Assessment of severity in colitis: A preliminary study

J. E. LENNARD-JONES[1], JEAN K. RITCHIE, W. HILDER, AND C. C. SPICER

From St Mark's Hospital, London, and the MRC Division of Medical Computing, Northwick Park Hospital, London

SUMMARY   Clinical features, alone and in combinations of up to five, noted early in patients during 181 admissions to one hospital for treatment of acute colitis over five years, have been correlated with the success or failure of drug treatment as judged by death during medical treatment or the need for urgent surgical treatment. Many of the 56 clinical features studied were of no value in predicting the outcome of the attack. The four features of greatest predictive value were the maximum daily body temperature, the maximum daily pulse rate, the bowel frequency and plasma albumin. The simplest and most discriminating prediction at the end of the first 24 hours in hospital was obtained by combining observations on maximum body temperature and the number of stools passed. Serial observations of temperature or pulse rate over the first four days in hospital gave more prognostic information than observations confined to the first day. The predictive value of certain other features, such as *x*-ray appearances, is described.

A simple classification of severity in acute colitis is proposed from these results as a basis for prospective testing in other hospitals.

The high mortality after surgical treatment for acute colitis in non-specialist hospitals suggests that the severity of the disease is often underestimated. Admissions for acute colitis are relatively uncommon and a survey of results in 34 hospitals over a long period showed that, in most hospitals, only one operation for acute colitis was performed each year (Ritchie, 1974). There appears to be a need for a simple classification of severity in colitis so that the severely ill patient can be recognized without delay by clinicians seeing this disease infrequently.

It would be helpful if existing classifications of severity could be simplified for general use. At St Mark's Hospital many patients with acute colitis are treated and an attempt has been made to correlate the clinical features soon after admission to hospital with the outcome of the attack. From this retrospective analysis it has been possible to define simple criteria associated with a failure to respond to medical treatment. The results of the study are presented in this paper. With the knowledge gained, a prospective study in district general hospitals is planned and these results will be published in due course.

Composition of the Series

A retrospective sample consisting of all patients

admitted to St Mark's Hospital during the period 1 January 1968 - 31 December 1972 was studied. In every patient the diagnosis of proctocolitis was established on the basis of clinical, radiological and pathological findings. Patients admitted for investigation, assessment or elective surgical treatment were excluded, resulting in 189 admissions (in 166 patients) for treatment of active colitis. The majority of patients with active colitis are managed at this hospital as outpatients and the series therefore includes only patients unresponsive to treatment outside hospital or patients with disease regarded as too severe for outpatient management.

Recording of Data

A proforma was designed for the analysis of the clinical features of the illness at the time of admission, the results of investigation and the outcome of the attack.

The following features were recorded in addition to age, sex, and date of admission to hospital.

A HISTORY TO ADMISSION (6)
Duration of illness from first symptom reasonably attributable to colitis: type of disease defined as: (a) 'first attack', continuous symptoms for less than one year; (b) 'continuous symptoms' for more than a year since onset; (c) 'intermittent', one or more periods

without symptoms: duration of attack: approximate weight loss: presence of anorexia or nausea: bowel frequency.

B PHYSICAL EXAMINATION (15)
Presence of severe abdominal tenderness or distension: anal complications: associated conditions such as erythema nodosum, pyoderma gangrenosum, arthritis, oral ulcers or monilia, eye lesions, evidence of liver disease or venous thrombosis: sigmoidoscopic appearance.

C WARD OBSERVATIONS (11)
Maximum temperature on each of first four days after admission: maximum pulse rate on each of first four days after admission: observed bowel frequency, consistency of the stool and the amount of blood during the first 24 hours.

D LABORATORY INVESTIGATIONS (13)
Haemoglobin, white cell count, ESR: serum proteins (total and albumin), blood urea and electrolytes and liver function tests.

E RADIOLOGY (8)
Plain abdominal radiograph—evidence of perforation, dilatation or mucosal islands: barium enema—extent of disease, granularity of mucosal line, ulceration.

F OUTCOME OF ADMISSION
(1) Medical response—discharged well or improvement followed by elective surgery; (2) medical failure—death during medical treatment or urgent operation required; (3) no medical treatment, early operation.

Method of Analysis

The data were recorded in a form suitable for punch card analysis. A comparison was made between the clinical features of the patients in the 144 admissions where medical response occurred (group 1) and the 37 admissions where medical treatment failed (group 2). For the purposes of the analysis the eight admissions in which the patient was judged so ill that immediate surgical treatment was undertaken have been considered separately (group 3).

Statistical Methods

The main statistical problem is to determine which subset best replaces the total number of 56 clinical

attributes recorded. Account has also to be taken of missing values since every attribute was not recorded in every patient. In some cases the absence of a result reflects the severity of the disease; for example, a barium enema is not performed in the very ill patient.

As the first step the individual attributes were examined to find those which differed significantly between the two outcome groups and to eliminate those which occurred too infrequently or in which the data were insufficient for useful analysis. Significance of differences for continuous variables was assessed by Student's t test and for discontinuous variables by the χ^2 test.

Having identified single attributes in which the two outcome groups differed, the predictive value of each attribute when present in an individual patient was calculated using Bayes theorem. For this purpose, continuous variables were divided into categories and treated as discrete.

It seemed likely that a combination of two or more attributes would give better prediction of medical success or failure than a single attribute. Two techniques were employed to find the best combinations of attributes:

STEPWISE LINEAR DISCRIMINANT ANALYSIS
The theoretical basis for using this technique for data combining quantitative and qualitative attributes and having no well defined statistical distribution is not at all well understood. Because of these difficulties it is not possible to assert that the results we have obtained are optimal in the sense that a linear discriminant function is optimal for distinguishing between normally distributed variables with a common variance/covariance matrix.

ANALYSIS OF SYMPTOM PATTERNS
An enumeration of all the possible subsets of up to five attributes (after categorization of continuous variables) to see which gives the best discrimination is a type of analysis of symptom pattern only feasible for up to four to five attributes as the number of possible patterns then becomes very large and the numbers in any given pattern becomes too small for useful results to be obtained (Spicer, Jones, and Lennard-Jones, 1973).

There remains the problem of giving expression to the predictive power of the attributes selected by the above procedures. If the theoretical basis of the discriminant function was known, this would not be necessary but in the present example this is far from being the case. The method used has been to express the proportion of patients in each pattern having an unfavourable outcome in terms of the logarithm of the odds on failure. Mathematical

expressions were fitted to the log-odds to assess the magnitude of the individual effects and inter-actions of the attributes of which the pattern is made up. The theory of this technique has been set out by Nelder and Wedderburn (1972) and more simply by Snedecor and Cochran (1967). Nelder (1974) has developed a set of computer programs for the calculations. This method is extremely flexible and enables the significance of high degrees of inter-actions between attributes to be tested and also the adequacy of the grouping of the quantitative variables. Some method of dealing with the latter problem is needed since it is clear that a very fine grouping, for example, by half a degree of tempera-ture, might lead to very small numbers in some categories while too coarse a grouping might lose information on the quantitative effects.

Missing values were treated in a number of ways. When each attribute was examined individually cases which had values missing for the attribute were omitted. When calculating the stepwise dis-criminant they were replaced by dummy values calculated by a method developed by Little (1973) based on that of Buck (1960) which allows unbiased estimates of the parameters to be made assuming that the missing values are not selected by some discriminating property. This was clearly not so in the case of the barium enema and its absence was explicitly included as a discriminating factor in the pattern analysis.

Results

OUTCOME OF THE ATTACK

The outcome of the admissions is shown in table I. It will be seen that medical response occurred in 144 of the 189 admissions (76·2%), medical treatment failed in 37 (19·6%) and immediate surgical treat-ment was undertaken on eight occasions (4·2%). There were four deaths, one during medical treat-ment and three after urgent surgical treatment.

The indication for immediate surgical treatment (group 3) in most cases was evidence on plain abdominal radiography of dilatation of the colon with mucosal projections into the lumen (islands)

and no analysis of the prognostic value of other clinical features in these patients is possible.

PREDICTIVE VALUE OF SINGLE ATTRIBUTES

Continuous variables were compared by calculating the mean in each group and comparing the signifi-cance of differences using Student's t test. Discon-tinuous variables were tabulated and the two groups were compared using the χ^2 test. Non-discriminating variables are shown in table II. Useful values were only found for the variables shown in tables III and IV. It will be seen that four attributes predicted that successful medical treatment was more likely than the prior probability of about 80%, and that 15 attributes suggested that failure of medical treat-ment was more likely than the prior probability of 20%. Single attributes available at the end of the first day gave no certainty of prediction, the greatest being 56%. Three attributes which became avail-able or occurred after admission had a predictive value greater than 80%. The most useful of these was a raised body temperature which persisted despite drug therapy and a temperature greater than 38°C on the fourth day was associated with about one chance in eight of successful medical treatment.

PREDICTIVE VALUE OF COMBINATIONS OF TWO OR THREE ATTRIBUTES FOR THE INDIVIDUAL PATIENT

Both discriminant and pattern analysis showed that the most useful combinations of attributes were derived from the maximum temperature or pulse rate (which are closely correlated) on any day, the observed bowel frequency on the day of admission and the plasma albumin level. It also appeared that there was little to be gained with the present number of observations by more than broad subdivisions of the values for temperature, pulse rate, bowel frequency and albumin level.

The data for maximum temperature and observed bowel frequency during the first 24 hours after admission can be presented in the form of a simple table (table V). It will be seen that if eight stools or fewer were passed on the first day and the maximum

Group 1		Group 2			Group 3
Medical Response		*Medical Failure*			*No Medical Treatment*
Discharged Well[1]	*Improved: Later Elective Surgery*	*Death during Medical Treatment*	*Urgent Operation; Survived*	*Died*	*Early Operation*
136	8	1	33	3	8

Table I   *The outcome of 189 attacks of colitis*
[1]One patient discharged unchanged.

History on Admission	Findings on Admission[1]	Laboratory Findings on Admission[1]
Age and sex of patient	Abdominal distension	Haemoglobin
Total length of history	Amount of blood in the stools	White cell count
Duration of present attack	Erythema nodosum, aphthous ulceration,	ESR
First attack or relapse	eye lesions, arthritis, liver disease, venous	
Stated bowel frequency	thrombosis	
Stated weight loss	Perianal disease	Plasma potassium
	Sigmoidoscopic appearance	Plasma urea
		Plasma bilirubin
		SGOT or SGPT
		Plasma alkaline phosphatase

Table II   Clinical attributes not of predictive value in assessing the likelihood of success or failure of medical treatment

[1]The following findings were significantly different but the separation achieved was not adequate for useful discrimination—plasma sodium, pyoderma gangrenosum, rectovaginal fistula and inflammatory polyps on sigmoidoscopy.

Attribute	$p(A/R)$[1]	$p(A/F)$[2]	$p(F/A)$[3]
			(Medical Failure) (%)
Formed stool	11	0	0
Bowel actions $\leqslant$ 3	23	8	7
Bowel actions $\leqslant$ 5	47	19	8
Anorexia or nausea	65	88	26
Abdominal tenderness	37	67	31
Bowel actions > 9	16	38	33
Pulse $\geqslant$ 100/min	27	57	36
Maximum temperature > 37·5°C	15	41	42
Maximum temperature > 38·0°C	6	27	56
Bowel actions > 12	3	19	55

Table III   Predictive factors at the end of the first 24 hours in hospital

[1]p(A/R) indicates the percentage of patients with the attribute shown among those who responded to medical treatment.
[2]p(A/F) indicates the percentage of patients with the attribute shown who failed to respond to medical treatment.
[3]p(F/A) indicates the percentage of patients who failed to respond to medical treatment when the attribute shown was present.
    The parts of the table above and below the horizontal line indicate values below and above the prior probability of medical failure (20%.)
    Bowel frequency refers to observed bowel frequency after admission to hospital. Temperature and pulse rate refer to the maximum value recorded during the 24 hours.

	$p(A/R)$	$p(A/F)$	$p(F/A)$ (Medical Failure) (%)
Barium Enema			
Normal or distal colitis	27	3	3
Substantial or extensive colitis	73	97	22
Ulceration	42	80	27
Albumin < 30 g/l	24	60	42
Oral monilia	9	27	43
Day 4			
Temperature > 37·5°C	10	42	52
Temperature > 38°C	1	19	88
Abdominal Radiograph			
Mucosal islands	1	16	75
Dilatation	2	24	75

Table IV   Predictive factors after the first day in hospital[1]
[1]Abbreviations and definitions are as in table III.

temperature was less than 38°C there was an excellent chance of successful medical treatment. Conversely, if both these values were exceeded, the probability of successful medical treatment was only one in five. Intermediate values indicated that the likelihood of successful medical treatment was less than the prior probability of 80%.

A combination of three variables is difficult to present in a simple form for routine clinical use. The probabilities are presented as a table in table VI. The data can also be presented in the form of a nomogram.

Discussion

The earliest study of predictive factors in colitis

Number of Bowel Actions	Body Temperature	
	≤ 38°C	*> 38°C*
0-5	4 (56)[1]	38 (8)
6-8	9 (43)	60 (5)
≥ 9	34 (32)	80 (5)

Table V   *The estimated percentage of patients failing to respond to medical treatment in an attack of colitis correlated with the maximum temperature and the observed bowel frequency during the first 24 hours in hospital*

[1]Figures in brackets show the actual number of cases in each group. Only cases with information on both attributes are included.

Bowel Actions per 24 Hours	Serum Albumin (g/l)	Pulse Rate	
		≤ 89	*≥ 90*
	< 30	11 (7)[1]	47 (8)
0-5	30-35	5 (18)	13 (11)
	> 35	1 (23)	6 (9)
	< 30	16 (5)	58 (12)
6-9	30-35	8 (19)	27 (12)
	> 35	3 (13)	22 (1)
	< 30	32 (8)	62 (12)
> 9	30-35	14 (8)	44 (10)
	> 35	7 (1)	22 (3)

Table VI   *Estimated percentage of patients failing to respond to medical treatment classified by frequency of bowel movements on day of admission, serum albumin, and pulse rate on first day after admission*

[1]Figures in brackets show the actual number of cases in each group. Missing values have been estimated by the method of Buck (1960).

(Cameron and Rippman, 1910) showed that a sustained temperature of 101°F or more was associated with a high mortality. In 1955 Truelove and Witts introduced a classification of the severity of colitis which has been widely used ever since. This classification takes into account five variables, including temperature, pulse rate, bowel frequency with the presence of blood in the stools, haemoglobin level, and ESR. These variables were chosen and critical values selected by the authors from their wide clinical experience of the disease. Three grades of severity were suggested and retrospective surveys have shown that these grades correlate well with the outcome of an attack (Edwards and Truelove, 1963; Watts, de Dombal, Watkinson, and Goligher, 1966). The difficulty with such a classification using five variables is that not all the conditions may be fulfilled at one time, particularly during the first few days after the patient is admitted to hospital. It would be a further advantage if clinical observations could be expressed in terms of the probability of medical response or failure.

A different approach has been adopted by Card (Jalan, Prescott, Sircus, Card, McManus, Falconer, Small, Smith, and Bruce, 1970, 1971) who asked a group of experienced clinicians to rank a series of case histories in order of increasing severity of the disease without knowledge of the outcome of the attack. There was substantial agreement between the clinicians and a communal ranking order was established. Several attributes were then chosen from the case histories and the correlation of each attribute with the consensus ranking order was calculated. A multiple regression of five variables with the consensus ranking of severity was calculated. The closest agreement was reached when all five attributes were included but the agreement was almost as good when the analysis was restricted to temperature, number of bowel movements, and ESR. The group of patients was arbitrarily divided into three approximately equal subgroups denoted as having severe, moderate or mild colitis. An individual patient could then be given a score y from the formula: $y = 0.84 x_1 + 0.12 x_2 + 0.35 x_3 - 0.59$, where x_1, x_2 and x_3 are variables derived from the measurement of temperature, number of bowel movements, and erythrocyte sedimentation rate, respectively. x_2 and x_3 are continuous variables in the range 0 to 4 derived from the original measurements by reading off corresponding values from a simple graph. x_1 assumes the value 3.13, 2.28 or 1.23 depending on whether the oral temperature was greater than 100°F, less than 100°F or normal. Patients were then classified as severe if y was greater than 2.5, mild if less than 1.328 and moderate in other cases. This classification was subsequently applied retrospectively to a series of patients and agreed well with the outcome of the illness.

The present study is based on a different premise. The outcome of the attack was defined first and those clinical features available during the first day and during the first few days after admission have been correlated with the outcome. Of necessity, this analysis reflects those clinical features which influence clinicians in their judgment about the management of each acute attack of colitis. This work must be regarded as a preliminary study and it is planned to test the conclusions prospectively in a further series of patients treated in non-specialist hospitals.

At the end of the first 24 hours in hospital no single attribute gave more than an indication that medical treatment might fail, though a normal consistency of stool and a low bowel frequency suggested that medical treatment was likely to succeed. The most useful predictive features at this early stage were the maximum temperature and the observed bowel frequency taken in combination. A patient with a bowel frequency of less than eight on the first day *and* a maximum temperature of

38°C or less had a high probability of responding to medical treatment. Conversely, a patient with a higher bowel frequency *and* a higher temperature had a four in five chance of failing to respond to drug therapy with a consequent need for operation.

After the first day in hospital a sustained fever (usually despite vigorous treatment) was of serious import so that a patient with a maximum temperature greater than 38°C on the fourth day had about one chance in eight of responding to medical treatment. The appearance of colonic dilatation, mucosal islands, or both, on a plain abdominal radiograph was of equally serious significance but this finding in part reflects a policy of management in the hospital.

A barium enema was most helpful if it showed a normal appearance of the colon or distal colitis. In this event medical treatment was usually successful. The finding of substantial or extensive colitis or the presence of colonic ulceration was associated with an increased likelihood of surgical treatment, as already shown by others (Watts *et al*, 1966; de Dombal, Geffen, Darnborough, Watkinson, and Goligher, 1968) but the predictive value of these findings was small in the individual patient. Avoidance of a barium enema in severely ill patients limits the clinical usefulness of this examination in acute colitis and therefore the predictive value of these two radiological features is underestimated.

We suggest that a combination of observed bowel frequency and maximum body temperature during the first day in hospital may provide a guide to the likelihood of successful medical treatment. Patients who pass no more than eight stools and have a fever of 38°C or less can be regarded as unlikely to require surgical treatment during that attack. Patients with a combination of more than eight stools in the first 24 hours and a maximum temperature greater than 38°C should be regarded as possible candidates for surgical treatment during that attack and surgical consultation should be arranged so that operation can be undertaken without delay if it proves necessary. A similar attitude should be taken to patients in whom fever persists for several days despite drug therapy or in whom colonic

dilatation or mucosal islands appear on serial plain abdominal *x*-ray examinations.

Attention is drawn to other clinical features such as a low plasma albumin, abdominal tenderness and oral monilia which are of less use as predictive attributes but which give some indication that an attack of colitis carries a greater risk than expected of unsuccessful medical treatment. Patients with any of these features or with intermediate bowel frequency and/or temperature should be observed carefully and treated vigorously until the outcome of medical treatment becomes clear.

References

Buck, S. F. (1960). A method of estimation of missing values in multi-variate data suitable for use with an electronic computer. *J. roy. statist. Soc.* B, **22**, 302-306.

Cameron, H. C., and Rippman, C. H. (1910). The post-mortem statistics of ulcerative colitis at Guy's Hospital from 1888 to 1907. *Guy's hosp. Rep.*, **64**, 353-371.

de Dombal, F. T., Geffen, N., Darnborough, A., Watkinson, G., and Goligher, J. C. (1968). Radiological appearances of ulcerative colitis: an evaluation of their clinical significance. *Gut*, **9**, 157-163.

Edwards, F. C., and Truelove, S. C. (1963). The course and prognosis of ulcerative colitis. Part I. *Gut*, **4**, 299-308.

Jalan, K. N., Prescott, R. J., Sircus, W., Card, W. I., McManus, J. P. A., Falconer, C. W. A., Small, W. P., Smith, A. N., and Bruce, J. (1970). An experience of ulcerative colitis. Part II. Short term outcome. *Gastroenterology*, **59**, 589-597.

Jalan, K. N., Prescott, R. J., Sircus, W., Card, W. I., McManus, J. P. A., Falconer, C. W. A., Small, W. P., Smith, A. N., and Bruce, J. (1971). Ulcerative colitis. A clinical study of 399 patients. *J. roy. Coll. Surg. Edinb.*, **16**, 338-355.

Little, R. (1973). Personal communication.

Nelder, J. A. (1974). GLIM manual, Oxford, Numerical Algorithms Group, 13 Banbury Road.

Nelder, J. A., and Wedderburn, R. W. M. (1972). Generalized linear models. *J. roy. statist. Soc. A.*, **135**, 370-384.

Ritchie, J. K. (1974). Results of surgery for inflammatory bowel disease: a further survey of one hospital region. *Brit. med. J.*, **1**, 264-268.

Snedecor, G. W., and Cochran, W. G. (1967). *Statistical Methods*, ch. 16, 6th ed. Iowa State University Press, Ames.

Spicer, C. C., Jones, J. H., and Lennard-Jones, J. E. (1973). Discriminant and Bayes analysis in the differential diagnosis of Crohn's disease and proctocolitis. *Meth. Inf. Med.*, **12**, 118-122.

Truelove, S. C., and Witts, L. J. (1955). Cortisone in ulcerative colitis: final report on a therapeutic trial. *Brit. med. J.*, **2**, 1041-1048.

Watts, J. McK., de Dombal, F. T., Watkinson, G., and Goligher, J. C. (1966). Early course of ulcerative colitis. *Gut*, **7**, 16-31.