Psychological Medicine, 2004, **34**, 1035–1045. © 2004 Cambridge University Press
DOI: 10.1017/S0033291703001508 Printed in the United Kingdom

Cluster analysis of key diagnostic variables from two independent samples of eating-disorder patients: evidence for a consistent pattern

DAVID CLINTON*, ERIC BUTTON, CLAES NORRING AND ROBERT PALMER

Division of Psychiatry, M57, Neurotec Department, Karolinska Institutet, Huddinge University Hospital, S-141 86, Sweden; University Department of Psychiatry, Brandon Unit, Leicester General Hospital, Leicester LE5 4PW, UK; Psychiatric Research Centre, Örebro County Council, PO Box 1613, S-701 16 Örebro, Sweden; University Department of Psychiatry, Brandon Unit, Leicester General Hospital, Leicester LE5 4PW, UK

ABSTRACT

Introduction. The optimal classification of eating disorders has been a matter of considerable debate. The present paper tackles this issue using cluster analysis with large independent samples of eating-disorder patients.

Method. Two samples of adult female patients from Sweden ($n=631$) and England ($n=472$) were classified on the basis of 10 key clinical variables of primary significance for diagnosing eating disorders. A separate series of cluster analyses were conducted on each sample.

Results. Results suggested that a three-cluster solution was optimal in both samples. The first cluster ('generalized eating disorder') was characterized by high levels of eating-disorder psychopathology on all variables except weight and menstrual functioning. The second cluster ('anorexics') was typified by low weight, amenorrhoea and the absence of binge eating, and seemed to correspond to the clinical picture of anorexia nervosa. The third cluster ('overeaters') was characterized by high weight and moderate levels of binge eating and compensatory behaviour.

Conclusions. Results suggest that patients presenting to eating-disorder services in different countries have clinical features that fall into very similar patterns. These patterns resemble, but are not identical to, existing diagnostic categories.

INTRODUCTION

Since the original key papers in the 1870s on what is now known as anorexia nervosa, a number of new terms have been introduced to describe a range of clinical conditions akin to anorexia nervosa. Russell's (1979) seminal paper on bulimia nervosa heralded a new wave of terminology to describe a related clinical problem, which shared some of the features of anorexia nervosa, but in which the dominant phenomenon centred around so-called 'binge eating'.

Nylander (1971) was one of the first authors to draw attention to the concept of a continuum of eating disorders in his research on dieting and feeling fat in a Scandinavian teenage school population. Button & Whitehouse (1981) extended such thinking in their study of British female college students. They identified a substantial number of young women who had many of the symptoms of anorexia nervosa without fulfilling strict criteria. They suggested that anorexia nervosa may be somewhat the tip of the iceberg and introduced the term 'subclinical anorexia nervosa' to describe this group.

* Address for correspondence: David Clinton, Division of Psychiatry, M57, Neurotec Department, Karolinska Institutet, Huddinge University Hospital, S-141 86, Sweden.
(Email: David.Clinton@neurotec.ki.se)

Holmgren and co-workers (1983) examined the entire eating-disorders spectrum, and introduced a continuum model of eating disorders (the ABC model), which included the so-called 'anorexic-like' patients. Szmukler (1983) also recognized this group, but preferred the term 'partial syndrome'. Such partialness of disorder also applies to bulimic disorders, as exemplified by the term 'binge-eating disorder' in which binge eating is present, but without the accompanying compensatory behaviours associated with bulimia nervosa. It is now widely recognized that there is a spectrum of problems aligned to the two major disorders. The current DSM system (APA, 1994) favours the term 'eating disorder not otherwise specified' often abbreviated as 'EDNOS'. This is a catch-all diagnosis which covers any eating disorder of clinical severity, the symptomatology of which does not meet criteria for anorexia nervosa or bulimia nervosa. The ICD-10 (WHO, 1992) contains the terms atypical anorexia nervosa and atypical bulimia nervosa to cover broadly the same range of disorders.

The diagnostic criteria for eating disorders used within the DSM and ICD systems tend to be based on clinical opinion and consensus. There are, however, variations and uncertainties even in relation to the definition of key features within these systems, such as bingeing and low weight. Overall, classification of the eating disorders is unsatisfactory, not least because in many clinical series a substantial proportion of patients have to be consigned to the residual category, EDNOS.

The question arises of whether prevailing diagnostic categories reflect the natural groupings of patients who seek help for eating problems. Empirical study of the patterns occurring in the real world may provide a useful commentary upon our familiar and well-used diagnostic concepts. Both factor analysis and cluster analysis have been used for many years by psychiatric researchers investigating problems of diagnosis (Everitt & Landau, 1998). Recent examples of factor-analytic approaches to diagnostic questions within the field of eating disorders are the studies by van der Ham *et al.* (1997) and Williamson *et al.* (2002). While factor analysis reveals patterns among variables, cluster analysis focuses on groupings of individuals. In the field of eating disorders, four relevant recent studies have used cluster analysis with eating-disorder patients. Hay *et al.* (1996) used cluster analysis to investigate the classification of bulimic disorders in a community sample of 250 young women. Stice & Agras (1999) subtyped 265 patients with bulimia nervosa into two subgroups using cluster analysis. Their findings were subsequently replicated by Grilo *et al.* (2001) in a sample of 48 patients with bulimia nervosa. Cluster analysis was also used by Mizes & Sloan (1998) to investigate subgroups within a sample of 53 EDNOS patients presenting to psychotherapy clinics for the treatment of an eating disorder.

These studies provide some useful empirical indicators about the classification of eating disorders, but they have important shortcomings. They focus on specific diagnoses (Hay *et al.* 1996), or subgroups of diagnoses (Mizes & Sloan, 1998; Stice & Agras, 1999; Grilo *et al.* 2001) rather than the entire spectrum of eating disorders. Samples have been small (van der Ham *et al.* 1997; Mizes & Sloan, 1998; Grilo *et al.* 2001) or based on non-clinical cases with doubtful relevance to clinical eating disorders (Bulik *et al.* 2000).

Another approach to studying questions related to classification is called taxometrics. However, where the previously mentioned methods are exploratory, trying to uncover naturally occurring groups of subjects or variables, taxometrics uses predefined groups (e.g. diagnoses) and tries to determine whether these groups are actually discrete classes or concepts with a more dimensional quality. There are three recent studies applying taxometric methods to the eating-disorder classification (Gleaves *et al.* 2000*a*, *b*; Williamson *et al.* 2002). The study by Williamson and co-workers is unique in that it first applies both exploratory and confirmatory factor analysis and then taxometric methods in an attempt to shed more light on the problem of eating-disorder classification. In a similar vein, Crow *et al.* (2002) used stepwise discriminant analysis to distinguish between groups of patients (total $n = 385$) with full and partial syndromes of anorexia nervosa, bulimia nervosa, and binge-eating disorder. They could differentiate the full syndromes from each other, but failed to discriminate the partial from the full syndromes for anorexia nervosa and binge-eating disorder. While these studies

give important indications about the underlying structure of the presently used eating-disorder classification, they are of less value in illuminating questions about the natural groupings of eating-disorder patients.

Our approach to the classification of eating disorders has been based on cluster analysis (Everitt *et al.* 2001). In the present study we attempted to explore natural groupings of patients who present to a wide range of eating-disorder services. We used data on key diagnostic variables routinely collected on series of patients presenting to a number of centres in Sweden and to one centre in England. There was sufficient overlap between the assessment methods and variables studied in both countries to enable us to make relevant comparisons. Primary aims of the study were to: (1) compare the cluster solutions based on 10 key diagnostic variables from two similar but unrelated datasets of eating-disorder patients; and (2) relate the most appropriate of these cluster solutions to the diagnostic eating-disorder classification of DSM-IV in order to explore the relevance of alternative diagnostic classifications.

METHOD

Participants

Two samples of female patients from Sweden ($n=631$) and England ($n=472$) were used in the study.

The Swedish sample was collected within the framework of the Co-ordinated Evaluation and Research at Specialist Units for Eating Disorders in Sweden (CO-RED) Project. This is a longitudinal naturalistic study of the treatment of eating disorders at 15 specialist centres. The units offer a wide variety of treatment forms such as in-patient, day patient, out-patient, individual psychotherapy, family and group therapy, psychoactive drugs, expressive forms of treatment, etc. At the conclusion of the 6-year data collection period in December 2002, 946 patients with eating disorders were included in the project. In order to conduct appropriate cluster analysis patients with missing data on one or more of the 10 variables used in the study were excluded; this left 631 patients. In the majority of excluded cases (71%) data were lacking on just one of the variables. The distribution of DSM-IV diagnoses was: anorexia nervosa, $n=137$ (22·8%); bulimia nervosa, $n=240$ (39·9%); binge-eating disorder, $n=31$ (5·2%); and eating disorder not otherwise specified (EDNOS), $n=193$ (32·1%). All subjects provided informed consent to take part in the CO-RED study. Age ranged from 14 to 49 years [mean (S.D.) = 24·5 (6·4) years], and all participants were female. Mean duration of eating disorder at presentation was 8·2 years (S.D. = 6·7 years).

The English sample came from Leicester, and was drawn from a series of patients referred to a specialized service for eating disorders within the British National Health Service. For some years the service has collected and electronically recorded standard clinical information on all patients assessed. This allowed for the identification of 890 potential participants. The data files and clinical records of these patients were used to extract data on the standard variables shared with the Swedish sample. Patients with missing data on one or more of the 10 standard variables, as well as those with a diagnosis other than an eating disorder were excluded in order to be consistent with the Swedish sample. This left 472 Leicester patients with complete data on all the relevant variables. Mean scores on clinical and demographic characteristics were very similar to another cohort of consecutive referrals to the same service previously published (Button *et al.* 1997). The sample was, therefore, reasonably representative of patients referred to the service. The distribution of DSM-IV diagnoses was: anorexia nervosa, $n=82$ (17·4%); bulimia nervosa, $n=163$ (34·5%); binge-eating disorder, $n=8$ (1·7%); and eating disorder not otherwise specified (EDNOS), $n=219$ (46·4%). Age ranged from 15 to 61 years [mean (S.D.) = 25·1 (7·8) years].

There were no systematic differences in terms of eating disorder, psychiatric or background variables between cases with and without complete data.

Instruments

Two separate semi-structured interviews with similar variables were used to assess relevant diagnostic variables. In the Swedish sample the Rating of Anorexia and Bulimia Interview (RAB) was used (Clinton & Norring, 1999; Nevonen *et al.* 2003). The RAB is a 56-item

semi-structured interview with graded response formats covering a wide range of eating-disorder symptoms, related psychopathology and background variables; it generates operational DSM-IV eating-disorder diagnoses, and is widely used in Sweden. It has satisfactory internal consistency and inter-rater reliability; kappa ranged from 0·47 to 0·92 (mean = 0·74) for the variables used in the present study. In the British sample the Clinical Eating Disorders Rating Instrument (CEDRI) was used (Palmer *et al.* 1987), a semi-structured interview similar to the RAB. It has demonstrated high inter-rater reliability, with kappa ranging from 0·73 to 1·0 (mean = 0·90) for the variables used in the present study. From these instruments, 10 essential clinical variables for the diagnosis of eating disorders according to DSM-IV were selected from both the RAB and CEDRI for subsequent cluster analysis. These variables were: BMI, fear of weight gain, restriction of food intake, avoidance of fattening foods, binge eating, self-induced vomiting, abuse of laxatives, compulsive exercise, amenorrhoea, and body-image disturbance. The phrasing of questions on the two interviews was similar and response formats were identical with the exception of questions pertaining to binge eating (RAB, 5-point graded scale; CEDRI, 4-point graded scale) and body-image disturbance (RAB, 3-point graded scale; CEDRI, 4-point graded scale).

Procedure

Data for both the Swedish and English samples were collected by staff from participating units. Interviewers had long experience in the assessment of eating disorders in a clinical setting using the respective instruments. For the most part interviewers were either qualified psychiatrists or clinical psychologists, although other professionals, such as experienced nurses and social workers, also took part. Training of interviewers took place at participating units and/or at centrally arranged project meetings and workshops (in the case of the Swedish data). Administration of measures took place at initial diagnostic assessment prior to treatment, or in the case of the Swedish series within 2–4 weeks of commencing treatment at the latest.

Data analysis

Prior to computation of cluster analysis all variables were standardized, and standard scores were used for subsequent cluster analyses. Cluster analysis was conducted in a series of three steps using SLEIPNER (Bergman & El-Khouri, 1998), a statistical package for person-based analysis, focusing on cluster analysis. Although other more common packages for cluster analysis are available, such as SPSS, SLEIPNER has the ability to identify outliers that may otherwise distort the normative pattern of results, and allows for a more flexible way of conducting non-hierarchical cluster analysis based on the initial results of hierarchical methods. Separate cluster analyses were conducted on each sample. We did not combine data from the two series. This allowed for a comparison of the pattern of results in one sample against the pattern of results in the other.

RESULTS

Step 1: Identification of outliers

In the first step residual analysis was conducted to identify outliers. This is an important step in cluster analysis, since it allows for the exclusion of statistically eccentric cases that may obscure the more normative patterns in the data. The procedure resulted in the identification of 30 outliers in each of the two samples. Outliers were a diverse group without common identifying features. Examples were: normal or overweight patients with amenorrhoea, bulimics with relatively normal BMI who reported low levels of compensatory behaviour, bulimics who reported high levels of compensatory behaviour but who were grossly overweight, low-weight anorexics of the bulimic subtype who reported little in the way of caloric restriction, EDNOS patients with a high degree of laxative abuse, and EDNOS patients with clear binge-eating symptoms and compensatory behaviour but little body image disturbance. These outliers were excluded from subsequent cluster analysis.

Step 2: Hierarchical cluster analysis

In the second step hierarchical cluster analysis was computed using Ward's method. The

Table 1. *Standard scores on essential clinical variables in relation to specific cluster solutions using Ward's method of hierarchical cluster analysis*

	n	BMI	Weight phobia	Binge eating	Restriction	Food avoidance	Vomiting	Laxative abuse	Compulsive exercise	Amenor-rhoea	Body image
Five-cluster solution											
Swedish sample											
Generalized eating disorder	211	0·02	0·19	0·33	0·44	0·54	0·38	−0·19	0·29	−0·51	0·25
Anorexics	178	−0·53	−0·13	−0·87	0·24	0·20	−0·76	−0·12	0·10	0·91	0·29
Overeaters	94	0·90	−0·10	−0·03	−0·95	−1·04	−0·51	−0·27	−0·55	−0·97	−0·69
Binge eaters	90	0·12	−0·27	0·88	−0·73	−0·83	1·01	−0·26	−0·37	0·40	−0·60
Laxative abusers	28	−0·16	0·61	0·36	0·68	0·82	0·47	3·93	0·21	0·05	0·55
English sample											
High-weight bingers	152	0·56	−0·43	0·33	−0·41	−0·62	0·16	−0·36	−0·63	−0·47	−0·31
Compulsive exercisers	93	−0·04	0·10	−0·58	0·14	0·41	−0·39	−0·36	0·56	−0·56	0·11
Anorexics	76	−0·94	−0·17	−0·92	0·10	0·09	−0·80	−0·40	−0·22	1·78	−0·43
Restricting bulimics	63	0·01	0·68	0·92	0·46	0·56	0·94	−0·20	0·53	0·07	0·49
Laxative abusers	58	−0·18	0·44	0·26	0·23	0·23	0·24	2·24	0·46	−0·28	0·67
Four-cluster solution											
Swedish sample											
Generalized eating disorder	211	0·02	0·19	0·33	0·44	0·54	0·38	−0·19	0·29	−0·51	0·25
Anorexics	178	−0·53	−0·13	−0·87	0·24	0·20	−0·76	−0·12	0·10	0·91	0·29
High-weight bingers	184	0·52	−0·18	0·41	−0·84	−0·94	0·23	−0·26	−0·46	−0·30	−0·64
Laxative abusers	28	−0·16	0·61	0·36	0·68	0·82	0·47	3·93	0·21	0·05	0·55
English sample											
Generalized eating disorder	156	−0·02	0·33	0·03	0·27	0·47	0·15	−0·29	0·55	−0·31	0·26
High-weight bingers	152	0·56	−0·43	0·33	−0·41	−0·62	0·16	−0·36	−0·63	−0·47	−0·31
Anorexics	76	−0·94	−0·17	−0·92	0·10	0·09	−0·80	−0·40	−0·22	1·78	−0·43
Laxative abusers	58	−0·18	0·44	0·26	0·23	0·23	0·24	2·24	0·46	−0·28	0·67
Three-cluster solution											
Swedish sample											
Generalized eating disorder	239	0·00	0·24	0·33	0·47	0·57	0·39	0·30	0·28	−0·44	0·28
High-weight bingers	184	0·52	−0·18	0·41	−0·84	−0·94	0·23	−0·26	−0·46	−0·30	−0·64
Anorexics	178	−0·53	−0·13	−0·87	0·24	0·20	−0·76	−0·12	0·10	0·91	0·29
English sample											
Generalized eating disorder	214	−0·06	0·36	0·09	0·26	0·41	0·17	0·39	0·52	−0·30	0·37
High-weight bingers	152	0·56	−0·43	0·33	−0·41	−0·62	0·16	−0·36	−0·63	−0·47	−0·31
Anorexics	76	−0·94	−0·17	−0·92	0·10	0·09	−0·80	−0·40	−0·22	1·78	−0·43
Two-cluster solution											
Swedish sample											
Restrainers	417	−0·23	0·08	−0·18	0·37	0·41	−0·10	0·12	0·20	0·13	0·28
High-weight bingers	184	0·52	−0·18	0·41	−0·84	−0·94	0·23	−0·26	−0·46	−0·30	−0·64
English sample											
Overeaters	366	0·19	0·03	0·19	−0·02	−0·02	0·17	0·08	0·05	−0·37	0·09
Anorexics	76	−0·94	−0·17	−0·92	0·10	0·09	−0·80	−0·40	−0·22	1·78	−0·43

resultant pattern of agglomeration was most interesting from five clusters down to two clusters. Results are presented in Table 1.

Looking first at the Swedish sample, the largest group in the five-cluster solution was what could be termed 'generalized eating disorder' ($n=211$). These patients presented with relatively normal BMI and relatively high levels of restriction, food avoidance, binge eating, vomiting, compulsive exercise and body-image disturbance, along with low levels of laxative abuse and amenorrhoea. The second largest cluster was labelled 'anorexics' ($n=178$), and was distinguished by amenorrhoea and low BMI, along with higher than average restriction and food avoidance, plus low levels of bingeing and vomiting. The third largest cluster was termed 'overeaters' ($n=94$), and was characterized by high BMI, average bingeing, along with low levels of restriction, compensatory behaviour and amenorrhoea. The fourth largest cluster was labelled 'bulimics' ($n=90$). These patients presented with the highest levels of binge eating, coupled with relatively average BMI, low restriction, and high levels of vomiting. The smallest cluster was termed 'laxative abusers' ($n=28$), and was typified by relatively low BMI, high fear of fatness, high restriction,

high levels of body-image disturbance, and markedly high levels of laxative abuse.

A generally similar pattern was found in the English sample. The largest cluster comprised 'high-weight bingers' ($n=152$) and was characterized by high BMI and binge eating, as well as low levels of restriction and compensation. Unlike the Swedish sample, the second largest cluster comprised what could be termed 'compulsive exercisers' ($n=93$), and was typified by high levels of compulsive exercise and avoidance of fattening food, along with low levels of binge eating and compensation. The third largest cluster was termed 'anorexics' ($n=76$) and was characterized by high levels of amenorrhoea and low BMI, as well as low levels of binge eating and compensatory behaviour. The fourth largest cluster was labelled 'restricting bulimics' ($n=63$) and was characterized by high levels of binge eating and vomiting, along with high restriction and food avoidance. Like the Swedish sample, the smallest cluster comprised 'laxative abusers' ($n=58$), and was typified by markedly high levels of laxative abuse, along with high to moderate levels of pathology on most other variables.

When the two closest clusters were agglomerated at the four-cluster level in the Swedish sample, 'bulimics' merged with 'overeaters' to produce a cluster of 'high-weight bingers' ($n=184$). These patients were characterized by high BMI, high levels of binge eating, and low levels of restriction and compensatory behaviour with the exception of vomiting which was moderate. In the English sample at the four-cluster level, 'compulsive exercisers' combined with 'restricting bulimics' to produce a new cluster that could also be described as 'generalized eating disorder' ($n=156$). This cluster was characterized by compulsive exercise, avoidance of fattening foods and restriction, along with above average levels of vomiting. Although levels of binge eating were around the mean for this group, examination of raw values indicated that 60% of these patients were binge eating, and 39% were bingeing sufficiently frequently to meet DSM-IV criteria for bulimia nervosa.

When the next two closest clusters were agglomerated at the three-cluster level in the Swedish sample, 'laxative abusers' merged with cases characterized by 'generalized eating disorder' ($n=239$). Because the former was such a small cluster the general pattern of results changed little, with the exception that the 'generalized eating disorder' cluster as a whole now exhibited considerably higher levels of laxative abuse. In the English sample, the 'generalized eating disorder' cluster also merged with the small cluster of 'laxative abusers' to produce a cluster of patients that further accentuated the 'generalized eating disorder' cluster ($n=214$), which now exhibited markedly higher levels of laxative abuse. Although the levels of binge eating in this cluster were near the mean, examination of raw scores indicated clear problems with this symptom (63% of the group were bingeing, and 43% met binge-eating criteria for bulimia nervosa).

When only two clusters were left in the Swedish sample, cases with 'generalized eating disorder' merged with the 'anorexics' to produce a cluster of 'restrainers' ($n=417$), characterized by moderately low BMI, along with relatively high restriction and food avoidance, plus moderately high levels of compulsive exercise and laxative abuse. In the English sample at the two-cluster level, cases with 'generalized eating disorder' merged with 'high-weight bingers' to produce a cluster of 'overeaters' ($n=366$).

Determination of the optimal number of clusters was based on both the interpretability of the specific cluster solution (i.e. how meaningful it appeared) and by using the variance ratio criterion (VRC; Calinski & Harabasz, 1974). The VRC is a statistical aid for determining an optimal number of clusters, and is computed by calculating a ratio of the total between group sum of squares (BGSS) to the total within group sum of squares (WGSS) in relation to number of (k) clusters and sample size (n). The formula used was:

$$\mathrm{VRC}=(\mathrm{BGSS}/k-1)/(\mathrm{WGSS}/n-k).$$

According to this method a statistically optimal number of clusters is reached at the point where the graph peaks. In both samples the statistical optimum appeared to be around three or four clusters. In the English sample the VRC peaked at three clusters and then declined, whereas in the Swedish sample the VRC peak came at four clusters. Of these two solutions, the three-cluster solution was chosen for further analysis based on the VRC, and the

Table 2. *Standard scores on essential clinical variables in relation to three-cluster solution following non-hierarchical relocation analysis*

Cluster labels	*n*	BMI	Weight phobia	Binge eating	Restriction	Food avoidance	Vomiting	Laxative abuse	Compulsive exercise	Amenorrhoea	Body image
Swedish sample											
Generalized eating disorder	216	0·03	0·39	0·60	0·47	0·48	0·68	0·32	0·26	−0·24	0·64
Overeaters	193	0·65	−0·36	0·28	−1·03	−0·88	0·01	−0·25	−0·49	−0·33	−0·63
Anorexics	192	−0·68	−0·07	−0·96	0·51	0·35	−0·78	−0·11	0·20	0·61	0·29
English sample											
Overeaters	171	0·55	−0·49	0·14	−0·59	−0·64	0·06	−0·23	−0·41	−0·47	−0·42
Generalized eating disorder	170	−0·01	0·60	0·38	0·49	0·60	0·40	0·38	0·54	−0·37	0·66
Anorexics	101	−0·91	−0·19	−0·89	0·18	0·09	−0·78	−0·25	−0·21	1·41	−0·41

interpretability of the solution. Although the four-cluster solution was also potentially interesting, it was not chosen for further analysis since the cluster of laxative abusers was relatively small and appeared to constitute more of a distinct subgroup of cases within the 'generalized eating disorder' category.

Step 3: Non-hierarchical cluster analysis

In the third step non-hierarchical cluster analysis using the relocation method was utilized to arrive at an optimal classification. This final step initially proceeded from the previous three-cluster solution using Ward's method. Using an iterative algorithm each case was examined in relation to cluster centroids in order to arrive at the optimal allocation of cases for a three-cluster solution. Conceptually, this step is akin to rotation in factor analysis. When relocation analysis is used in cluster analysis it tends to yield more homogeneous and conceptually distinct clusters compared to hierarchical methods. Results of this procedure, which can be considered as the final cluster results, are presented in Table 2.

Results of the non-hierarchical cluster analysis did indeed yield more homogeneous and distinct clusters. The three clusters were of more equal size. The 'bulimic' aspect of the 'generalized eating disorder' cluster in both samples was more evident in higher levels of binge eating compared to the hierarchical results. In many respects this cluster was now more classically bulimic and showed high levels of restriction and avoidance of fattening foods. Overall, the pattern of results in both samples after non-hierarchical cluster analysis was markedly similar. Nevertheless, there were some small differences between the two samples. English 'anorexics' tended to exhibit more of the physical symptoms (i.e. a tendency towards lower BMI and greater degree of menstrual dysfunction), whereas as their Swedish counterparts tended to exhibit more behavioural symptoms (i.e. a tendency towards greater restriction, avoidance of fattening foods, compulsive exercise and disturbed body image). Swedish patients in the 'generalized eating disorder' cluster tended to be characterized by high levels of binge eating and vomiting, while their English counterparts had a tendency to be somewhat more anorexic (i.e. engage in compulsive exercise and avoidance of fattening foods, as well as express a higher degree of weight phobia).

Finally, we compared the results of the non-hierarchical cluster analysis at step 3, with the original clinical diagnosis according to DSM-IV. Table 3 shows this comparison for each of the two samples.

There was a high degree of correspondence between the two samples and also substantial concordance between the clusters and clinical diagnosis. For the most part, patients with anorexia nervosa were found in the 'anorexic' cluster, patients with bulimia nervosa were found in the 'generalized eating disorder' cluster, and binge-eating disorder patients were found amongst the 'overeaters'. However, the patients diagnosed as EDNOS did differ somewhat between the two series. In the Swedish sample EDNOS patients tended to fall into the 'overeater' and 'anorexic' clusters, whereas in the English sample EDNOS patients tended to be found in the 'overeater' and 'generalized eating disorder' clusters.

Table 3. *Comparison of clusters and DSM-IV diagnoses*

DSM-IV diagnosis	Cluster label: Generalized eating disorder	Overeaters	Anorexics	Total
Swedish sample				
Anorexia nervosa	29 (21·2%)	3 (2·2%)	105 (76·6%)	137 (100·0%)
Bulimia nervosa	155 (64·6%)	78 (32·5%)	7 (2·9%)	240 (100·0%)
Binge-eating disorder	0 (0·0%)	29 (93·5%)	2 (6·5%)	31 (100·0%)
EDNOS	32 (16·6%)	83 (43·0%)	78 (40·4%)	193 (100·0%)
Total	216 (35·9%)	193 (32·1%)	192 (31·9%)	601 (100·0%)
English sample				
Anorexia nervosa	17 (22·4%)	2 (2·6%)	57 (75·0%)	76 (100·0%)
Bulimia nervosa	85 (55·6%)	63 (41·2%)	5 (3·3%)	153 (100·0%)
Binge-eating disorder	0 (0·0%)	7 (100·0%)	0 (0·0%)	7 (100·0%)
EDNOS	68 (33·0%)	99 (48·1%)	39 (18·9%)	206 (100·0%)
Total	170 (38·5%)	171 (38·7%)	101 (22·9%)	442 (100·0%)

EDNOS, Eating disorder not otherwise specified.

DISCUSSION

We have sought to contribute evidence relevant to the classification of eating disorders by examining natural groupings of patients on key diagnostic variables in two independent samples using cluster analysis. Strengths of the study were that samples were drawn from series of patients newly presenting to specialist eating-disorder services in two countries – England and Sweden – and that they were analysed separately. Subjects were also assessed using interview schedules that were similar, yet not identical, which suggests that results are not due to some specific feature of the interview methodology, and hints at greater generalizability. However, since the samples were of patients presenting to secondary services, they may not be representative of all of those suffering from eating disorders in their respective communities. This is a weakness with regard to investigating the nature of eating disorders, but does, nonetheless, mean that the patterns found are likely to be of relevance to clinicians. Another weakness was the high number of patients excluded because of incomplete data, although examination of cases with and without complete data revealed no systematic bias in terms of eating disorder, psychiatric or background variables.

Results suggested that a three-cluster solution with clusters of relatively equal size provided the most parsimonious classification of cases. The utility of the three-cluster solution was supported both statistically (i.e. using the VRC) and heuristically (i.e. the most interpretable). It also produced clusters with markedly similar groupings of patients in both samples using hierarchical and non-hierarchical techniques. Both samples generated a cluster that broadly corresponded to a diagnosis of anorexia nervosa without bingeing or vomiting. Membership of this 'anorexic' cluster was associated with a clinical diagnosis of anorexia nervosa in 55% of cases in the Swedish sample and 56% of the English sample. Conversely, of patients who received a clinical diagnosis of anorexia nervosa, three quarters fell into the 'anorexic' cluster and a quarter into the 'generalized eating disorder' cluster in both series. However, 40% of patients assigned to the 'anorexic' clusters had received a clinical diagnosis of EDNOS.

The other two clusters were also similar between the two samples, and we felt justified in giving them the same labels namely, 'generalized eating disorder' and 'overeaters'. The distinction between these two clusters seemed to be mainly one of severity and weight. Patients assigned to the 'generalized eating disorder' cluster were rated as having more severe symptoms in almost every respect except weight. In broad terms, these clusters correspond to bulimia nervosa with restrictive tendencies and to a grouping of other states characterized by overeating and high weight. Of those subjects who had a diagnosis of bulimia nervosa, 61% fell into the 'generalized eating disorder' clusters and conversely 62% of subjects in those clusters received a diagnosis of bulimia nervosa.

Only relatively few of the subjects in either series had received a clinical diagnosis of 'binge-eating disorder', perhaps because of an inconsistency of practice as to whether or not to use this still provisional diagnostic category. Nevertheless, of the 38 subjects who did receive a clinical diagnosis of binge-eating disorder, 36 (94%) were located in the 'overeaters' clusters.

When interpreting these results it should be borne in mind that the labelling of clusters is as much an art as a science. The use of common labels in the final three-cluster solution does not imply identity between the data-sets, and alternative labels could be applied. Nevertheless, our common labels do emphasize the continuity and surprising similarity between two independent samples of eating-disorder patients.

Direct comparisons between our results and previous research are complicated by methodological factors. Previous studies using cluster analysis with eating disorders have used more restricted samples, focusing on the identification of subtypes, and are not strictly comparable (Hay *et al*. 1996; Mizes & Sloan, 1998; Stice & Agras, 1999; Grilo *et al*. 2001). There are, however, some important similarities between our results and those of others who have used other methodological approaches. Using latent class analysis in a community sample of non-clinical cases, Bulik and co-workers (2000) found three general classes resembling anorexia nervosa, bulimia nervosa and binge-eating disorder. A similar conclusion could be drawn from Crow and co-workers (2002), who showed that patients with partial anorexia nervosa and binge-eating disorder could not be readily differentiated from those with full syndromes. Williamson and co-workers (2002), who examined the underlying structure of symptoms rather than groupings of individuals, found three factors: binge eating, fear of fatness/compensatory behaviour, and drive for thinness. They also performed taxometric analyses and found support for conceptualizing the bulimic disorders (bulimia nervosa and binge-eating disorder) as discrete syndromes, whereas the evidence concerning anorexia nervosa was inconclusive. A similar conclusion was drawn by Gleaves *et al*. (2000*b*), who found that bulimic-type anorexia nervosa could be conceptualized on a continuum with bulimia nervosa but qualitatively distinguished from anorexia nervosa (restricting type). Nevertheless, the present results are somewhat at odds with the findings of van der Ham and co-workers (1997) who found that the occurrence of bulimic symptoms were of greater relevance for distinguishing patients than anorexic symptoms. However, van der Ham and associates examined a relatively small number of eating-disordered adolescents, the majority of which presented with anorexic forms of eating disorders, and focused on patterns of results over an extended period of time rather than initial clinical presentation, which might explain the apparent discrepancy.

Our analysis produced clusters that are recognizable from a clinical perspective. Although the results do not correspond precisely to the diagnoses generated by established diagnostic systems, they broadly support the distinction between anorexia nervosa (restricting type), and bulimia nervosa. In the present study, analyses of two large and distinct samples produced closely similar results. Patients presenting to differing eating-disorder services in different countries had clinical features that fell into very similar patterns. This cross-sample similarity suggests that the patterns of symptoms reflected in the clusters are likely to be found in other samples. However, this might not be the case if the samples were drawn directly from the community rather than from the clinic. Nevertheless, the replication of clusters within the present study does allow some confidence that they may be meaningful, and not just chance findings. They may provide a parsimonious description of the symptomatology of clinical eating disorders. However, description is different to diagnosis.

Clinical diagnosis is a tool for use in aiding prognosis and treatment choice. The best description of a complex set of features may well involve dimensions or factors. Nevertheless, diagnosis favours categories over dimensions so that clinicians may assign an individual unambiguously. Clusters too involve individual membership. But clusters are different to categories. Categories are defined by their boundaries – by the fulfilment or non-fulfilment of particular criteria in the case of formal diagnoses. In regard to some eating-disorder features, such as weight loss, a dimensional description is clearly optimal, but nevertheless

a particular degree of weight or weight loss may need to be chosen as the boundary. Ideally, there should be a point of rarity at the boundary of diagnostic categories if it is not to be arbitrary. However, this ideal is often not the case, and may well not be so in this instance. Studies such as the present one may suggest alternative categorizations. Yet it is probably expecting too much in such a complex and poorly defined field as eating disorders for true categories to be 'discovered', or to reflect profound truths about underlying mechanisms. Rather they will need to be determined – provisionally – and then tested as to their utility (cf. Kendell & Jablensky, 2003). On the whole, such decisions have tended to emerge from clinical experience, and have only sometimes been the result of systematic research. Nevertheless, empirical evidence should be able to help to inform the decision as to where appropriate boundaries should be drawn, and, perhaps as importantly, where they should not be drawn. Our present findings suggest that there may be a special problem in categorizing eating disorders other than anorexia nervosa. Our 'generalized eating disorder' and 'overeating' clusters seem to be distinguished mainly by severity and weight. Such a distinction is unlikely to support clear boundaries.

A system of classification should provide categories that are mutually exclusive and – ideally – they should also be collectively exhaustive. The present eating-disorder categories largely fail these tests (Palmer, 2003). The categories overlap in practice, or are prevented from doing so only by arbitrary rules. Certainly over time an individual may move from one diagnosis to another. Furthermore, the eating-disorder diagnoses together cover the field only through the use of the wide catch-all category of EDNOS. Our findings as well as those of others (Bulik *et al.* 2000; Crow *et al.* 2002) could be seen as indicating that the present diagnostic classes are too narrowly defined, and that the relocation of a proportion of the patients today diagnosed as EDNOS would make sense from a clinical point of view. New categories would not need to be perfect to be an improvement. Clusters such as those that we have outlined may suggest new categories that could be clinically useful, such as 'generalized eating disorder'. To be useful, however, diagnostic categories need to be recognizable and workable. Our use of only key clinical characteristics necessary for establishing diagnosis and detectable at first interview is helpful in this respect. Yet the clusters found in the present study are far from being candidate syndromes or prototype disorders. To achieve such status, they would need to be shown to have consistent associations with other features, to have prognostic value and to be related to treatment outcome. They would then need to be defined in ways that could be used in clinical practice. A proper conservatism prevails in the definition and adoption of new syndromes. However, there has been some slow extension of the diagnostic canon over the years, most notably with bulimia nervosa (Russell, 1979), and the launching of binge-eating disorder as a candidate disorder. Moreover, there has been a good deal of tinkering involving changes of detail in successive revisions of diagnostic criteria. Hitherto, both the definition of new syndromes and their subsequent revision has been the work either of perceptive individual clinicians or of committees. This is likely to continue to be the case, although their decisions and proposals could benefit from a greater use of empirical research into classificatory systems. Diagnostic categories facilitate research, but research findings should also inform the definition of diagnostic categories.

The clusters emerging from our work broadly support current diagnostic categories. The three clusters – 'anorexic', 'generalized eating disorder' and 'overeaters' – resemble but are not identical with the diagnoses of anorexia nervosa (restricting type), bulimia nervosa and binge-eating disorder. Nevertheless, it remains an open question whether the patterns of symptoms defined by cluster analysis might lead to categories that would perform better than our present conventional categories. To this end our on-going work is focusing on, amongst other things, the relationship of DSM-IV diagnoses and statistically derived clusters to personality variables, co-morbidity and outcome.

ACKNOWLEDGEMENTS

The authors acknowledge the assistance of Dr Lennart Bodin, Örebro University, for his valuable comments on the application of cluster analysis techniques, as well as the generous

financial support of The Vardal Foundation – for Health Care Sciences and Allergy Research that helped to make this research possible.

REFERENCES

APA (1994). *Diagnostic and Statistical Manual of Mental Disorders* (*DSM-IV*) (4th edn). American Psychiatric Association: Washington, DC.

Bergman, L. & El-Khouri, B. M. (1998). *SLEIPNER: A Statistical Package for Pattern-oriented Analysis*. Department of Psychology, Stockholm University.

Bulik, C. M., Sullivan, P. F. & Kendler, K. S. (2000). An empirical study of the classification of eating disorders. *American Journal of Psychiatry* **157**, 886–895.

Button, E. J., Marshall, P. M., Shinkwin, R., Black, S. & Palmer, R. L. (1997). One hundred referrals to an eating disorders service: progress and service consumption over a 2–4 year period. *European Eating Disorders Review* **5**, 47–63.

Button, E. J. & Whitehouse, A. (1981). Subclinical anorexia nervosa. *Psychological Medicine* **11**, 509–516.

Calinski, T. & Harabasz, J. (1974). A dendrite method for cluster analysis. *Communications in Statistics* **3**, 1–27.

Clinton, D. N. & Norring, C. (1999). The Rating of Anorexia and Bulimia (RAB) Interview: development and preliminary validation. *European Eating Disorders Review* **7**, 362–371.

Crow, S. J., Agras, W. S., Halmi, K., Mitchell, J. E. & Kraemer, H. C. (2002). Full syndromal versus subthreshold anorexia nervosa, bulimia nervosa, and binge eating disorder: a multicenter study. *International Journal of Eating Disorders* **32**, 309–318.

Everitt, B. S. & Landau, S. (1998). The use of multivariate statistical methods in psychiatry. *Statistical Methods in Medical Research* **7**, 253–277.

Everitt, B. S., Landau, S. & Leese, M. (2001). *Cluster Analysis*, *4th edn*. Arnold: London.

Gleaves, D. H., Lowe, M. R., Snow, A. C., Green, B. A. & Murphy-Eberenz, K. P. (2000*a*). Continuity and discontinuity models of bulimia nervosa: a taxometric investigation. *Journal of Abnormal Psychology* **10**, 56–68.

Gleaves, D. H., Lowe, M. R., Green, B. A., Cororve, M. B. & Williams, T. L. (2000*b*). Do anorexia and bulimia nervosa occur on a continuum? A taxometric analysis. *Behavior Therapy* **31**, 195–219.

Grilo, C. M., Masheb, R. M. & Berman, R. M. (2001). Subtyping women with bulimia nervosa along dietary and negative affect dimensions: a replication in a treatment-seeking sample. *Eating and Weight Disorders* **6**, 53–58.

Hay, P. J., Fairburn, C. G. & Doll, H. A. (1996). The classification of bulimic eating disorders: a community-based cluster analysis. *Psychological Medicine* **26**, 801–812.

Holmgren, S., Humble, K., Norring, C., Roos, B.-E., Rosmark, B. & Sohlberg, S. (1983). The anorectic bulimic conflict. An alternative diagnostic approach to anorexia nervosa and bulimia. *International Journal of Eating Disorders* **2**, 3–14.

Kendell, R. & Jablensky, A. (2003). Distinguishing between the validity and utility of psychiatric diagnoses. *American Journal of Psychiatry* **160**, 4–12.

Mizes, J. S. & Sloan, D. M. (1998). An empirical analysis of eating disorder, not otherwise specified: preliminary support for a distinct sub-group. *International Journal of Eating Disorders* **23**, 233–242.

Nevonen, L., Broberg, A., Clinton, D. & Norring, C. (2003). A measure for the assessment of eating disorders: reliability and validity studies of the Rating of Anorexia and Bulimia interview – revised version (RAB-R). *Scandinavian Journal of Psychology* **44**, 303–310.

Nylander, I. (1971). The feeling of being fat and dieting in a school population. *Acta Sociomedica Scandinavica* **3**, 17–26.

Palmer, R. L. (2003). Concepts of Eating Disorders. In *Handbook of Eating Disorders* (2nd edn) (ed. U. Schmidt, J. Treasure and E. van Furth). John Wiley & Sons: Chichester and New York.

Palmer, R. L., Christie, M., Cordle, C., Davies, D. & Kenrick, J. (1987). The Clinical Eating Disorder Rating Instrument (CEDRI): a preliminary description. *International Journal of Eating Disorders* **6**, 9–16.

Russell, G. F. M. (1979). Bulimia nervosa: an ominous variant of anorexia nervosa? *Psychological Medicine* **9**, 429–448.

Stice, E. & Agras, W. S. (1999). Subtyping bulimic women along dietary restraint and negative affect dimensions. *Journal of Consulting and Clinical Psychology* **67**, 460–469.

Szmukler, G. I. (1983). Weight and food preoccupation in a population of English schoolgirls. In *Understanding Anorexia Nervosa and Bulimia: Fourth Ross Conference on Medical Research* (ed. J. G. Bergman), pp. 21–28. Ross Laboratories: Columbus, Ohio.

van der Ham, T., Meulman, J. J., van Strien, D. C., van Engeland, H. (1997). Empirically based subgrouping of eating disorders in adolescents: a longitudinal perspective. *British Journal of Psychiatry* **170**, 363–368.

WHO (1992). *International Classification of Diseases* (10th edn). World Health Organization: Geneva.

Williamson, D. A., Womble, L. G., Smeets, M., Netemeyer, R. G., Thaw, J. M., Kutlesic, V. & Gleaves, D. H. (2002). Latent structure of eating disorder symptoms: a factor analytic and taxometric investigation. *American Journal of Psychiatry* **159**, 412–418.