S.S.	D.F.	M.S.	V.R.	S.S.	D.F.	M.S.	V.R.
Between groups	30018·4	17	1765·8	2·11*	11694·3	16	730·9	1·79	41693·6	19	2194·4	1·25
Within groups	56108·0	67	837·4		25719·6	63	408·2		117680·8	67	1756·4	
	86126·4	84			37413·9	79			159374·4	86		

	Expt. 4				Expt. 5*a*				Expt. 5*b*			
	S.S.	D.F.	M.S.	V.R.	S.S.	D.F.	M.S.	V.R.	S.S.	D.F.	M.S.	V.R.
Between groups	2015·9	16	126·0	<1	14809·3	15	987·3	1·50	19928·4	3	6642·8	5·09*
Within groups	16339·5	61	267·9		39360·4	60	656·0		15647·6	12	1304·0	
	18355·4	77			54169·7	75			35576·0	15		

* Variance ratio significant ($0{\cdot}01 < P < 0{\cdot}05$).

The column headings, S.S., D.F., M.S., V.R., stand for sum of squares, degrees of freedom, mean square and variance ratio, respectively.

For each of the six experiments, an analysis of variance was carried out, to show whether there was any evidence of real variation in the mean count between sera (in Expt. 3), or between samples from the same pool of serum (in the other experiments). The results are shown in Table 9. Expt. 3 revealed no significant differences between sera. Differences between samples from the same pool were significant at the 5 % level in Expts. 1 and 5*b*, very nearly reached the 5 % level of significance in Expt. 2, and were not significant in Expts. 4 and 5*a*. These findings mean that, at least for Expts. 1 and 5*b*, the mean counts for different samples vary more than would be at all likely if this variation were due purely to the variation between replicates. It should be noted, however, that the first two samples in Expt. 5*b* gave some membranes with confluent areas and the recorded counts are consequently too low.

Variability of counts

In an egg neutralization test, the mean of the counts on about five eggs inoculated with the test serum plus virus, will be compared with the mean count on about five eggs inoculated with a normal serum plus virus. The simplest way to

https://doi.org/10.1017/S002217240003758X Published online by Cambridge University Press

assess the significance of the difference in means will be to perform a t-test. Since the effect of antibody is to reduce the count, the t-test will be one-sided; that is, the difference will be judged significant only when the test serum gives a lower mean count than the normal serum.

However, with two series of at most five counts in each, the standard error of the difference in means, which is used in the t-test, is rather inaccurately determined, and a more sensitive test would result if a better estimate of the standard error were available. We have, therefore, examined the ninety-two groups of

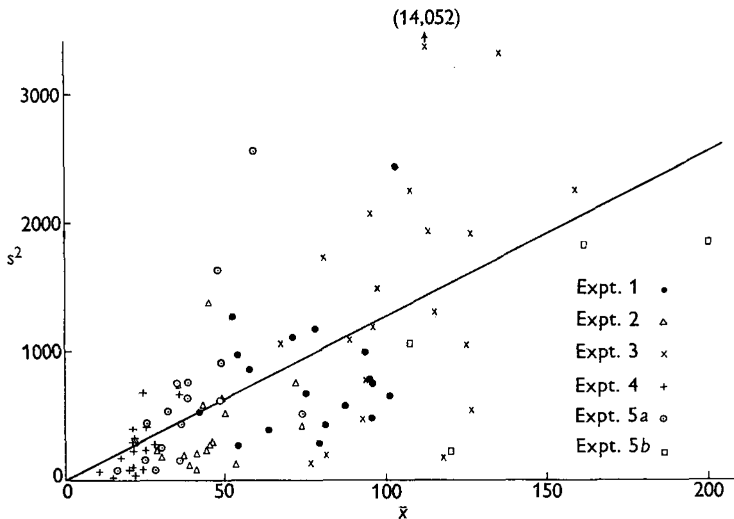

Fig. 1. Scatter diagram showing relationship between the sample mean and the sample variance estimate. The straight line is derived from the unbroken line in Fig. 2.

replicate counts given in Tables 4–8 with a view to obtaining some information about the variance between replicates which might be of use in future work. The four groups of counts obtained in Expt. 5*b* could perhaps have been omitted. Their variability will, however, be shown to be consistent with that of the remaining data, and since they provide some evidence of the variability to be expected among high counts they have been retained.

The most obvious feature of these data is that the variance increases appreciably with the mean. This may be seen from Fig. 1, in which, for each group of three to five counts, the estimate of variance (s^2 = sum of squares about the mean divided by one less than the group size) is plotted against the mean, $\bar{x}$. It therefore seems reasonable to try to predict the variance in terms of the mean count. If, from the present data, an estimate could be obtained of σ, the mean value of s^2 for a given value of $\bar{x}$, this estimate could be used in the calculation of the standard error of the difference between two means, $\bar{x}$ being calculated from both series pooled.

Since the scatter of the values of s^2 increases appreciably as $\bar{x}$ increases (cf. Fig. 1), the relationship between σ^2 and $\bar{x}$ is most easily estimated by making logarithmic transformations of both s^2 and $\bar{x}$. Now, $\log s^2$ is a biased estimate of $\log \sigma^2$. In order to apply regression methods to the logarithmic transforms, we must therefore

https://doi.org/10.1017/S002217240003758X Published online by Cambridge University Press

correct this bias so that we are working with unbiased estimates of log σ^2. This is achieved by adding a small quantity c_r to each value of log s^2, giving a corrected value

$$y = \log s^2 + c_r.$$

The correction term c_r depends only on the number, r, of counts in the group ($r = 3$, 4 or 5), and is obtained from Table 1 of Bartlett & Kendall (1946). (c_r is obtained by multiplying the entries in the second column of Bartlett & Kendall's Table 1 by $\log_{10} e = 0{\cdot}4343$, since we are working with logarithms to base 10, and taking $n = r - 1$; thus $c_3 = 0{\cdot}25$, $c_4 = 0{\cdot}16$, $c_5 = 0{\cdot}12$. These values are strictly correct only if the counts are normally distributed, a condition not fulfilled in the present data.)

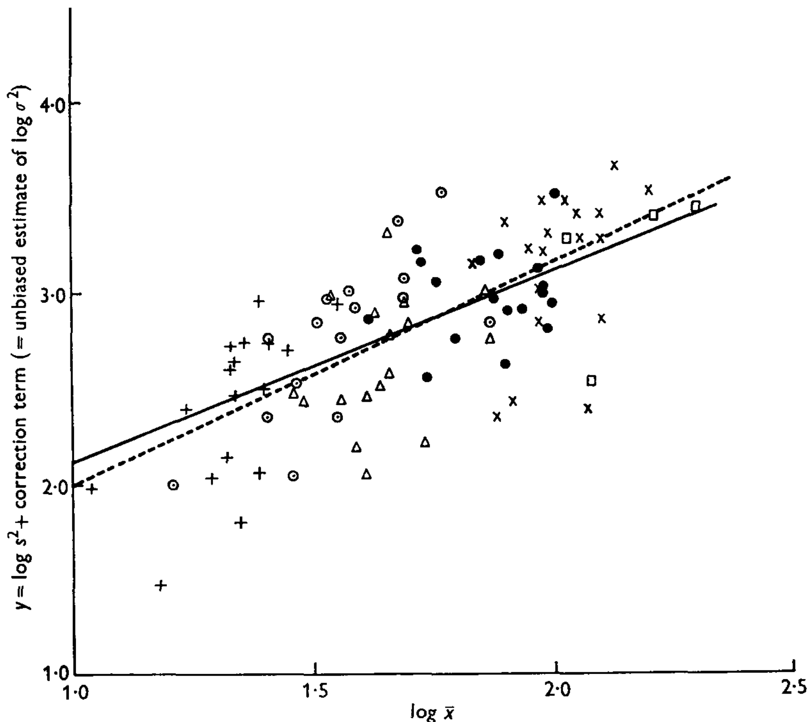

Fig. 2. Scatter diagram showing relationship between the logarithm of the sample mean and the estimate of the logarithm of the variance. Key to experiments as in Fig. 1. The broken line is the regression line of y on log $\bar{x}$, and the unbroken line the best-fitting line with a slope of unity.

The scatter diagram of y against log $\bar{x}$ is shown in Fig. 2, each point representing one group of replicate counts. The points for the six experiments clearly follow much the same trend. The dotted line shown in Fig. 2 is the regression line of y on log $\bar{x}$, in the calculation of which each value of y has been weighted in inverse proportion to its variance. (The weights are obtained by dividing the entries in the fourth column of Bartlett & Kendall's Table 1 by $(\log_{10} e)^2 = 0{\cdot}1886$; thus, for group size 3, 4 and 5, the weights are, respectively, 3·22, 5·67 and 8·22.) The analysis of variance in Table 10 shows that lines fitted separately for the six

https://doi.org/10.1017/S002217240003758X Published online by Cambridge University Press

experiments do not differ significantly either in slope or position, and we therefore regard the overall regression line as applying to all the experiments.

In the Appendix we give reasons for believing that the scatter of the values of y about this line is no more than might be expected from sampling theory. The point is of some interest, for if this scatter had been significantly greater than its expected value, it would not have been possible to postulate a single value of σ^2 for a given $\bar{x}$; we should have had to recognize that the relationship between σ^2 and $\bar{x}$ varied from one occasion to another.

Table 10. *Analysis of variance for regression of y on log* $\bar{x}$

	S.S.	D.F.	M.S.	V.R.
Deviations about individual regressions lines	59·7476	80	0·7468	1·00
Between slopes	4·6875	5	0·9375	1·26
Deviations about parallel regression lines	64·4351	85		
Between positions	0·2906	5	0·0581	<1
Deviations from single regression line	64·7257	90	0·7192	
Due to single regression	69·9763	1	69·9763	93·7*
	134·7020	91		

* Variance ratio significant ($P<0·001$).

Assuming, then, a linear functional relationship between the expected value of y (i.e. $\log \sigma^2$) and $\log \bar{x}$, we estimate this relationship from the overall regression line, the equation for which is

$$Y = 0·8520 + 1·1483 \log \bar{x}.$$

The regression coefficient, 1·1483, has a standard error of 0·1210 and so does not differ significantly from 1. It is convenient to assume a slope of 1; the best fitting straight line, shown as an unbroken line in Fig. 2, then has an equation

$$Y = 1·1059 + \log \bar{x}.$$

Using Y as an estimate of $\log \sigma^2$, we may write this in the form

$$\log \hat{\sigma}^2 = 1·1059 + \log \bar{x}, \quad (1)$$

whence

$$\hat{\sigma}^2 = 12·76\, \bar{x}. \quad (2)$$

The standard error of the coefficient 1·1059, in (1), is 0·0330. 95 % confidence limits for the coefficient 12·76, in (2), are therefore given by antilog $\{1·1059 \pm (1·987)(0·0330)\}$, i.e. 10·97 and 14·84.

The straight line given by (2) is shown in Fig. 1. We may note if the counts followed a Poisson distribution the mean value of s^2, for a given mean count of $\bar{x}$, would be equal to $\bar{x}$. Equation (2) shows that the counts forming the present data have a variance about thirteen times as great as that expected for a Poisson distribution. Similar analyses of counts obtained by a number of different workers (Armitage, 1957) have shown that in general the variance increases at a rather greater rate than proportionately to the mean.

https://doi.org/10.1017/S002217240003758X Published online by Cambridge University Press

Use of equation (2) *in a neutralization test*

Suppose that n_t and n_c eggs are inoculated with the test serum and virus, and with the control serum and virus, respectively. Let $\bar{x}_t$ and $\bar{x}_c$ be the mean counts for the test and control preparations, respectively, and let $\bar{x}$ be the overall mean of the n_t and n_c counts. We then calculate, by (2),

$$\hat{\sigma}^2 = 12{\cdot}76\bar{x},$$

and the standard error of the difference in means is

$$\hat{\sigma}\sqrt{\left(\frac{1}{n_t}+\frac{1}{n_c}\right)}.$$

Since we are only interested in cases where the test preparation gives a lower mean than the control, the difference $\bar{x}_c - \bar{x}_t$ is significant at the 5% level if it exceeds 1·64 times its standard error (assuming the difference to be normally distributed).

As an example, suppose the two sets of counts were as follows:

		Total	Mean
Test:	14, 19, 23, 23, 33	112	$\bar{x}_t = 22{\cdot}40$
Control:	25, 33, 61, 76	195	$\bar{x}_c = 48{\cdot}75$
		307	$\bar{x} = 34{\cdot}11$

$\bar{x}_c - \bar{x}_t = 26{\cdot}35$

$\hat{\sigma}^2 = (12{\cdot}76)\,(34{\cdot}11) = 435{\cdot}2$; $\hat{\sigma} = 20{\cdot}86$.

S.E. $(\bar{x}_c - \bar{x}_t) = (20{\cdot}86)\sqrt{(\frac{1}{5}+\frac{1}{4})} = 13{\cdot}99$.

$(\bar{x}_c - \bar{x}_t)/$S.E. $= 1{\cdot}88$, which is significant at the 5% level.

Smallest significant reduction in count

For any value, $\bar{x}_c$, of the control mean, we can determine what value of the test mean, $\bar{x}_t$, just differs significantly from $\bar{x}_c$ at the 5% level, when the test described above is carried out. Table 11 shows the values of $\bar{x}_t$ corresponding to five different values of $\bar{x}_c$. The third column of Table 11 shows the difference, $\bar{x}_c - \bar{x}_t$, as a percentage of the control mean, $\bar{x}_c$. The percentage reduction required for a given level of significance clearly falls as the control mean rises.

Table 11. *Reduction in mean pock count just reaching* 5% *level of significance* (*five eggs in each series*)

Control mean $\bar{x}_2$	Test mean $\bar{x}_1$	Percentage reduction
50	26·9	46
75	46·1	39
100	66·1	34
150	107·8	28
200	150·8	25

These percentage reductions are relevant for a comparison of two specified groups of counts. If many groups are compared, as in the experiments reported above, it would be misleading to test for significance by examining the difference between two selected groups. In Expt. 5*a* (Table 8), for example, the mean counts

https://doi.org/10.1017/S002217240003758X Published online by Cambridge University Press

for samples 10 and 15 differ by more than the minimum reduction tabulated in Table 11, but the analysis of variance revealed no significant difference between the sixteen means.

DISCUSSION

These experiments were made to provide some measure of the variability of pock counts using variola virus. As this virus was to be used for assessing antibody in human sera all of the replicate counts recorded were made in the presence of human serum. The variability of pock counts with vaccinia virus was noted by Keogh (1936) and more recently analysed by Overman & Tamm (1956) and Westwood, Phipps & Boulter (1957).

The chorio-allantoic membrane is not equally susceptible to different pox viruses (Fenner & McIntyre, 1956; Dumbell, Downie & Valentine, 1957); consequently the variation in pock counts may differ with different pox viruses. None the less, it seems likely that even with different viruses, the same non-specific factors may affect the uniformity of counts in groups of eggs receiving the same virus inoculum.

The considerable variation in individual pock counts within groups could not be accounted for by the slight variation in volume of inoculum in occasional eggs. Non-specific traumatic effects could not be attributed to unsuitable dropping fluids (Westwood *et al.* 1957) as none was used. Except in those instances where the membrane was badly scarred or when the embryo was dead at the time of examination we have recorded counts on all membranes and have not weighted the results in favour of those membranes giving counts within the 'normal' range as did Burnet & Faris (1942). Nor did we discard before counting, as did Westwood *et al.* (1957) in their work with vaccinia, all membranes showing non-specific lesions, encroachment of albumin sac, signs of haemorrhage or 'any other abnormality'. Exclusions of this kind are liable to subjective error and might be difficult to apply impartially. Not only did the presence of haemorrhage and non-specific lesions appear to depress the count of variola pocks, but those membranes in which the inoculated area lay close over the yolk sac usually showed lower counts than the mean of the group. However, these counts are included in our tables.

Overman & Tamm (1956) noted that membranes dropped 24 hr. before inoculation of vaccinia virus showed more non-specific lesions and fewer pocks than membranes inoculated after 1–3 hr. Westwood *et al.* (1957) found that even after 6 hr. the susceptibility of the chorio-allantois had already decreased. They suggested that peak sensitivity occurred between 1 and 4 hr. but that the peak time was unpredictable. In our own experiments with variola virus, as noted above, there was no great variability in susceptibility of chorio-allantois to variola virus at periods up to 4 hr. after dropping membranes.

The foregoing considerations force one to the conclusion that the mean count observed is lower than should be obtained under ideal conditions. Occasional eggs showed high counts, two to three times the mean, and therefore well outside Poissonian expectation. Such high pock counts are not in our view due to secondary spread nor to dispersion of clumped virus. The most likely explanation would

https://doi.org/10.1017/S002217240003758X Published online by Cambridge University Press

appear to be that, of the living variola virus particles inoculated on the chorio-allantois, only a proportion give rise to pocks and that occasional membranes of high susceptibility give these unusually high counts. Fenner & McIntyre (1956) who found wide scatter of pock counts due to myxoma virus, showed that only a proportion of rabbit-infective particles produced lesions in the egg. A similar difference in titre of cowpox virus for the rabbit skin and chorio-allantois has also been observed (Dumbell, Downie & Valentine, 1957).

In the course of our experiments we have aimed at a virus inoculum that would give a mean count of 50–100 pocks per membrane. Sometimes, however, the inoculum has been larger and it has proved possible because of their small size and discrete nature to count up to 200. Overman & Tamm (1956) found that the greatest precision in titrating vaccinia was obtained when the individual count fell between ten and twenty. Their work was done with membranes examined 3 days after inoculation. Westwood *et al.* (1957), who examined membranes 2 days after inoculation, when vaccinial lesions were smaller, found a linear relationship between virus dilution and mean pock counts up to seventy. The lesions produced by variola after 3 days are only about half the diameter of those produced by our strain of vaccinia after 2 days, and consequently the effect of overcrowding does not become serious until the mean count approaches 200. The data in Table 11 would seem to show that the higher counts possible when variola virus is used increased the significance of the percentage differences observed between mean counts with control and test sera in neutralization tests.

SUMMARY

1. An investigation into factors affecting the reliability and usefulness of the pock counting method as a means of measuring antibody against variola virus is given.

2. The use of dropping fluid, variation in the time between dropping the membranes and inoculation, and the presence of non-specific lesions on the membranes did not greatly affect the mean pock counts in our experiments.

3. Prolonging the time of contact *in vitro* of immune serum and virus had little effect when sera of high antibody content were used. With weaker sera the pock count was progressively diminished with increase in time of contact of serum and virus before inoculation.

4. Using a standard technique replicate experiments with serum-virus mixtures were made and the results analysed statistically. This analysis showed that a predictable relationship existed between mean pock count and variance; so that for a given control mean the minimum reduction in pock count due to immune serum which could be taken as significant at the 5 % level could be easily calculated.

We are indebted to Miss Irene Allen for help with the computations and to Mrs M. G. Young for preparing the diagrams.

https://doi.org/10.1017/S002217240003758X Published online by Cambridge University Press

REFERENCES

ARMITAGE, P. (1957). *J. Hyg., Camb.*, **55**, 564.
BARTLETT, M. S. & KENDALL, D. G. (1946). *J. R. statist. Soc.* Suppl. **8**, 128.
BOULTER, E. A. (1957). *J. Hyg., Camb.*, **55**, 502.
BURNET, F. M. (1936). *Spec. Rep. Ser. med. Res. Coun., Lond.*, no. 220. H.M. Stationery Office.
BURNET, F. M. & FARIS, D. D. (1942). *J. Bact.* **44**, 241.
BURNET, F. M., KEOGH, E. V. & LUSH, D. (1937). *Aust. J. exp. Biol. Med. Sci.* **15**, 296.
BURNET, F. M. & LUSH, D. (1939). *J. Path. Bact.* **48**, 275.
CHU, C. M. (1948). *J. Hyg., Camb.*, **46**, 49.
DOWNIE, A. W. (1951). *Lancet*, i, 419.
DOWNIE, A. W. & MCCARTHY, K. (1950). *Brit. J. exp. Path.* **31**, 789.
DOWNIE, A. W. & MCCARTHY, K. (1954). *Dynamics of Virus and Rickettsial Infections*, p. 194. New York: The Blakiston Co. Inc.
DUMBELL, K. R., DOWNIE, A. W. & VALENTINE, R. C. (1957). *Virology*, **4**, 467.
FENNER, F. & MCINTYRE, G. A. (1956). *J. Hyg., Camb.*, **54**, 246.
HAHON, N., LOUIE, R. & RATNER, M. (1957). *Proc. Soc. exp. Biol., N.Y.*, **94**, 697.
KEOGH, E. V. (1936). *J. Path. Bact.* **43**, 441.
MCCARTHY, K. & DOWNIE, A. W. (1948). *Brit. J. exp. Path.* **29**, 501.
MCCARTHY, K. & DOWNIE, A. W. (1953). *Lancet*, i, 257.
MCCARTHY, K. & GERMER, W. D. (1952). *Brit. J. exp. Path.* **33**, 529.
OVERMAN, J. R. & TAMM, I. (1956). *J. Immunol.* **76**, 228.
WESTWOOD, J. C. N., PHIPPS, P. H. & BOULTER, E. A. (1957). *J. Hyg., Camb.*, **55**, 123.

APPENDIX

Sampling theory for y

Table 10 shows a residual mean square for deviations of y about the overall regression line on log $\bar{x}$, equal to 0·7192 on 90 D.F. The system of weighting in this regression analysis is such that, if the original counts had been normally distributed, the expectation of this residual mean square would have been unity. However, the distributions of the counts are appreciably skew (which means that y and log $\bar{x}$ are correlated in samples from a given population, and tend to *decrease* the residual variation about the regression line), and have greater kurtosis than a normal distribution (which tends to *increase* the residual variance of y). In the absence of large samples it is difficult to estimate the skewness and kurtosis of the distributions. However, we have seen that for Expts. 2, 3, 4 and 5a there was no significant variation in mean count between groups of replicates. If we assume that in each of these experiments the groups are random samples from one population, the moments of the distributions may be estimated by combining the groups within each experiment.

From the pooled counts for each experiment, estimates were obtained of the second, third and fourth moments of the distribution, μ_2, μ_3 and μ_4. It may be shown that, for large samples, the variance of y is increased above its normal value by a factor

$$1+\tfrac{1}{2}\gamma_2,$$

where $\gamma_2=(\mu_4/\mu_2^2)-3$, a measure of kurtosis. Hence, in an experiment with k groups, $Sw(y-\bar{y})^2/(1+\frac{1}{2}\gamma_2)$ should be distributed as χ^2 on $k-1$ D.F., provided that the large-sample factor $1+\frac{1}{2}\gamma_2$ is valid for small samples. From the pooled counts

https://doi.org/10.1017/S002217240003758X Published online by Cambridge University Press

for each experiment, we have estimated γ_2 by g_2, the corresponding function of the sample moments, and calculated $\chi_1^2 = Sw(y-\bar{y})^2/(1+\frac{1}{2}g_2)$. The results are shown in Table 12. The values of χ^2 are all fairly well within the range of sampling variation. Their total is 52·12 on 66 D.F. ($P=0{\cdot}89$). This value is rather small, but not unreasonably so, in view of the approximations of the method. At any rate, there is no evidence of any variation in y above that expected by sampling theory.

A further check is provided by the correlation coefficient, r, between the values of y and $\log \bar{x}$ in these four experiments. In large samples, the expected correlation, ρ is given by

$$\rho^2 = \frac{\mu_3^2}{\mu_2(\mu_4 - \mu_2^2)}.$$

Again, it is uncertain how accurate this formula would be for small samples, but Table 12 shows that the observed values of r^2 agree remarkably well with the estimates, $\hat{\rho}^2$, of ρ^2, obtained from the sample moments.

Table 12. *Comparison of the observed and theoretical variation in y*

	Total variation of y				Correlation of y with $\log \bar{x}$		Variation of y about regression on $\log \bar{x}$	
Expt.	$Sw(y-\bar{y})^2$	$1+\frac{1}{2}g_2$	χ_1^2	D.F.	r^2	$\hat{\rho}^2$	χ_2^2	D.F.
2	12·6244	1·0821	11·67	16	0·0989	0·0830	11·47	15
3	20·6133	1·3151	15·67	19	0·1050	0·0848	15·32	18
4	19·6459	1·4057	13·98	16	0·3776	0·3869	14·19	15
5	20·5253	1·9012	10·80	15	0·5447	0·4250	8·55	14
			52·12	66			49·53	62

Finally, combining both these comparisons, the sum of squares about regression, for each experiment, may be compared with its expected value. The appropriate criterion is $\chi_2^2 = (1-r^2)\,\chi_1^2/(1-\hat{\rho}^2)$ on $k-2$ D.F. The values of χ_2^2 shown in Table 12 add to 49·53 on 62 D.F. ($P=0{\cdot}87$). In these four experiments, at least, there is no evidence of any variation in y, about the regression on $\log \bar{x}$, above that expected by sampling theory.

(*MS. received for publication* 5. VI. 57)

https://doi.org/10.1017/S002217240003758X Published online by Cambridge University Press