Long-Term Course of Binge Eating Disorder and Bulimia Nervosa: Relevance for Nosology and Diagnostic Criteria

Manfred M. Fichter, MD, Dipl.-Psych[1,2]*
Norbert Quadflieg, Dipl.-Psych[1]
Susanne Hedlund, PhD[2]

ABSTRACT

Objective: To present the twelve-year outcome of binge eating disorder (BED) in 68 female inpatients compared to bulimia nervosa, purging type (BN-P; N = 196).

Method: Self and expert ratings focused on the beginning of therapy and the 12-year follow-up.

Results: 36% of BED and 28.2% of BN-P patients still received an eating disorder diagnosis at follow-up. Differences between groups were small (Eating Disorder Inventory, Structured Inventory for Anorexic and Bulimic Syndromes, Hopkins Symptom Checklist, Beck Depression Inventory). Similar predictors for BED and BN-P were identified. Psychiatric comorbidity was the predominant predictor of poor outcome in both diagnoses. Predictors for BED outcome were body dissatisfaction, sexual abuse, and impulsivity; self-injury predicted BN-P outcome.

Conclusion: Course, outcome, and mortality were similar for BED and BN-P. Both disorders had psychiatric comorbidity as the main predictor of outcome, and there was a diagnostic shift between BED and BN-P over time, pointing to their nosological proximity. Data are relevant for the formulation of DSM-V and ICD-11 diagnostic criteria. © 2008 by Wiley Periodicals, Inc.

Keywords: eating disorder; Binge eating disorder; Bullimia nervosa; long-term outcome; prediction; nosology

(Int J Eat Disord 2008; 41:577–586)

Introduction

More than a decade has passed by since preliminary research criteria for binge eating disorder (BED) were introduced in the appendix of DSM-IV. In contrast to bulimia nervosa of the nonpurging type (BN-NP), BED patients do not exhibit regular inappropriate compensatory behavior to counteract weight gain. Although BN-NP constitutes a main DSM-diagnosis, it is rarely used clinically and overlaps with BED on logical grounds.[1] Research has focused more strongly on BED, and the clinical significance of a diagnosis of this kind generally seems beyond doubt.[2] In a recent review[3] on the diagnosis validity of BED broadening of certain criteria had been suggested. For the development of the DSM-V and ICD-11 eating disorder criteria, the existing nosological difficulties should be resolved. Comparative studies on exclusively clinical samples have not been conducted so far, but some community-based studies compared BED with BN-NP or bulimia nervosa, purging type (BN-P[4,5]). Other studies compared BED outpatients with outpatients suffering from BN-P or BN-NP.[6–8] In a community sample below age 40,[9] the BN participants reported more previous treatments for eating disorders than individuals with BED. Another community study on BED and BN patients aged 16 to 35 found no differences in age, history of anorexia nervosa (AN), and severe sexual abuse but reported a higher body weight for BED.[5,10] Fairburn et al.[5,10] reported no increased risk of sexual abuse in their seemingly less disturbed sample drawn from the community. One study reported higher age and weight in BED than BN-NP outpatients but no longer duration of eating disorder in any group.[7] Another large study from the community reported higher—although nonsignificant—age in BED than in BN and nearly identical duration of eating disorder both in BED and BN.[11] BED outpatients sometimes report less depression, less drive for thinness, and less interpersonal distrust compared

to age-matched BN outpatients.[6] Other studies found higher rates of major depression in bingers.[12,13] Two studies reported no crossover from BED to AN.[4,5] In conclusion, BED and BN showed more similarities than differences in cross-sectional studies on inpatients but not on community samples. The main difference seems to be a higher body weight in BED. In general, BED pathology as assessed in field studies tends to remit over 5 years, but low severity of BED pathology, low sample size, and high dropout rates constitute methodological problems in these studies. Overall, the evidence on similarities and differences of BED and BN is limited and still inconclusive.

This study aims to contribute to the resolution of nosological difficulties by following a large number of cases of BED and BN-P unselected for body weight over 12 years. Both samples represent rather severe cases of their respective diagnostic groups. They initially received extensive multimodal inpatient treatment including CBT. Assessments occurred at hospital intake (reported here), 2-/3-year follow-up, 6-year follow-up,[14,15] and 12-year follow-up (reported here), which is much longer than the typical follow-up period of 5 years.[5]

Method

Sample

Of 635 consecutively admitted eating disordered inpatients (592 females, 43 males), 68 females met the criteria for BED according to DSM-IV research criteria, and 196 females met the criteria of BN-P at intake. BN diagnoses followed DSM-III/III-R criteria at first and were confirmed later according to the DSM-IV criteria. Based on personal information (M.F.) from the architect of the BED criteria, R. Spitzer, the comprehensive data collection system for this study included information from medical charts and the structured inventory for anorexic and bulimic syndromes (SIAB) and was designed to allow for reliable post-hoc BED diagnoses when DSM-IV was published.

Inpatient treatment lasted on average 76.7 ± 40 days [mean $\pm$ standard deviation (SD)] for patients with BED and 95.5 ± 43 days for patients with BN-P [t (df = 262) = 3.2; $p < .01$]. Average age at admission for BED and BN-P, respectively, was 29.3 ± 8.4 and 25.6 ± 6.7 years [t (df = 98 (variances not equal)) = 3.2; $p < .01$]. Ages of onset were 17.7 ± 8.9 (BED) and at 17.6 ± 4.8 year (BN-P; t-test not significant). Duration of eating disturbance was 11.6 ± 7.3 and 8.1 ± 4.9 year [BED and BN-P, respectively; t (df = 86 (variances not equal)) = 3.7; $p < .01$].

Design

Data collection occurred in the hospital during inpatient treatment and, for the follow-up, by means of a comprehensive questionnaire and telephone interviews, conducted by trained clinicians. At the start of the study, the entire procedure was explained and written informed consent was obtained from each patient.

Assessments

Eating Disturbance. SIAB From this inventory the SIAB-EX interview for expert rating[16–18] was used at the 12-year follow-up, and its self-rating version (SIAB-S)[19] during index treatment. Diagnoses at the follow-up were derived from the computerized SIAB-EX algorithm. For the SIAB-EX, data from healthy controls were available: 202 females age 18 to 30 from a random general population sample who had never suffered from an eating disorder. The self-rating Eating Disorder Inventory (EDI)[20] was used at both time points. Additional questions aimed at body weight and other relevant issues (further treatments, comorbidity etc.) The Psychiatric Status Rating (PSR) yielded an overall assessment of eating disorder outcome at 12-year follow-up.[21] This scale includes six levels: 1 ("usual self"), 2 ("residual symptoms"), 3 ("marked symptoms"), 4 ("partial remission"), 5 ("meets criteria for a DSM-IV eating disorder diagnosis"), and 6 ("very severe DSM-IV eating disorder"). Patients with an eating disorder not otherwise specified (ED-NOS) causing only minor impairment were rated as "4", whereas patients with ED-NOS causing major impairment were rated as "5".

General psychopathology at admission and follow-up was measured using the Hopkins Symptom Checklist (SCL-90),[22] and the Beck Depression Inventory (BDI).[23] At the 12-year follow-up, the Structured Clinical Interview for DSM-IV (SCID-I;[24]) yielded assessments of lifetime and 1-month comorbidity.

Several questions assessed incidents of self-injury, shop-lifting (nonfood) and sexual promiscuity lifetime until admission and were combined into a measure of impulsivity.[25] Personality and family climate were assessed by standardized questionnaires.[17,26] Additional items used at admission covered mental illness in the patient's family of origin and social functioning. Sexual abuse was assessed by interview at the 12-year follow-up, rating the severity on a scale from 0 (no sexual abuse) to four (very severe abuse: frequent and violent).

Statistical Analyses. Analyses of variance (MANOVAs) with one between-subject factor (diagnosis) and one within-subject factor (time) preceded t-tests for simple group and time effects. For longitudinal comparisons, only sets of data complete for both time points were analyzed.

Controlling for weight and age yielded only minimally different results and we therefore decided to report the "pure" data.

ANOVAs with post hoc Scheffé range tests were used for comparing more than two groups.

For categorical variables cross-tabulations were used with χ^2-tests, and for tables including at least one cell with an expected frequency below 5 Fisher's exact test was applied. For ordered categories the Mann-Whitney-U-test was used.

Standardized mortality ratios were established on the basis of expected deaths between March 1987 and December 1999 in the West German female population, controlled for age groups as derived from mortality figures reported in the Federal Health Monitoring System of the Federal Statistical Office. Four outcome levels at 12-year follow-up summarized results across measures: (1) Good (0: No eating disorder present) versus poor diagnostic outcome (1: AN, BN, BED, ED-NOS present); (2) Good (0: PSR-rating 1 or 2) versus poor general severity outcome (1: PSR-rating 3 through 6); (3) Good (0: no binges at all) versus poor bingeing episode outcome [1: one or more binges occurred in the three months preceding follow-up)]; and (4) Good (0: no or small and infrequent binges, not meeting the DSM-IV definition) versus poor bingeing severity outcome (1: severe and frequent binges, meeting DSM-IV definition).

Potential outcome predictors were derived from the literature about BED[6,10,27-30] and BN.[31-33] This list included predictors concerning history of eating disorder, family of origin, psychiatric comorbidity, severity of eating disorder, personality, body dissatisfaction, degree of interoceptive awareness and history of severe sexual abuse. In the BED group, bivariate χ^2- and t-tests of potential predictors were computed for each outcome separately, and significant ($p < .05$) variables were introduced into stepwise logistic regression analyses in order to identify predictors of 12-year eating disorder outcome in BED. Logistic regression analyses with the same predictors as in BED for each outcome were then repeated for BN-P.

Results

The 12-year follow-up occurred at 12.6 ± 0.8 year (BED) and 12.4 ± 0.9 year (BN-P) after the end of index treatment (t-test not significant). Age at follow-up was 41.8 ± 8.1 year (BED) and 38.2 ± 6.5 year (BN-P) (t (df = 92 (variances not equal)) = 3.2; $p < .01$). At follow-up we reassessed 91% (62/68) of BED patients and 83% (163/196) of BN-P patients.

Two patients (2.9%) with BED and four patients (2.0%) with BN-P were deceased at 12-year follow-up.

Dropouts from the follow-up (excluding deceased individuals) did not differ systematically from study participants on relevant measures. At intake, patients with BED and BN-P did not differ regarding comorbid psychiatric disorders (alcohol, drugs, depression), obesity and sexual taboos in their family of origin, high risk regarding low "parental care" and high "parental control," childhood obesity, self-esteem, achievement orientation, social functioning, sexual problems in general and with men, extraversion, emotional lability, and previous eating disorder treatment.

None of the BED patients but 54 (27.6%) of the BN-P patients presented with a history of AN (χ^2 (df = 1) = 23.6; $p < .01$). More BED patients ($N = 14$ of 68; 20.6%) than BN-P patients ($N = 12$ of 193; 6.2%) reported an early onset of menarche (before or at age 11 year; χ^2 (df = 1) = 11.6; $p < .01$). Twenty-two of 54 (40.7%) BED patients had been pregnant at least once in their life compared to 28 of 118 (23.7%) BN-P patients (χ^2 (df = 1) = 5.2; $p < .05$). Impulsivity (combined scores for self-injury, nonfood shop lifting, and sexual promiscuity) was found in 45 patients with BED (66.2%) and 157 patients with BN-P (80.1%) (χ^2 (df = 1) = 5.5; $p < .05$) at the beginning of inpatient treatment. In the subcategories of impulsivity, significant differences were found for self-injury comparing BED ($N = 22$; 32.4%) and BN-P ($N = 94$; 48.0%) (χ^2 (df = 1) = 5.0; $p < .05$) and for nonfood shop lifting, BED: $N = 28$ (41.2%) and BN-P: $N = 116$ (59.2%); χ^2 (df = 1) = 6.6; $p < .05$.

Bingeing and Compensatory Behavior

BED patients binged less than patients with BN-P at intake (SIAB item binging mean = 2.6 ± 0.7 and 3.3 ± 0.8, respectively; $t = 6.4$; df = 262; $p < .01$). At 12-year follow-up rates of binges and compensatory behavior were low and did not differ between BED and BN.

Body Weight

BED patients were significantly heavier than patients with BN-P. A BMI of over 30 was found in 64.7% (44/68) of BED patients and 7.7% (15/196) of BN-P patients (χ^2 (df = 2) = 96.3; $p < 0.01$). BED patients managed to reduce their weight from an average BMI of 34.0 ± 9.0 at admission to 32.0 ± 9.2 at 12-year follow-up ($N = 61$), while BN-P patients gained some weight, from 21.6 ± 5.1 to 22.2 ± 5.3 ($N = 163$). MANOVA (df = 1,222) yielded significant main effects for group ($F = 145.8$;

TABLE 1. Eating disorder outcome at 12-year follow-up in female patients with binge eating disorder and bulimia nervosa, purging type (DSM-IV)

	Binge Eating Disorder N (%)	Bulimia Nervosa N (%)
Diagnostic 12-year outcome (DSM-IV; SIAB-EX)	(N = 64)	(N = 167)
No major eating disorder	43 (67.2)	117 (70.1)
BMI ≤ 30	20 (31.3)	111 (66.5)
BMI > 30	23 (35.9)	6 (3.6)
Anorexia nervosa, restricting type	0	1 (0.6)
Anorexia nervosa, binge eating/purging type	0	2 (1.2)
Bulimia nervosa, purging type	6 (9.4)	17 (10.2)
Bulimia nervosa, non-purging type	0	1 (0.6)
Binge eating disorder	5 (7.8)	3 (1.8)
Eating disorder, not otherwise specified	8 (12.5)	22 (13.2)
Deceased (χ^2 = 6.6, n.s., df = 7)	2 (3.1)	4 (2.4)
Any Eating Disorder Diagnosis (χ^2 n.s., df = 1)	19 (30.6)	46 (28.2)
Global 12-year outcome (PSR)	(N = 62)	(N = 163)
Good (rating 1 or 2)	23 (37.1)	84 (51.5)
Intermediate (rating 3 or 4)	23 (37.1)	40 (24.5)
Poor (rating 5 or 6) (n.s.)	16 (25.8)	39 (23.9)

n.s., not significant; PSR, psychiatric status rating.

$p < .01$) and time ($F = 4.9$; $p < .05$) as well as a significant interaction effect ($F = 15.0$; $p < .01$). Simple tests indicated significant ($p < .01$) differences between BED and BN at both time points ($t = 10.1$ and $t = 7.9$; df = 222) and significant ($p < .05$) changes over time ($t = 2.4$ for BED and $t = 2.2$ for BN-P).

Diagnostic Outcome of Eating Disorders

Table 1 summarizes eating disorder diagnoses at 12-year follow-up. BED patients never crossed over to AN. A similar percentage of both groups received the original diagnosis at 12-year follow-up (7.8% in BED and 10.2% in BN-P) or had no major eating disorder any more (67.2% in BED and 70.1% in BN-P). The rate of individuals with any eating disorder diagnosis was not significantly different between groups.

At the 12-year follow-up, the PSR indicated a somewhat better, but not significantly different outcome for BN-P (**Table 1**).

Twelve-Year Outcome of Eating Disorder Symptoms

Table 2 presents the SIAB-EX[16] outcome.

In spite of significant symptom reductions over time, most SIAB-EX scores remained significantly elevated after 12 year when compared to healthy controls. As our groups included also patients who still suffered from an eating disorder at the 12-year follow-up, we added a further analysis that included only patients defined as recovered (no DSM-IV eating disorder including ED-NOS) irre-

TABLE 2. Results [means and standard deviations (in parentheses)] of the structured interview for anorexic and bulimic syndromes (SIAB-EX) in female patients with binge eating disorder (BED) and bulimia nervosa (BN), purging type at 12-year follow-up

	A	B	C	D	E	ANOVA	ANOVA
	BED all	BN all	BED Recovered[a]	BN Recovered[a]	Healthy Controls	A vs. B vs. E df = 2,416, F[b]	C vs. D vs. E df = 2,312, F[c]
	N = 59	N = 158	N = 24	N = 89	N = 202		
Range from 0 (not present) to 4 (severely present)							
Body image and slimness ideal	0.7 (0.4)	0.7 (0.4)	0.5[d] (0.3)	0.4[d] (0.2)	0.3[e] (0.2)	60.0*	7.8*
General psychopathology and social integration	0.8 (0.7)	0.7 (0.5)	0.5[d] (0.6)	0.6[e] (0.5)	0.3[d] (0.4)	37.0*	11.5*
Sexuality	1.3 (1.1)	1.1 (1.1)	1.0[d] (1.0)	0.9[e] (0.9)	0.6[d] (0.8)	20.2*	6.0*
Bulimic symptoms	0.9 (1.0)	0.8 (1.1)	0.2 (0.3)	0.2 (0.3)	0.2 (0.3)	32.8*	1.4
Inappropriate Compensatory behaviors to counteract weight gain; fasting and substance abuse	0.2 (0.2)	0.2 (0.2)	0.1 (0.1)	0.1 (0.1)	0.1 (0.1)	6.6*	2.3
Atypical Binges	0.8 (0.8)	0.4 (0.6)	0.3 (0.4)	0.2 (0.3)	0.2 (0.4)	25.2*	1.5
Total Scale	0.7 (0.5)	0.6 (0.4)	0.4[d] (0.3)	0.4[d] (0.2)	0.3[e] (0.2)	69.6*	9.7*

df, degrees of freedom.
[a] Psychiatric Status Rating (PSR) = 1 or 2.
[b] Scheffé post-hoc tests indicated significant (p < .05) differences between controls and BN or BED for all subscales. Additionally, the post-hoc test between BED and BN was significant (p < .05) for 'Atypical Binges'.
[c,d,e] Refer to this ANOVA. Different letters after the means indicate groups which, according to post-hoc-Scheffé-tests, differed significantly (p < 0.05) from one another.
*p < .01.

TABLE 3. Results [means and standard deviations (in Parentheses)] of the eating disorder inventory (EDI) in female patients with binge eating disorder and bulimia nervosa, purging type (DSM-IV)

	Binge Eating Disorder $N = 50$		Bulimia Nervosa $N = 137$		MANOVA df $= 1,185$		
	Beginning of Inpatient Treatment	12-year Follow-Up	Beginning of Inpatient Treatment	12-year Follow-Up	F (Group)	F (Time)	F (Group × Time)
Total Score	84.3 (24.1)	49.1[a] (29.4)	86.5 (32.2)	33.8 (25.5)	2.8	278.4**	10.9**
Drive for Thinness	11.2 (4.8)	5.3[a] (5.5)	12.2 (5.5)	3.2 (4.3)	0.6	222.9**	9.2**
Bulimia	9.8[a] (4.7)	3.3 (4.7)	12.3 (5.0)	2.3 (4.0)	1.8	261.2**	11.1**
Body Dissatisfaction	23.6[a] (4.4)	17.1[a] (8.9)	16.1 (8.9)	7.9 (7.9)	53.3**	105.4**	1.3
Ineffectiveness	11.1 (7.4)	7.0 (6.8)	12.5 (7.0)	5.1 (5.4)	0.0	89.0**	7.2**
Perfectionism	5.1[a] (3.9)	4.0 (3.8)	7.0 (4.8)	4.9 (3.5)	5.7*	21.1**	2.1
Interpersonal Distrust	6.7 (4.5)	4.1 (3.4)	7.1 (4.5)	3.6 (3.8)	0.0	70.0**	1.5
Interoceptive Awareness	10.1[a] (5.9)	4.1 (5.3)	12.3 (6.0)	3.9 (5.0)	1.9	195.9**	5.7*
Maturity Fears	6.9 (4.7)	4.0 (3.2)	7.2 (5.4)	3.1 (2.6)	0.2	66.5**	1.9

All *t*-tests comparing means at beginning of treatment and 12-year follow-up within BED and BN-P were significant ($p < .05$). df, degrees of freedom.
[a] *t*-Test BED versus BN $p < .05$.
*$p < .05$; **$p < .01$.

TABLE 4. Psychiatric comorbidity (SCID-I, DSM-IV assessed at 12-year follow-up) in female patients with binge eating disorder (BED) and bulimia nervosa, purging type (BN) (DSM-IV)

	Lifetime			One month		
	BED $N = 59$, N (%)	BN $N = 158$, N (%)	χ^2 (df $= 1$)	BED $N = 59$, N (%)	BN $N = 158$, N (%)	χ^2 (df $= 1$)
Mood disorders	**40 (67.8)**	**109 (69.0)**	**n.s.**	**15 (25.4)**	**26 (16.5)**	**n.s.**
Major depressive episode	30 (50.8)	92 (58.2)	n.s.	9 (15.3)	17 (10.8)	n.s.
Dysthymic disorder	12 (20.3)	23 (14.6)	n.s.	6 (10.2)	8 (5.1)	n.s.
Bipolar I disorder	0 (0.0)	1 (0.6)	n.s.	0 (0.0)	1 (0.6)	n.s.
Cyclothymia	0 (0.0)	0 (0.0)	n.s.	0 (0.0)	0 (0.0)	n.s.
Anxiety disorder	**22 (37.3)**	**57 (36.1)**	**n.s.**	**15 (25.4)**	**35 (22.2)**	**n.s.**
Panic disorder	13 (22.0)	18 (11.4)	4.0*	9 (15.3)	10 (6.3)	4.3*
Agoraphobia	8 (13.6)	13 (8.2)	n.s.	6 (10.2)	10 (6.3)	n.s.
Social phobia	9 (15.3)	14 (8.9)	n.s.	8 (13.6)	5 (3.2)	8.2**
Specific phobia	4 (6.8)	14 (8.9)	n.s.	4 (6.8)	10 (6.3)	n.s.
Generalized anxiety disorder	9 (15.3)	22 (13.9)	n.s.	6 (10.2)	13 (8.2)	n.s.
Obsessive-compulsive disorder	6 (10.2)	12 (7.6)	n.s.	6 (10.2)	7 (4.4)	n.s.
Post-traumatic stress disorder	1 (1.7)	2 (1.3)	n.s.	1 (1.7)	2 (1.3)	n.s.
Substance use disorders	**11 (18.6)**	**57 (36.1)**	**6.1***	**5 (8.5)**	**23 (14.6)**	**n.s.**
Alcohol dependence	5 (8.5)	39 (24.7)	7.0**	3 (5.1)	15 (9.5)	n.s.
Alcohol abuse	1 (1.7)	6 (3.8)	n.s.	0 (0.0)	4 (2.5)	n.s.
Drug dependence	7 (11.9)	23 (14.6)	n.s.	2 (3.4)	4 (2.5)	n.s.
Drug abuse	0 (0.0)	4 (2.5)	n.s.	0 (0.0)	0 (0.0)	n.s.
Psychotic disorder	**0 (0.0)**	**0 (0.0)**	**n.s.**	**0 (0.0)**	**0 (0.0)**	**n.s.**
Any axis-I disorder	**43 (72.9)**	**126 (79.7)**	**n.s.**	**21 (35.6)**	**65 (41.1)**	**n.s.**
Adjustment disorder	0 (0.0)	0 (0.0)	n.s.	0 (0.0)	0 (0.0)	n.s.
Borderline personality disorder	**5 (8.5)**	**15 (9.5)**	**n.s.**	—	—	—

n.s., not significant; df, degrees of freedom; —, not assessed.
*$p < .05$; **$p < .01$.

spective of body weight (obesity). In comparison to healthy controls, recovered patients exhibited significantly elevated, i.e. pathological, scores in "body image and slimness ideal," "general psychopathology and social integration," "sexuality" and the total score at 12-year follow-up. Post-hoc range tests indicated no significant differences between BED and BN-P for "body image and slimness ideal," but both groups still differed from healthy controls. Additionally, significant differences between BED and BN-P emerged for "general psychopathology and social integration" and "sex-

uality." On these two scales, BED individuals did not differ significantly from healthy controls.

Course of Eating Disturbance

All individual scales of the EDI changed significantly over time.

There were group by time interactions for many subscales of the EDI (**Table 3**). BED patients typically exhibited less pathological scores than BN-P patients at first and more pathological ones after 12 year. Pairwise comparisons indicated that, at the

TABLE 5. Predictors of 12-year eating disorder outcome in female patients with binge eating disorder and bulimia nervosa, purging type (DSM–IV)

| | Binge Eating Disorder | | | | | Bulimia Nervosa | | | | |
| | Bivariate | | | Logistic Regression (N = 58; R^2 = 0.25) | | Bivariate | | | Logistic Regression (N = 154; R^2 = 0.05) | |
Predictor	Good Outcome (N = 43), N (%)	Poor Outcome (N = 19), N (%)	Statistics χ^2, df = 1	Wald Statistics	Odds Ratio (95 % CI)	Good Outcome (N = 117), N (%)	Poor Outcome, (N = 46), N (%)	Statistics, χ^2 df = 1	Wald Statistics	Odds Ratio (95 % CI)
Poor "Diagnostic Outcome"										
Psychiatric comorbidity	24 (55.8%)	17 (89.5%)	6.7*	4.6*	6.00 (1.17–30.95)	85 (72.6%)	40 (87.0%)	3.8[m]	4.4*	2.96 (1.07–8.21)
Severe sexual abuse	4 (10.3%) (N = 39)	7 (36.8%)	5.9*	4.1*	4.55 (1.04–19.90)	20 (17.9%) (N = 112)	12 (28.6%) (N = 42)	ns	ns	—
Self-injury	10 (22.3%)	10 (52.6%)	5.2*	ns	—	56 (47.9%)	27 (58.7%)	ns	ns	—
	Good Outcome (N = 39), N (%) or M (SD)	Poor Outcome, (N = 23), N (%) or M (SD)		Logistic Regression (N = 62; R^2 = 0.32)		Good outcome (N = 79), N (%) or M (SD)	Poor Outcome (N = 84) N (%) or M (SD)		Logistic Regression (N = 162; R^2 = 0.03)	
	Good Outcome (N = 39), N (%) or M (SD)	Poor Outcome, (N = 23), N (%) or M (SD)	Statistics	Wald Statistics	Odds Ratio (95 % CI)	Good outcome (N = 79), N (%) or M (SD)	Poor Outcome (N = 84) N (%) or M (SD)	Statistics	Wald Statistics	Odds Ratio (95 % CI)
Poor "General Severity Outcome"										
Psychiatric comorbidity	30 (76.9%)	11 (47.9%)	χ^2 = 5.5*, df = 1	4.5*	3.75 (1.11–12.68)	66 (83.5%)	59 (70.2%)	χ^2 = 4.0*, df = 1	3.8[m]	2.12 (0.99–4.52)
Body dissatisfaction	24.5 (3.4)	19.7 (6.7)	t = 3.2**, df = 60	8.9**	1.20 (1.07–1.36)	16.1 (9.2), (N = 78)	15.4 (8.7)	ns	ns	—
Body mass index	36.1 (7.7)	30.0 (10.1)	t = 2.7**, df = 60	ns	—	22.0 (6.2)	21.2 (3.9)	ns	ns	—
	Good Outcome (N = 32), N (%) or M (SD)	Poor Outcome (N = 27), N (%) or M (SD)		Logistic Regression (N = 37; R^2 = 0.26)		Good Outcome, (N = 100), N (%) or M (SD)	Poor Outcome (N = 58), N (%) or M (SD)		Logistic Regression (N = 89; ns)	
	Good Outcome (N = 32), N (%) or M (SD)	Poor Outcome (N = 27), N (%) or M (SD)	Statistics	Wald Statistics	Odds Ratio (95 % CI)	Good Outcome, (N = 100), N (%) or M (SD)	Poor Outcome (N = 58), N (%) or M (SD)	Statistics	Wald Statistics	Odds Ratio (95 % CI)
Poor "Bingeing Episode Outcome"										
Psychiatric comorbidity	17 (53.1%)	23 (85.2%)	χ^2 = 6.9**, df = 1	5.2*	13.09 (1.45–118.62)	72 (72.0%)	49 (84.5%)	χ^2 = 3.2[m] df = 1	ns	—
Self-injury	7 (21.9%)	13 (48.1%)	χ^2 = 4.5*, df = 1	ns	—	51 (51.0%)	27 (46.6%) (N = 57)	ns	ns	—
Emotional lability	6.6 (1.7) (N = 28)	7.8 (1.6) (N = 25)	t = 2.5*, df = 51	ns	—	7.0 (1.7) (N = 96)	7.1 (1.5) (N = 54)	ns	ns	—
Interoceptive awareness	8.8 (5.3)	11.8 (5.9)	t = 2.1*, df = 57	ns	—	12.3 (5.5)	12.1 (6.9) (N = 57)	ns	ns	—
Obesity of patient's father	15 (62.5%) (N = 24)	6 (31.6%) (N = 19)	χ^2 = 4.1*, df = 1	ns	—	23 (40.4%) (N = 57)	10 (26.3%) (N = 38)	ns	ns	—

TABLE 5. (Continued)

	Bivariate			Logistic Regression (N = 59; R^2 = 0.39)			Bivariate			Logistic Regression (N = 157; R^2 = 0.11)	
	Good outcome (N = 42), N (%) or M (SD)	Poor Outcome, (N = 17), N (%) or M (SD)	Statistics	Wald Statistics	Odds Ratio (95 % CI)		Good Outcome. (N = 126) N (%) or M (SD)	Poor Outcome, (N = 32), N (%) or M (SD)	Statistics	Wald Statistics	Odds Ratio (95 % CI)
Poor "Bingeing Severity Outcome"											
Impulsivity	23 (54.8%)	16 (94.1%)	χ^2 = 8.4**, df = 1	5.6*	13.60 (1.57–117.68)		100 (79.4%)	27 (84.4%)	ns	ns	—
Psychiatric comorbidity	24 (57.1%)	16 (94.1%)	χ^2 = 7.6**, df = 1	5.2*	12.37 (1.42–107.79)		91 (72.2%)	30 (93.8%)	χ^2 = 6.6**, df = 1	6.5*	7.15 (1.58–32.42)
Self - injury	10 (23.8%)	10 (58.8%)	χ^2 = 6.6**, df = 1	ns	—		65 (51.6%)	13 (40.6%)	ns	3.8[m]	0.44 (0.19–1.01)
Inefficiency	9.8 (7.5)	14.9 (6.3)	t = 2.5*, df = 57	ns	—		12.3 (6.6)	12.5 (8.0) (N = 31)	ns	ns	—

CI, confidence interval; —, not applicable.
[m] = p < .10; * p < .05; ** p < .01.

beginning of inpatient treatment, BED patients exhibited significantly lower scores for EDI-Bulimia, EDI-Perfectionism, and EDI-Interoceptive awareness. EDI-Body dissatisfaction was higher in BED at both time points. At the 12-year follow-up, BED patients showed higher scores for EDI-Drive for thinness.

Course of General Psychopathology

Both groups improved significantly over time on all scales (except SCL Somatization in BED which did not change significantly over time) of the SCL-90 (e.g. the global severity index for BED and BN-P, respectively, changed from 1.4 $\pm$ 0.7/1.5 $\pm$ 0.7 at the beginning of treatment to 0.9 $\pm$ 0.8/0.7 $\pm$ 0.6 at 12-year follow-up ($F_{\text{group by time}}$ = 7.7; p < .01; df = 1, 183); BED: t = 4.8; df = 50; p < .01; BN-P: t = 12.1; df = 133; p < .01). Most subscales showed significant group by time interactions, with BED scores over time again lying between the scores of BN-P. Diagnostic groups did not differ on most subscales of the SCL at either time point. Exceptions were SCL Obsessive compulsive symptoms at baseline and at follow-up the global severity index, SCL Somatization, SCL Anxiety, and SCL Phobic anxiety (simple group effects).

Depth of depression according to the BDI for BED (N = 50) dropped from 22.8 $\pm$ 8.8 at admission to 13.7 $\pm$ 11.3 at 12-year follow-up. In BN-P (N = 137) scores changed from 23.1 $\pm$ 11.5 to 10.1 $\pm$ 9.8. Neither the group by time interaction nor the simple group effects were significant. In both groups simple time effects were significant (p < .000).

Axis-I comorbidity (SCID-I) was high in both BED and BN-P (**Table 4**). Drug abuse or dependence mostly involved prescription drugs. Higher levels of anxiety disorders and lower levels of substance use disorder were found in BED.

Additional Treatment

For all hospital categories together, the mean number of inpatient treatment days over the 12-year observation period was 149 $\pm$ 292 days for BED (N = 67) and 154 $\pm$ 299 days for BN-P (N = 193) (t-test not significant). Eighty-five percent of patients with BED and 83% of patients with BN-P received at least one additional inpatient treatment during the 12-year follow-up period (χ^2-test not significant).

Predictors of 12-Year Eating Disorder Outcome

Bivariate analyses pointed to psychiatric comorbidity, self-injury, experience of violent and severe sexual abuse, and severity of sexual abuse as significant predictors of poor diagnostic outcome in

BED. Regarding sexual abuse, only the dichotomized experience of violent sexual abuse was retained, expecting it to have higher reliability. The regression resulted in psychiatric comorbidity and sexual abuse as being significant predictors (**Table 5**), but not self-injury. In BN-P, only psychiatric comorbidity reached significance.

Good and poor general severity outcome in BED (**Table 5**) were discriminated significantly by the indices for body mass index, body dissatisfaction (EDI), and psychiatric comorbidity (bivariate analyses). Psychiatric comorbidity before index treatment and higher body dissatisfaction at the beginning of inpatient treatment predicted a poor outcome. Entering only the BMI and comorbidity into the analysis showed similar results, a higher BMI and the presence of comorbidity predicting poor outcome. Analogous analyses for BN-P resulted in psychiatric comorbidity as a significant predictor for poor outcome.

Poor bingeing episode outcome—for any type of bingeing—(**Table 5**) was predicted by psychiatric comorbidity, self-injury, emotional lability, interoceptive awareness (EDI), and a history of obesity for the patient's father (bivariate analyses). In logistic regression, only comorbidity emerged as a significant predictor in BED. No significant predictor was found for BN-P. Psychiatric comorbidity, self-injury, impulsivity, and inefficiency (EDI) at admission to inpatient treatment were significant predictors in bivariate analyses for poor "bingeing severity outcome" (**Table 5**). Results were similar for BED and BN-P. In both groups, psychiatric comorbidity was one significant predictor, the other being impulsivity for BED and self-injury for BN-P.

Mortality

Two BED patients and four BN-P patients had died before the 12-year follow-up. For BED, the crude mortality rate was 2.9%, for BN-P it was 2.0% (χ^2-test not significant). The standardized mortality ratios were 2.29 (n. s., 95% CI: 0.00–5.45) and 2.36 (n. s., 95% CI: 0.05–4.67), respectively.

Conclusion

We compared the long-term course of two eating disorders (BED and BN) in rather homogeneous samples of severely disturbed patients who were consecutively admitted for inpatient treatment. Our study has several strengths: the prospective longitudinal approach; a long follow-up period of 12 year (with follow-ups also at 3 and 6 year),[14] the

longest as yet reported for BED; a very high participation rate after 12 year; a large sample size for BED and BN-P; and the use of standardized assessments for general and eating disorder pathology.

Limitations of our study include (1) possible errors or memory biases for retrospectively collected data; however, they would most probably not differ in any systematic way between diagnostic groups. (2) Additionally, the inpatient samples introduced here represent the severe end of the symptomatic spectrum and exclude less disturbed BED and BN-P individuals not seeking treatment or attending outpatient treatment. Selection bias cannot be ruled out having studied exclusively inpatients and not a representative community sample. On the other hand the German health system allows practically all residents access to inpatient treatment for eating disorders when needed and it is unlikely that a larger number of cases with severe BED did not get inpatient treatment. (3) Another limitation is that although we collected data on interim treatments, their effects over 12 years cannot be differentiated from specific long-term effects of the index treatment itself. (4) Methodologically, this study is not a strictly naturalistic one. (5) Our data were gathered from female participants only, and the generalizability of our results to the male BED population is unclear. As known from epidemiological studies the percentage of males is higher in BED than in BN.[34] (6) Finally, stepwise logistic regression analysis has its limitations especially when its results are interpreted. Since the causal relationship between predictor variables for the course of eating disorders is little understood we cannot rule out the possibility that our final predictive models failed to identify truly independent predictors. For example, variable B may lie on the path between variable A and the outcome; if so, even if A were a potent cause of the outcome, it would be dropped in an analysis that controlled for B. However, stepwise logistic regression is a useful, accepted and widely used procedure, which helps to better understand complex associations. These results should be interpreted diligently and need replication in other studies.

"The purpose of diagnosis is to predict a course of an illness and to prescribe treatment" ([35]; p.113). Our study addresses the long-term course of illness. Two of four conceptual models for BED described in the literature are relevant[36]: 'BED constitutes a distinct eating disorder' and 'BED constitutes a variant of BN-P'. Based on our data, BED seems to be a circumscribed illness closely related to BN-P. In our high severity sample, course and outcome of BED were very similar to, and by no means more positive

than, the course and outcome of BN-P: outcomes for eating disorder and general psychopathology were very similar for both groups; the same applied to mortality rates, which were the same in BN-P and BED and lower than in AN inpatients[37]; diagnostic shifts did occur from BED to BN-P and vice versa; and psychiatric comorbidity was comparable (anxiety and mood disorders) and constituted the main predictor for the course of both eating disorders. In addition, it is known from other studies that BED and BN-P moderately but significantly respond to certain antidepressants (AN does not respond) and exhibit high placebo response rates.[38] Generally, BED and BN-P also respond to psychotherapy approaches like CBT and IPT.[2] Kraepelinian reasoning would take into account the similar course of symptoms over time and conclude that BED and BN-P are located on a nosological continuum. Knowing that a history of AN constitutes the second most powerful predictor for BN-P but not for BED, the continuum might reach from AN over BN-P (primarily unidirectional shift) and from BN-P to BED (bidirectional shifts possible).

However, there are data from studies in the community which question the conclusion of relatedness of BN-P and BED.[5,10,32] In a five-year follow-up BED showed a higher recovery rate than BN. However, in one study,[5] BN-NP patients were included; overall, the average body weight for BED was lower and the sample was less chronically ill than our BED sample. There was an attrition rate of 23% in this British sample.[5] The attrition rate in the other community study on BED in New England[4] was very high (56%) so that a sensible interpretation of the results is hardly possible. Dropouts had reported more sexual abuse, which constituted a significant predictor of poor diagnostic outcome in BED in our study. Hence, the high attrition in these community samples may have established a positive selection bias for the follow-ups. It must be taken into account that according to the DSM-IV criteria two definitions for BED exist: (1) The research diagnostic criteria (appendix) which we used, and (2) The NOS-criteria for BED. The latter criteria are wider than those used by us since they do not require "marked distress" with respect to the eating binges. Apparently both community studies used the simpler definition of BED (BED as ED-NOS), while we used the more detailed "research diagnostic criteria for BED" supplied in the DSM-IV-TR appendix in our study. Not employing the additional criterion of "marked distress" increases the probability that mild cases of binge eating enter the sample. This is most likely the major reason for the differences found: Both community samples (using broad criteria) exhibited a relatively benign course of BED, while our sample of consecutively admitted patients, who all had marked distress concerning their bingeing, did not do as well over time. There are, however, two more recent American epidemiological studies differing in their results concerning BED from the older studies. Pope et al.[39] interviewed 888 first-degree relatives of 300 overweight or obese probands (150 with BED) and concluded from their family-epidemiological data that BED was at least as chronic as the well-validated disorders AN and BN and likely represents a stable syndrome. In addition, findings from the National Comorbidity Survey Replication[11] showed that when applying the strict BED criteria including marked distress regarding binge eating in a community study, BED emerges as a severe disorder comparable to BN. Our data[15,37] (except for AN) are in accordance with the conclusions of these more recent studies.

Diagnostic criteria for DSM-V and ICD-11 need to define whether BED should be seen as an eating disorder that is truly distinct from BN or as a subgroup of BN. We are aware that we presented exclusively prospective longitudinal data on BN-P in comparison to BED. We did not present data on BN-NP. However, in discussing the nosology of eating disorders with bingeing we must also take into account other eating disorders, which have already been defined such as BN-NP and AN. By definition (DSM-IV-TR) BED and BN-NP show the same kind of eating binges and the other diagnostic criteria are also very similar. One difference is that the research criteria for BED explicitly require the presence of "marked distress concerning binge eating", and that in BN-NP fasting and excessive exercise may occur. On these logical grounds it would not make sense to maintain both BED and BN-NP in future diagnostic criteria for these disorders. While it is not our aim to design new criteria for eating disorders, it is tempting to formulate criteria for a single diagnostic category which contains both BED and BN-NP. This category (BED/BN-NP) would not be very distant from the current BN-P (and possibly DSM-IV subthreshold bulimia nervosa),[40] but would be at a greater distance from AN (no history of AN in BED). While severe obesity is rarely seen in BN-P, obesity is frequently observed in BED and BN-NP. Thus, there would be a continuum concerning a patient's body weight from AN over BN-P to the combined category of BN-NP and BED.

for his valuable advice concerning the statistics of this article.

References

1. Cooper Z, Fairburn CG. Refining the definition of binge eating disorder and nonpurging bulimia nervosa. Int J Eat Disord 2003;34 Suppl:89–95.
2. Wilfley DE, Wilson GT, Agras WS. The clinical significance of binge eating disorder. Int J Eat Disord 2003;34 Suppl:S96–S106.
3. Latner JD, Courtney C. The diagnostic validity of the criteria for binge eating disorder. Int J Eat Disord 2008;41:1–14.
4. Cachelin FM, Striegel-Moore RH, Elder KA, Pike KM, Wilfley DE, Fairburn CG. Natural course of a community sample of women with binge eating disorder. Int J Eat Disord 1999;25:45–54.
5. Fairburn CG, Cooper Z, Doll HA, Norman P, O'Connor M. The natural course of bulimia nervosa and binge eating disorder in young women. Arch Gen Psychiatry 2000;57:659–665.
6. Raymond NC, Mussell MP, Mitchell JE, de Zwaan M, Crosby RD. An age-matched comparison of subjects with binge eating disorder and bulimia nervosa. Int J Eat Disord 1995;2:135–143.
7. Santonastaso P, Ferrara S, Favaro A. Differences between binge eating disorder and nonpurging bulimia nervosa. Int J Eat Disord 1999;2:215–218.
8. Barry DT, Grilo CM, Masheb RM. Comparison of patients with bulimia nervosa, obese patients with binge eating disorder, and non-obese patients with binge eating disorder. J Nerv Ment Dis 2003;9:589–594.
9. Striegel-Moore RH, Cachelin FM, Dohm FA, Pike KM, Wilfley DE, Fairburn CG. Comparison of binge eating disorder and bulimia nervosa in a community sample. Int J Eat Disord 2001;29:157–165.
10. Fairburn CG, Doll HA, Welch SL, Hay PJ, Davies BA, O'Connor ME. Risk factors for binge eating disorder: A community-based, case-control study. Arch Gen Psychiatry 1998;55:425–432.
11. Hudson JI, Hiripi E, Pope HG Jr, Kessler RC. The prevalence and correlates of eating disorders in the national comorbidity survey replication. Biol Psychiatry 2007;61:348–358.
12. Marcus MD, Wing RR, Ewing L, Kern E, Gooding W, McDermott M. Psychiatric disorders among obese binge eaters. Int J Eat Disord 1990;9:69–77.
13. Bulik CM, Sullivan PF, Kendler KS. Medical and psychiatric morbidity in obese women with and without binge eating. Int J Eat Disord 2002;1:72–78.
14. Fichter MM, Quadflieg N, Gnutzmann A. Binge eating disorder: Treatment outcome over a 6-year course. J Psychosom Res 1998;44:385–405.
15. Fichter MM, Quadflieg N. Twelve-year course and outcome of bulimia nervosa. Psychol Med 2004;8:1395–1406.
16. Fichter MM, Herpertz S, Quadflieg N, Herpertz-Dahlmann B. Structured interview for anorexic and bulimic disorders for dsm-iv and icd-10: Updated (third) revision. Int J Eat Disord 1998;3:227–249.
17. Fichter M, Quadflieg N. The structured interview for anorexic and bulimic disorders for DSM-IV and ICD-10 (SIAB-EX): Reliability and validity. Eur Psychiatry 2001;16:38–48.
18. Fichter MM, Elton M, Engel K, Meyer AE, Mall H, Poustka F. Structured Interview for Anorexia and Bulimia nervosa (SIAB): Development of a new instrument for the assessment of eating disorders. Int J Eat Disord 1991;10:571–592.
19. Fichter MM, Quadflieg N. Comparing self- and expert rating: A self-report screening version (SIAB-S) of the structured interview for anorexic and bulimic syndromes for DSM-IV and ICD-10 (SIAB-EX). Eur Arch Psychiatry Clin Neurosci 2000;250:175–185.
20. Garner DM, Olmsted MP, Polivy J. Development and validation of a multidimensional eating disorder inventory of anorexia nervosa and bulimia. Int J Eat Disord 1983;2:14–34.
21. Herzog DB, Keller MB, Lavori PW, Ott IL. Short-term prospective study of recovery in bulimia nervosa. Psychiatry Res 1988;1:45–55.
22. Franke GH.SCL-90-R Symptom-Checkliste von L. R. Derogatis - Deutsche Version. 2. vollständig überarbeitete und neu normierte Auflage. 2002. Göttingen, Beltz Test GmbH. Ref Type: Serial (Book, Monograph).
23. Beck AT, Ward CH, Mendelson M, Mock J, Erbaugh J. An inventory for measuring depression. Arch Gen Psychiatry 1961; 4:561–571.
24. Wittchen HU, Wunderlich U, Gruschwitz S, Zaudig M.Strukturiertes Klinisches Interview für DSM-IV (SKID). 1997. Göttingen, Bern, Toronto, Seattle, Hogrefe.
25. Fichter MM, Quadflieg N, Rief W. Course of multi-impulsive bulimia. Psychol Med 1994;3:591–604.
26. Fahrenberg J, Hampel R, Selg H.FPI-R Das Freiburger Persönlichkeitsinventar. 7. überarbeitete und neu normierte Auflage. 2001. Göttingen, Bern, Toronto, Seattle, Hogrefe.
27. Peterson CB, Crow SJ, Nugent S, Mitchell JE, Engbloom S, Mussell MP. Predictors of treatment outcome for binge eating disorder. Int J Eat Disord 2000;28:131–138.
28. Stice E, Presnell K, Spangler D. Risk factors for binge eating onset in adolescent girls: A 2-year prospective investigation. Health Psychol 2002;2:131–138.
29. The McKnight Investigators. Risk factors for the onset of eating disorders in adolescent girls: Results of the McKnight longitudinal risk factor study. Am J Psychiatry 2003;160:248–254.
30. Lowe MR, Caputo GC. Binge eating in obesity: Toward the specification of predictors. Int J Eat Disorder 1991;2:248–254.
31. Halmi KA, Agras WS, Mitchell J, Wilson GT, Crow S, Bryson SW, et al. Relapse predictors of patients with bulimia nervosa who achieved abstinence through cognitive behavioral therapy. Arch Gen Psychiatry 2002;12:1105–1109.
32. Fairburn CG, Welch SL, Doll HA, Davies BA, O'Connor ME. Risk factors for bulimia nervosa. A community-based case-control study. Arch Gen Psychiatry 1997;6:509–517.
33. Keel PK, Mitchell JE. Outcome in bulimia nervosa. Am J Psychiatry 1997;3:313–321.
34. Spitzer RL, Yanovski S, Wadden T, Wing R, Marcus MD, Stunkard A, et al. Binge eating disorder: Its further validation in a multi-site study. Int J Eat Disord 1993;2:137–153.
35. Stunkard AJ, Allison KC. Binge eating disorder: Disorder or marker? Int J Eat Disord 2003;34:107–116.
36. Devlin MJ, Goldfein JA, Dobrow I. What is this thing called BED? Int J Eat Disord 2003;34:2–18.
37. Fichter MM, Quadflieg N, Hedlund S. Twelve-year course and outcome predictors of anorexia nervosa. Int J Eat Disord 2006; 39:87–100.
38. Halmi KA. The multimodal treatment of eating disorders. World Psychiatry 2005;4:69–73.
39. Pope HG Jr, Lalonde JK, Pindyck LJ, Walsh T, Bulik CM, Crow SJ, et al. Binge eating disorder: A stable syndrome. Am J Psychiatr 2006;163:2181–2183.
40. Le Grange D, Binford RB, Peterson CB, Crow SJ, Crosby RD, Klein MH, et al. DSM-IV threshold versus subthreshold bulimia nervosa. Int J Eat Disord 2006;39:462–467.