

## ORIGINAL ARTICLE

# Discriminant histological features in the diagnosis of chronic idiopathic inflammatory bowel disease: analysis of a large dataset by a novel data visualisation technique

S S Cross, R F Harrison

*J Clin Pathol* 2002;55:51–57

**Background/Aims:** The histopathological assessment of endoscopic colorectal biopsies is important in the distinction between normality and chronic idiopathic inflammatory bowel disease, and between ulcerative colitis and Crohn's disease, in subjects with symptoms of bowel dysfunction. This study aims to use carefully defined histopathological observations on a large study population to produce systems that improve classification into these diagnostic categories.

**Methods:** Eight hundred and nine endoscopic colorectal biopsies with verified outcomes (165 normal, 473 ulcerative colitis, 171 Crohn's disease) were examined by a single experienced histopathologist and 20 defined features were recorded for each case using a novel graphical interface with reference images of each feature. These features, together with age and sex, were used to produce and test statistical classifiers using logistic regression and a novel growing cell structure technique.

**Results:** The distinction between chronic idiopathic inflammatory bowel disease and normality was made with a good level of performance by both statistical classifiers (with areas under the receiver operating characteristic curves above 0.80). The growing cell structure system selected features as discriminant that agreed with the published literature. Logistic regression produced a more variable selection of discriminant features because of the high correlation between many features. The distinction between ulcerative colitis and Crohn's disease was performed less accurately, with areas under the receiver operating characteristic curves of about 0.70. Again the features selected as discriminant broadly agreed with those in the published literature.

**Conclusions:** Histopathological examination of endoscopic colorectal biopsies is an effective method of distinguishing between subjects with chronic idiopathic inflammatory bowel disease and normality, but less good at distinguishing between ulcerative colitis and Crohn's disease. The features selected as discriminant in this large statistical analysis broadly agree with those published in the literature from more qualitative studies.

See end of article for authors' affiliations

Correspondence to: Dr S S Cross, Academic Unit of Pathology, Section of Oncology and Pathology, Division of Genomic Medicine, University of Sheffield Medical School, Beech Hill Road, Sheffield S10 2UL, South Yorkshire, UK; s.s.cross@sheffield.ac.uk

Accepted for publication 21 May 2001

The distinction between chronic idiopathic inflammatory bowel disease (CIIBD) and normality, and between ulcerative colitis and Crohn's disease in known cases of CIIBD, is very important to patient management. Histopathological assessment of endoscopic colorectal biopsies often plays a large part in these diagnostic processes. Although the accuracy and reliability of histopathology diagnosis is reasonable in this context,<sup>1–3</sup> it is not perfect, and some studies have sought to improve this by statistical analysis of defined observations.<sup>1–7</sup> However, many of these studies have had relatively small numbers of cases, usually 40–50<sup>4–5</sup> (the largest being 128),<sup>7</sup> and different selections of features, many of which have been poorly defined. In our study, we use a large (809 cases) study population, making carefully defined observations entered using a novel graphical interface and use two classification techniques, one novel and unique to our laboratory, to investigate the most discriminant histopathological features for these diagnostic processes.

## MATERIALS AND METHODS

### Study population

In total, we studied 809 large bowel endoscopic biopsies reported in the department of histopathology, Royal Hallamshire Hospital, Sheffield between 1990 and 1995 (inclusive). Biopsies originating in diverted bowel, rectal stumps, or pouches were excluded, as were those with a diagnosis of neoplasia. The diagnosis was confirmed by the finding of typical endoscopic appearances seen on video photographs in the

clinical notes, subsequent bowel resection, pattern of disease on radiological investigation, or microbiological culture results. In cases without confirmation by subsequent resection specimens this final diagnostic outcome was made with review of the patient's case notes.<sup>3</sup> The biopsies were a mixed population of single distal biopsies and colonoscopic series from initial presentation and follow up of disease. By verified outcome, the study population contained 165 biopsies from normal subjects, 473 from patients with ulcerative colitis, and 171 with Crohn's disease.

### The observed features

The biopsies were examined (blind to all clinical details) by a single experienced observer (SSC) using a computer interface that implements the British Society of Gastroenterology (BSG) guidelines for the initial biopsy diagnosis of suspected CIIBD,<sup>1</sup> with digitised images representing examples of each histopathological feature.<sup>8</sup> Some of the features are dichotomous variables—for example, the presence or absence of mucosal granulomas, whereas others are ordinal categories, such as mucin depletion classified into none, mild, moderate, or severe. Table 1 lists the observed features and their coding. The observer recorded each feature by choosing the digitised

**Abbreviations:** BSG, British Society of Gastroenterology; CIIBD, chronic idiopathic inflammatory bowel disease; GCS, growing cell structure; LP, lamina propria; ROC, receiver operating characteristic

**Table 1** The observed features (input variables) and their coding

| Variable no. | Input variable                          | Type                 | Values     |
|--------------|---|----------------------|------------|
| 1            | Age                                     | Real number          | 16 to 86   |
| 2            | Female sex                              | Binary               | 0, 1       |
| 3            | Mucosal surface                         | Ordinal, categorical | 0, 1, 2, 3 |
| 4            | Crypt architecture                      | Ordinal, categorical | 0, 1, 2, 3 |
| 5            | Crypt profiles                          | Real number          | 2 to 8     |
| 6            | Increased LP cellularity                | Binary               | 0, 1       |
| 7            | Superficially increased LP cellularity  | Binary               | 0, 1       |
| 8            | Increased lymphoid aggregates           | Binary               | 0, 1       |
| 9            | Patchy increased LP cellularity         | Binary               | 0, 1       |
| 10           | Transmucosally increased LP cellularity | Binary               | 0, 1       |
| 11           | Cryptitis: extent                       | Ordinal, categorical | 0, 1, 2, 3 |
| 12           | Cryptitis: polymorphs                   | Ordinal, categorical | 0, 1, 2, 3 |
| 13           | Crypt abscesses: extent                 | Ordinal, categorical | 0, 1, 2, 3 |
| 14           | Crypt abscesses: polymorphs             | Ordinal, categorical | 0, 1, 2, 3 |
| 15           | Increased LP polymorphs                 | Binary               | 0, 1       |
| 16           | Epithelial changes                      | Ordinal, categorical | 0, 1, 2, 3 |
| 17           | Mucin depletion                         | Ordinal, categorical | 0, 1, 2, 3 |
| 18           | Intraepithelial lymphocytes             | Binary               | 0, 1       |
| 19           | Subepithelial collagen                  | Binary               | 0, 1       |
| 20           | LP granulomas                           | Binary               | 0, 1       |
| 21           | Submucosal granulomas                   | Binary               | 0, 1       |
| 22           | Basal histiocytes                       | Binary               | 0, 1       |

LP, lamina propria.

image that was most similar to the image seen down the light microscope and using a mouse to click on this, which recorded the observation in the computerised database. The observations were spread over a period of nine months, with no more than 30 biopsies observed in a single day. The histopathological diagnosis at the initial reporting of each specimen was recorded, and after making all the observations on a patient the observing histopathologist (SSC) recorded a diagnosis. The sensitivities and specificities were calculated for these diagnoses.

#### Partitioning of the dataset

The dataset was analysed in two parts. The whole dataset was used with a dichotomous outcome of normal or CIIBD (with Crohn's disease and ulcerative colitis cases combined as a single category). The order was randomised and the first 540 cases were used as the training set and the other 269 cases as the test set. Three hundred and seventy cases from the whole dataset were selected with a dichotomous outcome of Crohn's disease or ulcerative colitis. These cases were defined by their outcome and the presence of active inflammation (as indicated by polymorphs in the lamina propria (LP)). The order of these was randomised and the first 247 cases were used as the training set and the other 123 cases as the test set. The input data in both partitions were normalised to the interval (0, 1). This corresponds only to a shift in origin for the binary features and can be justified for the categorical features on the grounds that they are ordinal, that increasing value corresponds to an increasing effect,<sup>9</sup> and that the distance between each ordinal category has been designed to be equal in the BSG guidelines.<sup>1</sup>

#### Logistic regression analysis

Logistic regression was performed using a main effects only model implemented within the Statistical Package for Social Sciences (<http://www.spss.com>). For both partitions of the dataset four models were created—all input variables entered together, forward conditional stepwise analysis, backward conditional stepwise analysis, and use of the features selected as significant by the BSG guidelines in their authors' meta-analysis of the literature.<sup>1</sup> For each model, both the area under the receiver operating characteristic (ROC) curve for the test set (using the method of Hanley and McNeil),<sup>10</sup> and the odds ratios for each input variable were calculated.<sup>11</sup> The sen-

sitivities and specificities were calculated for the points on the ROC curves that gave the greatest overall accuracy.

#### Growing cell structure system analysis

Our novel variant of the growing cell structure (GCS) system performs as a conventional statistical classifier, but also produces visualisation of each input variable in relation to the known outcome. A detailed explanation of our GCS system has been given elsewhere with a clear analogy of its function.<sup>12</sup> The process is simple and divided into two steps, namely: sorting of cases into different groups and calculating predictive probabilities for each group. The cases are sorted into different groups using just the input data, without reference to the known outcome, as an unsupervised process. Each node in the network is a collection of cases which are grouped together because they have features that are more similar to each other than to the rest of the cases—for example, they all have pronounced cryptitis and severe mucin depletion. The network starts with three nodes and the cases are distributed between these three nodes so that the cases at each node are as similar to each other as possible; that is, the differences between cases in each node (error) are minimised. Once the cases have been distributed between the initial three nodes a fourth node is added, adjacent to the two nodes that have most error in them, and the cases are then redistributed between the four nodes. The process of adding new nodes is continued until adding a node does not reduce the overall error across the whole network. It should be emphasised that the whole process of sorting the cases to the nodal network is made without reference to the outcome and without any decisions by the user as to which features are more likely to be significant than others. The distribution of each input variable can be visualised on the final network using a colour contour display of the mean value for each node. This produces a series of colour contour maps for all the input variables, which can be compared with each other. Once the network of nodes has been constructed, and the cases distributed among them, then the known outcomes are used to calculate the probabilities of each outcome for each node. These probabilities are calculated using Bayes's theorem implemented by a Parzen window method.<sup>13</sup> In our study, the outcomes were CIIBD or normal and Crohn's disease or ulcerative colitis so—for example, the calculated probability could be the probability of a patient having Crohn's disease if the input variables for the biopsy

**Table 2** Odds ratios (with 95% confidence intervals (CI)) from logistic regression for all the input variables entered together from the training set of normal versus chronic idiopathic inflammatory bowel disease

| Input variable                          | Odds ratio | 95% CI                    |
|---|------------|---------------------------|
| Age                                     | 0.62       | 0.16 to 2.48              |
| Female sex                              | 0.52       | 0.31 to 0.87              |
| Mucosal surface                         | 52.88      | 0.60 to 4640              |
| Crypt architecture                      | 21.21      | 0.79 to 572               |
| Crypt profiles                          | 0.07       | 0 to 5318                 |
| Increased LP cellularity                | 0.01       | 0 to >10 <sup>6</sup>     |
| Superficially increased LP cellularity  | 8769.19    | 0 to >10 <sup>6</sup>     |
| Increased lymphoid aggregates           | 398.54     | 0 to >10 <sup>6</sup>     |
| Patchy increased LP cellularity         | 129.22     | 0 to >10 <sup>6</sup>     |
| Transmucosally increased LP cellularity | 28.13      | 0 to >10 <sup>6</sup>     |
| Cryptitis: extent                       | 0.00       | 0 to 7.59                 |
| Cryptitis: polymorphs                   | 95.59      | 0 to >10 <sup>6</sup>     |
| Crypt abscesses: extent                 | 26.28      | 0 to >10 <sup>6</sup>     |
| Crypt abscesses: polymorphs             | 0.03       | 0 to 353                  |
| Increased LP polymorphs                 | 13817.62   | 12.77 to >10 <sup>6</sup> |
| Epithelial changes                      | 6.02       | 0.01 to 4421              |
| Mucin depletion                         | 0.43       | 0.01 to 36.26             |
| Intraepithelial lymphocytes             | 8040       | 0 to >10 <sup>6</sup>     |
| Subepithelial collagen*                 | N/A        | N/A                       |
| LP granulomas                           | 750.40     | 0 to >10 <sup>6</sup>     |
| Submucosal granulomas                   | 21.72      | 0 to >10 <sup>6</sup>     |
| Basal histiocytes                       | 2.07       | 0.17 to 25.29             |

\*Submucosal collagen was not seen to be thickened in this study. N/A, not applicable; LP, lamina propria.

**Table 4** Odds ratios (with 95% confidence intervals (CI)) from logistic regression for all the input variables entered together from the training set of Crohn's disease versus ulcerative colitis

| Input variable                          | Odds ratio | 95% CI                   |
|---|------------|--------------------------|
| Age                                     | 1.88       | 0.32 to 10.97            |
| Female sex                              | 0.38       | 0.18 to 0.80             |
| Mucosal surface                         | 0.68       | 0.06 to 7.37             |
| Crypt architecture                      | 4.24       | 0.49 to 36.85            |
| Crypt profiles                          | 0.82       | 0.01 to 53.96            |
| Increased LP cellularity                | 22.59      | 0.36 to 1431             |
| Superficially increased LP cellularity  | 0.02       | 0.00 to 1.26             |
| Increased lymphoid aggregates           | 0.24       | 0.02 to 3.09             |
| Patchy increased LP cellularity         | 0.02       | 0.00 to 0.55             |
| Transmucosally increased LP cellularity | 0.45       | 0.01 to 22.45            |
| Cryptitis: extent                       | 9.46       | 0.92 to 97.38            |
| Cryptitis: polymorphs                   | 0.87       | 0.07 to 11.07            |
| Crypt abscesses: extent                 | 0.02       | 0.00 to 0.77             |
| Crypt abscesses: polymorphs             | 4.45       | 0.47 to 42.47            |
| Increased LP polymorphs                 | 0.01       | 0.00 to 0.71             |
| Epithelial changes                      | 2.24       | 0.68 to 7.41             |
| Mucin depletion                         | 0.95       | 0.15 to 5.98             |
| Intraepithelial lymphocytes             | 0.91       | 0.00 to 782              |
| Subepithelial collagen*                 | N/A        | N/A                      |
| LP granulomas                           | 0.08       | 0.01 to 0.49             |
| Submucosal granulomas                   | 0.01       | 0.00 to >10 <sup>6</sup> |
| Basal histiocytes                       | 0.74       | 0.17 to 3.11             |

\*Submucosal collagen was not seen to be thickened in this study. N/A, not applicable; LP, lamina propria.

features placed that case on that node in the network. Again, these probabilities are visualised as a colour contour display on the network. Visual comparison of the input feature overlays with the probability overlays indicates which input features have an association with a particular outcome.<sup>12</sup> The GCS system was used on both partitions of the dataset with all input variables included and just the BSG selected variables. The sensitivities and specificities were calculated for the points on the ROC curves that gave the greatest overall accuracy; that is, the points where the highest percentages were correctly classified regardless of prevalence (equal weighting to false positives and false negatives).

**Human diagnosis**

After the observer had made and recorded all the observations on a case he recorded his own diagnosis, within the categories recommended in the BSG guidelines,<sup>1</sup> still without any knowledge of the clinical details of the case.

**RESULTS**

Tables 2–7 and figs 1 and 2 summarise the results.

**DISCUSSION**

Our study examines the histological features that are discriminant in two decisions, namely: (1) Does an endoscopic colorectal biopsy show features of CIIBD or not? (2) If so, does it show features of Crohn's disease or ulcerative colitis? Two complementary methods are used to investigate these classifications.

Both methods give a good performance on the CIIBD/normality classification with areas under the ROC curves above 0.80, sensitivities of 76–80%, and specificities of 86–98% (table 6). This level of performance suggests that the input features that were included in our study do contain important discriminant value. By looking at the colour map of the posterior probability of CIIBD from the GCS system (fig 1), and comparing this with the colour maps for the input features, discriminant features can be selected by the fact that they have similarly distributed colour maps to the probability map. From this analysis, the following features appear to be discriminant for CIIBD: increased age, abnormal mucosal surface, abnormal crypt architecture, reduced crypt profiles, increased LP cellularity, patchy increase in LP cellularity, transmucosal increase in LP cellularity, extent of cryptitis and number of

**Table 3** Odds ratios (with 95% confidence intervals (CI)) from logistic regression for all the input variables selected by the labelled methods from the training set of normal versus chronic idiopathic inflammatory bowel disease

| Forward conditional entry of variables |                        | Backward conditional entry of variables |                                  | BSG selected variables          |                                 |
|--|------------------------|---|----------------------------------|---------------------------------|---------------------------------|
| Input variable                         | Odds ratio (95% CI)    | Input variable                          | Odds ratio (95% CI)              | Input variable                  | Odds ratio (95% CI)             |
| Female sex                             | 0.55 (0.33 to 0.92)    | Female sex                              | 0.52 (0.31 to 0.88)              | Mucosal surface                 | 18.34 (0.30 to 1127)            |
| Crypt architecture                     | 70.90 (5.58 to 900)    | Mucosal surface                         | 28.75 (0.43 to 1904)             | Crypt architecture              | 23.08 (1.04 to 512)             |
| Increased LP cellularity               | 3.67 (1.66 to 8.10)    | Crypt architecture                      | 18.50 (1.08 to 317)              | Crypt profiles                  | 0.42 (0 to 14239)               |
| Cryptitis: extent                      | 0.02 (0.00 to 1.15)    | Increased lymphoid aggregates           | 3.76 (1.61 to 8.80)              | Increased LP cellularity        | 5.26 (2.36 to 11.70)            |
| LP polymorphs                          | 2710 (14.21 to 517224) | LP polymorphs                           | 618.38 (36.46 to 10487)          | Patchy increased LP cellularity | 1.76 (0.45 to 6.94)             |
|  |                        | Intraepithelial lymphocytes             | 1823.29 (0 to >10 <sup>6</sup> ) | LP granulomas                   | 312.59 (0 to >10 <sup>6</sup> ) |
|  |                        | LP granulomas                           | 553.74 (0 to >10 <sup>6</sup> )  |                                 |                                 |

LP, lamina propria.

**Table 5** Odds ratios (with 95% confidence intervals (CI)) from logistic regression for all the input variables selected by the labelled methods from the training set of Crohn's disease versus ulcerative colitis

| Forward conditional entry of variables |                      | Backward conditional entry of variables |                      | BSG selected variables                  |                               |
|--|----------------------|---|----------------------|---|-------------------------------|
| Input variable                         | Odds ratio (95% CI)  | Input variable                          | Odds ratio (95% CI)  | Input variable                          | Odds ratio (95% CI)           |
| Crypt architecture                     | 5.00 (1.28 to 19.54) | Female sex                              | 0.44 (0.22 to 0.90)  | Mucosal surface                         | 0.98 (0.12 to 7.77)           |
| Superficially increased LP cellularity | 0.11 (0.01 to 1.35)  | Crypt architecture                      | 4.90 (1.22 to 19.71) | Crypt architecture                      | 4.85 (0.66 to 35.7)           |
| Patchy increased LP cellularity        | 0.27 (0.13 to 0.55)  | Superficially increased LP cellularity  | 0.07 (0 to 1.02)     | Crypt profiles                          | 0.63 (0.01 to 34.19)          |
| LP granulomas                          | 0.08 (0.02 to 0.41)  | Patchy increased LP cellularity         | 0.11 (0.03 to 0.39)  | Patchy increased LP cellularity         | 0.39 (0.12 to 1.31)           |
|  |                      | Cryptitis: extent                       | 6.85 (1.28 to 36.73) | Transmucosally increased LP cellularity | 1.31 (0.35 to 4.96)           |
|  |                      | Increased LP polymorphs                 | 0.05 (0 to 0.75)     | LP granulomas                           | 0.09 (0.02 to 0.48)           |
|  |                      | LP granulomas                           | 0.07 (0.01 to 0.37)  | Submucosal granulomas                   | 0.01 (0 to >10 <sup>6</sup> ) |

BSG, British Society of Gastroenterology; LP, lamina propria.

**Table 6** Areas under the receiver operating characteristic (ROC) curves for logistic regression and the growing cell structure (GCS) system for the test set of the normal versus CIIBD data set

|                            | Logistic regression |                     |                      |              | GCS system    |              | Human performance |                    |
|----------------------------|---------------------|---------------------|----------------------|--------------|---------------|--------------|-------------------|--------------------|
|                            | All variables       | Forward conditional | Backward conditional | BSG selected | All variables | BSG selected | Initial diagnosis | Observer diagnosis |
| Area under ROC curve       | 0.8436              | 0.8680              | 0.8678               | 0.8775       | 0.8393        | 0.8532       | N/A               | N/A                |
| SE of area under ROC curve | 0.0245              | 0.0221              | 0.0221               | 0.0211       | 0.0230        | 0.0231       | N/A               | N/A                |
| Sensitivity for CIIBD      | 78%                 | 79%                 | 80%                  | 80%          | 76%           | 77%          | 84%               | 82%                |
| Specificity                | 89%                 | 89%                 | 89%                  | 89%          | 86%           | 89%          | 98%               | 99%                |

BSG, British Society of Gastroenterology; CIIBD, chronic idiopathic inflammatory bowel disease; N/A, not applicable; SE, standard error.

**Table 7** Areas under the receiver operating characteristic (ROC) curves for logistic regression and the growing cell structure (GCS) system for the test set of the Crohn's disease versus ulcerative colitis (UC) data set

|                            | Logistic regression |                     |                      |              | GCS system    |              | Human performance |                    |
|----------------------------|---------------------|---------------------|----------------------|--------------|---------------|--------------|-------------------|--------------------|
|                            | All variables       | Forward conditional | Backward conditional | BSG selected | All variables | BSG selected | Initial diagnosis | Observer diagnosis |
| Area under ROC curve       | 0.7564              | 0.7596              | 0.7812               | 0.7771       | 0.7202        | 0.7334       | N/A               | N/A                |
| SE of area under ROC curve | 0.0464              | 0.0460              | 0.0437               | 0.0442       | 0.0466        | 0.0465       | N/A               | N/A                |
| Sensitivity for UC         | 84%                 | 63%                 | 69%                  | 67%          | 66%           | 74%          | 65%               | 61%                |
| Specificity                | 62%                 | 87%                 | 78%                  | 79%          | 66%           | 62%          | 98%               | 76%                |

BSG, British Society of Gastroenterology; N/A, not applicable; SE, standard error.

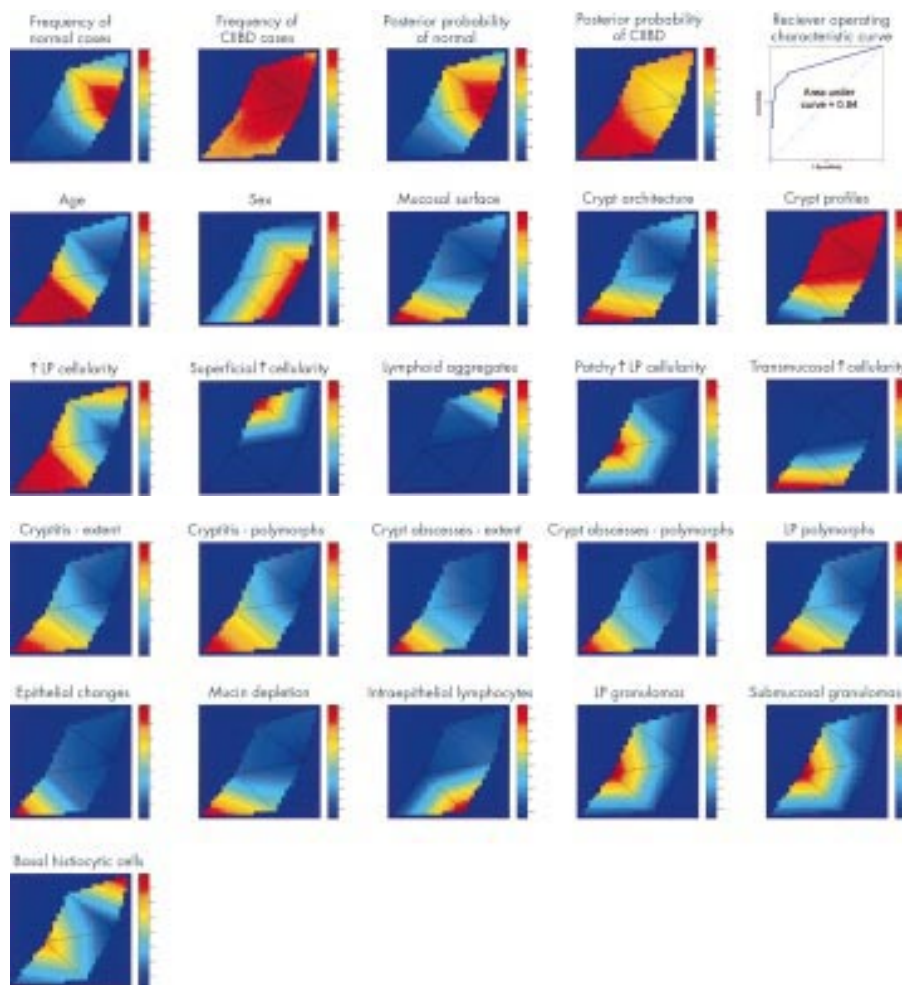
polymorphs, extent of crypt abscesses and number of polymorphs, LP polymorphs, epithelial degeneration and ulceration, mucin depletion, increased intraepithelial lymphocytes, LP granulomas, and submucosal granulomas. All these features agree with the canonical information in the published literature.<sup>1</sup>

When looking at the odds ratios for the input variables for the logistic equation, which produces a similar level of classification performance, a rather different picture emerges. With all input features entered together, only one input feature has an odds ratio with the lower 95% confidence interval above one, namely: increased LP polymorphs (table 2). Thirteen other features have odds ratios above one but with wide 95% confidence intervals, extending to below one, so they are not significant. This is probably because of the high degree of correlation between many of the input features so that one feature, such as increased LP polymorphs, may have a very high odds ratio, but all the other input variables that are highly correlated with it will have lower odds ratios. An advantage of the GCS system is that the initial clustering of cases in the network is unsupervised, so that all discriminant input variables can be discerned, even if they are highly correlated. However, the input features that have odds ratios above one

(but without a lower 95% confidence interval above one) do correspond closely to features that are cited in the literature as being discriminant.<sup>1</sup> The logistic regression methods using selected variables (table 3) produce a slightly different selection of variables with forward and backward conditional entry, but LP polymorphs, abnormal crypt architecture, and male sex feature in both selections.

The distinction between ulcerative colitis and Crohn's disease in cases known to be CIIBD appears to be a more difficult classification task, with a lower area under the ROC curves for both logistic regression and the GCS system (in the range 0.72–0.78, table 7). Looking at the colour map overlays from the GCS system (fig 2), the following input features have maps that correspond closely with the area of a high probability of Crohn's disease—patchy increase in LP, LP granulomas, and submucosal granulomas. These input features correlate with the area of high probability of ulcerative colitis—abnormal mucosal surface, abnormal crypt architecture, reduced crypt profiles, transmucosal increase in cellularity, extent and number of polymorphs in cryptitis, extent and number of polymorphs in crypt abscesses, increased LP polymorphs, degenerate or ulcerated surface epithelial cells, and pronounced mucin depletion. The following features appear to





**Figure 1** The frequency, probability, and input variable maps for the growing cell structure (GCS) network used to distinguish between normality and chronic idiopathic inflammatory bowel disease (CIIBD). The black triangular grid represents the network of nodes, with a node at each angle. The values for the parameter displayed on each map are shown as a colour contour overlay, with dark blue representing low values and red representing high values. The two top left maps show the values of the frequency of normal and CIIBD cases on each node after distribution of all the test set cases. It can be seen that the normal cases are almost all clustered on a single node on the right hand side of the map, whereas the CIIBD cases are distributed widely over the whole map. However, when the posterior probabilities are displayed (next two maps), the posterior probability of CIIBD is much lower at the node where the probability of normality is highest. If the colour map for the posterior probability of CIIBD is taken as a reference, then any input variable maps with a similar distribution of colours are predictive of the outcome of CIIBD. It can be seen that mucosal surface, crypt architecture, low number of crypt profiles, increased lamina propria (LP) cellularity of all patterns of distribution, cryptitis, crypt abscesses, LP polymorphs, epithelial changes, mucin depletion, intraepithelial lymphocytes, granulomas, and basal histiocytic cells all have distributions that are similar to at least part of the distribution of the probability of CIIBD, and are therefore likely to be predictive for that outcome. The receiver operating characteristic curve for the classification of the independent test set is also shown.

be non-discriminant because they cover areas of high probability for both ulcerative colitis and Crohn’s disease—age, sex, increased lymphoid aggregates, intraepithelial lymphocytes, and basal histiocytic cells. The selection of these features by the GCS system in our study is again in broad agreement with the published literature.<sup>1</sup>

In the logistic regression method, the outcome of Crohn’s disease was coded as zero and that of ulcerative colitis as one. Thus, in tables 4 and 5 any odds ratios above one will indicate an input feature that was discriminant for ulcerative colitis. In table 4 no features have odds ratios with lower 95% confidence intervals above one (indicating significant discriminant value for a diagnosis of ulcerative colitis). The following features in table 4 have odds ratios with an upper 95% confidence interval below one (indicating significant discriminant value, at a 5% level, for a diagnosis of Crohn’s disease): female sex, patchy increased LP cellularity, extent of crypt abscesses, increased LP polymorphs, and LP granulomas. With forward conditional entry of variables, an abnormal glandular architecture favours ulcerative colitis and a superficial increase in LP cellularity, patchy increase in LP cellularity, and LP granulomas favour

Crohn’s disease. Backward conditional entry adds the extent of cryptitis as a feature favouring ulcerative colitis and female sex and increased LP polymorphs as features favouring Crohn’s disease. As for the CIIBD/normality decision, the logistic regression does not identify all the features that might be discriminant if they are highly correlated with other significant input features.

Both techniques in our study identify histological features that are discriminant for the distinction between CIIBD and normality, and between Crohn’s disease and ulcerative colitis in cases of CIIBD. The specific features that they identify show some variation between the techniques, with the GCS system identifying more features than logistic regression, because in this last technique only one feature may be selected from a group of highly correlated variables. The features that these techniques identify are in agreement with the published literature<sup>1</sup> and no additional discriminant features are revealed. Thus, our study provides confirmation of existing knowledge, and it does so in the context of a large dataset with many cases and carefully defined features. Tables 6 and 7 show some interesting differences between the trained classification

### Take home messages

- Histopathological examination of endoscopic colorectal biopsies is useful to distinguish between normal subjects and those with chronic idiopathic inflammatory bowel disease
- It is less good at distinguishing between ulcerative colitis and Crohn's disease

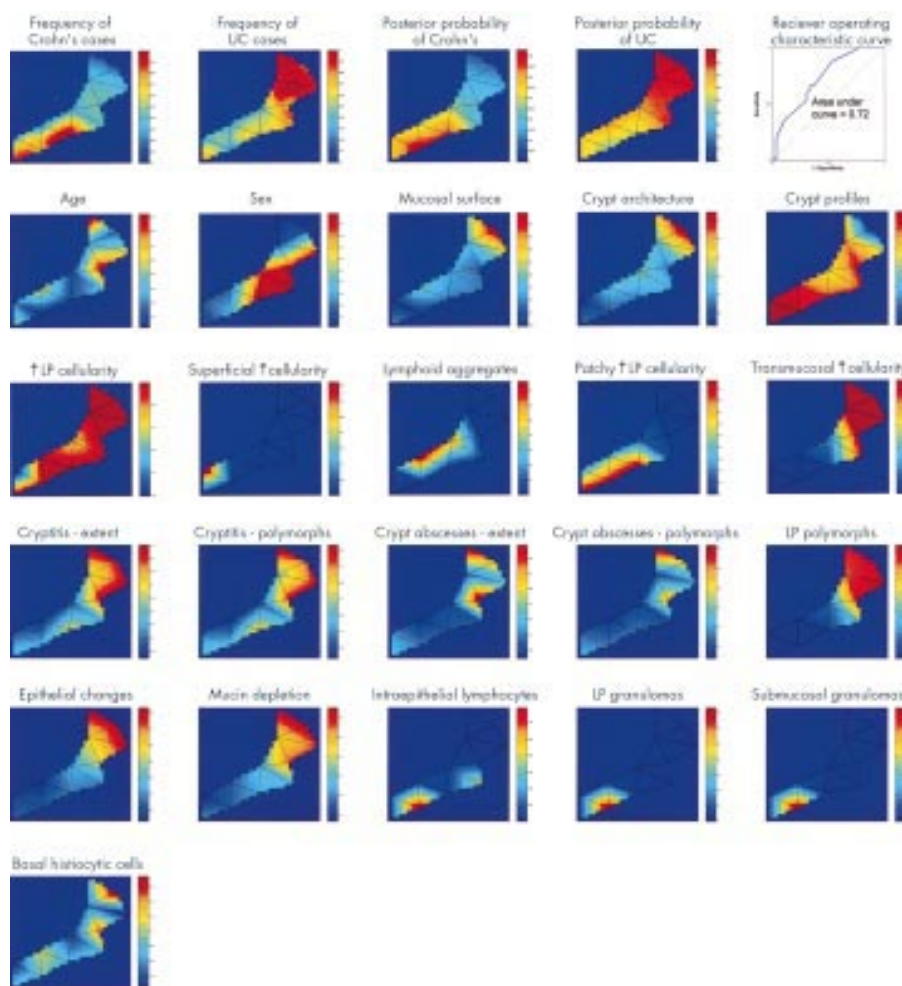
systems and human performance. For the distinction between normality and CIIBD, both the observing histopathologist, and the initial reporting histopathologist had slightly higher sensitivities and much higher specificities than any of the classification techniques. This suggests that either the histopathologists must be using some diagnostic features that are not encoded by the defined observed features in our study or that they are integrating them in a more subtle way than the classification techniques. Visual images contain immense amounts of data and encoding these into discrete variables, even as many as were used in our study, leads to a considerable loss of information.<sup>14</sup> It may be that even more categories need to be introduced for each feature to provide sufficiently rich information for the classification systems to produce optimal predictive power.<sup>15</sup> For the distinction between ulcerative colitis and Crohn's disease (table 7) the human performance differs between the initial reporting and the observing his-

topathologists. Whereas the sensitivities for ulcerative colitis, and specificity of the observing pathologist, are similar to the classification techniques, the specificity for Crohn's disease of the initial reporting histopathologists is much higher (98%) than any other method. The initial reporting pathologists had all the clinical information that was on the histopathology request form available to them when making the diagnosis. This may have contained information about small bowel radiological studies or other investigations that may have facilitated the diagnosis. Because the observing pathologist's performance was not better than any of the classification techniques, this suggests that there are not histopathology features additional to those included in our study that are of discriminatory value. However, the relatively modest performance of any system in the Crohn's disease from ulcerative colitis distinction suggests that histopathology will never provide perfect classification of these diseases, and that other investigations are required to add information.

### Authors' affiliations

**S S Cross**, Academic Unit of Pathology, Section of Oncology and Pathology, Division of Genomic Medicine, University of Sheffield, Sheffield S10 2UL, South Yorkshire, UK

**R F Harrison**, Department of Automatic Control and Systems Engineering, University of Sheffield



**Figure 2** The frequency, probability, and input variable maps for the growing cell structure (GCS) network used to distinguish between ulcerative colitis and Crohn's disease. The triangular grid and the colour contour maps are constructed as described for fig 1. If the colour map for the posterior probability of Crohn's disease is taken as a reference, then any input variable maps with a similar distribution of colours—for example, patchy increase in lamina propria (LP) cellularity—are predictive of the outcome of Crohn's disease. The receiver operating characteristic curve for the classification of the independent test set is also shown.

## REFERENCES

- 1 **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.
- 2 **Dundas SAC**, Dutton J, Skipworth P. Reliability of rectal biopsy in distinguishing between chronic inflammatory bowel disease and acute self-limiting colitis. *Histopathology* 1997;**31**:60–6.
- 3 **Dube AK**, Cross SS, Lobo AJ. Audit of the histopathological diagnosis of non-neoplastic colorectal biopsies: achievable standards for the diagnosis of inflammatory bowel disease. *J Clin Pathol* 1998;**51**:378–81.
- 4 **Cook MG**, Dixon MF. An analysis of the reliability of detection and diagnostic value of various pathological features in Crohn's disease and ulcerative colitis. *Gut* 1973;**14**:255–62.
- 5 **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.
- 6 **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.
- 7 **Le Berre N**, Heresbach D, Kerbaol M, *et al*. Histological discrimination of idiopathic inflammatory bowel disease from other types of colitis. *J Clin Pathol* 1995;**48**:749–53.
- 8 **Cross SS**, Dube AK, Underwood JCE, *et al*. A visual image-based data entry system for reporting non-neoplastic endoscopic colorectal biopsies. *J Pathol* 1997;**181**:17A.
- 9 **Tarassenko L**. *A guide to neurocomputing applications*. London: Arnold, 1998.
- 10 **Hanley JA**, McNeil BJ. The meaning and use of the area under a receiver operating characteristic (ROC) curve. *Radiology* 1982;**143**:29–36.
- 11 **Feinstein AR**, ed. Multiple logistic regression. In: *Multivariable analysis*, 1st ed. New Haven: Yale University Press, 1996:297–330.
- 12 **Walker AJ**, Cross SS, Harrison RF. Visualisation of biomedical datasets by use of growing cell structure networks: a novel diagnostic classification technique. *Lancet* 1999;**354**:1518–21.
- 13 **Parzen E**. On estimation of a probability density function and mode. *Annals of Mathematical Statistics* 1962;**33**:1065–76.
- 14 **Cross SS**. Grading and scoring in histopathology. *Histopathology* 1998;**33**:99–106.
- 15 **Morris JA**. Information and observer disagreement in histopathology. *Histopathology* 1994;**25**:123–8.



## Direct Access to Medline

**Medline**

Link to Medline from the homepage and get straight into the National Library of Medicine's premier bibliographic database. Medline allows you to search across 9 million records of bibliographic citations and author abstracts from approximately 3,900 current biomedical journals.

**[www.jclinpath.com](http://www.jclinpath.com)**