Psychological Medicine, 2001, **31**, 881–890. Printed in the United Kingdom

Long-term outcome of anorexia nervosa in a prospective 21-year follow-up study

B. LÖWE,[1] S. ZIPFEL, C. BUCHHOLZ, Y. DUPONT, D. L. REAS AND W. HERZOG

From the Department of General Internal and Psychosomatic Medicine, University of Heidelberg, Medical Hospital, Heidelberg, Germany

ABSTRACT

Background. Given our poor understanding of the very long-term course of anorexia nervosa, many questions remain regarding the potential for recovery and relapse. The purpose of the present study was to investigate long-term outcome and prognosis in an anorexic sample 21 years after the initial treatment.

Method. A multidimensional and prospective design was used to assess outcome in 84 patients 9 years after a previous follow-up and 21 years after admission. Among the 70 living patients, the follow-up rate was 90%. Causes of death for the deceased patients were obtained through the attending physician. Predictors of a poor outcome at the 21-year follow-up were selected based on the results of a previous 12-year follow-up of these patients.

Results. Fifty-one per cent of the patients were found to be fully recovered at follow-up, 21% were partially recovered and 10% still met full diagnostic criteria for anorexia nervosa. Sixteen per cent were deceased, due to causes related to anorexia nervosa. The standardized mortality rate was 9·8. The three groups also showed significant differences in psychosocial outcome. A low body mass index and a greater severity of social and psychological problems were identified as predictors of a poor outcome.

Conclusions. Recovery is still possible for anorexic patients after a period of 21 years. On the other hand, patients can relapse, becoming symptomatic again despite previously achieving recovery status. Only a few patients classified as having a poor outcome were found to seek any form of treatment, therefore, it is recommended that these patients should be monitored regularly and offered treatment whenever possible.

INTRODUCTION

Anorexia nervosa (AN) is a complex illness known to be associated with a high risk of morbidity and mortality. Despite a good chance for recovery over the long-term, many patients will take a chronic or lethal course (Beumont *et al*. 1993; Sullivan *et al*. 1998; Walsh & Devlin, 1998; Becker *et al*. 1999; Fichter & Quadflieg, 1999). Estimates of mortality range from 0 to 22% and the most common causes of death are complications due to starvation or suicide (Nielsen *et al*. 1998; Crow *et al*. 1999; Emborg, 1999; Herzog *et al*. 2000). A recent review of 108 follow-up studies reported that full recovery can be expected in approximately 45% of anorexic patients, whereas 33% improve and 19·8% remain chronic (Steinhausen, 1999).

The course of anorexia nervosa is often marked by different periods of relapse, remission, and crossover into bulimia nervosa, characteristics sometimes missed by studies relying on a cross-sectional design. Although some have argued that recovery after 12 years is rare (Hsu, 1988), others have found a polarization effect, with both a high rate of full recovery and mortality after a 12-year period (Deter & Herzog, 1994). Studies including a follow-up of over 20 years have also documented a variable course, with an increasing mortality

[1] Address for correspondence: Dr. Bernd Löwe, Department of General Internal and Psychosomatic Medicine, University of Heidelberg, Medical Hospital, Bergheimer Strasse 58, D–69115 Heidelberg, Germany.

rate in particular (Theander, 1985; Ratnasuriya *et al.* 1991; Crisp *et al.* 1992). These findings argue for longer-term follow-ups that can assess the probability of remission after a decade or more of illness.

Due to the severe psychological, social and financial impact of anorexia nervosa and its sequelae, much interest exists in better delineating the psychosocial outcome of this disorder, as well as investigating factors that may be helpful in predicting outcome. It has been shown that a high percentage of anorexic patients suffer from additional psychiatric disorders at follow-up (Steinhausen, 1999). For example, one 10-year follow-up demonstrated high rates of affective disorders (29%), anxiety disorders (33·9%), and alcohol abuse (3·2%) (Halmi *et al.* 1991). Despite a great deal of attention from researchers, predicting the outcome of anorexia nervosa has been met with considerable difficulty and no single prognostic factor has been clearly and consistently demonstrated across outcome studies.

The present study describes the clinical course and outcome of 84 patients examined in a longitudinal design 21 years after first in-patient treatment. Results of the 12-year follow-up have been published previously in this journal (Herzog *et al.* 1997*a*). In an attempt to improve upon previous methodological shortcomings, we used again a multi-dimensional design, which allowed for the simultaneous assessment of eating disorder outcome, mortality, additional psychiatric diagnoses, psychosocial functioning and medical outcome. Psychological, social and medical variables, which are indicators for a poor prognosis during the long-term course, were identified. In an attempt to delineate better the individual course of anorexia, changes in recovery status between the 12-year and 21-year follow-ups were described.

METHOD

Subjects

A well-documented sample of female AN patients receiving in-patient treatment between 1971–1980 at the University Medical Hospital in Heidelberg were followed-up after an average (S.D.) of 21·3 (2·9) years following first admission. The original sample of 103 patients included 12 males and seven females with severe somatic disorders (three diabetes mellitus; one Crohn's disease; one tuberculosis; one hepatitis B; one acute intermittent porphyria). All males and the seven females with severe somatic illnesses were excluded to yield a representative and homogenous sample, consisting of 84 female anorexics with no severe somatic disorders. This patient sample was previously investigated after 3·6 and 11·7 years (Deter & Herzog, 1994; Herzog *et al.* 1997*a*, *b*). At initial assessment, all patients met Feighner diagnostic criteria for anorexia nervosa (Feighner *et al.* 1972), and on retrospective analysis, DSM-IV criteria (American Psychiatric Association, 1994; Herzog *et al.* 1997*b*). Thirty-six patients (42·9%) showed binge-eating/purging behaviour and 48 patients (57·1%) showed restricting behaviour. Mean body mass index (S.D.) was 13·3 (2·0) kg/m^2 (Table 1). Average length of hospital stay was 3 months. Treatment was conducted by an multidisciplinary and specialized eating disorder team using an integrated approach, including behavioural and psychodynamic components. Family therapy was provided, whenever possible.

After an average of 3·6 years (T1), 75·9% of the patient sample underwent a first follow-up assessment. In a second investigation (T2) after a mean follow-up period of 11·7 years, we were able to acquire information regarding the entire patient sample (Herzog *et al.* 1997*a*, *b*). For the current follow-up assessment, patients were contacted by sending a letter to their last known address. If a patient had moved, the address was obtained at the registration office, a databank which contains address information for German citizens. At follow-up, 16·7% ($N = 14$) of the patients were deceased and of the living patients, 90% ($N = 63$) completed a psychiatric interview, a physical evaluation and standardized psychological questionnaires. The remaining seven patients either refused to participate ($N = 3$) or did not answer our letters ($N = 4$). Due to the German registration office databank, we knew that these patients were still alive. At the previous 12-year follow-up, three of these seven patients had a good outcome, three had an intermediate outcome and one had a poor outcome. Informed consent was obtained from all patients participating in the study. The research team was blind to clinical records and results of earlier follow-up investigators. The majority of the patients (*N*

Table 1. *Patient sample at first admission* (*T*0). 12-*year follow-up* (*T*2) *and* 21-*year follow-up* (*T*3)

	T0	T2	T3
Years of initial/follow-up assessment	1971–80	1988–90	1997–98
Number (%) of surviving patients	84 (100)	75 (89·3)	70 (83·3)
Number (%) of surviving patients followed-up	—	75 (100)	63 (90.0)
Mean (S.D.) length of follow-up, years	—	11·7 (2·4)	21·3 (2·9)
Mean (S.D.) age, years	20·7 (6·0)	32·5 (6·1)	42·0 (6·5)
Mean (S.D.) BMI, kg/m^2	13·3 (2·0)	19·6 (3·3)	20·2 (3·1)

$= 41$; 65·1%) were assessed directly by two members of the research team, including a medical doctor (B. L.) and a clinical psychologist (C. B.) at either the hospital or at the patient's home, in the event they were unable to personally come to the hospital. Information regarding the remaining patients was obtained via a telephone interview, questionnaires, and, with informed consent, from their current physician ($N = 22$; 34·9%). A comparison of the patients who completed a direct assessment *versus* the patients who were indirectly followed-up revealed no significant differences between body mass index (BMI), age, or sociodemographic characteristics at initial assessment and follow-up. Additionally, there were no significant differences in outcome grouping at the 12-year follow-up. Information regarding the deceased patients was acquired through the attending physician. Sample characteristics across the three follow-up assessments (T0, T2, and T3) are illustrated in Table 1.

Measures

At follow-up, patients completed a German version of the Eating Disorders Longitudinal Interval Follow-up Evaluation (LIFE Eat II) (Kächele, 1999), which is a psychiatric interview based on the LIFE II (Keller *et al.* 1987). The inclusion of a section for eating disorder symptomatology on the LIFE II was originally done by Herzog *et al.* (1993). Inter-rater-reliability for the LIFE Eat II was conducted in a German multi-centre study on 256 formerly anorexic patients (Kächele, 1999; Metzger *et al.* 1999). Reliability coefficients were $r = 0{\cdot}79$ for anorexia nervosa (Psychiatric Status Rating Scale), $r = 0{\cdot}79$, for bulimia nervosa (Psychiatric Status Rating Scale), and $r = 0{\cdot}62$ for major depression. The clinical psychologist who performed the interviews in our follow-up study (CB) also participated in the multi-centre study. Supplementary to assessing eating disorder symptomatology and social and demographic information, this interview assessed additional psychiatric diagnoses. Psychiatric comorbidity was confirmed using the International Diagnostic Checklists for DSM-IV (IDCL) (Hiller *et al.* 1997).

Outcome was defined according to the Psychiatric Status Rating Scale for Anorexia Nervosa (PSR) (Herzog *et al.* 1993; Kächele, 1999), which is part of the LIFE Eat II and is based on DSM-IV diagnostic criteria. Good outcome was defined as a PSR-level of 1; intermediate outcome as a PSR-level of 2, 3, or 4 and poor outcome as PSR-level 5 or 6. Patients receiving a PSR-level of 5 or 6 by definition also met full DSM-IV criteria for AN. Patients who had died during the follow-up period due to causes related to AN were classified in the poor outcome group. Eating disorder symptomatology present during the 3 months prior to the assessment was used for classification purposes. For the 12-year follow-up investigation of this patient group (T2), outcome was defined by using the Morgan-Russell General Outcome Categories (Morgan & Russell, 1975), which bases outcome on weight and menstrual status. Although different measures were used to determine outcome at T2 and T3, a comparison of outcome at these two follow-up periods was still possible, as the core variables of eating disorder outcome, namely menstrual and weight status, are the central part of both measures.

Psychosocial functioning was assessed using the Global Assessment of Functioning (GAF) Scale (Endicott *et al.* 1976; American Psychiatric Association, 1994). Additionally, a thorough evaluation of medical and treatment history as well as a blood test were taken. The standardized psychological questionnaires included the

German versions of the Beck Depression Inventory (BDI) (Hautzinger *et al*. 1995), the Eating Disorder Inventory (EDI) (Thiel & Paul, 1988), and the Symptom Check List–Revised (SCL-90-R) (Francke, 1995). Body image was assessed with the German-language Body Image Questionnaire (FKB-20) (Clement & Löwe, 1996). This instrument has been found to have good validity and a stable factor structure, consisting of the following two factors: perception of body dynamics and negative evaluation of the body. To facilitate the assessment, all patients received the questionnaires via mail and completed them before the interview.

Statistical analyses

Standardized mortality rate was calculated according to the Subjects-years-method and the German death statistics (Deutsches Statistisches Bundesamt, 1970–1997; Armitage & Berry, 1996). Differences between outcome groups were analysed using one-way analyses of variance (ANOVA) according to the general linear model. Fischer's exact tests (2-tailed) were performed for categorical data. Paired *t* tests were used to compare body mass index and EDI scores across T2 and T3. To control for type 1 error, a Bonferroni-adjustment was done. Ordered logistic regressions (Hosmer & Lemeshow, 1989; SAS Institute, 1995) were conducted to identify predictors of outcome. The criterion variable for this analysis was outcome group at the 21-year follow-up (T3) and has three categories (good, intermediate, or poor). The poor outcome group included the living patients with AN and the patients who died due to AN. Predictor variables assessed at the 12-year follow-up (T2) were chosen based on findings from previous reports and included body mass index, severity of psychological symptoms, severity of social symptoms (ANSS) (Deter, 1992), the eight subscales of EDI and eight laboratory findings. Separate analyses were performed for each prognostic factor in order to obtain a respectable degree of accuracy considering the size of our patient sample. However, an analysis of the intercorrelations of the predictors at T2 was done before computing the ordered logistic regressions. These ordered logistic regressions complete the results of our analyses on the basis of initial variables (T0), which showed that a long duration of illness before first hospitalization, a low body-mass index, an inadequate weight gain during first hospitalization, the binge-eating/purging subtype and severe psychological or social problems (ANSS) predicted a poor outcome at T3 (Zipfel *et al*. 2000). All statistical analyses were carried out using the SAS System for Windows, Release 6.12 (SAS Institute, 1994*a*, *b*).

RESULTS

General outcome

At the 21-year follow-up, 14 (16·7%) of the 84 patients were deceased. Of these patients, 12 had died due to causes directly related to AN and had met full DSM-IV criteria for AN at the time of death (AN-binge-eating/purging type $N = 11$; AN-restricting type $N = 1$). Causes of death for the 12 patients dead due to AN included: infection (bronchial pneumonia and sepsis; $N = 4$), complications due to dehydration and electrolyte imbalance ($N = 3$) and suicide ($N = 2$). One patient died due to generalized peritonitis after small intestinal perforation. Two patients died in an extremely malnourished state, but the exact causes of death are unknown. Two additional patients (2·6%) died due to causes unrelated to AN (asthmatic attack and metastatic rectal carcinoma). In the period since the 12-year follow-up (Herzog *et al*. 1997*a*), six patients had died. The observed rate of death in our patient sample was 9·8 times greater than the expected mortality rate according to the German population (standardized mortality rate = 9·8).

Of the 63 living patients who were examined, eight met full DSM-IV criteria for AN at follow-up (PSR-level of 5 or 6). Similar to the patients deceased due to AN-related causes, most patients were classified as AN-binge-eating/purging type ($N = 7$) in comparison to AN-restricting type ($N = 1$). Although three patients showed severe eating disorder behaviours at follow-up, these patients did not meet full criteria for AN. Thus, they were classified as eating disorder not otherwise specified (EDNOS) (PSR-level of 4). For the remaining 52 patients, no eating disorder diagnoses was given (PSR-level of 1, 2 or 3).

When considering the 14 dead and 63 examined living patients together, the following pattern of outcome is revealed: 50·6% ($N = 39$) of the patients showed a good outcome and

https://doi.org/10.1017/S003329170100407X Published online by Cambridge University Press

Table 2. *Psychosocial outcome at follow-up*

	Outcome group								
	Good ($N = 39$)		Intermediate ($N = 16$)		Poor ($N = 8$)		Total ($N = 63$)		Group differences*
	N	(%)	*N*	(%)	*N*	(%)	*N*	(%)	*P*
Marital status									
Never married	4	(10·3)	4	(25·0)	3	(37·5)	11	(17·5)	
Divorced/separated/widowed	3	(7·7)	2	(12·2)	2	(25·0)	7	(11·1)	0·06
Married/living with significant other	32	(82·0)	10	(62·5)	3	(37·5)	45	(71·4)	
Living arrangements									
Alone	4	(10·3)	5	(31·3)	4	(50·0)	13	(20·6)	
Living with spouse/partner	30	(76·9)	6	(37·5)	2	(25·0)	38	(60·3)	< 0·001
Living with family members	5	(12·8)	5	(31·3)	2	(25·0)	12	(19·1)	
Children									
Children	29	(74·4)	12	(75·0)	2	(25·0)	43	(68·3)	0·03
No children	10	(25·6)	4	(25·0)	6	(75·0)	20	(31·7)	
Employment									
Able to work	32	(82·1)	11	(68·8)	2	(25·0)	45	(71·4)	< 0·001
Not able to work	7	(17·9)	5	(31·3)	6	(75·0)	18	(28·6)	
GAF Scale									
Mean (S.D.) GAF scores	73·7	(12·2)	66·6	(14·5)	39·4	(15·2)	67·2	(17·3)	< 0·001

* Fisher's Exact Test (2-tailed) for marital status, living arrangements, children, employment. One-way analysis of variance according to the general linear model for Global Assessment of Functioning (GAF) Scale.

achieved a full recovery at the time of follow-up (PSR = 1); 20·8% ($N = 16$) of the patients were partially recovered at follow-up, comprising the intermediate outcome group (PSR = 2, 3, or 4); 26% ($N = 8$ living; $N = 12$ dead) of the patients were classified as having a poor outcome (PSR = 5 or 6 dead due to AN); and, 2·6% ($N = 2$) of the patients died due to causes unrelated to AN.

All the patients in the poor outcome group, including the deceased patients, continued to meet full diagnostic criteria for AN at the time of death or at follow-up. Mean body mass index (S.D.) differed significantly between the three outcome groups in the expected way (fully recovered, 21·6 (2·3); partially recovered, 19·7 (2·1); not recovered, 15·3 (2·7); $F_{2,60} = 26{\cdot}1$, $P < 0{\cdot}001$).

Psychosocial outcome

Significant differences in terms of psychosocial functioning existed between the three outcome groups of living patients (Table 2). Specifically, a direct relationship between eating disorder outcome and psychosocial outcome existed, such that poor eating disorder outcome was related to poor psychosocial outcome. Patients with a good eating disorder outcome were found to have a better living situation, including living with someone, having more children, and being able to work. For the total patient sample, 60·3% were found to be living with a partner, 68·3% had at least one child and 71·4% were able to work. Mean level of functioning as measured by the Global Assessment of Functioning (GAF) Scale differed significantly between the outcome groups. The low score of the poor outcome group indicates the severity of overall impairment associated with chronic AN.

Additional psychiatric diagnoses

At follow-up, mood disorders were found to be the most common additional psychiatric disorders, with 17·5% of the living patient sample meeting DSM-IV criteria for depression. Anxiety disorders were diagnosed in 15·9% of the patients, including 6·4% meeting criteria for obsessive–compulsive disorder. Substance related disorders, were diagnosed in 11·1% of the patients and 6·4% met criteria for a personality disorder (Table 3). The percentages of total mood disorders and total substance related disorders differed significantly between the three outcome groups, with higher proportions of disorders in the poor and intermediate outcome groups, compared with the good outcome group. No patients met criteria for bulimia nervosa at follow-up. Since the last follow-up, two patients have attempted suicide. One of these patients,

https://doi.org/10.1017/S003329170100407X Published online by Cambridge University Press

Table 3. *Additional psychiatric diagnoses at follow-up*

	Outcome group								
	Good ($N = 39$)		Intermediate ($N = 16$)		Poor ($N = 8$)		Total ($N = 63$)		Group differences*
Psychiatric diagnosis (DSM-IV)	*N*	(%)	*N*	(%)	*N*	(%)	*N*	(%)	*P*
Total mood disorders	3	(7·7)	5	(31·3)	3	(37·5)	11	(17·5)	0·02
Major depressive disorder, recurrent (296.3x)	1	(2·6)	5	(31·3)	3	(37·5)	9	(14·3)	
Major depressive disorder, single episode (296.2x)	1	(2·6)	0	(0)	0	(0)	1	(1·6)	
Dysthymic disorder (300.4)	1	(2·6)	0	(0)	0	(0)	1	(1·6)	
Total anxiety disorders	4	(10·3)	3	(18·8)	3	(37·5)	10	(15·9)	0·13
Panic disorder without agoraphobia (300.01)	2	(5·1)	0	(0)	0	(0)	2	(3·2)	
Panic disorder with agoraphobia (300.21)	0	(0)	1	(6·3)	0	(0)	1	(1·6)	
Agoraphobia without history of panic disorder (300.22)	0	(0)	0	(0)	1	(12·5)	1	(1·6)	
Generalized anxiety disorder (300.02)	1	(2·6)	0	(0)	0	(0)	1	(1·6)	
Anxiety disorder not otherwise specified (300.00)	1	(2·6)	0	(0)	0	(0)	1	(1·6)	
Obsessive–compulsive disorder (300.3)	0	(0)	2	(12·5)	2	(25·0)	4	(6·4)	
Total substance-related disorders	2	(5·1)	1	(6·3)	4	(50·0)	7	(11·1)	< 0·001
Alcohol abuse (305.00)	1	(2·6)	0	(0)	3	(37·5)	4	(6·4)	
Cannabis abuse (305.20)	1	(2·6)	1	(6·3)	0	(0)	2	(3·2)	
Sedative dependence (304.10)	0	(0)	0	(0)	1	(12·5)	1	(1·6)	
Total schizophrenia and other psychotic disorders	1	(2·6)	0	(0)	0	(0)	1	(1·6)	†
Schizoaffective disorder (295.70)	1	(2·6)	0	(0)	0	(0)	1	(1·6)	
Personality disorders	1	(2·6)	2	(12·5)	1	(12·5)	4	(6·3)	†
Borderline personality disorder (301.83)	0	(0)	1	(6·3)	0	(0)	1	(1·6)	
Histrionic personality disorder (301.50)	0	(0)	1	(6·3)	0	(0)	1	(1·6)	
Avoidant personality disorder (301.82)	0	(0)	0	(0)	1	(12·5)	1	(1·6)	
Dependent personality disorder (301.6)	1	(2·6)	0	(0)	0	(0)	1	(1·6)	

* Fisher's Exact Test (2-tailed) for total mood disorders, total anxiety disorders, and total substance-related disorders.
† Group differences not computed due to small cell sizes.

who was dependent on benzodiazepine, met full criteria for AN and the other showed no symptoms of AN, but attempted suicide in a depressive crisis.

Results from the psychological questionnaires revealed significant differences in scores on most dimensions between the outcome groups. Specifically, significant differences existed in body image (FKB-20, EDI), eating behaviour and preoccupation with thoughts of food (EDI). Further, non-eating disorder related differences existed between groups, such as depression, anxiety and somatization (BDI, SCL-90-R). It is important to note that on average, the good outcome group obtained scores in the normative range.

In terms of additional treatment received for AN since T2 (12-year follow-up), four patients (6·4%) had received further in-patient treatment and 17 patients (27%) had received out-patient treatment. Of the patients receiving in-patient treatment, three still met full AN criteria, and one was classified as having an 'intermediate' outcome. Only one of the patients receiving out-patient treatment belonged to the poor outcome group (12·5%). In contrast, seven patients (23·1%) in the good outcome group and nine patients (43·8%) in the intermediate outcome group received out-patient treatment.

Changes in outcome between T2 and T3

A non-significant difference in mean BMI (S.D.) for the total group was found between the 12-year follow-up and 21-year follow-up (19·8 kg/m² (3·2); 20·2 kg/m² (3·1); respectively; $T_{2,60} = 0{\cdot}83$; $P = 0{\cdot}41$). Additionally, no significant differences existed on any of the scales of the Eating Disorder Inventory (EDI) between follow-up assessments.

Another picture emerges, however, upon the examination of individual outcome of the patients between the follow-up periods (Table 4). Of the 41 patients who participated in the present study that had demonstrated a good outcome at the 12-year follow-up, 30 maintained their good outcome status. Ten patients showed partial symptoms, and one patient fully relapsed, with a co-morbid drug abuse problem. Of the 21

https://doi.org/10.1017/S003329170100407X Published online by Cambridge University Press

Table 4. *Comparison of outcome at the* 12*-year and* 21*-year follow-ups*

	21-year follow-up					
	Good ($N = 39$)	Intermediate ($N = 16$)	Poor, not recovered ($N = 8$)	Poor, dead due to AN ($N = 12$)	Dead due to other causes ($N = 2$)	Drop out ($N = 7$)
12-year follow-up						
Good ($N = 44$)	30	10	1	—	—	3
Intermediate ($N = 21$)	9	5	2	—	2	3
Poor						
Not recovered, living ($N = 10$)	—	1	5	3	—	1
Dead due to AN ($N = 9$)	—	—	—	9	—	—
Dead due to other causes ($N = 0$)	—	—	—	—	—	—

Table 5. *Ordered logistic regressions for predicting outcome group at T*3 *on the basis of predictor variables at T*2

Predictor variables at T2	Odds ratio*	(95% CI)	*P*
Body Mass Index (kg/m²)	0·68	(0·55–0·84)	< 0·001
Severity of psychological symptoms (psychological ANSS†)	1·30	(1·16–1·47)	< 0·001
Severity of social problems (social ANSS†)	1·25	(1·10–1·42)	< 0·001
Eating Disorder Inventory (EDI†)			
Ineffectiveness	1·20	(1·07–1·35)	0·003
Perfectionism	1·18	(1·01–1·37)	0·042
Interpersonal distrust	1·21	(1·03–1·44)	0·023
Interoceptive awareness	1·16	(1·02–1·31)	0·021
Laboratory findings§			
Haemoglobin (mmol/l)	0·46	(0·23–0·91)	0·025
Alkaline phosphatase (U/l)	1·02	(1·01–1·04)	0·013

* The odds ratio indicates the odds of being classified as having a worse (> 1) or better (< 1) outcome for each one-unit increase in the predictor variables.

† ANSS refers to the Anorexia Nervosa Symptom Score (Deter, 1992). Scale scores are included in the analysis in steps of 5 points.

‡ The subscales 'Drive for Thinness', 'Bulimia', 'Body Dissatisfaction' and 'Maturity Fears' at T2 are not significant predictors of outcome at T3.

§ Serum creatinine, potassium, albumin, glutamic-pyruvic transaminase (GPT/ALT) and bilirubin, and blood leukocyte count at T2 are not significant predictors of outcome at T3.

patients with partial recovery at the earlier follow-up assessment, a bipolar trend was observed, with nine patients achieving a full recovery and two patients meeting full criteria for AN. Two other patients were deceased at follow-up, but the causes of death were unrelated to AN. The majority of the 10 patients that were classified as having a poor outcome at the 12-year follow-up, remained chronically ill after 21 years, with only one patient achieving a partial recovery. Three patients had died due to AN-related causes and five patients continued to meet full criteria for AN.

The ordered logistic regressions conducted on BMI, anorexia nervosa symptom scores (ANSS), the eight subscales of the EDI, and eight laboratory findings at T2 showed that during the long-term course of AN, a low BMI and severe psychological and social problems were the most important predictors of a poor outcome (Table 5). Anaemia (low haemoglobin) and an elevated level for alkaline phosphatase were the identified laboratory predictors of a poor outcome. Additionally, elevated scores on four out of the eight EDI-subscales were significant in predicting poor outcome. However, it is important to note that the EDI-subscales and the laboratory findings were not as strong in predicting outcome as the BMI and the psychological and social problems. The analysis of intercorrelations between the predictors revealed non-significant coefficients for 60% of the intercorrelations, whereas 40% of the intercorrelation coefficients were significantly different from zero (e.g. the highest intercorrelation was found between the EDI-subscales 'Ineffectiveness' and 'Interpersonal distrust'; $r = 0{\cdot}74$; $P < 0{\cdot}001$).

https://doi.org/10.1017/S003329170100407X Published online by Cambridge University Press

DISCUSSION

The present study examined the clinical course and outcome of 84 anorexia nervosa patients 21 years after first in-patient treatment and 9 years after a previous follow-up. In terms of eating disorder symptoms, it was found that more than half of the patients achieved a complete recovery. However, a significant number of patients took a chronic or lethal course. Specifically, one-quarter had either died due to AN-related causes or continued to meet full criteria, and approximately one-quarter of the patients exhibited partial AN symptoms. The low attrition rate (10%) and high rate of direct follow-up support the representativeness of these findings to adult patients treated in an in-patient setting. Outcome at the 12-year follow-up for the seven patients who did not participate in the current study was found to correspond well to the overall sample (Table 4). Therefore, there is no evidence to suggest that these patients had a less favourable course of illness.

Our results are more promising than those by Ratnasuriya *et al.* (1991), who investigated long-term outcome using a similar average length of follow-up (20·2 years). In their study, only 30% of the patients were categorized as having a good outcome. It is possible that differences in methodology or patient sample can explain the different outcome results. One unlikely explanation, however, is that our patient sample was less severe at admission, as the eating disorders unit at our medical hospital traditionally provides treatment to severely ill patients (mean BMI at T0 = 13·3 kg/m²; see Table 1). Also in comparison with other studies, our patient sample presented at first assessment with a particularly low body weight (Herzog *et al.* 1992). In contrast, our recovery rate was less than that reported by Theander (1985), who found a recovery rate of 76% after 24 years. This result might reflect differences in outcome criteria between the two studies. In terms of mortality, our crude mortality rate of 18·2% corresponds well to the long-term mortality rates reported by Ratnasuriya *et al.* (17·5%) and Theander (15·9%). However, it is possible that the exclusion of the patients with severe somatic disorders in our study may have resulted in a lower mortality rate than if we had included them in our analyses. The standardized mortality rate of 9·8 indicates the severity of AN regarding medical complications and mortality.

Defining outcome remains an important methodological consideration. In the present study, outcome was based on the PSR-scales for anorexia nervosa (Herzog *et al.* 1993; Kächele, 1999). Despite the difficulties associated with a direct comparison of outcome between the Morgan–Russell scales and the PSR-scales, we used the latter due to their ability to provide a finer classification of outcome groups, including an intermediate group. An analysis of outcome using the Morgan–Russell scales to classify patients yielded no significant change in group membership.

As part of our methodological design, we also examined additional psychiatric diagnoses and psychosocial outcome at follow-up. The comparison of the psychosocial measures of outcome revealed significant differences across outcome groups. Even after the small sample size of the groups was considered, significant differences existed. Patients with a good outcome were comparable to a normal sample in terms of psychosocial functioning, number of children and the ability to work. In contrast, patients with existing AN showed significant impairment in most areas of life. Additional psychiatric co-morbidity was highest in the poor outcome group, but clinically relevant symptoms were also found in the intermediate group. Similar to other studies of psychiatric co-morbidity (Halmi *et al.* 1991; Hsu *et al.* 1992; Braun *et al.* 1994; Rastam *et al.* 1995; Thornton & Russell, 1997), the most common diagnoses were affective disorders (17·5%), anxiety disorders (15·9%, including 6·4% obsessive–compulsive disorders) and substance-related disorders (11·1%). However, these results were lower than the percentages in the 10-year follow-up study by Halmi *et al.* (1991). This might be due to the longer follow-up period and the higher percentage of recovered patients in our study. It is also possible that additional psychiatric disorders are more likely to occur concomitantly with onset or in the first years after onset, with stabilization occurring over the long-term as the psychological and social situation of the patient improve. It is of interest, however, that not one of the patients met criteria for (normal weight) bulimia nervosa diagnosis at follow-up. However, bingeing and purging behaviours were

present in our sample, as evidenced by the high percentage of chronic anorexic patients and patients dead due to AN, which met the criteria for AN-binge-eating/purging type.

In terms of treatment, the patients of the intermediate outcome group were found to have received a high rate of treatment during the time between follow-up period. In contrast, the chronic group received a low rate of eating disorder treatment, especially out-patient treatment. Altogether, only 50% of the chronic anorexic patients had received treatment during the past 9 years. It should be noted, however, that it is unknown whether these patients refused further treatment or whether these patients did not have access to appropriate treatment. Because our study was not a treatment effectiveness study, we cannot determine whether recovery was directly attributable to the effect of psychotherapeutic interventions.

It is worth noting that the patients with the intermediate outcome at 12 years showed the most change in terms of outcome at 21 years, with half of the patients achieving full recovery. The finding that only one out of 10 chronically ill patients at 12 years improved (moving from not recovered to partially recovered status) is similar to the results of the long-term study by Theander (1985). Despite the encouraging finding that approximately 73% of the recovered patients at 12 years maintained their good outcome, a significant portion of recovered and partially recovered patients at 12 years became symptomatic again at 21 years. Altogether, a slight polarization effect in outcome was demonstrated, such that there were increases in the portions of both the good and the poor outcome groups.

Our attempt to identify predictors of outcome showed that a low body mass index and severe psychological and social problems were associated with a chronic course. As these same variables when measured at admission were shown to be significant predictors of a poor outcome (Zipfel *et al*. 2000), it is probable that they are valid outcome predictors during the entire course of anorexia nervosa. Since these variables are likely to change over time, they represent an important addition to the previously identified predictors of a chronic course (e.g. long duration of illness, inadequate weight gain, etc.), which are unchanging and likely to be obtained during the initial admission for treatment when the patient meets full criteria for anorexia nervosa. The laboratory findings that anaemia and elevated alkaline phosphatase are predictive of a poor outcome may reflect the malnutrition status of severely ill patients with its metabolic consequences for bone, marrow, and liver. Since we performed separate analyses for each prognostic factor, the intercorrelations between the prognostic factors were not considered in our results.

Conclusion

In summary, approximately one-half of our anorexic samples achieved full recovery from eating disorder symptoms, psychological symptoms and problems with social integration (partner, family, career) after 21 years. Conversely, there was also a risk of chronicity, as evidenced by a high mortality rate, poor psychosocial functioning, and the presence of additional psychiatric disorders. Our data suggest that even 12 years after admission, patients can either continue to improve or relapse. Body mass index and the severity of social and psychological problems are important indicators for the severity of the illness. The low rate of treatment received by the chronically ill patients may suggest they failed to receive a sufficient amount of treatment. This lack of treatment may be due to either the patient's failure to seek treatment or due to the inability to locate clinicians willing and able to provide appropriate treatment. As it is a realistic goal that a chronic eating disorder can be effectively treated or even prevented, it is recommended that treatment providers continue to monitor these chronically ill patients, providing treatment whenever possible.

We thank Beate Wild, Birgit Müller and Dr Timo Specht for their help with different phases of the study. Our sincere thanks to our former patients and their doctors, who made this work possible.

This study was supported by the German Ministry of Technology and Research (grant 701-629-3) and the medical faculty of the University of Heidelberg (project 161/97).

REFERENCES

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

Armitage, P. & Berry, G. (1996). *Statistical Methods in Medical Research*. Blackwell: Oxford.

Becker, A. E., Grinspoon, S. K., Klibanski, A. & Herzog, D. B. (1999). Eating disorders. *New England Journal of Medicine* **340**, 1092–1098.

Beumont, P. J., Russell, J. D. & Touyz, S. W. (1993). Treatment of anorexia nervosa. *Lancet* **341**, 1635–1640.

Braun, D. L., Sunday, S. R. & Halmi, K. A. (1994). Psychiatric comorbidity in patients with eating disorders. *Psychological Medicine* **24**, 859–867.

Clement, U. & Löwe, B. (1996). *Fragebogen zum Körperbild* (*FKB-20*) [*Body Image Questionnaire* (*BIQ-20*)]. Hogrefe: Göttingen, Germany.

Crisp, A. H., Callender, J. S., Halek, C. & Hsu, L. K. (1992). Long-term mortality in anorexia nervosa. A 20-year follow-up of the St George's and Aberdeen cohorts. *British Journal of Psychiatry* **161**, 104–107.

Crow, S., Praus, B. & Thuras, P. (1999). Mortality from eating disorders – a 5- to 10-year record linkage study. *International Journal of Eating Disorders* **26**, 97–101.

Deter, H. C. (1992). The anorexia nervosa symptom score: a multidimensional tool for evaluating the course of anorexia nervosa. In *The Course of Eating Disorders. Long-term Follow-up Studies of Anorexia and Bulimia Nervosa* (ed. W. Herzog, H. C. Deter and W. Vandereycken), pp. 40–52. Springer: New York.

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

Deutsches Statistisches Bundesamt (1970–1997). *Altersspezifische und Standardisierte Sterbeziffern* [*Age-specified and Standardized Mortality Rates*]. Deutsches Statistisches Bundesamt: Wiesbaden, Germany.

Emborg, C. (1999). Mortality and causes of death in eating disorders in Denmark 1970–1993: a case register study. *International Journal of Eating Disorders* **25**, 243–251.

Endicott, J., Spitzer, R. L., Fleiss, J. L. & Cohen, J. (1976). The global assessment scale. A procedure for measuring overall severity of psychiatric disturbance. *Archives of General Psychiatry* **33**, 766–771.

Feighner, J. P., Robins, E., Guze, S. B., Woodruff, R. A., Jr, Winokur, G. & Munoz, R. (1972). Diagnostic criteria for use in psychiatric research. *Archives of General Psychiatry* **26**, 57–63.

Fichter, M. M. & Quadflieg, N. (1999). Six-year course and outcome of anorexia nervosa. *International Journal of Eating Disorders* **26**, 359–385.

Francke, G. H. (1995). *SCL-90-R. Die Symptom-Checkliste von Derogatis. Deutsche Version* [*SCL-90-R. Symptom Check List-Revised. German Version*]. Beltz Test: Göttingen, Germany.

Halmi, K. A., Eckert, E., Marchi, P., Sampugnaro, V., Apple, R. & Cohen, J. (1991). Comorbidity of psychiatric diagnoses in anorexia nervosa. *Archives of General Psychiatry* **48**, 712–718.

Hautzinger, M., Bailer, M., Worall, H. & Keller, F. (1995). *Beck-Depressions-Inventar* (*BDI*) *Testhandbuch, 2. Auflage* [*Beck Depression Inventory* (*BDI*) *German Version*]. Huber: Bern.

Herzog, D. B., Sacks, N. R., Keller, M. B., Lavori, P. W., von Ranson, K. B. & Gray, H. M. (1993). Patterns and predictors of recovery in anorexia nervosa and bulimia nervosa. *Journal of the American Academy of Child and Adolescent Psychiatry* **32**, 835–842.

Herzog, D. B., Greenwood, D. N., Dorer, D. J., Flores, A. T., Ekeblad, E. R., Richards, A., Blais, M. A. & Keller, M. B. (2000). Mortality in eating disorders: a descriptive study. *International Journal of Eating Disorders* **28**, 20–26.

Herzog, W., Rathner, G. & Vandereycken, W. (1992). Long-term course of anorexia nervosa: a review of the literature. In *The Course of Eating Disorders. Long-term Follow-up Studies of Anorexia and Bulimia Nervosa* (ed. W. Herzog, H. C. Deter and W. Vandereycken), pp. 15–29. Springer: New York.

Herzog, W., Deter, H. C., Fiehn, W. & Petzold, E. (1997*a*). Medical findings and predictors of long-term physical outcome in anorexia nervosa: a prospective, 12-year follow-up study. *Psychological Medicine* **27**, 269–279.

Herzog, W., Schellberg, D. & Deter, H. C. (1997*b*). First recovery in anorexia nervosa patients in the long-term course: a discrete-time survival analysis. *Journal of Consulting and Clinical Psychology* **65**, 169–177.

Hiller, W., Zaudig, M. & Mombour, W. (1997). *Internationale Diagnosen Checklisten* (*IDCL*) *für DSM-IV* [*International Diagnostic Checklists* (*IDCL*) *for DSM-IV*]. Hogrefe: Göttingen.

Hosmer, D. W. & Lemeshow, S. (1989). *Applied Logistic Regression*. John Wiley: New York.

Hsu, L. K. (1988). The outcome of anorexia nervosa: a reappraisal. *Psychological Medicine* **18**, 807–812.

Hsu, L. K., Crisp, A. H. & Callender, J. S. (1992). Psychiatric diagnoses in recovered and unrecovered anorectics 22 years after onset of illness: a pilot study. *Comprehensive Psychiatry* **33**, 123–127.

Kächele, H. für die Studiengruppe MZ-ESS (1999). Eine multizentrische Studie zu Aufwand und Erfolg bei psychodynamischer Therapie von Eßstörungen [A multicentre study investigating the cost and effectiveness of a psychodynamically-oriented treatment of eating disorders – study design and initial results]. *Psychotherapie, Psychosomatik, Medizinische Psychologie* **49**, 100–108.

Keller, M. B., Lavori, P. W., Friedman, B., Nielsen, E., Endicott, J., McDonald, Scott, P. & Andreasen, N. C. (1987). The Longitudinal Interval Follow-up Evaluation. A comprehensive method for assessing outcome in prospective longitudinal studies. *Archives of General Psychiatry* **44**, 540–548.

Metzger, R., Krämer, B. & Kordy, H. (1999). Reliabilität des LIFE II [Reliability of LIFE II]. Unpublished manuscript. Center for Psychotherapy Research: Stuttgart.

Morgan, H. G. & Russell, G. F. (1975). Value of family background and clinical features as predictors of long-term outcome in anorexia nervosa: four-year follow-up study of 41 patients. *Psychological Medicine* **5**, 355–371.

Nielsen, S., Moller-Madsen, S., Isager, T., Jorgensen, J., Pagsberg, K. & Theander, S. (1998). Standardized mortality in eating disorders – a quantitative summary of previously published and new evidence. *Journal of Psychosomatic Research* **44**, 413–434.

Rastam, M., Gillberg, I. C. & Gillberg, C. (1995). Anorexia nervosa 6 years after onset: Part II. Comorbid psychiatric problems. *Comprehensive Psychiatry* **36**, 70–76.

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

SAS Institute (1994*a*). *SAS/STAT User's Guide, Version 6, 4th edn, Volume 1*. SAS Institute Inc: Cary, NC.

SAS Institute (1994*b*). *SAS/STAT User's Guide, Version 6, 4th edn, Volume 2*. SAS Institute Inc: Cary, NC.

SAS Institute (1995). *Logistic Regression. Examples using the SAS System, Version 6, 4th edn*. SAS Institute Inc: Cary, NC.

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

Sullivan, P. F., Bulik, C. M., Fear, J. L. & Pickering, A. (1998). Outcome of anorexia nervosa: a case–control study. *American Journal of Psychiatry* **155**, 939–946.

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 Psychiatric Research* **19**, 493–508.

Thiel, A. & Paul, T. (1988). Entwicklung einer deutschsprachigen Version des Eating-Disorder-Inventory (EDI) [Development of the German version of the Eating Disorder Inventory (EDI)]. *Zeitschrift für Differentielle und Diagnostische Psychologie* **9**, 267–278.

Thornton, C. & Russell, J. (1997). Obsessive–compulsive comorbidity in the dieting disorders. *International Journal of Eating Disorders* **21**, 83–87.

Walsh, B. T. & Devlin, M. J. (1998). Eating disorders: progress and problems. *Science* **280**, 1387–1390.

Zipfel, S., Löwe, B., Reas, D. L., Deter, H.-C. & Herzog, W. (2000). Long-term prognosis in anorexia nervosa: lessons from a 21-year follow-up study. *Lancet* **355**, 721–722.