Pure & Appl. Chem., Vol.52, pp.2485-2494.
Pergamon Press Ltd. 1980. Printed in Great Britain.

ASSESSING THE ANALYTICAL QUALITY OF THE CLINICAL LABORATORY

T. P. Whitehead

Department of Clinical Chemistry, Wolfson Research Laboratories,
Queen Elizabeth Medical Centre, Edgbaston, Birmingham, UK.

Abstract - The clinical laboratory services of the UK National Health Service (NHS) is the fastest growing expenditure area of the Service. The clinical chemistry services are probably the largest of the disciplines making up the clinical laboratory service regarding workforce, workload and expenditure. In the early years of the NHS there was evidence of poor analytical performance by the clinical chemistry laboratories. This was shown by external quality assessment (EQA) techniques. By continuous and frequent use of EQA where portions of the same material are sent to over 400 laboratories and analyses compared it has been shown that the analytical performance has considerably improved. The reasons for this are probably the adoption of methods shown to perform well on EQA and also the widespread use of internal quality control techniques. The difficulties of providing appropriate material for EQA and the method of dealing with poor performance are also discussed.

INTRODUCTION

Approximately four per cent of the costs of the National Health Service Hospital Costs are spent on clinical laboratories. Thus, in the current year approximately 250 million pounds will be spent on laboratory services. The labour force is now said to be higher than that of the Scientific Civil Service. It has, for many years in a situation of escalating costs been the forerunner, growing faster than any of the other services. This is a reflection of the increased use of laboratory investigation in the diagnosis and treatment of patients. Within the various disciplines comprising the clinical laboratory services the discipline of clinical chemistry has probably the largest workforce, workload and investment in capital equipment. It is the youngest of the disciplines only having a history of some four decades. The laboratories comprising the discipline are virtually all situated in hospitals although providing some general practioner services. There are approximately 400 such laboratories. The labour force includes graduates in science and medicine and technologists.

A basic role of the clinical chemistry laboratory is to provide reliable data on the composition of specimens obtained from patients as an aid to the diagnosis and treatment of disease.

During the last three decades there has been abundant evidence both from the UK and from many other parts of the world that where portions of the same material are analysed by a number of laboratories that the variance in the results obtained indicate a situation which is unacceptable on the basis of both clinical and analytical criteria.

For example, in 1952 Professors Wootton and King from the Royal Postgraduate School of Medicine in London distributed a serum to laboratories in the UK for calcium analysis. One half of the laboratories stated that the level of the calcium was "within the normal range", one quarter stated it to be above normal, one quarter to be below normal. Similar surveys in the U.S.A. resulted in both Federal and State regulations regarding the accreditation of laboratories. In the U.S.A. Senate such words as scandalous and criminal were used to describe some of the results produced at high cost to the patient in the U.S.A.

The first individuals to practice clinical chemistry were graduates in medicine and biochemistry and the technologists were frequently recruited from other disciplines such as histopathology. Thus there was not a sound basis of analytical chemistry in the subject, the introduction gradually over the years of more chemists has been a stimulus to better analytical training for all entering the discipline.

This paper describes the observation of the analytical quality of clinical chemistry laboratories in the U.K. and summarises the endeavours made to improve the situation.

QUALITY ASSURANCE

All the measures taken to ensure good clinical chemistry practice from preparation of the patient before collection of a specimen to correct interpretation of a result are described as quality assurance techniques. Changes in the composition of biological specimens produced by incorrect collection procedures or inappropriate specimen handling can far outweigh the analytical variance. Assuring quality in patient investigation is not solely an analytical problem.

INTERNAL QUALITY CONTROL

Clinical chemistry laboratories in the U.K. now normally practise some form of internal quality control, usually the performance of assays on quality control material at the same time as patients' material. This type of quality control has been introduced as a result of active teaching for all grades of staff, the inclusions of questions regarding the techniques in examinations and active research and development resulting in many publications.

EXTERNAL QUALITY ASSESSMENT

An important stimulus to such activity has been external quality assessment. This is where portions of the same material are sent to several laboratories for particular analyses and the results compared. This paper describes the author's experience with the U.K. National Quality Assessment Scheme (NQAS) in Clinical Chemistry.

The Scheme is sponsored by the Department of Health and Social Security (DHSS) who also sponsor schemes in the other disciplines. The Scheme is operated from the Wolfson Research Laboratories, Department of Clinical Chemistry, Queen Elizabeth Medical Centre, Birmingham. It is financed by the DHSS and employs six people.

THE UNITED KINGDOM NATIONAL QUALITY CONTROL SCHEME (UKNQCS)

The main objectives of the Scheme when it started in 1969 were:

1. To send at fourteen-day intervals a portion of the bulk human serum and, on occasions, non-human serum, to all those hospital laboratories in the U.K. which perform clinical chemistry analysis.

2. To assess results from fifteen of the more commonly performed analyses. A laboratory which did not routinely perform all of the fifteen analyses, would not necessarily be excluded from participating in the Scheme.

3. To return results from participating laboratories to the organizing laboratory quickly; the results from all laboratories to be available to the participants within ten days of the specimen arriving in the participating laboratories.

4. To make participation voluntary and preserve anonymity.

5. To present the results in a manner that would enable the participants to make judgements of their performance, particularly in relation to the analytical method used.

6. To assess the role of automation, analytical methods, laboratory workload, and other factors possibly affecting the variance of reuslts.

7. To determine whether any improvement in precision and accuracy in the hospital laboratories of the U.K. occurred as a result of such frequent surveys.

ORGANISATION OF THE SCHEME

The Scheme has been in operation for eleven years. The distribution of serum specimens to 200 laboratories in the U.K. began in July 1969; at the present time the participants number 420. There is reason to think that the vast majority of laboratories within the NHS performing clinical chemical analysis have entered the scheme and approximately 90 per cent of participating laboratories return the results for each distribution of serum. At the present time the survey regularly includes fifteen different chemical determinations. These are serum sodium, potassium, chloride, urea, glucose, calcium, phosphate, iron, total protein, albumin, bilirubin, alkaline phosphatase, cholesterol, uric acid, and creatinine. Surveys of blood lead, thyroid function tests and certain enzymes are also being carried out. The computer in the author's laboratory has been programmed to perform virtually all the clerical tasks involved in the scheme and the survey involves approximately four hours of computer time each fortnight.

The difference between the NQAS and previous surveys was the frequency of survey, twenty times a year, approximately five times as frequent as many other schemes. Also the relatively short delay, ten days, between receipt of the specimen and the receipt of the results of the survey. Both these differences were made possible by the use of computer facilities.

The material distributed

Obtaining sufficient suitable material of acceptable quality for distribution is a major problem. Virtually all the material used at the present time has been lyophilized usually by commercial organisations. It is not always of human origin, animal material is suitable for some tests. Confidence by the participants in the organization of such schemes is primarily based upon the quality of the material distributed; it is common for laboratories with poor performance to blame the material. Providing the quality of the materials used has to be based upon the fact that certain participants consistently achieve an excellent analytical performance on material for an individual survey which has been randomly distributed. The problem of how well reconstituted lyophilyzed material simulates patients' serum will be dealt with later.

Timetable of distribution of sera and results

Labelling for distribution is facilitated by the computer line printer which draws on a disc file of addresses and code numbers of the participating laboratories. The computer also prepares the document to be returned by the participants. This document (Fig. 1) indicates which analyses are to be performed, which units are to be used and also gives any special instructions regarding reconstitution of the material. The documentation and material are packed in a protective polystyrene box.

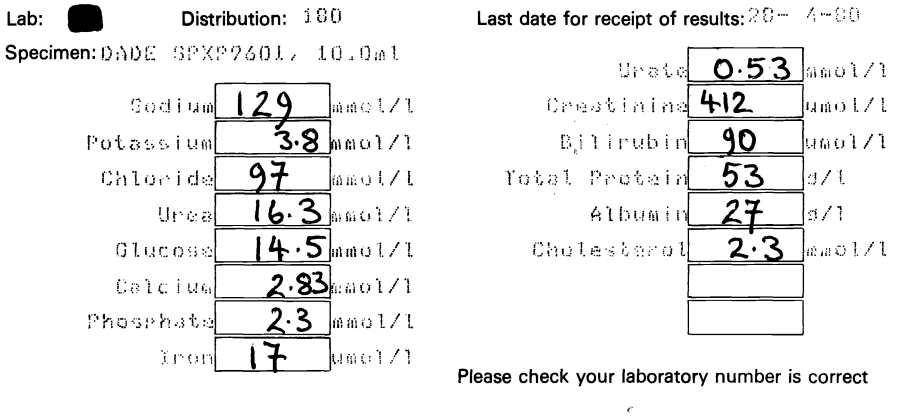

Fig. 1 Computer printed document sent with specimen and retained by participants

All packs are posted on a Saturday and almost without exception arrive in the participating laboratories early on the following Monday morning (day 1). The laboratories perform the analyses listed and return the results before the following Monday (day 8). On day 8 all results are put into the computer and following verification they are analysed on the Tuesday (day 9). Computer printout of results are posted to the participating laboratories on the Wednesday (day 10). This timing of distribution enables at least 90 per cent of all participating laboratories to be included in the printout prepared on day 9.

Format of report to participants

Each participating laboratory has its own computer printed report. It is important that such documentation has information of educational importance comparing performance by different analytical methods. It is also important that the statistical methods used are clearly understood by all participants. It is frequently those laboratories with least understanding of the statistics used who need the most help in improving their performance.

367 LABS PARTICIPATED IN THE SCHEME

	NA	K	CL	UREA	GLUC	CALC	PHOS	FE
NO. OF RESULTS	384	384	273	374	367	339	287	227
MEAN VALUE	139.0	2.84	106.4	6.0	3.7	2.02	1.95	28.3
STD. DEVIATION	1.9	0.14	2.7	0.5	0.4	0.14	0.12	3.0
COEFF. OF VAR.	1.3	4.8	2.5	7.7	11.4	6.7	6.3	10.6

RECALCULATED RESULTS EXCLUDING THOSE OUTSIDE ± 3 SD IN THE ABOVE CALCULATIONS:—

NO. OF RESULTS	379	378	270	369	363	335	284	220
MEAN VALUE	139.0	2.83	106.4	6.0	3.7	2.02	1.95	28.6
STD. DEVIATION	1.6	0.09	2.4	0.4	0.3	0.08	0.10	2.4
COEFF. OF VAR.	1.1	3.3	2.3	6.3	9.3	4.0	5.2	8.3

Fig. 2 Example of calculation of overall means and recalculation

The following is a description of the information provided in the computer printout. The computer lists the results attributed to each laboratory so that they may be checked for clerical errors by the participating laboratory. The mean, standard deviation, and coefficient of variation for each determination is calculated and printed. After removal of all results outside three standard deviations either side of the mean, these statistics are recalculated and these are termed the recalculated mean and standard deviation. An example of this portion of the printout is shown in Fig. 2. This technique eliminates those results which are probably due to random mistakes. Following the statistical calculations, there are printouts of histograms of the reported results for each determination. An example of such a histogram is shown in Fig. 3. The range of the histogram corresponds to the recalculated means, ±2 SD. A result within the limits is shown by a cross and a result outside these limits by a dot. These limits are not 'limits of acceptability' but are a convenient method of presenting the results, which enables each participant's results to be related to all other results.

The computer disc file contains information regarding the analytical methods in use in the participating laboratories for each determination and the results are classified according to the methods in use. These are presented as statistical summaries. Only results used in the calculation of the recalculated mean are included. The mean, standard deviation, and coefficient of variation of the results of each method are calculated and a summary is typed and included in the report received by each participant. Table 1 illustrates the format used for the presentation of the results according to analytical method, glucose is used as an example.

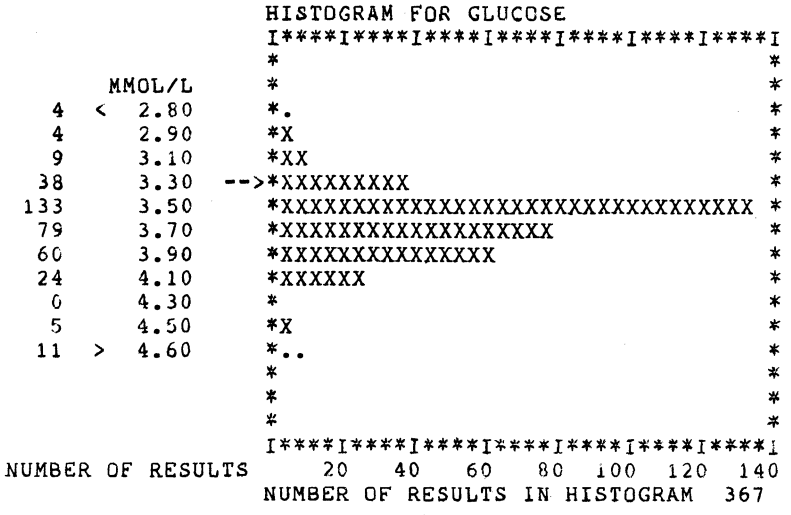

Fig. 3 Example of a histogram for glucose. The arrow indicates the actual value.

TABLE 1. Example of print out for glucose showing the statistics of individual methods

RESULTS FOR GLUCOSE (MMOL/L)
(EXCLUDING VALUES OUTSIDE +/- 3S.D.)

	NO.	MEAN	S.D.	C.V.
OVERALL	363	3.66	0.34	9.26
-->MANUAL-GLUCOSE OXIDASE	69	3.70	0.46	12.57
OTHER	51	3.62	0.36	9.87
AUTOANALYZER I REDUCTION	18	3.83	0.47	12.27
AA II OR SMA REDUCTION	14	3.91	0.32	8.25
AA I GLUCOSE OXIDASE	99	3.69	0.25	6.66
AA II OR SMA GLUCOSE OX.	62	3.68	0.29	7.86
BECKMAN GLUCOSE ANALYZER	48	3.52	0.24	6.78
YOUR RESULT:		3.20		

Calculation of the variance index

If EQAS are to play a part in improving the performance of the participating laboratories then it is the author's experience that some type of performance scoring is essential.

The performance of individual laboratories should be compared with the performance of all participating laboratories. This means that all laboratories should know what the best laboratories are capable of. Mere declaration of the percentage of laboratories achieving results within a certain acceptable coefficient of variance does not communicate such information. Performance in any individual laboratory and in all laboratories taking part in a scheme will alter with time, knowledge of the magnitude of such changes is also important. Laboratories in EQAS should be provided with information indicating quantitative relations of performance of individual laboratories to methods, apparatus, workload, staffing, etc.

It is for these reasons that the variance index was devised.

Calculation of the variance index (VI)

The VI is calculated on the results obtained from the participating laboratories for a particular determination. First, the mean value obtained by laboratories using the same method is calculated. Previously the type of

analytical methods used by participants for individual determinations have been classified and those using the same or similar methods are grouped together for calculation of the method mean. The participant is required to agree to such classification. For some determinations participants may use methods which cannot be classified in this way and their results cannot be used in VI calculations. More than 90 per cent of the 420 participating laboratories use methods which can be used for VI calculations.

The calculation only uses those values which fall within the mean ±3 SD for the results returned by participants for this method. This is to avoid incorporating results which are random mistakes, such as those occurring in clerical transcription into the method mean calculation and thus falsely distorting the value.

The method mean ($\bar{x}_m$) is subtracted from the result of an individual laboratory (x) and the percentage variation from the method mean calculated.

$$\%\text{Variation} = V = \frac{x-\bar{x}_m}{\bar{x}_m} \times 100 \text{ (The sign is ignored).}$$

The VI is calculated from this figure by dividing it by the chosen coefficient of variation (CCV) given in Table 2. To avoid decimal points this figure is multiplied by 100.

$$\text{Variance Index} = VI = \frac{V}{CVV} \times 100.$$

Obviously, the lower the VI, the closer the result is to be method mean. The CCV values shown in Table 2 are the lowest CVs obtained for particular determinations during the first two years of the scheme. They are kept constant so that improvements in the performance of laboratories can be detected.

TABLE 2. Chosen coefficient of variation used in VI calculations

Determination	Coefficient of variation (%)	Determination	Coefficient of variation (%)
Sodium	1.6	Uric Acid	7.7
Potassium	2.9	Creatinine	8.9
Chloride	2.2	Bilirubin	19.2
Urea	5.7	Total Protein	3.9
Glucose	7.7	Albumin	7.5
Calcium	4.0	Alkaline phosphatase	19.6
Phosphorus	7.8	Cholesterol	7.6
Iron	15.0		

Because the coefficient of variation and not standard deviation is used in the calculation of VI when the mean value falls outside the limits listed in Table 3 the VI is not calculated. It is particularly important to avoid VI calculations on low mean values for serum determinations with a high variance such a bilirubin, alkaline phosphatase, and iron.

A formal definition of variance index is 'the difference between the result obtained by a participant and the calculated method mean expressed as a percentage of the mean, divided by a chosen coefficient of variation for that determination; the resultant figure is multiplied by 100'.

TABLE 3. Range of values used in the VI calculation

Determination	Low	High	Units
Sodium	110.0	160.0	mmol/l
Potassium	1.5	8.0	mmol/l
Chloride	65.0	130.0	mmol/l
Urea	2.5	66.7	mmol/l
Glucose	0.8	22.2	mmol/l
Calcium	1.0	4.0	mmol/l
Phosphorus	0.6	3.9	mmol/l
Iron	3.6	53.6	µmol/l
Urate	179.0	893.0	µmol/l
Creatinine	62.0	1770.0	µmol/l
Bilirubin	9.0	342.0	µmol/l
Total protein	40.0	100.0	g/l
Albumin	15.0	60.0	g/l
Alkakine phosphatase	6.0	100.0	KA units/100 ml
Cholesterol	1.3	12.9	

TABLE 4. Running over-all mean VIS for laboratories with different workloads (the figures in brackets are the number of tests performed each year)

All laboratories	68	Size 3 (101-185,000)	66
Size 1 (50,000)	83	Size 4 (185,000)	61
Size 2 (50-100,000)	71		

Variance index score (VIS)

The performance of an individual laboratory in several analyses of different material for the same substance may be expressed as the mean variance index. In the UKQAS we use the term variance index score (VIS). To avoid incorporating very high VI values in the score, possibly due to a clerical error, VI values greater than 400 are treated as 400.

The mean VIS may be calculated for different determinations and several distributions, the resultant calculation is *the over-all variance index score*. In practice it has been found useful to calculate the *'running' over-all variance index score*. In this the over-all variance index score for the most recent forty analyses is calculated. Where the score for more recent results are added, the appropriate number of the earliest results are dropped out of the calculation.

Running variance index score

Figs. 4 to 7 illustrate the running VI score graphs which are drawn on the computer graph plotter.

Each square represents a distribution of serum which was used for the calculation of running VI score. The graphs involve a period of two years. A break in the graph indicates that a laboratory did not return a result for that particular determination.

The worst 5 per cent and the best 5 per cent lines delineate the area in which 90 per cent of laboratories have VIS.

Fig. 4 illustrates the results from a laboratory where performance was poor for the first six months of the period, gradually improved in the next six months, remained static as regards performance in the next six months, and then dramatically improved within the last few weeks.

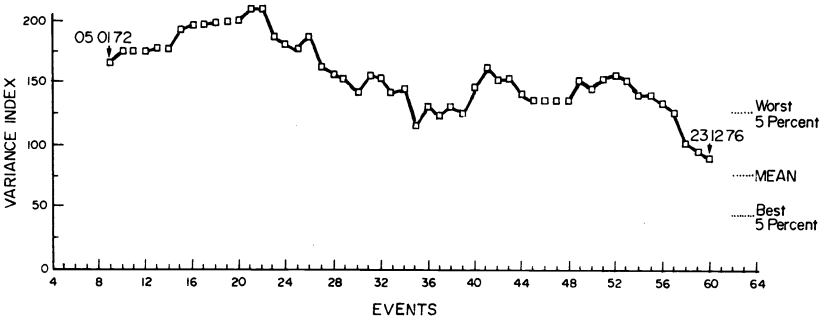

Fig. 4 Example of the running variance index score computer printout

Fig. 5 shows the results from a laboratory with a very consistent performance kept up over a period of two years.

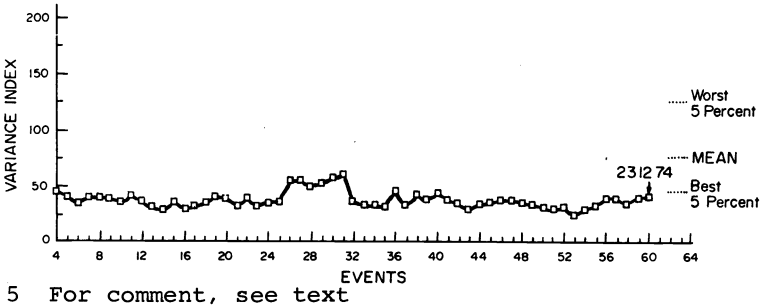

Fig. 5 For comment, see text

Fig. 6 shows the results from a laboratory which had an average performance which dramatically deteriorated over a period of a few weeks due to staff difficulties and then improved when these problems had been solved.

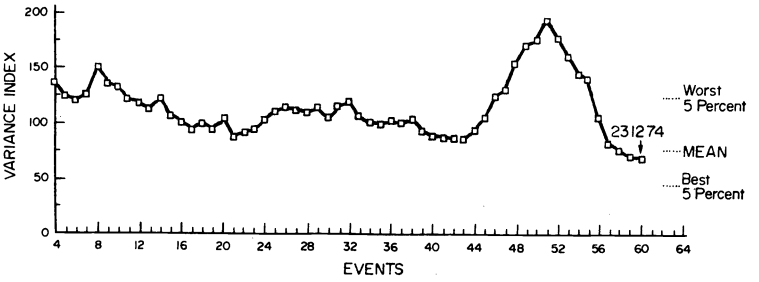

Fig. 6 For comment, see text

Fig. 7 is a typical running VIS graph.

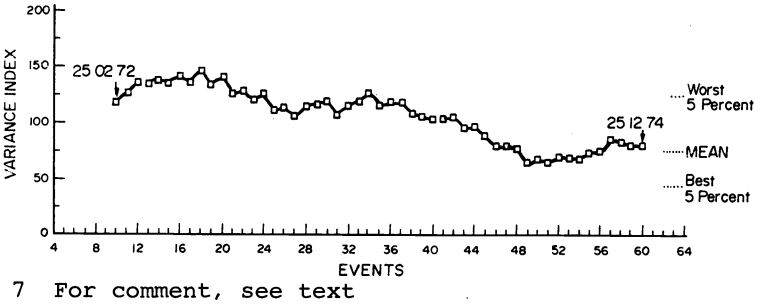

Fig. 7 For comment, see text

Throughout the scheme it has been shown that the laboratories with the smallest workload have the highest VIS. The current mean values for various levels of workload are shown in Table 4. There is good evidence that those laboratories making extensive use of automatic methods of analysis have the lowest VIS. Smaller laboratories using more manual methods are less precise.

At intervals, a table of the type illustrated in Table 5 is distributed to individual laboratories. It shows the VIS for the individual determinations and the mean VIS for each determination for all laboratories.

TABLE 5. Example of results of a participating laboratory showing the VIS for individual determinations compared with the mean VIS for all participating laboratories

Determination	No. of possible results	No. of results returned	Mean VIS	Mean VI for all labs.
Sodium	34	34	46	86
Potassium	29	28	42	88
Chloride	20	11	61	80
Urea	34	34	21	79
Glucose	33	33	21	85
Calcium	34	34	45	78
Phosphate	34	32	33	75
Iron	20	19	90	77
Uric acid	33	33	28	79
Creatinine	32	32	51	98
Bilirubin	31	23	25	96
Total protein	33	31	47	75
Albumin	33	33	17	79
Alkaline phosphatase	28	26	38	79
Cholesterol	33	32	56	88

In this way participants can judge which determinations are making significant contributions to their over-all VIS.

IMPROVEMENT IN PERFORMANCE

The mean VIS has consistently fallen over the years since it was introduced in late 1972. Fig. 8 illustrates the fall from November 1972 to June 1975. Over the last five years the fall has remained consistent but slower and the mean VIS for all laboratories is approximately 60% of that in 1972.

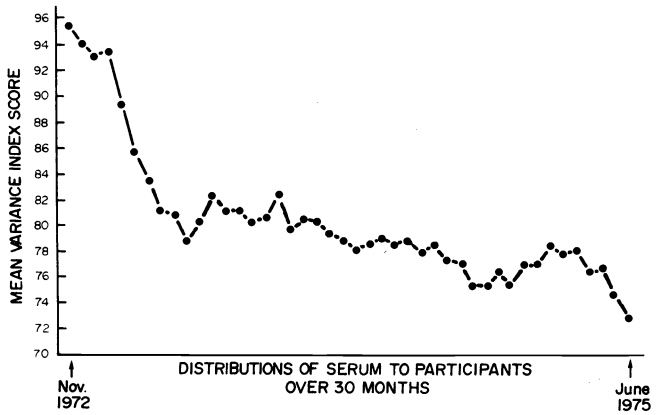

Fig. 8 Changes of Mean VIS with time - mean of all laboratories in the UKNQAS

The reasons for the improvement are possibly two fold. Firstly, there have been considerable changes in the methods used by laboratories for nearly all determinations surveyed. Manual reduction methods for glucose, and used by a quarter of participants in 1972, inherently imprecise have been completely replaced by 1979 with glucose oxidase methods, many automated. Salt fractionation methods of albumin determination a difficult method used by one third of the participants in 1972 has been abandoned for the simple and precise bromocresol green method. There are several other such changes and it is known that the EQAS has helped to persuade laboratories to change to better methods.

Secondly, there has been an improvement in the precision of methods in constant use during the eleven years of the scheme. For example, flame photometry of sodium and potassium has shown consistent improvement and the between laboratory coefficient of variation is consistently less than ±2.0%.

DEALING WITH POORLY PERFORMING LABORATORIES

Confidentiality has been preserved throughout the course of the Scheme. Three years ago the professional bodies associated with clinical laboratory work set up panels to consider the performance of individual laboratories. The results of each laboratory are considered by a panel of four laboratory workers. They decide whether the performance of a laboratory is acceptable, if not, then a letter is sent pointing out the poor performance and offering help. So far this technique has been successful and it is not envisaged that their will need to be compulsory participation such as in the U.S.A. and Germany.

Choice of the target values

The use of the method mean as the target value on which to judge performance has been critisised. For some determinations such as enzyme the use of values provided by reference laboratories is mor appropriate.

There is evidence that for the ions sodium, potassium, chloride and calcium the method mean on the Scheme lie very close to the values obtained by definitive methods.

A combination of consensus and reference laboratories is gradually being adopted.

CHOICE OF MATERIAL FOR DISTRIBUTION

The Scheme relies on lyophilyzed material. There is now increasing evidence that some materials lack commutability, that is the ability to simulate fresh patients serum and this is certainly a problem with certain methods. The addition of preservatives and other substances to lyophilyzed material may also produce problems in certain determinations. Organisers of schemes like the UKNQAS have to continually be concerned with research into the properties of lyophilyzed material. An additional problem is that the ability to obtain and lyophilyze material in large batches is solely based in commercial organisations. There is no doubt that studies of material properties is required if further development in survey work are to take place.

ACKNOWLEDGEMENTS

This work would not have been possible without the help of many colleagues and the support of the Department of Health & Social Security.