J Clin Pathol 1993;46:1046–1050

Development of an internal quality assessment scheme in a clinical bacteriology laboratory

C E Constantine, M Amphlett, M Farrington, D F J Brown, S Messer, A Rampling, R E Warren

Abstract

***Aim*—To develop an internal quality assessment (IQA) scheme in a clinical bacteriology laboratory.**

***Methods*—Over 24 months, 1230 diagnostic specimens, representing 0·42% of laboratory workload, were anonymised and resubmitted for analysis. Six hundred and twenty one (48·7%) of these gave positive culture results; 44 fecal and upper respiratory specimens were "spiked" (artificially inoculated) to increase the proportion of positive samples.**

***Results*—Discrepancies between IQA and clinical sample results occurred in 188 cases (14·8%): 76·6% of these were in culture results, 13·3% in microscopy performance, and 10·1% in clerical recording. The culture discrepancy rate for each positive sample was lowest for wound (17·5%) and urine (18·1%) specimens, and highest for faeces (34·9%) and upper respiratory (37·7%) samples. Discrepancies in several areas responded to staff training and improvement in technical methods.**

***Conclusions*—An IQA programme of this type assesses the reproducibility of tests within a diagnostic laboratory when analysing common specimen types and organisms. It permits blind assessment of many areas of diagnostic work that are not readily amenable to other quality assurance methods, and it raises the awareness of all staff to the importance of quality in every aspect of specimen and data processing.**

(*J Clin Pathol* 1993;46:1046–1050)

Quality assessment (QA) is a system in which specimens of undisclosed content are introduced to a laboratory and examined by the same staff using the same procedures as they would normally use to examine patients' specimens of the same type.[1-3] External quality assessment (EQA) schemes bring advantages of unbiased testing and standardisation, comparison with laboratory peers, and dependable follow up of deficiencies. Only a limited number and range of test samples can be circulated in such schemes; for example, EQA specimens comprise only about 0·1% of workload in our laboratory. During the past 15 years UK National External Quality Assessment schemes (UKNEQAS) have developed to cover most areas of technical practice.[1 2] In contrast, formal internal quality assessment (IQA) has not featured prominently in the work of clinical microbiology laboratories.

In the United Kingdom recent impetus has been given to microbiological quality in its broadest sense by national initiatives such as laboratory accreditation,[4] value for money assessment by the Audit Commission,[5] and medical audit.[6] In late 1988 and 1989 the performance of the Clinical Microbiology and Public Health Laboratory, Addenbrooke's Hospital (CMPHL) fell below the national mean in the UKNEQAS categories of general bacteriology and acid-alcohol fast bacilli. Prompted by these developments we introduced a formal IQA scheme involving blind reprocessing of a proportion of routine diagnostic specimens. Our aim was to reprocess enough specimens to test the whole of the laboratory's service, without excessively increasing workload. This paper describes the implementation of the scheme, some modifications made as a result of our initial experience, the results during the first 24 months, and some changes to routine laboratory practice stimulated by IQA.

Methods

Twenty four and a half whole-time equivalent medical laboratory scientific officer (MLSO) and medical laboratory assistant (MLA) staff work in the bacteriology laboratory of the CMPHL. During the 24 months covered by this report, the CMPHL processed 301 266 bacteriological specimens.

The scheme was introduced in May 1989. Modifications to administrative details, to

Clinical Microbiology and Public Health Laboratory, Addenbrooke's Hospital, Cambridge, CB2 2QW
C E Constantine
M Amphlett
M Farrington
D F J Brown
S Messer
A Rampling
R E Warren
Correspondence to: Dr M Farrington

Accepted for publication 1 July 1993

Table 1 Discrepancy rates by specimen type in five assessment periods over 24 months

			Discrepancy rates (%):						
Period	Dates (month/year)	No of specimens submitted	Wound	Urine	Sputum	Faeces	Upper respiratory	Genital	Overall
1	05/89—10/89	202	14·7	22·5	13·3	28·6	16·0	10·7	17·8
2	11/89—12/89	122	33·3	15·9	7·1	11·1	11·1	31·6	18·9
3	01/90—05/90	298	1·8	12·9	17·5	25·0	8·1	17·5	12·4
4	06/90—11/90	351	20·8	15·0	20·0	0·0	14·0	2·0	13·4
5	12/90—05/91	301	11·6	16·0	24·4	17·4	11·4	9·1	15·0

methods of resubmission, and to the balance of specimen types submitted followed reviews in October and December 1989 (table 1). Subsequently the scheme ran without significant changes to established procedures. Further assessments were performed in May and November 1990 and May 1991. Two MLSOs 1 and the Quality Assessment Officer (QAO, MLSO 2 grade) ran the scheme part-time. One MLSO 3 and one consultant medical microbiologist scored and analysed results.

We resubmitted about 14 specimens each week to the routine diagnostic bacteriology laboratory. Resubmission occurred within 24 hours of processing the original sample, and whenever possible on the same day. The original sample was known as the "clinical" specimen and the reprocessed sample as the "IQA" specimen. Generally the IQA MLSOs 1 retrieved suitable specimens from refrigerated storage, assigned unique IQA scheme numbers, kept record sheets of the clinical and IQA specimen pairs, and wrote fresh request cards. Each card carried the IQA number, details of recent antibiotic treatment, and brief clinical information. Rephrasing of these details prevented easy recognition.

Large volume, cloudy, freshly voided urine samples were chosen for reprocessing on the day of receipt after numbering by the laboratory reception staff. They were mixed, then at least 5 ml was transferred to a sterile universal container. The IQA MLSOs 1 inoculated throat swabs to two sets of plates. One set was submitted for examination by the routine laboratory bench staff (the "IQA" plates), and the second set (the "reference" plates) were incubated and examined by the IQA staff who used the results for investigation of the causes of discrepancies.

Pus and fecal samples and other swabs were recovered from refrigerated storage after initial processing, transferred to fresh containers, and relabelled. Initial attempts to achieve consistent results when sputum samples were split between two universal containers on receipt were not successful. Therefore, sputum specimens were retrieved from refrigerated storage soon after homogenisation with dithiothreitol and samples were transferred to other sterile universal containers. After the first two assessment periods, IQA sputum samples were not homogenised again before reprocessing.

Some fecal and upper respiratory specimens were "spiked" with pathogens to increase the numbers of positive specimens of these types. This was unnecessary in other areas where a higher proportion of routine specimens gave naturally positive culture results. The QAO compared results from the "spiked" specimens with results of cultures set up by the IQA staff at the same time (known as "reference" plates).

Staff at each laboratory bench processed the IQA specimens according to routine methods, and entered the results on to the laboratory computer. IQA record sheets, request forms, and computer generated reports were returned to the MLSO 3 who compared the results of the IQA specimen with the clinical specimen result recorded on the laboratory computer.

For the period covered by this report, assessments of "culture" and "microscopy" were performed as shown in table 2. A specimen was considered "positive" if organisms were reportedly isolated from either the "IQA" or the "clinical" sample, or both, or if the specimen was "spiked". "Clerical" discrepancies included errors of specimen registration by clerical staff, and of result entry by MLSO staff.

The MLSO 3 noted details of inconsistent results on a form, and passed it to the MLSO 2 in charge of the appropriate section of the laboratory. After immediate investigation, the MLSO 2 filled in the report section of the form. The completed form was returned to the MLSO 3 who prepared a weekly graphical summary of discrepancy rates. This was circulated to the senior medical, MLSO, and scientific staff and a copy posted on the laboratory notice-board. At the end of each assessment period the consultant microbiologist and the MLSO 3 prepared performance summaries. These included tabulation of results for the whole period, and analyses of discrepancy rates overall and those for culture positive samples only. IQA results were also discussed regularly in laboratory technical meetings that were held weekly between senior medical, MLSO, and scientific staff, and proposals for improving routine laboratory methods and the IQA scheme itself came regularly from all grades of staff.

Results

Between May 1989 and April 1991, 1274

Table 2 Definitions of microscopy and culture discrepancies.

Specimen type	*Microscopy*	*Culture*
Wound	Gram stain presence *v* absence of pus cells or organism type	Presence *v* absence of named organism in report, including components of mixtures. Antimicrobial susceptibility result differences
Urine	Wet preparation ≥10-fold difference in quantitation of pus cells or red cells Crystals, casts seen *v* not seen	Presence *v* absence of named organism in report, including components of mixtures Antimicrobial susceptibility result differences. Antibacterial activity found *v* not found
Sputum	Gram stain presence *v* absence of pus cells or organism type	As wound
Faeces	Parasite concentration presence *v* absence of parasite	As wound
Upper respiratory	Not performed	As wound
Genital	Wet preparation presence *v* absence of *Trichomonas* sp or yeasts	As wound

specimens (including 44 "spiked" samples) were resubmitted to the diagnostic laboratory in the IQA scheme. This represented 0·42% of total workload. Overall, there were 188 discrepancies (14·8%) between IQA and clinical tests. Of these, 25 (2·0% of IQA samples) had discrepancies in microscopy, 145 (11·3%) in culture, and 19 (1·5%) in clerical recording. Reports from the "clinical" or "IQA" specimen included the isolation of named organisms in 577 of the specimen pairs. Including the "spiked" samples, 621 specimens were therefore classified as "positive" (48·7%). Overall, 23·2% positive specimens gave discrepant culture results.

Although the range of discrepancy rates seen among the specimen types was relatively small, the proportion of positive cultures submitted varied nearly threefold from 24·7% (genital specimens) to 71·9% (wound specimens). Furthermore, the culture discrepancy rate per positive sample varied over twofold from 17·5% (wound specimens) to 37·7% (upper respiratory specimens) (table 3). Overall, most discrepancies were between culture results (76·6%), but microscopy discrepancies were commonest with urine specimens (30·6% of discrepancies in that area) and genital specimens (14·3%), whereas clerical discrepancies were more of a problem with sputum (21·9%) and genital (19·0%) specimens.

For managerial reasons, the five assessments were made at irregular intervals during the first 24 months of the programme (table 1). Table 1 also shows the discrepancy rates for each specimen type among the assessment periods, and several trends are evident. Wide variations are seen in the results for wound and genital specimens, broadly consistent results for the urine and upper respiratory specimens, and small trends downward for fecal specimens and upward for sputum specimens. Over the five assessment periods there was a sustained decline from about 5% to 1% in microscopy discrepancies, and a small continuing downward trend in culture discrepancies from about 14% to 10%. Clerical discrepancies, however, rose from under 2% to 4% (data not shown).

In some areas there were substantial changes in performance related to the activities of the IQA scheme. For example, discrepancies in streptococcal recognition/grouping expressed as a percentage of the total number of streptococci recognised and grouped in the five time periods were 26·8%/8·8%, 11·1%/11·2%, 21·1%/0%, 23·1%/5·7% and 27·1%/0%, respectively. Thus although overall problems with streptococcal work fell only slightly, there was a large fall in the proportion of Lancefield grouping errors in the last three assessment periods. Discrepancies with yeast recognition (expressed as a percentage of the total number of yeasts reported in each assessment period) fell consistently from 50% to 16·6%, to 8·5%, and then 6·7% in the first four periods of assessment. In the last period, however, these discrepancies rose again to 23%.

Discussion

An overall discrepancy rate of 14·8% in an IQA scheme over a two year period is at first sight disturbing. It is especially noteworthy in the light of our laboratory's sustained good results in the UKNEQAS bacteriology schemes during the same period, with performance in all sections except parasitology remaining above the national mean throughout 1990 and 1991. In any IQA scheme based on specimen resubmission it is often impossible to define one result as "correct" when discrepancies occur. Therefore, we classified differences in technical interpretation between results of a sample pair as "discrepancies" rather than "errors". Reference plates were helpful when investigating the causes of discrepancies, but could not be used to define errors because they were read by senior staff under different conditions.

Each specimen resubmitted in IQA assesses more individual procedures than are performed on a single UKNEQAS sample, because microscopy and antimicrobial susceptibility testing are combined with organism isolation and identification in most cases, and each UKNEQAS specimen is tested only once. IQA therefore measures the consistency of repeated performances of the same test, whereas EQA measures the accuracy of a single test performance against an externally proved standard. Assessments of IQA results can be more locally sensitive, but also more stringent because allowances do not have to be made for variation in technical methods and clinical interpretation between different laboratories, and differences of quantitation are included.

IQA by blind testing of unknown samples is highly cost effective.[7] Surprisingly, we have been unable to find any other reports that have described resubmitting clinical

Table 3 Summary of performance in IQA scheme

Specimen type	*Wound*	*Urine*	*Sputum*	*Faeces*	*Upper respiratory*	*Genital*	*Overall*
Total number submitted	199	449	175	95	174	182	1274
Number "spiked"	0	0	1	24	19	0	44
Number discrepant	27	72	32	15	21	21	188
Discrepancy rate (% specimens)	13·6	16·0	18·3	15·8	12·1	11·5	14·8
Microscopy (number (% discrepancies))	0	22 (30·6)	0	0	NT	3 (14·3)	25 (13·3)
Culture (number (% discrepancies))	25 (92·6)	45 (62·5)	25 (78·1)	15 (100)	20 (95·2)	14 (66·7)	145 (76·6)
Clerical (number (% discrepancies))	2 (7·4)	5 (6·9)	7 (21·9)	0	1 (4·8)	4 (19·0)	19 (10·1)
Positivity (%)	71·9	55·2	50·9	45·3	30·5	24·7	48·7
Culture discrepancy rate/positive (%)	17·5	18·1	28·1	34·9	37·7	31·1	23·2

NT = not tested.

specimens as a basis for IQA in bacteriology. Only one partly comparable study of bacteriological QA has been published in which 623 simulated specimens were introduced blind to 19 Canadian laboratories.[8 9] The authors assessed only isolation and identification of bacteria, but the culture discrepancy rate at CMPHL of 11·3% between paired samples compares favourably with the reported error rate from the 19 laboratories of 42% for single examinations. This result prompted the authors to comment ruefully that "... bacteriology remains partly an art as well as a science ..." ! More studies are required to establish acceptable performance standards in this field.

Most discrepancies in our scheme (76·6%) were in the area of culture results. This was presumably because there are more technical and interpretative procedures needed to generate culture results than are involved in microscopy results or clerical performance. After the first assessment in October 1989, MLSO staff throughout the laboratory were given additional training in recognition of yeast and streptococcal colonies and in streptococcal grouping. Improvements were seen in the percentage of streptococci correctly grouped and in yeast colonies recognised. Little impact was made, however, on streptococcal recognition, and yeast discrepancies became more frequent again in period 5. This suggests that successful educative programmes must be repeated if improvements are to be maintained. Microscopy discrepancies were most common on the urine bench. Investigation of this during period 2 showed these were mainly transposition errors caused by poor positioning of the microtitre trays used to hold urine samples in the inverted microscope. Improving the equipment produced a sustained fall in microscopy errors, and this was responsible for most of the sustained improvement in microscopy results. Many of the discrepancies with culture results on the urine bench in period 1 were with the bioassay for antimicrobial activity which had just been changed from a well-diffusion to a spot-inoculum assay. Very few discrepancies were seen once MLSO staff had gained experience with the method. The high rates of clerical discrepancies with sputum (especially in period 5) and genital specimens, and the low rates with fecal and upper respiratory specimens remain unexplained.

Several practical features of specimen resubmission have emerged. Appropriate, but different, clinical details must be provided for the specimen pairs to ensure that the same tests are performed on both specimens. IQA and clinical specimen pairings would be more difficult to spot if the IQA sample was delayed by one or more days after the clinical sample, but relative organism numbers may alter after overnight refrigerated storage. Whenever possible we now resubmit all swab specimens on the same day as the original. To disguise pairings further we now choose multiple wound specimens of the same type for reprocessing on the same day. When we first introduced inoculation of "reference" plates they were usually read at the time of result collation, which was often one day later than the IQA and clinical specimen plates. This extra incubation apparently increased the recovery of certain components of heavily mixed growths. All plates are now read on the same day and all IQA culture plates are now stored in a refrigerator until matching of results has been completed. Although repeated inoculation of swabs does not reduce the density of streptococcal growth,[10] inclusion of selective media in the first set of inocula might inhibit growth on non-selective plates inoculated later. We have developed a formal scoring scheme that accounts for this possibility, and are currently assessing its value.

Roughly equal proportions of positive test samples should be submitted to each bench area, and this should be kept constant between assessment periods. Although negative samples from normally sterile areas can test only practical technical performance and clerical work, negative samples from areas with a normal flora do test technical recognition. "Spiking" of specimens is a partial solution, but the advantages of paired IQA and clinical sample submission are thereby lost, and "spiking" is not always technically straightforward. For example, we found some fecal samples to be naturally antagonistic to added salmonellae and shigellae.

All external schemes suffer from the disadvantage that EQA specimens are impossible to camouflage and may be treated differently. Black and Dorse proved the importance of this factor by submitting "mock" clinical samples indistinguishable from routine requests, and lyophilised specimens containing the same organisms.[9] More work was performed on the lyophilised specimens, and laboratories were much more successful at identifying their contents correctly.

IQA that involves reprocessing specimens has additional advantages to those of EQA schemes such as UKNEQAS. Larger numbers of samples can be handled, and flexibility and rapid local responsiveness are possible.[2] IQA can be used to monitor the whole laboratory from specimen registration through to issue of reports, and the process lends itself to the definition of performance standards. Certain types of test, specimen, and organism that are not suited to freeze-dried national distribution can be included, such as urine microscopy. Use of clinical material ensures that pathogen and normal flora quantitation is naturally appropriate.

By developing our own IQA programme, and by publicising the results while encouraging feedback of technical problems and suggestions from all levels of staff, awareness of the importance of quality has increased throughout the laboratory. We believe this approach leads to a constructive attitude among all staff, and avoids the adversarial nature of schemes that, for example, involve surreptitious introduction of specimens from fictitious senders. At the CMPHL the IQA

scheme is viewed positively by MLSO staff, and adverse comments have centred on the increased workload generated.

We have found it impossible consistently to ascribe most discrepancies to "methodological" or "interpretative" groups that might respond respectively to improved technical methods or to training and supervision. Thus although our current scoring scheme takes account of potential clinical importance, results are still collated quarterly and patterns of discrepancies sought by inspection. Other recent developments include: widening the use of specimen "spiking" to other laboratory areas and modifying "spiking" techniques; introducing IQA to other specimens and laboratory areas, principally blood cultures; and bringing medical microbiological comments on reports into review. These developments, the results of the scoring system, and an assessment of the costs of running the scheme will be the subject of a later report.

Although this IQA programme has been developed within a United Kingdom diagnostic laboratory to meet local needs, we believe the underlying principles and details of its organisation will be widely applicable. We have found the scheme to be a thought-provoking and powerful method for assessing the performance of the day to day work of the whole laboratory.

We are grateful to Diana Haynes and Enid Walpole for their expert technical work, to Dr P D Meers for helpful discussions, and to all the staff of the CMPHL for their enthusiastic participation in this scheme.

1 Snell JJS. External quality assessment. In: Snell JJS, Farrell ID, Roberts C, eds. *Quality control: principles and practice in the microbiology laboratory*. London: Public Health Laboratory Service, 1991:18–23.
2 Snell JJS, Hawkins JM. Quality assurance—achievements and intended directions. *Rev Med Microbiol* 1992;**3**: 28–34.
3 Bartlett RC. Quality assurance in the clinical microbiology laboratory. In: Balows A, ed. *Manual of clinical microbiology*. 5th edn. Washington, DC: American Society for Microbiology, 1991:36–43.
4 Audit Steering Committee. Royal College of Pathologists accreditation pilot study: a year later. *J Clin Pathol* 1991;**44**:172–3.
5 Audit Commission. *Pathology services: a management review*. London: HMSO, 1991.
6 Farrington M. Medical audit in clinical microbiology in the United Kingdom. *Rev Med Microbiol* 1992;**3**: 104–11.
7 Bartlett RC, Rutz CA, Konopacki N. Cost effectiveness of quality control in bacteriology. *Am J Clin Pathol* 1982; 77:184–90.
8 Black WA, Dorse SE. A regional quality control program in microbiology. I. Administrative aspects. *Am J Clin Pathol* 1976;**66**:401–6, 407–15.
9 Black WA, Dorse SE. A regional quality control program in microbiology. II. Advantages of simulated clinical specimens. *Am J Clin Pathol* 1976;**66**:407–15.
10 Ross PW. The isolation of Streptococcus pyogenes from throat swabs. *J Med Microbiol* 1977;**10**:69–76.