THE ERROR OF THE RED CELL COUNT

BY

ROSEMARY BIGGS AND R. L. MACMILLAN

From the Department of Pathology, Radcliffe Infirmary, Oxford

(RECEIVED FOR PUBLICATION, JULY 25, 1948)

The staffs of haematological laboratories spend much of their time counting red cells. The precision of this estimation is therefore of practical importance, and many experiments have been designed to establish the error of the technique. In spite of this, the most widely divergent views are still held as to the reliability of single observations.

In 1881 Lyon and Thoma showed that the standard error of counts made on the same sample of blood was roughly proportional to the square root of the number of cells counted. Thus in a count of 100 cells the standard error would be roughly $\sqrt{100}$, and repeat counts on the same sample might by chance vary from 80 to 120 cells, or, in counts of 500 cells, the standard error would be approximately $\sqrt{500}$ or 4.5 per cent. In 1906-7 "Student" confirmed and stated more exactly the findings of Lyon and Thoma when he showed that the scatter of cells in a haemocytometer followed the Poisson distribution, in which the standard error is equal to the square root of the mean for the distribution. Thus, if 500 counts are made on one sample of blood and the mean value for each is 100 cells, the standard error is $\sqrt{100}$ and the majority of counts would show a range from 80 to 120 cells. From these observations it appears that there is an irreducible error in making a red-cell count which is approximately proportional to the square root of the number of cells counted.

In 1935 and 1940 Berkson, Magath, and Hurn again demonstrated the Poisson distribution of the red cells in a haemocytometer and made a thorough analysis of the other errors inherent in this technique. They showed that in a normal count the standard error is 7.7 per cent. An error of this magnitude implies that two counts on a normal sample of blood may differ by up to 1,100,000 ($\pm \sqrt{2S}$). In 1947 Hynes showed that the distribution of cells varies from one end of the

coverslip to the other, and unless there is a reasonable length of coverslip on either side of the ruled area the error may be larger.

Wintrobe (1946) on the other hand maintained that repeat counts on a single sample should agree to within 200,000 cells, implying a standard error of 1.4 per cent, and several workers have published results in which this error has been achieved or even bettered (Mayers, 1922; Smith, 1931; Wintrobe, 1934; Price-Jones and others, 1935). This error is considerably less than that due to the Poisson distribution alone. Wintrobe (1946) does not recognize this distribution because he holds that the number of cells enumerated in the five separate squares of a red-cell count should not differ by more than 18. A greater internal variation is thought to indicate incomplete mixing of the contents of the pipette, and according to this criterion counts showing a greater variability should be rejected. In the Poisson distribution, on the other hand, differences of up to 40 cells might occur by chance. Wintrobe's opinion is apparently shared by the many skilled workers who habitually apply this criterion of adequate mixing. In fact, it is probably true that in spite of Berkson's careful experiments and statistical analysis many believe that Wintrobe's assessment of the error is more nearly true.

Experiments

We have therefore made several experiments in an attempt to find the origin of this apparently irreconcilable discrepancy.

Experiment 1.—Two reliable and skilled technicians made ten counts on a single sample of normal blood, using five pipettes and five counting chambers. In all these experiments one pipette was always used with the same counting chamber. The cells in a volume of 1/10,000 c.mm. of blood in five large squares were counted. The values for each square were recorded

TABLE I
THE NUMBERS OF RED CELLS RECORDED FOR EACH OF THE COUNTS IN EXPERIMENT 1

Observer	Pipette									
	1		2		3		4		5	
1	503	481	467	498	496	496	494	497	496	492
2	481	481	489	482	482	477	480	478	465	458

separately and the workers were requested to apply Wintrobe's criterion for complete dispersion of the cells in the pipette; that is, to reject counts showing a variation greater than 18 cells between the five squares. They were aware that they were dealing with only one sample of blood, and were encouraged to compare results. The figures obtained showed close agreement (Table I). The first technician's counts had a coefficient of variation of 1.9 per cent and those of the second 2.1 per cent. In this experiment the technicians separately obtained a precision closely approaching Wintrobe's requirements. There was, however, a fairly large difference between the two

sets of results, one giving a mean value of 477 and the other of 492.

Experiment 2.—Five technicians made five counts on one sample of normal blood using five pipettes and five counting-chambers. The blood was divided into five fractions, which the technicians had no reason to believe were derived from the same person. They did not compare results, but were requested to use Wintrobe's criterion for complete mixture of the contents of the pipettes. From Table II it will be seen that these counts show far greater variation than those of the first experiment. The coefficient of variation is now 7.6 per cent.

TABLE II
THE NUMBERS OF RED CELLS RECORDED FOR EACH OF THE COUNTS IN EXPERIMENT 2

Observer	Pipette				
	1	2	3	4	5
1	430	404	437	453	517
2	475	405	455	412	443
3	437	440	508	435	488
4	488	449	485	463	505
5	458	397	458	436	447

Experiment 3.—Five doctors who, though competent in the technique, were not engaged in red-cell counting as part of their daily work, made ten counts on the sample of blood used in Experiment 2, using ten pipettes and counting-chambers. These workers inevitably made their counts more slowly than the technicians, and—choosing the squares at random—were unable to obtain a close agreement between the counts for the different squares; a difference of up to 40 cells was allowed. The results show an even greater variation, and the standard error is 9.5 per cent (Table III).*

* That this error is greater than that recorded by Berkson may probably be due to two main factors: (1) Berkson used an electric counter and enumerated the cells from photographs, whereas in these experiments the cells were counted by eye. (2) Berkson's experiments were made by one observer, and our experiments show that an error of 3 per cent may be attributed to differences between observers.

TABLE III
THE NUMBERS OF RED CELLS RECORDED IN EACH OF THE COUNTS IN EXPERIMENT 3

Observer	Pipette									
	1	2	3	4	5	6	7	8	9	10
1	427	372	418	440	349	484	430	416	449	464
2	434	420	385	472	415	420	415	396	439	424
3	480	421	473	496	474	411	472	423	502	488
4	451	369	500	464	444	410	422	396	459	471
5	462	453	450	520	489	409	508	347	440	391

Discussion

These experiments show that the magnitude of the error depends entirely on the condition of the experiment. The main components of error from an analysis of variance are shown in Table IV.

TABLE IV

THE COMPONENTS OF ERROR DERIVED FROM AN ANALYSIS OF VARIANCE IN EXPERIMENTS 1, 2, AND 3

Experiment	Total error %		Error due to observers %	Error due to pipettes and counting chambers %	Random error %
	Separate observers	Both observers			
I	1.9 2.1	3.6	2.0	0	2.5
II	7.6		2.9	4.8	5.1
III	9.5		3.1	4.2	7.8

The striking feature is the progressive increase in random error. In experiment 1, which is similar to those described by Wintrobe (1934) and Price-Jones and others (1935), the error is small if the results obtained by the two technicians are considered separately. The random error is well below that of the Poisson distribution (4.5 per cent). There is no component of error referable to the calibration and filling of pipettes and calibration of counting chambers, factors well recognized as a source of error in haematological technique. Moreover, it can be shown statistically that there is less than a 1/100 probability of obtaining mean values as widely divergent as 477 and 492 from the same sample of blood when the counts for each observer are so uniform.

Since there are good reasons for suspecting that the standard error recorded in the first experiment is not a true estimate, and since workers not engaged in red-cell counting as a routine are unable to obtain any close agreement, it is perhaps of interest to examine the process of training which leads to the making of uniform counts. This has been described by Emerson (1921), who recommended his medical students to make repeated counts on one person's blood until they achieved an agreement of 200,000 cells. He says, "Some students attain this accuracy quickly. Some, however, repeat this daily counting for twenty or thirty or even more days before their work is satisfactory to themselves or to us. By this time

they have certainly learned wherein lies the error of their technique. It is of interest that the most careful ones sometimes make the greatest errors since they take too much time where speed is essential." It seems probable that a training of this sort leads to an unconscious bias in favour of agreement between counts, a bias that was well demonstrated in Experiment 1.

In Experiment 2 there was no reason to obtain any agreement between the five sub-samples of blood, but the values for the last four squares of each count were biased by that of the first square (because an agreement between the five squares of 18 cells or less was required). In the third experiment the full range of random variation was shown.

It seems probable that any artificial reduction in the true variability will lead to a decrease rather than an increase in precision; nevertheless it is possible that a skilled observer trained to make uniform counts on a single sample of blood might be able to select from each counting chamber squares typical of the whole field, which would estimate the mean value more precisely than a random selection. Experience in many sampling problems, however, suggests that subjective selection of this kind is a frequent source of unconscious bias. To test this point a further experiment was made.

Experiment 4.—Five technicians made routine red cell counts on eighteen different samples of blood and recorded separately the figures for each square. The counting chambers were then taken and the cells in all

TABLE V

ONE OF THE 18 COUNTS MADE IN EXPERIMENT 4.

(In the top row are the separate counts for each of the five squares obtained by the technicians. The counts for the separate squares of the whole field are recorded below.)

Technician's counts	80	82	79	82	77
Counts for whole field	77	69	62	69	70
	64	94	77	71	67
	85	73	66	67	72
	62	89	65	72	77
	64	64	67	77	62

the squares in the ruled area were counted. The results of one such experiment are shown in Table V. The mean value per square obtained as an average for the whole field is clearly a better estimate of the true mean value than that derived from any five

squares. Nevertheless if the five squares were chosen at random their mean value should not differ significantly from that for all the squares. When the mean values for each selection of five squares obtained by the technicians are compared with the corresponding mean for the whole field, several of the eighteen samples show a large difference, and statistical tests show that the dispersion about the field means is significantly greater than would be expected for the means of random samples. In other words, on an average, a random selection of five squares would give a more trustworthy estimate of the field mean than would the set of five selected by one of these technicians; the insistence on agreement between counts for the five selected squares has achieved a bias in estimation rather than an increase in precision.

Conclusions

It is concluded that : (1) the training of workers who make routine red cell counts should be a training in accurate counting rather than a training in achieving agreement between replicate counts ; (2) in making a red cell count a difference

of up to 40 cells should be allowed between the five separate squares counted ; (3) from the most reliable estimates the standard error of the red-cell count lies between 8 and 10 per cent.

We should like to thank Dr. R. G. Macfarlane for his continued interest and advice, Mr. Finney for assistance with the statistical analysis, and graduate assistants and technicians of the department of haematology for their patience and willing co-operation in carrying out the experiments.

REFERENCES

- Berkson, J., Magath, T. B., and Hurn, M. (1935). *J. Amer. statist. Ass.*, **30**, 414.
 Berkson, J., Magath, T. B., and Hurn, M. (1940). *Amer. J. Physiol.*, **128**, 309.
 Emerson, C. P. (1921). "Clinical Diagnosis." Philadelphia, p. 459.
 Hynes, M. (1947). *J. clin. Path.*, **1**, 25.
 Lyon, J. F., and Thoma, R. (1881). *Virchows Arch.*, **84**, 131.
 Mayers, L. H. (1922). *Arch. intern. Med.*, **30**, 478.
 Price-Jones, C., Vaughan, J. M., and Goddard, H. M. (1935). *J. Path. Bact.*, **40**, 503.
 Smith, C. (1931). *Arch. intern. Med.*, **47**, 206.
 "Student" (1906-7). *Biometrika*, **5**, 351.
 Wintrobe, M. M. (1934). *Arch. intern. Med.*, **54**, 256.
 Wintrobe, M. M. (1946). "Clinical Haematology." Philadelphia, p. 253.