

A reproducible grading scale for histological assessment of inflammation in ulcerative colitis

K Geboes, R Riddell, A Öst, B Jensfelt, T Persson, R Löfberg

Abstract

Background—Evaluation of histological activity in ulcerative colitis needs to be reproducible but has rarely been tested. This could be useful both clinically and in clinical trials.

Aim—To develop reproducible criteria which are valid in the assessment of acute inflammation (activity) and chronicity, and to evaluate these features in an inter-observer variability study.

Methods—A six grade classification system for inflammation was developed which could also be fine tuned within each grade. The grades were: 0, structural change only; 1, chronic inflammation; 2, lamina propria neutrophils; 3, neutrophils in epithelium; 4, crypt destruction; and 5, erosions or ulcers. Ninety nine haematoxylin-eosin sections from endoscopically inflamed and non-inflamed mucosa from patients with distal ulcerative colitis were assessed in two separate readings by three pathologists independently and without knowledge of the clinical status. Interobserver agreement was compared pairwise using kappa statistics.

Results—Initially, kappa values between the observers were 0.20, 0.42, and 0.26, which are too low to be of value. Following development of a semiquantitative pictorial scale for each criterion, kappa values improved to 0.62, 0.70, and 0.59. For activity defined by neutrophils between epithelial cells, kappa values were 0.903, 1.000, and 0.907. Complete agreement was reached in 64% of samples of endoscopically normal and in 66% of endoscopically inflamed tissue. Neutrophils in epithelium correlated with the presence of crypt destruction and ulceration.

Conclusion—A histological activity system was developed for ulcerative colitis that showed good reproducibility and modest agreement with the endoscopic grading system which it complemented. It has potential value both clinically and in clinical trials.

(Gut 2000;47:404–409)

Keywords: ulcerative colitis; biopsy; inflammation; scoring system

The diagnosis of idiopathic inflammatory bowel diseases (IBD) is usually based on a combination of clinical, radiological, endoscopic, and microscopic criteria. The reproducibility of the microscopic criteria, used for diagnosis of ulcerative colitis (UC), has been

examined in several studies.^{1–4} Overall, the most important features for discriminating between normal tissue and IBD, and UC and Crohn's disease are architectural abnormalities and inflammatory changes.^{5–8} Structural changes of the crypt architecture and basal plasmacytosis have good predictive diagnostic value for IBD versus non-IBD.^{8–10} UC is a chronic relapsing condition. Hence, it is not only important to reach a correct diagnosis but also to assess disease activity. This is done with the aid of clinical parameters, sometimes combined into clinical indices, and by using endoscopy, often supplemented by biopsy. It has been shown that treatment can alter the microscopic diagnostic features and signs of activity.^{11–14} Although histological activity may not be important in the everyday management of patients, which often depends on general well being, it can be important in the choice of drug treatment and in monitoring drug therapy, and it may be particularly important in the design and evaluation of any study aiming to show a therapeutic benefit. Patients with residual microscopic acute inflammation are more likely to relapse.^{15 16}

Different histological scoring systems have been designed for assessment of disease activity in UC.^{14 17–19} Usually they combine chronic and more acute changes, and epithelial as well as inflammatory features.²⁰ Microscopic activity is based on the presence of neutrophils or defined as unequivocal damage of the surface and crypt epithelium typically in conjunction with neutrophils.²¹ The use of neutrophils as an indicator of disease activity is supported by studies of leucocyte scanning.^{22 23} Histologically, neutrophils appear to be the effector cell causing epithelial damage.²⁴ Histological scoring systems are used in many clinical drug trials but data regarding the reproducibility of the scores and the possible relation between the occurrence of erosions and ulcers and the presence of neutrophils within the epithelium are limited.^{16 17} Hence the purpose of our study was to evaluate a microscopic scoring system designed for UC based on different grades of activity and to assess its reproducibility. The scoring system was designed to aid in assessing the effectiveness of therapy in UC.

Material and methods

Following a literature search for different features used to diagnose and assess activity in UC and different scoring systems available, a new classification and grading system for

Abbreviations used in this paper: IBD, inflammatory bowel disease; UC, ulcerative colitis.

GI pathology Unit, KU Leuven, Belgium
K Geboes

McMaster University, Medical Center, Hamilton, Canada
R Riddell

Medilab AB and Karolinska Institute, Stockholm, Sweden
A Öst

Astra Zeneca AB, Sweden
B Jensfelt
T Persson

Department of Gastroenterology, Karolinska Institute, Huddinge University Hospital, Sweden
R Löfberg

Correspondence to:
Dr K Geboes, Department of Pathology, University Hospital, KU Leuven, Minderbroedersstraat 12, 3000 Leuven, Belgium.
Email: Karel.Geboes@uz.kuleuven.ac.be

Accepted for publication
5 April 2000

Table 1 Different grades used for evaluation of disease severity in ulcerative colitis

Grade 0	Structural (architectural change)
Subgrades	
0.0	No abnormality
0.1	Mild abnormality
0.2	Mild or moderate diffuse or multifocal abnormalities
0.3	Severe diffuse or multifocal abnormalities
Grade 1	Chronic inflammatory infiltrate
Subgrades	
1.0	No increase
1.1	Mild but unequivocal increase
1.2	Moderate increase
1.3	Marked increase
Grade 2	Lamina propria neutrophils and eosinophils
2A Eosinophils	
2A.0	No increase
2A.1	Mild but unequivocal increase
2A.2	Moderate increase
2A.3	Marked increase
2B Neutrophils	
2B.0	None
2B.1	Mild but unequivocal increase
2B.2	Moderate increase
2B.3	Marked increase
Grade 3	Neutrophils in epithelium
3.0	None
3.1	< 5% crypts involved
3.2	< 50% crypts involved
3.3	> 50% crypts involved
Grade 4	Crypt destruction
4.0	None
4.1	Probable—local excess of neutrophils in part of crypt
4.2	Probable—marked attenuation
4.3	Unequivocal crypt destruction
Grade 5	Erosion or ulceration
5.0	No erosion, ulceration, or granulation tissue
5.1	Recovering epithelium+adjacent inflammation
5.2	Probable erosion—focally stripped
5.3	Unequivocal erosion
5.4	Ulcer or granulation tissue

	Surface epithelium	Crypt epithelium	Crypt abscesses
1 0	X		
2 0		X	
3 0	X	X	
4 0			X
5 0	X		X
6 0		X	X
7 0	X	X	X

Figure 1 Different combinations of locations of neutrophils within the epithelium. Seven possible combinations were identified.

assessment of progressive inflammation and activity was conceived (table 1). The underlying hypothesis is that the different major grades and subgrades are progressive and correlate with increasing disease severity or activity.

Furthermore, the position of the neutrophils between the epithelial cells was scored separately in grade 3. Seven possible combinations were identified (fig 1). The purpose of this was to examine the relationship between location of neutrophils and occurrence of erosions or ulcers and crypt destruction. If such a relationship exists, the location of the neutrophils would be important for evaluation of medical treatment. In addition, the technical quality of the sections and orientation of the samples and their effect on accurate microscopic analysis were evaluated.

The different major grades and subgrades were clearly defined: grade 0.0 indicated the absence of any abnormality; grade 0.1 indicated a solitary architectural abnormality such as one definitely abnormal crypt or inappropriate spacing; grade 0.3 was used when the lesions were severe and diffuse; and grade 0.2 was any

abnormality between a single lesion and diffuse lesions. Any increase in chronic inflammation in the lamina propria infiltrate indicated a grade 1 score. Subgrade 1.1 indicated a mild but unequivocal increase and subgrade 1.3, a marked increase in chronic inflammation. The presence of granulocytes within the lamina propria or between epithelial cells was scored in grades 2 and higher. Subgrade 2B.1 was scored when one or more neutrophils were present in the lamina propria. For the assessment of subgrades, the worst area of the biopsy, and not the average aspect, was used. A series of photographs was constructed to define the upper and lower limits of each subgrade (fig 2).

The results were analysed in three different ways. In the first analysis the activity was assessed according to the six different major grades. In the second analysis grades 0 and 1 were combined in one group and the others in a second group. In the third analysis major grades 0, 1, 2A, and 2B were combined in one group and grades 3, 4, and 5 in a second group. The cut off between the two groups was the presence of neutrophils either in the lamina propria in the second analysis or between epithelial cells in the third analysis. The purpose of the different analyses was to examine possible differences between a simple and more complex scoring system.

A total of 99 haematoxylin-eosin stained sections from biopsies obtained in patients with established distal UC were examined. All patients were selected for a controlled clinical trial. The biopsies were obtained in endoscopically inflamed (n=68) and non-inflamed mucosa (n=31). In 66 cases more than one sample from the same region was available. Two separate readings were performed. In the first reading three different semiserial sections (5 µm) of each biopsy were assessed by three pathologists (KG, RR, AO). In the second session the same section was examined by all three pathologists to minimise any effect caused by differences between sections. Both readings were performed independently and without knowledge of the endoscopic status. The first reading was followed by analysis of areas of disagreement related to sectioning and the use of semiserial sections. Possible sources of error and areas of disagreement were identified and clarified by re-examination of the slides over a multiheaded microscope. The grading scale was adapted and refined. In particular the boundaries within each grade between different degrees of severity and the features allowing identification of a precise lesion such as crypt destruction and erosion were discussed and defined more precisely.

The second reading was followed by a statistical analysis of the data and re-examination of the slides for identification of sources of disagreement. Overall interobserver agreement was compared pairwise using kappa statistics. A kappa value of ≤ 0.5 was considered "poor", 0.51–0.6 "moderate", 0.61–0.8 "good", and > 0.8 "excellent". In addition, the relation between microscopic and endoscopic findings was studied, and p values were calculated for

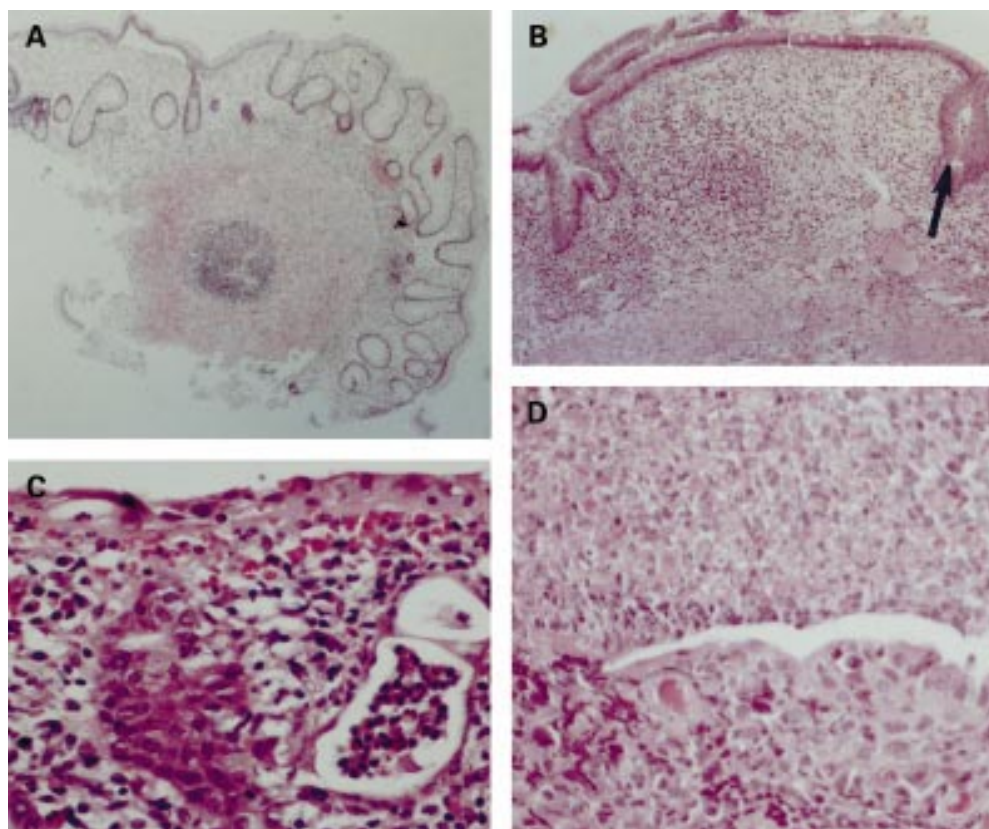


Figure 2 Microscopic photographs illustrating the different grades as well as the lower and upper borders. (A) Grade 0.3: the section is of good quality, the architecture is severely and diffusely altered but the lamina propria cellular infiltrate is not increased in intensity. (B) Grade 3.3: neutrophils are infiltrating the crypt epithelium in >50% of the crypts (arrow indicates neutrophils in lumen of crypt). (C) Grade 5.1: the surface is covered by attenuated cells, characteristic of a recovering epithelium. In addition there is severe adjacent inflammation. (D) Grade 5.3: unequivocal erosion because of the loss of surface epithelial cells, adjacent recovering epithelium, and severe inflammation (haematoxylin-eosin).

Table 2 Results of the first and second readings for the three observers (A, B, and C) (mean grades)

	A		B		C	
	1st	2nd	1st	2nd	1st	2nd
All biopsies	3.02	3.22	2.68	3.17	2.33	3.17
Healthy tissue	1.66	1.50	1.66	1.54	0.66	1.54
Inflamed tissue	3.62	4.05	3.14	3.97	3.06	3.97

Kappa values: A:B=0.42 (1st), 0.62 (2nd); A:C=0.26 (1st), 0.70 (2nd); B:C=0.20 (1st), 0.59 (2nd).

In the first reading, three different semiserial sections were examined (total, n=95; healthy tissue, n=29; inflamed tissue, n=66). In the second reading the same section was examined by all three pathologists (total, n=86; healthy tissue, n=28; inflamed tissue, n=58).

the different locations of neutrophils in the epithelium.

Results

INITIAL READING

The results of the first reading are summarised in table 2. Analysis of interobserver differences allowed a distinction between four different categories: (a) observer errors—that is, when one or more observers had missed a feature; (b) errors due to the use of three different semiserial sections; (c) differences due to the quality of the section; and (d) disagreement resulting from differences in interpretation.

Following a discussion, the grading scale was adapted and the borders between the subgrades were defined more clearly. It was decided to use only one serial section in the

second reading to be read by all three pathologists independently to exclude differences resulting from the use of different sections. For assessment of different subgrades of severity it was decided to construct a series of photographs showing the upper and lower limits of a given feature, rather than to use a typical example of a certain degree of severity, as this was one of the reasons for the many differences in interpretation. Finally, it was also decided to assess the quality of the sample to see if technical quality had an important effect on analysis.

SECOND READING

The results of the second reading are also summarised in table 2. Agreement was excellent for major grades 0 and 5. Complete agreement for the final score (major grade and subgrades) was reached in 65% of samples of good quality examined by all three observers (n=86). For biopsies obtained in endoscopically normal tissue complete agreement was reached for the final score in 64% of samples and for biopsies obtained in endoscopically inflamed tissue complete agreement was reached in 66% of cases (n=58) (figs 3, 4). A difference of one major grade between one observer and the two others was noted in 18 cases (21%) and a two grade difference was noted in nine cases. A large discrepancy was found in two cases, graded as 5 by two of the three observers in one case, and as 5 and 3 by two observers in the second case.

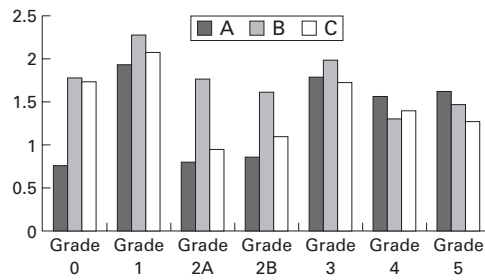


Figure 3 Interobserver variability for the different grades of the scoring system. Mean scores of the three observers (A, B, and C) are shown for the second reading. Only biopsies from endoscopically involved areas are included. Observer agreement is high for grades 3, 4, and 5.

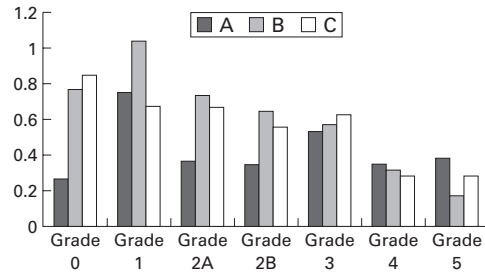


Figure 4 Interobserver variability for the different grades of the scoring system for biopsies obtained in endoscopically healthy tissue. Mean scores of the three observers (A, B, and C) are shown for the second reading. Agreement is excellent for grades 3 and 4.

Thirty one samples were found to be of good quality by all three examiners whereas 36 were of substandard quality and 22 of poor quality (13 of these were not examined by all three observers because of poor quality). For the overall scores, kappa values were 0.489, 0.434, and 0.670 for the poor specimens and 0.666, 0.706, and 0.510 for the good samples.

Distinction between the groups with and without granulocytes was excellent, with kappa values of 0.815, 0.945, and 0.869. Agreement on the diagnosis of activity, defined by the presence of unequivocal damage of surface or crypt epithelium in conjunction with neutrophils (major grades 3, 4, and 5), was also excellent, with kappa values of 0.903, 1.000, and 0.907.

For samples of good quality from endoscopically uninfamed tissue (n=28), kappa values for agreement between pathologists and endoscopy were, respectively, 0.613, 0.545, and 0.386, indicating moderate agreement. The average grade scored by the pathologists for these samples was below 2 (presence of granulocytes in the lamina propria) or more. The mean values were 1.07, 1.13 and 1.07.

For specimens of good quality obtained in endoscopically inflamed tissue, kappa values for agreement between pathology and endoscopy were 0.648, 0.808, and 0.570, respectively, indicating moderate to excellent agreement. Mean scores for inflamed tissue, given by each of the pathologists, were 3.25, 2.75, and 3.00. The endoscopy score for inflamed tissue correlated well with grade 5 and grades 3 and 2B ($p < 0.02$) for all assessors. Grade 2A, eosinophils within the lamina propria only, also correlated with the endoscopy score for each of

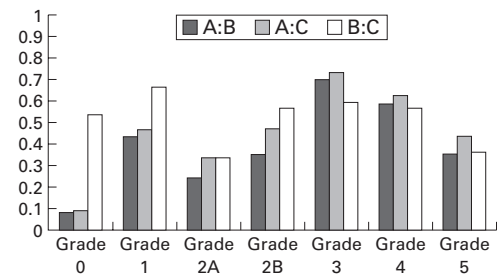


Figure 5 Interobserver variability in the assessment of activity of ulcerative colitis on rectal biopsies for the second reading. Kappa values for the different grades of the scoring system are shown for the three observer pairs (A:B, A:C, and B:C). Agreement is poor for assessment of structural changes without associated inflammatory changes (that is, quiescent colitis) for two pairs of observers.

Table 3 Differences in the assessment of subgrades

Grade	Complete agreement	One grade difference	Two grade difference
Grade 0	23%	51%	27%
Grade 1	48%	49%	3%
Grade 2A	35%	46%	19%
Grade 2B	48%	36%	16%
Grade 3	66%	31%	3%
Grade 4	66%	17%	17%

Intragrade differences and variability of intensity for grades 0, 1, 2, 3, and 4. Results are expressed as percentage of the total number of biopsies scored for a particular item.

the observers ($p < 0.11$). However, this does not mean that there was complete agreement between the endoscopy score and histology. Endoscopically inflamed tissue appeared normal for all three observers in three cases (4%), and in endoscopically uninfamed tissue, microscopic active inflammation was still present in 13/28 (46%) cases according to all three observers. No abnormalities were present in 11/28 samples (39%). In the remainder, structural changes or an increase in intensity of mononuclear cells was observed by at least one of the three observers.

Kappa values for the different major grades (0, 1, 2A, 2B, 3, 4, and 5) are shown in fig 5. Values were more than 0.4 and usually more than 0.5 for grades 1, 3, and 4. Analysis of the results for intragrade differences and variability of the subgrades (for example, 1.1 v 1.2 v 1.3) for major grades 0 to 4 are shown in table 3. For major grades 4 and 5, complete agreement was reached in 66% of cases for the presence, absence, or type of crypt destruction and in 44% for the presence, absence, or type of erosion or ulceration. In 17% of cases one grade difference was noted for crypt destruction but in only 14% the lesion was considered to be absent. In 22% of cases presenting with an erosion, the lesion was not recognised as such by one of the observers.

Disagreement was due to observer errors, interpretation errors, or a combination of both. The total number of observer errors was 4.67% of all observations when the location of neutrophils was not included and 3.06% of all observations when this item was included. Observer errors were found in all major grades: 10% for grade 0; 2% for grade 1; 8% for grade 2A; 6% for grade 2B; 1% for grade 3; and 3% for grades 4 and 5.

Table 4 Neutrophils in epithelium and the presence of an erosion or ulceration

Grade	Location 1+3+5+7	Location 2+4+6
5.0	8	5
5.1	3	0
5.2	16	2
5.3	14	0
5.4	6	0

Neutrophils in the surface epithelium (1, 3, 5, and 7) or crypt epithelium (2, 4, and 6) and the presence of an erosion or ulceration (major grade 5): the numbers indicate the number of biopsies involved and assessed by one observer. Neutrophils in the crypts were less common when an erosion was present.

CORRELATION BETWEEN THE LOCATION OF NEUTROPHILS IN THE EPITHELIUM AND OCCURRENCE OF CRYPT DESTRUCTION, EROSIONS, AND ULCERATIONS

Correlation of the location of neutrophils in the epithelium (major grade 3) with major grades 4 (crypt destruction) and 5 (erosion) showed that both crypt destruction and erosion were significantly more common when neutrophils were present in the epithelium. For major grade 4 indicating crypt destruction, p values between the presence of the lesion and neutrophils in the crypt epithelium (locations 2+3+6+7 *v* 1+4+5) were 0.18, 0.026, and 0.062, respectively, for the three different observers. For grade 4, crypt destruction and crypt abscesses (locations 4+5+6+7 *v* 1+2+3), p values were 0.0002, 0.015, and 0.0001. This indicates that the presence of crypt abscesses implies crypt destruction. For grade 5 (erosions or ulcerations), the correlation with neutrophils in the surface epithelium (locations 1+3+5+7 *v* 2+4+6), p values were 0.0045, 0.16, and 0.0057. This means that the presence of neutrophils in the epithelium implies the likelihood of an erosion or ulceration (table 4).

Discussion

Overall, complete agreement for the final score was reached by all three observers for 56/86 samples (65%) of good or acceptable quality. The mean grades for the 28 samples obtained in endoscopically normal tissue (1.07, 1.13, and 1.07) and for the 58 samples obtained in endoscopically inflamed tissue (3.25, 2.75, and 3.00) were comparable. Kappa values were 0.62, 0.70, and 0.59, indicating moderate to good agreement. Observer errors were noted in 3.5% of cases. The improvement following the second reading indicates the importance of precise definitions.

Technical quality and orientation improve microscopic assessment and final grading. Disagreement is partly due to the presence of more than one sample and discontinuity of the lesions when samples are compared. Discontinuity of the lesions has been reported for UC and accurate assessment of microscopic disease activity may need analysis of more than one biopsy sample.^{12 13} As only one final score combining the major grade and the subgrade was given, some observers gave an average score when the lesions were discontinuous while others scored the worst. It was finally decided that the worst score should be used.

Kappa values improve and agreement becomes excellent when classes or grades are collapsed. When disease activity was defined by

the presence of unequivocal damage of the epithelium with neutrophils, kappa values were 0.903, 1.000, and 0.907. This confirms earlier studies showing that neutrophils can be assessed reproducibly and that interobserver agreement is good for histological features associated with neutrophils.^{1 14 17} The reproducibility of a simple scoring system is thus high, as expected. However, the purpose of the study was to construct a more refined scoring system, allowing us to define more precisely disease activity and eventually guide long term therapy; our study showed moderate to excellent agreement for such a complex system.

In the past, several studies have found variable results when comparing endoscopy scores and histology in UC.^{20 25} In the present study, the correlation between endoscopy and histology was generally good for endoscopically inflamed mucosa. The correlation was worse for endoscopically uninfamed mucosa. This is not surprising as endoscopy and histology do not assess the same features. Focal active inflammation is likely to be missed by endoscopy and biopsies thus add an additional dimension regarding the presence of inflammation. Therefore, it seems appropriate to use both endoscopy and histology for the assessment of disease activity and extent. Persistent microscopic lesions, in the absence of endoscopic lesions, may indicate increased likelihood of relapse.^{15 16}

Analysis of areas of disagreement after the first reading showed that there were only a few slides where neutrophils in the lamina propria was the only sign of activity, and was not associated with neutrophils infiltrating surface or crypt epithelium. Variations resulted from failure of one or more pathologists to detect a lesion or from interpretation of neutrophils in endothelial lined channels as being normal or abnormal. Observer errors diminished but were still noted after the second reading. These were mostly due to poor technical quality but in five good quality specimens the final grade of the biopsy was influenced by observer error.

The number of observer errors for erosions or ulcerations was small. Disagreement was related mainly to differences in interpretation between a genuine erosion or an artifactual stripping of epithelium. The presence of recovering epithelium, defined as attenuated (flat or cuboidal) surface epithelium cells, was therefore included in the assessment form to improve recognition of an erosion. The presence of adjacent severe inflammation was added to exclude as far as possible other potential causes of superficial damage. We also included a category of probable erosion versus focal stripping. Inclusion of the latter category was responsible for a two grade difference in the final score in 5/12 cases where such a major difference was found. Hence it may not be appropriate to include this step in grade 5.

Some interpretation problems were not solved after the first reading. Evaluation of the lamina propria mononuclear infiltrate remained a problem. This has also been observed in other studies.²⁶ It has been shown that the lamina propria cellularity is an important feature of IBD but the boundaries of the

normal cellular lamina propria infiltrate need clarification and standardisation.^{27, 28} Identification of eosinophils in the lamina propria is another source of disagreement, responsible for 8% of observer errors. Yet some data in the literature indicate that eosinophils are important for the pathogenesis of CD and UC.²⁸⁻³⁰ Therefore, we propose to include eosinophils in histological scoring systems assessing disease activity in IBD.

For grade 0, complete agreement was reached in only 22% of cases after the second reading. Similar disagreement was also noted after the first reading. Following the analysis a series of photographic standards was constructed for comparison, representing not an average picture but the boundaries between the different categories. Data from the second reading showed that the disagreement was partly due to observer errors and partly to differences in interpretation. Disagreement in structural changes was mainly due to the availability of more than one sample from the same patient and the same area. Orientation of the sections is another source of disagreement. Bifid crypts and increased distance between the crypt base and muscularis mucosae were used as the main criteria, but these features cannot be assessed adequately in transverse sections or sections cut tangentially. Variability in crypt diameter and intercryptal distance are clearly more difficult parameters to use.

A good correlation was found between the location of neutrophils and occurrence of crypt destruction, erosions, and ulcers. Our data confirm that neutrophils between epithelial cells are related to epithelial cell damage. Neutrophils in the surface epithelium correlate with the occurrence of erosions. Crypt abscesses imply crypt destruction. Reduction or disappearance of neutrophils in the epithelium in consecutive biopsies is thus most likely a sign of reduction of disease activity and could indicate the efficacy of a given treatment.

In conclusion, our scoring system is sufficiently reproducible that it may be of value in clinical trials and in the routine assessment of clinical activity. Technical quality, both in terms of quality of sections and good orientation, can improve the quality of reporting and is therefore strongly encouraged. Some histological features, especially structural abnormalities, need more accurate definition. The presence of eosinophils has to be included in any grading system assessing activity in UC. The presence of neutrophils in the epithelium correlates well with the development of lesions such as crypt destruction and erosions or ulcers. Although endoscopy is a very important tool for diagnosis and assessment of disease activity and there is a good correlation between endoscopy and histology, the combination of endoscopy and biopsy provides a better indication of activity than endoscopy alone, especially in endoscopically non-inflamed mucosa.

The study was supported by a grant from Astra-Zeneca, Sweden. The IOIBD Pathobiology section provided the opportunities for preparatory discussions.

- Jenkins D, Balsitis M, Gallivan S, *et al.* Guidelines for the initial biopsy diagnosis of suspected chronic idiopathic inflammatory bowel disease. The British Society of Gastroenterology Initiative. *J Clin Pathol* 1997;50:93-105.
- Riddell RH. Histopathology of ulcerative colitis. In: Allan RW, Rhodes JM, Hanauer SB, *et al.*, eds. *Inflammatory bowel disease*, 3rd ed. New York: Churchill Livingstone, 1997: 291-309.
- Goldman H. Colonic mucosal biopsy in inflammatory bowel disease. *Surg Pathol* 1991;4:3-23.
- Theodossi A, Spiegelhalter DJ, Jass J, *et al.* Observer variation and discriminatory value of biopsy features in inflammatory bowel disease. *Gut* 1994;35:961-8.
- Rubio CA, Johansson C, Uribe A, *et al.* A quantitative method of estimating inflammation in the rectal mucosa: IV. Ulcerative colitis in remission. *Scand J Gastroenterol* 1984;19:525-30.
- Seldenrijk CA, Morson BC, Meuwissen SGM, *et al.* Histopathological evaluation of colonic mucosal biopsy specimens in chronic inflammatory bowel disease: diagnostic implications. *Gut* 1991;32:1514-20.
- Rubio CA, Johansson C, Kock Y. A quantitative method of estimating inflammation in the rectal mucosa. II. Normal limits in symptomatic patients. *Scand J Gastroenterol* 1982; 17:1077-81.
- Tanaka M, Riddell RH, Saito H, *et al.* Morphologic criteria applicable to biopsy specimens for effective distinction of inflammatory bowel disease from other forms of colitis and Crohn's disease from ulcerative colitis. *Scand J Gastroenterol* 1999;34:55-67.
- Nostrant TT, Kumar NB, Appelman HD. Histopathology differentiates acute self-limited colitis from ulcerative colitis. *Gastroenterology* 1987;92:318-28.
- Schumacher G, Kollberg B, Sandstedt B. A prospective study of first attacks of inflammatory bowel disease and infectious colitis. *Scand J Gastroenterol* 1994;29:318-32.
- Janssens J, Geboes K, Delanote C, *et al.* 5-amino-salicylic acid (5-ASA) enemas are effective in patients with resistant ulcerative colitis. *Gastroenterology* 1983;84:1198.
- Bernstein CN, Shanahan F, Anton PA, *et al.* Patchiness of mucosal inflammation in treated ulcerative colitis: a prospective study. *Gastrointest Endosc* 1995;42:232-7.
- Kleer CG, Appelman HD. Ulcerative colitis: patterns of involvement in colorectal biopsies and changes with time. *Am J Surg Pathol* 1998;22:983-9.
- Odze R, Antonioli D, Peppercorn M, *et al.* Effect of topical 5-aminosalicylic acid (5-ASA) therapy on rectal mucosal biopsy morphology in chronic ulcerative colitis. *Am J Surg Pathol* 1993;17:869-75.
- Wright R, Truelove SC. Serial rectal biopsy in ulcerative colitis during the course of a controlled therapeutic trial of various diets. *Am J Dig Dis* 1964;11:847-57.
- Riley SA, Mani V, Goodman MJ, *et al.* Microscopic activity in ulcerative colitis: what does it mean? *Gut* 1991;32:174-8.
- Riley SA, Mani V, Goodman MJ, *et al.* Comparison of delayed release 5 aminosalicylic acid (mesalazine) and sulphasalazine in the treatment of mild to moderate ulcerative colitis relapse. *Gut* 1988;29:669-74.
- Gomes P, Claire du Boulay, Smith CL, *et al.* Relationship between disease activity indices and colonoscopic findings in patients with colonic inflammatory bowel disease. *Gut* 1986;27:92-5.
- Florén CH, Benoni C, Willén R. Histologic and colonoscopic assessment of disease extension in ulcerative colitis. *Scand J Gastroenterol* 1987;22:459-62.
- Danielsson A, Hellers G, Lyrenas E, *et al.* A controlled randomized trial of budesonide versus prednisolone retention enemas in active distal ulcerative colitis. *Scand J Gastroenterol* 1987;22:987-92.
- Goldman H. Interpretation of large intestinal mucosal biopsy specimens. *Hum Pathol* 1994;25:1150-9.
- Stahlberg D, Veress B, Mare K, *et al.* Leucocyte migration in acute chronic inflammatory bowel disease: comparison of histological assessment and TC-99m-HMPAO labeled leucocyte scan. *Am J Gastroenterol* 1997;92:283-8.
- Weldon MJ, Masoomi AM, Britten AJ, *et al.* Quantification of inflammatory bowel disease activity using technetium-99m HMPAO labeled leucocyte single photon emission computerized tomography (SPECT). *Gut* 1995;36:243-50.
- Keren DF, Appelman HD, Dobbins III WO, *et al.* Correlation of histopathologic evidence of disease activity with the presence of immunoglobulin-containing cells in the colons of patients with inflammatory bowel disease. *Hum Pathol* 1984;15:757-63.
- Powell-Tuck J, Day DW, Buckell NA, *et al.* Correlations between defined sigmoidoscopic appearances and other measures of disease activity in ulcerative colitis. *Dig Dis Sci* 1982;27:533-7.
- Tsang P, Rotterdam H. Biopsy diagnosis of colitis: possibilities and pitfalls. *Am J Surg Pathol* 1999;23:423-30.
- Jenkins D, Goodall A, Drew K, *et al.* What is colitis? Statistical approach to distinguish clinically important inflammatory change in rectal biopsy specimens. *J Clin Pathol* 1988; 41:72-9.
- Dubucquoi S, Janin A, Klein O, *et al.* Activated eosinophils and interleukin 5 expression in early recurrence of Crohn's disease. *Gut* 1995;37:242-6.
- Wright R, Truelove SC. Circulating and tissue eosinophils. *Am J Dig Dis* 1966;11:831-46.
- Oshitani N, Campbell A, Kitano A, *et al.* In situ comparison of phenotypic and functional activity of infiltrating cells in ulcerative colitis mucosa. *J Pathol* 1996;178:95-9.