Psychological Medicine, 2004, **34**, 1395–1406. © 2004 Cambridge University Press
DOI: 10.1017/S0033291704002673 Printed in the United Kingdom

Twelve-year course and outcome of bulimia nervosa

MANFRED M. FICHTER* AND NORBERT QUADFLIEG

Department of Psychiatry, University of Munich; Klinik Roseneck–Hospital for Behavioural Medicine, Prien, Germany

ABSTRACT

Background. Since little is known about the long-term course and outcome of bulimia nervosa, the authors designed a 12-year prospective longitudinal study with five cross-sectional assessments based on a large sample of consecutively treated females with bulimia nervosa (purging type) (BN-P).

Method. One hundred and ninety-six females with BN-P were assessed (1) at the beginning of treatment, (2) at the end of treatment, (3) at 2-year follow-up, (4) at 6-year follow-up, and (5) at 12-year follow-up. In self-ratings as well as expert ratings based on interviews we assessed specific eating-disorder and general psychopathology.

Results. The general pattern of results over time showed substantial improvement during treatment, slight (in most cases non-significant) decline during the first two years after the end of treatment, and further improvement and stabilization until 12-year follow-up. At that point the majority of patients (70·1%) showed no major DSM-IV eating disorder, 13·2% had eating disorders not otherwise specified, 10·1% had BN-P and 2% had died. Very few had undergone transition to anorexia nervosa or binge-eating disorders. Logistic regression analyses showed that psychiatric co-morbidity was the best and most stable predictor for eating-disorder outcome at 2, 6 and 12 years.

Conclusions. Course and outcome of BN-P was generally more favourable than for anorexia nervosa.

INTRODUCTION

While the course of anorexia nervosa – recently reviewed by Steinhausen (2002) – is quite well documented and while numerous papers have described the short-term course of bulimia nervosa, very little is known about the long-term outcome. Bulimia (nervosa) was first described by Russell in 1979 and included in DSM-III in 1980 and thus constitutes a relatively recent addition to the medical literature. The DSM-IV stated that the long-term course of bulimia nervosa is not known (APA, 1994). Recent overviews concluded that the long-term outcome of bulimia nervosa is still unclear (Keel & Mitchell, 1997; Quadflieg & Fichter, 2003).

* Address for correspondence: Dr Manfred M. Fichter, Professor of Psychiatry, Klinik Roseneck affiliated with the University of Munich, Am Roseneck 6, 83209 Prien, Germany.
(Email: mfichter@schoen-kliniken.de)

Some studies on the longer-term course and outcome of bulimia nervosa describe the course of eating disorders without sufficient distinction between anorexia and bulimia nervosa. In other studies patients have been assessed at variable rather than fixed follow-up intervals (Abraham *et al.* 1983, 1–6 years; Swift *et al.* 1985, 2–5 years; Hsu & Sobkiewicz, 1989, 4–6 years; Mitchell *et al.* 1989, 2–5 years; Fallon *et al.* 1991, 2–9 years). Only three follow-up studies covering more than 10 years have been published. Of those, two studies reported on a limited number of bulimia nervosa patients (Collings & King, 1994, $n=44$; Abraham, 1998, $n=48$), while the third study by Mitchell and colleagues (Keel *et al.* 1999, 2000; Crow *et al.* 2002) covered an 11·5-year course of 175 women. The sample of Collings & King (1994) originally had not been intended for a

longitudinal study; moreover, the sample's participation in a controlled drug trial most likely led to narrow inclusion criteria that limited generalizability. The authors concluded that a majority of their bulimia nervosa patients showed a full or partial recovery after 10 years. In a stepwise logistic regression, unexpectedly, a family history of alcohol abuse was one of the predictors of favourable outcome, although heavy drinking in patients themselves was 'associated with continued symptoms at follow-up' (p. 86). In Keel and colleagues' (1999) sample about 70% showed full or partial remission and 30% still had an eating disorder or a relapse at 11·5-year follow-up. Poor outcome was associated with a longer duration of symptoms at baseline assessment and a lifetime history of substance use. Similarly, 73% of Abraham's (1998) sample showed recovery.

The aim of this study was to describe the longer-term (12-year) course and outcome of bulimia nervosa based on a large patient sample and to identify risk factors for an unfavourable course.

METHOD

Sample characteristics

Of 635 consecutively admitted patients with eating disorders (592 females and 43 males) 196 females fulfilled the diagnostic criteria for bulimia nervosa (purging type, BN-P) according to DSM-IV (APA, 1994). The patients were treated between September 1985 and June 1988 at the Roseneck Hospital, a 350-bed medical centre for biobehavioural treatment in Upper Bavaria, Germany. Upon admission patients filled out an extensive 50-page questionnaire including the Structured Inventory for Anorexic and Bulimic Syndromes, Self-Rating Version (SIAB-S, see below); the Eating Disorder Inventory (EDI); questions concerning details of eating binges and other eating-disorder symptoms; and questions about the development of the eating disturbance and general psychopathology (Hopkins Symptom Checklist, SCL 90) as well as onset, severity and duration of symptoms. In addition, the therapists diagnosed each patient according to DSM-III/III-R (APA, 1980, 1987) criteria since DSM-IV criteria were not available at the time of in-patient treatment. During 6- and 12-year follow-up we assessed patients using DSM-IV criteria for BN-P to ensure that these criteria had been met at the time of hospital admission[1]†. Average age at admission was $25{\cdot}6 \pm 6{\cdot}7$ years (mean $\pm$ S.D.). The in-patient treatment[2] lasted on average $95{\cdot}5 \pm 42{\cdot}6$ days. Details of the sample and treatment are given in Fichter & Quadflieg (1997).

Patients were assessed (1) at the beginning of intensive (in-patient) treatment, (2) at the end of intensive treatment, (3) at 2-year follow-up, (4) at 6-year follow-up, and (5) at 12-year follow-up. Each follow-up consisted of two steps: all participants received a comprehensive questionnaire, which contained the scales and items described below. After a patient had sent back the questionnaire, she was contacted for a detailed interview covering mostly the same areas as the questionnaire. Interviews were conducted by clinical psychologists and physicians especially trained for this task. The interviewers at the three follow-ups knew details from the patients' medical charts and their questionnaire. Patients were introduced to the aims of the prospective longitudinal study and informed consent was obtained from each patient.

Measures and instruments

Eating disturbance

The SIAB-EX interview for expert rating (Fichter *et al.* 1991, 1998; Fichter & Quadflieg, 1999, 2001*a*) was used at all follow-ups and its self-rating version (SIAB-S) (Fichter & Quadflieg, 2000) was used at 2-year and 12-year follow-ups. Participants also took the SIAB-S at the beginning and the end of intensive treatment. SIAB items and mean total scores range from 0 (symptoms not present) to 4 (very severe symptoms). For the SIAB subscales, data on healthy controls are available: 202 females aged 18–30 years from a random general population sample who had never suffered from an eating disorder (Fichter *et al.* 1998; Fichter & Quadflieg, 2001*b*).

The EDI, developed by Garner *et al.* (1983), was used as a self-rating instrument at all five points of measurement. All items of the EDI were assessed on a six-point scale ranging from never to always. Following the instructions of

† The notes will be found on p. 1404.

https://doi.org/10.1017/S0033291704002673 Published online by Cambridge University Press

Garner *et al.* (1983) the answers were recoded to a 0–3 format, with higher values indicating more disturbed eating. Additional questions addressed bodyweight and other relevant areas (treatment, other disorders, etc.).

The Psychiatric Status Rating Scale (PSR) developed by Herzog *et al.* (1988) was used as a global (eating disorder) outcome scale ranging from 1 (usual self) to 6 (eating disorder diagnostic criteria fulfilled, severe).

General psychopathology was measured using the Hopkins Symptom Checklist (SCL-90) (Derogatis *et al.* 1974) with items ranging from 0 (not at all) to 4 (very much). Lifetime psychiatric co-morbidity before index treatment (as used as a predictor in logistic regression analysis) was assessed by the therapists during inpatient treatment and at the 2-year follow-up (see Fichter & Quadflieg, 1997). At 12-year follow-up the Structured Clinical Interview for DSM-IV (SCID-I; First *et al.* 1996; Wittchen *et al.* 1997) yielded assessments of lifetime and 1-month co-morbidity. Borderline personality disorder was assessed using data gathered from the SIAB-EX and additional questions from a diagnostic checklist.

Statistical analyses

Means are presented with standard deviations. Multivariate analyses of variance (MANOVAs) with repeated measures were calculated based on five time-points (df = 4): beginning of therapy, end of therapy, 2-year follow-up, 6-year follow-up and 12-year follow-up. *Post hoc* Scheffé range tests were calculated when appropriate. For longitudinal comparisons, only sets of data complete for all time-points were analysed. Logistic regression analysis with all predictors entered in one step yielded data on outcome prediction. Standardized mortality ratio was computed on the basis of expected deaths between January 1987 and September 1999 in the West German female population controlled by age groups as derived from mortality figures reported in the Federal Health Monitoring System of the Federal Statistical Office (http://www.gbe-bund.de).

RESULTS

Follow-up assessments occurred at $2{\cdot}0 \pm 0{\cdot}7$ years (2-year follow-up), $6{\cdot}2 \pm 0{\cdot}9$ years (6-year follow-up) and $12{\cdot}4 \pm 0{\cdot}9$ years (12-year follow-up) after the end of index treatment. Completion rates were high at the 2-year follow-up (99·0%; 194/196), the 6-year follow-up (95·4%; 185/194 alive) and the 12-year follow-up (84·9%; 163/192 alive). At all follow-ups (2 years, 6 years and 12 years, respectively) we assessed 175 (89·3%), 137 (69·9%) and 139 (70·9%) individuals by questionnaire and interview, 9 (4·6%), 28 (14·3%) and 19 (9·7%) by interview alone, 3 (1·5%), 4 (2·0%) and 1 (0·5%) by questionnaire and short interview, 6 (3·1%), 15 (7·7%) and 3 (1·5%) by short interview, and 1 (0·5%) individual by questionnaire only. At 2-year follow-up one patient (0·5%) could not be reached and one patient (0·5%) refused to participate. At 6-year follow-up six patients (3·0%) could not be reached and three patients (1·5%) refused to participate. At 12-year follow-up three (1·5%) patients could not be reached and 26 (13·3%) refused participation; the 29 patients who refused to participate or could not be located did not differ systematically from study participants with regard to the relevant variables [age at index treatment, age at onset of eating disorder, duration of eating disorder, duration of index treatment, depth of depression at admission (BDI; Beck *et al.* 1961)], and EDI scales at admission (total score, drive for thinness, bulimia, body dissatisfaction).

Course of eating disturbance

EDI self-ratings typically showed substantial improvement in eating disturbance during therapy, a slight (in most cases non-significant) decline during the first 2 years after the end of therapy and further improvement and stabilization in years 3–12 after the end of treatment (see Table 1 and Fig. 1). Other self-ratings not shown in Table 1 (SIAB-S, SCL-90-R) confirmed this pattern. Body image and ideal of slimness ($F = 114{\cdot}6$, df = 4, $p < 0{\cdot}001$) as well as bulimic behaviour ($F = 103{\cdot}4$, df = 4, $p < 0{\cdot}001$) both showed a definite and significant drop in severity at 12-year follow-up compared with any previous time-point (Scheffé tests).

Bodyweight

The body mass index (BMI) increased only slightly across the observation period from 21·1

https://doi.org/10.1017/S0033291704002673 Published online by Cambridge University Press

Table 1. *Course of eating disorder symptoms in female patients with bulimia nervosa (purging type; DSM-IV) according to the eating disorder inventory (EDI) (values shown are means with standard deviations in parentheses; n = 93)*

Scale	Start of treatment	End of treatment	Follow-up 2 years	Follow-up 6 years	Follow-up 12 years	MANOVA *F* (time)
Total score	88·9[a] (31·7)	46·7[b,c] (32·3)	55·3[b] (33·2)	41·9[c,d] (29·2)	35·1[d] (24·7)	78·6***
Drive for thinness	12·5[a] (5·5)	6·8[b,c] (5·5)	7·5[b] (6·0)	5·1[c,d] (5·7)	3·3[d] (4·2)	62·8***
Bulimia	12·5[a] (4·7)	3·3[b] (4·3)	6·1[c] (5·8)	4·0[b] (5·1)	2·4[b] (4·0)	93·6***
Body dissatisfaction	16·7[a] (8·5)	10·4[b,c] (9·2)	12·2[b] (9·1)	10·2[b,c] (8·4)	8·9[c] (8·2)	24·8***
Ineffectiveness	12·8[a] (6·8)	5·7[b,c] (6·1)	7·5[b] (7·0)	5·1[c] (5·8)	5·3[b,c] (5·3)	39·1***
Perfectionism	6·8[a] (4·8)	5·4[b] (3·7)	5·7[a,b] (3·9)	5·2[b] (3·7)	4·7[b] (3·3)	8·0***
Interpersonal distrust	7·3[a] (4·8)	4·8[b] (4·4)	4·8[b] (4·2)	3·9[b,c] (4·2)	3·3[c] (3·6)	22·7***
Interoceptive awareness	12·9[a] (5·9)	5·9[b,c] (5·7)	7·2[b] (6·2)	5·4[b,c] (5·4)	4·0[c] (4·8)	58·3***
Maturity fears	7·4[a] (5·5)	4·5[b] (4·1)	4·4[b] (4·1)	3·1[b] (2·8)	3·2[b] (2·7)	28·0***

Different letters after the means indicate different groups (Scheffé test).
*** $p < 0{\cdot}001$.

(±4·5) at the end of treatment to 22·1 (±5·3) at 12-year follow-up. Obesity (BMI >30) was found in 6·3% of participants (12/192) at 2-year follow-up, in 6·0% (11/182) at 6-year follow-up and in 8·6% (14/163) at 12-year follow-up. Very low weight (BMI <17·5) was found in 6·3% (12/192) at 2-year follow-up, in 6·6% (12/182) at 6-year follow-up and in 4·9% (8/163) at 12-year follow-up.

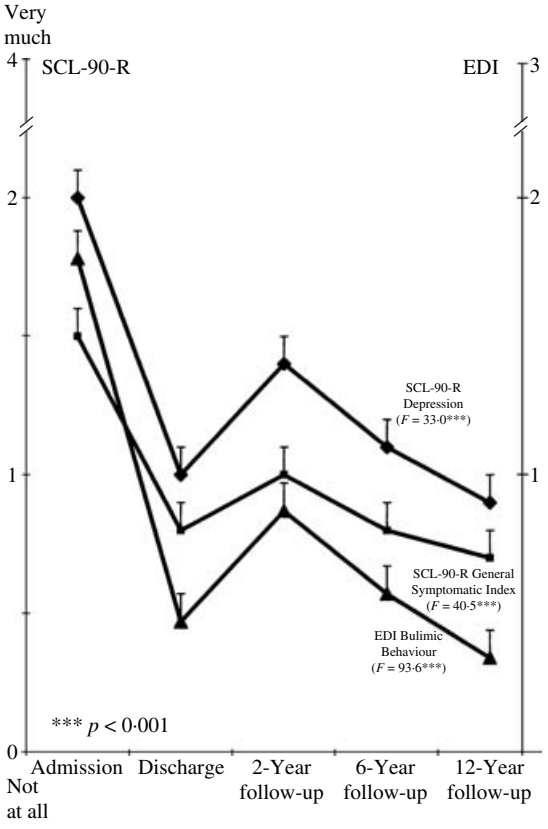

FIG. 1. Course of self-rating scales [Hopkins Symptom Checklist, Eating Disorder Inventory (EDI)] over a 12-year period in females with bulimia nervosa (EDI bulimia score was standardized by dividing by the number of items in the scale).

Diagnostic outcome of eating disorders

At 12-year follow-up a total of 167 patients were assessed for eating-disorder diagnoses. One hundred and seventeen (70·1%) were recovered and received no DSM-IV eating-disorder diagnosis [anorexia nervosa (AN), bulimia nervosa (BN), binge-eating disorder (BED), eating disorder not otherwise specified (EDNOS)]. Six of these recovered patients were obese (BMI >30). At 12-year follow-up, one patient (0·6%) had restricting type AN, two (1·2%) had binge-eating/purging type AN, 17 (10·1%) had purging type BN, one (0·6%) had non-purging type BN, three (1·8%) had BED, and 22 (13·2%) had EDNOS. Four patients were deceased at 12-year follow-up, two of them in follow-up years 1–6 and two in years 7–12[3].

Based on the total sample of 196 females with BN crude mortality rate was 2·0% (4/196) and the standardized mortality ratio was 2·36 (95% CI 0·05–4·67). Table 2 presents diagnostic outcome at all three follow-ups. The percentage of patients meeting criteria for BN decreased substantially over the 12 years. Very few patients developed AN or binge-eating disorder but some developed EDNOS.

https://doi.org/10.1017/S0033291704002673 Published online by Cambridge University Press

Table 2. *Diagnostic outcome (SIAB-EX; DSM-IV) in female patients with bulimia nervosa (purging type; DSM IV) (n = 162)**

	Follow-up 2 years	Follow-up 6 years	Follow-up 12 years
No major eating disorder	86 (53·1%)	108 (66·7%)	107 (66·0%)
Anorexia nervosa, restricting type	1 (0·6%)	2 (1·2%)	1 (0·6%)
Anorexia nervosa, binge-eating/purging type	2 (1·2%)	5 (3·1%)	2 (1·2%)
Bulimia nervosa, purging type	48 (29·6%)	34 (21·0%)	16 (9·9%)
Bulimia nervosa, non-purging type	7 (4·3%)	1 (0·6%)	1 (0·6%)
Binge-eating disorder	0	2 (1·2%)	3 (1·9%)
Eating disorder, not otherwise specified	12 (7·4%)	2 (1·2%)	22 (13·6%)
Obesity only (BMI > 30)	6 (3·7%)	6 (3·7%)	6 (3·7%)
Deceased	0	2 (1·2%)	4 (2·5%)
	100·0%	100·0%	100·0%

* Only data that were complete for all three time-points were included.

Table 3. *Twelve-year outcome according to the structured inventory for anorexic and bulimic syndromes (expert-rating SIAB-EX) in female patients with bulimia nervosa (purging type; DSM-IV) and healthy controls (values shown are means with standard deviations in parentheses)*

Subscale	A Bulimia nervosa†, recovered (n = 114)	B Bulimia nervosa†, all (n = 158)	C Healthy controls† (n = 202)	A *v.* C *t* test max. df = 289 *t* =	B *v.* C *t* test max. df = 358 *t* =
Body image and slimness ideal	0·5 (0·3)	0·7 (0·4)	0·3 (0·2)	5·8***	9·0**
General psychopathology and social integration	0·6 (0·5)	0·7 (0·5)	0·3 (0·4)	5·4***	7·5**
Sexuality	1·0 (1·0)	1·1 (1·1)	0·6 (0·8)	3·7***	5·4**
Bulimic symptoms	0·5 (0·8)	0·8 (1·1)	0·2 (0·3)	3·2***	6·8**
Inappropriate compensatory behaviours to counteract weight gain, fasting and substance abuse	0·1 (0·1)	0·2 (0·2)	0·1 (0·1)	N.S.	3·3**
Atypical binges	0·3 (0·5)	0·4 (0·6)	0·2 (0·4)	N.S.	3·6**
Total scale	0·5 (0·3)	0·6 (0·4)	0·3 (0·2)	6·5***	10·1**

† Scores of the SIAB-EX range from 0 (symptom not present) to 4 (symptom very severely present). The table shows average raw subscale scores.

df, degrees of freedom; N.S., not significant.

** $p < 0{\cdot}01$; *** $p < 0{\cdot}001$.

Twelve-year outcome of eating-disorder symptoms (SIAB-EX)

Table 3 presents the outcome assessed with the updated SIAB-EX (Fichter *et al.* 1998). Although our group of BN patients showed a significant symptom reduction over time, all SIAB-EX scores were still significantly elevated at the 12-year follow-up in relation to healthy controls. As this group included also patients who still had an eating disorder at the 12-year follow-up, we carried out another analysis including only patients defined as recovered. Comparing healthy controls with 114 recovered BN patients revealed no significant differences for the subscales 'inappropriate compensatory behaviours' and 'atypical binges'. In comparison with healthy controls at 12-year follow-up, recovered patients exhibited significantly elevated scores in the SIAB-EX subscales 'body image and slimness ideal', 'general psychopathology and social integration', 'sexuality' and 'bulimic symptoms'.

Bingeing and vomiting. At 12-year follow-up 36 of 163 patients (22·1%) binged at least twice a week, 18·4% binged less than twice a week and 59·5% had not binged at all in the preceding three months. Vomiting was observed in 20·8% at least twice a week, in 11·3% less than twice a week and not at all in 67·9% ($n = 159$).

Social integration and sexuality. As assessed in a SIAB subscale these aspects showed a significant improvement during treatment and

https://doi.org/10.1017/S0033291704002673 Published online by Cambridge University Press

continuous improvement during the follow-up period ($F=31{\cdot}4$, df $=4$, $p<0{\cdot}001$). The percentage of patients who had a more or less stable partnership increased steadily over time from 51·4% at the beginning of index treatment to 62·0% at the 2-year follow-up, 73·4% at the 6-year follow-up and 74·4% at the 12-year follow-up. While 18·1% suffered from amenorrhoea at the beginning of treatment this was the case in only 2·6% 12 years later.

Course of general psychopathology

The course over time was similar for all scales of the SCL-90. There was substantial and significant improvement during treatment, some decline until 2-year follow-up, improvement to post-treatment levels until 6-year follow-up and further improvement until 12-year follow-up.

Concerning the SCL depression scale the score at the beginning of treatment differed significantly from all other time points (Scheffé test). Discharge scores differed from scores at 2-year follow-up which in turn differed from scores at 12-year follow-up. Scores of the general symptomatic index at admission were different from all other time-points and 2-year follow-up scores differed from 12-year follow-up scores.

Psychiatric co-morbidity

Lifetime co-morbidity as assessed with SCID-I at 12-year follow-up was high at 79·7%; 1-month overall co-morbidity was 41·1%. Mood disorders were found in 69·0% (lifetime) and 16·5% (current 1-month), mostly major depression (58·2% and 10·8%) and dysthymic disorder (14·6% and 5·1%, lifetime and current respectively). Anxiety disorders were diagnosed in 36·1% (lifetime) and in 22·2% (current). Substance use disorders were found in 36·1% (lifetime) and 14·6% (current), most diagnoses being substance dependence [24·7% (lifetime) and 9·5% (current), alcohol dependence 14·6% (lifetime) and 2·5% (current), drug dependence]. Drug abuse or dependence mostly involved prescription drugs. Borderline personality disorder was diagnosed in 9·5%. No psychotic disorders were detected.

Additional treatment

Interim treatment may of course affect the course of illness. We therefore assessed in-patient treatments during the follow-up intervals, and calculated the mean number of days (mean $\pm$ S.D.) per year for treatment in different types of hospitals. For all hospital categories together the mean in-patient treatment was $15{\cdot}1 \pm 37$ days per year for the 2-year follow-up, $9{\cdot}5 \pm 29$ days per year for the 2- to 6-year follow-up and $6{\cdot}4 \pm 14$ days per year for the 6- to 12-year follow-up period. Altogether 140 of 158 patients (88·6%) received at least one in-patient treatment during the 12-year follow-up period (in addition to the index treatment); the majority ($n=120$) were treated in hospitals for somatic medicine (internal medicine, surgery, etc.); 47 received in-patient behavioural or analytic psychotherapy and 35 received in-patient psychiatric care. The average length of stay was longest in psychiatric hospitals followed by in-patient behavioural or analytic psychotherapy, in-patient rehabilitation and in-patient somatic medicine. The number of persons admitted per year to any type of in-patient institution was 31·5 for the first 2 follow-up years, 22 for the 2- to 6-year follow-up period and 18·5 for the 6- to 12-year follow-up period. Thus patients in our cohort received continuously less in-patient treatment over the 12-year follow-up period.

Predictors of outcome

The following seven possible predictors were selected on the basis of the literature (Fairburn *et al.* 1997; Keel & Mitchell, 1997; Vaz, 1998; Steinhausen, 1999; Quadflieg & Fichter, 2003) and were assessed at the beginning of index treatment:

(1) Presence of lifetime psychiatric co-morbidity prior to index treatment.
(2) Positive history of AN.
(3) Presence of childhood obesity.
(4) Higher age at onset of eating disorders (in years).
(5) Longer duration of illness (in years).
(6) High frequency of binges.
(7) Treatment for eating disorders before index treatment.

Logistic regression analyses were calculated with the criterion being whether or not an eating disorder was present at a follow-up (AN, BN, BED or EDNOS). We did this separately for each of the three follow-ups (2-year, 6-year, 12-year). In order to check these data for their stability, we repeated the same analyses using

https://doi.org/10.1017/S0033291704002673 Published online by Cambridge University Press

the PSR, which takes into account more strongly the eating disorder symptom severity; the results concerning the PSR as criterion [PSR score 1–4 (=0) *v.* 5–6 (=1)] at follow-up are presented in square brackets.

Lifetime psychiatric co-morbidity was by far the strongest and most stable predictor of unfavourable outcome, i.e. presence of eating-disorder diagnosis at follow-up; the odds ratios (ORs) for the three time-points were as follows: 2·53 (95% CI 1·06–6·06, $p<0·05$) for the 2-year follow-up, 2·81 (95% CI 1·02–7·71, $p<0·05$) for the 6-year follow-up and 2·52 (95% CI 0·93–6·80, $p<0·10$) for the 12-year follow-up. [With the PSR as criterion the ORs were 3·55 (95% CI 1·34–9·41, $p<0·05$) for the 2-year follow-up, 2·40 (95% CI 0·88–6·58, $p<0·10$) for the 6-year follow-up, and 3·71 (95% CI 1·16–11·91, $p<0·05$) for the 12-year follow-up.]

A positive history of AN was a much weaker predictor for poor outcome; the ORs from the three time-points were as follows: non-significant for the 2-year follow-up, 2·05 (95% CI 0·94–4·45, $p<0·10$) for the 6-year follow-up, and non-significant for the 12-year follow-up. [With the PSR as criterion the ORs were 2·19 (95% CI 1·03–4·65, $p<0·05$) for the 2-year follow-up, non-significant for the 6-year follow-up, and 2·38 (95% CI 1·03–5·50, $p<0·05$) for the 12-year follow-up.]

Childhood obesity was a significant predictor only at the 2-year follow-up with ORs of 2·86 (95% CI 1·02–8·06, $p<0·05$); all other tests with eating-disorder diagnosis (or PSR) were non-significant.

A higher age at onset of the eating disorder (in years) was a significant predictor for the presence of an eating-disorder diagnosis for the 12-year follow-up only (OR 1·01, 95% CI 1·01–1·16, $p<0·05$). Although this was statistically significant it has little practical relevance; it indicates that if the onset of an eating disorder were 1 year later, the risk of also having an eating-disorder diagnosis at 12-year follow-up would increase by only 1%.

Longer duration of eating disorder (in years), a higher frequency of binges and having undergone treatment for the eating disorder prior to the index treatment were not significant predictors for having an eating-disorder diagnosis at any of the three follow-ups. The same was true for all tests using the PSR rating at follow-up as the criterion.

DISCUSSION

Our large sample of women with BN (purging type) according to DSM-IV originated from consecutive admissions to in-patient treatment. We thus have a homogeneous sample with limited self-selection bias. The strength of our study consists in (*a*) the prospective approach with five cross-sectional assessments; (*b*) a large sample size of $n=196$; (*c*) a long follow-up period of 12 years; and (*d*) the use of interview as well as self-rating data. We also obtained a high participation rate over time and succeeded in re-interviewing 85% of our sample after 12 years (99% at 2-year and 95% at 6-year follow-up). This compares very favourably with other studies with long follow-up periods [Keel *et al.* 1999 (79·0% after 11·5 years); Collings & King, 1994 (90% after 10 years, using selected drug trial sample); Reas *et al.* 2000 (45% after 9 years)].

Our study also suffers from several limitations.

(1) Although telephone interviews are 'the next best thing to being there' (Tausig *et al.* 1988) these young patients might conceal more relevant information than in a face-to-face interview. However, when addressing this issue in a subsample we did not find evidence for this reporting bias (Fichter & Quadflieg, 1997). Paulsen *et al.* (1988), using the Schedule of Affective Disorders and Schizophrenia–Lifetime Version (SADS-L), and also Reich & Earls (1990) came to the same conclusion. Nevertheless, we cannot completely rule out that personal interviews could possibly have supplied more relevant information. However, very personal topics such as sexual behaviour can apparently be explored more easily by phone compared to face-to-face[4].

(2) We studied in-patients and not out-patients. During the time of recruitment for this study in-patient treatment was the major form of treatment for BN in Germany and at that time outpatient facilities for the treatment of BN were hardly available. Patients in our sample were quite chronic and severely disturbed; they received intensive treatment based on the knowledge and experience of that time (multi-modal

https://doi.org/10.1017/S0033291704002673 Published online by Cambridge University Press

behavioural treatment). Moreover, differences in the health-care system between countries must be taken into account for understanding this issue.

(3) Some data of our study were collected retrospectively (e.g. previous history of AN or mental disorder, course of illness over the time between assessments, lifetime co-morbidity). A considerable amount of literature on memory and possible biases of retrospective assessment is available (i.e. Friedman, 1993). Thus Klesges *et al*. (1995) reported from the second National Health and Nutrition Examination Survey (NHANES) that 'up to 31% of adults in this sample may have underreported dietary intake' (p. 438). In doing five sequential cross-sectional assessments we attempted to reduce the problems associated with long-term recall.

(4) Patients received treatment at the inception of the study and some received further treatment during the 12-year follow-up interval. We cannot comment on the long-term effects of our treatment and our study is not a purely naturalistic study. However, we did gather information on major treatments during the follow-up intervals.

(5) Our findings do not automatically generalize to men with BN.

Course of eating disturbance and bodyweight

Generally, our data showed substantial and highly significant improvement during inpatient treatment, slight decline during the following 2 years and further improvement and stabilization until the last assessment 12 years after treatment. Bodyweight increased slightly but significantly during the 12-year follow-up period. This can be interpreted as increasing acceptance of one's own bodyweight and appearance. This interpretation is supported by the results of several self-rating scales: the EDI 'drive for thinness', EDI 'body dissatisfaction', and SIAB subscale 'body image and ideal of slimness' all indicated a significant improvement regarding body perception and acceptance. The increase in weight in our sample was less than the average weight increase over that time period in women of the same age in the general population (Hebebrand *et al*. 1996).

Concerning diagnostic outcome, 10·1% of the patients in our study were diagnosed as having BN at 12-year follow-up. This matches the findings of the two other studies reporting on longer follow-up periods of 10–11 years [Collings & King, 1994 (9·1%); Keel *et al*. 1999 (11·0%)]. Crossover from BN to AN was rather limited in our study (1·8% at 2-year follow-up, 4·3% at 6-year follow-up and 1·8% at 12-year follow-up). Only one study with a shorter observation period reported a considerably higher crossover from BN to AN (Abraham *et al*. 1983). Other studies with a shorter follow-up period also reported low crossover rates from BN to AN ranging from 0 to 7% (Swift *et al*. 1985, 1987; Hsu & Sobkiewicz, 1989; Fallon *et al*. 1991; Johnson-Sabine *et al*. 1992; Fairburn *et al*. 1995). EDNOS including BED were found in 15% of our participants at the 12-year follow-up. Most of these patients exhibited bulimic symptoms and it can be assumed that they are at high risk of developing the full set of BN-P symptoms under stress (Keel & Mitchell, 1997; Quadflieg & Fichter, 2003). Keel *et al*. (1999) reported 18·5% of EDNOS (including BED) after 11·5 years.

Recovery occurred not only during treatment or during the period of 1 or 2 years after treatment; a considerable proportion recovered during later follow-up years. Reports on the long-term course of AN (Theander, 1970, 1985; Ratnasuriya *et al*. 1991) indicate that with more follow-up years there was a corresponding increase in remission rates but also a considerable increase in mortality. In contrast to AN, much less is known about the long-term course of BN and no data are available on the course of BN beyond 12 years. The 12-year outcome of our sample with BN-P was generally somewhat more favourable than that of AN (Theander, 1985, 1996; Deter & Herzog, 1994). BN and AN are similar in that the number of recovered or improved patients increased as did mortality over time. Nevertheless BN remains a serious disorder deserving early and intense intervention. Herzog *et al*. (1999) reported full recovery in only 53% of females treated for BN during the first 2 years of follow-up, leaving a considerable number of women suffering from severe symptoms after the end of treatment. Future studies should examine whether recovery can still occur more than 12 years after treatment.

Outcome of eating-disorder symptomatology

Even in patients classified as 'recovered' some minor symptoms may remain. Ideally, patients fully recovered from BN should have the same scores as healthy controls on scales measuring symptoms and attitudes relevant to this disorder. Using a healthy control group of women at the same age as our female patients with BN-P, we could show that both groups did not differ regarding 'inappropriate compensatory behaviours to counteract weight gain, fasting and substance abuse' and 'atypical binges'. However, scores regarding 'bulimic symptoms', 'body image disturbance and ideal of slimness', 'general psychopathology', and impaired 'sexuality and social integration' were still elevated at 12-year follow-up. For the BN-P total group at 12-year follow-up bulimic symptoms as measured by the SIAB-EX were four times higher than for the healthy controls, indicating that some core disturbance of BN is still present, although to a smaller degree. However, 'inappropriate compensatory behaviour' was at a level almost as low as in the healthy control group. Compared with healthy controls, behaviours and attitudes regarding sexuality of BN-P patients still showed some impairment at 12-year follow-up. Social functioning exhibited a general improvement in our study, similar to others (Fallon *et al.* 1991; Johnson-Sabine *et al.* 1992; Maddocks *et al.* 1992). After 11·5 years social adjustment, functioning at work, leisure activities and family relationships were found to be greatly improved (Keel *et al.* 2000). Crow *et al.* (2002) found a higher percentage of amenorrhoea at follow-up than we did.

Course of general psychopathology

In comparison to scores of women aged 30–39 years from a representative sample from the German population (Franke, 2002) the SCL-90 scores of our BN-P sample were still higher even at 12-year follow-up. Apparently both in eating-disorder and general psychopathology, the average patient reaches an improved level of well-being but remains more disturbed than women who have never had an eating disorder.

Mortality

Mortality in our study was 0% at 2-year follow-up, 1·1% at 6-year follow-up and 2·0% at 12-year follow-up. Other studies with a follow-up period of 10 years or more reported a crude mortality of 2·2% (Collings & King, 1994), 2·3%, (Abraham, 1998) and 0·6% (Keel *et al.* 1999). For shorter follow-up periods reports of crude (non-standardized) death rates for patients with BN ranged from 1·1 to 5·8% (Mitchell *et al.* 1989; Fallon *et al.* 1991; Maddocks *et al.* 1992; Fairburn *et al.* 1995). Only one study reported standardized mortality rates for BN (Patton, 1988). The rate was much higher (9·38) than in our study (2·36). In both studies the elevation of standardized mortality rates was not statistically significant. An update of reviews on the mortality of eating disorders is warranted for BN. Mortality for BN-P is clearly lower than that for AN. Of 11 cases with an eating disorder who died, as reported by Keel *et al.* (2003), only one was diagnosed with BN while the remaining 10 suffered from either restricting AN ($n=5$) or binge-purging AN ($n=5$).

Predictors of outcome

It is a common concept that co-morbid psychiatric disorders may complicate treatment and unfavourably influence the course of BN. In our study lifetime psychiatric co-morbidity was the strongest and most stable predictor for poor outcome concerning eating-disorder diagnoses at 2-, 6- and 12-year follow-up. Although our sample was quite large it was not large enough to allow meaningful predictor analyses for the breakdown of co-morbid psychiatric diagnoses such as mood disorder, anxiety disorder and substance use disorder. Keel & Mitchell (1997) concluded in their review that depression and substance use were not predictors of outcome in BN. Other studies by Herzog *et al.* (1999) and Reas *et al.* (2000) confirmed this conclusion for longer follow-up periods. However, Bulik *et al.* (1998) reported that major depression before treatment was associated with an unfavourable course. Other larger studies with shorter follow-up intervals also suggested that co-morbid depression (Herzog *et al.* 1988; Keller *et al.* 1989) or a history of substance/alcohol abuse (Abraham *et al.* 1983; Mitchell *et al.* 1986; Fallon *et al.* 1991; Herzog *et al.* 1991) predicted negative outcome. In addition the other two studies with a long follow-up period (>10 years) reported

https://doi.org/10.1017/S0033291704002673 Published online by Cambridge University Press

that substance use (Keel *et al.* 1999) and depression, respectively, predicted a poor outcome (Collings & King, 1994). Analyses using structural equation modelling based on our sample (Fichter *et al.* 2003) took into account the 6 years post-treatment and demonstrated that specific eating-disorder pathology was mainly influenced by specific eating-disorder pathology at earlier time-points and not by (non-eating-specific) general psychopathology, and vice versa.

Concerning late age at onset as predictor of poor outcome, reviews concluded that data on the relevance of age at onset for the prognosis of BN are equivocal. The majority of studies found age at onset to be no predictor (Fichter & Quadflieg, 1995; Keel & Mitchell, 1997; Quadflieg & Fichter, 2003). Keel & Mitchell (1997) did not find age at onset to be a predictor of outcome after 11·5 years, while Collings & King (1994; follow-up >10 years) report that patients who recovered had been younger at the onset of their eating disorder than those who did not recover. According to our data a higher age at onset was a rather unstable predictor; it was a non-significant predictor for the 2- and 6-year follow-up and only a statistically significant but weak predictor at 12-year follow-up.

In summary, lifetime psychiatric co-morbidity was, of all possible predictors which we analysed, by far the most stable and strongest predictor of an unfavourable course of the eating disorder. Our regression analyses, however, explained only 8·3–11·3% of the total variance. The short-term course could not be better predicted than the long-term course. Future research may want to take into account other possible predictors which are harder to measure, such as family interaction, chronic problems in living, quality of interim treatment, and so on. At this point in time the clear distinction cannot be made whether the course of BN is simply difficult to predict or whether there are other variables besides psychiatric co-morbidity that would add significantly to the power of prediction for the long-term course of BN.

ACKNOWLEDGEMENTS

This research was funded by the Wilhelm-Sander-Stiftung, Munich, Germany, grant number 1991.004.2 (6- and 12-year course) and the German Bundesministerium für Bildung, Forschung und Technologie (BMBF) (2-year course, FKZ 0702623-8). We especially thank Dipl.-Psych. Heidelinde Krenn and Dipl.-Psych. Dr Ursula Wunderlich for their sensitive way of conducting interviews with the patients. We thank Dr Susanne Hedlund for her valuable advice in formulating the manuscript. We thank Professor Dr Johannes Hebebrand, University of Marburg, for making available the complete tables of BMI percentiles from the German National Nutrition Survey.

DECLARATION OF INTEREST

None.

NOTES

[1] For this decision we also took into account the self-ratings at admission and chart information. Because of the severity of the eating disorder at the beginning of in-patient treatment, all 196 patients also had met DSM-IV criteria then.

[2] The duration of treatment can be explained by the special character of the German health care system: at the time of the study special out-patient treatment facilities for eating disorders were rare and in-patient treatment was quite cost-effective (US$140 per day) including room, board and all treatment expenses.

[3] Case 1 died aged 30 of pneumonia 54 months after index treatment (no major eating disorder at 2-year follow-up); case 2 died aged 28 of encephalitis and pneumonia with cardiac problems 57 months after index treatment (she had fulfilled criteria for BN-P at 2-year follow-up); case 3 died aged 41 from suicide, 101 months after index treatment (alcohol dependence and very low bodyweight) and case 4 died aged 58, 135 months post-treatment (from the sequelae of a stroke, having suffered from a metabolic syndrome and obesity for several years).

[4] In addition to these data, at 12-year follow-up 27 patients were assessed twice (once by telephone, once face-to-face) by two different interviewers blind to each other. For 18 patients the first interview was scheduled by phone and the second face-to-face. For nine patients the sequence was reversed. There were no significant effects of interview modality or sequence. Pearson correlations of the total score and of the subscales of the SIAB-EX were medium to high, ranging from

$r = 0{\cdot}53$ to $r = 0{\cdot}93$, with most coefficients being higher than 0·70. Mean scores did not differ significantly (t test).

REFERENCES

Abraham, S. (1998). Sexuality and reproduction in bulimia nervosa patients over 10 years. *Journal of Psychosomatic Research* **44**, 491–502.

Abraham, S. F., Mira, M. & Llewellyn-Jones, D. (1983). Bulimia: a study of outcome. *International Journal of Eating Disorders* **2**, 175–180.

APA (1980). *Diagnostic and Statistical Manual of Mental Disorders* (*DSM-III*). American Psychiatric Association: Washington, DC.

APA (1987). *Diagnostic and Statistical Manual of Mental Disorders–Revised* (*DSM-III-R*). American Psychiatric Association: Washington, DC.

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

Beck, A. T., Ward, C. H., Mendelson, M., Mock, J. & Erbaugh, J. (1961). An inventory for measuring depression. *Archives of General Psychiatry* **4**, 561–571.

Bulik, C. M., Sullivan, P. F., Joyce, P. R., Carter, F. A. & McIntosh, V. V. (1998). Predictors of 1-year treatment outcome in bulimia nervosa. *Comprehensive Psychiatry* **39**, 206–214.

Collings, S. & King, M. (1994). Ten-year follow-up of 50 patients with bulimia nervosa. *British Journal of Psychiatry* **164**, 80–87.

Crow, S., Thuras, P., Keel, P. K. & Mitchell, J. E. (2002). Long-term menstrual and reproductive function in patients with bulimia nervosa. *American Journal of Psychiatry* **159**, 1048–1050.

Derogatis, L. R., Liberman, R. S., Rickels, K., Uhlenhutz, E. H. & Cor, L. (1974). The Hopkins Symptom Checklist (HSCL). A self-report-symptom inventory. *Behavioural Science* **19**, 1–4.

Deter, H. C. & Herzog, W. (1994). Anorexia nervosa in a long-term perspective: results of the Heidelberg–Mannheim study. *Psychosomatic Medicine* **56**, 20–27.

Fairburn, C. G., Norman, P. A., Welch, S. L., O'Connor, M. E., Doll, H. A. & Peveler, R. C. (1995). A prospective study of outcome in bulimia nervosa and the long-term effects of three psychological treatments. *Archives of General Psychiatry* **52**, 304–312.

Fairburn, C. G., Welch, S. L., Doll, H. A., Davies, B. A. & O'Connor, M. E. (1997). Risk factors for bulimia nervosa. A community-based case-control study. *Archives of General Psychiatry* **54**, 509–517.

Fallon, B. A., Walsh, T., Sadik, C., Saoud, J. B. & Lukasik, V. (1991). Outcome and clinical course in in-patient bulimic women: a 2- to 9-year follow-up study. *Journal of Clinical Psychiatry* **52**, 272–278.

Fichter, M. M., Elton, M., Engel, K., Meyer, A. E., Mall, H. & Poustka, F. (1991). Structured Interview for Anorexia and Bulimia Nervosa (SIAB): development of a new instrument for the assessment of eating disorders. *International Journal of Eating Disorders* **10**, 571–592.

Fichter, M. M., Herpertz, S., Quadflieg, N. & Herpertz-Dahlmann, B. (1998). Structured Interview for Anorexic and Bulimic Disorders for DSM-IV and ICD-10: updated (third) revision. *International Journal of Eating Disorders* **24**, 227–249.

Fichter, M. M. & Quadflieg, N. (1995). Comparative studies on the course of eating disorders in adolescents and adults. Is age at onset a predictor of outcome? In *Eating Disorders in Adolescence. Anorexia and Bulimia Nervosa* (ed. H.-C. Steinhausen), pp. 301–337. De Gruyter: New York.

Fichter, M. M. & Quadflieg, N. (1997). Six-year course of bulimia nervosa. *International Journal of Eating Disorders* **22**, 361–384.

Fichter, M. M. & Quadflieg, N. (1999). *Strukturiertes Inventar für Anorektische und Bulimische Essstörungen* (*SIAB*). Hogrefe Verlag: Göttingen, Germany.

Fichter, M. M. & Quadflieg, N. (2000). 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). *European Archives of Psychiatry and Clinical Neuroscience* **250**, 175–185.

Fichter, M. M. & Quadflieg, N. (2001*a*). The Structured Interview for Anorexic and Bulimic disorders for DSM-IV and ICD-10 (SIAB-EX): reliability and validity. *European Psychiatry* **16**, 38–48.

Fichter, M. M. & Quadflieg, N. (2001*b*). Das Strukturierte Interview für Anorektische und Bulimische Ess-Störungen nach DSM-IV und ICD-10 zur Expertenbeurteilung (SIAB-EX) und dazugehöriger Fragebogen zur Selbsteinschätzung (SIAB-S). [The structured interview for anorexic and bulimic eating disturbances according to DSM-IV and ICD-10 for expert rating (SIAB-EX) and the corresponding questionnaire for self-rating (SIAB-S)]. *Verhaltenstherapie* **11**, 314–325.

Fichter, M. M., Quadflieg, N. & Rehm, J. (2003). Predicting the outcome of eating disorders using structural equation modelling. *International Journal of Eating Disorders* **34**, 292–313.

First, M. B., Gibbon, M. & Spitzer, R. L. (1996). *User's Guide for the SCID-I Structured Clinical Interview for DSM-IV* (*SCID-I Version 2.0, February 1996, Final Version*). Biometrics Research Department, New York State Psychiatric Institute: New York.

Franke, G. H. (2002). *SCL-90-R Symptom-Checkliste von L. R. Derogatis–Deutsche Version–(2.A)*. Beltz Test: Göttingen, Germany.

Friedman, W. J. (1993). Memory for the time of past events. *Psychological Bulletin* **113**, 44–66.

Garner, D. M., Olmstead, M. P. & Polivy, J. (1983). Development and validation of a multidimensional eating disorder inventory of anorexia nervosa and bulimia. *International Journal of Eating Disorders* **2**, 14–34.

Hebebrand, J., Himmelmann, G. W., Heseker, H., Schäfer, H. & Remschmidt, H. (1996). Use of percentiles for the body mass index in anorexia nervosa: diagnostic, epidemiological, and therapeutic considerations. *International Journal of Eating Disorders* **19**, 359–369.

Herzog, D. B., Dorer, D. J., Keel, P. K., Selwyn, S. E., Ekeblad, E. R., Flores, A. T., Greenwood, D. N., Burwell, R. A. & Keller, M. B. (1999). Recovery and relapse in anorexia and bulimia nervosa: a 7·5-year follow-up study. *Journal of the American Academy for Child and Adolescent Psychiatry* **38**, 829–837.

Herzog, D. B., Keller, M. B., Lavori, P. W. & Ott, I. L. (1988). Short-term prospective study of recovery in bulimia nervosa. *Psychiatry Research* **23**, 45–55.

Herzog, T., Hartmann, A., Sandholz, A. & Stammer, H. (1991). Prognostic factors in out-patient psychotherapy of bulimia. *Psychotherapy and Psychosomatics* **56**, 48–55.

Hsu, L. K. G. & Sobkiewicz, T. A. (1989). Bulimia nervosa: a four- to six-year follow-up study. *Psychological Medicine* **19**, 1035–1038.

Johnson-Sabine, E., Reiss, D. & Dayson, D. (1992). Bulimia nervosa: a follow-up study. *Psychological Medicine* **22**, 951–959.

Keel, P. K., Dorer, D. J., Eddy, K. T., Franko, D., Charatan, D. L. & Herzog, D. P. (2003). Predictors of mortality in eating disorders. *Archives of General Psychiatry* **60**, 179–183.

Keel, P. K. & Mitchell, J. E. (1997). Outcome in bulimia nervosa. *American Journal of Psychiatry* **154**, 313–321.

Keel, P. K., Mitchell, J. E., Miller, K. B., Davis, T. L. & Crow, S. J. (1999). Long-term outcome of bulimia nervosa. *Archives of General Psychiatry* **56**, 63–69.

Keel, P. K., Mitchell, J. E., Miller, K. B., Davis, T. L. & Crow, S. J. (2000). Social adjustment over 10 years following diagnosis with bulimia nervosa. *International Journal of Eating Disorders* **27**, 21–28.

Keller, M. B., Herzog, D. B., Lavori, P. W., Ott, I. L., Bradburn, I. S. & Mahoney, E. M. (1989). High rates of chronicity and rapidity of relapse patients with bulimia nervosa and depression. *Archives of General Psychiatry* **46**, 480–481.

Klesges, R. C., Eck, L. H. & Ray, J. W. (1995). Who underreports dietary intake in a dietary recall? Evidence from the second national Health and Nutrition Examination Survey. *Journal of Consulting and Clinical Psychology* **63**, 438–444.

Maddocks, S. E., Kaplan, A. S., Woodside, D. B., Langdon, L. & Piran, N. (1992). Two-year follow-up of bulimia nervosa: the importance of abstinence as the criterion of outcome. *International Journal of Eating Disorders* **12**, 133–141.

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.

Mitchell, J. E., Pyle, R. L., Hatsukami, D., Goff, G., Glotter, D. & Harper, J. (1989). A 2–5-year follow-up study of patients treated for bulimia. *International Journal of Eating Disorders* **8**, 157–165.

Patton, G. C. (1988). Mortality in eating disorders. *Psychological Medicine* **18**, 947–951.

Paulsen, A. S., Crowe, R. R., Nayes, R. & Pfohl, B. (1988). Reliability of the telephone interview in diagnosing anxiety disorders. *Archives of General Psychiatry* **45**, 62–63.

Quadflieg, N. & Fichter, M. M. (2003). The course and outcome of bulimia nervosa. *European Child & Adolescent Psychiatry* **12** (Suppl. 1), 99–109.

Ratnasuriya, R. H., Eisler, I., Szmukler, G. I. & Russell, G. F. M. (1991). Anorexia nervosa: outcome and prognostic factors after 20 years. *British Journal of Psychiatry* **158**, 495–502.

Reas, D. L., Williamson, D. A., Martin, C. K. & Zucker, N. L. (2000). Duration of illness predicts outcome for bulimia nervosa: a long-term follow-up study. *International Journal of Eating Disorders* **27**, 428–434.

Reich, W. & Earls, F. (1990). Interviewing adolescents by telephone: is it a useful methodological strategy? *Comprehensive Psychiatry* **31**, 211–215.

Russell, G. F. M. (1979). Bulimia nervosa: an ominous variant of anorexia nervosa. *Psychological Medicine* **9**, 429–449.

Steinhausen, H. C. (1999). Eating disorders. In *Risks and Outcome in Developmental Psychopathology* (ed. H.-C. Steinhausen and F. Verhulst), pp. 210–230. Oxford University Press: Oxford.

Steinhausen, H. C. (2002). The outcome of anorexia nervosa in the 20th century. *American Journal of Psychiatry* **159**, 1284–1293.

Swift, W. J., Kalin, N. H., Wamboldt, F. S., Kaslow, N. & Ritholz, M. (1985). Depression in bulimia at 2- to 5-year follow-up. *Psychiatry Research* **16**, 111–122.

Swift, W. J., Ritholz, M., Kalin, N. H. & Kaslow, N. (1987). A follow-up study of thirty hospitalized bulimics. *Psychosomatic Medicine* **49**, 45–55.

Tausig, J. E., Freeman, M. S. W. & Freeman, E. W. (1988). The next best thing to being there: conducting the clinical research interview by telephone. *American Journal of Orthopsychiatry* **58**, 418–427.

Theander, S. (1970). Anorexia nervosa. A psychiatric investigation of 94 female cases. *Acta Psychiatrica Scandinavica* (Suppl. 214).

Theander, S. (1985). Outcome and prognosis in anorexia nervosa and bulimia: some results of previous investigations compared with those of a Swedish long-term study. *Journal of Psychosomatic Research* **19**, 493–508.

Theander, S. (1996). Anorexia nervosa with an early onset: selection, gender, outcome, and results of a long-term follow-up study. *Journal of Youth and Adolescence* **25**, 419–430.

Vaz, F. J. (1998). Outcome of bulimia nervosa: prognostic indicators. *Journal of Psychosomatic Research* **45**, 391–400.

Wittchen, H. U., Wunderlich, U., Gruschwitz, S. & Zaudig, M. (1997). *Strukturiertes Klinisches Interview für DSM-IV* (*SKID*). Hogrefe: Göttingen, Germany.