J. clin. Path. (1969), **22**, 278-284

Recommended scheme for the evaluation of instruments for automatic analysis in the clinical biochemistry laboratory[1]

Several new instruments for automatic analysis are now becoming commercially available and may be of value to National Health Service laboratories. Although manufacturers may have arranged an evaluation of these in the prototype stage it will be necessary to make a complete and independent assessment of production models. This schedule outlines the points which should be considered in this assessment. It has been written with currently available analytical systems in mind, and may need modifying if it is to form the basis of evaluation of instruments which will be introduced in the future. Some of the tests described may not be applicable to all machines but are nevertheless listed for completeness.

A technical evaluation of the mechanical and electronic components of instruments of this type is beyond the scope of most clinical laboratories. Work of this type should be delegated to some more specialized establishment and will not be considered in detail here. The first section of this schedule lists the information which is required in order to provide a basis for comparison between the different systems of analysis. This is followed by a description of the tests to be performed on the machine by the laboratory making the assessment. In order to provide a ready basis for comparison between instruments, it is essential that these tests are performed in the manner described and the report of the assessment must provide clear factual answers to specific questions; wherever possible these should be expressed in tabular form.

Since the terms used in automatic analysis may not be familiar to all readers a glossary is given at the end of this schedule and many of the definitions are further discussed in the text. In particular, it has been found necessary to assign specific meanings to the words 'specimen' and 'sample'.

[1]The scheme was prepared by a working party of the Laboratory Equipment and Methods Advisory Group consisting of P. M. G. Broughton, M. A. Buttolph, A. H. Gowenlock, D. W. Neill, and R. G. Skentelbery. This report was received for publication on 1 March 1969.

MANUFACTURERS' DATA

Manufacturers should be asked to provide factual data about the specification of their instruments. This information should be recorded in a standard form to provide specific factual information about the instrument, as below:

1 GENERAL (*a*) What is the capital cost of the machine?

(*b*) What basic facilities are included and what facilities can be provided as optional extras?

(*c*) How many channels can be provided?

(*d*) What is the throughput time for each recommended analytical method?

(*e*) What is the sampling rate for each recommended analytical method?

(*f*) Are any claims made for the accuracy, precision, etc, of the system? If so, give details.

(*g*) What analyses have been performed on the instrument and are instructions provided for carrying them out?

(*h*) Is the instrument designed to work solely with methods which do not involve the removal of protein?

(*i*) What standards are provided or recommended for use with the instrument?

(*j*) Can the instrument be modified to permit different analyses to be performed on it? If so, how long does this take?

(*k*) What space and service does the instrument require?—floor area, bench space, shelves, drainage, electric power, water supply, gas, etc. What are the dimensions and weight of the machine?

(*l*) Will users of the machine require special training to operate the instrument and maintain it? If so, is any provision made for this?

2 SPECIMENS (*a*) What type of specimen is required?—whole blood, separated plasma, serum or centrifuged blood with sampling from the supernatant plasma or serum.

(*b*) What other materials can be analysed? *eg*, urine, cerebrospinal fluid, etc.

(*c*) What is the size of a machine batch?

(*d*) What are the limitations on specimen volume?

(*e*) What is the range (if any) of sample volume?

(*f*) How can the sample volume be changed and recalibrated?

(*g*) How does the sampler function and what precautions are needed in its use?

(*h*) Is there any alarm system to detect blockage of the sampler by clots?

3 REAGENTS (*a*) How many reagent dispensers can be accommodated in each channel and of what volumes?

(*b*) What range of reagent volumes can be dispensed? Are the volumes fixed or variable? How is the volume of each dispenser changed, and how is it indicated?

(*c*) Within what time intervals after sampling can reagents be added?

(*d*) How is mixing of the sample with reagents achieved?

(*e*) Will the tubing, and other components exposed to reagents, withstand strong acids and alkalis? What reagents and solvents are specifically prohibited?

4 COLORIMETER (*a*) Is this a single or double beam instrument? What is the wavelength range, the band width, and the method of wavelength selection?

(*b*) Is a single cuvette or a number of matched ones provided? Of what is the cuvette made? What is its volume, and what volume of solution (including any used for rinsing) is needed to make each measurement? Can bubbles of air enter the cuvette, and, if so, what effect will they have on the measurement? What is the light-path length of the cuvette and are cuvettes with other path lengths available? If so, give details.

(*c*) How is the output signal timed in relation to cuvette filling and emptying? Is this instantaneous or integrated over a stated period?

(*d*) Can the time interval between the addition of the last reagent and the reading of the colour be varied, and, if so, within what limits?

(*e*) What is the form of output, in what units, and are there facilities for other forms? How many significant figures are given in the output?

(*f*) If the output is not in concentration units, how is concentration calculated?

(*g*) Over what range of concentration are results presented for each method?

(*h*) In multichannel instruments how are the outputs of the channels coordinated and presented?

(*i*) Does the method of standardization depend on a two-point setting or on a calibration curve?

(*j*) How frequently should the system be re-standardized and the reagent blank reset?

(*k*) What provision is there for correction for non-linearity in the calibration curve?

(*l*) What provision is made for blank determinations? Can the blank reading be automatically subtracted from that of the test?

5 PERIPHERAL FACILITIES Are the following facilities included? If so, give full details of their specifications: (*a*) provision for heating or incubation, (*b*) flame photometer, (*c*) fluorimeter, (*d*) other detector systems (*eg*, ultraviolet spectroscopy, atomic absorption), and (*e*) automatic specimen identification.

EVALUATION OF INSTRUMENTS IN THE CLINICAL LABORATORY

1 GENERAL This section concerns assessment of the performance of instruments in the environment for which they are intended, namely, the clinical laboratory. It defines certain machine characteristics and lays down standard methods for their determination applicable to all instruments for automated biochemical analysis. The basis of all instrument assessment must be comparative, in terms of costs, facilities, and performance, and all data must be expressed in a form which permits ready comparison of discrete systems with each other, with continuous flow methods and conventional methods.

Performance is to be assessed on the instrument as installed and operated as prescribed by the manufacturer or his agent. No modification should be made to the instrument or to the analytical methods prescribed. The analytical methods should be those suggested by the manufacturer, but these should, as far as possible, be the same for all discrete systems tested. The adequacy of performance can only be judged in comparison with the alternative methods of analysis which are available. Performance data must therefore be obtained under standardized, well defined, and reproducible conditions.

As many different determinations as possible should be tried on the instrument, but results should be obtained using one at least of each of the three following classes of test: (*a*) a simple method, using only the basic facilities of the instrument, *eg*, total protein, and sodium and potassium where a flame photometer is supplied; (*b*) a more complex method, using several reagents and facilities, *eg*, urea by the Berthelot reaction; (*c*) a complex method using as

many facilities as possible, including controlled temperature and time, *eg*, an enzyme analysis.

2 ASSESSMENT CRITERIA The ultimate aim of any analytical system is to enable the operator to produce as rapidly as possible results which are precise and accurate at an economic cost. Details of precision and accuracy tests are given below, as are methods for determining the related parameters, carryover, and sample contamination. First, however, a simple check on the performance of the system is carried out.

3 OVERALL PERFORMANCE (*a*) When some experience has been gained in the use of the system select one analytical method from each of the three groups (*a*), (*b*), and (*c*) described in paragraph 1 of this section.

(*b*) Set up and standardize the system according to the manufacturer's instructions.

(*c*) Analyse in succession 20 specimens, or one machine batch, whichever is the less, of a serum for which the concentration of the constituent being determined is known. This may be pooled or commercial serum, of human or animal origin, but the concentration of the constituent under test should be near the middle of the range of the analytical method.

Inspection of the results should reveal any serious problems such as drift or gross lack of precision or accuracy.

This test is carried out for the benefit of the tester, so that time and effort are not expended on a system which gives an obviously unacceptable performance. It is not necessary to report the results of this test if they are satisfactory.

4 CARRYOVER AND SAMPLE CONTAMINATION Sometimes referred to as interaction, carryover is a factor which has become of increasing importance with the introduction of mechanized systems for analysis. It concerns the influence of the concentration of the test substance in one sample upon the result obtained for the following sample, and is defined as follows. If a sample A is followed by a sample B and the recorded results for A and B are a and b, and the true values are a′ and b′ respectively, then the carryover, K, is the difference between the true and apparent values of B (b − b′) divided by the difference between the recorded value for A and the true value for B (a − b′)

$$\text{thus } K = \frac{b - b'}{a - b'}$$

In an automated analytical system, carryover can occur where solutions derived from different samples take a common path, for example, the flowcell of a colorimeter. There may be a number of such common paths in a particular system, each giving rise to a certain amount of carryover. For the purpose of evaluation it is only necessary to measure the total carryover in the system, and directions for doing this are given below. For design or research purposes it may be desirable to determine the carryover at a particular stage in the system. In this case, a technique similar to that described for the measurement of cross contamination (paragraph 6) can be used together with the equation given above.

Carryover depends not only on the design of the analytical system, but also on the physical characteristics of the solutions involved, and will thus have a different value for each analytical method.

Whereas carryover arises when a single determination is carried out on a series of different specimens, cross contamination is of importance only when the same specimens are sampled a number of times.

Systems based on the continuous flow principle can analyse a specimen for several constituents simultaneously, the analyses being carried out 'in parallel', and only one sample being removed from each sample cup. Using a discrete system, the same analyses would be carried out sequentially, the analyses being performed 'in series', and a number of samples being removed from each sample cup. Each time the probe enters a cup to remove a sample, part of the preceding specimen, adhering to the probe, is deposited in the cup. The volume of specimen thus deposited is called the 'cross contamination'.

The same sampling procedure is used in most discrete systems. The sample is drawn into the sampler probe, and is then ejected into a mixture tube, together with a fixed volume of diluent solution. Any diluent solution adhering to the probe when it enters the next sample cup may be deposited therein. The volume of diluent thus deposited is called the 'sample diluent contamination'.

5 DETERMINATION OF TOTAL CARRYOVER This should be determined for each method by analysing alternating groups of three specimens of serum. A serum (A) with a high value for the determination should be used for the first three specimens and a serum (B) with a low value for the next three as follows:

$$A_1\ A_2\ A_3\ B_1\ B_2\ B_3$$

The corresponding recorded values are

$$a_1\ a_2\ a_3\ b_1\ b_2\ b_3$$

The carryover between specimen A_3 and specimen B_1 is given by

$$K = \frac{b_1 - b_3}{a_3 - b_3}$$

(b_3 is assumed to be the 'true' value for serum B since it is preceded by two specimens of equal concentration and the effect of carryover should be negligible).

The concentration difference between sera A and B should be sufficient to permit accurate measurement of K.

Carry out 10 such tests for each analytical method. Results should state each value of a_3, b_3, and K.

6 DETERMINATION OF CROSS CONTAMINATION AND SAMPLE DILUENT CONTAMINATION These parameters are best measured using a radioisotope technique because of the accuracy, ease of operation, and extreme sensitivity it affords.

To measure the cross contamination between sample cups, *ie*, the volume of specimen carried from one cup to the next, the following procedure should be used:

(*a*) Prepare two aqueous solutions of bovine albumin containing 10 g/100 ml and 2 g/100 ml respectively.

(*b*) To half of each solution add a suitable quantity of protein labelled with, for example, ^{131}I, and mix well.

The amount of labelled protein added will depend upon the sample volume used and the limits of detection required.

(*c*) Arrange sample cups, filled to the normal level (see below) in the following sequence:

Cup No.[1]	1	2	3	4	5	6
g/100 ml	10[2]	10[2]	10[2]	2	2	2
Cup No.	7	8	9	10	11	12
g/100 ml	2[2]	2[2]	2[2]	10	10	10

[1]Each cup will have a corresponding mixture tube to receive the sample.
[2]Indicates a labelled specimen.

(*d*) Operate the sampling procedure in the normal way using water as the diluent.

(*e*) Transfer the contents of sample cups 4 and 10 and of mixture tubes 3 and 9 to suitable containers for radiation measurement.

(*f*) Ensure that the height of the liquid in all containers is the same to avoid geometric errors, and measure the radioactivity in each. The cross contamination between the sample cups is:

$$\frac{\text{Activity in cup 4}}{\text{Activity in tube 3}} \times \text{volume of sample dispensed}$$

and

$$\frac{\text{Activity in cup 10}}{\text{Activity in tube 9}} \times \text{volume of sample dispensed.}$$

Cross contamination could be expressed as the volume of the preceding specimen deposited in each cup as a percentage of the volume of specimen already present in the cup. However, this parameter (percentage cross contamination) is not a constant, being greater the less specimen there is in each cup. Thus cross contamination should be reported in terms of volume.

Cross contamination arises principally from the adhesion of specimen to the outer surface of the sample probe. The volume adhering will depend on the depth of probe entering the specimen, which in turn depends on the depth of specimen in the cup. For this reason the cups should all be filled to the same level, and the volume of specimen (bovine albumin solution) used, and the depth to which the probe enters the specimen, should be reported.

Operations (*c*) to (*f*) should be carried out five times and all the values for cross contamination obtained should be reported, together with the volume of sample taken and the dilution.

A similar technique is used to determine sample diluent contamination.

(*a*) Prepare an aqueous solution containing 7 g/100 ml of bovine albumin.

(*b*) Prepare an aqueous solution containing a suitable quantity of a labelled compound, for example, $K^{131}I$. This solution, which is used as sample diluent, should have a viscosity similar to that of water.

(*c*) Fill 10 sample cups to the normal level (see below) with bovine albumin solution. Each cup will have a corresponding mixture tube to receive the sample.

(*d*) Operate the sampling procedure in the normal way, using the labelled solution prepared as in (*b*) as diluent.

(*e*) Transfer the contents of the sample cups and mixture tubes to suitable containers for radiation measurement.

(*f*) Ensure that the height of liquid in all containers is the same to avoid geometric errors, and measure the radioactivity in each. The sample diluent contamination is given by:

$$\frac{\text{Activity in sample cup}}{\text{Activity per ml of diluent}}$$

Report the volume of specimen (bovine albumin solution) in the sample cup, the depth to which the sampler probe enters the specimen, the volume of sample taken, and the dilution.

List all the values obtained for the sample diluent contamination.

7 ACCURACY AND PRECISION If one specimen is analysed a number of times, and the 'true' result of the determination is known, then the amount by which the mean of the results obtained departs from the 'true' result is a measure of the accuracy of the determination. The assessment of accuracy can be approached in two different ways. In general it is possible to obtain either an aqueous standard solution of accurately known composition or a specimen of pooled serum which has been analysed a number of times. Analytical accuracy is a function not only of machine performance, but also of the chemical method used, and of the accuracy with which calibration solutions have been standardized.

If one specimen is analysed a number of times, then the degree to which the results vary from the mean of the results indicates the precision. This is normally measured by calculation of the standard deviation and coefficient of variation, reference to which may be found in any standard book on statistics.

If a number of specimens of the same serum are analysed in succession in a single batch, the results are obtained under the best possible conditions. Although such within-batch precision may be satisfactory, it is essential to measure the between-batch precision to obtain results relevant to normal operation. Furthermore, in some tests the precision may vary with concentration and tests at several different concentrations are therefore needed. Similarly, accuracy may vary between batches or at different concentrations and a range of tests should be performed.

8 DETERMINATION OF ACCURACY (*a*) Several specimens of serum will be needed. These are prepared by reconstituting a sufficient number of bottles from the same batch of commercial control serum and mixing. Aliquots are then deep-frozen, and must be used within one week of preparation. At least three such preparations should be made for each analytical method.

(*b*) The manufacturer's values for these sera should be checked by analysing 20 samples or one machine batch, whichever is the less, by established methods. Twenty samples or one machine batch, of the same sera can then be analysed with the system under test. Report all three sets of values, *ie*, those given by the serum manufacturer, those found by established methods, and those found using the system under test. This latter set of values should be reported in the order in which the results are produced so that any systematic variation may be detected. A brief description should be given of the established method used.

(*c*) The linearity of response over the complete concentration range should be measured with aqueous solutions for urea, sodium, potassium, chloride, bicarbonate, calcium, phosphate, glucose, and total protein. The results will also serve as a check on the accuracy of any reference serum used.

(*d*) The limits of linearity of response of all analytical methods tested should also be assessed using commercial sera. Take two preparations with 'high' and 'low' concentrations of the constituent in question, and prepare a series of specimens of different concentrations by mixing them in varying ratios.

(*e*) At least 100 specimens from patients should be analysed by conventional methods and those under test. Results by the two methods should be compared by plotting regression lines and calculating correlation coefficients. Preliminary experiments should be carried out to ensure that the calibration solutions used give the same result when analysed by each method.

9 DETERMINATION OF PRECISION (*a*) For each analytical method, the between-batch precision should be determined for at least three different concentrations, including those which could be most critical in clinical situations. Thus at least three specimens of serum (pooled serum or commercial control serum) will be needed: the concentration of the constituent in question must be known approximately, and these concentrations should fall in the ranges given below.

(*b*) Analyse one specimen of each serum in each of a number of batches of specimens analysed over a period of several days. At least 20 replicate analyses should be made on each specimen, preferably one per day. If time does not permit this, the between-batch variation may be obtained by analysing replicates each morning and afternoon for 10 days. Results should state the mean value, number of analyses, and standard deviation, together with the time over which results were obtained. It is important that specimens are arranged within the batches to which they are added in a random manner.

(*c*) Several serum specimens should be analysed and all results should be calculated and reported, but those at the concentrations given in the Table are particularly important.

(*d*) The within-batch precision is determined using the results obtained in the experiment of paragraph 8 (*b*). Calculate the mean value and standard deviation and report the number of samples analysed.

10 RUNNING COSTS These must be standardized under several well defined headings, and should be

TABLE

CONCENTRATIONS OF SERUM SPECIMENS

Sample	*Concentration*		
Urea (mg/100 ml)	30-40	70-80	160-180
Glucose (mg/100 ml)	40-50	140-160	200-240
Na (m-equiv/l)	120-125	135-140	150-155
K (m-equiv/l)	2·5-3·0	4·5-5·0	6·5-7·0
Cl (m-equiv/l)	80-85	100-105	115-120
CO_2 (m-equiv/l)	10-15	20-25	30-35
Total protein (g/100 ml)	4-4·5	6-6·5	8-8·5
Albumin (g/100 ml)	2-2·5	3-3·5	4-4·5
Calcium (mg/100 ml)	7-8	9·5-10·5	12-13
Phosphorus (mg/100 ml)	2-2·5	4-5	7-9
Uric acid (mg/100 ml)	4-5	6-7	8-9
Cholesterol (mg/100 ml)	200-250	300-350	400-450
Akaline phosphatase (K-A units)	10-15	25-30	40-50
GOT (Karmen units)	15-30	60-70	100-120

reported as 'mean cost per sample'. Show the period of time to which this costing refers, and the number of samples subjected to each analytical method during this period.

(*a*) Reagents and distilled water.

(*b*) Calibration serum (if used).

(*c*) Consumables, such as sample cups, test tubes, recorder charts, etc.

(*d*) Cost of repairs and maintenance. Assess these in relation to any maintenance contract.

(*e*) Detailed notes must be kept of the time spent by staff in operating the machine, preparing specimens, recording results, preparing reagents, etc. It may not be possible to arrive at an exact cost for these, but an approximate figure should be given for the number of staff, and their grades or special skills, required for normal operation of the instrument.

11 RECORD OF MACHINE PERFORMANCE It is essential that a record be kept of machine performance. All failures should be recorded, together with the steps taken to rectify them. Report the total time during which the machine could not be used because of breakdown, and also the total time of successful operation.

12 SUBJECTIVE REPORT Expressions of opinion should not be recorded elsewhere in the report but, where necessary, a subjective report should be made. This should not include comments on matters on which factual results are available, such as carry-over and accuracy, except where it is felt that the factual results presented do not give a true picture of the capabilities of the system.

Comments on the attitude and ability of the manufacturer and service engineers may be made here. The ease with which technicians learned to operate the system, and the usefulness of any instruction manual or other similar data should be reported.

It may also be useful to include a brief discussion of the ease with which the system can be incorporated into the routine of the laboratory.

This document is reproduced by permission of the Department of Health and Social Security.

Requests for reprints should be sent to Mr M. A. Buttolph, Department of Health and Social Security, Scientific and Technical Branch, 14 Russell Square, London WC1.

GLOSSARY

THE ACCURACY of a determination is the agreement between the result found and the true or most probable value of the quantity measured.

BETWEEN-BATCH PRECISION See 'precision'.

CARRYOVER is the degree to which the result obtained is influenced by the concentration of the preceding sample. This is also known as 'interaction', and is defined in mathematical terms in paragraph 4 of the section 'Evaluation of instruments in the clinical laboratory'.

CHANNEL The modules or components involved in one analytical method are collectively termed a 'channel'. A single-channel system will, without modification, analyse for one constituent in each specimen. A multichannel system will, without modification, carry out two or more analyses concurrently on the same specimen.

CONTAMINATION See 'sample contamination'.

A CONTINUOUS FLOW analytical system is one in which solutions derived from succeeding samples all flow along the same path throughout the entire analytical process. See also 'discrete'.

CROSS CONTAMINATION See 'sample contamination'.

A CUVETTE is an optical cell in which the optical density of coloured solutions is measured.

A DISPENSER is a device for dispensing automatically a preset volume of liquid.

A DISCRETE analytical system is one in which the solutions derived from different samples are contained, during part or all of the analytical process, in separate vessels. See also 'continuous flow'.

A MACHINE BATCH is the maximum number of specimens that can be sampled by the system without further attention from the operator.

PLASMA is the clear portion of unclotted blood from which the cells have been removed.

PRECISION is the agreement between a series of measurements of the same quantity. When a group of identical specimens is analysed in sequence the agreement found

is the 'within-batch precision'. The 'between-batch precision' is the agreement found when identical specimens are added to different groups of specimens and analysed. The precision is expressed as a standard deviation.

SAMPLE is that part of a specimen (q.v.) on which the analysis is performed.

SAMPLE CONTAMINATION is the contamination of the specimen during sampling, and is measured as the volume of foreign material added to the specimen. 'Cross-contamination' is the volume of the preceding specimen deposited in the sample cup during sampling. 'Sample diluent contamination' is the volume of sample diluent deposited in the sample cup during sampling.

SAMPLE CUP is the vessel containing the specimen (q.v.) from which the sample (q.v.) is taken.

SAMPLER is a device for taking automatically a preset volume of sample.

SAMPLING RATE is the rate at which samples are taken.

SERUM is the clear portion of clotted blood from which the cells and clot have been removed.

SPECIMEN is the bulk material available for analysis.

THROUGHPUT TIME is the interval elapsing between sampling and the output of the result.

WITHIN-BATCH PRECISION See 'precision'.