J. Hyg., Camb. (1969), **67**, 335
Printed in Great Britain

Overlap and the errors of plaque counting

II. The bias of the variance and the concealment of errors

By D. W. HOWES

Viral Products Section, National Biological Standards Laboratory, Commonwealth Department of Health, Melbourne, Victoria, Australia

and S. FAZEKAS de St GROTH

Division of Animal Genetics, Commonwealth Scientific and Industrial Research Organization, Sydney, New South Wales, Australia

(*Received* 30 *September* 1968)

INTRODUCTION

Procedures for correcting overlap bias of plaque counts, described in the preceding paper (Howes, 1969), yield more reliable estimates of the numbers of plaque-forming units actually present in samples. However, to take full advantage of the method, it is also necessary to specify the errors associated with plaque counting.

Plaque numbers on replicate cultures are usually assumed to follow a Poisson distribution, but this assumption can be correct only where no overlaps occur. Deviations from this simplest model become of practical as well as theoretical importance where observed counts are noticeably biased by plaque overlap.

Because substantial overlapping is a feature of many plaque assay systems, a study of its effect upon the apparent errors of plaque counts was undertaken.

THEORETICAL

Derivation of equations

The fraction of plaques expected to become undetectable owing to overlapping is an exponential function of plaque number. High counts will therefore suffer a disproportionately greater reduction than low counts and the observed distribution will be negatively skewed. The effect on the variance is greater than that on the mean, so that the ratio of variance to mean for observed plaque numbers will be substantially less than the Poissonian expectation of 1. This is shown in Fig. 1 for two theoretical distributions.

The approximate relationship between the true and observed plaque numbers is

$$C = (1/K)\,(1-e^{-KN}), \tag{1}$$

where C is the observed and N the true count, while K is the assay constant (as defined in the companion paper).

The expectation of C (= estimated mean C) is

$$E\{C\} = (1/K)[1-E(e^{-KN})], \tag{2a}$$

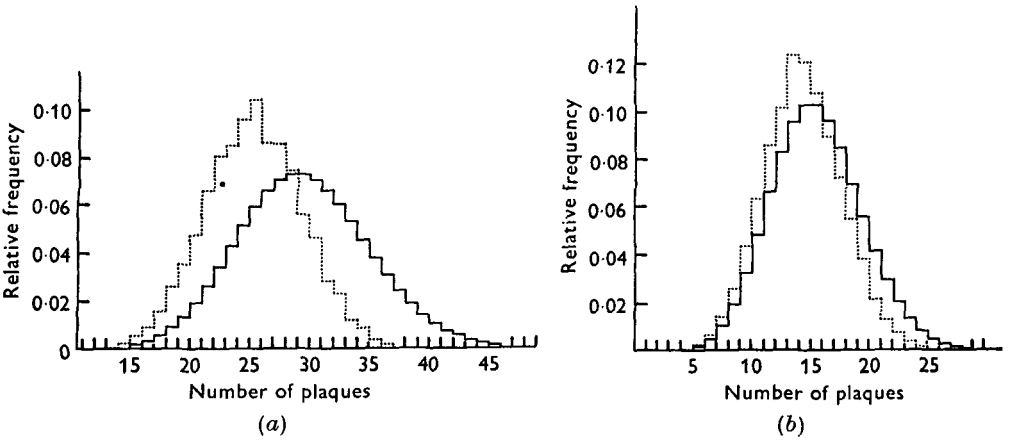

Fig. 1. Displacement and compression of hypothetical ideal plaque-count distributions by plaque overlapping. Assay constant, $K = 0{\cdot}01$. True count distributions shown by solid lines, with mean of 30 (*a*) and 15 (*b*) suffer displacement and compression by overlapping, becoming observed count distributions shown by interrupted lines.

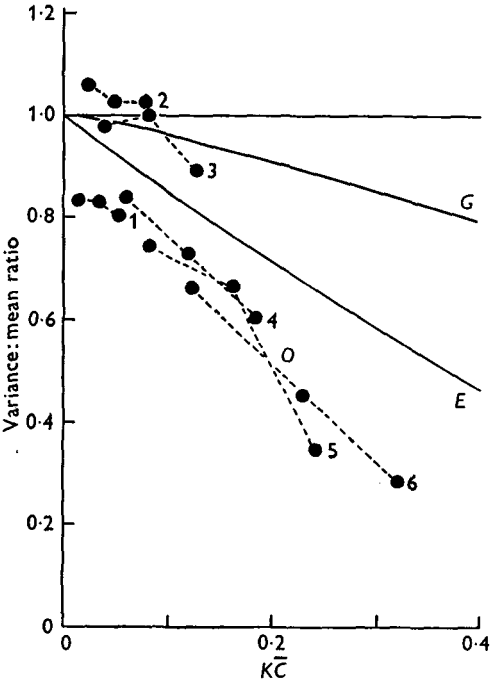

Fig. 2. Predicted and observed changes in the variance:mean ratios of plaque-count distributions produced by plaque overlapping. The predicted overlap biases of mean observed counts, for values of $K\bar{C}$ up to 0·4, are shown by the ratio of observed to true mean counts, line *G*. The variance to mean ratios predicted by equations (2b) and (3b) are shown by line *E*. Experimental values for observed variance to mean ratios are linked by broken lines, *O*. Each group of 30 replicate cultures inoculated with the Saukett strain of virus (Expt. 1) provided one value for each of the three counting times. The position of each group in the range of serial 1·6-fold dilutions is indicated by number, number 1 representing the most dilute inoculum.

https://doi.org/10.1017/S0022172400041735 Published online by Cambridge University Press

which becomes

$$E\{C\} = (1/K)\{1-\exp[-N(1-e^{-K})]\}. \tag{2b}$$

The variance of observed counts may be calculated as

$$V\{C\} = E\{C^2\} - [E\{C\}]^2,$$

and accordingly

$$V\{C\} = (1/K^2)\{\exp[-N(1-e^{-2K})] - \exp[-2N(1-e^{-K})]\}. \tag{3a}$$

Expanding the exponential terms, (3a) gives

$$V\{C\} = Ne^{-2KN} \tag{3b}$$

Variance $\{C\}/N$ is a single valued function of KN. Hence the variances of any two count distributions will be in the same proportion to their means if the value of $K\bar{C}$ is the same for both, even though the mean number, size, and morphology of plaques are different.

The progressive fall in the observed variance to mean ratio which is expected to occur as $K\bar{C}$ increases is shown in Fig. 2. This may be compared with the fall in the ratio C/N, which shows the expected progressive increase in the overlap biases as KC increases. Observed variances may be greater than those predicted by equation (3*b*) as, for a given true count, the numbers of plaques obscured by overlapping are also randomly distributed. However, this effect will be small relative to the random sampling error of true counts.

The variance of N may be derived from equation (3*b*) as

$$V(N) = e^{2KN}V(C) \tag{4}$$

Equation (4) might be used to prepare a table of correction factors corresponding to values of $K\bar{N}$ or $K\bar{C}$, and thus estimate the variance of true counts.

EXPERIMENTAL RESULTS

An experimental evaluation of the above relationship was carried out using the plaque counting data provided by the experiments described in the preceding paper.

Estimation of variances

The data from all experiments agreed with the expectation that, where some plaques become undetectable due to overlapping, the variance of a distribution of observed counts will be reduced by a larger factor than the mean count.

However, in only one of the four experiments (Expt. 1, Saukett) did the variance to mean ratio for low values of KC closely approach the ideal value of 1. In this experiment, observed decreases in the ratio paralleled those predicted by equation (3*b*) for an ideal system (Fig. 2). In the other experiment with this virus strain (Expt. 4) most of the initial variances of observed plaque count distributions were substantially lower than expected, while for the two experiments with the type 1 LSc-2ab strain, they were considerably higher than expected. In assessing *observed* variance to mean ratios for conformity with equation (3*b*), they were therefore expressed as a fraction of the estimated ratio for *true* counts, which was

https://doi.org/10.1017/S0022172400041735 Published online by Cambridge University Press

calculated using the least biased data obtained at the first counting of each group of replicate cultures. This value was obtained by correcting the mean count by the methods described in the preceding paper, and by correcting the observed variance by means of equation (4). These normalized variance to mean ratios are shown in Fig. 3.

Although considerable scatter is evident, the adjusted ratios agree reasonably well with those predicted by equation (3*b*) for values of $K\bar{C}$ up to about 0·2. Above this value a trend towards lower values than those predicted was evident. This was to be expected since the counts were outside the range set for the validity of the correcting formula.

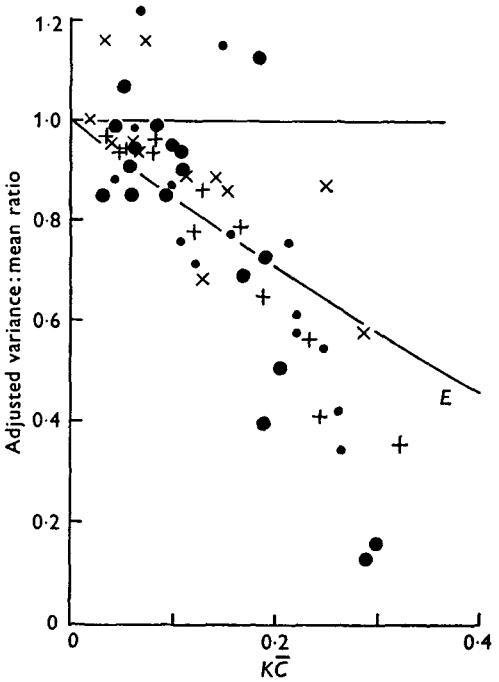

Fig. 3. 'Adjusted' variance: mean ratios for all experiments. Ratios, adjusted as described in the text, are plotted against $K\bar{C}$, and are compared with expected values shown by line *E*. ●, Expt. 2, LSc-2ab; •, Expt. 3, LSc-2ab; +, Expt. 1, Saukett; ×, Expt. 4, Saukett.

Diminution of observed variances with time

The progressive concealment of the true errors of plaque counting, which is the result of the drop in variance to mean ratios, is best illustrated by the unadjusted ratios shown in Table 1 for the LSc-2ab strain of virus (Expt. 2).

Repeated counting of the same plaques showed that, in all but one instance, an increase in the overlap bias caused by an increase in plaque size led to a fall in the variance to mean ratio. The data show that at or near the upper limit to the acceptable counting range ($KC = 0{\cdot}18$) overlapping may reduce the ratio by more than a half, and where the limit is greatly exceeded this reduction may be by as much as sevenfold.

https://doi.org/10.1017/S0022172400041735 Published online by Cambridge University Press

Heterogeneity of cultures

It is of practical importance to find out whether the apparent heterogeneity of cultures in their sensitivity to virus, as indicated by the high variance to mean ratios at the first counting time, was due, partly or wholly, to counting plaques too early. This can be determined by examining the effect of adding the relatively small counts of late-appearing or 'new' plaques to the counts of early-appearing or 'original' plaques.

Table 1. *The overlap biases of variances of observed plaque-count distributions*

Observed variance : mean ratios for the LSc-2ab strain of virus (Expt. 2) demonstrate the fall in the observed variance : mean ratio which accompanies an increase in plaque size. Corrected ratios show that correction procedures will give more realistic if inexact estimates of the ratios for the true counts. Corrections become inadequate at about $K\bar{C} = 0{\cdot}18$. Mean plaque size for the three counting times were 2·8, 6·7, and 9·3 mm. respectively.

Rel. virus conc.	Counting time	$K\bar{C}$	Variance : Mean Ratios: Observed $= \frac{V(C)}{\bar{C}}$	Corrected $= \frac{V(N)}{\bar{N}}$
1·11	1	0·04	1·86	1·99
	2	0·20	1·01	1·42
	3	0·30	0·32	0·55
1·00	1	0·04	2·34	2·48
	2	0·19	0·98	1·35
	3	0·29	0·32	0·53
0·56	1	0·02	2·47	2·54
	2	0·10	2·38	2·81
	3	0·19	1·86	2·58
0·50	1	0·02	1·63	1·66
	2	0·09	1·48	1·71
	3	0·17	1·20	1·58
0·28	1	0·01	1·43	1·45
	2	0·06	1·32	1·44
	3	0·11	1·31	1·56
0·25	1	0·01	1·50	1·52
	2	0·05	1·62	1·75
	3	0·10	1·44	1·68

For a system in which there was neither heterogeneity nor overlap, the addition of new to original plaques would lead to identical increases in both the mean and the variance, and their ratio would remain constant at the value of 1. For a system free of heterogeneity but subject to overlap bias, new plaques would be more readily obscured because of their smaller size, and the ratio would fall slightly on adding the two sets of counts.

In practice there will be a tendency for this expected fall in the variance to mean ratios to be masked by the increased counting uncertainty which accompanies an increase in overlap bias, and by any increase in the heterogeneity of cultures with respect to late-appearing plaques. In the present study the Saukett strain

https://doi.org/10.1017/S0022172400041735 Published online by Cambridge University Press

gave an example of such behaviour, and the addition of new to original plaques led to slight increases in the values of variance to mean ratios (Table 2).

The greater uncertainty associated with the counting of LSc-2ab plaques would be expected to lead to somewhat greater rises in variance to mean ratios on adding new to original plaques. However, pronounced falls were observed, which reflect a strong negative correlation between the counts of original and new plaques on

Table 2. *Changes in variance: mean ratios produced by adding new plaques to those previously present*

(Data used in the calculation of the variance: mean ratios shown in the body of the table were progressive total counts for each culture for two or three counting times. The mean diameter of the original plaques at each time is shown at the bottom of each column. New plaques present at the third counting time in Expts. 3 and 4 were not counted.)

		Counting time				
Virus strain	Relative virus concentration	1st	2nd	3rd	1st	2nd
		Experiment 2 (15)*			Experiment 3 (15)	
LSc-2ab	1·39	—	—	—	2·56	1·25
	1·25	—	—	—	1·06	0·55
	1·11	1·86	1·49	1·54	1·41	0·90
	1·00	2·34	1·76	1·66	1·94	1·39
	0·56	2·47	1·79	1·74	1·33	1·05
	0·50	1·63	1·36	1·31	1·40	0·81
	0·28	1·43	1·00	1·04	2·11	1·51
	0·25	1·50	1·26	1·20	1·03	0·54
	0·14	1·15	1·10	1·10	0·88	0·92
	0·12	0·59	0·31	0·41	0·67	0·56
	Mean plaque Diam.	2·8 mm.	6·7 mm.	9·3 mm.	3·5 mm.	7·7 mm.
		Experiment 1 (30)			Experiment 4 (13-15)	
Saukett	0·50	0·66	0·71	0·70	—	—
	0·31	0·74	0·72	0·71	—	—
	0·28	—	—	—	0·56	0·69
	0·25	—	—	—	1·09	1·05
	0·20	0·84	1·00	0·99	—	—
	0·14	—	—	—	0·61	0·88
	0·12	0·98	0·84	0·84	0·71	0·86
	0·08	1·06	1·15	1·22	—	—
	0·06	—	—	—	0·50	0·68
	Mean plaque Diam.	4·7mm.	7·0 mm.	9·0 mm.	2·7 mm.	5·2 mm.

* Numbers in parentheses indicate the number of replicate cultures per group.

individual cultures (Table 2). This can only occur if a substantial part of the apparent heterogeneity in the sensitivities of cultures to virus, which was observed at the first counting time, was due to variation between cultures in the time of appearance of plaques. In such circumstances the use of early counting to avoid overlap bias will increase apparent heterogeneity and, although intended to decrease the total error, may actually increase it.

DISCUSSION

The reliability of plaque assays can be established only by repeatedly assaying the same virus suspension. Yet, in practice the determination of variance to mean ratios is often taken as a valid alternative. The latter approach is certainly simpler, but it has two disadvantages. Since it considers only some of the possible sources of experimental error, it will always give rather optimistic assessments of the reliability of a system. Even more damaging is the common assumption underlying this form of analysis, that the theoretical distribution of plaque numbers is Poissonian, and that the variance therefore equals the mean. This can hold only for counting techniques for which every object will always be counted even though overlapping occurs, and is therefore not true of plaque counting. In practice the adoption of the Poissonian model may be justified where the overlap bias of observed counts is very small. Where this is not so, allowance should be made for the disproportionate reduction of variances. Failing to do so can lead to gross underestimation of errors, and to statistically invalid and possibly misleading interpretation of data. The methods developed above permit compensation to be made for this bias and the estimation of true errors.

Additional requirements where true errors are being assessed are that cultures be identified by code numbers, and that inocula be allocated at random. Without such safeguards errors due to heterogeneity in materials or environmental conditions may be greatly underestimated. Even in carefully designed and conducted assays the heterogeneity of cultures may prove a major source of error. This was the case with the LSc-2ab strain of poliovirus analysed in this study. Heterogeneity could be attributed largely to variation in the time of appearance of plaques, and not to variation in the sensitivity of cultures to the initiating virus particles. This means that estimates of virus concentration provided by such an assay system will carry smaller errors if counting is delayed, and if the observed counts falling within the acceptable counting range are corrected for the overlap biases which are the consequence of delayed counting.

The approach to the analysis of errors used here, and the methods for correcting overlap bias should prove useful where the precision of other plaque assay systems is being evaluated. The conclusion of the present study, that variance to mean ratios for observed counts are unsatisfactory estimators of the true errors, is true for all systems. Such devices should never be used where the overlap biases of observed counts have not been studied.

SUMMARY

The overlapping of plaques compresses their distributions and reduces observed variances. For a given distribution the reduction of the variance is substantially larger than that of the mean. To derive the error of a plaque assay from the assumption that the variance equals the mean may therefore lead to serious overestimation of the precision of the assay.

Procedures for estimating the true error of plaque assays are developed, and their use is illustrated on experimental material.

REFERENCE

HOWES, D. W. (1969). Overlap and the errors of plaque counting. 1. The overlap biases of observed counts and their correction. *J. Hyg., Camb.* **67**, 317.