Audit in histopathology: Description of an internal quality assessment scheme with analysis of preliminary results

J A Zuk, W E Kenyon, M W Myskow

Abstract

In the first six months of a formal Internal Quality Assessment Scheme operating in the Department of Histopathology, Broadgreen Hospital, Liverpool, 1005 items of data were gathered, relating to 80 cases. The scheme entails a random 2% sample of biopsy specimens being selected, each case being analysed using a structured proforma, and a numerical scoring system being allocated to all aspects of specimen handling. Technical and secretarial performance was good while the quality of clinical information provided by the requesting clinician was poor. There was a wide variation in reporting times, which related partly to the complexity of the specimen, and partly to the degree of supervision required by the reporting pathologist. Month by month analysis of reporting times showed a significant increase in reporting times associated with rotation of junior staff, but not with periods of annual leave. Pathologist performance scores were good, but close examination of components of the overall score for an individual pathologist indicated occasional areas of weakness (such as the adequacy of the macroscopic report).

It is concluded that this scheme is worthwhile and its practice will be continued indefinitely. The comprehensive nature of the analysis allows for the formal identification of areas of work which need improvement, and the allocation of a formal numerical score allows improvement in these areas to be monitored. The monthly meetings provide a means whereby performance scores are fed back to the participating pathologists, and they are also of general educational value regarding histological reporting practice. It is intended that the scheme be extended to include the assessment of special stains, frozen sections, and adequacy of the report delivery service. The system is easily adaptable for use within other histopathology departments.

It is vital that high standards of histopathology reporting are achieved and maintained, with appropriate measures being introduced towards improvement where required. The

report entered into the patient's notes is still widely held to be of "gold standard" and often includes the primary diagnosis on which patient management depends.¹

External Quality Assessment Schemes (EQAS) contribute towards the maintenance of standards, but such schemes have been introduced relatively slowly into histopathology compared with other disciplines, although one which assesses technical performance in immunohistochemistry has existed for some time.² Such EQAS are also of educational value and help towards uniformity and consistency of reporting. The requirement for maintenance of the quality of our services has been recently reinforced by proposals for the reorganisation of the NHS,³ which will create a more competitive atmosphere and will also emphasise the need for all doctors to participate in medical audit.⁴

EQAS arguably evaluate the maximal potential standard which any participating laboratory or individual is capable of reaching, much like the situation of an examination. This contrasts with internal quality control (or assessment) which is concerned with frequent, regular review of all laboratory procedures.¹ It examines what is actually happening within these departments under normal functional circumstances. Because many histopathology departments include junior medical staff, it could be anticipated that the maintenance of standards of various aspects of a pathologist's work, in particular surgical reporting, is likely to prove difficult to achieve. Furthermore as is clearly stated in the recently published Royal College of Pathologist's *Code of Practice for Pathology Departments*, the responsibility for the maintenance of standards of all aspects of work lies with the head of the department.⁵

We therefore decided to introduce an Internal Quality Assessment Scheme (IQAS) to evaluate and monitor the quality of all aspects of work in our department, with emphasis on the performance of the histopathologist. As well as having other benefits, we hoped our scheme would highlight any deficiencies within our service which would be amenable to change and improvement, this being a major criterion of effective medical audit.⁶

Scheme design

THE DEPARTMENT

Our Scheme was devised and developed to meet the particular needs of the Histopath-

Department of
Histopathology,
Broadgreen Hospital,
Thomas Drive,
Liverpool L14 3LB

J A Zuk
W E Kenyon
M W Myskow

Correspondence to:
Dr Myskow

Accepted for publication
29 August 1990

BROADGREEN HISTOPATHOLOGY INTERNAL QA ASSESSMENT CHECKLIST		Assessment Date..... Month Assessed..... Assessor..... Pathologist(s) Assessed..... Lab. Number.....	
COMMENTS		SCORES	
I <i>Assessment of Requesting Consultant</i> (a) <i>Consultant's name</i> (b) <i>Request Form</i> – clerical details 			

Assessment checklist

ology Department of Broadgreen Hospital, a designated teaching hospital and district general hospital for East Liverpool, with a workload for the year ending March 1989 of 7800 requests. Medical staffing consists of two consultants and usually one each of senior house officer, registrar, and senior registrar (all juniors except SHO rotate to other hospitals).

All aspects of work are fully computerised, the system having been developed and piloted at this site and designed to run in conjunction with the McDonnell Douglas Patient Information System (Homer; McDonnell Douglas, Hemel Hempstead, England). On receipt, individual specimens are categorised as follows: urgent, taken out of main stream and handled individually (category I); priority, inpatient diagnostic biopsies (category II); routine, all others (category III). For analysis purposes, category III is divided into biopsy specimens and resections. Categorisation was introduced to take account of the fact that specimen types I and II are those on which immediate patient management decisions are most frequently made.

ASSESSMENT SCHEME

Operating regularly since July 1989, this consists of a monthly retrospective analysis of a random 2% of surgical cases. These are selected from the month falling six weeks before the assessment date, and each month is con-

secutively analysed. The figure of 2% was considered to be high enough to be representative while still being a compromise between the detail of our analysis and the practicality of its implementation in view of the time required.

A nominated person (usually the senior registrar) coordinates the Scheme, organises case selection and retrieval of all material relevant to each case, including histological sections, the report, and a computer printout of itemised data concerned principally with the recorded times of key events involved in report generation as follows:

- (a) request date (REQ)
- (b) date and time slides first presented to pathologist (TO PATH)
- (c) date and time report presented for typing (TO TYPE)
- (d) date and time report typed (TYPED)
- (e) date and time final report is signed and approved by pathologist (REP APPR).

Cases are then distributed to one of several possible auditors, these being pathologists from within our department not involved with the case being reviewed and who had at least extensive post-primary MRCPath experience. An assessment checklist (figure) for completion accompanies each case. Fourteen variables are individually scored, the scheme emphasising analysis of pathologist performance (nine variables, three concerning key times in report

generation) but also assessing performance of requesting clinician (clerical and clinical details supplied), MLSOs (section and staining quality), and typists.

All aspects of the pathologist's routine daily performance in diagnostic histopathology are scored from initial macroscopic description and block selection to clarity and content of microscopic report, together with final diagnosis reached, accuracy of SNOMED usage, and accuracy of patient clerical details on final signed report. Scores are weighted to account for the magnitude and seriousness of omissions or errors. Total scores are calculated, the maximum being 7, although for some specimens this is reduced if the assessment categories of block selection and macroscopic description are not applicable.

Several key time intervals are also calculated from computer data and recorded (with one day deducted for each Saturday, Sunday, and public holiday). Any action recorded as having occurred later than 16.30 h was regarded as having happened first thing next morning, as 16.30 h is the deadline for delivery of completed reports from our office.

Auditors are given seven to 10 days to complete their assessment after which all medical staff meet to discuss the cases, allowing for direct feedback of performance which takes about an hour a month. Information relating to technical and clerical performance is presented to laboratory staff immediately afterwards, as required.

Preliminary results

The following is an analysis of preliminary results obtained from the first six months of our Scheme. One thousand and five items of data derived from 80 cases were analysed. A similar analysis will be performed on a regular six monthly basis. Specimen categories were as follows: category I (one case); category II (10 cases); category III (69 cases). For purposes of analysis, category III was divided into biopsy specimens ($n=44$) and resections ($n=25$). Student's t test was used for statistical analyses.

REQUESTING CLINICIAN

Specimens were received from 27 individual sources. The overall average score for clinical details supplied was 67.5% (averages for individual sources ranging 25–100%). Clinical details were good in 50% of cases, adequate in 35%, and inadequate or absent in 15%.

For clerical details, the average score for all 27 sources was 81% (ranging from 50–100% for each individual). This information was incomplete or inaccurate in 19% of cases.

TECHNICAL PERFORMANCE

In 79 cases there was sufficient material to assess quality of histological sections, with an average score of 78%, section quality being assessed as good (81% of cases), adequate (16%), and poor (3%). Fifteen cases were thus determined to be of suboptimal quality for the following reasons: "holes" ($n=5$); folds ($n=5$); debris ($n=2$); "chatters" ($n=1$); "scores" ($n=1$); and absence of full transverse section

Table 1 Total specimen scores obtained (expressed in percentages)

Specimen type	Average score (whole group of pathologists)	Range of averages for individual pathologists
All specimens (80)	87	80–91
Category I (1)	100	—
Category II (10)	92	85–100
Category III (44) (biopsy)	87	82–96
Category III (25) (resection)	84	76–97

($n=1$). Suboptimal sections were from 10% of category II specimens, 14% of category III (biopsy), and 36% of category III (resection) specimens, this reflecting the nature of tissue cut and larger blocks numbers derived from resections.

Quality of staining (haematoxylin and eosin and special stains) was evaluated in 66 cases (omitted in first, pilot month of study), with an average score of 92%. Quality was assessed as only adequate in 8% and poor in a further 2% of cases. Monthly variability for section and staining quality was 62–92% and 79–100% respectively. There was quite random monthly fluctuation—that is, there was no obvious change in performance as the scheme evolved.

PATHOLOGIST PERFORMANCE

The 80 cases were reported by seven pathologists—two consultants, one senior registrar, three registrars and one SHO, the registrar numbers reflecting the rotation of junior staff. Of the 63 cases reported by juniors, 40 of 44 diagnoses reached and reports issued by SHO and registrars, and two of 19 cases reported by the senior registrar were under direct consultant supervision. The remaining 9% of cases reported by SHO and registrars were under senior registrar supervision. Macroscopic description and block selection were almost always unsupervised.

(i) Pathologist scores

The overall average total scores for the group of seven pathologists with ranges of average scores for individuals are detailed in table 1. There is a general trend between scores and specimen type with better average performance for priority specimens (categories I and II), but this was insignificant (statistics not applicable to category I as only one case). Greatest inter-individual variability in performance was for category III (resection) specimens. For all data

Table 2 Scores for individual variables of pathologist performance (expressed in percentages)

Variable analysed	Number of cases to which this applies	Average score (all pathologists)	Range of average scores for individual pathologists
Microscopic clarity and content	80	79	67–95
Macroscopic clarity and content	31	71	33–100
Final diagnosis	80	89	67–100
SNOMED	80	92.5	67–100
Clerical detail check	80	96	84–100
Block selection	31	94	86–100

Table 3 Average individual variable scores by specimen category (expressed in percentages)

Variable analysed	Specimen category		
	II	III Biopsy	III Resection
Microscopic clarity and content	90	81	70
Diagnosis	100	84	92
SNOMED	90	91	96
Clerical detail check	90	98	96

relating to performance, only one serious error was found ("serious omission" for macroscopic description of category III resection specimen). More detailed analysis of the individual variables of pathologist performance which contribute to the final report was performed (table 2) in an attempt to identify specific areas of weakness. Despite the range of average individual scores for a final diagnosis of 67–100%, there was never any serious disagreement from the diagnosis stated, and even minor deviation from excellence was down marked as we consider this to be the most important part of the entire report.

A wide range of average scores for individual pathologists was found for microscopic clarity and content (range of 28%) and even more so for macroscopic clarity and content (range of 67%). The latter accounts for the wide range of average individual pathologists' scores for category III resection specimens (table 1), an area obviously requiring improvement.

Accuracy of SNOMED categorisation also shows substantial interindividual variability (table 2) which may reflect junior staff's lack of familiarity with the system. The figures for clerical detail check clearly indicate unnecessary individual carelessness at the vital stage of final report checking before signing.

Variables of pathologist performance applicable to all specimen categories were further analysed according to category (table 3). There exists an obvious trend between average scores for microscopic clarity and content and

specimen type, with a lower average score for resections compared with category III biopsy specimens (not significant; $0.5 \leq p \leq 0.1$) compared with category II specimens. Possible relevant factors accounting for this include longer report length of microscopic report for resections and the fact that all staff are acutely aware of the prime importance of category II specimen reports.

For all specimen types, average scores for diagnosis were quite consistent and high. SNOMED and clerical detail check errors were slightly more common for category II specimens, which may reflect the introduction of errors as the pathologist attempts to expedite these important reports.

(ii) Pathologist reporting times

The figures presented in table 4 concern the total time interval between the time a request is made and the time that report is completed and approved by the pathologist. This shows a wide absolute range (one to 10 days) for all specimens and is at best a very crude indicator of performance as it is subject to many variables beyond the pathologist's control (fixation, processing, and especially specimen delivery). As expected, shorter times were found for higher priority specimens. We provide a service for our own hospital and for two peripheral hospitals and general practitioners, with an average of 3.7 days (request date to report approved) for specimens from within our own hospital and 4.3 days for all other sources.

The calculated time interval, which is a direct measure of pathologist performance, is that time interval between when sections are first presented until the time the written histology report is presented for typing (TO PATH/TO TYPE) (table 5). Again there is an obvious trend, with more rapid reporting of higher priority specimens (categories I and II) and for category III biopsy specimens (not significant) as would be anticipated, as this was the reason for the initial introduction of specimen categorisation. Although these seem excessive—for example, category II—average 1.3 days—as previously stated, any action occurring after 16.30 h was recorded as having occurred the following day, and many reports are submitted for typing after this deadline.

Our average figures for all pathologists for the reporting interval TO PATH/TO TYPE are interpreted as satisfactory. For each specimen type, however, there exists considerable interindividual variation for this calculated interval. At all times report quality remained top priority, and we endeavour to strike a balance between quality and speed. Fifty seven per cent of all cases were reported by juniors under direct supervision and the effect of such supervision on the reporting time TO PATH/TO TYPE was therefore further analysed (table 6). For supervised and unsupervised specimens there was a general correlation between specimen priority and speed of reporting, in terms of both average time and mode. We interpret the necessary delay introduced in reporting category III

Table 4 Time intervals between request date and date and time of report approval (expressed in days)

Specimen type (number of specimens)	Average for all pathologists	Range of averages for all pathologists	Mode	Absolute range
All (80)	4.0	3.3–4.4	3	1–10
Category I (1)	1.0	—	—	—
Category II (10)	2.6	2.2–3.0	2	2–4
Category III (44) biopsy	3.8	3.0–4.2	3	1–7
Category III (25) resection	5.0	4.2–6.5	3	3–10

Table 5 Time interval (TO PATH/TO TYPE) (expressed in days)

Specimen type	Average for all pathologists	Range of averages for individual pathologists	Mode	Range
All	1.7	1.0–2.5	1	0–7
Category I	0	—*	—	—
Category II	1.3	0.0–2.0	1	0–3
Category III biopsy	1.5	1.0–2.1	1	0–4
Category III resection	2.1	1.2–4.0	1, 3	0–7

*NB Only one specimen in category I.

Table 6 Time interval (TO PATH/TO TYPE) comparing supervised and unsupervised reports (in days)

Specimen type	Average time		Mode		Range	
	Unsupervised	Supervised	Unsupervised	Supervised	Unsupervised	Supervised
All	1.3	2.0	1	2	0-4	0-7
Category II	0.75	1.67	1	2	0-1	0-3
Category III biopsy	1.25	1.79	1	2	0-4	0-4
Category III resection	1.55	2.43	0	1,2,3	0-4	0-7

Table 7 Average monthly scores for pathologist performance according to specimen type (expressed as a percentage)

Month of study	Average score			
	All specimens	Category II	Category III biopsy	Category III resection
May 1989	75	80	81	71
July 1989	91	—	89	100
Aug 1989	81	95	77	71
Sept 1989	89	100	92	84
Oct 1989	87	95	86	86
Dec 1989	94	90	92	96

(biopsy and resection) specimens as reasonable, but find the approximate doubling of this average time interval for category II specimens unacceptable. All three supervisors performed similarly with little interindividual variability in average scores and reporting times for the specimens they supervised.

(iii) Effect of stage of study

For each month of the study, the average scores for all pathologists for all specimens, subdivided according to specimen type, are presented in table 7. The monthly variability for average scores was 75–94% for all specimens (80–100% for category II, 81–94% for category III (biopsy) and 71–100% for category III (resections). We interpret this as a reasonable degree of variability, and there was no obvious trend between scores in any specimen category and month of study, except that all scores tended to be low for the first (pilot) month of study (May 1989). Since our scheme was introduced in July 1989, it suggests that it has proved beneficial to pathologist performance, but this requires further study.

The average monthly pathologist reporting times (TO PATH/TO TYPE), subdivided according to specimen category, are shown in table 8. For all specimens together, the monthly variability was 1.3–2.2 days (category

II 0–2 days, category III biopsy specimen 1.1–1.8 days, and category III resection specimen 1.2–4.3 days). Longest average reporting times were in the month of October, being accounted for by category II (average of two days) and in particular category III resection specimens (4.3 days). This coincided with our acquisition of two new junior pathologists in October which may account for this observed report delay. Confirmation of this explanation will depend on an analysis of future trends.

The average of 2.0 days for category II specimens during October 1989 is unacceptably high, requiring a change in practice, around the time of rotation of new Registrars into the Department each October. Note, however, that scores for performance (table 7) were maintained during this period.

TYPIST PERFORMANCE

One of the major factors determining final times of report generation from any department is typist performance (table 9). We interpret our typists' performance as being excellent, with 86% of all reports typed before 16.30 h on the same day the pathologist's report is presented to them. All priority specimens are typed the same day, which necessarily introduces delay for the other specimen types.

PERFORMANCE OF AUDITORS

Finally, in an attempt to confirm the validity of our assessment and scoring system, the marking performance of all four auditors was studied (table 10). As in any marking system, subjective interindividual variation in scoring will be found. Greatest degrees of such variability were found for the assessment of the following: clinical information supplied; staining quality; macroscopic description; and use of SNOMED. This variability reflects, at least partly, personal preferences of individual auditors. During our monthly audit meetings, however, discussion is aimed at encouraging reasonable uniformity in the assessments made by the auditors.

There was little interindividual variability in the marking of variables of final diagnosis (range of 16%), macroscopic clarity and content (13%), and clerical information check. Because we regard these as the most important qualitative aspects of all reports issued and as these key variables of pathologist performance were marked quite consistently, we conclude that our assessment and scoring system is indeed valid.

Discussion

We introduced our IQAS as a method of medical audit whereby we could assess, monitor, and evaluate the histopathology service we provide. Our process covers the performance of pathologists, technicians, typists and the clinicians initiating requests. It is therefore in some ways similar to a system already described,⁷ but differs in the greater depth of analysis undertaken in our scheme.

Our system fulfills the major principles of

Table 8 Average monthly specimen reporting times (TO PATH/TO TYPE) according to specimen type (expressed in days)

Month of study	Average reporting time			
	All specimens	Category II	Category III biopsy	Category III resection
May 1989	1.9	0	1.8	2.0
July 1989	1.8	—	1.7	2.5
Aug 1989	1.3	1.25	1.1	2.0
Sept 1989	1.3	1	1.4	1.2
Oct 1989	2.2	2	1.7	4.3
Nov 1989	1.6	1.5	1.5	1.7

Table 9 Typist performance

Specimen type	Day report typed (%)	
	Same day	Following morning
All specimens	86	14
Category II	100	0
Category III biopsy	86	14
Category III resection	80	20

the process of effective audit⁸ in that it is relevant, objective, quantified and repeatable with formal identification of areas requiring improvement. It also provides a method for reassessment of performance once appropriate changes have been implemented. As with all valid audit systems, this final step of reassessment is of utmost importance, resulting in the "closing of the audit feedback loop".⁶ We intend performing the monthly assessment process as routine on an indefinite basis, with statistical analysis of all data accumulated over each six monthly period, as we illustrate here.

We have no doubt that the introduction of our IQAS has been a worthwhile activity with its benefits greatly outweighing its main disadvantage, which is the use of valuable time. Because all activities within the department are being monitored, this in itself is a major incentive towards optimal performance by all concerned. With slight modification the scheme is adaptable for use within other histopathology departments, and already others from within our region have expressed interest.

Our study has provided verification that, in our opinion, most pathologists' activities in our department are performed satisfactorily in terms of both accuracy and speed. The system has allowed for the formal identification of the need for improvement and action required in certain areas of work. Our monthly meetings provide a means whereby there is direct feedback of these faults to the appropriate pathologist. Specifically, we have identified important individual weaknesses in clarity and content of the microscopic report, and even more so in macroscopic description of resection specimens. Appropriate action required includes further education and closer supervision, as well as making individuals aware of their particular weaknesses. We have also identified unnecessary carelessness by certain people in the use of SNOMED coding and, even more importantly, clerical detail checking, especially for category II (priority biopsy)

specimens. Alerting the pathologists to these particular faults should in itself result in improvement. The speed of completion of supervised category II reports, particularly at times of acquisition of new junior staff, requires improvement, and supervisors are now well aware of the need for more rapid and punctual supervision. Reanalysis of future results will enable us to check that improvement has occurred in the correct directions.

As well as highlighting faults, our monthly meeting is of general educational value, often leading to wider discussion of certain issues. It helps towards uniformity and consistency of reporting, which is important with junior staff rotating through the department. The possibility of introducing set reporting formats for most specimen types has been raised, these formats being written guidelines which would act as a check to ensure that all necessary positive and negative points are included in a report, but without completely abolishing individual style.

One of the major findings of this study has been the formal identification of poor performance of certain clinicians, in terms of both clerical and clinical details supplied on request forms. As well as being potentially dangerous, such practice wastes the time of clerical staff and pathologists. About one fifth of all requests had inaccurate or absent clerical details, but this practice is by no means restricted to our hospital.⁹ As always, it lies with the pathologist to educate clinicians in the dangers of this and we intend presenting the actual figures at our hospital "grand round" on a six monthly basis. Letters stating the need for improvement could also be sent to poor performers.

In future we intend to maintain the same general format of our IQAS, but our system is not intended to be static. In addition to our current practice, we will also undertake the following:

- (i) A retrospective analysis of performance during the six months before the introduction of our scheme to establish whether its implementation has resulted in improvement.
- (ii) Extensive assessment of performance of reporting category I (high priority) specimens (as only one was included in this survey); assessment of frozen sections.
- (iii) Assessment of reporting a randomly selected proportion of cases from each body region, systematically and consecutively.
- (iv) Analysis of the use of further work by the pathologist (special stains, immunohisto-

Table 10 Performance of auditors

	Variable scored											
	Requesting consultant			MLSOs			Pathologist					
	Total	Clerical	Clinical	Total	Section Quality	Stain Quality	Total	Micro	Macro	Diagnosis	SNOMED	Clerical
		Information	Information									Blocks
Range of average scores given by auditors	62-77	80-88	50-76	75-92	72-86	75-100	83-91	73-86	58-91	84-100	80-100	88-100
												83-100

chemistry, etc.) in terms of its contribution to final diagnosis and effect on reporting times.

- (v) Monitoring our delivery services, both to and from our department, especially as one of the major findings of the Royal College of Pathologists' pilot scheme in laboratory accreditation was delay in the report despatch for no good reason.⁹

1 Lee FD. External quality assessment in histopathology: an overview. *J Clin Pathol* 1989;42:1009–11.

- 2 Reynolds GJ. External quality assurance assessment in immunohistochemistry. *Histopathology* 1989;15:627–33.
- 3 Department of Health. *Working for Patients*. London: HMSO, 1989.
- 4 Department of Health. *Working for Patients. Medical audit (working paper No 6)*. London: HMSO, 1989.
- 5 Royal College of Pathologists. *Codes of practice for pathology departments*. London: RCP, 1989:7.
- 6 Smith T. Medical audit. *Br Med J* 1990;300:65.
- 7 Ramsay AD, Gallagher PJ. Quality control of surgical pathology in peer review – the Southampton experience. *J Pathol* 1989;158:343A.
- 8 Shaw CD, Costain DW. Guidelines for medical audit: seven principles. *Br Med J* 1989;299:498–9.
- 9 The College Accreditation Steering Committee. Royal College of Pathologists' United Kingdom pilot study of laboratory accreditation. *J Clin Pathol* 1990;43:89–91.