J. Hyg., Camb. (1981), **87**, 25
Printed in Great Britain

Antibodies to *Eubacterium* and *Peptostreptococcus* species and the estimated probability of Crohn's disease

BY J. P. VAN DE MERWE
Department of Medical Microbiology, Erasmus University, Rotterdam, The Netherlands

P. I. M. SCHMITZ
Institute of Biostatistics, Erasmus University, Rotterdam, The Netherlands

AND F. WENSINCK
Department of Medical Microbiology, Erasmus University, Rotterdam, The Netherlands

(*Received* 2 *January* 1980)

SUMMARY

Anaerobic coccoid rods belonging to species of *Eubacterium* and *Peptostreptococcus* agglutinate more frequently with sera from patients with Crohn's disease than with sera from patients suffering from other diseases and from healthy subjects. Results of agglutination tests with four strains of coccoid anaerobes were used to estimate the probability that a patient suffers from Crohn's disease. The data on healthy subjects and patients with Crohn's disease were subjected to logistic discriminant analysis. With the methods and interpretation described, 52 % of the patients with Crohn's disease were recognized as 'definite' or 'probable' Crohn's disease and 14 % as 'suspected'. Only 1 % of the healthy subjects were classified as 'suspected' and none as 'definite' or 'probable' Crohn's disease.

INTRODUCTION

Recently, antibodies to anaerobic gram-positive coccoid rods belonging to species of *Eubacterium* and *Peptostreptococcus* were found in sera of a considerable percentage of patients with Crohn's disease (Wensinck, 1975, 1976; Wensinck & Van De Merwe, 1981). In healthy subjects and patients with other diseases than Crohn's disease (CD) these antibodies were found less frequently.

Agglutination of four strains of coccoid anaerobes was used to establish the diagnosis of CD. Tests with these strains showed different sensitivities and specificities and a simple interpretation of the agglutination results was not possible. The results with sera from standard groups of patients with CD and of healthy subjects were therefore subjected to discriminant analysis. The methods used and the interpretation of the results in terms of probability that a patient suffers from CD are reported in this paper.

MATERIALS AND METHODS

Agglutination reactions

The techniques used to demonstrate agglutinins to *Eubacterium contortum* (strains Me_{44} and Me_{47}), *E. rectale* (strain Me_{46}) and *Peptostreptococcus productus* (strain C_{18}) were described by Wensinck & Van De Merwe (1981). The results of agglutination reactions were scored as negative (0) or positive (1, 2 or 3, according to strength).

Patients and control subjects

Between 1 October 1975 and 1 February 1978, all consecutive patients with CD in the Departments of Internal Medicine and Surgery were studied. The data obtained at the first presentation of the patient were used. The diagnosis of CD* was based on generally accepted clinical, radiological and histological criteria (Lennard-Jones, Lockhart-Mummery & Morson, 1968; Kirsner, 1975). Patients in whom the differential diagnosis between CD and ulcerative colitis could not be made in the period mentioned were not included. A group of 114 patients with CD was thus collected, 43 men with a median age of 33 years (range 14–68) and 71 women with a median age of 29 years (range 17–74). Forty-three had ileal disease, 25 ileocolonic and 46 colonic disease. Fifty-three patients had undergone intestinal resections. Nineteen patients were taking salicylazosulphapyridine, ten were on corticosteroids and eight on corticosteroid enemas.

The group of healthy subjects consisted of 95 volunteers of the Red Cross Blood Transfusion Service, Rotterdam, 53 male with a median age of 38 years (range 23–64) and 42 female with a median age of 35 years (range 28–62). The agglutination reactions of patients and controls are given in Table 1.

Statistical methods

The interpretation of any test result rests upon three factors: the initial likelihood that a disease is present on the basis of the clinical evidence at hand (*a priori* probability) and the likelihood that a given test result occurs in the population with the disease and in that without the disease, respectively. Combining these likelihoods, usually calculated by applying Bayes' theorem, results in the *a posteriori* probability of the disease. For our results logistic discriminant analysis (Anderson, 1972) was used for discrimination between patients with CD and healthy subjects (HS) on the basis of the agglutination reactions. The *a posteriori* probability of CD can be written as:

$$\Pi(\mathrm{CD}|x_1, x_2, x_3, x_4) = \frac{1}{1+\exp(\beta_0+\beta_1 x_1+\beta_2 x_2+\beta_3 x_3+\beta_4 x_4)}. \quad (1)$$

In this formula, x_1–x_4 are the results of the agglutination reactions with strains Me_{44}, C_{18}, Me_{46} and Me_{47}, respectively. The expression is known as the multivariate logistic function. The coefficients β_0–β_4 were estimated with the maximum

* Diagnoses were established by M. Van Blankenstein and J. Dees, Department of Internal Medicine II (Professor Dr M. Frenkel), University Hospital Dijkzigt, Erasmus University, Rotterdam.

Table 1. *Frequency distribution of agglutination reactions in* 114 *patients with Crohn's disease (CD) and* 95 *healthy subjects (HS)*

Agglutination*	CD	HS	Agglutination*	CD	HS
0 0 0 0	21	75	2 0 0 0	1	5
0 0 0 1	0	1	2 0 0 1	1	0
0 0 0 2	2	1	2 0 0 2	2	0
0 0 0 3	1	0	2 0 1 0	1	0
0 0 1 0	1	0	2 0 3 0	1	0
0 0 2 0	1	0	2 1 3 0	1	0
0 0 3 0	3	0	2 1 3 2	1	0
0 0 3 1	1	0	2 2 0 0	3	1
0 0 3 2	1	0	2 2 3 0	2	0
0 1 0 0	0	2	2 3 0 1	1	0
0 1 0 1	0	1	2 3 3 2	1	0
0 1 0 2	1	0	3 0 0 0	3	2
0 1 1 0	1	0	3 0 0 1	1	0
0 1 3 2	1	0	3 0 0 2	1	0
0 2 0 0	2	0	3 0 1 0	2	1
0 2 3 0	1	0	3 0 3 0	2	0
0 3 0 0	1	1	3 1 0 2	2	0
0 3 0 1	2	0	3 1 2 0	1	0
0 3 0 2	1	0	3 1 2 3	1	0
0 3 1 0	1	0	3 1 3 0	2	0
0 3 3 0	1	0	3 1 3 3	2	0
1 0 0 0	3	4	3 2 0 2	1	0
1 0 0 2	0	1	3 2 0 3	1	0
1 0 2 0	1	0	3 2 3 0	1	0
1 2 3 2	1	0	3 3 0 0	3	0
1 3 0 0	2	0	3 3 0 1	1	0
1 3 0 2	1	0	3 3 0 3	1	0
1 3 1 0	1	0	3 3 2 3	1	0
1 3 2 2	1	0	3 3 3 0	9	0
1 3 3 0	1	0	3 3 3 1	1	0
1 3 3 1	1	0	3 3 3 2	4	0
1 3 3 2	1	0	3 3 3 3	3	0

* Agglutination reactions with strains Me_{44}, C_{18}, Me_{46} and Me_{47}, respectively, according to strength.

likelihood method (Cox, 1970) and the estimates are denoted by b_0–b_4, respectively. The estimated *a posteriori* probability of CD can thus be written as:

$$P(\text{CD}|x_1, x_2, x_3, x_4) = \frac{1}{1+\exp(b_0+b_1x_1+b_2x_2+b_3x_3+b_4x_4)}. \quad (2)$$

The estimates b_0–b_4 are obtained from the reference samples (n_1 patients with CD and n_2 healthy subjects) with *a priori* probabilities $P(\text{CD})$ and $P(\text{HS})$ proportional to the sample sizes n_1 and n_2:

$$P(\text{CD}) = \frac{n_1}{n_1+n_2}; \quad P(\text{HS}) = \frac{n_2}{n_1+n_2}. \quad (3)$$

If *a posteriori* probabilities have to be calculated from other *a priori* probabilities $Q(\mathrm{CD})$ and $Q(\mathrm{HS})$, b_0 is replaced by d_0:

$$d_0 = b_0 - \ln \frac{n_2}{n_1} + \ln \frac{Q(\mathrm{HS})}{Q(\mathrm{CD})}. \tag{4}$$

Allocation rules and interpretation

After calculation of the estimated *a posteriori* probability of CD, several allocation rules can be used. It could be stated, for example, that a subject with agglutination reactions $x\ (= x_1, x_2, x_3, x_4)$ is classified as:

$$\left.\begin{aligned} &\text{non-CD if } 0 \leqq P(\mathrm{CD}|x) < 0{\cdot}5, \\ &\text{CD if } 0{\cdot}5 \leqq P(\mathrm{CD}|x) \leqq 1. \end{aligned}\right\} \tag{5}$$

When we classify as 'CD' when the *a posteriori* probability exceeds a certain value, say 0·8, and classify as 'non-CD' when the probability is below another value, say 0·2, we have an allocation rule with a region of doubt. Thus:

$$\left.\begin{aligned} &\text{non-CD if } 0 \leqq P(\mathrm{CD}|x) < 0{\cdot}2, \\ &\text{no decision (doubt) if } 0{\cdot}2 \leqq P(\mathrm{CD}|x) < 0{\cdot}8, \\ &\text{CD if } 0{\cdot}8 \leqq P(\mathrm{CD}|x) \leqq 1. \end{aligned}\right\} \tag{6}$$

On the basis of the results of the analysis (see Results), we divided the possible results of *a posteriori* probabilities in four classes with the following interpretation:

$$\left.\begin{aligned} &\text{no support for CD if } 0 \leqq P(\mathrm{CD}|x) < 0{\cdot}8, \\ &\text{suspected CD if } 0{\cdot}8 \leqq P(\mathrm{CD}|x) < 0{\cdot}95, \\ &\text{probable CD if } 0{\cdot}95 \leqq P(\mathrm{CD}|x) < 0{\cdot}99, \\ &\text{definite CD if } 0{\cdot}99 \leqq P(\mathrm{CD}|x) \leqq 1. \end{aligned}\right\} \tag{7}$$

RESULTS

Estimation of the coefficients and application of the allocation rules

With the results of the agglutination reactions x_1 (Me_{44}), x_2 (C_{18}), x_3 (Me_{46}) and x_4 (Me_{47}) of 114 patients with CD and 95 healthy subjects, the estimates b_0–b_4 of the coefficients β_0–β_4 in (1) were:

$$b_0 = 1{\cdot}33. \tag{8}$$

$$\left.\begin{aligned} b_1 &= -0{\cdot}45, \\ b_2 &= -0{\cdot}95, \\ b_3 &= -2{\cdot}32, \\ b_4 &= -1{\cdot}04. \end{aligned}\right\} \tag{9}$$

The coefficient b_0 is connected with *a priori* probabilities:

$$P(\mathrm{CD}) = 114/209 = 0{\cdot}55; \quad P(\mathrm{HS}) = 0{\cdot}45.$$

For equal *a priori* probabilities $Q(\text{CD}) = Q(\text{HS}) = 0{\cdot}5$, coefficient b_0 is replaced by d_0 according to (4), thus:

$$d_0 = 1{\cdot}51. \tag{10}$$

The other coefficients remain unchanged. The 95 % confidence intervals for these coefficients are:

$$\left.\begin{aligned} -0{\cdot}86 &< \beta_1 < -0{\cdot}05 \\ -1{\cdot}51 &< \beta_2 < -0{\cdot}38 \\ -4{\cdot}23 &< \beta_3 < -0{\cdot}41 \\ -1{\cdot}77 &< \beta_4 < -0{\cdot}31 \end{aligned}\right\}. \tag{11}$$

From (11) it is concluded that coefficients β_1–β_4 differ significantly from 0 ($\alpha = 0{\cdot}05$). With the estimated coefficients (9) and (10) *a posteriori* probabilities were calculated for all possible agglutination reactions (Table 2). The results for the elements from the two samples of patients with CD and healthy subjects are given in Table 3. When allocation rule (5) is used, classifications are obtained as summarized in Table 4. If allocation (6) is used, a classification is obtained as given in Table 5. From this table it is evident that about 20 % of the elements in both samples are not classified.

With the use of allocation rule (7) *a posteriori* probabilities are obviously interpreted more realistically (Table 6). In this situation, 42 % of the patients with CD are classified as 'definite CD', 10 % as 'probable CD', 14 % as 'suspected CD' and 34 % as 'no support for CD'. Of the healthy subjects, 99 % are classified in this latter category, whereas only 1 % is classified as 'suspected CD' and none as 'probable' or 'definite CD'.

Verification

The results of allocation were presented for the elements, which were also used for estimation of the coefficients of the logistic function. It would be better to evaluate the classification in new samples from the patients with CD and healthy subjects. As these samples were not available, a split-sample method was used. The sample consisting of 114 patients with CD was randomly divided into two groups of 57 subjects. The sample consisting of 95 healthy subjects was divided into a group of 48 and one of 47 subjects. With the agglutination reactions of the first subsample of 57 patients with CD and 48 healthy subjects, new coefficients were estimated. The results, for equal *a priori* probabilities were:

$$\left.\begin{aligned} d_0 &= 1{\cdot}75, \\ b_1 &= -0{\cdot}56, \\ b_2 &= -0{\cdot}94, \\ b_3 &= -2{\cdot}01, \\ b_4 &= -1{\cdot}66. \end{aligned}\right\} \tag{12}$$

The *a posteriori* probabilities for the elements of both groups, as determined with these coefficients, are given in Table 7. From this table it is seen that the results

Table 2. A. posteriori *probabilities P(CD|x) of CD at* a priori *probability of* 0·5 *for all possible agglutination results with strains* Me_{44}, C_{18}, Me_{46} *and* Me_{47}

Agglutination*	P(CD|x)	Agglutination	P(CD|x)	Agglutination	P(CD|x)	Agglutination	P(CD|x)
0 0 0 0	0·18	1 0 0 0	0·26	2 0 0 0	0·35	3 0 0 0	0·46
0 0 0 1	0·38	1 0 0 1	0·50	2 0 0 1	0·61	3 0 0 1	0·71
0 0 0 2	0·64	1 0 0 2	0·73	2 0 0 2	0·81	3 0 0 2	0·87
0 0 0 3	0·83	1 0 0 3	0·89	2 0 0 3	0·92	3 0 0 3	0·95
0 0 1 0	0·69	1 0 1 0	0·78	2 0 1 0	0·85	3 0 1 0	0·90
0 0 1 1	0·86	1 0 1 1	0·91	2 0 1 1	0·94	3 0 1 1	0·96
0 0 1 2	0·95	1 0 1 2	0·97	2 0 1 2	0·98	3 0 1 2	0·99
0 0 1 3	0·98	1 0 1 3	0·99	2 0 1 3	0·99	3 0 1 3	0·99
0 0 2 0	0·96	1 0 2 0	0·97	2 0 2 0	0·98	3 0 2 0	0·99
0 0 2 1	0·98	1 0 2 1	0·99	2 0 2 1	0·99	3 0 2 1	1
0 0 2 2	0·99	1 0 2 2	1	2 0 2 2	1	3 0 2 2	1
0 0 2 3	1	1 0 2 3	1	2 0 2 3	1	3 0 2 3	1
0 0 3 0	1	1 0 3 0	1	2 0 3 0	1	3 0 3 0	1
0 0 3 1	1	1 0 3 1	1	2 0 3 1	1	3 0 3 1	1
0 0 3 2	1	1 0 3 2	1	2 0 3 2	1	3 0 3 2	1
0 0 3 3	1	1 0 3 3	1	2 0 3 3	1	3 0 3 3	1
0 1 0 0	0·36	1 1 0 0	0·47	2 1 0 0	0·58	3 1 0 0	0·69
0 1 0 1	0·62	1 1 0 1	0·72	2 1 0 1	0·80	3 1 0 1	0·86
0 1 0 2	0·82	1 1 0 2	0·88	2 1 0 2	0·92	3 1 0 2	0·95
0 1 0 3	0·93	1 1 0 3	0·95	2 1 0 3	0·97	3 1 0 3	0·98
0 1 1 0	0·85	1 1 1 0	0·90	2 1 1 0	0·93	3 1 1 0	0·96
0 1 1 1	0·94	1 1 1 1	0·96	2 1 1 1	0·98	3 1 1 1	0·98
0 1 1 2	0·98	1 1 1 2	0·99	2 1 1 2	0·99	3 1 1 2	0·99
0 1 1 3	0·99	1 1 1 3	1	2 1 1 3	1	3 1 1 3	1
0 1 2 0	0·98	1 1 2 0	0·99	2 1 2 0	0·99	3 1 2 0	1
0 1 2 1	0·99	1 1 2 1	1	2 1 2 1	1	3 1 2 1	1
0 1 2 2	1	1 1 2 2	1	2 1 2 2	1	3 1 2 2	1
0 1 2 3	1	1 1 2 3	1	2 1 2 3	1	3 1 2 3	1
0 1 3 0	1	1 1 3 0	1	2 1 3 0	1	3 1 3 0	1
0 1 3 1	1	1 1 3 1	1	2 1 3 1	1	3 1 3 1	1
0 1 3 2	1	1 1 3 2	1	2 1 3 2	1	3 1 3 2	1
0 1 3 3	1	1 1 3 3	1	2 1 3 3	1	3 1 3 3	1
0 2 0 0	0·60	1 2 0 0	0·70	2 2 0 0	0·78	3 2 0 0	0·85
0 2 0 1	0·81	1 2 0 1	0·87	2 2 0 1	0·91	3 2 0 1	0·94
0 2 0 2	0·92	1 2 0 2	0·95	2 2 0 2	0·97	3 2 0 2	0·98
0 2 0 3	0·97	1 2 0 3	0·98	2 2 0 3	0·99	3 2 0 3	0·99
0 2 1 0	0·94	1 2 1 0	0·96	2 2 1 0	0·97	3 2 1 0	0·98
0 2 1 1	0·98	1 2 1 1	0·99	2 2 1 1	0·99	3 2 1 1	0·99
0 2 1 2	0·99	1 2 1 2	0·99	2 2 1 2	1	3 2 1 2	1
0 2 1 3	1	1 2 1 3	1	2 2 1 3	1	3 2 1 3	1
0 2 2 0	0·99	1 2 2 0	1	2 2 2 0	1	3 2 2 0	1
0 2 2 1	1	1 2 2 1	1	2 2 2 1	1	3 2 2 1	1
0 2 2 2	1	1 2 2 2	1	2 2 2 2	1	3 2 2 2	1
0 2 2 3	1	1 2 2 3	1	2 2 2 3	1	3 2 2 3	1
0 2 3 0	1	1 2 3 0	1	2 2 3 0	1	3 2 3 0	1
0 2 3 1	1	1 2 3 1	1	2 2 3 1	1	3 2 3 1	1
0 2 3 2	1	1 2 3 2	1	2 2 3 2	1	3 2 3 2	1
0 2 3 3	1	1 2 3 3	1	2 2 3 3	1	3 2 3 3	1
0 3 0 0	0·79	1 3 0 0	0·86	2 3 0 0	0·90	3 3 0 0	0·94
0 3 0 1	0·92	1 3 0 1	0·94	2 3 0 1	0·96	3 3 0 1	0·98
0 3 0 2	0·97	1 3 0 2	0·98	2 3 0 2	0·99	3 3 0 2	0·99
0 3 0 3	0·99	1 3 0 3	0·99	2 3 0 3	1	3 3 0 3	1
0 3 1 0	0·97	1 3 1 0	0·98	2 3 1 0	0·99	3 3 1 0	0·99

Table 2 (*cont.*)

0 3 1 1	0·99	1 3 1 1	0·99	2 3 1 1	1	3 3 1 1	1
0 3 1 2	1	1 3 1 2	1	2 3 1 2	1	3 3 1 2	1
0 3 1 3	1	1 3 1 3	1	2 3 1 3	1	3 3 1 3	1
0 3 2 0	1	1 3 2 0	1	2 3 2 0	1	3 3 2 0	1
0 3 2 1	1	1 3 2 1	1	2 3 2 1	1	3 3 2 1	1
0 3 2 2	1	1 3 2 2	1	2 3 2 2	1	3 3 2 2	1
0 3 2 3	1	1 3 2 3	1	2 3 2 3	1	3 3 2 3	1
0 3 3 0	1	1 3 3 0	1	2 3 3 0	1	3 3 3 0	1
0 3 3 1	1	1 3 3 1	1	2 3 3 1	1	3 3 3 1	1
0 3 3 2	1	1 3 3 2	1	2 3 3 2	1	3 3 3 2	1
0 3 3 3	1	1 3 3 3	1	2 3 3 3	1	3 3 3 3	1

* Agglutination reactions with strains Me_{44}, C_{18}, Me_{46} and Me_{47} respectively.

Table 3. *Frequency distribution of* a posteriori *probabilities* $P(CD|x)$ *of CD at* a priori *probability of* 0·5 *in patients with CD and healthy subjects*

$P(CD\|x)$	Crohn's disease	Healthy subjects
0–< 0·10	0	0
0·10–< 0·20	21	75
0·20–< 0·30	3	4
0·30–< 0·40	1	8
0·40–< 0·50	3	2
0·50–< 0·60	0	0
0·60–< 0·70	6	2
0·70–< 0·80	5	3
0·80–< 0·90	9	0
0·90–< 0·95	7	1
0·95–< 0·99	11	0
0·99–1	48	0
Total	114	95

Table 4. *Classification matrix for allocation rule* (5)

		Population of origin	
		CD	HS
Population of allocation	CD	86	6
	HS	28	89
	Total	114	95

Table 5. *Classification matrix for allocation rule* (6)

		Population of origin	
		CD	HS
Population of allocation	CD	75	1
	Doubt	18	19
	HS	21	75
	Total	114	95

Table 6. *Frequency distribution of* a posteriori *probabilities* $P(CD|x)$ *in four classes for allocation rule* (7) *in patients with CD and healthy subjects (HS)*

$P(CD\|x)$	Interpretation	Number (%) of CD	Number (%) of HS
0–< 0·80	No support for CD	39 (34)	94 (99)
0·80–< 0·95	Suspected CD	16 (14)	1 (1)
0·95–< 0·99	Probable CD	11 (10)	0 (0)
0·99–1	Definite CD	48 (42)	0 (0)
Total		114 (100)	95 (100)

Table 7. *Frequency distribution of* a posteriori *probabilities of CD,* $P(CD|x)$ *for the elements from the first sub-sample of* 57 *patients with CD and* 48 *HS and the second sub-sample of* 57 *patients with CD and* 47 *HS, with use of coefficients* (12)

	Sub-sample 1		Sub-sample 2	
$P(CD\|x)$	CD	HS	CD	HS
0–< 0·10	0	0	0	0
0·10–< 0·20	9	38	12	37
0·20–< 0·30	0	3	3	1
0·30–< 0·40	1	3	0	4
0·40–< 0·50	1	1	2	2
0·50–< 0·60	2	0	1	0
0·60–< 0·70	0	0	0	0
0·70–< 0·80	4	2	3	1
0·80–< 0·90	5	1	2	2
0·90–< 0·95	5	0	5	0
0·95–< 0·99	6	0	6	0
0·99–1	24	0	23	0
Total	57	48	57	47

of both groups are similar. It is concluded, therefore, that the coefficients (9) and (10) obtained with the original samples n_1 and n_2 may be used.

A posteriori *probability of CD at* a priori *probability* $\neq 0{\cdot}5$

The estimations of probabilities of CD and the corresponding interpretations in this paper are based on *a priori* probabilities of 0·5. For application at other *a priori* probabilities for all individuals, coefficient d_0 as well as the interpretations of the *a posteriori* probabilities have to be adjusted. For situations with considerable individual variation of *a priori* probabilities, the coefficient d_0 cannot be adjusted individually because a reliable interpretation of the *a posteriori* probabilities can be made only on the basis of an evaluation of the results. With the present material this is not possible.

DISCUSSION

With the methods described, 52% of patients with known CD could be recognized with the agglutination reactions as 'definite CD' or 'probable CD'. None of the healthy subjects was classified in these categories. Moreover, 14% of the patients with CD were classified as 'suspected CD' compared to only 1% of

healthy subjects. About one-third of the patients with CD were classified as 'no support for CD' compared to 99% of healthy subjects. From these results we conclude that application of the agglutination reactions in combination with the interpretation given yields an improvement in discrimination of individuals into groups of patients with CD and healthy subjects. If this system is used for diagnostic purposes, it should discriminate between 'CD' and 'non-CD'. The extrapolation from the group 'healthy subjects' to 'non-CD' is only allowed when there are no relevant differences in agglutination reactions between these groups. It has been shown previously that this is true for a large number of 'control' diseases, but in cirrhosis of liver and coeliac disease the results require a different interpretation (Wensinck & Van De Merwe, 1981).

It should be noted that the test is evaluated with the samples which were also used for estimation of the coefficients of the discriminant function. This difficulty is partly overcome by verification via the split-sample method, but ideally the predictive value of a diagnostic test for clinical use should be evaluated via a comparison of the patient population with a similar population that only differs by the absence of the particular disease. The final classification of patients as 'CD' or 'non-CD', however, requires long-term studies as the average period of time between onset of symptoms of Crohn's disease and diagnosis is 4 years (Dyer & Dawson, 1970; Brandes & Eulenburg, 1976; Mekhjian *et al.* 1979).

REFERENCES

ANDERSON, J. A. (1972). Separate logistic discrimination. *Biometrika* **59**, 19.

BRANDES, J.-W. & EULENBURG, F. (1976). Der lange Weg zur Diagnose Morbus Crohn. *Zeitschrift für Gastroenterologie* **14**, 400.

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

DYER, N. H. & DAWSON, A. M. (1970). Diagnosis of Crohn's disease. A continuing source of error. *British Medical Journal* **i**, 735.

KIRSNER, J. B. (1975). Problems in the differentiation of ulcerative colitis and Crohn's disease of the colon: the need for repeated diagnostic evaluation. *Gastroenterology* **68**, 187.

LENNARD-JONES, J. E., LOCKHART-MUMMERY, H. E. & MORSON, B. C. (1968). Clinical and pathological differentiation of Crohn's disease and proctocolitis. *Gastroenterology* **54**, 1162.

MEKHJIAN, H. S., SWITZ, D. M., MELNYK, C. S., RANKIN, G. B. & BROOKS, R. K. (1979). Clinical features and natural history of Crohn's disease. *Gastroenterology* **77**, 898.

WENSINCK, F. (1975). The faecal flora of patients with Crohn's disease. *Antonie van Leeuwenhoek* **41**, 214.

WENSINCK, F. (1976). Faecal flora of Crohn's patients. Serological differentiation between Crohn's disease and ulcerative colitis. In *The Management of Crohn's Disease*, pp. 103–5. Amsterdam: Excerpta Medica.

WENSINCK, F. & VAN DE MERWE, J. P. (1981). Serum agglutinins to *Eubacterium* and *Peptostreptococcus* species in Crohn's and other diseases. *Journal of Hygiene* **87**, 13.