An assessment of the operation of an external quality assessment (EQA) scheme in histopathology in the South Thames (West) Region: 1995-1998

R L Carter, M G Cook

Abstract

Aims—To describe the design and organisation of a voluntary regional external quality assessment (EQA) scheme in histopathology, and to record the results obtained over a three year period.

Methods—A protocol is presented in which circulation of EQA slides alternated with teaching sessions. Procedures for the choice of suitable cases, evaluation of submitted diagnoses, and feedback of results to participants are described. The use of teaching sessions, complementary to the slide circulations, and dealing with current diagnostic problems is also outlined. **Results**—Participation rates in the nine slide circulations varied between 66% and 89%, mean 85%. Overall scores were predictably high but 4% of returns, from 10 pathologists, were unsatisfactory. These low scores were typically isolated or intermittent and none of the participants fulfilled agreed criteria for chronic poor performers.

Conclusions—This scheme has been well supported and overall performances have been satisfactory. The design was sufficiently discriminatory to reveal a few low scores which are analysed in detail. Prompt feedback of results to participants with identification of all "incomplete" and "wrong" diagnoses is essential. Involvement of local histopathologists in designing, running, and monitoring such schemes is important.

(*J Clin Pathol* 1998;51:910-913)

Keywords: external quality assessment; general histopathology

External quality assessment (EQA) schemes in clinical chemistry, microbiology, and haematology have been functioning in the United Kingdom for many years,¹ but the operation of similar procedures for histopathology continues to present problems.²⁻⁵ Schemes have been described to evaluate technical aspects of histopathology⁶ and to assess working procedures within individual laboratories in the form of internal or local audits^{7,8}; but it is in the appraisal of diagnostic ability that the difficulties become apparent. What is involved is essentially judgement of the interpretative skills of individual consultants, an evaluation of opinions rather than results. "Identical aliquots, objective instrumental evaluation, mean

and standard deviations do not exist for pieces of tissue as they do for tubes of blood."⁹

Papers on quality control in diagnostic histopathology appeared in the 1970s^{10,11} and in 1983 the Royal College of Pathologists advocated the wider use of slide clubs as a "means for encouraging quality assurance and medical audit in histopathology."¹² Three years later, the External Quality Assurance Forum in Histopathology, in which the Royal College of Pathologists was represented, recommended that regional EQA schemes should be introduced for all consultant histopathologists.¹³ Wide objectives were set out, together with a working protocol. (This proposed transition from slide club to formal EQA had, incidentally, been anticipated by histopathologists in south west England.¹⁴) Further appraisals of the need for, and conduct of, EQA histopathology were made^{15,16} and increasing numbers of schemes were set up. A valuable account, based on four years' experience, was published from the east of Scotland.¹⁷

In this paper we describe the format of an EQA scheme in histopathology used in the South Thames (West) region, and present results obtained over a three year period.

South Thames (West) Regional EQA scheme in histopathology

We prepared draft protocols and discussed them with local consultant histopathologists at regional meetings between November 1993 and May 1994. A final document was submitted to the (then) South West Thames Regional Health Authority in the form of an application for a research and development (R&D) grant. After external review the scheme was accepted and R&D funds awarded for two years, later extended for a third. One of us (RLC) served as co-ordinator. The protocol was ratified by the Royal College of Pathologists, and appropriate CME credits agreed.

The scheme was planned in the form of alternating slide sessions with examination of circulated cases, and teaching sessions in which current problems in diagnostic histopathology were discussed by speakers drawn, according to topic, from within the region or outside. All consultant histopathologists in the region were expected to take part in these activities, together with any long term locum consultants. Potential participants were assigned a random identifying number to ensure anonymity, which was strictly preserved. The scheme ran between May 1995 and March 1998. Between 41

Department of
Histopathology, Royal
Surrey County
Hospital, Guildford,
Surrey GU2 5XX, UK
R L Carter
M G Cook

Correspondence to:
Professor Carter.

Accepted for publication
11 August 1998

and 49 histopathologists were eligible to take part, based in 16 hospitals.

EQA SLIDE SESSIONS

The main features are set out below, and certain individual aspects are considered further in the discussion. Ten cases were circulated at each session, five being supplied in turn by each of two hospitals in the area. Almost all hospitals in the region thus contributed material during the three year time span. The cases were selected from two or more arbitrarily chosen consecutive working days by the coordinator, who visited the relevant departments, together with one or more of the local consultants. Care was taken to choose material that was representative of local practice, the only consistent constraint being that the block should contain an adequate volume of tissue. Clinical information was usually taken verbatim from the request forms, though grossly obscure or inadequate details were clarified. The original macroscopic descriptions were also recorded.

Twenty two unstained sections were prepared from each of the 10 cases and sent to the coordinator, who was responsible for making up complete sets and dispatching them to the participating hospitals together with the clinical documentation. One complete slide set was usually sent to each department. The designated first and last slides from each case were retained as a check for consistency of material, particularly from small or focal lesions. The slides were circulated between three and four weeks before a regional meeting. Answers from each participant, identified by their code number, were requested (and usually obtained) one week before the meeting.

The answers took the form of brief reports. A bald list of diagnoses was discouraged. The main differential diagnoses were required where appropriate. Indications for additional histopathological investigations to establish or clarify a particular diagnosis such as microbial stains, immunofluorescence, or immunohistochemistry were also required. Comments indicating a need to discuss individual cases with colleagues were acceptable. Participants also had the option to indicate personal "areas of exclusion" to cover types of material which they rarely, if ever, examined.

The individual results from each circulation, collated by the coordinator, were presented at the regular regional histopathologists' meeting. A simple scheme for scoring individual answers was used as follows: "correct and complete", 2; "adequate but incomplete", 1; "wrong", 0. The coordinator's reasons for scoring answers as "1" or "0" were set out at each meeting and were occasionally modified in the light of general discussion. It was agreed that three substandard scores in four consecutive slide circulations (see Results) would be taken to indicate chronic poor performance.

Within a week of each meeting, all histopathologists included in the scheme received two documents compiled by the coordinator: a summary of the overall results for the 10 circulated cases, and the participant's own indi-

vidual score sheet, identified by code number, which included a record of CME credits. All answers scored as "1" or "0" were annotated to indicate why such scores had been assigned. Potentially dangerous misdiagnoses were highlighted. Individuals who did not take part in a particular circulation still received both documents, their personal score sheet noting "no return received" and nil CME credits. Individual score sheets also came to be used to draw the pathologist's attention to any personal trend for unsatisfactory rates of participation and/or suboptimal scores as the number of circulations increased—see Results.

A policy with respect to chronic poor performers, as defined above, was agreed along the following lines. Any such identification by the organiser was first required to be confirmed by an additional, independent opinion from the College Regional Adviser who would convene a subgroup of two consultant histopathologists, one of whom worked outside the region. The need for a prompt response was emphasised. If the organiser's conclusions was confirmed, then a confidential "Dear Colleague" letter would be sent to the individual involved. The issue would be regarded as resolved if satisfactory results were obtained in the next two slide circulations. If the individual did not participate in the next two circulations or continued to perform unsatisfactorily, the organiser would send a second "Dear Colleague" letter and at the same time inform the College Regional Adviser, requesting that the matter should be referred directly to the College. It was never necessary to take this potentially very serious step, one consequence of which would be an inevitable loss of anonymity (see also Discussion).

TEACHING SESSIONS

These activities were regarded as complementary to the slide circulations. Likely topics were discussed beforehand with participants and were chosen with particular emphasis on common diagnostic problems. Five slide seminars were held which dealt with the following subjects: intraduct proliferations in the breast, cervical intraepithelial neoplasia, reactive lymphadenopathies, diagnosis of early prostatic cancer, and melanocytic skin lesions. Two of these sessions included previous circulation of slides. The results were collated by the coordinator and sent out to participants after discussion by the guest speakers. The scores of these more specialised cases were not included in the cumulative scores for the general slide circulations.

Anonymity was regarded as an essential feature of both aspects of this EQA scheme. An EQA secretary was responsible for initially assigning random identifying numbers to each participant, and the coordinator did not have access to the matched list of numbers and names. Results prepared by the coordinator were posted to individual pathologists by the EQA secretary. Correspondence with participants was occasionally required, usually to discuss marking of a particular diagnosis. Such letters were (by agreement) addressed to the

EQA secretary in the first instance and the coordinator replied in an anonymised "Dear Colleague" format.

All aspects of the scheme were reviewed annually with participants at a regional meeting.

Results

PARTICIPATION

Nine circulations of general EQA slides were completed. Between 41 and 49 histopathologists were eligible to take part, the numbers fluctuating because individuals retired or moved to new posts outside the region. The actual numbers of participants in each circulation varied between 31 and 42 (66–89%), with a mean of 85%.

An acceptable level of individual participation was proposed in the original protocol as submission of diagnoses from at least three of four general slide circulations. This approach was later modified for two reasons: the scheme ran for longer than was originally anticipated, and a shifting proportion of people joined or left it at various times. A reasonable view can, however, be obtained by examining participation rates among 43 individuals who were available to take part in seven, eight, or nine circulations. Of these, 17 (43%) participated in all of them compared with three pathologists who took part in three and one who only participated on a single occasion. These four individuals (9%) were manifest non-participants.

PERFORMANCE

Three hundred and thirty eight individual returns were received, of a possible total maximum of 410. The scores for each were expressed as percentages and grouped in three broad categories: 90–100, 71–89, and ≤ 70 , corresponding to "good," "adequate," and "unsatisfactory." The overall distribution of these categories in the nine circulations was as follows: 90–100 ($n = 249$) 74%; 71–89 ($n = 76$) 22%; ≤ 70 ($n = 13$) 4%. The range of variation for each of these categories across the nine circulations was as follows: 90–100 varied between 62% and 83%, 71–89 varied between 15% and 29%, and ≤ 70 varied between 0% and 7%. A general tendency for high scores was predicted, given the nature of the material circulated, but the scheme also identified a small group of low scores.

LOW SCORERS

Further analyses were made of the 13 low scores to determine whether they could be construed as evidence of poor performance, originally defined in this protocol as three sub-standard scores in four consecutive slide circulations. The 13 low scores, recorded as $\leq 70\%$, were distributed as follows: 70% (8), 65% (2), 60% (1), 55% (2). They were submitted by 10 participants: eight individuals produced one isolated low score: 70% (6), 65% (1), and 55% (1); one individual submitted two non-consecutive low scores in eight returns (55%, 70%), and another submitted three non-consecutive low scores in seven returns (65%,

60%, and 70%). Seven of these 10 participants were regular or reasonably regular participants in whom it was possible to assess overall patterns of performance. All seven sets of returns included high scores (90–100%) before or after the low scoring return, and six of the seven participants achieved high scores in at least 50% of these returns. Low scores were not concentrated in any one slide circulation.

Possible reasons for the low scores were assessed by examining the distribution of individual diagnoses scored as "1" (incomplete) or "0" (wrong) (see above) within each low scoring return. "Incomplete" answers were somewhat more frequent than "wrong" answers, at 31 (24%) *v* 27 (21%). The latter included six diagnoses which were regarded as potentially dangerous. These misdiagnoses were reported by five individuals in the low scoring group and their attention was specifically directed to them by the coordinator on the score sheets. It should, however, be noted that isolated, potentially dangerous diagnostic errors were also recorded from eight other participants whose returns achieved overall high scores in most or all of the slide circulations (see Discussion). A few unsatisfactory performances, typically isolated or intermittent, have thus been identified within the working context of this EQA scheme, but the individuals concerned do not fall within the agreed definition of chronic poor performers. Dispatch of formal "Dear Colleague" letters, as laid out in the protocol, was not therefore required.

Discussion

The format adopted for this EQA scheme, in which general slide circulations alternated with formal teaching sessions, does not appear to have been reported previously. The protocol for the EQA slide circulations was similar to that advocated by the Royal College of Pathologists¹³ and others^{17,18} but a few specific features may be noted. Involvement of the regional histopathologists in all stages of the scheme was considered to be of prime importance. Imposition of EQA schemes on what has been ironically described as "those on the diagnostic factory floor"⁴ is not an acceptable approach, even if it could be made to operate. Local involvement in the present scheme took the form of discussions of the initial protocol design, review of results from each slide circulation, and an annual assessment of the EQA scheme as a whole. The choice of material was sometimes slanted by the coordinator to include cases that were relevant to a recent teaching session or to discussions at a previous review meeting. A simple manual scoring system (as published before^{13,17}) was used, rather than a computerised scheme.¹⁹ Predominantly high scores were anticipated, given the type of material circulated, and this pattern was confirmed; but the system also revealed small numbers of low scores, suggesting that reasonable discrimination had been achieved.

Several general questions are raised again in this account, one of which is the approach to "wrong" answers in EQA slide circulations.

They represent a spectrum ranging from fine morphological distinctions between different entities which are clinically inconsequential to major errors which would be of direct clinical significance. The latter are rarely caused by a failure to recognise an obscure entity: "the plain truth is that most errors that arise in diagnostic practice involve what must be regarded as relatively mundane material."¹⁶ (It is this mundane, or familiar, material that should form the basis for general EQA schemes rather than the more unusual cases which are better suited for slide clubs.) One disquieting, but perhaps not surprising, finding from the present scheme was that occasional wrong diagnoses with potentially serious clinical implications appeared in returns which (overall) still achieved "good" or "adequate" scores. Negative scoring¹³ for such diagnoses would not have revealed them, and results from "good performers" still need to be carefully assessed.

It is, however, the "poor performers" who continue to attract debate. How should they be defined? How far can extrapolation be made from the contrived context of an EQA scheme to the conditions of daily diagnostic practice? And once identified, what action should be taken? The reasons for poor performances are diverse and the increasing pressures from growing diagnostic workloads, depleted staffing levels, and extra administrative requirements are self evident. The conclusion that chronic substandard performances in an EQA scheme may indicate a genuine diagnostic problem is the most that can reasonably be drawn. They cannot, however, be ignored, and a considered response by the local coordinator should be made. All inadequate or wrong diagnoses in every circulation were identified to each participant in this scheme on their score sheets (irrespective of overall "good," "adequate," or "unsatisfactory" scores). The few low scorers have been fully described here but, as already noted, they did not fall into the defined category of chronic poor performers. It was thus never necessary to use the system of graded "Dear Colleague" letters which was described in the protocol. There is no question that a precipitate or ill judged response to low scores in an EQA scheme could cause devastating professional and personal damage to the individual in question. The manifest injustice of this situation is enhanced by the fact that participation in local EQA schemes is still voluntary, though mandatory requirement to take part is expected within the next two or three years.²⁰

Other topics of current concern include the funding of histopathology EQA schemes and measurement of their outcome or effectiveness.^{21 22} Funding for the scheme reported here by regional R&D money is unlikely to be generally applicable, and it is essential that the appropriate NHS trust managers include EQA when planning resources for histopathology. (They also need to understand the applications, advantages, and limitations of such activities.) Any valid "measurement" of the outcome of EQA schemes is very difficult to envisage, and qualitative end points

are equally hard to define. Improved diagnostic accuracy in histopathology clearly contributes to the care and clinical outcome of patients, often decisively, but it is only one among several other determining factors.

These and other limitations of current EQA schemes in histopathology must, however, be put in context. Experience with such schemes is relatively short and the subject is still developing.²³ A new sector of CPA (UK) Ltd—CPA (EQA)—has recently assumed responsibility for "oversight" of EQA schemes in histopathology.²⁴ One consequence may be the emergence of more standardised protocols; alternatively, stimulating opportunities might arise for comparing various different formats for histopathology EQA across the country.

We thank Mr R Stamp MRCPATH and Mrs Sally Bishop of the Research and Development Directorate of the former South West Thames RHA (now NHS Executive South Thames) for helpful discussions in planning this EQA scheme.

- Whitehead TP, Woodford FP. External quality assessment of clinical laboratories in the United Kingdom. *J Clin Pathol* 1981;34:947-57.
- du Boulay C. Continuing medical education for pathologists: an evaluation of the Royal College of Pathologists' Wessex pilot scheme. *J Clin Pathol* 1997;50:1022-6.
- Furness P. Diagnostic histopathology external quality assessment—a threat or a promise? *ACP News*, Summer 1997:27-9.
- Slater D, Harvey L. Histo- and cytopathology EQA schemes—poor indicators of poor performance? *ACP News*, Winter 1996/97:12-14.
- Underwood J. The future of EQA in histopathology: education or performance surveillance? *ACP News*, Winter 1988:16-19.
- Reynolds GJ. External quality assurance and assessment in immunocytochemistry. *Histopathology* 1989;15:627-33.
- Ramsay AD. Locally organised medical audit in histopathology. *J Clin Pathol* 1991;44:353-7.
- Thomas J St J, Lessells AM, McIntyre MA, et al. Prospective study of quantitative aspects of audit in a large general histopathology laboratory. *J Clin Pathol* 1991;44:928-31.
- Morson BC, Frei JV. Medical audit of rectal biopsy diagnosis of inflammatory bowel disease [letter]. *J Clin Pathol* 1982;35:476.
- Owen DA, Tighe JR. Quality evaluation in histopathology. *BMJ* 1975;i:149-50.
- Langley FA. Quality control in histopathology and diagnostic cytology. *Histopathology* 1978;2:3-18.
- Royal College of Pathologists. Quality assurance and medical audit in histopathology. *Bull R Coll Pathol* 1983;42:9.
- Royal College of Histopathologists. External quality assessment scheme in histopathology. *Bull R Coll Pathol* 1986;56:3-4.
- Sherwood AJ, Hunt AC. An external quality assessment scheme in histopathology. *J Clin Pathol* 1984;37:409-14.
- Lee FD, Burnett RA. Quality assurance in histopathology. *J Pathol* 1987;152:247-51.
- Lee FD. External quality assessment in histopathology: an overview. *J Clin Pathol* 1989;42:1009-11.
- Klys HS, Lessells AM. External quality assurance in histopathology: experience of the east of Scotland scheme. *J Clin Pathol* 1992;45:288-91.
- Furness P. Standard operating procedures for diagnostic histopathology external quality assessment schemes. *Bull R Coll Pathol* 1995;92:22-5.
- Furness PN, Lauder I. Response analysis in histopathology external quality assessment schemes. *J Clin Pathol* 1993;46:357-63.
- Royal College of Pathologists. Continuing Medical Education. *Bull R Coll Pathol* 1998;102:28-9.
- Flynn FV. The Royal College of Pathologists' CME scheme. *Postgrad Med J* 1996;72(suppl 1):S24-6.
- Bignall J. Continuing medical education: shibboleth or necessity? *Lancet* 1998;351:1004.
- du Boulay C. Audit of continuing medical education for pathologists: strategies and implications. *J Clin Pathol* 1996;49:100-1.
- Department of Health. *Oversight of provision of external quality assessment schemes in histopathology, cytopathology, cytogenetics and molecular genetics for pathology laboratories*. London: DOH, 1998:EL(98)2.