The Prediction of Success or Failure of Medical Treatment for Acute Colitis

C. C. SPICER, M R C P, M R C Computer Unit, Clinical Research Centre, Northwick Park, Harrow, and J. E. LENNARD-JONES, M D, F R C P, Physician, St Mark's Hospital, London, and Reader in Medicine, The London Hospital Medical College

Acute colitis is a potentially dangerous disease and failure to respond to drug therapy necessitates surgical treatment by colectomy.

A survey in non-specialist hospitals suggests that the severity of colitis is often under-estimated (Ritchie, 1974) and this may contribute to a high mortality after colectomy. There is a need for a simple classification of severity in colitis so that the severely ill patient can be recognised without delay by clinicians unfamiliar with the disease. The present investigation aims to develop simple criteria of severity that can become widely known by general practitioners and doctors working in hospitals. This article presents preliminary data based on a retrospective analysis. A prospective study in District General Hospitals is planned as the next stage.

THE DECISION PROCESS
The first step in management of acute colitis is to allocate the patient to the disease class ‘acute non-specific colitis’ since the treatment of other acute inflammations of the colon is different. Once this decision has been taken the first treatment decision is between immediate surgical treatment and drug therapy. The indications for immediate surgical treatment include very severe ulceration of the colon with involvement of the muscle layers or serosa, perforation or severe haemorrhage. If these clinical features are absent, most patients are treated initially with drugs.

The outcome of surgical treatment is recovery or death. The outcome of medical treatment can be improvement, failure to improve or death. If the patient fails to improve, a further decision has to be taken between continued medical treatment with a change of drug therapy, or colectomy. This process of decision is repeated as often as necessary until the outcome of the illness is determined.
OBJECT OF THE PRESENT STUDY
The object of the present study is to recognise those clinical attributes at the end of the first 24 hours in hospital and a few days later which forecast most accurately the success of medical treatment, as judged by recovery, or its failure, as judged by all other outcomes. It is hoped that future studies will enable these attributes to be used as the basis for a decision procedure of the kind described by Emerson (1975) (see page 238).

METHOD OF INVESTIGATION
A retrospective analysis was made by Dr J. K. Ritchie of the clinical features and outcome of all 189 acute attacks of colitis treated in the wards of St Mark's Hospital during the years 1968–1972. Most patients with acute colitis are treated as out-patients and this series therefore includes only patients with disease unresponsive to treatment outside hospital, or patients with disease regarded as too severe for out-patient management.

Data concerning the clinical features of the illness were recorded in a form suitable for punch card analysis. In all, 56 attributes were analysed derived from the clinical history (6), physical examination (15), sigmoidoscopy (3), ward observations over four days (11), haematology (4), biochemistry (9) and radiology (8). A comparison was made between the clinical features of the 144 attacks of colitis in which medical treatment succeeded and the 37 attacks in which medical treatment failed. Eight patients were so seriously ill on admission as to need an immediate operation and were not included in the analysis.

The statistical analysis was carried out in two stages, the first to see whether a small number of clinical variables would suffice for prediction, and the second to estimate the quantitative effects of these variables on the risk of coming to operation.

There is no single approach to the problem of reducing the number of variables. Two main methods were used—

(a) a simple enumeration of all subsets of variables up to 5 variables and selection of those that best distinguished the patients who came to operation;

(b) stepwise linear discriminant analysis that selects in succession, the best, 1, 2, 3, . . . etc. variables for distinguishing the two groups.

Details of this technique are given by Dixon (1973) who describes the appropriate computer program.

The use of these two approaches has been described by Spicer et al. (1973). The linear discriminant does not directly find the set that gives the least
misclassification except when the variables have a certain specified (multivariate normal) distribution. As it has been successfully used in similar medical applications it seemed worth trying in the present case as a cross check on the more direct analysis. Both these two methods selected the same set of discriminating characters.

Once a small number of variables has been identified which carry most of the predictive information on the course of the disease, it is desirable to derive, if possible, a quantitative relationship between their values and the chances of a successful outcome. The simplest way of doing this is to draw up a table of all possible combinations of their respective values and take the observed proportion of patients who have had operations in each combination as the expected risk. This is not a very good method if the number of patients is small, since the risks in individual categories are often based on very small numbers and are consequently unreliable.

It is better, if possible, to express the risks mathematically in terms of a set of parameters, fewer than the total number of categories. If it is found that only one parameter is required for each value of the different clinical variables, the analysis is equivalent to the common assumption of independence. However, more complex degrees of dependence can be taken into account by adding further parameters and the statistical significance of such additions assessed. The statistical methods for doing this are described by Snedecor and Cochran (1967), Cox (1970) and Nelder and Wedderburn (1972). Nelder (1974) gives details of computer programs for performing the calculations. In statistical terminology, the analysis is equivalent to that of a multi-way contingency table.

RESULTS
The main results of this study are summarised in Table 1, which shows the proportion of patients expected to come to operation classified according to their maximum temperature and number of bowel actions in the first 24 hours after admission to hospital. Of the clinical data available at this stage these two variables alone are sufficient, and the addition of others does not improve the accuracy of predictions. The analysis also shows that there is nothing to be gained with the present data by increasing the precision of the temperature scale or the counted number of bowel movements beyond the broad groupings in Table 1. If much larger numbers of patients were available other variables might be revealed as having some predictive capacity but the increase in accuracy could only be marginal. Similarly, larger numbers would define more accurately the dependence of prognosis on temperature and bowel movements.
Table 1. Percentage of patients at 24 hours after admission who do not require subsequent operation, classified by frequency of bowel movements and maximum temperature during the preceding 24 hours

| Bowel actions per 24 hours | Temperature |
|---------------------------|-------------|
|                           | 97°-100° | > 100° |
| 0-5                       | 96       | 62     |
| 6-8                       | 91       | 40     |
| ≥9                        | 66       | 20     |

Table 2. Percentage of patients not requiring subsequent operation, classified by frequency of bowel movement on first day, serum albumin and pulse rate on fourth day after admission

| Bowel Actions per 24 hours | Serum Albumin g/100 ml | Pulse rate ≤ 89 | > 90 |
|---------------------------|------------------------|-----------------|------|
| 0-5                       | <3.0                   | 89              | 53   |
|                           | 3.0-3.5                | 95              | 87   |
|                           | >3.5                   | 99              | 94   |
| 6-8                       | <3.0                   | 84              | 42   |
|                           | 3.0-3.5                | 92              | 73   |
|                           | >3.5                   | 97              | 78   |
| ≥9                        | <3.0                   | 68              | 38   |
|                           | 3.0-3.5                | 86              | 56   |
|                           | >3.5                   | 93              | 78   |

At four days after admission the temperature is still an important predictor but can be replaced by the pulse rate, with which it is closely correlated. In addition to this, the level of serum albumin, whose values are known on the fourth day, is also important, the prognosis being worse as the albumin levels decline.

It is rather confusing to express this in tabular form, as there are now three variables, but Table 2 shows that all three appear to affect the frequency of operation to some extent. A further statistical analysis confirms that there are
indeed highly significant effects of each and there is some evidence that a high pulse rate and low serum albumin reinforce one another’s unfavourable effects.

**DISCUSSION**

The results given in Table 1 provide useful and easily memorised working rules for assessing prognosis at an early stage. The results in Table 2 are more complicated, but demonstrate the value of the serum albumin level. It is planned to carry out a prospective survey which can test the reliability of these findings and collect the additional information required to incorporate them in a decision procedure.

We also feel that the methods of analysis used were of considerable value in reducing the number of tests to a manageable level without a significant loss of information. The more elaborate statistical methods of testing the assumption of independence between the variables and in assessing the statistical uncertainties of the data seem to be worth applying in other similar studies.

**Acknowledgements**

We wish to acknowledge the help of Dr J. K. Ritchie, who abstracted the case reports, and Mr Warren Hilder, who collaborated in the computing and statistical analysis.

**References**

Cox, D. R. (1970) *The Analysis of Binary Data*. London: Methuen.

Dixon, W. J. (Ed.) (1973) *Biomedical Computer Programs*. Los Angeles: University of California Press.

Emerson, P. A. (1975) *Journal of the Royal College of Physicians of London*, 9, 238.

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

Nelder, J. A. and Wedderburn, R. W. M. (1972) *Journal of the Royal Statistical Society, A*, 135, 370.

Ritchie, J. K. (1974) *British Medical Journal*, 1, 264.

Spicer, C. C., Jones, H. and Lennard-Jones, J. E. (1973) *Methods of Information in Medicine*, 12, 118.

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