Hay, P. J. ; Fairburn, C. G.; Doll, H. A.
The classification of bulimic eating disorders: a community-based cluster analysis study, *Psychological Medicine*, 1996; 26(4):801-812.

Psychological Medicine, 1996, **26**, 801–812.

The classification of bulimic eating disorders: a community-based cluster analysis study

P. J. HAY,[1] C. G. FAIRBURN AND H. A. DOLL

From the Department of Psychiatry, University of Oxford

SYNOPSIS There is controversy over how best to classify eating disorders in which there is recurrent binge eating. Many patients with recurrent binge eating do not meet diagnostic criteria for either of the two established eating disorders, anorexia nervosa or bulimia nervosa. The present study was designed to derive an empirically based, and clinically meaningful, diagnostic scheme by identifying subgroups from among those with recurrent binge eating, testing the validity of these subgroups and comparing their predictive validity with that of the DSM-IV scheme.

A general population sample of 250 young women with recurrent binge eating was recruited using a two-stage design. Four subgroups among the sample were identified using a Ward's cluster analysis. The first subgroup had either objective or subjective bulimic episodes and vomiting or laxative misuse; the second had objective bulimic episodes and low levels of vomiting or laxative misuse; the third had subjective bulimic episodes and low levels of vomiting or laxative misuse; and the fourth was heterogeneous in character. This cluster solution was robust to replication. It had good descriptive and predictive validity and partial construct validity.

The results support the concept of bulimia nervosa and its division into purging and non-purging subtypes. They also suggest a possible new binge eating syndrome. Binge eating disorder, listed as an example of Eating Disorder Not Otherwise Specified within DSM-IV, did not emerge from the cluster analysis.

INTRODUCTION

Until recently, research on eating disorders has focused on the two well recognized symdromes, anorexia nervosa and bulimia nervosa. Clinical experience, however, indicates that a substantial number of those who present for treatment of an eating disorder do not fulfil diagnostic criteria for either of these disorders (Mitchell *et al.* 1986; Bunnell *et al.* 1990; Hall & Hay, 1991). A particularly common group are those who have recurrent episodes of uncontrolled over-eating ('binges') yet do not have bulimia nervosa (Clinton & Glant, 1992; Fichter *et al.* 1993).

Two sources of evidence point to the importance of this group. First, it is now appreciated that a quarter or more of those who present for the treatment of obesity have recurrent bulimic episodes similar to those seen among patients with bulimia nervosa (Gormally *et al.* 1982; Hudson *et al.* 1988; de Zwaan *et al.* 1992; Wilson *et al.* 1993). These patients differ from those obese patients who do not binge: for example, they are more dissatisfied with their shape and weight, and they have higher levels of general psychiatric symptoms (Marcus *et al.* 1985, 1988, 1990*a*; Prather & Williamson, 1988; Wadden *et al.* 1993; Specker *et al.* 1994; Telch & Agras, 1994). In addition, it has been suggested that they may respond less well to treatment (Keefe *et al.* 1984; Marcus *et al.* 1990*b*) although the evidence for a differential treatment response is weak and some controlled studies have found no differences between those who do and do not binge (La Porte, 1992; Wadden *et al.* 1992; Yanovski, 1993).

The importance of this group is also highlighted by research on the prevalence of bulimia nervosa. Community studies have found that the prevalence of recurrent binge eating in the

[1] Address for correspondence: Dr Phillipa Hay, Department of Psychiatry, University of Adelaide, Royal Adelaide Hospital, North Terrace, Adelaide, SA 5001, Australia.

absence of bulimia nervosa may be as high as the prevalence of bulimia nervosa itself (Beglin, 1990; Bruce & Agras, 1992; Spitzer *et al.* 1992*a*).

On these grounds it was suggested by Spitzer and colleagues that a third eating disorder be included in DSM-IV (Spitzer *et al.* 1992*a*, 1993). This disorder, termed 'binge eating disorder', was designed to describe those individuals with recurrent bulimic episodes who do not meet diagnostic criteria for anorexia nervosa or bulimia nervosa. The suggestion attracted criticism on a number of grounds, the main one being that it was premature to demarcate this particular group from among all those who binge eat, given the absence of good descriptive and therapeutic studies on the entire population (Fairburn *et al.* 1993). It was eventually decided not to include binge eating disorder as a new diagnosis in DSM-IV. Instead, it was listed as an example of Eating Disorder Not Otherwise Specified with reference to patients who have recurrent episodes of binge eating (as defined in the DSM-IV) in the absence of the regular use of inappropriate compensatory behaviours characteristic of bulimia nervosa. Suggested specific criteria were included in Appendix B reserved for categories meriting further study (American Psychiatric Association, 1994, p. 550 and p. 729).

The aim of the present study was to derive an empirically-based scheme for classifying those with recurrent binge eating and to compare the predictive validity of this scheme with that of the DSM-IV scheme. Two aspects of the method are of particular note. First, the sample was community-based and designed to be composed of a broad range of those who binge eat, including those with bulimia nervosa; and secondly, the subjects were assessed using a standardized investigator-based interview.

METHOD

Design

A representative community-based sample of young women with recurrent binge eating was recruited. To identify subgroups among the sample it was subjected to a cluster analysis based on current eating disorder features. The descriptive and construct validity of the cluster solution was evaluated. The sample was followed up 1 year later to test the predictive validity of the cluster solution, and to compare its predictive validity with that of the DSM-IV scheme (American Psychiatric Association, 1994).

Subjects and definitions

The aim was to recruit a broad sample of those with recurrent binge eating. There has been considerable uncertainty over how best to define a binge (Fairburn & Wilson, 1993). There is consensus among technical and lay definitions that there should be a sense of loss of control (Beglin & Fairburn, 1992*a*), but there is less agreement over whether the amount eaten should also be large (Fairburn *et al.* 1986; Garner *et al.* 1991). For the purpose of this study, binge eating was defined as an episode of eating that was associated with a sense of loss of control at the time and that was viewed by the subject as large (i.e. the definition included both 'subjective' and 'objective' bulimic episodes – see Assessment below). This choice of definition allowed us to investigate subjects with bulimic episodes of varying size. Objective bulimic episodes accord with the definition of a binge in DSM-IV (American Psychiatric Association, 1994). 'Recurrent' was defined as having had at least 12 bulimic episodes (as defined above) over the previous 3 months. While the DSM-III-R and DSM-IV definitions of bulimia nervosa require binges to occur on average at least twice this rate, there are no data which support this specific threshold. Indeed, there is evidence that those who binge on average once a week do not differ from those who binge more often with respect to their specific eating disorder features, general psychopathology and outcome (Wilson & Eldredge, 1991). The study was confined to women aged between 16 and 35 years, since this is the group among whom eating disorders, such as bulimia nervosa and anorexia nervosa, appear to be most common. Subjects were excluded if they were pregnant, or had medical problems or treatments that were likely to affect their eating or weight.

Recruitment and follow-up procedure

The subjects were recruited from women aged 16 to 35 years whose names were listed on the case registers of 19 general practices located in urban and rural areas of Oxfordshire. This mode of recruitment is a good way of obtaining

a local general population sample since practically all of the British population is registered with a general practitioner (Office of Health Economics, 1987). These women were sent a self-report questionnaire (Fairburn & Beglin, 1995) (see below) designed to identify eating disorder features. Subjects who did not return the questionnaire were sent a second one. Respondents who reported recurrent episodes of over-eating were invited to be interviewed. Identifying the response rate in a study of this type is difficult since general practitioner registers are invariably overinclusive (i.e. they include the names of people who have subsequently moved away from the area). It is, therefore, not possible to conclude that non-response is indicative of the subject having received the questionnaire and chosen not to take part. For this reason a detailed pilot study was undertaken to evaluate this method of recruitment (Beglin, 1990). It involved attempting to trace all non-respondents to determine who had moved away from the area. Once their absence had been taken into account, the response rate was found to be 86%. The subjects were re-interviewed 1 year later to assess change in their eating disorder features.

Assessment of eating disorder features

The eleventh edition of the Eating Disorder Examination (EDE; Cooper & Fairburn, 1987; Fairburn & Cooper, 1993) was used to measure the behaviour and attitudes characteristic of those with eating disorders. The EDE is an investigator-based interview that assesses both the frequency of key behaviours (e.g. various forms of over-eating, self-induced vomiting, laxative and diuretic misuse) and the severity of other important aspects of the characteristic psychopathology of eating disorders. Five subscales may be derived from its ratings (Restraint, Shape Concern, Weight Concern, Bulimia and Eating Concern). The overall severity of eating disorder psychopathology may be quantified using the global EDE score (the mean of the five EDE subscale scores (Fairburn & Cooper, 1993)). Tests of the reliability and validity of the EDE support its use (Cooper *et al.* 1989; Wilson & Smith, 1989; Rosen *et al.* 1990; Beumont *et al.* 1993).

The EDE distinguishes four forms of over-eating; objective bulimic episodes, subjective bulimic episodes, episodes of objective over-eating and episodes of subjective over-eating. These are mutually exclusive, although it is not uncommon for subjects to have episodes of more than one type. The distinction between the four types is based upon the presence or absence of two features, loss of control during the episode (required for both types of bulimic episode) and the consumption of an objectively large amount of food (required for objective bulimic episodes and episodes of objective over-eating). 'Loss of control' is rated as present if the subject would have had difficulty either preventing the episode from starting or stopping once it had started. The amount eaten is classed as 'large' when it clearly exceeds the quantity that would usually be eaten under the circumstances. This decision is made by the interviewer following specific guidelines.

Fully operational eating disorder diagnoses may be derived on the basis of EDE ratings. Specific definitions have been devised for the DSM-IV concepts of anorexia nervosa, bulimia nervosa and binge eating disorder (Fairburn & Cooper, 1993). For the purposes of this study bulimia nervosa was divided into two subtypes: a purging type in which the 'purging' (vomiting or the misuse of laxatives or diuretics) occurred on average at least once weekly for 3 months and a non-purging type in which purging did not occur at all. Fourteen subjects with episodes of purging that occurred less than weekly were excluded from the analyses relevant to the evaluation of the DSM-IV scheme (but not the cluster analyses) since their status in this scheme is unclear. An operational definition of eating disorder not otherwise specified (EDNOS) was also devised which was designed to identify those who were likely to have a clinically significant disturbance of eating.

The self-report case-finding questionnaire (the Eating Disorder Examination-Questionnaire version or EDE-Q) was derived from the EDE. It has been validated against the EDE and has been shown to provide a good measure of many eating disorder features (Fairburn & Beglin, 1995). With respect to its assessment of binge eating, it is overinclusive and, as a result, the EDE-Q is vulnerable to generating false positive cases (Beglin & Fairburn, 1992*b*). This may be a limitation inherent to self-report questionnaires in general rather than being specific to this particular instrument (Wilson, 1993).

Assessment of other features

General psychiatric symptoms were measured using the Brief Symptom Inventory (Derogatis & Melisaratos, 1983). Social functioning was assessed using the British adaptation of the Social Adjustment Scale (Cooper *et al.* 1992). Self-esteem was assessed with the Robson 12-item questionnaire (Robson, 1989). Parental and personal histories of obesity were obtained by interview. Parental obesity was regarded as present when subjects reported that their parent's highest ever weight was equivalent to a body mass index of 30 or more. Personal obesity was rated as present when they provided convincing evidence of having been significantly overweight prior to the onset of any sustained disturbance in eating (for example, such that they could not take part in sport, or such that a doctor had advised them to lose weight). Histories of psychiatric disorder were obtained using the Structured Clinical Interview for DSM-III-R (SCID; Spitzer *et al.* 1992*b*) and the Family History Research Diagnostic Criteria (Endicott *et al.* 1978).

Ethics

The study was approved by the Psychiatric Research Ethics Committee of Oxfordshire Regional Health Authority.

Statistical analyses

The first step of the cluster analysis was a principal components analysis for the purpose of reducing the dimensionality of the data (Everitt & Dunn, 1991). The second step was a Ward's cluster analysis (Ward, 1963; Statistical Analysis System (SAS) Release 6.07, 1989). A combination of clinical judgement, subjective inspection of the dendrogram structure, and the results of the pseudo-F statistic, the pseudo-t^2 statistic and Cubic Clustering Criterion (CCC) were used to choose the optimal cluster solution. In addition, a general rule was applied that a meaningful cluster should consist of at least 10 subjects (Cyr *et al.* 1986; Welch *et al.* 1990). The third step involved evaluating the robustness of the Ward's solution by performing two further cluster analyses: a Ward's analysis repeated on a 75% random sample and a hierarchical complete linkage cluster analysis (SAS, 1989). The unweighted Kappa coefficient (Everitt, 1989) was used to assess the reliability of the cluster solutions.

The χ^2 test and analyses of variance with Tukey's multiple comparison procedure, as appropriate, were used to test for differences between clusters and between the DSM-IV diagnostic groups. Because this involved multiple testing across the same subjects, only significance levels which reached $P < 0{\cdot}01$ are reported. Prior to the analysis of variance, variables that were not normally distributed were transformed to rank normal.

Univariate tests of construct validity were carried out using variables external to the clustering solution. A non-parametric discriminant function analysis, the k-nearest neighbour method (SAS, 1989), was also performed. Its purpose was to test the validity of distinguishing between clusters on the basis of features external to the cluster analysis, when these features were examined from a multivariate rather than a univariate point of view.

The relative predictive validities of the original cluster solution and the DSM-IV scheme were evaluated by canonical discriminant function analyses (Grayson *et al.* 1990). This technique identifies linear combinations of quantitative (explanatory) variables that summarize best the differences between groups. The changes in the five EDE subscale scores over the year of follow-up were entered as the explanatory variables. The eigenvalues and F-statistic results were compared for each diagnostic scheme to assess which had the better predictive validity.

RESULTS

The sample

Seven hundred and fifty-five subjects were selected for interview on the basis of their responses of the EDE-Q. Of these, 142 (19%) could not subsequently be traced, and 76 (10%) declined to take part. The remaining 537 (71%) subjects completed the recruitment interview. Of these, 255 (47%) did not meet the inclusion criteria for this study (i.e. they were false positives) and 21 (4%) met the exclusion criteria, the most common reason being pregnancy. Of the remaining 261 subjects, 11 (4%) were excluded because of uncertainty over the exact frequency of their episodes of overeating or

Table 1. *The eating disorder features of the subjects in the four Ward clusters*

	Cluster 1 ($N = 30$)	Cluster 2 ($N = 86$)	Cluster 3 ($N = 30$)	Cluster 4 ($N = 102$)	F*	Tukey
Objective bulimic episodes†						
Median	28·5	27·0	0·0	8·0		
Mean	43·8	29·3	9·6	11·7	26·08	1 > 4, 1 > 3
S.D.	46·9	18·0	17·2	14·6		2 > 4, 2 > 3
Subjective bulimic episodes†						
Median	14·5	0·0	117·5	13·0		
Mean	19·9	9·6	119·4	16·3	72·21	1 < 3, 2 < 3
S.D.	23·3	17·9	48·2	16·2		2 < 4, 3 > 4
Self-induced vomiting†						
Median	76·0	0·0	0·0	0·0		
Mean	89·5	10·4	1·6	2·7	61·37	1 > 4, 1 > 2
S.D.	74·4	24·5	4·6	9·8		1 > 3, 2 > 4
Laxative or diuretic misuse†						
Median	22·0	0·0	0·0	0·0		
Mean	53·6	7·6	17·1	2·0	14·33	1 > 3, 1 > 2
S.D.	72·4	14·6	38·9	4·9		1 > 4
Shape concern‡						
Mean	3·8	3·7	3·4	2·7	17·93	1 > 4, 2 > 4
S.D.	1·4	1·1	1·1	1·1		
Weight concern‡						
Mean	3·9	3·4	3·0	2·5	16·93	1 > 3, 1 > 4
S.D.	1·4	1·2	1·3	1·0		2 > 4
Restraint‡						
Mean	3·7	3·4	2·5	2·0	28·40	1 > 3, 1 > 4
S.D.	1·1	1·0	1·3	1·3		2 > 3, 2 > 4
Global EDE score‡						
Mean	3·2	3·1	2·7	2·0	45·25	1 > 4, 2 > 4
S.D.	0·8	0·7	0·8	0·7		3 > 4

Note: where data were highly skewed the median is presented as well as the mean.
* ANOVA F statistic, df = 3 and $P < 0·001$ for all tests, Tukey test $P < 0·01$.
† Number of episodes over last 3 months.
‡ EDE = Eating Disorder Examination subscale (Fairburn & Cooper, 1993).

purging behaviour. Thus, the final sample consisted of 250 subjects with recurrent binge eating.

The mean age of the subjects was 24·7 years (S.D. = 5·5). One hundred and twelve (45%) were married or living as married, 128 (51%) were single, and 10 (4%) were divorced or separated. Their parental social class distribution was as follows (Office of Population Censuses and Surveys, 1980): 100 (40%) I or II; 107 (43%) III; and 43 (17%) IV or V.

By definition, all the subjects had recurrent episodes of binge eating. One hundred and eighty-two (73%) had objective bulimic episodes with a mean weekly frequency of 2·5 episodes (S.D. 2·1); 149 (60%) had subjective bulimic episodes with a mean weekly frequency of 3·8 episodes (S.D. 3·9); 31 (12%) had episodes of objective over-eating; 88 (35%) practised self-induced vomiting; and 87 (35%) misused laxatives or diuretics. The weight distribution of the sample was as follows (expressed as body mass index (BMI; Garrow, 1988)): 16 (7%) < 20·0; 126 (57%) 20·0–24·9; 46 (21%) 25·0–29·9; 29 (13%) 30·0–39·9; and 5 (2%) > 40·0.

Ninety-three subjects (37%) met the EDE operational definition of bulimia nervosa; none met the criteria for anorexia nervosa; and 63 (25%) met the criteria for binge eating disorder. A further 60 (24%) subjects met criteria for EDNOS but not binge eating disorder. Thirty-four subjects (14%) fell below threshold criteria for EDNOS.

The principal components analysis

Data on 22 EDE items were entered into the principal components analysis. These items represented a comprehensive range of eating disorder behaviours and attitudes. (Details of these items are available to interested readers

upon request.) The first seven principal components which resulted were selected for use as the variables in the cluster analyses. Each had an eigenvalue greater than 1·0 and together they explained more than 80% of the variance.

Ward's cluster analysis

The pseudo-F statistic suggested a six- or two-cluster solution and the pseudo-t^2 statistic suggested a four- or six-cluster solution. The CCC did not suggest any particular clustering solution. The six-cluster solution was evaluated. It contained two clusters that had fewer than 10 subjects, one with only two subjects and one with nine subjects. The first two subjects were outliers and had atypical clinical characteristics with an unusually high frequency of episodes of objective over-eating and low scores on other eating disorder features. They were excluded from further consideration. The latter nine subject cluster joined a cluster with 21 subjects to form a branch of the cluster dendrogram at a higher level. The subjects in these two clusters had similar EDE subscale scores. It was decided to combine these two clusters thereby generating a four-cluster solution.

Replication of Ward's cluster solution

The replication of Ward's method on a 75% random sample ($N = 198$) produced a four-cluster solution, with clusters that had very similar characteristics to those produced using the full sample. The assignment of subjects to the original Ward cluster set was compared with the assignment of subjects to the corresponding new cluster sets. There was a 77% level of agreement of subject assignment (kappa = 0·65).

The pseudo-F statistic for the complete linkage clustering suggested an eight-cluster solution. For this solution the three largest groups had similar symptom profiles to three of the Ward clusters and, if regrouped, a further three clusters were similar to the fourth Ward cluster. The level of agreement of assignment of subjects to clusters with similar symptom profiles was 80% (kappa = 0·70).

Internal validation: descriptive validity

The eating disorder features of the four Ward clusters are shown in Table 1. In clinical terms they may be described as follows.

Cluster 1 ($N = 30$)

These subjects had the most severe eating disorder. They were characterized by having a high frequency of vomiting (83% had at least weekly episodes, only one had no vomiting) and laxative or diuretic misuse (47% had at least weekly episodes, 11 had none). Both objective and subjective bulimic episodes were common (66% had at least weekly objective bulimic episodes, 30% had none; 50% had at least weekly subjective bulimic episodes, 43% had none). Other eating disorder features, including dietary restraint, were prominent.

Cluster 2 ($N = 86$)

These subjects had frequent objective bulimic episodes (92% had at least weekly episodes) but few subjective bulimic episodes (65% had none, and only 29% had at least weekly episodes). More than half (61%) did not vomit (only 23% had at least weekly vomiting) and 62% did not misuse laxatives or diuretics (21% had at least weekly episodes). The levels of dietary restraint and concern about shape and weight were high.

Cluster 3 ($N = 30$)

These subjects had a high frequency of subjective bulimic episodes, all at least weekly. In contrast, 60% had no objective bulimic episodes and only 30% had at least weekly episodes. The majority (87%) did not vomit and 66% did not misuse laxatives or diuretics. Levels of dietary restraint and shape and weight concern were high.

Cluster 4 ($N = 102$)

This group was heterogeneous. It was composed of subjects with subjective bulimic episodes (61% had at least weekly episodes, 30% had none) and/or objective bulimic episodes (45% had at least weekly episodes, 37% had none). The majority (79%) did not vomit and 77% did not misuse laxatives or diuretics. Other eating disorder features were mild in severity.

The four clusters differed significantly in their eating disorder features (see Table 1). Comparisons between cluster 4 and clusters 1 and 2 were particularly likely to reach statistical significance. Clusters 1 and 2 were the most similar and differed significantly only on levels of self-induced vomiting and laxative or diuretic misuse.

Table 2. *Present state and background features of subjects in the four Ward clusters*

	Cluster 1 Mean (S.D.)	Cluster 2 Mean (S.D.)	Cluster 3 Mean (S.D.)	Cluster 4 Mean (S.D.)	*N*	ANOVA *F*	Tukey $P < 0.01$
Present state features							
Age (years)	24·3 (5·4)	24·7 (5·4)	25·2 (5·4)	24·6 (5·7)	248	2·22	NS
Social adjustment[a]	1·6 (0·5)	1·4 (0·5)	1·4 (0·5)	1·2 (0·3)	208	8·94*	1 > 4, 2 > 4
Level of psychiatric symptoms[b]	1·5 (0·8)	1·1 (0·9)	1·1 (0·9)	0·7 (0·6)	227	6·05*	1 > 4
Self-esteem[c]	34·8 (17·7)	41·5 (15·4)	41·1 (16·7)	48·5 (12·3)	194	6·86*	1 < 4
Body mass index[d]	22·7 (2·3)	25·7 (6·0)	26·7 (6·2)	25·1 (5·3)	220	2·15	NS
Alcohol use (units/week)	5·5 (6·7)	6·6 (10·4)	7·6 (13·3)	5·8 (8·3)	248	0·15	NS
Background features							
Onset age[e]	17·2 (4·2)	16·0 (4·6)	16·3 (5·6)	17·1 (4·4)	238	1·02	NS
Max alcohol use ever (units/week)	20·7 (28·1)	19·7 (28·5)	25·8 (28·4)	15·8 (25·6)	247	0·75	NS
	N (%)	*N* (%)	*N* (%)	*N* (%)	*N*	χ^2 (df = 3)	*P*
History of overweight[f]	2 (7)	16 (19)	4 (13)	18 (18)	248	2·58	NS
History of anorexia nervosa	4 (13)	9 (11)	5 (17)	11 (11)	248	1·36	NS
History of purging[g]	28 (93)	55 (64)	8 (27)	30 (29)	247	52·94	< 0·001
Dieting before binge eating[h]	12 (67)	33 (59)	8 (57)	25 (57)	132	1·60	NS
History of depression[i]	9 (30)	22 (26)	12 (40)	19 (19)	248	6·22	NS
Family eating disorder	3 (10)	15 (17)	3 (10)	10 (10)	248	2·94	NS
Parental obesity (BMI > 30)	13 (43)	33 (38)	11 (37)	40 (46)	248	0·32	NS
Family history of depression[j]	8 (27)	28 (33)	8 (27)	25 (25)	248	1·56	NS
Family alcohol/substance abuse[j]	8 (27)	26 (30)	11 (37)	27 (26)	248	1·32	NS

[a] Total role area score derived from the Social Adjustment Scale (Cooper *et al.* 1982).
[b] General severity index (GSI) derived from the Symptom Checklist-53 (Derogatis & Melisaratos, 1983).
[c] Derived from Robson (1989) questionnaire.
[d] Body mass index = weight divided by height squared (in kg/m^2).
[e] Onset of a sustained (at least 3 months) eating disorder feature.
[f] Prior to the onset of a sustained eating disorder feature.
[g] Purging refers to a history of self-induced vomiting and/or laxative or diuretic misuse sustained for at least 3 months.
[h] Having a history of restrictive dieting prior to the onset of objective bulimic episodes.
[i] Diagnoses assessed using sections from the Structured Clinical Interview for DSM-III-R (Spitzer *et al.* 1992*b*).
[j] Family History Research Diagnostic Criteria (Endicott *et al.* 1978).
* $P < 0.001$.

External validation: construct validity

The construct validity of the four Ward clusters was evaluated against external characteristics (i.e. those not used in the clustering solution). The results are shown in Table 2.

There were significant differences between the cluster groups in their levels of general psychiatric disturbance, self-esteem and social adjustment. Those in clusters 1 and 2 were also more likely to have a history of laxative or diuretic misuse. There were no significant differences between the groups in any of the family or personal history variables. There was a trend for subjects in Cluster 1 to weigh less than the subjects in the other three clusters ($P < 0.1$).

The discriminant function analysis of construct validity was performed on 16 of the 17 variables listed on Table 2 and 158 (63%) subjects. This was because of the high number of missing values for the variable 'dieting preceding binge eating'. In the analysis the total misclassification rate of actual and predicted cluster groupings was as follows: cluster 1, 0%; cluster 2, 35%; cluster 3, 16%; and cluster 4, 56%.

External validation: predictive validity

Two hundred and seven (83%) of the original 250 subjects were reassessed 1 year later, 206 of whom were members of the Ward's cluster solution. Of the 43 subjects who did not complete a follow-up assessment, 22 (9%) declined to be interviewed, 19 (8%) could not be contacted and two (0·8%) were physically unwell. There were no significant differences in social class, age, age at onset of eating disorder and severity of eating disorder features between those who were followed-up and those who were not.

The subjects who were followed up were subdivided using two different classificatory schemes based on their original eating disorder

Table 3. *Outcome of the subjects in the four Ward cluster groups and the DSM-IV diagnostic groups at one year*

	Cluster 1 $N = 22$	Cluster 2 $N = 70$	Cluster 3 $N = 29$	Cluster 4 $N = 85$
Global EDE score, mean (S.D.)	2·8 (1·1)	2·2 (1·2)	1·9 (1·1)	1·6 (0·9)*
Outcome category[b]†, *N* (%)				
Bulimia nervosa	10 (46)	15 (21)	5 (17)	6 (7)
EDNOS	8 (36)	21 (30)	8 (28)	21 (25)
Non-case	3 (14)	26 (37)	10 (35)	39 (46)
Recovered	1 (5)	8 (11)	6 (21)	19 (22)

	DSM-IV diagnostic groups[a]				
	BN-P $N = 50$	BN-NP $N = 15$	BED $N = 51$	EDNOS $N = 49$	Non-case $N = 29$
Global EDE score, mean (S.D.)	2·5 (1·2)	2·2 (1·0)	1·7 (1·1)	2·0 (0·9)	1·2 (0·8)**
Outcome category[b]‡‡, *N* (%)					
Bulimia nervosa	20 (40)	2 (13)	6 (12)	5 (10)	1 (4)
EDNOS	14 (28)	7 (47)	14 (28)	16 (33)	4 (14)
Non-case	11 (22)	5 (33)	23 (45)	20 (41)	14 (48)
Recovered	5 (10)	1 (7)	8 (16)	8 (16)	10 (35)

[a] BN-P = purging bulimia nervosa; BN-NP = non-purging bulimia nervosa; BED = binge eating disorder; EDNOS = Eating Disorder Not Otherwise Specified; non-case = subjects with recurrent binge eating falling below threshold for a disorder of clinical severity.

[b] Subjects were divided into four outcome categories: bulimia nervosa (there were no subjects with anorexia nervosa); EDNOS; 'recovered' – defined as having had no objective or subjective bulimic episodes and no purging over the preceding 3 months, and having a global EDE score within 1 S.D. of the mean of a local general population sample of women (Beglin, 1990); and 'non-case' which was an intermediate group not meeting criteria for the other three categories.

* Tukey statistic: 1 > 4, 2 > 4, $P < 0{\cdot}01$. ** Tukey statistic: BN purging > Binge Eating Disorder and non-cases, $P < 0{\cdot}01$.

† $\chi^2 = 26{\cdot}8$, df = 9, $P = 0{\cdot}002$. †† $\chi^2 = 36{\cdot}5$, df = 12, $P < 0{\cdot}001$.

features. The first was Ward's clustering solution and the second was the DSM-IV scheme. Outcome was described in two ways. First, the global EDE score was used to provide a measure of the overall severity of eating disorder features. Secondly, subjects were divided into four outcome categories: bulimia nervosa or anorexia nervosa (there were no subjects with the latter diagnosis); EDNOS; 'recovered' – defined as having had no objective or subjective bulimic episodes and no purging over the preceding 3 months, and having a global EDE score within 1 S.D. of the mean of a local general population sample of women (Beglin, 1990); and an intermediate group not meeting criteria for the other three categories.

The relative outcomes of the groups derived from the cluster solution are shown in Table 3. The majority of subjects (82%) in Ward's cluster 1 met criteria for an eating disorder at follow-up whereas half (51%) in cluster 2 did so. The proportion with bulimia nervosa in the two groups was 40% and 20% respectively. Thirteen (45%) of those in cluster 3 and 27 (32%) of those in cluster 4 met criteria for an eating disorder case. The overall difference in outcome of the four clusters was statistically significant ($\chi^2 = 26{\cdot}8$, df = 9, $P = 0{\cdot}002$). A similar pattern was seen in the global EDE scores. Those in clusters 1 and 2 had significantly higher scores than those in cluster 4 (ANOVA $F = 10{\cdot}3$, df = 3, $P < 0{\cdot}001$, Tukey test $P < 0{\cdot}01$).

There were also significant differences in outcome between the DSM-IV diagnostic categories. Forty per cent of those with purging bulimia nervosa had bulimia nervosa at follow-up. A similar proportion of those with purging and non-purging bulimia nervosa met criteria for an eating disorder (68% and 60% respectively) whereas 40% and 43% of the binge eating disorder and EDNOS cases did so. The overall difference in outcome was statistically significant ($\chi^2 = 36{\cdot}5$, df = 12, $P < 0{\cdot}001$). Those in the purging bulimia nervosa group had significantly higher global EDE scores at follow-up than those in the binge eating disorder and non-case groups (ANOVA $F = 7{\cdot}85$, df = 4, $P < 0{\cdot}001$).

The changes in the five EDE subscale scores over the year of follow-up were entered as explanatory variables in two canonical discriminant function analyses, one based on

Ward's solution and the other based on the DSM-IV scheme. The sizes of the eigenvalues and levels of statistical significance of the *F*-statistic reflected the between-group differences in symptom change. Large between-group differences would support different groups following distinct courses. The scheme with the higher eigenvalue (0·12) and higher level of statistical significance ($P = 0{\cdot}0006$) on the first canonical variable was Ward's cluster solution. It was therefore regarded as having the better predictive validity. The eigenvalue for the DSM-IV scheme was 0·08 ($P = 0{\cdot}2$).

DISCUSSION

The first goal of this study was to identify clinically meaningful subgroups among those with recurrent binge eating using cluster analysis. Cluster analysis is a well established and appropriate statistical technique for suggesting possible diagnostic categories since it attempts to divide individuals into clusters of those with similar features.

The present sample was of sufficient size for a cluster analysis. Although not all subjects could be traced and 10% declined to take part, it was likely to be reasonably representative of young women with recurrent binge eating, and certainly much more so that clinical samples given their vulnerability to bias (Fairburn *et al.* 1995). The only group of such young women not represented in the sample was those with the bulimic form of anorexia nervosa. Such subjects are uncommon in community samples (King, 1989; Meadows *et al.* 1986) and are therefore difficult to recruit in large numbers. Relevant features of the specific psychopathology of eating disorders were used to cluster the subjects, and external features (including outcome at 1 year) were used to validate the cluster solution.

Ward's method produced a four-cluster solution which was replicated in a random 75% sample of the subjects and partially replicated using another cluster analysis method. The first cluster was composed of subjects with a severe eating disorder. The high frequency of self-induced vomiting and purging was their most notable feature, although they also had a high level of dietary restraint and concern about shape and weight. Their clinical picture resembled the purging form of bulimia nervosa except that their binges were not necessarily large.

The subjects in the second Ward cluster resembled the first in that they also had a severe eating disorder and they had a high frequency of objective bulimic episodes, but fewer vomited or abused laxatives or diuretics. Many could be said to have the 'non-purging' type of bulimia nervosa in that while they did not purge they had high levels of dietary restraint and extreme concerns about shape or weight. They did not resemble 'binge eating disorder', as described by Spitzer and colleagues (1992*a*, 1993) and outlined in the DSM-IV (American Psychiatric Association, 1994), in view of the severity of their other symptoms. In particular, unlike those with binge eating disorder, but like those with purging bulimia nervosa, they had a high level of dietary restraint, i.e. they used an 'inappropriate compensatory behaviour' (American Psychiatric Association, 1994).

The most distinctive feature of the subjects in the third cluster was a high frequency of subjective bulimic episodes. They did not resemble 'binge eating disorder' as defined in the DSM-IV as their bulimic episodes were not objectively large. At present subjects with this constellation of symptoms would be categorized within 'Eating Disorder Not Otherwise Specified' in the DSM-IV scheme or Atypical Bulimia Nervosa in the ICD-10 scheme (World Health Organization, 1992). Almost nothing has been written on subjects of this type (Fairburn & Walsh, 1995).

The subjects in the fourth cluster were the least disturbed and most heterogeneous group. Their bulimic episodes tended to occur in the context of few other features of an eating disorder.

The evaluation of construct validity revealed no significant differences between the cluster groups in the frequency of the family or personal history variables studied. This suggests that these factors may be general risk factors for recurrent binge eating, rather than specific risk factors for particular eating disorders. This, however, needs to be tested in case–control studies. The evaluation of construct validity did, however, find significant differences between the groups in current levels of general psychopathology, self-esteem and social adjustment. Those in the first cluster had the most severe

general psychiatric symptoms, lowest self-esteem and poorest social adjustment, while those in the fourth cluster had the least severe levels of psychosocial disturbance. The construct validity of the first and third cluster groups was supported by a multivariate analysis.

There was good support for the predictive validity of the groupings generated by the cluster analysis since there were significant differences between them at 1 year follow-up. Similarly, there was support for the predictive validity of the DSM-IV scheme. The predictive validity of the cluster solution was superior to that of the DSM-IV scheme since the change in severity of eating disorder features on discriminant function analyses significantly discriminated between the Ward's clusters but not between the DSM-IV diagnostic groups.

The findings of the study have five main implications for the classification of those with recurrent binge-eating. First, the study failed to provide evidence to support the construct 'binge eating disorder' since none of the cluster groups resembled this condition. It could be argued that this is because the sample was too young and did not include enough women with moderate to severe obesity. Studies of patients with binge eating disorder suggest that binge eating disorder, in contrast to bulimia nervosa, mainly affects women in their thirties and forties and those who are obese (e.g. Wilson *et al.* 1993; Brody *et al.* 1994). Spitzer & colleagues (1992*a*) also found binge eating disorder to be common in women attending hospital-affiliated weight control programmes. However, in the non-patient community samples binge eating disorder occurred in only 4·4% of obese (BMI > 27·5) subjects and of the 19 cases, eight had never been obese. The findings of the present study, and those of another community-based study (Dansky & Brewerton 1996), suggest that this association, namely with obesity, may reflect referral bias rather than true characteristics of the disorder. In the present study, the mean age of those with DSM-IV binge eating disorder was 25·4 years (S.D. 5·9) and 30 (48%) had a BMI of > 25.

Secondly, the findings provide support for the DSM-IV practice of subtyping bulimia nervosa according to the presence or absence of purging, since the first two Ward clusters resembled these two categories. These clusters differed from each other mainly in terms of the severity of eating disorder features rather than the features themselves. There were no differences between them with respect to levels of general psychopathology, self-esteem or social adjustment. On the other hand support was provided for retaining the distinction between them, as the two clusters differed in their outcome at 1 year.

The third finding is that the defining features of the first group (that which most closely resembled the purging type of bulimia nervosa) did not require the episodes of binge eating to be objectively large. This broad definition of binge eating is in line with earlier recommendations (Fairburn *et al.* 1986; Garner *et al.* 1991), and with the lay understanding of the term (Beglin & Fairburn, 1992*a*).

The fourth finding of note is that one of the clusters (cluster 3) is not represented in current diagnostic systems. This cluster was characterized by recurrent subjective bulimic episodes, a low frequency of vomiting and intermediate levels of dietary restraint in the presence of moderately high levels of general psychiatric symptoms. This disorder was not transitory in that almost half of the group still had an eating disorder 1 year later. Further research is needed on the characteristics and course of this group.

Finally, the findings are consistent with the hypothesis that bulimic eating disorders occur on a continuum, ranging from mild eating disorders (those with subjective bulimic episodes and a low frequency of purging behaviours) through to severe eating disorders (those with objective bulimic episodes and purging). Thus, it could be argued that there is a single bulimic eating disorder, rather than two or three different bulimic eating disorders.

Before the findings of this study can be used to justify adjusting current diagnostic practices, they will need to be replicated using a sample from a different population. Replication would provide support for the stability of the cluster solution. Any demonstration that the cluster groups have a differential response to one or more specific treatments would also add to their validity.

Financial support for the study was provided by a programme grant (18157) from the Wellcome Trust supplemented by a grant (9202) from the John D. and Catherine T. MacArthur Foundation Research

Program. Both grants were held by C.G.F., the latter obtained in collaboration with G. Terence Wilson (of Rutgers University), Judith Rodin (of Yale University) and P.J.H. P.J.H. also received personal support from a Nuffield Medical Fellowship (1991–4) and a Girdler's Junior Research Fellowship to Green College, Oxford (1991–3). C.G.F. is supported by a Wellcome Trust Senior Lectureship (13123) which also provides support for H.A.D.

We are grateful to Zafra Cooper and David Garner for their comments on the manuscript. We are also grateful to Faith Barbour, Jenny Burton, Beverley Davies, Valerie Dunn, Jan Francis, Patricia Norman, Marianne O'Connor, Sue Shaw, Angela Tremayne, Sarah Welch and Valerie Williams, who assisted with recruitment and assessment of the subjects.

REFERENCES

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

Beglin, S. J. (1990). Eating Disorders in Young Adult Women. Ph.D. thesis, University of Oxford.

Beglin, S. J. & Fairburn, C. G. (1992*a*). What is meant by the term 'Binge'? *American Journal of Psychiatry* **149**, 123–124.

Beglin, S. J. & Fairburn, C. G. (1992*b*). Evaluation of a new instrument for the detection of eating disorders in community samples. *Psychiatry Research* **44**, 191–201.

Beumont, P. J. V., Kopec-Schrader, E. M., Talbot, P. & Touyz, S. W. (1993). Measuring the specific psychopathology of eating disorder patients. *Australian and New Zealand Journal of Psychiatry* **27**, 506–511.

Brody, M. L., Walsh, B. T. & Devlin, M. J. (1994). Binge eating disorder: reliability and validity of a new diagnostic category. *Journal of Consulting and Clinical Psychology* **62**, 381–386.

Bruce, B. & Agras, W. S. (1992). Binge eating in females: a population-based investigation. *International Journal of Eating Disorders* **12**, 365–373.

Bunnell, D. W., Shenker, I. R., Nussbaun, M. P., Jacobson, M. S. & Cooper, P. J. (1990). Subclinical versus formal eating disorders: differentiating psychological features. *International Journal of Eating Disorders* **9**, 357–362.

Clinton, D. N. & Glant, R. (1992). The eating disorders spectrum of DSM-III-R. Clinical features and psychosocial concomitants of 86 consecutive cases from a Swedish urban catchment area. *Journal of Nervous and Mental Disease* **180**, 244–250.

Cooper, P., Osborn, M., Gath, D. & Feggetter, G. (1982). Evaluation of a modified self-report measure of social adjustment. *British Journal of Psychiatry* **141**, 68–75.

Cooper, Z. & Fairburn, C. G. (1987). The Eating Disorder Examination: a semistructured interview for the assessment of the specific psychopathology of eating disorders. *International Journal of Eating Disorders* **6**, 1–8.

Cooper, Z., Cooper, P. J. & Fairburn, C. G. (1989). The validity of the Eating Disorder Examination and its subscales. *British Journal of Psychiatry* **154**, 807–812.

Cyr, J. J., Atkinson, L. & Haley, G. A. (1986). A replicated cluster solution in a heterogenous psychiatric population. *Journal of Clinical Psychology* **42**, 92–99.

Dansky, B. S. & Brewerton, T. D. (1996). The National Women's Study: relationship of crime victimization and PTSD to bulimia nervosa. *International Journal of Eating Disorders* (in the press).

Derogatis, L. R. & Melisaratos, N. (1983). The brief symptom inventory: an introductory report. *Psychological Medicine* **13**, 595–605.

de Zwaan, M., Nutzinger, D. O. & Schoenbeck, G. (1992). Binge eating in overweight women. *Comprehensive Psychiatry* **33**, 256–261.

Endicott, J., Andreason, N. & Spitzer, R. (1978). *Family History – Research Diagnostic Criteria (FH-RDC) 7/18/78*. Biometrics Research. New York State Psychiatric Institute: New York.

Everitt, B. S. (1989). *Statistical Methods for Medical Investigations*. Oxford University Press: New York.

Everitt, B. S. & Dunn, G. (1991). *Applied Multivariate Data Analysis*. Edward Arnold: London.

Fairburn, C. G. & Beglin, S. J. (1995). The assessment of eating disorders: Interview or self-report questionnaire? *International Journal of Eating Disorders* **16**, 363–370.

Fairburn, C. G. & Cooper, Z. (1993). The Eating Disorder Examination, (12th edn.). In *Binge Eating: Nature, Assessment and Treatment* (ed. C. G. Fairburn and G. T. Wilson), pp. 317–360. Guilford Press: New York.

Fairburn, C. G. & Walsh, B. T. (1995). Atypical eating disorders. In *Comprehensive Textbook of Eating Disorders and Obesity* (ed. K. D. Brownell and C. G. Fairburn), pp. 135–140. Guilford Press: New York.

Fairburn, C. G. & Wilson, G. T. (1993). Binge eating: definition and classification. In *Binge Eating: Nature, Assessment and Treatment* (ed. C. G. Fairburn and G. T. Wilson), pp. 3–14. Guilford Press: New York.

Fairburn, C. G., Cooper, Z. & Cooper, P. J. (1986). The clinical features and maintenance of bulimia nervosa. In *Handbook of Eating Disorders: Physiology, Psychology and Treatment of Obesity, Anorexia and Bulimia* (ed. K. D. Brownell and J. P. Foreyt), pp. 389–404. Basic Books: New York.

Fairburn, C. G., Welch, S. L. & Hay, P. J. (1993). The classification of recurrent overeating: the 'binge eating disorder' proposal. *International Journal of Eating Disorders* **13**, 155–159.

Fairburn, C. G., Welch, S. L., Norman, P. A., O'Connor, M. E. & Doll, H. A. (1995). Bias and bulimia nervosa: how typical are clinic cases? *American Journal of Psychiatry* (in the press).

Fichter, M. M., Quadflieg, N. & Brandl, B. (1993). Recurrent overeating: An empirical comparison of binge eating disorder, bulimia nervosa, and obesity. *International Journal of Eating Disorders* **14**, 1–16.

Garner, D. M., Shafer, C. L. & Rosen, L. W. (1991). Critical appraisal of the DSM-III-R diagnostic criteria for eating disorders. In *Child Psychopathology: Diagnostic Criteria and Clinical Assessment* (ed. S. R. Hooper, G. W. Hynd and R. E. Mattison), pp. 261–303. Erlbaum: Hillsdale, NJ.

Garrow, J. S. (1988). *Obesity and Related Diseases*. Churchill Livingstone: Edinburgh.

Gormally, J., Black, S., Daston, S. & Rardin, D. (1982). The assessment of binge-eating severity among obese persons. *Addictive Behaviours* **7**, 47–55.

Grayson, D., Bridges, K., Cook, D. & Goldberg, D. (1990). The validity of diagnostic systems for common mental disorders: a comparison between the ID-CATEGO and the DSM-III systems. *Psychological Medicine* **20**, 209–218.

Hall, A. & Hay, P. J. (1991). Eating disorder patient referrals form a population region 1977–1986. *Psychological Medicine* **21**, 697–701.

Hudson, J. I., Pope, H. G. Jr, Wurtman, J., Yurgelun-Todd, D., Mark, S. & Rosenthal, N. E. (1988). Bulimia in obese individuals: relationship to normal weight bulimia. *Journal of Nervous and Mental Disease* **176**, 144–152.

Keefe, P. H., Wyshogrod, D., Weinberger, E. & Agras, W. S. (1984). Binge eating and outcome of behavioural treatment of obesity: a preliminary report. *Behavior Research and Therapy* **22**, 319–321.

King, M. B. (1989). *Eating Disorders in a General Practice Population. Prevalence, Characteristics and Follow-up at 12 to 18 months. Psychological Medicine Monograph Supplement 14*, Cambridge University Press: Cambridge.

La Porte, D. J. (1992). Treatment response in obese binge eaters: Preliminary results using a very low calorie diet (VLCD) and behaviour therapy. *Addictive Behaviors* **17**, 247–257.

Marcus, M. D., Wing, R. R. & Lamparski, D. M. (1985). Binge

eating and dietary restraint in obese patients. *Addictive Behaviors* **10**, 163–168.

Marcus, M. D., Wing, R. R. & Hopkins, J. (1988). Obese binge eaters: affect, cognitions, and response to behavioral weight control. *Journal of Consulting and Clinical Psychology* **3**, 433–439.

Marcus, M. D., Wing, R. R., Ewing, L., Kern, E., Gooding, W. & McDermott, M. (1990*a*). Psychiatric disorders among obese binge eaters. *International Journal of Eating Disorders* **9**, 69–77.

Marcus, M. D., Wing, R. R., Ewing, L., Kern, E., McDermott, M. & Gooding, W. (1990*b*). A double-blind placebo-controlled trial of fluoxetine in the treatment of obese binge eaters and non-binge-eaters. *American Journal of Psychiatry* **147**, 876–881.

Meadows, G. N., Palmer, R. L., Newball, E. U. M. & Kenrick, J. M. T. (1986). Eating attitudes and disorder in young women: a general practice based survey. *Psychological Medicine* **16**, 351–357.

Mitchell, J. E., Davis, L., Goff, G. & Pyle, R. (1986). A follow-up study of patients with bulimia. *International Journal of Eating Disorders* **5**, 441–450.

Office of the Population of Census and Surveys (1980). *Classification of Occupations and Coding Index*. Her Majesty's Stationery Office: London.

Office of Health Economics (1987). *Compendium of Health Statistics*. Office of Health Economics: London.

Prather, R. C. & Williamson, D. A. (1988). Psychopathology associated with bulimia, binge eating, and obesity. *International Journal of Eating Disorders* **7**, 177–184.

Robson, P. J. (1989). Development of a new self-report questionnaire to measure self-esteem. *Psychological Medicine* **19**, 513–518.

Rosen, J. C., Vara, L., Wendt, S. & Leitenberg, H. (1990). Validity studies of the Eating Disorder Examination. *International Journal of Eating Disorders* **9**, 519–528.

Statistical Analysis System (SAS) Institute Inc. (1989). *SAS/Stat User's Guide Version 6*, 4th edn, Vol. 1 and 2. SAS Institute Inc.: Cary, North Carolina.

Specker, S., de Zwaan, M., Raymond, N. & Mitchell, J. (1994). Psychopathology in subgroups of obese women with and without binge eating disorder. *Comprehensive Psychiatry* **35**, 185–190.

Spitzer, R. L., Devlin, M., Walsh, B. T., Hasin, D., Wing, R., Marcus, M., Stunkard, A., Wadden, T., Yanovski, S., Agras, S., Mitchell, J. & Nonas, C. (1992*a*). Binge eating disorder: a multisite field trial of the diagnostic criteria. *International Journal of Eating Disorders* **11**, 191–203.

Spitzer, R. L., Williams, J. B. W., Gibbon, M. & First, M. B. (1992*b*). The structured clinical interview for DSM-III-R (SCID) I. History, rationale, and description. *Archives of General Psychiatry* **49**, 624–629.

Spitzer, R. L., Yanovski, S., Wadden, T., Wing, R., Stunkard, A., Devlin, M., Mitchell, J., Hasin, D. & Horne, R. L. (1993). Binge eating disorder: its further validation in a multisite study. *International Journal of Eating Disorders* **13**, 137–153.

Telch, C. F. & Agras, W. S. (1994). Obesity, binge eating and psychopathology: are they related? *International Journal of Eating Disorders* **15**, 53–61.

Wadden, T. A., Foster, G. D. & Letizia, K. A. (1992). Response of obese binge eaters to treatment by very low calorie diet combined with behavior therapy. *Journal of Consulting and Clinical Psychology* **60**, 808–811.

Wadden, T. A., Foster, G. D., Letizia, K. A. & Wilk, J. E. (1993). Metabolic, anthropometric, and psychological characteristics of obese binge eaters. *International Journal of Eating Disorders* **14**, 17–25.

Ward, J. H. (1963). Hierarchical grouping to optimize an objective function. *Journal of the American Statistical Association* **58**, 236–244.

Welch, G., Hall, A. & Renner, R. (1990). Patient subgrouping in anorexia nervosa using a psychologically-based classification. *International Journal of Eating Disorders* **9**, 311–322.

Wilson, G. T. Assessment of binge eating. (1993). In *Binge Eating: Nature, Assessment and Treatment* (ed. C. G. Fairburn and G. T. Wilson), pp. 227–249. Guilford Press: New York.

Wilson, G. T. & Smith, D. (1989). Assessment of bulimia nervosa: an evaluation of the Eating Disorder Examination. *International Journal of Eating Disorders* **8**, 173–179.

Wilson, G. T. & Eldredge, K. (1991). Frequency of binge eating in bulimia nervosa: diagnostic validity. *International Journal of Eating Disorders* **10**, 557–561.

Wilson, G. T., Nonas, C. A. & Rosenblum, G. D. (1993). Assessment of binge eating in obese patients. *International Journal of Eating Disorders* **13**, 25–33.

World Health Organization (1992). *The ICD-10 Classification of Mental and Behavioural Disorders. Clinical Descriptions and Diagnostic Guidelines*. World Health Organization: Geneva.

Yanovski, S. Z. (1993). Binge eating disorder: current knowledge and future directions. *Obesity Research* **4**, 306–324.