An approach to quality and performance control in a computer-assisted clinical chemistry laboratory

P. E. UNDRILL AND S. C. FRAZER

From the Department of Chemical Pathology, University of Aberdeen, Foresterhill, Aberdeen AB9 2ZD, UK

SUMMARY A locally developed, computer-based clinical chemistry laboratory system has been in operation since 1970. This utilises a Digital Equipment Co Ltd PDP 12 and an interconnected PDP 8/F computer. Details are presented of the performance and quality control techniques incorporated into the system. Laboratory performance is assessed through analysis of results from fixed-level control sera as well as from cumulative sum methods. At a simple level the presentation may be considered purely indicative, while at a more sophisticated level statistical concepts have been introduced to aid the laboratory controller in decision-making processes.

The analysis of patient samples in the routine laboratory serving the Aberdeen Royal Infirmary and other hospitals has, for several years, been assisted by a PDP 12 computer (Undrill, 1974). A central concept of this system is the use of visual displays of various types to disseminate information pertinent to laboratory status at any given time. An interconnected PDP 8/F aids the construction of work schedules via a multidisplay network (Undrill and Gibson, 1978) and allows the analysis of numerical data passed to it from time to time during a working session of the laboratory. In terms of quality control functions, these data are generally analysed interactively using a visual display with graphical capabilities. The overall computer system is shown in Figure 1.

The laboratory employs two widely used controlling techniques:
(1) Fixed-level bovine sera pools, usually at two distinct levels in the analytical range, are inserted alternately every tenth sample at a point midway between successive pairs of drift standards used to monitor the stability of the analytical equipment.
(2) A cumulative sum is calculated as the additive sum of the excess of the mean of patient sample assays over a known expected normal value.

The analysis protocols may be separated into those applying to: (i) immediate (real-time) analyses, which have to be rigidly formalised; and (ii) retrospective analyses, in which time-consuming investigations that give additional information can be incorporated to yield more widely based comparisons.

Methods and their presentation

REAL TIME The information available to the laboratory concerning the performance of its Auto-Analyzer systems is derived from retransmitting potentiometers on each chart recorder and is shown on the visual display units, as in Figure 2. The first of these is a dynamically refreshed point-plot display relayed by closed-circuit television to several points within the laboratory, and the second is a conventional character display. The former has part of its display area reserved for a representation of the signals received by the computer from any selected analytical channel. The display channel may be altered either from the control terminal or from one of the several push-button control stations situated throughout the laboratory. The incoming signals may be retrospectively displayed for up to 5 minutes, and the display is ongoing with a repeat time of about 10 minutes. The laboratory status section is furnished with indicators which show the following conditions: (i) If a channel is switched into the computer. (ii) If a method is recognised by the computer as in operation by reason of the occurrence of a recognisable signal peak. (iii) The result and sequence identifier (that is, turntable position) for the most recently handled sample. (iv) Any informative error condition. (There is a priority system dictating the hierarchy for error erasure.) (v) A figure giving the excess of the current method mean over a predetermined value. This is presented to two decimal places and is calculated from results lying within accepted 'normal' limits for each method. For certain methods, for example, glucose...
An approach to quality and performance control in a computer-assisted clinical chemistry laboratory 101

Fig. 1 A computer system for clinical chemistry based upon a DEC PDP 12 and PDP 8/F interconnected processor configuration.

Fig. 2a Laboratory status display.

Fig. 2b Quality control tabulated display.
and cholesterol, this concept of normality is not considered rigid enough for the calculation to be applied. (vi) An alternative display format allowing the operator to view any portion of computer main memory. This has been found to be of exceptional value in development work and fault-finding situations.

The importance of item (v) is that it may give some indication of method performance, under the constraint that sufficient values have been registered. Our practice is to consider 50 or more results as sufficient for this purpose. The parameter is not wholly dependent on any direct laboratory or external calibration. Its use, on a cumulative basis, by forming an arithmetic sum from session to session, the csum, has been postulated as a method of internal quality control (Wootton, 1968), and its analysis will be pursued at greater length in another section.

Positive correlation between equipment performance and the mean of the 'normal' range of patient results has been reported by Whitehead et al. (1968), Talley (1969), and Taylor and Carter (1973), although strong doubts as to its value for this purpose have been expressed by Kilgariff and Owen (1968) and Reed (1970).

In a study (Undrill, 1974) of the short-term value of this parameter as a method of positive control, that is, at a within-session level compared with fixed-level methods, little significant correlation was found. Assuming Gaussian statistics and a zero correlation hypothesis, only in two cases out of 14 was there any reason to doubt this hypothesis. For these two cases, the probability of the non-obeyance of this hypothesis was found to be not greater than 74% and 68% respectively.

Simulation studies by Reed (1970) had also indicated the sensitivity of the method to distribution changes. Techniques based on $\chi^2$ calculations presented by Gindler et al. (1971) or those related to analysis of variance (Amenta, 1968) are less sensitive to distribution parameters and should be more successful for short-term detection of analytical malfunction.

The divergence of opinion as to the usefulness of this technique is likely to be dependent on the degree of distribution invariance in the data analysed by the various authors. For this reason, and because of the short period of our own study (two months), it was decided to retain the feature in the laboratory information regime and to continue its analysis on a much longer time scale.

The second display tabulates the results pertaining to the 'known value' quality control samples. Using the complete display area available, a synopsis of 16 analytical methods is provided. With such controls inserted after every nine patient samples, a summary of the performance covering an analytical run length of 200 is provided for each method. This is normally sufficient for the twice daily working sessions.

At an individual result level, three levels of potential error indication can be appended to a result produced by the computer system: (i) an identified input signal fault, relating to noise, shape, or time of peak occurrence; (ii) a fault in the recalibration drift standards, either as above or as per the discrimination rules associated with the drift correction algorithm; and (iii) a fault in the initial calibration standards, either as a signal fault or as an error in the parameters of the calibration curve (for example, from a standard sample inserted out of sequence).

**POST ANALYSIS**

Numerical data are passed from time to time to the PDP 8/F via a data buffer. The transfer of any item is verified by passing it back to the source computer, the next item being transmitted only in the event of a true match. The overall data transfer rate is around 5K baud. Each production of a batch of results for reporting also releases those data for future transmission to the PDP 8/F system. Up to 700 results can be stored in the PDP 12 before transmission becomes necessary, and the two machines are linked through the transfer software for about 5 seconds in order to swap the data.

The quality control assessment is in two major sections, being applied either to the fixed-level quality control standards or to cumulative sum data. The storage of data is arranged as in Fig. 3, the primary storage level for each method being an array of 1680 results which is effectively filled so that it contains the most recent 1680 transferred results. Similar arrays exist for the most recent 252 values for high and low controls, and cumulative sum. A single record is set aside for a cumulative histogram of result distribution used to provide information on the setting of 'normal' limits. Sundry individual records are allocated to temporary work files, intermediate statistic, and fixed data files. Computer control sequences allow the detailed analysis of these data arrays associated with the 16 analytical methods.

**FIXED-LEVEL CONTROLS**

The basic form of display is as in Fig. 4a, showing the most recent 216 points on a graph having the assayed value as its ordinate and reporting dates as the abscissa. The initial display contains: (i) limit lines set at levels determined by the analyst; (ii) the mean and number of control sample results transferred in the most recent analytical session; and (iii) the mean and standard deviation over a preset
An approach to quality and performance control in a computer-assisted clinical chemistry laboratory

**Fig. 3** Data file structure.

### Input data (16 methods)

<table>
<thead>
<tr>
<th>Input data</th>
<th>Fixed data</th>
<th>Intermediate data</th>
<th>Cumulative data</th>
<th>Overall quality control file</th>
</tr>
</thead>
</table>

**Fig. 4a** Primary display of fixed level control.

**Fig. 4b** Alteration of window position and width.
window applied to results both within and without the laboratory limits.

Interactive processing begins with the ability to move the analysis window over the extent of the graph, or to alter the window dimension stepwise, each step corresponding to one analytical session. The new means and standard deviations are calculated as appropriate with each change of window disposition (Fig. 4b).

If \( m_1, \sigma_1, n_1, \) and \( m_2, \sigma_2, n_2 \), are the mean, variance, and number of results in the reference and final window position, then a Student's \( t \) statistic can be calculated as:

\[
t = \frac{|m_1 - m_2|}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}
\]

This value and its corresponding level of statistical significance, associated with a nil difference hypothesis, are registered on the display for both those values within and without controlling limits.

The above analysis may be considered as applying, in an overall manner, to limited subsets of displayed data. An alternative approach is to apply discriminant analysis to data points bounded by the window. Discriminatory methods may be subdivided into two classes (Westgard et al., 1977): (i) those techniques for which the probability of false rejection, \( P_{fr} \), increases as the number of points considered, that is, the window length, increases; and (ii) those techniques for which \( P_{fr} \) stays constant with varying window length.

The first of these criteria is commonly used to test an individual measurement against a previously determined distribution, and the second is more suitable when small numbers of results are grouped together and used as a single control observation. Each form of analysis can be applied to the windowed points, and the criteria adopted were chosen from a comprehensive list subjected to simulation studies by Westgard et al. (1977). These are given in the Table.

Having positioned the window, the class of discriminant feature is selected, with the result that true or false indicators are displayed for each of the discriminants. Limit lines are drawn across the data at either (a) 1, 2, and 3 standard deviations from the parent population mean, or (b) appropriate values for tests 1 and 2 of class 2 (Table) computed from the number of values in the group and the level of \( P_{fr} \). The latter may be selected from 5\%, 1\%, or 0.2\%.

In Fig. 5a a class 1 analysis is shown, points lying outside 2.5\( \sigma \) being arbitrarily excluded by vertical editing. These points are shown by smaller display points. Figure 5b shows a class 2 analysis, a complete period of one week being removed from the parent distribution by horizontal editing to meet a temporary local requirement. As seen, each set of results of these discriminant features, applied to the test window, is shown in the top left-hand corner of the display with the test's numerical bandwidth.

### CUMULATIVE SUM

Notwithstanding earlier comments relating to the potency of the cusum method, its value as a long-term indicator of analytical function change, or of distribution change, can be investigated by graphical techniques well suited to interactive interrogation. Also, any trend towards larger analytical runs may lead to a greater usefulness of the cusum as a short-term indicator.

Basically, the underlying assumption is that, for certain methods, the 'normal' patient distribution is invariant, and deviations of a calculated mean are attributable to equipment variations. Arithmetically, the process may be described as formulating a running sum, \( C_n \), where

\[
C_n = \sum_{j=1}^N c_j \quad \text{and} \quad c_n = \frac{1}{M} \sum_{i=1}^M r_i - R_m.
\]

\( \sum \) and \( \sum \) refer to summations between and within analytical sessions, respectively, with \( M \) values in a session. \( R_m \) is the appropriate method distribution mean and \( R_{iL} \leq r_i \leq R_{iH} \), \( R_{iH} \) and \( R_{iL} \) being truncation limits relating to 'normal' values for the method. \( N \) values are utilised when \( N \leq M \).

There appear to be two distinct levels at which to analyse the cusum curve: first, examining sections of the curve in relation to other sections, and, secondly, investigating the performance of the most recently generated curve section. The former method may be called slope analysis.

Sections of the curve are identified, interactively, with markers, as in Fig. 6a, and from the bounded

---

**Table** Discriminant analysis criteria

<table>
<thead>
<tr>
<th>Class</th>
<th>Discriminant analysis criteria</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1 value outside ( m \pm 3\sigma )</td>
</tr>
<tr>
<td></td>
<td>2 successive values unilaterally outside ( m \pm 2\sigma )</td>
</tr>
<tr>
<td></td>
<td>4 successive values unilaterally outside ( m \pm \sigma )</td>
</tr>
<tr>
<td></td>
<td>10 successive values unilaterally biased about ( m )</td>
</tr>
<tr>
<td>2</td>
<td>1 value outside appropriate limit</td>
</tr>
<tr>
<td></td>
<td>2 successive values unilaterally outside appropriate limit</td>
</tr>
<tr>
<td></td>
<td>Group mean outside appropriate limit of deviation from population mean</td>
</tr>
<tr>
<td></td>
<td>Group range exceeds acceptable limit</td>
</tr>
</tbody>
</table>

Class 1 Tests on individual controls or subgroups of controls treated as individual measurements. \( m \) and \( \sigma \) are the mean and standard deviation of the parent distribution. Probability of false rejection increases with number of controls in windowed group.

Class 2 Tests on subgroups as distributions, the parameters of which are compared with that of the parent distribution. Probability of false rejection is constant with subgroup dimension.
points a best straight line spline section is computed. For the \( k \)th sector the values \( a_k \) and \( b_k \) are calculated for the relation \( y = a_k x + b_k \), and from the distribution of points about this ideal line the standard error of estimate \( S_{ek} \) is found,

\[
S_{ek} = \sqrt{\frac{\sum_{j}(y_j - \hat{y}_j)^2}{N_k - 2}}
\]

where \( y_j \) and \( \hat{y}_j \) are the expected and observed values of \( y \), and \( \sum_j \) extends over the \( N_k \) of the \( k \)th sector points. An estimate of the standard error of \( a_k \), \( \sigma_{a_k} \), is given by

\[
\sigma_{a_k} = \frac{S_{ek}}{\sum_j x_j^2 - N_k \bar{x}_k^2}
\]

where \( \bar{x}_k \) is the mean of \( x \) values in the \( k \)th sector, and from this a 't' statistic can be calculated which relates the difference of the slopes of the two sections to the standard error of their slopes:

\[
t = \frac{|a_1 - a_2|}{\sigma_{a_k} \times \sigma_{a_2}}
\]

Dependent on the degrees of freedom in each section, the probability value for the slope of the two sections being identical is displayed. Any significant difference will indicate an underlying change in the constructional parameters, which may be of analytical origin.

Under ideal conditions, the trend of \( C_s \) with time should be linear and, if \( R_m \) has been chosen correctly, parallel with the abscissa. Although any straight line will indicate a static situation, a gradient will tend to mask day-to-day variations dependent on the ordinate sensitivity and the visual acuity and inferential capabilities of the observer. In Fig. 6b the systematic day-to-day error in \( R_m \) is removed by normalising the cusum curve. This can be carried out for any chosen period along the time axis. This normalised curve produces information on a more applicable \( R_m \) over the period of interest.
To analyse the curve under the second concept, use is made of a V-shaped mask, which may be called a sector of surety. This mask is moved point by point along the curve, starting from the most recently generated end. The principle of the mask is described by Belz (1973) and shown in Figure 7a.

PRQ is a V-shaped mask which is placed on the cusum curve AA' such that its semi-axis BOB' intersects the curve at some point O. The region of interest is to the right-hand side of O, that is, in the direction OX. If the cusum line crosses out of the minor enclosure of the mask, loss of control is likely to have occurred at O. It is clear that the shape of the mask is dependent on the method parameters and the level of statistical inference selected.

It can be shown, to a good approximation, that

$$|OR| = \frac{2 \sigma \chi^2}{\delta^2} \cdot \log_e \left(\frac{\alpha}{2}\right) \text{ and } |\tan \theta| = \frac{\delta}{2k}$$

where $\alpha$ is the false positive acceptance factor associated with the mask, $\sigma$ is the standard error on the mean of data associated with each point of the cusum line, and $\delta$ is the actual difference that is being used as the control index. In this case $\delta$ is taken as a deviation from zero.

For the second expression, which describes the semi-angle, $\theta$, of the sector, $K$ is a normalisation factor relating unit ordinate to unit abscissa.

In the implementation of the method, statistical considerations show that $|OX|$ should not exceed $3 \times |OR|$, otherwise the validity of the rejection
An approach to quality and performance control in a computer-assisted clinical chemistry laboratory

Fig. 7a  Sector of surety mask.

becomes questionable. To compute the above parameters, $\sigma^2$ can be derived from the overall standard deviation for the method and the mean number of samples per session. A level of $\alpha$ that may be acceptable is 5%, and $\delta$ can be obtained from the fixed-level controls by postulating an acceptable variation from day to day. A useful value for $\delta$ is 50% of the mean range that is normally attributable to the fixed-level controls of an individual method. Computer interpretation involves placing a marker along BOB', and as this is moved point by point over the cusum curve, the sector of surety is erased and redrawn at each move. Observation will show those points outwith the sector and the appropriate point of divergence on the abscissa, as in Figure 7b. The magnitudes of $\theta$ and OR are seen to change radically between methods and mirror the greater stability of some methods.

In general, it has been found by Belz (1973) that the cusum method is more effective than the Shewhart (1931) limit procedures in recognising changes in control level, at least when the changes are small, that is, of the order of $\frac{1}{3} \sigma^2$. The average run length before detection of such changes can be reduced by a factor of 3 and this suggests its use as a control measure.
Discussion

A comprehensive system of quality controlling techniques has been described and applied to a service clinical chemistry laboratory. These relate to real-time, fixed-level methods and cumulative sum calculations. Retrospective analysis techniques are separated into those which attempt to detect the long-term and the more immediate changes. These are adapted and utilised as shown to aid the laboratory control staff. They are implemented on a dual processor laboratory computer system at Aberdeen.

The real-time sequences are built in the machine language protocols of the laboratory monitoring computer, in this case a mixture of PDP 8 and LINC assembly code. Retrospective data analysis is programmed in standard FORTRAN IV as supported by the DEC OS/8 operating system, modified where necessary to take advantage of specific device functions. The modifications are in the form of simple machine language subroutines, the use of which is allowed in most FORTRAN IV implementations. Therefore they are transferable, for the most part, to any small computer with suitable peripherals, irrespective of whether or not the data acquisition process is computer-based. The precise form of visualisation will depend on the equipment at hand.

All of the retrospective techniques are suitable for interactive interrogation; however, a simplified subset can be adapted to provide permanent copy for record purposes. A computer system is not a necessary prerequisite for these analyses, but its existence saves much labour and so the probability of utilisation of advanced quality control techniques, given the normal pressures of work on laboratory staff.

References


Requests for reprints to: Dr P. E. Undrill, Department of Chemical Pathology, University of Aberdeen, University Medical Buildings, Foresterhill, Aberdeen AB9 2ZD, UK.