STATISTICAL CONTROL IN HAEMATOLOGY

By H. O. LANCASTER, M.B., B.S., B.A.
From the School of Public Health and Tropical Medicine, Sydney

(With 1 Figure in the Text)

§1. INTRODUCTORY

The possibilities of statistical control in bacteriology are now widely appreciated, but there has been no systematic attempt to apply the same statistical techniques to the similar problems of counting in haematology. Since Plum (1936) has described the evolution of the blood-counting apparatus and a very comprehensive review of the applications of statistics to bacteriology has been given by Eisenhart & Wilson (1943), chief consideration will here be given to the special problems arising in haematology because of certain departures from the ideal conditions usually presumed to hold in the derivation of the Poisson law. After a brief historical review of the chief contributions to the theory of blood counting, this paper contains a discussion of the adequacy of the Poisson law for the distribution of cells in the haemocytometer chamber and the measurement of the goodness of fit by the χ^2 test in large samples. In counting the red cells a fundamental condition necessary for the derivation of the Poisson law does not hold. It is, however, shown that allowance can be made for this 'crowding effect', and reasons are given for believing that the effect is not likely to invalidate the methods of 'statistical control', introduced. Illustrative examples from the author's own counts are treated by these methods. The results of the application of statistical control to the work of technicians are then described.

§2. HISTORICAL

Poisson (1837), the celebrated French mathematician, first used the distribution that goes by his name to describe the frequency of occurrence of rare events. Abbé (1878) derived the same law for the distribution of the cells over the squares of a haemocytometer chamber. He did not give any experimental verification of the theory as he was more interested in the accuracy of his pipettes and haemocytometer from the volumetric point of view. Lyon & Thoma (1881) supplied an experimental verification, but statistical techniques at the time were too undeveloped to make the verification completely convincing, as there were no criteria to judge whether divergences of the observed results from the theoretical could be explained as due merely to random fluctuations. It seems that all these experimental results were unknown to Bortkiewicz (1898), whose much-quoted example of the numbers of fatalities from the kick of a horse in the different army corps of the Prussian army is rather artificial.

https://doi.org/10.1017/S0022172400015175 Published online by Cambridge University Press

'Student' (1907) wished to estimate the error of his counts of yeast cells and also to find the most efficient dilution for obtaining pure subcultures from a single cell. Numerous authors have since pointed out that he was unaware of the previous work in the field and did not use the term 'Poisson distribution' when describing his results. The importance of his work was recognized by Goodall (1908) in a very favourable review notice, but for many years the possibilities of a statistical approach to the problem of blood counting were overlooked. Although Lucy Whitaker (1914) warmly attacked Student's paper, to which attack Student (1919) later replied, the fundamental importance of the Poisson distribution became generally recognized, especially in the field of counting in bacteriology, largely as a result of Student's paper. In fact, R. A. Fisher now rates it as the most important discrete distribution in biological work. Fisher, Thornton & MacKenzie (1922; Fisher, 1948) supplied useful methods by which the consistency of a small number of parallel counts could be tested, which are, essentially, new applications of the χ^2 test for goodness of fit. At the same time they introduced the concept of 'statistical control' of laboratory work. These methods of control appear never to have been applied to the counting of blood cells. Besides being a convenient check on current laboratory work, Fisher showed that the methods could be used to test the consistency of published data and gave examples of agreement in the count of parallel bacterial plates that were so close that the results could only have occurred with excessive rarity under conditions of strictly random sampling and were, therefore, suspect. Finally, a very full and authoritative review of the statistical problems of bacteriology with a complete bibliography may be found in the paper of Eisenhart & Wilson (1943). Recently Berkson and his co-workers have shown by photographic methods that many individuals systematically count too low and have attempted to have the over-strict criteria of a 'good count' modified so that technicians may avoid the temptation to count erroneously in order to bring the dispersion within pre-conceived bounds (Berkson, Magath & Hurn, 1935, 1939; Magath, Berkson & Hurn, 1936).

§3. LARGE SAMPLE THEORY: THE POISSON DISTRIBUTION

If it can be assumed that: (i) complete mixing has been attained, (ii) there is no cohesion between cells, and (iii) the fact of one cell falling on to a square of a haemocytometer chamber does not affect the chances of another doing so, then we may expect that the probability of any given square receiving i cells will be given by

$$p_i = e^{-\mu}\mu^i/i\,!, \tag{1}$$

where μ is the mean number of cells which fall on a square. A proof of this formula will be found in Yule & Kendall (1950). The large sample theory follows directly from this. If we know the true value of μ, then we may calculate the expected frequencies of the numbers of squares of the haemocytometer containing no cell, one cell, two cells and so on. A test of goodness of fit follows immediately, since the expected values can all be specified. But we usually do not know the true value μ, so we estimate it from the data by the observed mean and perform the χ^2 test for goodness of fit with the degrees of freedom two less than the number of classes.

Example. The number of cells on the squares of a Neubauer haemocytometer chamber were counted, with the results set out in Table 1.

Table 1. *A test of the goodness of fit of the Poisson distribution to an observed red cell count by means of* χ^2

No. of cells on the square	Observed frequency	Expected frequency	Observed − expected	Contribution to χ^2
0	11	10·740	0·260	0·006
1	36	38·851	−2·851	0·209
2	76	70·273	5·727	0·467
3	80	84·737	−4·737	0·265
4	74	76·634	−2·634	0·091
5	58	55·445	2·555	0·118
6	38	33·429	4·571	0·625
7	17	17·275	−0·275	0·004
8	6	7·812	−1·812	0·420
9	3	3·140	−0·804	0·135
10	0	1·136 (4·804)		
11	1	0·528		
Total	400	400·000	0·000	2·340

Mean number of cells per square = 3·6175.
Variance = 3·5308.
χ^2 for 8 degrees of freedom = 2·340, $P = 0·97$.

The test of goodness of fit shows that the Poisson law describes the distribution in this case adequately. In fact, the χ^2 of 2·34 for 8 degrees of freedom would be exceeded by chance in some 97 % of cases. (This count was a selected example and so this excellent fit need not cause surprise.)

An outstanding feature of the Poisson is that the variance has an expectation equal to the mean. The significance of the difference of the observed variance from its expected value can be tested by a well-known use of χ^2. We take

$$\chi^2 = \Sigma(x-\bar{x})^2/\bar{x}$$
$$= n\Sigma x^2/\Sigma x - \Sigma x, \tag{2}$$

or
$$\chi^2 = (n-1)\text{ (estimated variance)/(mean)}. \tag{3}$$

For counts over 400 squares,

$$\chi^2 = 399\text{ (estimated variance)/(mean)}. \tag{4}$$

In the example shown in Table 1, the mean and variance are 3·6175 and 3·5308. χ^2 is 389·4 with 399 degrees of freedom, quite close to the expected value of 399. The deviation of the variance in this case from its expectation is readily explained by 'sampling error'.

§4. EXPERIMENTAL OBSERVATIONS OF THE CROWDING EFFECT AND THE DIMINISHED VARIANCE DUE TO CROWDING

It is to be noted that, in the above example, the mean is relatively low, for the normal red cell count results in a mean of 7 or 8 cells per small haemocytometer square. Berkson *et al.* (1935), working at an average cell density of 5–7 cells per

https://doi.org/10.1017/S0022172400015175 Published online by Cambridge University Press

square, were able to demonstrate a constant crowding effect which resulted in a deficiency of squares with very high or very low cell counts, due to an exchange of red cells between the squares, the cells being crowded off the squares with a high count on to the neighbouring squares. The assumption made by these authors, that the variance is likely to be a constant proportion of the theoretical value, cannot be accepted without qualification, though their estimate that the variance was 85 % of the mean was a reasonable figure at the cell density at which they were working. On the theoretical grounds discussed in a later paragraph, it seems evident that the variance would be more closely represented by a quadratic function of the mean than by a linear function, the coefficient of the term in the square of the mean being small and negative, so that the variance approximates closely to the mean when the mean is low, but falls away from it as the mean rises. From a consideration of 32 counts of the whole 400 small squares of the haemocytometer chamber, two regression lines relating the variance to the observed mean have been determined, first using a formula of the type

$$v = a_1 m + a_2 m^2, \tag{5}$$

where m is the observed mean; and secondly using a formula

$$v = m - bm^2. \tag{6}$$

The appropriate equations, fitted by the method of least squares, were

$$v = 0{\cdot}999474m - 0{\cdot}020457m^2, \tag{7}$$

and

$$v = m - 0{\cdot}020534m^2. \tag{8}$$

Equations (7) and (8) give practically identical values for the expected value of the variance. Equation (8) is to be preferred because of its greater simplicity. The value of b, empirically obtained, agrees roughly with that of the theoretical discussion given later in this paper.

Using (8), the expected variance has been computed for each of the counts in Table 2 and for each of the counts of Berkson *et al.* (1935), shown in Table 3. In both cases the counts have been rearranged in order of increasing mean. The regression formulae, (7) and (8), fit the series of the author's counts well. This can be seen by the generally small value of the differences between the observed and the expected variances in the fifth column of Table 2. For the series of Berkson *et al.* the agreement is fair, but the fit would have been closer with a larger value of b. In Table 3 the expected variance was computed with a value of b obtained by the method of least squares from Berkson's data,

$$v = m - 0{\cdot}026324m^2. \tag{9}$$

The residual sums of squares about the regression line (or the sum of the deviations from the regression line taken with a positive sign) are not greatly reduced by the use of this new regression formula.

Table 4 shows the expected variance, as given by (8), for selected values of the mean.

https://doi.org/10.1017/S0022172400015175 Published online by Cambridge University Press

Table 2. *A comparison of the observed variance with the calculated, using the regression formula of* (8) *for the author's own counts*

No. of count	Mean	Observed variance	Expected variance	Difference: observed − expected variance
1	0·31	0·31	0·31	0
2	0·32	0·36	0·32	+0·04
3	0·35	0·36	0·35	+0·01
4	0·37	0·36	0·37	−0·01
5	0·71	0·74	0·70	+0·04
6	0·72	0·65	0·71	−0·06
7	0·80	0·79	0·79	0
8	0·84	0·86	0·83	+0·03
9	0·86	0·82	0·84	−0·02
10	0·88	0·87	0·86	+0·01
11	0·93	0·86	0·91	−0·05
12	1·21	1·16	1·18	−0·02
13	1·42	1·43	1·38	+0·05
14	1·63	1·69	1·57	+0·12
15	1·69	1·69	1·63	+0·06
16	1·78	1·48	1·71	−0·23
17	1·93	2·09	1·85	+0·24
18	2·35	1·92	2·24	−0·32
19	2·45	2·35	2·33	+0·02
20	2·87	2·85	2·70	+0·15
21	2·92	2·47	2·74	−0·27
22	3·62	3·53	3·35	+0·18
23	5·36	4·35	4·77	−0·42
24	5·47	5·05	4·86	+0·19
25	5·50	5·60	4·88	+0·72
26	6·61	5·10	5·71	−0·61
27	6·69	5·47	5·77	−0·30
28	6·70	5·83	5·78	+0·05
29	7·19	6·63	6·13	+0·50
30	7·31	6·01	6·22	−0·21
31	8·04	6·70	6·72	−0·02
32	8·05	6·84	6·72	+0·12

Table 3. *The counts of Berkson, Magath & Hurn* (1935), *rearranged in order of increasing mean*

Four hundred small squares counted. The expected variances have been computed from the regression equations (8) and (9).

Observed mean	Observed variance	Expected variance by (8)	Observed − expected variance (8)	Expected variance by (9)	Observed − expected variance (9)
5·09	4·37	4·56	−0·19	4·41	−0·04
5·31	4·75	4·73	+0·02	4·57	+0·18
5·49	4·62	4·87	−0·25	4·70	−0·08
5·75	4·37	5·07	−0·70	4·88	−0·51
6·03	5·95	5·28	+0·67	5·07	+0·88
6·04	5·38	5·29	+0·09	5·08	+0·30
6·31	4·80	5·49	−0·69	5·26	−0·46
6·49	5·38	5·62	−0·24	5·38	0·00
6·75	5·71	5·81	−0·10	5·55	+0·16
6·77	5·20	5·83	−0·63	5·66	−0·46
Σd			−2·02		0·03
$\Sigma \lvert d \rvert$			3·58		3·07
Σd^2			1·99		1·61

https://doi.org/10.1017/S0022172400015175 Published online by Cambridge University Press

Table 4. *Some tabulated values of the expected variance computed by equation* (8)

Mean	Expected variance	Mean	Expected variance
0·5	0·49	4·5	4·08
1·0	0·98	5·0	4·48
1·5	1·45	5·5	4·88
2·0	1·92	6·0	5·26
2·5	2·37	6·5	5·63
3·0	2·81	7·0	5·99
3·5	3·25	7·5	6·35
4·0	3·67	8·0	6·69

§5. SYSTEMATIC CHANGES IN THE OBSERVED FREQUENCIES AS SHOWN BY THE χ^2 TEST FOR GOODNESS OF FIT

If χ^2 be computed as in the example of Table 1, the fit is usually found to be good for low values of the mean, say four red cells per small haemocytometer square, but a systematic departure from theory occurs with higher values of the mean; there are then fewer squares than expected containing numbers of cells differing widely from the mean. There is naturally a complementary excess of squares containing approximately the same number of cells as the mean; thus there is a smaller variance than that given by the Poisson law. In such a case the differences between observed and expected frequencies would show systematic effects. There would be an excess of minus signs at both ends of column (4) in a table such as Table 1 and an undue number of plus signs in the middle of the range. Further, the contributions to χ^2—and hence the total χ^2 of column (5)—would be too large in general. Berkson (1938) has drawn attention to the relative insensitivity of χ^2, used in the classical test of goodness of fit (as in Table 1), to this 'crowding effect'.

§6. A THEORETICAL CONSIDERATION OF 'CROWDING' IN RED CELL COUNTING

In bacteriology the bacteria are so small that the presence of one in a certain volume of fluid does not affect the probability of another being present in the same volume to any appreciable degree. There may be differences in the opportunity of developing a colony or there may be a difficulty in counting the colonies which develop, since overlap of colonies may occur in the plate poured if the inoculum be too heavy and spreading inefficient. But in other applications of the Poisson law to practical problems, the effect of crowding may be more apparent. Garwood (1948) has discussed the problem of overlap in bombing, and Armitage (1949) a similar problem of overlap in counting dust particles, in which case only one of two particles falling together can be counted. Our problem is different from that of either of these authors since the overlapping observation (i.e. the blood cell) is not lost but tends to move from a crowded square to one less crowded, and so contributes doubly to the reduction of the variance below the theoretical value specified by the Poisson law.

§7. MATHEMATICAL DISCUSSION OF 'CROWDING'

Suppose that there are a number of smooth circular disks (the red cells) of radius r which readily slip over one another but which do not slip on the surface of the haemocytometer. It will be assumed that when one disk has fallen to the surface of the haemocytometer and another falls directly on to it the second moves away from the first along the line of centres until the two are just in contact, both flat on the surface; further, that the surface is divided up by parallel lines a distance h apart and that h is greater than $2r$.

The problem is now: given that all positions of the centre of the lower disk within a square are equally likely, what is the probability of a second disk slipping over the 'northern' boundary of the square if it comes in contact with the lower cell? Once we have determined the probability of the upper cell in contact with a lower passing over the 'northern' boundary of the square, we can find the approximate probability of any cell passing out of a haemocytometer square if it comes in contact with a lower cell as it settles down on to the haemocytometer. After obtaining this probability, an attempt will be made to relate it to the diminution of the observed variance. The area between the lines bounding the square to north and south is called a strip.

It may be assumed without loss of generality that the first disk has its centre always along a fixed line perpendicular to the boundary. As a convention, a cell is considered to be within the strip if it is intersected by the 'northern' boundary, and outside the strip unless it is entirely within the 'southern'.

Fig. 1 shows the centre, C, of the lower cell internal to the northern boundary (case 1), and external to the northern boundary (case 2). The 'critical' area is shaded. If the centre of the upper cell falls into this critical shaded area, it will be initially 'within' the strip but will pass out of it over the northern boundary. Suppose now that C is at a distance s north of the northern boundary. The area shaded is then

$$(r-s)\sqrt{\{4r^2-(r-s)^2\}}.$$

Since all values of s are equally likely for all positions of C from $(-h+r)$ to r, the expectation of the shaded area is given by

$$\begin{aligned} E\text{ (critical area)} &= \int_{-r}^{r}(r-s)\sqrt{\{4r^2-(r-s)^2\}}\,ds/h \\ &= \int_{0}^{2r}x\sqrt{(4r^2-x^2)}\,dx/h \\ &= 8r^3/3h. \end{aligned} \tag{10}$$

Each square of a haemocytometer has four edges to which a similar reasoning, with modifications necessary because of the convention noted above, can be applied. For the square, the expectation of the total critical area is approximately four times that which we have calculated for the northern boundary of a strip of the same width. If then λ is the probability of an upper cell coming in contact with a lower and passing over a boundary of the square, it is given by the ratio of the expectation of the critical area to the area of the square

$$= 4(8r^3/3h)/h^2 = 32r^3/(3h^3). \tag{11}$$

https://doi.org/10.1017/S0022172400015175 Published online by Cambridge University Press

It is important to note that if r is small relatively to h, then λ is inversely proportional to h^3, since r is a constant when consideration is limited to the discussion of red cells.

In our case, with the red cells counted on the small squares of the haemocytometer,

$$h = \tfrac{1}{20} \text{ mm.} = 50\,\mu, \quad r = 4\mu. \tag{12}$$

So

$$\lambda = 32r^3/3h^3 = 0{\cdot}00546. \tag{13}$$

We have now to relate λ to the fall in the variance due to the crowding effect. The effect of the crowding on the variance may be calculated as follows: let λ be the probability of the upper cell passing over the boundary of the small squares ($\frac{1}{20}$ mm. side) after coming in contact with another specified cell. The distribution of the cells in the fluid over the squares before the cells settle will be given by the Poisson formula

$$P_i = e^{-m}m^i/i\,!. \tag{14}$$

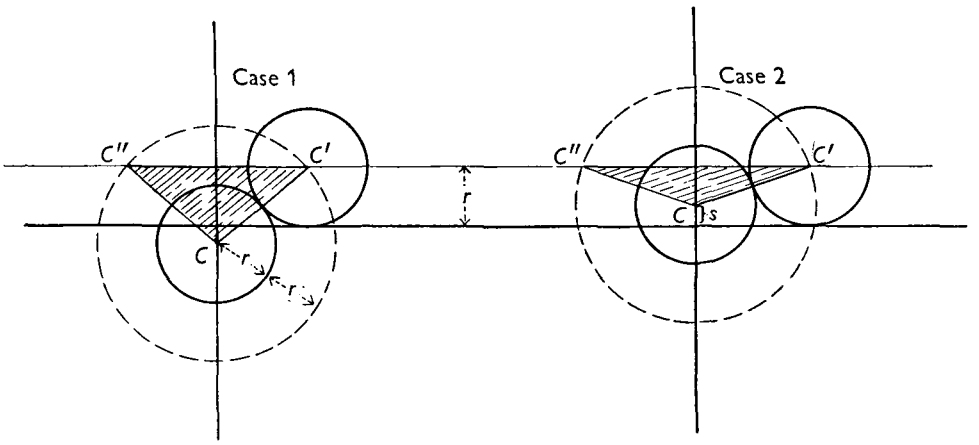

Fig. 1. *The critical area.* Under the assumptions made in the text, if the centre C' of any cell falls within the large circle, it will be in contact with the 'lower' cell with centre C. The 'northern' boundary has been drawn heavily. At a distance r north of the boundary is a faintly drawn line, such that if the centre of a cell is to the north it will be considered external to the 'strip' by the blood-counting conventions. If the centre of the upper cell falls in the cross-hatched area, the upper cell will pass over the northern boundary.

After crowding, squares with j cells initially will have an expectation of the loss of $j(j-1)\lambda$ cells. The expectation of the number of cells exchanged will be

$$\sum_{j=0}^{\infty} j(j-1)\lambda P_j = m^2\lambda. \tag{15}$$

As an approximation it is assumed:

(i) That each square receives cells independently of the number it already contains, so that each square receives an expectation of $m^2\lambda$ cells.

(ii) That only one cell is lost or gained by each square and that this occurs with a frequency equal to the expectation.

Squares finally containing j cells will be composed of four groups:

(i) Those originally containing j cells which neither lost nor gained any cell.

https://doi.org/10.1017/S0022172400015175 Published online by Cambridge University Press

(ii) Those originally containing j cells which have lost one cell and regained one.

(iii) Those which originally had $(j-1)$ cells and gained one cell.

(iv) Those which originally had $(j+1)$ and lost one cell. We suppose that these four groups have frequencies,

$$F_{j1}, F_{j2}, F_{j3}, F_{j4}.$$

Then the variance of the number of the cells on the squares of the haemocytometer after rearrangement is given by

$$v = \sum_{j=0}^{\infty} \sum_{i=1}^{4} j^2 F_{ji} - m^2, \tag{16}$$

since the mean of the distribution is unchanged.

But

$$F_{j1} = (1-\lambda m^2)\{1 - j(j-1)\lambda\} P_j, \tag{17}$$

$$F_{j2} = j(j-1)\lambda P_j \lambda m^2, \tag{18}$$

$$F_{j3} = \lambda m^2 P_{j-1}\{1 - (j-1)(j-2)\lambda\}, \tag{19}$$

$$Fj_4 = P_{j+1}(1-\lambda m^2)(j+1)j\lambda. \tag{20}$$

Neglecting terms containing the square of λ and using (15), the contribution to the variance of the squares containing j cells will be

$$(j^2 - m^2) \sum_{i=1}^{4} F_{ji}.$$

But the sum of these quantities is readily evaluated if we neglect the terms containing λ^2 and notice that

$$\left.\begin{aligned} j^2 &= j(j-1) + j = (j-1)(j-2) + 3(j-1) + 1, \\ j^3(j-1) &= j(j-1)\{(j-2)(j-3) + 5(j-2) + 4\}, \\ (j+1)j^3 &= (j+1)j\{(j-1)(j-2) + 3(j-1) + 1\}. \end{aligned}\right\} \tag{21}$$

Hence,

$$v = m - 2\lambda m^2 \tag{22}$$

gives an approximation to the relation between the variance after rearrangement and the mean. For red cell counting on the small squares, we have already found a value of λ, namely 0·00546. The theoretical regression equation is therefore given by

$$v = m - 0{\cdot}01092 m^2. \tag{23}$$

The coefficient of m^2 is about half that experimentally determined and so the theory must be regarded as giving a not unreasonable account of the settling of the cells and the diminished variance, considering the approximations involved.

If, for a square of $k \times \frac{1}{20}$ mm. side, λ' be defined as the probability that a cell will fall on another given cell placed at random in the square and pass over a boundary, and $m'(=k^2 m)$ as the expected mean, then

$$\lambda' = 32r^3/(3k^3h^3) = \lambda/k^3, \tag{24}$$

and the variance v' is given by

$$v' = k^2 m - 2k\lambda m^2. \tag{25}$$

https://doi.org/10.1017/S0022172400015175 Published online by Cambridge University Press

So that $$v/m = 1 - 2\lambda m, \tag{26}$$

but $$v'/m' = 1 - 2\lambda m/k. \tag{27}$$

Berkson and his collaborators are therefore open to criticism in using the same ratio between variance and mean for the small haemocytometer squares and for groups of them. An alternative derivation of an equation equivalent to (27) is given in the Appendix to this paper. Equation (27) also suggests that the diminution in the variance between sets of sixteen small squares is not likely to be more than 2 %.

On the basis of Berkson's empirical result that the diminution in variance for single squares is about 15 %, the diminution for sets of sixteen small squares will be about 4 % (see Appendix, p. 417). This discrepancy between the two values was to be expected in view of the discrepancy between the coefficients of m^2 in (9) and (23). The use of χ^2 in the next section, as a measure of dispersion in sets of counts from groups of sixteen small squares will therefore not be appreciably invalidated by the crowding effect. It will be neglected in the following discussion of the test of consistency of parallel counts, on the understanding that the counts are from areas of sixteen small haemocytometer squares.

§8. THE TEST OF CONSISTENCY OF PARALLEL COUNTS

The test of the goodness of fit by the large sample methods of §3 are in any case time-consuming and break down if only the counts from a few areas are available. An alternative test due to R. A. Fisher (Fisher *et al.*, 1922; Fisher, 1948) is available. Suppose n parallel counts have been made and the counts are given individually by x_i, where i can take the values 1, 2, 3, . . . , n. The test of consistency of results is to take

$$\chi^2 = \Sigma(x_i - \bar{x})^2/\bar{x}, \tag{28}$$

with $(n-1)$ degrees of freedom. Under the conditions (i), (ii) and (iii) of §3 above, the χ^2 obtained will be exceeded by chance in a certain proportion of cases, which is the probability of χ^2 derived from the tables. We then expect to find that a χ^2 giving a probability of say 0·1 will be exceeded in 10 % of the tests, whereas a χ^2 giving a probability of 0·3 will be exceeded in 30 % of the tests, and so on. Thus the probability range may be split up in any arbitrary fashion to obtain an expected frequency for χ^2 between the limits of any division. For example, in 20 % of the cases we expect χ^2 to correspond to a probability between 0·1 and 0·3.

Example. In a set of five counts of red cells each over the area of sixteen small squares of a Neubauer haemocytometer, the following counts were obtained: 40, 49, 31, 40 and 37. χ^2 (given by (28) with 4 degrees of freedom) is equal to 4·2944. By consulting the tabulated values of χ^2, in the tables of Fisher & Yates, for example, we find that this value of χ^2 has a probability between 0·3 and 0·5 and so this set of counts would be referred to this class in the 'control chart'.

In Table 5 are summarized the results of the author's own red cell counts, treated by the methods described. A convenient set of probability classes is indicated—0·0–0·1, 0·9–1·0, and 0·1 (0·2) 0·9. In the second column are shown the corresponding values of χ^2. In the third and fourth columns are given the observed and expected frequencies. In Table 6 are given the author's findings for white cell

https://doi.org/10.1017/S0022172400015175 Published online by Cambridge University Press

Table 5. *The red cell counts of the author treated by the χ^2 test*

Probability class	Values of χ^2 defining the probability class	Observed frequency	Expected frequency
1·0–0·9	0–1·064	6	6·4
0·9–0·7	1·064–2·195	15	12·8
0·7–0·5	2·195–3·357	14	12·8
0·5–0·3	3·357–4·878	13	12·8
0·3–0·1	4·878–7·779	14	12·8
0·1–0·0	7·779–	2	6·4
Total		64	64·0

Sets of five parallel red blood cell counts have been collected into classes according to the probability resulting from the χ^2 test used as a measure of dispersion.

The resultant χ^2 has 4 degrees of freedom. 1·064 is the 0·9 or 90 % point for χ^2 with 4 D.F.

Table 6. *The white cell counts of the author, treated by the χ^2 test*

Probability class	Values of χ^2	Observed frequency	Expected frequency
1·0–0·9	0–0·584	2	2·4
0·9–0·7	0·584–1·424	5	4·8
0·7–0·5	1·424–2·366	8	4·8
0·5–0·3	2·366–3·665	3	4·8
0·3–0·1	3·665–6·251	5	4·8
0·1–0·0	6·251–	1	2·4
Total		24	24·0

Sets of four parallel white blood counts have been collected into classes according to the probability resulting from the χ^2 test used as a measure of dispersion.

counting. The observed frequencies of the probability classes from both the red and the white cell counts are in accord with expectations. There is, therefore, no evidence to assert that the experimental techniques of mixing and spreading on the haemocytometer chamber or counting were unsatisfactory. In these two tables, there are tabulated only sets of counts with the same number of parallel counts. This is an unnecessary restriction, since the boundaries of the probability classes are independent of the number of replicates. It is therefore legitimate to use the results from experiments in which the numbers of replicates vary from experiment to experiment. The different types of cell counts should of course be kept separate; there should be one control chart for red cells, another for white cells. An unpublished random sampling experiment by the author indicates that the χ^2 test gives satisfactory classification of the counts with a mean as low as five cells if the number of parallel counts are four. Therefore, the great majority of the white cell counts and all the red cell counts likely to be met with in clinical practice can be included in the control chart.

The χ^2 variable, though continuous, is used to approximate to a discrete variable. It might be supposed therefore that some irregularities would occur. This has not been found to be so, except for duplicate counts, where there does not seem to be any suitable adjustment that will give the expected numbers in both the terminal classes. The technique is unsuitable for such cases.

There is an alternative method of assessing control, which in general is not so valuable as the method outlined above, but it is of some value in the last-mentioned case. The additive property of χ^2 may be used. The χ^2 for each experiment and the corresponding number of degrees of freedom may both be summed over a series of experiments. If this gives an unduly high value for χ^2, it indicates that the agreement between parallel plates is unduly poor and an experimental reason for this should be sought. On the other hand, an unduly low χ^2 suggests that parallel counts are not all independent because the observer may, consciously or otherwise, have allowed a knowledge of the previous counts to affect his count of the later replicates. This occurs frequently; examples are given below.

§9. THE VALUE OF STATISTICAL CONTROL IN THE LABORATORY

Berkson has commented on the unsatisfactory nature of the criteria used for a 'good' count. The most common type of criterion in red cell counting is that the range of the five parallel counts of sixteen small haemocytometer squares must not exceed some arbitrary number. This range is usually laid down regardless of the mean of the counts. It would be possible, of course, to devise a method of statistical control using the range and mean, but we have not done so for three reasons:

First, it is difficult to see what adjustments are necessary to allow for continuity in such a measure as the range divided by the estimated standard deviation.

Secondly, the method uses only the two extreme observations and so does not take note of all the available information.

Thirdly, the method is more susceptible to manipulation, conscious or unconscious, on the part of the technician than the χ^2 test.

Further, the χ^2 test is a constant reminder that the accuracy of the counts measured by the coefficient of variation is directly proportional to the square root of the number of cells counted. For the coefficient of variation is $100/\sqrt{m}$. The attempt to discourage the spuriously close agreement which is often given between counts by technicians, is by no means academic. For, far from increasing accuracy, the conscious choice of squares or areas to be counted may reduce it as follows. If the technician sees two areas in good apparent agreement and then seeks other areas not widely different, he is in effect basing his count on the first two areas. Further, if the technique is good, there is no justification for believing that the difference between the observed and true mean of a set of parallel counts is correlated to any appreciable extent with the variance. It is better practice to lay down beforehand which areas will be examined.

The spuriously low variances obtained by conscious choice of squares lead to a further fallacy. Counts made on the same individual at different times or on different individuals are compared by means of a t-test or similar techniques which make use of these spurious observed variances. It is therefore asserted that there is a significant difference when such may not be the case.

These methods of control would also be useful in an investigation into the literature of haemocytometer counts, with the object of examining the internal consistency of the counts made by various authors.

§10. SOME ILLUSTRATIVE EXAMPLES OF STATISTICAL CONTROL

In Table 7 are given the results obtained by two medical graduates not yet specially trained in haematology. The few counts (white cell counts) from Dr H. are in general agreement with the theory. Dr M.'s red cell counts are in good agreement with the theory, but there is a slight excess of white cell counts in the probability class 0·1 to 0·0, showing that there is a relatively large dispersion in too many of his counts.

Table 7. *Statistical control applied to the counts of two observers, medical graduates*

Probability class	Dr H. (white blood cell counts)	Dr M. (white blood cell counts)	Dr M. (red blood cell counts)	Expected percentage
1·0–0·9	0	7	15	10
0·9–0·7	2	9	16	10
0·7–0·5	2	11	15	20
0·5–0·3	1	13	16	20
0·3–0·1	4	13	22	20
0·1–0·0	1	13	16	10
Total	10	66	100	100

Table 8. *Statistical control applied to observers, experienced technical assistants, A and B*

Probability class	Red blood cell count of technician A	White blood cell counts, of A October 1947	White blood cell counts, of A January 1948	White blood cell counts, of A February 1948	Red blood cell counts (1st series) of technician B	Red blood cell counts (2nd series) of B
1·0–0·9	15	8	3	8	24	13
0·9–0·7	23	6	17	17	25	31
0·7–0·5	24	12	14	19	24	22
0·5–0·3	11	20	19	14	25	20
0·3–0·1	18	20	25	33	2	7
0·1–0·0	9	34	22	9	—	7
Total number of counts	100	100	100	100	100	100
Total χ^2	360·96	542·04	423·18	336·42	173·28	316·82
Degrees of freedom	400	300	300	300	400	400

In table 8 the results from two experienced technicians are examined in some detail. The distribution of A's red cell counts evidently differs from the theoretical only by sampling errors. All the χ^2 values have been totalled to give a total of 361·0 for 400 degrees of freedom, in excellent agreement with the theory. The white cell counts were not done so well however. In his first hundred white cell counts there is an excessive number (34 instead of 10) of sets of counts in the class with probability 0·1 to 0·0. Investigation showed that the acetic acid mixture was not sufficiently strong to haemolyse completely the red cells and that clumps of

white cells were often present on the haemocytometer slide. The next set of white counts by the same observer, A, is better, but it was not till after a further strengthening of the white cell fluid to obtain complete haemolysis that satisfactory results, in accord with theory, were obtained. The total χ^2 for these 100 sets of white cell counts has 300 degrees of freedom, and there has been a fall in χ^2 from the unduly high figure of 542 to 336. This last figure is quite in keeping with the theory as can be inferred from Table 9.

Table 9. *Significance levels of χ^2 appropriate to red and white cell counting, and of the ratio of variance to mean*

Probability of the value of χ^2 or the ratio being exceeded by chance	399 degrees of freedom, or single count of 400 squares		255 degrees of freedom, or single count of 256 squares		63 degrees of freedom, or single count of 64 squares		599 degrees of freedom, or single count of 600 squares	
	χ^2	Variance /Mean	χ^2	Variance /Mean	χ^2	Variance /Mean	χ^2	Variance /Mean
0·995	329·1	0·825	199·7	0·783	37·0	0·588	512·7	0·856
0·975	345·1	0·865	212·2	0·832	42·5	0·675	532·6	0·889
0·025	455·8	1·142	300·6	1·179	86·3	1·370	668·2	1·116
0·005	474·6	1·189	315·9	1·239	94·6	1·502	690·9	1·153

These figures have been computed using Fisher's approximation, that $\sqrt{(2\chi^2)} - \sqrt{(2n-1)}$ is a normal deviate with unit standard deviation.

Observer B usually picked his areas to avoid excessive dispersion. This is quite apparent in his first series and somewhat less so in his second series, for which he had been instructed not to select areas at the time of the count but to count on areas fixed beforehand. Even his second series has a total χ^2 significantly low at the 0·995 level of significance. It may be thought that this observer, by zealous shaking or some other device, had reduced his variance and that the Poisson law does not hold in his counts. But in a series of 25 white cell counts with complete enumeration of the cells in each of the 64 small squares, he gave variances completely in agreement with the Poisson. In two other white cell counts, with a high mean, he obtained an excessively high dispersion, evidently owing to an inability to count the number of white cells accurately on a square with more than, say, 15 cells present. The crowding effect in white cell counting is trivial.

SUMMARY

The value of 'statistical control' in haematology has been demonstrated. After a short historical survey of the chief contributions to the theory of the technique of counting and of the application of the Poisson distribution in biology, the statistical theory of the distribution of cells in the haemocytometer is briefly described and examples are given of large sample methods of testing goodness of fit. Some of the author's counts are considered.

In the red cell counts a 'crowding' effect is noted by which the variance between the counts on the individual squares on the haemocytometer is reduced. This effect is trivial in white cell counting owing to the small proportion of the area

occupied, but is important in reducing the variance between the numbers of red cells on individual small haemocytometer squares. A quadratic function adequately describes the regression of the variance on the mean for fixed size of haemocytometer square, for the usual range of red cell density. This function, calculated on the basis of the author's own counts also describes Berkson's findings on crowding adequately and is not, considering the approximations involved, inconsistent with the theoretical discussion on crowding. Reasons are given for believing that the crowding effect is only trivial if blocks of 16 squares are compared in red cell counting.

The methods of statistical control, suggested by R. A. Fisher's work, are introduced. It is shown that they give satisfactory results when applied to some of the author's own counts. The methods are used to discuss the consistency of the counts of certain medical graduates and technicians. The methods are suitable for a review of the literature with the object of examining the internal consistency of the counts made by various authors.

APPENDIX

The variance of the number of cells falling on to a block of squares (Alternative treatment)

Equations (26) and (27) show that for constant values of the size and density of blood cells and the size of the small haemocytometer squares, the ratio between variance and mean of a series of counts depends on the number of small squares over which each count is made. In view of the approximations in the argument of §7 it is of some interest to derive this result by an alternative method. In the present argument only the *empirical* result of §4 is used, that the variance is related to the mean by the quadratic expression (6), no assumption being made about the value of the coefficient b.

Since the cells slip from a crowded square on to adjacent squares, the numbers of cells on adjacent squares are correlated. We suppose that this correlation is represented by the coefficient, ρ. It is supposed that there is no transfer from a square to one diagonally adjacent and that there is no correlation between squares not in contact. Let the variance after reduction by the crowding effect when the cell density is m per square be σ^2.

It is easily seen that there are $2k(k-1)$ pairs of correlated variables, where the variables are the numbers of cells per square. The variance of the total number of cells on a block of k^2 squares is therefore given by

$$V(k^2) = \{k^2 + 4k(k-1)\rho\}\sigma^2. \tag{29}$$

When n is indefinitely large, there is an area not including a margin of width, say, two red cell diameters, from which no cells are lost or exchanged with neighbouring blocks. From the marginal strip there may be some exchange with neighbouring blocks. The dimensions of the inner area are of order k^2 and of the strip, k. Thus the ratio of the expected variance allowing for crowding to that given by the Poisson distribution approaches unity as n becomes indefinitely large. As $k \to \infty$, therefore,

$$V(k^2)/(k^2 m) \to 1. \tag{30}$$

https://doi.org/10.1017/S0022172400015175 Published online by Cambridge University Press

Equating the two limiting values of $V(k^2)$ as $k \to \infty$,

$$(k^2 + 4k^2\rho)\,\sigma^2 = k^2 m, \tag{31}$$

$$4\rho = m/\sigma^2 - 1,$$

$$\rho = \tfrac{1}{4}\{m/\sigma^2 - 1\}. \tag{32}$$

The value of ρ from (32) may be substituted in (29), so that, for finite n,

$$V(k^2) = \{k^2 + k(k-1)\,(m/\sigma^2 - 1)\}\sigma^2$$

$$= k\sigma^2 + k(k-1)\,m. \tag{33}$$

Using the empirical regression equation (6),

$$V(k^2)/(k^2 m) = \{k\sigma^2 + k(k-1)\,m\}/k^2 m$$

$$= \{k(m - bm^2) + k(k-1)\,m\}/k^2 m$$

$$= 1 - bm/k. \tag{34}$$

This corresponds to the result (27) in which $b = 2\lambda$. The case where $k = 4$ is of special interest to us. If the variance of the cells per square is 85 % of the theoretical Poisson variance at a certain cell density, then $bm = 0{\cdot}15$ and bm/k is $0{\cdot}04$; thus the variance of blocks of 16, or 4^2, squares is reduced only by 4 % below the theoretical Poisson value.

REFERENCES

Abbé, E. (1878). Ueber Blutkörper-Zählung. *S.B. Jena. Ges. Med. Naturw.* no. 48.

Armitage, P. (1949). An overlap problem arising in particle counting. *Biometrika*, **36**, 257.

Berkson, J. (1938). Some difficulties of interpretation encountered in the application of the chi-square test. *J. Amer. statist. Ass.* **33**, 526.

Berkson, J., Magath, T. B. & Hurn, M. (1935). Laboratory standards in relation to chance fluctuations of the erythrocyte count as estimated with the hemocytometer. *J. Amer. statist. Ass.* **30**, 414.

Berkson, J., Magath, T. B. & Hurn, M. (1939). The error of estimate of the blood cell count as made with the hemocytometer. *Amer. J. Physiol.* **128**, 309.

Bortkiewicz, L. (1898). *Das Gesetz der kleinen Zahlen.* Leipzig.

Eisenhart, C. & Wilson, P. W. (1943). Statistical methods and control in bacteriology. *Bact. Rev.* **7**, 57.

Fisher, R. A. (1948). *Statistical Methods for Research Workers*, 10th ed., London: Oliver and Boyd.

Fisher, R. A., Thornton, H. G. & Mackenzie, W. A. (1922). The accuracy of the plating method of estimating the density of bacterial populations. *Ann. appl. Biol.* **9**, 325.

Fisher, R. A. & Yates, F. (1948). *Statistical Tables for Biological, Agricultural and Medical Research*, 3rd ed. London: Oliver and Boyd.

Garwood, F. (1947). The variance of the overlap of geometrical figures with reference to a bombing problem. *Biometrika*, **34**, 1.

Goodall, A. (1908). Review of Student (1906). *Folia haemat., Lpz.*, **5**, 430.

Lyon, J. F. & Thoma, R. (1881). Ueber die Methode der Blutkörperzählung. *Virchows Arch.* **84**, 131.

Magath, T. B., Berkson, J. & Hurn, M. (1936). The error determination of the erythrocyte count. *Amer. J. clin. Path.* **6**, 568.

Plum, P. (1936). Accuracy of haematological counting methods. *Acta med. scand.* **90**, 342.

Poisson, S. D. (1837). *Recherches sur la probabilité des jugements*, pp. 205 *et seq.* Paris.

'Student' (1907). On the error of counting with a haemocytometer. *Biometrika*, **5**, 351.

'Student' (1919). An explanation of deviations from Poisson's law in practice. *Biometrika*, **12**, 211.

Whitaker, L. (1914). On Poisson's law of small numbers. *Biometrika*, **10**, 36.

Yule, G. U. & Kendall, M. G. (1950). *An Introduction to the Theory of Statistics*, 14th ed. London: Griffin.

(*MS. received for publication* 7. VI. 50.)