Bulletin of the World Health Organization, 58 (2): 313-320 (1980)

A statistical procedure for quality control in diagnostic laboratories

M. J. GODDARD [1]

The accuracy of results from any epidemiological project depends upon the performance of all stages of data collection and handling. One link in this chain may involve microscopic inspection of specimens and the purpose of this article is to describe means of assessing the accuracy of this stage. Two aspects merit consideration: the design of a sampling system and the collection and interpretation of the results. The importance of design is stressed, but the considerable variation in individual circumstances rules out a full treatment of the topic. A method of record keeping is described together with a method for graphical presentation of results. A procedure is outlined for detecting unacceptable performances and some of the statistical considerations are discussed.

The large amounts of effort and money that are being invested in epidemiological and clinical research have resulted in the introduction of new methods and devices capable of increasing accuracy and reproducibility. However, there are many steps in any project between the drawing of specimens from a field population and the publication of results. Given that precise results can be obtained in only a few stages of any process, the validity of any set of data is determined by whichever step introduces the most error.

This article is concerned primarily with the detection and assessment of errors made by microscopists examining specimens in diagnostic laboratories. Even with strictly controlled sample preparation and microscopes capable of excellent resolution, a microscopist may produce incorrect results for a multitude of reasons, both psychological and physical.

One way of determining the frequency and magnitude of such incorrect results is by performing checks on a sample of a technician's counts. Regimens to perform this task routinely have been termed "acceptance sampling" schemes and have originated in industrial situations. The implementation of such schemes is fairly simple, the cost requirement is quite low, and results are readily obtained, summarized, and understood. Such tools offer the same kind of benefit as an insurance policy: when things "go well" there is no worry, but when a technician begins to give poor results this will quickly become evident.

One example of an "acceptance sampling" or quality control scheme has been used in St Lucia in the West Indies (*1*). A much more traditional scheme, whose statistical properties are better known, is presented below. The design of such a sampling scheme is discussed and suggestions are made on the method of recording results. The graphical presentation of results is also outlined. Finally, a statistical test that may be applied to the data obtained is described.

DESIGN IN QUALITY CONTROL

The topic of design, although dealt with only briefly here, is of considerable importance. Prospective users of quality control schemes should be particularly wary of employing strict statistical tools on haphazardly operated sampling schemes. The considerations of design in quality control are not unlike those relating to the operation of clinical trials. Recourse to qualified statistical advice, where possible, is strongly recommended. A "well designed" scheme may be scientifically correct but very difficult to operate. A statistician will be able to devise a sampling scheme that is not only valid but also capable of comfortable operation.

By way of example, here is one scheme used by the Tropical Research Centre in Ndola, Zambia, for assessing various indicators of the intensity of infection of *Schistosomiasis haematobium*. It is based on a technique suggested by Davis (*2*). Slides prepared by this method are wet when read and thus checking must be done soon after the original count is made. As counts are performed, results are entered in the technician's results book using a standard code. (Stenographers' spiral pads have been found useful here.) When 5 slides have been read, the chief technician, by means of a random number, selects one to be checked

[1] Research Fellow, Department of Medical Statistics and Epidemiology, London School of Hygiene and Tropical Medicine, Keppel Street (Gower Street), London WC1E 7HT, England.

(taking care to avoid seeing the results of the first count). The check counts are recorded in a separate results book and the collation of results is performed later.

The selection of the random number is performed according to a strict statistical procedure. Tables of random digits exist (*3,4*) and special dice can now be obtained to generate numbers easily. In the above scheme, 1 in 5 slides is checked. It was felt that by the time the technician had read more than 5 slides, the slide first read would have become too dry to be checked. For other diagnostic techniques, there may be other criteria that restrict the sampling fraction to certain values.

Special care is recommended for counts made at the beginning and end of examination periods. When a technician has finished counting with fewer than 5 slides, a random number is still drawn and the first few slides from the next examination period are checked to make up a full complement. If 3 slides remain and the random number is 2, the appropriate slide is checked on the same day. If, however, the random number is 5, then the second slide in the next period is checked. The temptation to ignore the few remainders should be avoided: the chance of an erroneous count in these samples may exceed that for all other samples.

RECORD KEEPING

It is important to recognize that there are two possible goals in using a statistical quality control scheme. For many, the graphical display of the results will be sufficient. This will enable administrators to decide whether the quality of microscopy is adequate and may discourage microscopists from producing poor performances (which are readily observed by all). In other instances, it may be convenient not only to use the display aspects but also to take advantage of statistical tools to detect poor microscopy. The graphical aspects of the technique may be termed the "empirical" phase, and the testing aspects the "statistical" phase. Researchers wishing to use statistical techniques must be aware of the empirical aspects, but those solely interested in descriptive aspects need not understand the statistical section.

In order to implement any of the procedures to be described, the investigator will be required to make at least 2 and at most 4 strictly defined decisions. It is convenient to group the first 2 (in order of presentation) and call them empirical decisions. These will be made regardless of the purpose of the application. The second 2 decisions might be described as statistical and are only required of investigators interested in performing statistical tests. Some of these decisions require the selection of parameter values. The term "scheme" will apply to the procedure as specified by one set of parameters, and the phrase "choice of scheme" should be taken to mean the choice of a set of parameters.

All laboratories have specified methods of recording and transcribing data from which epidemiological or therapeutic decisions are made. In order to implement a quality control scheme, no alteration in any on-going method is suggested, but the preparation of a special register for the tabulation of results of checking (not for recording all the new data) is outlined.

It is recommended that for each technician a separate book or section in a book be reserved. A binder is convenient for this purpose as the size can be increased as records accumulate and changes in staff will not involve much bookkeeping. It is further advocated that sheets be numbered as they are added to the master records.

For each slide checked, an entry is made in the register. An example is shown in Table 1 and each column is described below.[a]

Column 1. The date need not be noted on each entry.

Column 2. The microscopist's results are recorded. This value should not be known before the checking technician performs the second count. A standard code will facilitate later inspection of the results.

Column 3. The check count is entered using the same code as in column 2.

Column 4. The difference (i.e., column 2–column 3) is calculated and recorded. This is not directly required for the procedure to be described, but may prove useful for later assessment of the microscopist's performance.

Column 5. The accuracy of the microscopist's result is marked. The selection of what is accepted and what is rejected requires the *first empirical decision.* The basis for this decision depends upon the diagnostic method employed and the laboratory policy. It is strongly recommended that the senior technician, in collaboration with other interested staff, draw up a strict outline of which types of error are acceptable and which are not. In quantitative counts, for example, an error of, say, 5 may be unimportant if the true count is 500, but very important if the true count is 2. An undercount may be more dangerous than an overcount (particularly if a check shows evidence of infection in a slide originally read as uninfected). A prepared outline of the components of this decision will aid the supervising technician and will enable

[a] Columns 8 and 9, which concern the abbreviated procedure, are discussed on pages 317-319.

different laboratories to form a basis for comparing error rates of technicians. Lastly, by expressing a quantitative result in qualitative form, some information is lost, a point that is discussed later under "Final comments".

Column 6. For each slide checked that is deemed acceptable a score of −1 is entered. For each error, a positive value (b) (5 in this example) is assigned. As familiarity with the record keeping system is gained, it will become unnecessary to complete both columns 5 and 6.

The choice of the value for b will require the *second empirical decision*. The investigator wishing to initiate a quality control scheme will need to decide on the proportion of errors that he deems to be tolerable. Let this proportion be p_0. When a technician appears to be making errors (as judged by the proportion of checks being rejected) consistent with an underlying error rate equal to, or less than p_0, his overall performance is taken as acceptable. The value of p_0 must be set before the procedure is begun, in order to fix the value of b. The value of p_0 may be arbitrarily chosen, or a pilot study may be undertaken to determine a suitable value. (Different values of p_0 were tried in St Lucia and the value of 1/30 was chosen.)

Once p_0 is chosen, the value of b will depend upon whether or not the quality control scheme is used for descriptive or statistical purposes. For the latter case, the choice of b is discussed below under "Statistical aspects". When the sole intention of the exercise is to collect data and plot the results, the choice of b is readily made. If p_0 is an error rate of 1 in n, then $b = n-1$. For $p_0 = 1/30$, we have $b = 29$. The reason for this choice of b will be clearer after the next column is described.

Column 7. The cumulative score is calculated, being the sum of the previous total score and the current score. This "cuscore" is set at 0 when the scheme is initiated, or a new technician begins, and continues as long as the records are kept.

Table 1. Example of a register sheet for checks on the performance of a microscopist

Date (1)	Micro-scopist's count (2)	Checker's count (3)	Difference (2) − (3) (4)	Mark (5)	Score (6)	Cuscore (7)	Abbreviated procedure (see pp. 317-319) Abbreviated cuscore (8)	Abbreviated procedure (see pp. 317-319) No. wrong (9)	No. of checks (10)
18 May	1	1	0	✓	−1	−1		0	1
	110	105	5	✓	−1	−2		0	2
	15	15	0	✓	−1	−3		0	3
	0	1	−1	x	+5	2	5	1	4
	18	19	−1	✓	−1	1	4	1	5
	50	45	5	✓	−1	0	3	1	6
	0	0	0	✓	−1	−1	2	1	7
	8	8	0	✓	−1	−2	1	1	8
	10	10	0	✓	−1	−3	0	1	9
	150	160	−10	✓	−1	−4		1	10
19 May	30	28	2	✓	−1	−5		1	11
	90	85	5	✓	−1	−6		1	12
	0	0	0	✓	−1	−7		1	13
	340	520	−180	x	+5	−2	5	2	14
	15	15	0	✓	−1	−3	4	2	15
	3	3	0	✓	−1	−4	3	2	16
	120	10	110	x	+5	1	8	3	17
	110	115	−5	✓	−1	0	7	3	18
	60	70	−10	✓	−1	−1	6	3	19
	20	20	0	✓	−1	−2	5	3	20
	0	0	0	✓	−1	−3	4	3	21
	.	.	.	.	.	.	.	.	.
	.	.	.	.	.	.	.	.	.
	.	.	.	.	.	.	.	.	.

The choice of $b = 29$ for $p_0 = 1/30$, referred to above, should be clearer now. When a microscopist makes errors consistently more often than 1 in 30, then the cuscore will tend to remain positive. A good performance, with an error rate of 1/30 or less, will lead to near zero or negative cuscores.

Column 10. The cumulative number of checks performed is recorded continuously (regardless of the purpose for which the data are collected). As with the cuscore, it is wrong to start again at the beginning of a new year, or at the beginning of each new page. The values are tabulated from the initiation of the scheme or the start of a microscopist and always increase.

The use of a cuscore, as tabulated in column 7, will facilitate graphical presentation of results but does not directly provide the current number of errors made. If the cumulative score after check number m is given by S_m in column 7, then the number of errors made is

$$(S_m + m) / (b + 1)$$

and the error rate is

$$(S_m + m) / (m(b + 1)).$$

GRAPHICAL DISPLAY

The columns of figures, set out as in Table 1, facilitate the construction of graphs to summarize results. It is convenient to plot the cuscore from column 7 along the y axis. The example provided in Table 1 is plotted in Fig. 1. It is seen that incorrect counts result in large jumps up and correct counts in small jumps down. Poor performances will produce upward trends, better performances will result in downward slopes. The example displayed is hypothetical and has a large number of errors (3 in 21 checks) to demonstrate the nature of the path traced.

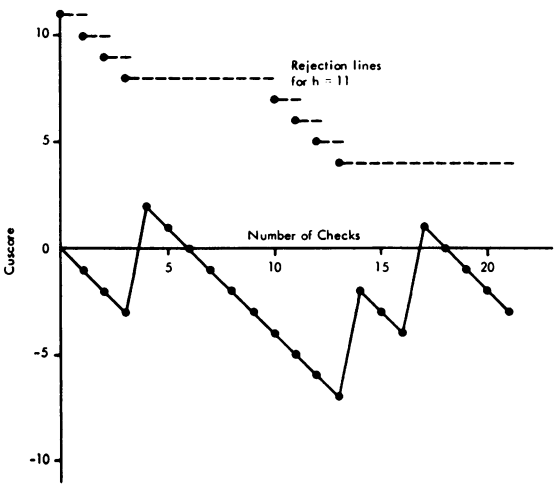

Fig. 1. Cuscore graph of example for which the register sheet is shown in Table 1 (Full procedure).

As records accumulate over a long period for established microscopists, it will be necessary to continue the graphs on further sheets. The horizontal line $y = 0$ will not be needed on all graphs. In plotting the results for a good technician, the cuscore will tend toward the lower right corner. In this case the particular negative value of y at the end of one sheet will have to be taken into account in deciding upon the upper value of y on the next page: this will also depend on the value of b used. Further comments on the choice of the scale are relevant to those wishing to employ statistical tests and are discussed under "Final comments".

STATISTICAL ASPECTS

It is emphasized that the principal benefits of this exercise are to be found in the disciplines of making the empirical decisions described above, in collecting the data, and in displaying the results. An administrator may wish to decide on a non-statistical basis whether or not the quality of microscopy is adequate. In similar situations in industrial settings, statistical tests have been studied that will detect poor performances. It is a simple matter to perform these tests, but the decision to do so will affect the procedure, even for empirical purposes, in that b is not taken as the obvious $n-1$ for $p_0 = 1/n$; rather, b is assigned an arbitrary value based on statistical properties.

In some circumstances, a laboratory may be interested purely in the statistical aspects and may wish to dispense with the graphical procedure. It is possible to effect a considerable saving in graphing, and the main register does not then need to be as complete as usual. This situation will be considered after the full approach has been described.

The full procedure

The second empirical decision described on page 315 involved the selection of a proportion of errors deemed to be tolerable for the laboratory. Although this is naturally thought of as an upper limit, it is quite possible that a microscopist may occasionally have an observed error rate greater than p_0 and yet, over a longer period, may still perform with an overall error rate that is acceptable. This is due to the random way in which errors occur: one does not expect every thirtieth count to be wrong—but 1 in 30 "on

average". Thus, while acceptable performances are reflected by cuscore charts sloping downwards, even an acceptable performance may occasionally slope upwards. It is the aim of a statistical test to differentiate between upward jumps due to random fluctuations in an acceptable performance and upward jumps due to an unacceptable performance. It is worth bearing in mind as well that any criterion will occasionally adjudge what is basically an acceptable performance to be unacceptable. (Unfortunately, it can be proved mathematically that for a given chart any criterion will eventually be exceeded for any non-zero value of p_0.) It is the aim of any well chosen scheme to minimize the chances of this occurring and maximize the chance of an unacceptable performance being quickly detected.

Instead of calculating the probability of rejection for schemes with different underlying error rates (denoted by p), a simpler procedure, based on the average run length, or "ARL", is used. The number of checks performed until a reasonable criterion is exceeded should depend on the value of p and may therefore vary. Of interest are the mean and the standard deviation of this distribution. The choice of scheme will be based on the ARL for various values of p: when $p = p_0$, a large ARL is desired and when p exceeds p_0 a low ARL is sought.

Page (*5*) considered procedures appropriate to these circumstances. The criterion on which a test is based is the difference between the current cuscore and the lowest previous cuscore. In Table 1, after 21 checks the cuscore was −3 and the lowest previous cuscore −7 giving a difference of 4. An upper limit, h, for this value is taken as the criterion. Any microscopist's performance where this difference equals or exceeds h is deemed unacceptable.

The test is readily applied to the graph of results if two parallel lines separated by h units (on the vertical scale) are etched on clear plastic. The lower line is placed on the lowest point of the cuscore graph. If the graph surpasses the upper parallel line to the right of this lowest point, the microscopist is deemed to be making errors more frequently than specified by p_0. Positions of the rejection lines for each check are given by the dashed lines in Fig. 1.

In order to obtain values of b and h for use in this procedure, the investigator is required to make two more decisions. Given that a laboratory has determined the upper level of errors (p_0), the ARL for the scheme at this value must be chosen. This is the *first statistical decision*. Naturally, one requires this ARL to be fairly large in order to avoid declaring acceptable performances "not acceptable" too often.

The *second statistical decision* can be made in one of two ways. In the simplest version, a value of p (say p_1) that is deemed as "not acceptable" is chosen and the ARL for this value of p_1 is specified. In Table 2, 3 sets of values of ARL are given for $b = 3$, 5, and 10 and for various values of h. For example, taking $p_0 = 1/30$ and arbitrarily setting the ARL for this value of p_0 at 2000, then to obtain an ARL of about 100 when $p_1 = 1/10$, it will be necessary to use $b = 5$ and $h = 11$, as in Table 1 and Fig. 1.

The second approach to this problem is as follows. Given that a value of p_0 and its associated ARL have been decided, then there are various combinations of b and h that can be used. Rather than deciding on an ARL at a given value of p_1, the investigator may compare the decreases in ARL as the probability of a mistaken reading increases. In general at higher values of b, the drop in ARL as p increases is steeper. For large values of b and h, however, there may be difficulties in the graphing of the cuscore.

Table 2 presents both the values of the ARL and the standard deviations of these distributions, when the SD exceeds the ARL by more than 1. (The standard deviations were obtained from equation 14 on page 368 of Ewan & Kemp, *6*.) The variance of the distribution is the sum of the squares of the ARL and another, possibly much smaller, factor. Thus standard deviations are equal to, or slightly larger than, the associated ARL. (As the run length is necessarily positive, the distribution is positively skewed and very large run lengths may occasionally occur.) Average run lengths exceeding 10 000 are not shown in Table 2.

As the choice of p_0 is often an arbitrary decision, then so too are the choices of b and h. In establishments seriously seeking to pursue a quality control scheme, a trial period of data collection will make it possible to decide on the most suitable values for these parameters.

An abbreviated procedure

Any laboratory wishing solely to perform the statistical procedures and not to use the full display features described above can employ a modified technique that greatly reduces the need for graphing (though all checks must still be performed). The outline for this work is given in an article by Ewan & Kemp (*6*).

In this method of presenting results, cuscores are kept only after errors have been made and only as long as they are not negative; an example of the abbreviated form of recording is given in columns 8 and 9 of Table 1. Notice that, once the full cuscore is dropped, the current proportion of errors cannot be obtained and a new column, the cumulative number wrong, has to be added.

The modification pertains mostly to the graphical presentation where one need only consider a chart of vertical height h units. Much less time has to be spent

Table 2. Average run lengths for various schemes[a]

Values of h	Values of p: 1/100	1/90	1/80	1/70	1/60	1/50	1/40	1/30	1/25	1/20	1/15	1/10
A. Table for b = 3												
4	3467	2820	2240	1727	1280	900	587	340	242	160	95	47
5	5101	4141	3281	2521	1861	1301	841	481	339	221	128	61
6	9816	7936	6256	4776	3496	2416	1535	855	590	374	208	92
7								4767	2799	1467	646	211
8								8491	4879	2479	1040	312
9										5143	1976	515
10											4701	949
11											8401	1496
12												2505 (2507)
13												4294
14												6977
B. Table for b = 5												
6	2140	1746	1392	1078	804	570	376	222	160	108	66	34
7	2632	2144	1706	1318	979	691	453	265	190	127	77	39
8	3436	2792	2215	1705	1262	885	575	332	235	155	92	45
9	4986	4038	3190	2443	1795	1248	800	452	316	204	117	55
10	9224	7413	5801	4389	3176	2162	1346	729	493	306	166	72
11						8892	4642	2029	1210	650	299 (301)	108 (110)
12							6461	2760	1618	849	377 (379)	128 (130)
13							9657	3998	2290	1163	493 (495)	156 (159)
14								6308	3490	1692	673	194 (197)
15									5828	2636	958 (960)	246 (249)
16										4258	1384 (1386)	311 (315)
17										5945	1827 (1829)	376 (381)
18										8603	2469 (2471)	460 (464)
19											3399 (3402)	565 (570)
20											4719 (4721)	694 (699)
C. Table for b = 10												
11	1146	941	757	592	448	323	219	134	100	70	45	25 (27)
12	1255	1030	827	646	487	350	236	144	106	74	47 (49)	26 (28)
13	1390	1139	912	711	535	383	257	155	114	79	50 (52)	27 (29)
14	1561	1276	1021	793	594	424	283	169 (171)	124	85	53 (55)	29 (31)
15	1784	1456	1161	900	672	477	315	187 (189)	136	92 (94)	57 (59)	30 (33)
16	2087	1699	1351	1043	775	547	358 (360)	210 (212)	151 (153)	102 (104)	62 (65)	32 (35)
17	2523	2047	1621	1245	919	643	417	241	171 (173)	114 (116)	68 (71)	35 (38)
18	3203	2587	2037	1554	1137	787	502 (504)	283 (285)	198 (201)	129 (132)	76 (79)	38 (42)
19	4412	3536	2759	2081	1501	1020	636	347 (349)	238 (241)	152 (155)	87 (90)	42 (46)
20	7160	5644	4319	3183	2234	1467	876 (878)	453 (455)	300 (303)	184 (188)	101 (105)	47 (52)
21				6934	4460	2662 (2664)	1432 (1434)	660 (662)	411 (414)	236 (240)	122 (126)	53 (59)
22				8152	5212	3087	1641 (1643)	744 (746)	458 (461)	259 (264)	131 (136)	56 (62)

[a] Values of the ARL exceeding 10 000 have been left blank. The standard deviation of the run distribution (which is not less than the ARL) is shown in parentheses below the ARL if it exceeds the ARL by more than one.

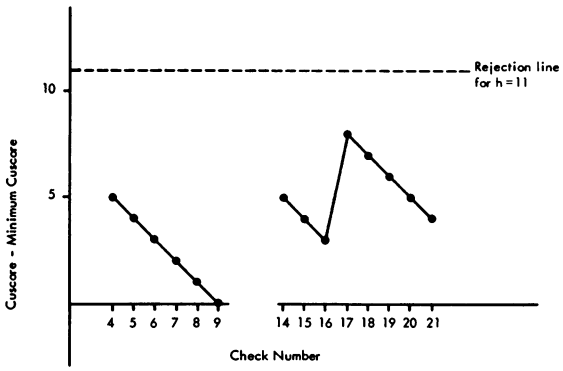

Fig. 2. Cuscore graph of example for which the register sheet is shown in Table 1 (Abbreviated procedure).

in graphing as only portions of the full charts are displayed. Obviously, the impact of a long decreasing graph is lost when this approach is used (Fig. 2).

Final comments

Under "Graphical display", the position of the *y* axis on continuation sheets was briefly discussed. When a statistical approach is undertaken, this question has a ready solution. For the full presentation, the maximum value of the cuscore axis should be the current minimum cuscore plus the value of the criterion in *h* units. The situation for an abbreviated statistical display is even simpler: only the values from 0 to *h* inclusive are required.

Page (*5*) presents a table from which values of *b* and *h* can be obtained, his calculations being based on an approximate formula given by Burman (*7*). The tables shown here are based on Burman's paper rather than Page's as with high-speed computers readily programmed, it was just as simple to obtain exact values as are shown (after rounding) in Table 2.

Lastly, it was noted earlier that using a qualitative result for a quantitative measure was inefficient. This is undoubtedly true, but the loss in efficiency is offset by added robustness of the procedure. To have used a quantitative measure would have required a knowledge of the variance of this measure. This will depend on many factors and will vary from disease to disease and technique to technique. Given that the variance is known, if it does follow a complicated pattern then statistical methods would have to be adapted to allow for this. The technique advocated avoids these problems and, though not efficient, has wider application.

ACKNOWLEDGEMENTS

I am indebted to Dr A. Davis, Parasitic Diseases Programme, WHO, for his interest and encouragement in the applications of these techniques to medical situations. Both Professor P. Armitage and Professor M. J. R. Healy read early versions of this script and made several helpful suggestions for which I am most thankful.

RÉSUMÉ

MÉTHODE STATISTIQUE DE CONTRÔLE DE LA QUALITÉ DES TRAVAUX DE LABORATOIRE CLINIQUE

Dans l'article ci-dessus, l'auteur décrit une méthode permettant de contrôler la qualité des examens au microscope effectués par des techniciens de laboratoire clinique. Pour un surplus minime de temps et d'argent, il est possible de dresser un bilan précis des travaux qui met clairement en évidence, pour chaque technicien, la qualité de ses prestations récentes et celle des efforts fournis au cours de toute la période étudiée.

Il convient de mettre au point très soigneusement la procédure qui sera appliquée à cette fin de manière à en assurer la valeur scientifique sans compliquer son exécution. Pour illustrer les points à prendre en considération, l'auteur donne en exemple la méthode de contrôle appliquée par un laboratoire particulier. Il est bien évident que chaque technique diagnostique a ses exigences, et il est donc recommandé de solliciter l'avis d'un statisticien pour qu'il mette au point une méthode de contrôle appropriée.

Une fois faites les vérifications des travaux selon le schéma prévu, il n'est pas difficile d'en réunir les résultats et de les traduire en tableaux et registres éloquents au moyen d'un simple procédé arithmétique. Le mode d'établissement de la "note cumulative" (désignée en anglais par le terme abrégé "*cuscore*") est décrit, et un exemple de présentation graphique de cette évaluation est donné. Le tableau correspondant offre un témoignage précis de l'efficacité du technicien pour chaque série d'opérations considérée, tout en permettant d'apprécier cette efficacité pour toute la période précédente (tableau 1). A une efficacité médiocre correspond une courbe ascendante tandis qu'une courbe descendante traduit un bon rendement. Ces courbes permettent au personnel administratif d'apprécier aisément les résultats du travail du personnel technique de laboratoire. La méthode présente aussi l'avantage de susciter une émulation parmi les techniciens, dans la mesure où ils se montrent

soucieux d'obtenir la courbe la plus satisfaisante ou à tout le moins d'éviter un classement en queue de peloton.

Une méthode statistique destinée à faciliter le dépistage rapide d'un technicien dont le travail doit être considéré comme inefficace est décrite. Elle est facile à appliquer à partir du tableau des "notes cumulatives". Malheureusement, toute courbe exprimant l'ensemble des "notes cumulatives" peut à un moment donné être déclarée inacceptable, même lorsque l'efficacité du technicien en cause est généralement acceptable. Le nombre des contrôles à opérer avant que le rejet soit décidé afin de pallier cet inconvénient dans une certaine mesure est appelé ARL (*Average Run Length*). Les différentes valeurs d'ARL—avec les écarts types correspondants, qui sont indiqués entre parenthèses—sont données dans le tableau 2 pour plusieurs systèmes de contrôle de la qualité. Ceci devrait permettre aux chercheurs qui souhaitent améliorer un système de contrôle de vérifier les valeurs d'ARL pour les taux d'erreur acceptables et de voir comment ces valeurs décroissent lorsque les taux d'erreur augmentent.

REFERENCES

1. BARTHOLOMEW, R. W. & GODDARD, M. J. Quality control in laboratory investigations on *Schistosomiasis mansoni* on St Lucia, West Indies: a staff assessment scheme. *Bulletin of the World Health Organization,* **56**: 309-312 (1978).
2. DAVIS, A. Comparative trials of antimonial drugs in urinary schistosomiasis. *Bulletin of the World Health Organization,* **38**: 197-199 (1968).
3. FISHER, R. A. & YATES, F. *Statistical tables for biological, agricultural and medical research.* Oliver and Boyd, 1963.
4. Rand Corporation. *A million random digits with 100 000 normal deviates.* New York, The Free Press, 1955.
5. PAGE, E. S. Continuous inspection schemes. *Biometrika,* **41**: 101-115 (1954).
6. EWAN, W. D. & KEMP, K. W. Sampling inspection of continuous processes with no autocorrelation between successive results. *Biometrika,* **47**: 363-380 (1960).
7. BURMAN, J. P. Sequential sampling formulae for a binomial population. *Journal of the Royal Statistical Society,* **8** (Suppl.): 98-103 (1946).