Ann. Ist. Super. Sanità, vol. 31, n. 1 (1995), pp. 71-75

EQA for general haematology in the United Kingdom

S. Mitchell LEWIS

Department of Haematology, Royal Postgraduate Medical School, London, UK

Summary. - In the United Kingdom the national quality assessment scheme (NEQAS) in haematology organizes regular surveys for blood counts, blood films, reticulocyte counts, cytochemistry, identification of abnormal haemoglobins, HbA_2 and HbF quantitation, G6PD screening tests, assays of serum vitamin B_{12}, folate and ferritin. For most tests there has been significant reduction in inter-laboratory variance despite occasional blunders. This illustrates the role of NEQAS in improving the standard of laboratory practice in the UK. The problems in equating analysis of NEQAS survey materials with routine laboratory specimen are discussed.

Key words: EQA in haematology, blood counts, blood films.

Riassunto *(Valutazione esterna di qualità per l'ematologia generale nel Regno Unito).* - Nel Regno Unito lo schema nazionale di valutazione esterna di qualità (NEQAS) in ematologia prevede l'organizzazione, ad intervalli di tempo regolari, di indagini per i conteggi delle cellule del sangue, l'esame degli strisci di sangue, la conta dei reticolociti, la citochimica, l'identificazione di emoglobine anomale, il dosaggio di HbA_2 ed HbF, i test di screening per la G6PD, i dosaggi di vitamina B_{12}, folato e ferritina nel siero. Per la maggior parte dei test si è verificata una significativa riduzione della variabilità interlaboratori, nonostante occasionali errori grossolani. Questo dimostra la validità del NEQAS nel migliorare lo standard di qualità delle prestazioni dei laboratori nel Regno Unito. Vengono discussi i problemi che insorgono nell'assimilare le analisi effettuate su materiali di controllo del NEQAS con quelle sui campioni osservati nella routine del laboratorio.

Parole chiave: valutazione esterna di qualità in ematologia, conteggi delle cellule del sangue, esami dello striscio di sangue.

Introduction

The organization structure of the national external quality assessment scheme (NEQAS) in the UK is described in this issue (see Lewis [1]). There are now approximately 650 participants in the General Haematology Scheme in the UK who are registered for one or more of the tests which constitute the regular surveys, as indicated in Table 1. In this article some specific aspects of UK NEQAS will be described.

Blood count

To ensure adequate stability of the blood cells the blood is collected into ACD or CPD solution in which the red cells can be maintained for at least 3-4 weeks, or the blood is fixed to maintain all the cells in a condition suitable for counting (see Lewis, in this issue [1]).

At 3-4 weekly intervals each laboratory receives two samples for blood count; these are selected so as to give maximum information that will help to identify the reason for any discordant results. Thus paired samples might be identical in order to check on imprecision, or the samples may be at different levels of concentration to check on linearity of response or at levels of concentration which are clinically important, e.g. at the margins of the normal reference range. For the blood count equine blood is particularly useful as the red cells have MCV of 48 fl (horse) and 58 fl (donkey), thus providing a model for human microcytosis.

An important purpose of NEQAS is to ensure interlaboratory harmonisation of results of tests in practice. Unfortunately, cell counters based on different principles are likely to give different results on the preserved blood used in the NEQAS survey, especially in measurements of cell volume. Accordingly, it is necessary to analyse the results for each group of systems separately and to assess performance by participants according to their own instrument class as well as by comparison with all classes combined.

Provision of analyser-compatible NEQAS material is becoming increasingly difficult for the sophisticated systems which perform full differential leucocyte counts on EDTA blood, using cell size as well as a number of physical and chemical reactions to distinguish the types of cells. Attempts are being made, with varying degrees of success, to produce artificial materials which simulate at least some of the leucocytes, in at least some of the systems. Further studies are necessary to develop a suitable comprehensive material.

Vitamin B_{12}, folate and ferritin

For many years it has been recognised that determination of vitamin B_{12} and folate by microbiological assay is a complicated procedure requiring meticulous attention to technical details. The introduction of radioisotope (RID) assay was a technical advance in as far as the tests were simplified and came within the facilities available in many district general hospitals.

In the UK NEQAS some 250 laboratories take part in this section of the surveys. For these surveys the same serum samples are used for serum vitamin B_{12} and folate assays, whole blood samples for red cell folate and selected sera for serum ferritin.

For serum B_{12} there is an overall CV of 20% (range 15 - 33%) with method-related differences in median and SD. Ratio analysis has shown both consistent bias and random error. These observations suggest that the faults are essentially due to the methods used rather than to individual participants. Furthermore there are differences in the assessment by participants of the clinical significance of their results. This appears due not only to "correct" interpretation of a quantitatively incorrect measurement but also to incorrect interpretation of a correct results (Fig. 1). This appears to be due largely to lack of understanding of the theory of reference values as described by the International Council for Standardization in Haematology [2] and by using population reference ranges derived from various textbooks and other sources without taking account of method differences. There has been an improvement in red cell folate assays; an average CV of 30 - 35% in previous years has now reduced to about 16% in the last few surveys. By contrast, measurement of red cell folate is still unsatisfactory with an average CV of 35% and wide deviations between the different methods used by the participants.

Ferritin assays over the past decade show improvement, which has coincided with a changing pattern of methods and also with the availability of a reference standard. Interpretation of results is less satisfactory. Thus, for example, in a recent survey a specimen with a median value of 17 µg/l was reported as normal by 38% of the participants; one with a value of 40 µg/l was reported as low by 5%. In response to a questionnaire the lower limit of normal was variously stated to be 9 - 41 µg/l for men and 3 - 20 µg/l for women. As described above in the case of vitamin B_{12}, here, too there is need to establish normal reference ranges in accordance with ICSH recommendations.

HbA_2 and HbF

Approximately 280 laboratories now take part. When these trials were started the results showed wide variations with serious differences from results obtained by four referee laboratories. Thus, for example, the distribution of reported values of HbA_2 in one sample was 0 - 7% when the reference value was 2.3 ± 0.6% and in the other sample the values were 0.1 - 15% when the reference value was 4.6 ± 0.3%. Furthermore, x-y (Youden) plots showed

Table 1. - Test registration and survey programmes in UK NEQAS

	no. of labs	no. of surveys	no. of samples distributed to participants
Blood count (including platelet counts)	609 (*)	12	24
Blood films (including parasites)	500	4	16
Reticulocyte count	450	4	8
G6PD	208	4	12
Abnormal haemoglobins HbA_2/HbF	287	4	12
Vitamin B_{12}/serum & red cell folate	290	11	33
Ferritin	186	4	8
Cytochemistry	280	4	8

(*) with a total of 1000 blood counting systems

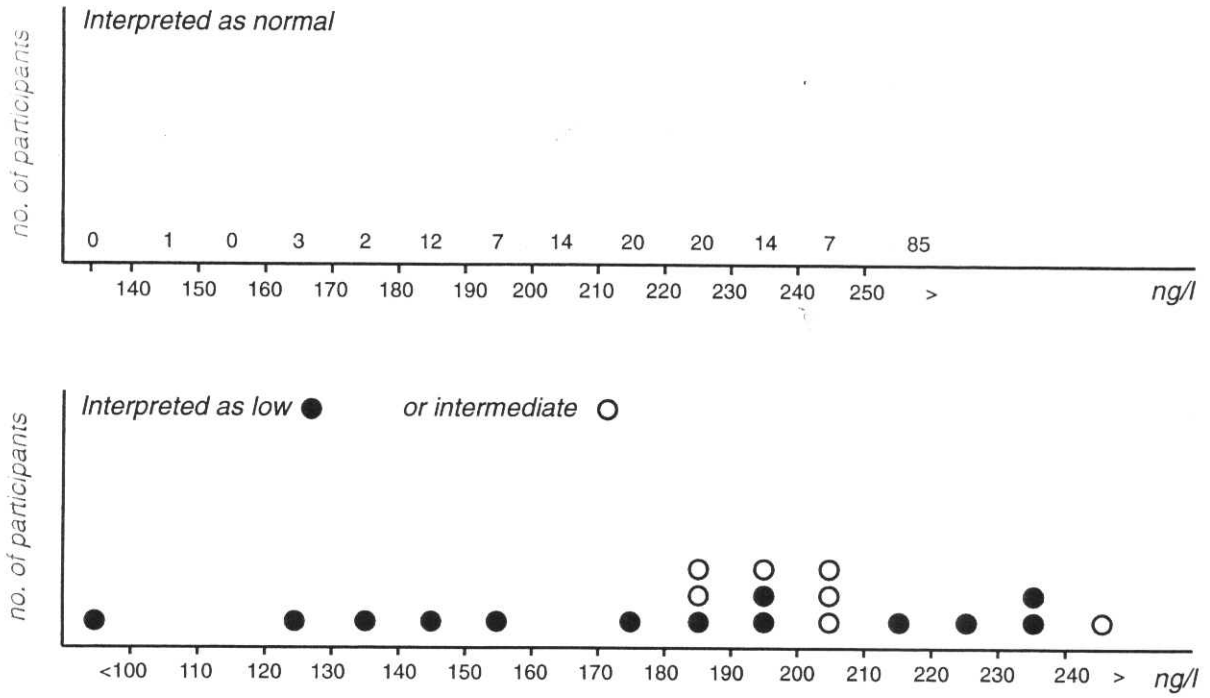

Fig. 1. - Results of vitamin B_{12} survey. The median value for vitamin B_{12} on the specimen was 232 ng/l (SD 38 ng/l). Note the range of results reported by the participants, and also discrepancies in some cases between measurement and interpretation.

Reproduced by permission of the Royal Society of Edinburgh from *Proceedings of the Royal Society of Edinburgh, Section B: Biological Sciences*, vol. 101 (1993), pp. 283-310.

intralaboratory errors in measurement to be random rather than consistent bias. Technical workshops were organised, the frequency of surveys was increased and participants were provided with samples of control material when necessary. The result has been a significant improvement in performance (see Lewis, in this issue [1], Fig. 2).

The CV for HbF has remained 40-60% in normal samples and 35% with increased levels. It is hoped that the recent availabilities of international reference reagents of HbA_2 and HbF will lead to improvement in performance.

Glucose-6-phosphate dehydrogenase

The specimens used are either normal human blood (5.5 - 7.5 IU/g Hb), G6PD deficient sheep blood (0.7 - 1.5 IU/g Hb), or a mixture of human and sheep blood to simulate an intermediate level (2.0 - 4.5 IU/g Hb). The surveys have shown that when correctly performed both dye decolorisation and fluorescent spot screening tests accurately distinguish grossly deficient samples from normal; however the dye decolorisation method is over-sensitive at the lower level of the normal range giving false positive readings; whereas it is more sensitive for identifying the intermediate state than the commercially available fluorescent kit (Table 2).

This demonstrates the role of NEQAS in identifying method (or kit) problems as distinct from unsatisfactory performance by an individual laboratory.

Blood films

There are four aspects in blood film surveys, namely, technical skill in staining, differential count quantification, reporting on blood cell morphology and opinion of haematological diagnosis. The use of unstained blood films to assess technical aspects of staining by participants may sometimes be unsuccessful as delays beyond 6 - 8 hours in staining the film, even when they are prefixed, may result in unsatisfactory staining. Accordingly, in most NEQAS surveys films have been fixed and stained in the organising laboratory by a standardised Romanowsky method.

Material selected for the surveys has included normal blood films, where the main emphasis has been on reliability of differential cell count counting, but the majority of films have come from patients with blood disease (Table 3). Some of these have been "bread and butter" haematological conditions, whereas others have been unusual or even rare problems, included to stimulate and provoke participants as well as to provide material for education and to give participants the opportunity to build up departmental slide libraries. In addition, at intervals, separate surveys consist of films for identifying and quantifying blood parasites.

Table 2. - G6PD surveys. Cumulative results in 18 tests

G6PD values (U/gHb at 30°C)	Dye decolorization kit				Flourescent spot kit			
	Tot. no.	Low	Equiv.	Normal	Tot. no.	Low	Equiv.	Normal
		%				%		
0.8 - 1.9 (deficient)	127	100	0	0	352	94	6	0
2.0 - 5.0 (intermediate)	64	91	6	3	178	46	22	32
> 5.0 (normal)	322	22	24	54	870	0	1	99

Participants report up to five outstanding morphological abnormalities including an estimate of any abnormal differential leucocyte distribution. Results are scored on the basis of a consensus of participant results and a consensus of comments from a selected group of referees. The laboratory from which the material originated provides a report of the definitive diagnosis (if this has been established) or the likely diagnosis from all the retrospective information, together with an educational commentary.

In general, the standard of morphological identification and interpretation is high, and the surveys are a popular part of the NEQAS programme. Nonetheless (or perhaps as a consequence), they have stimulated considerable debate. This relates especially to the amount of information which should accompany the films: it has been argued that these exercises are intended for assessing morphological skill and thus should not be biased by providing "leading" information whether by way of blood count data or clinical comment. The contrary view is that without the information which would normally be provided with routine blood film requests, the exercise becomes unrealistic and thus is contrary to the principles of EQA.

The essence of the problems is that blood films serve as the interface between technical performance, professional expertise and clinical judgement. This is coupled with the problem of how to establish the correct answer for each film, and whether EQA should have predominantly educative or regulatory function. In the UK there is universal agreement that the primary function of NEQAS is educational, and separate from the demands of accreditation.

As an increasing number of laboratories install automated systems, it is necessary to asses the future role of the blood film in the diagnostic laboratory and the importance of this component of quality assessment programme; as fewer films are examined, the need for external quality assessment becomes more rather than less important. There is need to maintain a high level of competence in morphological identification of haematological abnormalities; this must include an ability to screen normal films with a high level of specificity and abnormal films with a high level of sensitivity.

It is, however, a sobering thought that in the Canadian experience [3] a 32% error rate was found to be due to the use of poor microscopes which were dusty, incorrectly illuminated and inadequately maintained. Attention to these factors reduced the failure rate to 6%!

Reticulocyte counts

This represents one of NEQAS's less successful present activities. Pre-stained films are distributed and participants carry out microscopic counts by their usual method. For over ten years the CV has remained unchanged: 35 - 40% when the reticulocyte count is 1%, 20 - 30% at 10%; 15 - 20% at 15%. This is due largely to the fact that in most laboratories so few reticulocytes are included in the count especially when the reticulocyte count is low, that the Poisson distribution error is excessive. This could be reduced by 5%, or less, by counting greater numbers in accordance with the recommended reference method [4]. However, this is time consuming and impractical in a routine service laboratory.

Another source of error is inability to recognise late reticulocytes which have only a few minute particles of stained RNA; the reference method includes all cells with two or more stained platelets, and recommends the use of phase microscopy for their easier detection [4].

The fact that the reticulocyte counts are a labour intensive test, and have such a large variance in practice has resulted in its diminishing use albeit acknowledged in principle that it should have an important role as

Table 3. - Diagnoses in blood films distributed in NEQAS surveys during the past three years

Normal
Pelger-Huet anomaly
Iron deficiency anaemia
Heterozygous β-thalassaemia
HbS/β-thalassaemia
HbC disease
Sickle cell disease
Immune haemolytic anaemia
Hereditary elliptocytosis
Lead toxicity
Microangiopathic haemolytic anaemia
Pernicious anaemia
Alcholism
Abeta-lipoproteinanaemia
Refractory anaemia
Myelofibrosis
Myeolodysplastic syndrome
Thrombocythaemia
Hairy cell leukaemia
Lymphoma
Chronic lymphocytic leukaemia
Acute myeloid leukaemia
Chronic myeloid leukaemia
Acute lymphocytic leukaemia
Sezary syndrome
Hyperosinophilic syndrome
Myelomonocytic leukaemia
Megakaryoblastic leukaemia
Carcinoma: Leuco-erythroblastic anaemia
Leukaemoid reaction
Blood parasites
Various types of Plasmodium
Trypanosomiasis
Microfilaria
Not present

measure of erythropoietic function. Unfortunately the results in NEQAS surveys do not inspire participants to use this test extensively.

The recent advent of automated methods for counting reticulocytes by fluorescent flow cytometry may change this situation as many thousands of cells can be rapidly counted with the high level of precision and reliability which should give this test the respect for its clinical utility which it deserves. This new development challenges ICSH to provide suitable specimens with an accurately measured and stable reticulocyte population.

Conclusions

The limitations of EQAS must be recognised alongside its uses. In theory, EQA tests should be representative of the routine practice in a laboratory. In reality, NEQAS specimens are usually distinguished from routine specimens and in many laboratories results on the NEQAS specimens represent the "best performance" mean of several successive measurements rather than the result obtained in a single shot as is the normal procedure.

Another problem with the blood count is that EQA surveys require material which is of necessity preserved, by contrast to the fresh EDTA blood normally used. This creates an artificial situation - in some cases unsatisfactory performance in a NEQAS test is not reflected in tests on EDTA specimens and conversely, when a counter is recalibrated to give satisfactory results with the preserved blood it will then give discrepant measurements, especially of red cell size parameters, on the routine specimens. This is especially apparent in laboratories with two different systems functioning in parallel.

However, as NEQAS has been shown to play an important part in achieving national harmonisation of laboratory results, efforts should be made by manufacturers to ensure that their instruments are able to accomodate the more commonly used EQA procedures, and to achieve inter-instrument comparability of results.

Despite these limitations, it is reassuring to see the overall progress in reliability of various blood count parameters since NEQAS was established in the UK. This reflects refined instrumentation, more reliable calibration, standardised methods, etc. The reliability of haemoglobinometry in the UK is demonstrated by a CV of 1.5 - 1.8%. This is probably the minimum CV for this test as performed in the routine laboratory. It might be agreed that there is no need to continue to include tests that are so well performed in the regular surveys. However, in each survey there is an occasional blunder sometimes as a result of transcription error but also in some cases due to instrument failure or technical faults.

It is problematic whether NEQAS can claim responsibility for the improvement or whether it acts as a window, merely observing the improvement; there is no doubt however that it does provide the awareness when there is a problem and a stimulus for its resolution.

Submitted on invitation.
Accepted on 16 October 1994.

REFERENCES

1. LEWIS, S.M. 1995. Quality assurance programmes in the United Kingdom. *Ann. Ist. Super. Sanità* **31**(1): 53-59.

2. INTERNATIONAL FEDERATION OF CLINICAL CHEMISTRY/INTERNATIONAL COMMITTEE FOR STANDARDIZATION IN HAEMATOLOGY. 1987. On the theory of reference values. Part I. The concept of reference values. *J. Clin. Chem. Clin. Biochem.* **25:** 337-342.

3. PANTALONY, D., WOOD, D. & JACOBS, W. 1986. *Proficiency testing in haematologic morphology - a ten year experience.* 21. Congress of International Society of Haematology, Sydney. (Abstracts 250).

4. DACIE, J.V. & LEWIS, S.M. 1995. *Practical haematology.* 8th ed. Churchill Livingstone, Edinburgh. pp. 65-68.