LABOUR SAVING METHODS IN CELL COUNTS

By BARNET WOOLF

Department of Public Health and Social Medicine, University of Edinburgh

THE following problem was put to me by Dr R. H. Girdwood of the Department of Medicine at Edinburgh. He wished to follow changes in the megakaryocyte content of the bone marrow in some of his patients. Marrow smears are contaminated with blood, but it is safe to assume in the cases under consideration that the amount of marrow tissue is proportional to the number of nucleated white cells. The relative megakaryocyte content, R, is therefore $\frac{P}{Q}$, where P is the megakaryocyte count and Q the white cell count in a convenient number of microscope fields. This is the method used by Dameshek and Miller (1946). Now Q is extremely large compared with P, making it necessary to count many thousands of cells to include a moderate number of megakaryocytes. The problem is, can one save time and labour by counting the white cells in a fraction $\frac{1}{a}$ of the fields, so that $R = \frac{P}{aQ}$; what would be the standard deviation of a value of R so obtained; and what are the best values of P and Q to secure a desired precision in R with a minimum amount of counting?

I propose to extend the scope of the problem. Cell counts usually have one of the following objectives:—

(*a*) Estimation of the number of cells in a given volume of fluid or tissue, as in the ordinary blood erythrocyte count.

(*b*) Estimation of the relative frequencies of two or more kinds of cell, as in differential leucocyte counts, platelet ratio determinations and the megakaryocyte method described above.

(*c*) Estimation of the fraction of a given species of cell showing atypical characteristics, as in counts of the percentage of reticulocytes, siderocytes or punctate basophilia in the blood erythrocytes.

I propose to deal only with objectives (*b*) and (*c*). In objective (*b*), the desired result is a ratio. If P and Q are the numbers counted of the two types of cell, the answer, R, is $\frac{P}{Q}$. In objective (*c*), the desired result, F, is the fraction $\frac{P}{P+Q}$. F cannot exceed unity, while R may have any numerical value.

It is usual to express the range of uncertainty of a measurement in terms of its standard deviation. There are various algebraically

equivalent formulæ for the standard deviation of a ratio or of a fraction. The forms most useful for the present purpose are :—

$$\frac{\text{Standard Deviation of R}}{\text{R}} = \sqrt{\frac{1}{\text{P}} + \frac{1}{\text{Q}}} \quad . \quad . \quad . \quad (1)$$

$$\frac{\text{Standard Deviation of F}}{\text{F}(1-\text{F})} = \sqrt{\frac{1}{\text{P}} + \frac{1}{\text{Q}}} \quad . \quad . \quad . \quad (2)$$

These formulæ bring out clearly how the precision of an estimate depends on the numbers of the two types of cell that are counted.

Now it can be shown that when P and Q are sampled on different scales, so that

$$\text{R} = \frac{\text{P}}{a\text{Q}} \quad \text{or} \quad \frac{a\text{P}}{\text{Q}}$$

and

$$\text{F} = \frac{\text{P}}{\text{P}+a\text{Q}} \quad \text{or} \quad \frac{a\text{P}}{a\text{P}+\text{Q}}$$

formulæ (1) and (2) are still true and applicable. It can further be proved that, for any given value of the standard deviation, the number of cells required to be counted is a minimum when P = Q. Hence the simple practical rule :—

> " Choose that value of $\frac{1}{a}$ which will make the counts of the two kinds of cells as near equality as possible."

It is easy to see the commonsense justification of this rule. When $\text{R} = \frac{\text{P}}{\text{Q}}$, there is no point in laboriously measuring Q to an accuracy of 1 in 1000 when P is accurate only to within 1 in 10.

When there is a big disparity in the frequencies of the two kinds of cell, the time saved by this method of " balanced sampling " is very great. For instance, the siderocyte content of a healthy blood may be as low as 2 per 1000 red corpuscles. A count of 10,000 cells would give about 20 siderocytes and 9980 normal erythrocytes. The standard deviation of the estimate of F will be proportional to $\sqrt{\frac{1}{20} + \frac{1}{9980}}$, which is approximately $\sqrt{0{\cdot}0501}$, or about 0·224. If the procedure is adjusted to yield equal numbers, by counting normal cells in about $\frac{1}{500}$ the area used for sampling siderocytes, a total count of 100 cells will give a standard deviation proportional to $\sqrt{\frac{1}{50} + \frac{1}{50}}$, which is $\sqrt{0{\cdot}04}$ or 0·200, greater precision being obtained with only 1 per cent. of the labour of counting. On the other hand, there is little advantage to be gained if the ratio of the two kinds of cell is less than 4 : 1, at which point there is an economy of 36 per cent.

Table I shows, at different levels of cell ratio, the number of cells that must be enumerated in a " straight " count to give an estimate as accurate as that obtained with a total count of 100 cells by the method of balanced sampling. In their marrow smears, with a frequency ratio of about 5,000 : 1, Damashek and Miller (1946) counted at least half a million cells. The table shows that they could have done as well by balanced sampling with a total count of about 400.

In general, if $\frac{1}{a}$ is chosen to make P = Q, the total number of cells required to be counted in balanced sampling is the fraction $\frac{4a}{(a+1)^2}$ of a number needed in a direct count, for the same precision of estimate.

The practical details of the method vary with the type of tissue preparation. I will discuss three cases: films counted under the high power of the microscope, films counted under low power, and cell counts in a ruled hæmocytometer.

TABLE I

Relative Counts by Direct Method and Balanced Sampling for Equal Precision

Frequency Ratio of Two Kinds of Cell.	Direct Count as Percentage of Balanced Sampling Count.
1 : 1	100
2 : 1	113
3 : 1	133
4 : 1	156
5 : 1	180
6 : 1	204
7 : 1	229
8 : 1	253
9 : 1	278
10 : 1	303
15 : 1	427
20 : 1	551
30 : 1	801
40 : 1	1,051
50 : 1	1,301
100 : 1	2,550
200 : 1	5,050
500 : 1	12,550
1,000 : 1	25,050
2,000 : 1	50,050
5,000 : 1	125,050
10,000 : 1	250,050

In preparations counted under oil immersion, there are usually relatively few cells in each field, so that a large number of fields have to be explored. I will take an imaginary case in which each field contains about 15 of the more abundant cells, while the rarer cells occur at the rate of about one in every two fields. Approximate equality of the two counts will be obtained by counting the abundant cells in one out of every thirty of the fields scanned for the rarer cells.

Suppose it has been decided, on the principles discussed below, that sufficient accuracy can be attained with a count of 200 of each kind of cell. That means using about 400 fields for the rare cells, and about 13 or 14 for the common kind. The fields are counted in groups of 30, out of which a definite one, say the fifteenth, is used for the abundant cells. This goes on until the count of the rarer cells reaches 200. It is important that the count shall not stop there, but shall continue until the last group of 30 fields is completed, and any additional cells found must be included in the reckoning. The general rule is that the fields must be counted in groups containing an average of at least ten of each kind of cell, and the count must end with a complete group. In accordance with this rule, if in the example just discussed there are only eight of the commoner cells in each field, it is best still to operate with groups of 30 fields, and to count 2 fields in each group—say the tenth and twentieth—for the common cells.

Now suppose that, with an average of about 15 common and ½ rare cells per field, the degree of accuracy aimed at can be attained with a count of 100 of each kind of cell. This involves exploration of about 200 fields, of which about seven will be used for counting abundant cells. Now 7 fields would probably not be enough to even out the bias that may arise from local variations of cell density in the preparation. The general rule should be to use at least 10 fields. In the case under consideration, the fields could be counted in groups of 20, using one in each group for full enumeration. This would involve counting more of the abundant cells than is theoretically necessary, in order to promote fair sampling. Every case can be dealt with by these four rules—decide on the minimum number to be counted on the principles discussed below, take the fields in groups containing a minimum of 10 cells of each kind, use at least 10 fields for the commoner type of cell, if in doubt increase the numbers to be counted. When there are large numbers of cells in each field, the method of sampling recommended for low power counting may be adopted.

Balanced sampling is facilitated when the counting is done in a hæmocytometer. The rarer cells will be counted in the large squares, while the more abundant cells will be enumerated in an appropriate number of the small squares. Suppose, for example, that each large square is divided into 16 small squares. When the ratio of frequencies is about 16 : 1, one small square is counted for the abundant cells in each large square counted for rare cells. If the ratio is about 8 : 1, two small squares are counted in each large square. When the ratio is about 30 : 1, one small square is used in each pair of large squares. If the large squares average less than 10 rare cells, it is still necessary to count them in groups. But there is no need to bother about even sampling of the common cells, because that is assured automatically.

I now come to counting under the low power. The essential distinction between low and high power counting does not, of course, arise from the degree of magnification, but from the fact that under

the low power a sufficient number of the rarer cells will usually be found in a very few fields. Suppose that the frequency ratio is about 100 : 1, and the count of the rarer cells can be completed in 5 fields. One is then faced with the problem of counting the abundant cells in an area equal to $\frac{1}{20}$ of a field, and doing this, moreover, in such a way as to ensure even sampling. One solution would be to decrease the field areas by counting under a high power. Another would be to insert a mask in the eyepiece of the microscope, cutting off all except a small square in the middle of the field. But the most satisfactory procedure would be to equip the microscope with a ruled sampling plate. A standard ocular micrometer grid could be inserted, and

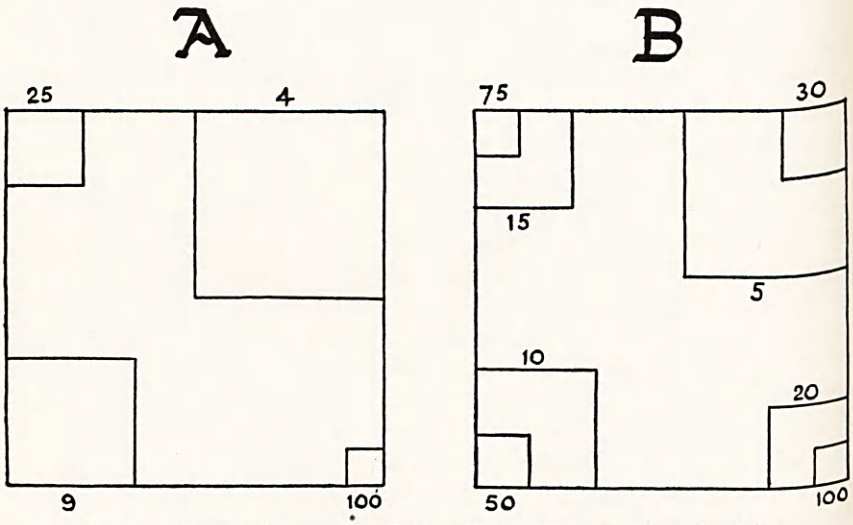

FIG. 1.—Designs for sampling plates. The numbers give the ratio of the area of the small to the total square.

used as recommended for the rulings in a hæmocytometer. Or a specially designed sampling plate could be made. Suitable patterns for such plates are shown in Fig. 1.

In pattern A there is a large square for counting the rarer cells, and 4 smaller squares, any of which may be chosen for sampling the more abundant cells. Convenient values for the sides of the small squares would be $\frac{1}{2}$, $\frac{1}{3}$, $\frac{1}{5}$ and $\frac{1}{10}$ the side of the large square, giving sampling ratios of 1 : 4, 1 : 9, 1 : 25 and 1 : 100. Other ratios can readily be obtained. For instance, a ratio of 50 : 1 would be achieved by counting abundant cells in the $\frac{1}{25}$ square once in every pair of large squares scanned for rare cells. A greater choice of ratios is offered if one utilises each corner for more than one smaller square, as in pattern B. This pattern also illustrates how to obtain any desired

ratio by making the side of the smaller square proportional to the square root of the sampling fraction. For the ratios shown in the figure the sides of squares are as follows :—

Large Square / Small Square	5	10	15	20	30	50	70	100
Side of Small Square	·4472	·3162	·2582	·2236	·1826	·1414	·1195	·1000

Other ratios or designs could easily be evolved to suit any particular requirements. These sampling plates are quite easy to make. The design should be drawn large in Indian ink on white card, and reduced photographically to the required size on to sensitised film, which can then be cut to a disc to rest on the diaphragm in the microscope ocular, where it can be kept flat if necessary by a metal ring, or it could be cemented to a glass disc. The best film is the 35 mm. high contrast, high resolution kind used for micro-copying of documents. Most university libraries have the film and apparatus for making the finished plate from the original drawing, or it could be done by any good commercial photographer. Or one could use the special Graticule photographic plates. It is easy to have several positives made at the same time to keep as spares. In counting with a sampling plate, it is still necessary to take units of area containing at least ten of each kind of cell, but as in the hæmocytometer, fair sampling of the abundant cells is automatically secured.

It may seem that one cannot fix the sampling ratio until one knows what the answer to the count is going to be. But in practice no difficulty need arise. One usually has a fair idea beforehand of the relative frequency of the two kinds of cells, or can get a near enough figure by counting a field or two. If the sampling ratio is badly misjudged, the count should be continued until the required number has been obtained of the kind of cell whose count is lowest, always of course finishing with a complete group of fields. The only penalty of a bad judgment is a small increase in the numbers to be counted, with of course a corresponding gain in precision. The size of the unit group of fields can also be adjusted during the count. If, for example, one decides on a unit group of three squares of fields, and it appears that the average content is only about seven cells, one would make the count cover an even number of such groups, which is equivalent to increasing the unit to six squares or fields.

The number of cells to be counted depends on the least difference in which one is interested. There are commonly two kinds of question to which a count is expected to supply the answer. These are :—

(*a*) Does the ratio or fraction in this specimen differ by more than a specified amount from some fixed value, usually the accepted figure for the " normal " ?

(*b*) Is there a genuine difference in the relative cell frequencies in two or more specimens, which will often be successive readings on the same patient ?

A guide to the numbers required to answer the first type of question is given in Table II. Suppose, for example, that the result of the count is a value 10 per cent. above the figure one has chosen to regard as the normal. The table shows that this result cannot be trusted unless at least 800 cells of each kind have been counted, making a

TABLE II

Size of Count to Detect Departures from a Fixed Standard Value

Limits Outside which a Difference is Significant.	Number of Cells of Each Kind to be Counted.	
Per cent.	Odds of 20 to 1.	Odds of 100 to 1.
110—91	800	1400
115—87	360	630
120—83	200	350
125—80	130	230
130—77	90	160
140—71	50	90
150—66	40	60
200—50	12	20

total count of 1600. Even with this number, one would make a wrong decision about once in 20 times, because that is the frequency with which a deviation of 10 per cent. or more occurs by chance in a count of this size. To be sure of being right 99 times out of 100, one would have to count 1400 cells of each kind, or a total of 2800. It is, however, very rarely that one is interested in differences as small as this. For most physiological measurements, 10 per cent. is well within the range of normal variation. It may be that an excess value is of no clinical significance unless it is 50 per cent. or more above the normal standard. In that case, one need only count 40 cells of each kind to be right about 19 times out of 20, or 60 cells of each kind to be right about 99 times out of 100. With a count of this size, any result between 150 per cent. and 66 per cent. of normal is dismissed as " not significantly different from 100 per cent." If one wishes to work to closer limits, say between 130 per cent. and 77 per cent., one must count at least 320 cells—160 of each type. In routine clinical work it is probably a safe general rule to count about 100 of each type of cell and to regard all results between 140 and 71 per cent. as not significantly different from the standard. In the rare cases where narrower limits are desirable, the size of count needed can be obtained from Table I, or from the formulæ given below.

The numbers required when working to limits not given in Table II can readily be calculated from the following formulæ. In all cases

P = Number of each type of cell to be counted, so that total count is 2P.

x = Required percentage difference *above* normal. For instance, if the limit is to be 125 per cent., $x = 25$.

y = Required percentage difference *below* normal. For instance, if the limit is to be 80 per cent., $y = 20$.

Then, for odds of 20 to 1 against a mistake :—

$$P = \frac{80{,}000}{x^2} \quad \text{or} \quad \frac{8(100-y)^2}{y^2} \quad . \quad . \quad . \quad (3)$$

And for odds of 100 to 1 against a mistake :—

$$P = \frac{140{,}000}{x^2} \quad \text{or} \quad \frac{14(100-y)^2}{y^2} \quad . \quad . \quad . \quad (4)$$

The value of y corresponding to a given value of x is given by the formulæ :—

$$y = \frac{100x}{100+x} \quad . \quad . \quad . \quad . \quad . \quad (5)$$

In practice, it is convenient to round off the value of P found from formula (3) or (4) to the next highest multiple of ten. When P works out at 20 or less, add on 50 per cent.

Table III shows the least numbers that must be counted when one is interested in establishing a disparity between two specimens. The

TABLE III

Size of Count to Detect Differences Betweeen Two Estimations

Limits Outside which a Difference is Significant.	Number of Cells of Each Kind to be Counted.	
Per cent.	Odds of 20 to 1.	Odds of 100 to 1.
105—95	6,730	11,780
110—91	1,770	3,100
115—87	830	1,450
120—83	490	860
125—80	330	580
130—77	280	420
135—74	190	330
140—71	150	260
150—66	110	190
$166\frac{2}{3}$—60	70	120
175—57	60	110
200—50	40	70

first column shows the least difference that is to be detected, one of the two estimations being expressed as a percentage of the other. Column 2 shows the smallest number to be counted of each type of cell to ensure an answer that will be right about 19 times out of 20, and column 3 the numbers to give a trustworthy result 99 times out of 100. If more cells are counted, the chance of being wrong will be

correspondingly decreased. The table is used in this way. If one is not concerned with increases of less than 50 per cent., one counts 190 cells of each kind, making a total of 380 cells in each count. If the results in the two specimens show a smaller disparity than 150 : 100, the difference between them is regarded as not significant. The majority of cases that arise in practice will be covered by the following simple rules :—

Total Count in Each Specimen.	Minimum Detectable Disparity.
200 cells	$1\frac{3}{4} : 1$
400 ,,	$1\frac{1}{2} : 1$
800 ,,	$1\frac{1}{3} : 1$

The numbers required to establish minimum differences not shown in Table III may be calculated from the following formula :—

$$P = \frac{4+4z+2z^2}{z^2} \times T \quad . \quad . \quad . \quad . \quad (6)$$

where P = Number of each type of cell to be counted.

z = Difference expressed as a decimal. For instance when disparity is 125 : 100, $z = 0{\cdot}25$. When it is 80 : 100, $z = 0{\cdot}20$.

T = 4 for odds of 20 to 1.

= 7 for odds of 100 to 1.

As before, values of P should be rounded off to the next highest multiple of ten. When P is 20 or less, add on 50 per cent.

It must be emphasised that the numbers in Tables II and III refer only to counts made by balanced sampling. For any other method, much larger numbers must be counted to get the same precision.

Even in a research problem, it is seldom worth while counting more than 800 or 1000 cells in any one specimen. It may seem from Table III that much greater accuracy can be obtained with marathon counts, but this is usually purely illusory. By increasing the numbers counted, one can go on improving the precision of the count *on the individual slide or preparation*. But this does not necessarily mean a better estimate of the cell distribution in the patient's blood or tissues. Duplicate blood punctures, especially in peripheral sites, may give quite discordant cell counts. There is no reason to suppose that there are no local differences in the relative numbers of megakaryocytes and white cells in the bone marrow. No cell count, however extensive, can eliminate the other sources of error inherent in the technique of specimen taking. Where extreme accuracy of cell counting is wanted, the only reliable method is to take a number of independent specimens

and count a moderate number of cells in each, pooling the numbers to obtain the final answer. Several duplicate punctures on patients for each estimation will of course usually be quite impracticable. But any research worker using cell counts can at least test his technique by taking several independent specimens from a human guinea-pig, and finding out how much they differ among themselves, and then regulate his size of count accordingly. If the unavoidable errors in taking the sample give rise to a standard deviation of say 10 per cent., it is mere self-deception and waste of labour to count 10,000 cells and regard the result as accurate to within a standard deviation of 2 per cent.

The method of balanced sampling has two important advantages. One is obvious, the great economy of labour when cell frequencies are widely different, exemplified by the figures in Table I. The other is the fact that, whatever the relative cell frequencies in the preparation, the same number of cells counted always gives an estimate with approximately the same precision. This obviates the need to calculate standard deviations and significances of differences for each individual case. If in a given type of estimation one always counts the same number of cells, then one is always working to a known degree of precision and can see at once whether apparent differences are or are not statistically significant. If one wants to increase or decrease the precision desired, the number to be counted can be read off from Table II or III, or easily calculated from the formulæ.

It will be seen from Table I that when the frequency ratio of the two kinds of cell is less than 4 : 1, there is little advantage in balanced sampling over a direct count. In such cases, it would be simpler to count a specified number of cells. It is desirable to choose the numbers to be counted so as to maintain the principle that all estimations shall have the same relative precision. When the result is to be expressed as a fraction, $\frac{P}{P+Q}$, this can be effected by using the figures given in Tables II and III, counting a total of twice the number stated for each level of accuracy. In this way, if the total count in balanced sampling is standardised at 200 cells, it will still be 200 when direct counting is used for low frequency ratios. When the result is to be expressed as a ratio, $\frac{P}{Q}$, rather more will require to be counted. The exact numbers can be calculated by multiplying the standard total by the percentages shown in Table I—113 per cent. when the frequency ratio is 2 : 1, 133 per cent. when it is 3 : 1, and 156 per cent. when it is 4 : 1. A simpler procedure would be to add on 50 per cent. when the frequency ratio is between 2 : 1 and 4 : 1, and 15 per cent. when it is between 1 : 1 and 2 : 1.

Incidentally, Tables II and III can be used as a guide to the precision of a count of a single type of cell, such as the ordinary erythrocyte count in blood. The limits will apply when half the stated numbers of cells are counted. For instance, if one wishes to detect all red cell

counts that are less than 80 per cent. of a fixed normal, with only one chance in 100 of being deceived, Table II indicates that one must count at least $\frac{1}{2} \times 230$, or 115 cells.

Although I have dealt only with cell counts, it is obvious that the method of balanced sampling can be used in many other kinds of problem, including the selection of experimental and control groups in clinical experiments and trials. The method has, in fact, been used by Drillien (1947) in a study on prematurity and infant mortality, in which her control group was a 1 in 10 sample of full term babies. The relevant formulæ are given in a brief note (Woolf, 1947) to one of her papers. The method described in this paper is a special application of a general theory of weighted and stratified sampling, which will be published, with proofs of the various formulæ, in a statistical journal.

Summary

A method of differential cell counting is described, which involves sampling the more common type of cell in a smaller area than the rare type. When there are great disparities in the relative frequencies, enormous economies in time and labour are possible. Full practical instructions are given, including tables and formulæ for finding easily the standard deviation of the result of the count.

It is a pleasure to thank Dr R. H. Girdwood for bringing this problem to my notice, and Professor F. A. E. Crew, F.R.S., for his interest and support.

REFERENCES

Damashek, W., and Miller, E. B. (1946), *Blood*, **1**, 27.
Drillien, C. M. (1947), *Journ. Obstet. and Gynæc.*, **54**, 300, 443, 452.
Woolf, B. (1947), *Journ. Obstet. and Gynæc.*, **54**, 322.