MEDICAL PRACTICE

*Occasional Reviews***Diagnosis of inflammatory bowel disease: an international multicentre scoring system**

SUSAN E CLAMP, JOHANNES MYREN, IAN A D BOUCHIER, GEOFFREY WATKINSON, FRANCIS T DE DOMBAL

Abstract

Inflammatory bowel disease is being studied by the World Organisation of Gastroenterology (OMGE) Research Committee in a survey (currently) affecting 1696 patients presenting to 30 hospitals in 16 countries. A computer-aided diagnostic prediction system (when tested in 1056 patients) showed an accuracy of match between computer prediction and clinical diagnosis in 94% of these patients. Most centres, however, do not possess appropriate computing facilities. A simple diagnostic "scoring system," based on likelihood ratios, has therefore been developed, using a series of numbers set out on a single sheet of paper. The overall accuracy of the simple scoring system in the same 1056 patients was

equivalent (93%) to the computer-aided system. In a further series of 510 "new" patients from eight fresh centres the diagnostic scoring system made a firm prediction in 490 cases (96%). Where made, the accuracy of match between the scoring system prediction and clinical diagnosis was 96%. It is suggested that this simple system, which requires no analytical resource and could be used by any gastroenterologist, might be of some value in clinical research and clinical practice.

Introduction

Since the pioneer studies that identified and categorised colonic Crohn's disease,¹ the problems in discriminating between patients with ulcerative colitis and those with colonic Crohn's disease have given rise to considerable controversy. Goodman and Sparberg² observed that this differentiation "is one of the commonest diagnostic problems in clinical gastroenterology"; and Johnson and Roth³ neatly summarised the problem by pointing out that the same patient may fit one category better by clinical features, the other by radiographic criteria, and pathologically show the features of both. As a result, not surprisingly the proportion of patients with "unclassifiable" inflammatory bowel disease has recently been placed as high as 10-20%.⁴

This difficult problem was recently studied in a multinational survey conducted under the sponsorship of the research committee of the Organisation Mondiale de Gastroentérologie (OMGE).⁵ In the course of this multinational study from 16 countries case data of 1056 patients with inflammatory bowel disease were subjected to computer analysis, and a computer "diagnostic prediction" was subsequently compared with the diagnosis established in the particular centre.

A surprisingly high accuracy of match between clinical diagnosis and computer prediction was noted (86% in the 1056

This work was performed under the auspices of the Research Committee of the World Organisation of Gastroenterology (OMGE) (chairman, Professor I A D Bouchier; secretary, Dr F T de Dombal; former OMGE president, Dr G Watkinson; principal study co-ordinator, Professor J Myren). Data analysis was carried out by the clinical information science group, St James's Hospital, Leeds LS9 7TF.

Department of Surgery, St James's University Hospital, Leeds LS9 7TF

SUSAN E CLAMP, World Organisation of Gastroenterology research assistant

FRANCIS T DE DOMBAL, MD, reader in clinical information science

Ullevål Sykehus, Oslo Kommune, Oslo, Norway

JOHANNES MYREN, MD, professor of medicine

Department of Medicine, University of Dundee and Ninewells Hospital, Dundee

IAN A D BOUCHIER, MD, professor of medicine

Gartnavel General Hospital, Glasgow G12 0YN

GEOFFREY WATKINSON, MD, consultant physician

patients when computer analysis was restricted to clinical data and 93% when computer analysis was undertaken on clinical and investigative data). As emphasised in the OMGE study,⁵ current definitions of ulcerative colitis and Crohn's disease are quite arbitrary, and the "eventual diagnoses" represent merely the views of a multinational team of concerned, participating, experienced clinicians. Nevertheless, in the absence of a definitive test for ulcerative colitis or Crohn's disease, the predictive system would enable *any* clinician to decide (in well over 90% of cases) how an individual case would be diagnosed in the centres concerned in the OMGE survey.

One major barrier to any wider use of the predictive system, however, concerns the computer. Not all centres have access to computer facilities, and many clinicians are understandably reluctant to subject patient data to computer analysis where this may have a bearing on subsequent management. For these reasons, a simplified "scoring system" has been derived from the data studied by the OMGE Research Committee, which may be operated by a clinician using a single sheet of paper, without recourse to the computer, in one or two minutes. Such a system, and its application to a total of 1566 patients with inflammatory bowel disease, forms the basis of this report.

Methods

PATIENTS AND CONDUCT OF STUDY

The 1056 patients (471 with (predominantly colonic) Crohn's disease and 585 with ulcerative colitis), whose data formed the basis of the simplified scoring system, presented to a total of 30 hospitals in 16 countries during 1976-8 with "active" inflammatory bowel disease. Wherever possible data collection proceeded on a prospective, unselected basis. All data were collected using a common protocol,⁵ and were forwarded to a central data processing team in Leeds—to be subsequently analysed with the aid of a desk-top computer (WANG 2200T) using programs written by the OMGE Research Committee survey team. Data collection in each of the 30 hospitals was comprehensive and thorough: the number of potential datapoints was 67 584 (64 per patient) and the data actually recorded totalled 62 699 items (93% of the possible tally).

Once a scoring system had been devised from these data its accuracy was then tested against a further series of 510 patients presenting to a further eight centres (during 1978-80) in Europe and around the Mediterranean. Conditions for acceptance into the study and data collection procedures were identical in both periods.

SCORING SYSTEM

The scoring system was derived from a series of "likelihood ratios" in the manner shown in the table. The scoring system thus represents a series of ratios between the conditional probabilities for each attribute state in ulcerative colitis and Crohn's disease. As regards age, for example, for some categories, such as patients aged between 20 and 49, there is little difference between the patients with ulcerative colitis and those with Crohn's disease. Such items therefore score 0 on the simplified scoring system. For patients in the under-20 age group the ratio is 1.5:1 in favour of Crohn's disease. This is rounded

to a value of "+1." In the group aged 50-59 a similar ratio (1.5:1 this time in favour of ulcerative colitis) is rounded to "-1." In the over-70s (a rare category) the ratio of 2:1 in favour of ulcerative colitis patients is rounded to "-2."

As regards a second category, rectal bleeding, the table shows that the proportion of patients with Crohn's disease without bleeding (72%) is almost exactly six times the proportion of patients with ulcerative colitis (12%). This is rounded to "+6" on the scoring

Attribute	Score	Attribute	Score
Age 27	0	Radiology	
Duration 2 years	0	Site: R colon	+1
Family history nil	0	Site: L colon	-1
Past history nil	-1	Stenosis	+4
Pain right low quad	+10	Endoscopy	
Intermittent pain	0	Patchy changes	+16
Bowels $\times$ 3-4/day	0	Biopsy	
No blood per rectum	+6	Granuloma	+27
No mucus per rectum	+3	Haemoglobin 14.7	0
No complications	0	White cell count 11.7	0
Emaciated	+2	Albumin 3.8	0
No tenderness	-1		
No distension	0	Total (investigative)	+47
Mass	+10	Prediction (investigative) Crohn's	
Total (clinical)	+29		
Prediction (clinical) Crohn's		Total (overall)	+76
		Prediction (overall) Crohn's disease	

FIG 1—Case history showing use of scoring system set out in appendix. Typical case of Crohn's disease diagnosed on both clinical and investigative data.

FIG 2—Retrospective series of 1056 patients. Scattergraph comparing total scores (using clinical data alone) for each of 585 patients with ulcerative colitis and 471 patients with Crohn's disease.

Example of derivation of scoring system from "likelihood ratios"

	Crohn's disease (n = 471)	Ulcerative colitis (n = 585)	Ratio	Score
Age				
<19	12.3*	7.9*	1.557	+1
20-29	36.7	27.0	1.359	0
30-39	21.8	27.2	1.248	0
40-49	12.5	14.8	1.184	0
50-59	8.1	12.2	1.506	-1
60-69	6.2	7.0	1.129	0
>70	2.4	3.7	1.542	-1
Blood in stools				
None	71.9	11.9	6.042	+6
Slight	18.8	40.6	2.159	-2
Severe	9.3	47.4	5.097	-5

*Figures are percentages of patients in each category.

system. Slight rectal bleeding gives a ratio of just over 2:1 in favour of ulcerative colitis (rounded to "-2"), and severe rectal bleeding (47% as against 9%) is treated similarly and rounded to "-5" by the system. The entire system is set out in detail as an appendix to this paper.

Its operation is simple (fig 1). The clinician specifies each patient's symptoms, signs, and investigative data, and adds or subtracts the whole numbers as indicated. The overall patient score for each new patient is the sum total (positive or negative) of the individual scores. Obviously a high negative score favours ulcerative colitis and a high positive score Crohn's disease.

Results

RETROSPECTIVE CASES

The scoring system for inflammatory bowel disease was first applied to the "original" group of 1056 cases derived from 30 hospitals in 16 countries, these being the cases from which the scoring system was originally derived. The results are shown in figs 2 and 3. Figure 2 shows that, by using clinical data alone, good differentiation was achieved between the two groups of patients with ulcerative colitis and Crohn's disease. Taking $\geq +5$ as an indicator of Crohn's disease and ≤ -5 as an indication of ulcerative colitis, 396 patients with Crohn's disease (84%) and 509 patients with ulcerative colitis (87%) were correctly assigned to their respective categories by the scoring system. The overall accuracy of match between scoring system and clinical diagnosis was therefore 86% (905 out of 1056 cases).

Figure 3 shows improved differentiation when investigative data are added to the scoring system. If a score of $\geq +10$ is taken to indicate Crohn's disease (less being taken to indicate ulcerative colitis) 438 patients with Crohn's disease (92.9%) and 545 patients with ulcerative colitis (93.2%) are correctly assigned—as judged by the eventual clinical diagnosis. The overall accuracy of match between system prediction and eventual clinical diagnosis is 93% (983 of 1056 patients).

FIG 3—Same series as fig 2. Total scores of system using scores for clinical and investigational data for each patient.

These figures in respect of the simplified system's predictive accuracy are of interest because they parallel closely the accuracy of match between the computer-aided prediction and eventual clinical diagnosis in the original OMGE survey.

PROSPECTIVE CASES

Nevertheless, a predictive system that is tested merely using case data used for its construction is incompletely tested. The simplified scoring system was applied, therefore, to a further series of 510 cases collected from eight different centres during a different period (1978-80). Of these patients, 120 were eventually categorised as having Crohn's disease and 390 ulcerative colitis (figs 4 and 5).

These further cases illustrate comparable differentiation to that in the retrospective series between the two groups in 510 new cases. By using clinical data alone, a firm prediction may be made by the simplified system when the overall score exceeds $+10$ (Crohn's disease) or falls below 0 (ulcerative colitis). Where made (in 85% of cases), such a prediction is 95% accurate.

When investigative data are added (and a score of $\geq +20$ taken to indicate Crohn's disease, < 10 to indicate ulcerative colitis), an even

better differentiation is obtained. In 4% of cases (20 out of 510) no firm prediction is made by the system. In the remaining cases (the overwhelming majority, 96%) a firm prediction is made, and is accurate in 479 out of 490 cases, 96% of the total. This figure equals the accuracy of match between computer-aided prediction and eventual clinical diagnosis in the 510 cases.

FIG 4—Further series of 510 cases from eight different centres. Overall scores for each patient, considering clinical data alone.

FIG 5—Same analysis as fig 4. Overall scores, however, result from considering clinical and investigational data.

Discussion

The simple scoring system described in the preceding paragraphs is neither comprehensive nor novel. The system is limited in that it merely discriminates between ulcerative colitis and Crohn's disease—rather than "diagnosing" either of these complaints. There is, for example, nothing in the present system that would distinguish Crohn's disease from intestinal tuberculosis; and it may well be that this system would be totally

unsuitable for use in primary health care or in countries where tuberculosis of the intestine is a common problem.

Moreover, systems that discriminate between ulcerative colitis and Crohn's disease are scarcely new. Perhaps the best system propounded since Lockhart *et al* first described colonic Crohn's disease¹ is that set out by Lennard-Jones.^{6,7} Indeed, a glance at the appendix below shows that the present system suggests (in many instances) similar categories to those described by Lennard-Jones.

Nevertheless, the present system does have some novel features. Firstly, it is a "numerate" system with "weightings" attached to each category. Hence it may be of value in difficult cases where the various modalities, such as x-ray or endoscopy, give conflicting indications, allowing a clinician to select *on balance* the most likely disease from the overall score.

Secondly, and perhaps more importantly, the present system *has been thoroughly validated* by using around 1600 cases from 30 hospitals in 16 countries. This extensive validation has shown a 96% match (in new cases) between the scoring system's firm prediction and the clinical diagnosis established some time later by the clinicians concerned. Moreover, the stability of this predictive accuracy in varying centres implies that these results should be capable of being reproduced by most clinicians, even those with only moderate experience of inflammatory bowel disease.

Thirdly, it might be argued that the scoring system outlined might be of some practical value both in research and clinical management. For example, as regards research studies, inflammatory bowel disease has undoubtedly been bedevilled over the years by descriptions of patients from different centres using slightly different diagnostic criteria. Centres reporting future series of cases might find the scoring system outlined a useful (and geographically stable) "double check," to ensure that patients in their series, using their own personal diagnostic criteria, conform broadly to those seen in other centres around the world.

Fourthly, as regards individual cases and their clinical management, the scoring system might be of potential value by enabling an individual clinician to compare a particular case with a rather large pool of cases from around the world and thereby arrive at a "working" diagnosis of either ulcerative colitis or Crohn's disease.

In making such an assertion it is, however, necessary to discuss two potential drawbacks of this (or any other) system for diagnosing inflammatory bowel disease. Firstly, even in this day and age, the diagnostic categorisation of inflammatory bowel disease is still somewhat arbitrary; and what this scoring system (or any other) uses as a "gold standard" is therefore not an absolute but an arbitrary diagnosis.

Nevertheless, in the case of the present survey each individual diagnosis was established by an experienced clinician (for a list of participating centres see below). Since there was little geographical variation, this scoring system might therefore be said—at worst—to represent in formal terms the consensus view of 100 experienced clinicians in 16 different countries.

Secondly, it may be argued that in the fullness of time the diagnoses established as a basis for the present series will change significantly. Follow-up data (gathered over a two- or four-year period) on the original series of 1056 patients indicate, however, that up to the present the number of reclassified patients has been encouragingly small (well under 5% of the original series). Thus, while the scoring system outlined here may change slightly in time (as additional cases are collected and analysed, and as further follow-up in the medium to long term takes place), present data indicate that any such changes will be relatively small.

For these reasons, despite its undoubted imperfections, the World Organisation of Gastroenterology Research Committee recommends (pro tem) the use of this scoring system for the two conditions discussed.

We are particularly grateful to the many clinicians in various centres

who contributed data without which the present survey would have been quite impossible.* Mrs D Blackband provided valuable secretarial help. The work was aided in part by research grants from the Medical Research Council of the UK and Pharmacia Ltd, which are acknowledged with gratitude.

Appendix

SCORING SYSTEM USED IN PRESENT ANALYSES

Age:		Upper 1/2	-2
< 19	+1	Left 1/2	-3
50-59	-1	Central	+6
> 70	-2	Nil	-1
Duration:		Abdominal findings:	
1-3 months	-2	Distension	+2
3-6 months	-1	Mass	+10
Family history:		Radiology:	
Ulcerative colitis	-2	Normal	-3
Crohn's disease	+4	Continuous	-1
Patient's history:		Segmental	+11
Appendicitis	+3	Site:	
Anal fissure	+7	Jejunum	+7
Fistula	+4	Ileum	+31
Nil	-1	R colon	+1
Site of pain:		L colon	-1
R lower quad	+10	Rectum	-3
L lower quad	-1	Findings:	
Right 1/2	+2	Stenosis	+4
Left 1/2	-6	Ulcers	-1
Central	+2	Dilatation	+4
Nil	-1	Fistula	+6
Type of pain:		Skip lesions	+8
Severe	+2	Endoscopy:	
Steady	+2	Normal	+12
Bowels:		Ulcers	-1
Normal	+1	Stenosis	+2
× 1/day	+3	Bleeding	-4
× > 10/day	-2	Diffuse	-2
Blood:		Patchy	+16
Nil	+6	Biopsy:	
Slight	-2	Normal	+5
++	-5	Ulcers	-3
Mucus:		Giant cell	+20
Nil	+3	Granuloma	+27
Slight	-1	Mucosal	-1
++	-2	Transmural	+16
Complications		Laboratory tests:	
Perianal	+7	Hb < 10 g/dl	-1
Fistula	+8	WCC > 20 000	+1
Systemic	+1	Alb > 5 g/dl	-1
Nutrition:		Plat < 150 000	-6
Emaciated	+2	> 400 000	+1
Tenderness:		Iron 20-40	+1
R lower quad	+10		

*A full list of clinicians is set out in *Scandinavian Journal of Gastroenterology* 1979;special supplement 56:2-3. Participating centres include: *Austria*: Universitätsklinik für Gastroenterologie und Hepatologie, Wien; *Belgium*: Ziekhuis St Elisabeth, Antwerp; *Canada*: St Michael's Hospital, Toronto; *Civic Hospital*, Ottawa; *Hospital St Justine*, Hospital St Luc, and *Hôtel Dieu*, Montreal; *Centre Hospitalier Universitaire*, Sherbrooke; *Clinical Research Centre*, Dalhousie University, Halifax; *Czechoslovakia*: Hospital Bulovka, Prague; *E Germany*: Wilhelm-Pieck-Universität, Rostock; *W Germany*: Krankenhaus Bad Caanstatt, Stuttgart; *Medizin Klinik mit Poliklinik*, der Universität Erlangen-Nürnberg, Erlangen; *Hungary*: Leading County Hospital, Miskolc, Borsod; *Israel*: Chaim Sheba Medical Centre, Tel-Hashomer; *Assaf Harofeh Government Hospital*, Serfin; *Italy*: Università di Bologna; *Istituto di Clinica Medica I*, Milan; *Norway*: Ulleval Hospital, Rikshospitalet, Aker Hospital, and Kroghstøtten Hospital, Oslo; *Haukeland Hospital*, Bergen; *County Hospital*, Fredrikstad; *County Hospital*, Kristiansand; *Regional Hospital*, Trondheim; *County Hospital*, Akershus; *Poland*: Polish Association of Gastroenterology, Katowice; *Portugal*: Faculdade de Medicina de Lisboa, Lisbon; *Scotland*: Victoria Infirmary, Western Infirmary, and Stobhill General Hospital, Glasgow; *Sicily*: Istituto di Patologia Medica 3°, Palermo; *Spain*: Hospital Santa Cruz y san Pablo, Barcelona; *United States*: Mayo Clinic, Rochester.

References

- ¹ Lockhart-Mummery HE, Morson BC. Crohn's disease (regional enteritis) of the large intestine and its distinction from ulcerative colitis. *Gut* 1960;1:87-105.
- ² Goodman MJ, Sparberg M. *Ulcerative colitis*. New York: John Wiley, 1978.
- ³ Johnson WD, Roth JLA. Diagnosis and differential diagnosis of chronic ulcerative colitis and Crohn's colitis. In: Kirnser JB, Shorter RG, eds. *Inflammatory bowel disease*. Philadelphia: Lea and Febiger, 1975:201-25.
- ⁴ Morson BC. Histopathology. In: Engel A, Larssen T, eds. *Regional enteritis*. (Fifth Skandia International Symposium.) Stockholm: Nordiska Bokhandels Forlag, 1971:15-33.
- ⁵ Myren J, Bouchier IAD, Watkinson G, de Dombal FT. Inflammatory bowel disease—An OMGE Survey. *Scand J Gastroent* 1979;14(Suppl 56):2-27.
- ⁶ Lennard-Jones JE. Crohn's disease: natural history and treatment. *Postgrad Med J* 1968;44:674.
- ⁷ Lennard-Jones JE. Definition and diagnosis. In: Engel A, Larssen T, eds. *Regional enteritis*. (Fifth Skandia International Symposium.) Stockholm: Nordiska Bokhandels Forlag, 1971:105 et seq.

(Accepted 2 October 1981)

Contemporary Themes

Care of the aged in France

N K CONI

The French and British populations contain similar proportions of aged citizens,¹ but largely because of language difficulties² most British doctors remain ignorant of how the French health and social services are evolving to meet the challenge of the rapid growth of these proportions. I visited establishments in and around four centres of geriatric excellence in southern France and talked to general practitioners, doctors in both general and geriatric medicine, psychiatrists, administrators, nurses, social workers, and remedial therapists in order to share experience and seek fresh ideas. The centres I chose were Toulouse, Limoges, Lyons, and Montpellier, but I must emphasise that this sample may be highly unrepresentative because there seems to be even more regional variation in the provision of services in France than in Britain. I make no attempt to comment on clinical practices, though considerable differences exist.³ Many facets of the British system seem superior to those found in France, but my observations are largely confined to those areas where French experience seems to have certain lessons to offer. Some of these lessons may seem rather political but, as Virchow said, "Medicine is a social science and politics nothing else but medicine on a large scale," and it is unlikely that the change of government in France will rapidly invalidate these observations.

The qualitative factor in scale of provision

Though nowhere near the top of the international league, the French as a people are probably overmedicalised and the British are certainly undermedicalised. In 1976-7 France was credited with 105.5 hospital beds per 10 000 population, while England had 83.1 (Scotland is far better provided for).⁴ The disparity is still greater in terms of expenditure per head on health care and days per head spent in hospital.⁵ The stock of residential accommodation also appears to be more generous in France, where there are over 360 000 beds, which represents 47

per thousand persons over 65, whereas in Britain most localities fall well below the DHSS norm of 25 per thousand.⁶

This difference in quantity makes a tremendous impact on quality. To a Frenchman, waiting a year or more for elective hip replacement surgery is unthinkable. To a French general practitioner in some (but not all) areas the admission of a demented elderly patient to a psychiatric hospital needs one quick telephone call with an immediately favourable response. The frail, isolated, incompetent old lady who is falling about and going off her feet is accepted unquestioningly by the acute medical service. The painful exercise of rationing health care, so familiar to the British general practitioner and geriatrician, scarcely arises in France. The onerous responsibility of providing a service for a given catchment population, as well as to the individual patient, is totally foreign to French geriatric practice. There is an imbalance of beds favouring the acute rather than the extended care sector,⁷ which may result in the elderly patient being taken through a tragic "circuit touristique" of transfers from one institution to another, sometimes finishing up far from home; but these are also disgraceful features of the British scene. At least the two-year wait in totally unsuitable housing with totally inadequate support is generally avoided. The inevitable blocking of acute beds may be as high as 15%⁸ but seems to be tolerated because of the adequate supply. There is little evidence of waste—for instance, a general medical service in the university hospital of Limoges, dealing with an elderly population (21% over 65), maintains an average duration of stay of 8.5 days. There is a policy to reduce the number of hospital beds in France, and no new beds, whether private or public, may be opened without closing old ones on at least a knock-for-knock basis. It may also be noted that the French economy has permitted the construction of general hospitals of 1200 beds and more, which seem to function perfectly satisfactorily.

Financing health care

Health care is, in general, financed by a system of reimbursement by the Sécurité Sociale, a scheme to which substantial contributions are made by both employer and employee. This reimbursement is generally of the order of 70% and is often augmented to 90% or so by occupation-linked insurance

Newmarket General Hospital, Newmarket, Suffolk CB8 7JG

N K CONI, MRCP, FRCP(C), consultant physician in geriatric medicine, Cambridge Health District