

Feasibility of new prognostic classification for rectal cancer

S A C DUNDAS, R W LAING, A O'CATHAIN, I SEDDON, *D N SLATER,
T J STEPHENSON, J C E UNDERWOOD

*From the Departments of Pathology and Community Medicine, Sheffield University Medical School, Sheffield, and the *Department of Histopathology, Rotherham District General Hospital, Rotherham*

SUMMARY Sixty slides from 60 blocks taken from 30 colonic carcinomas were circulated twice to six histopathologists of varying experience. Five of the six pathologists showed a good to excellent intraobserver agreement for assessment of the character of the invasive margin ($0.44 < \kappa < 0.82$), which was not significantly affected by sampling ($0.40 < \kappa < 0.56$, comparing both slides from each tumour) or observer (five of six pathologists agreeing on 46 of 60 slides). Pathologists were unreliable in assessing peritumoural lymphocytic infiltrates, with only two pathologists achieving moderate levels of intraobserver agreement ($-0.03 < \kappa < 0.52$). The interobserver agreement for peritumoural lymphocytic infiltrates was also low ($\kappa < 0.29$) between the three most experienced pathologists. The assessment of peritumoural lymphocytic infiltrates was significantly affected by sampling, the two pathologists with the lowest intraobserver variation achieving κ values of 0.21 and 0.10 between the 30 paired slides from each tumour.

The character of the invasive margin was reliably assessed, was not dependent on sample, and added useful prognostic information, but peritumoural lymphocytic infiltration is not a reproducible observation and may therefore not add useful prognostic information in routine use.

Dukes' classification of colorectal cancer has been used by pathologists and surgeons as a prognostic guide for assessing the benefit of surgical procedures and advising patients for more than 55 years.¹ The original Dukes' classification depended on objective anatomical variables but later Dukes modified his classification to take account of more subjective criteria (tumour grade). This later classification has been criticised because pathologists vary in their interpretation of tumour grade,² and only a small proportion of patients are placed into either an excellent or poor prognostic group.³

A careful retrospective study of anatomical and subjective variables which were thought to be potentially related to prognosis was carried out by Jass *et al.*³⁻⁶ Cox regression analysis was used to identify the minimum number of variables giving the most useful prognostic information. Four variables with an important and independent influence were identified: (i) number of lymph nodes with metastatic tumour; (ii) local spread; (iii) character of invasive margin;

and (iv) peritumoural lymphocytic infiltration. The first two variables are largely objective observations and are also used in the traditional Dukes' classification. The second two are subjective variables which were assessed by Jass alone and no attempt was made to assess the interobserver reproducibility of these variables. In a recent editorial it was suggested that this classification should be adopted for routine use⁷ but doubt has been raised about the reproducibility of the peritumoural lymphocytic infiltrate variable.⁸ It is essential that the intra- and interobserver reliability of the proposed subjective variables is assessed before this new system is introduced for routine use.

Material and methods

A heterogeneous group of histopathologists, three with a special interest in gastrointestinal pathology and three with other special interests, were invited to take part in the study. Experience varied between three and 22 years and all but two had passed the final MRCPath examination. A professional statistician was recruited and participated in the study design. Each pathologist was sent a protocol defining the

objectives of the study and a copy of the proposed prognostic classification of rectal cancer.³

Thirty rectal or sigmoid colon carcinomas were received fresh in the laboratory. These were flushed through with 10% formalin and immersed in formalin for at least three days. Sections were cut (1 cm) using a macrotome.⁹ Two blocks from the advancing edge of the tumour were taken by two of us (SD or RL), these were routinely processed and sections stained with haematoxylin and eosin. These 60 slides were labelled by a third party and circulated among the six histopathologists. The histopathologists were asked to state for each section whether the advancing edge of the tumour was "expanding" or "infiltrating" and whether there was a conspicuous peritumoural lymphocytic infiltrate according to the criteria defined in the proposed classification. The slides were relabelled and recirculated at least one month later and the pathologists asked to repeat their observations.

Cohen's κ statistic,¹⁰ a measure of the proportion of agreement having allowed for that expected by chance, was used to assess the reproducibility of results between different readings of the same slides and between different pathologists. Values of κ greater than 0.75 indicate excellent agreement beyond chance expectation and values less than or equal to 0.40 poor agreement.¹¹

Results

CHARACTER OF INVASIVE MARGIN:

INTRA-OBSERVER VARIABILITY

Table 1 shows the κ statistic for each pathologist compared with himself. The pathologists are listed in order of years of experience.

The table shows that five of the six pathologists showed good to excellent agreement between the first and second assessment of the invasive margin. There was a good correlation between experience and consistency in assessing the character of the invasive margin. Pathologist 6 placed nearly all the cases into one category, this resulted in a low value for his κ statistic as agreement is easy to reproduce by chance.

Table 1 *Measure of agreement between two assessments of the same slides for each pathologist*

Pathologist	κ statistic	95% confidence interval
1	0.82	(0.67, 0.97)
2	0.62	(0.38, 0.87)
3	0.57	(0.34, 0.80)
4	0.52	(0.25, 0.79)
5	0.44	(0.29, 0.76)
6	-0.03	(-0.51, 0.45)

Table 2 *Sampling error measured by comparing assessment of character of invasive margin between two slides from different areas of same tumour*

Pathologist	κ statistic	95% confidence interval
1	0.56	(0.25, 0.86)
2	0.58	(0.20, 0.96)
3	0.40	(0.01, 0.78)
4	0.42	(0.05, 0.79)
5	0.46	(0.11, 0.81)
6	-0.00	(-0.10, 0.10)

INTER-OBSERVER VARIABILITY

The three pathologists with a special interest in gastrointestinal pathology agreed in 47 of the 60 slides (78%) with a κ statistic of 0.66, indicating a good level of agreement. The three most experienced pathologists agreed on the character of the invasive margin in 47 of the 60 slides (78%) with a κ statistic of 0.63, indicating a good level of agreement. The three least experienced pathologists agreed on 40 of the 60 slides (66%) with a κ statistic of 0.32 indicating a poor level of agreement.

SAMPLING ERROR

The κ statistics for the six pathologists comparing their assessment of the character of the invasive margin between two slides from the same tumour is shown in table 2.

There was a good level of agreement between two slides prepared from blocks taken from different areas of the same tumour. This confirms that the character of the invasive margin is reproducible between different blocks of the same tumour.

PERITUMOURAL LYMPHOCYTIC INFILTRATION:

INTRA-OBSERVER VARIABILITY

Table 3 shows the κ statistic for each pathologist compared with himself.

The table shows that there was a generally poor level of agreement for each pathologist compared with himself. Only two pathologists achieved a good level of agreement between the two assessments.

INTER-OBSERVER VARIABILITY

There was a generally poor level of agreement between observers in assessing the presence or absence of a

Table 3 *Measure of agreement between two assessments of the same slides for each pathologist*

Pathologist	κ statistic	95% confidence interval
1	0.29	(0.03, 0.54)
2	0.48	(0.21, 0.76)
3	0.20	(-0.11, 0.51)
4	-0.03	(-0.26, 0.19)
5	0.52	(0.30, 0.74)
6	0.09	(-0.31, 0.74)

peritumoural lymphocytic infiltrate. The three most experienced pathologists agreed in only 28 of the 60 slides with a κ statistic of 0.29, indicating a poor level of agreement. The three least experienced pathologists agreed on 37 of the 60 slides with a κ statistic of 0.30.

SAMPLING ERROR

As the pathologists were poor at agreeing with themselves as to the presence or absence of a peritumoural lymphocytic infiltrate it is difficult to assess sampling error between different areas of the same tumour. The κ values between two slides from different areas of the same tumour for the two pathologists with the lowest levels of intraobserver variation (pathologists 5 and 2, table 3), however, were 0.21 and 0.10, respectively. This suggests that this feature is also variable between different areas of the same tumour.

Discussion

In a recent editorial it was stated that the "post-Dukes era" has arrived.⁷ It was also recommended in another editorial that pathologists should adopt the new "prognostic classification of rectal cancer" proposed by Jass *et al.*¹² The difficulties of applying the new classification system in practice have been highlight-

ed,⁸ drawing attention particularly to difficulties in assessing the presence or absence of a "peritumoural lymphocytic infiltrate".

Our results show that pathologists are able reliably and reproducibly to assess the character of the invasive margin of a tumour and that this observation is not significantly affected by the site of sampling of the invasive tumour margin. Specialist gastrointestinal pathologists were not significantly better than non-specialists, though more experienced pathologists were more likely to agree with themselves and with others than the less experienced. Addition of this observation to the traditional Dukes' classification adds useful prognostic information and therefore this information should be sought routinely.

We have shown that pathologists are not reliable in assessing the presence or absence of a "peritumoural lymphocytic infiltrate". None of the observers in our study achieved the low degree of intraobserver variability for this feature reported by Jass *et al.*⁴ Even among the three most experienced pathologists in our study the degree of interobserver agreement was little better than that expected by pure chance. Some of the interobserver variability may arise from the designation of this feature as a "peritumoural lymphocytic infiltrate" as plasma cells, histiocytes, eosinophils, and

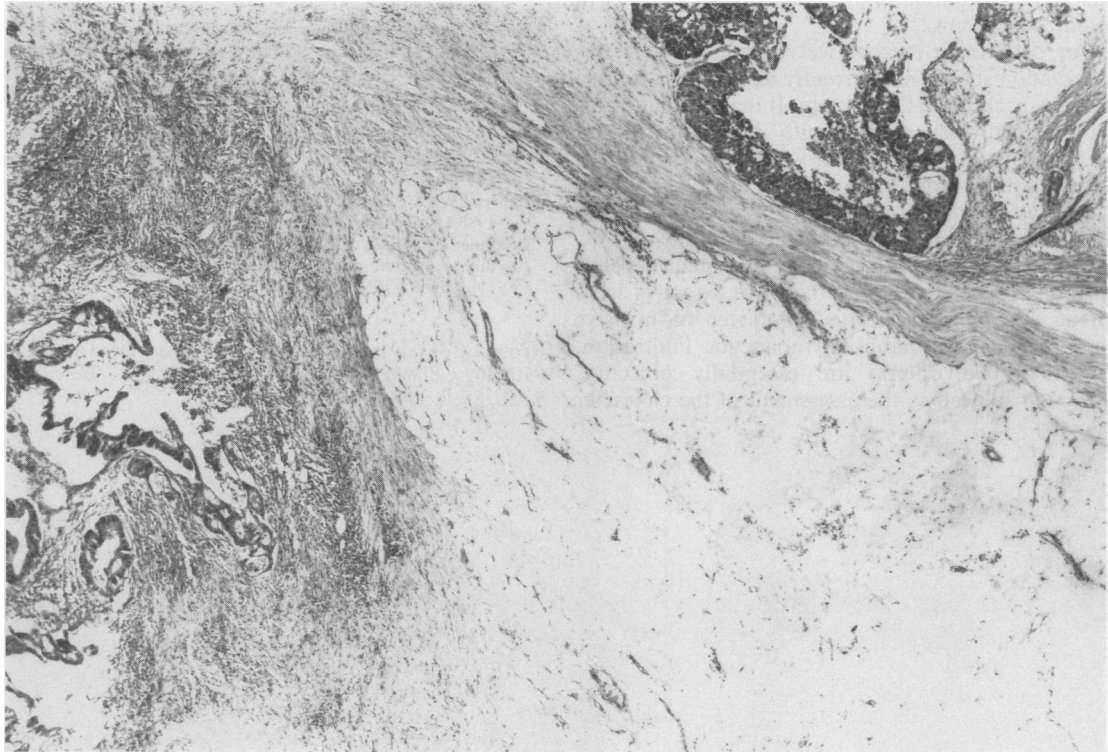


Figure Advancing tumour margin with dense inflammatory infiltrate investing left part of margin but no infiltrate to the right.

neutrophils are also allowed. Jass has commented, however, that it is his rule never to diagnose a "peritumoural lymphocytic infiltrate" unless the expanding margin or tips of invasive tongue of tumour are invested by a distant cellular lamina that includes a population of lymphocytes (though not necessarily vast numbers). Doubtful cases are always downgraded leaving a small group (20–25%) with an excellent prognosis. Assessment is based on the worst area (Jass JR, personal communication).

Our results also suggest the "peritumoural lymphocytic infiltrate" depends on sampling. Jass *et al* do not state whether the effect of sampling was assessed in their study, but as the work was carried out on archival material this was probably not possible. Jass *et al* state that assessment is based on the worst area but they do not state how large this area is. One of the sections in our study showed a dense inflammatory infiltrate in front of the advancing margin of one half of the section and no infiltrate on the other (figure). The more extensively a tumour is sampled the greater the chance of finding a similar "worst" area. Previous authors have also found sampling a problem when assessing peritumoural lymphocytic infiltrates. Skinner *et al* found it necessary to examine 10 to 20 slides from six to 10 blocks from each tumour to quantitate peritumoural infiltrates of macrophages in colonic neoplasms.¹³

For both these reasons we recommend that this observation should not currently be used in a routine prognostic classification system. It may be that if this feature is further defined with good illustrations pathologists may be able to recognise it reliably. But as pathologists showed a high degree of intraobserver variation for this feature it is possible that even with further clarification it cannot be used routinely.

The new prognostic classification system depends on four prognostic criteria: these are extent of local spread, lymph node status, character of invasive margin, and peritumoural lymphocytic infiltration. The first two criteria are essentially objective. Although subjective, the assessment of the character

of the invasive margin is reliable, it is not dependent on sample and adds useful prognostic information. Peritumoural lymphocytic infiltration, as currently defined, is not a reproducible observation and therefore may not add useful prognostic information in routine practice.

We thank Dr J Jass for his helpful comments and Dr D D Giri for randomising and labelling the slides.

References

- 1 Dukes CE. Histological grading of rectal cancer. *Proc Roy Soc Med* 1937;**30**:371–6.
- 2 Thomas GDH, Dixon MF, Smeeton NC, Williams NS. Observer variation in the histological grading of rectal carcinoma. *J Clin Pathol* 1983;**36**:385–91.
- 3 Jass JR, Love SB, Northover JMA. A new prognostic classification of rectal cancer. *Lancet* 1987;i:1303–6.
- 4 Jass JR, Atkin WS, Cuzick J, *et al*. The grading of rectal cancer: historical perspectives and a multivariate analysis of 447 cases. *Histopathology* 1986;**10**:437–59.
- 5 Jass JR. Lymphocytic infiltration and survival in rectal cancer. *J Clin Pathol* 1986;**39**:585–9.
- 6 Jass JR, Morson BC. Reporting colorectal cancer. *J Clin Pathol* 1987;**40**:1016–23.
- 7 Anonymous. Any advance on Dukes? [Editorial]. *Lancet* 1987;i:1301.
- 8 Penman HG. Reporting colorectal cancer. *J Clin Pathol* 1988;**41**:358.
- 9 Chan KW, Boey J, Wong SKC. A method of reporting radial invasion and surgical clearance of rectal carcinoma. *Histopathology* 1985;**9**:1319–27.
- 10 Fleiss JL. *Statistical methods for rates and proportions*. 2nd ed. New York: John Wiley & Sons, 1981.
- 11 Silcocks PBS. Measuring repeatability of histological diagnosis—a brief review with some practical examples. *J Clin Pathol* 1983;**36**:1269–75.
- 12 Ewen SWB. Colorectal excision specimens and prognosis. *J Pathol* 1988;**154**:206–7.
- 13 Skinner JM, Jarvis LR, Whitehead R. The cellular response to human colonic neoplasms: macrophage numbers. *J Pathol* 1983;**139**:97–103.

Requests for reprints to: Dr S A C Dundas, Department of Pathology, Sheffield University Medical School, Beech Hill Road, Sheffield S10 2RX, England.