The Course of Illness Following Inpatient Treatment of Adults with Longstanding Eating Disorders: A 5-Year Follow-Up

KariAnne R. Vrabel, CandPsych1*
Jan H. Rosenvinge, PhD2
Asle Hoffart, PhD1
Egil W. Martinsen, MD, PhD3
Øyvind Rø, MD, PhD1

ABSTRACT

Objective: The objectives of this article were to study the course and outcome of longstanding eating disorders (ED) 5 years after completing treatment and to identify subgroups of patients with different course and outcome.

Method: A total of 77 patients with a mean age of 30 years were assessed at the beginning and end of in-patient therapy and at 1-, 2- and 5-year follow-up, respectively.

Results: Of the 90% who participated in the 5-year follow-up, 46 patients (61%) had improved, and 30 (39%) did not meet diagnostic criteria for an ED. Cluster analysis identified a group of patients with no improvement over time.

Conclusion: Overall, the course is favourable, but a subgroup of patients with no improvement over time may need intensified treatment efforts.

Keywords: anorexia nervosa; bulimia nervosa; chronic eating disorder; follow-up; outcome; in-patient treatment

Introduction

Follow-up studies have shown that 30–75% of patients recover from anorexia nervosa (AN) and between 50 and 70% from bulimia nervosa (BN). This variability across studies is related to (1) patient characteristics like age and admission status, (2) the duration of follow-up, (3) dropout, attrition, and diagnostic crossovers, (4) methodology (i.e. questionnaire vs. interview, telephone vs. in-person, individual factors vs. global scores), and (5) outcome definitions, notably the definition of remission and recovery. To reduce this variability, several authors recommend specific guidelines for conducting follow-up studies. These focus on the use of prospective data collection with repeated follow-up measure points, standard diagnostic research criteria to define cases, a sufficient length of time to account for the full course of eating disorders (ED), as well as measuring multiple aspects of outcome status. Several authors have also highlighted the importance of standardized definition of outcome criteria, recovery or remission. Within the field of ED, the question of how to define outcome and recovery is nevertheless unresolved.

Most follow-up studies on ED have been conducted on adolescents or adults in their early twenties. More knowledge is needed on the course of illness for older adults with severe, long-standing ED as these patients may differ from younger patients with respect to a longer duration of illness, sociodemographic variables (i.e. age, marital status, social class), diagnostic and course variables (i.e. family history of ED, age at onset, diagnosis at onset, diagnostic crossovers) as well as with respect to work functioning, personality disorders, frequency of ED symptoms, and psychiatric comorbidity.

Many patients with AN and BN experience recurrent relapses after treatment and many do not respond to the first treatment trial. We have limited knowledge about the course and optimal treatment for these nonresponders. The same is the case for patients with comorbid symptom- or personality disorders, who may be difficult to treat yet frequently seen in clinical practice.

Diagnostic shifts between the different ED diagnoses over time are common. Moreover, the empirical support to the current diagnostic division of ED is mixed, and the division does not capture the...
diversity of disorders of eating experienced by clinicians and sufferers. In the literature there is an increasing emphasis on common psychopathological features in the various ED, and a transdiagnostic theory and treatment model has been proposed. In the present study we therefore have pooled and analyzed patients with AN, BN, and EDNOS in one sample. Following the logic of a transdiagnostic approach a cluster analytic strategy is then appropriate. Using this strategy several cross-sectional studies have revealed different subgroups, which resemble, but are not identical to existing diagnostic categories. Most follow-up studies describe the course and outcome of ED in terms of average mean group scores. This may mask a possible heterogeneity in course and outcome as subgroups of patients may show different courses regarding rate and stability of change. Here, cluster analysis may be more sensitive. The present study is, to our knowledge, the first to use cluster analysis to identify subgroups in terms of their course and outcome.

The aim of this study is to (1) report 5-year course and outcome of adults hospitalized for long-standing ED, who had comorbid disorders as well as a history of treatment failures and dropouts and (2) to identify ED subgroups with different course and outcome.

Method

Participants

From 1998 to 2000, 92 patients were admitted to a specialized ED-unit at a psychiatric center (Modum Bad, Norway). Following clinical evaluation, patients were allocated to specific inpatient treatment programs for AN and BN. The programs were symptom-focused, multi-component and lasted 22–23 weeks for AN and 15 weeks for BN. For EDNOS the treatment program lasted 22–23 weeks for subthreshold AN patients with normal weight and 15 weeks for BN with low symptom frequency others. The admission criteria were symptoms of AN, BN, and EDNOS that impaired daily functioning, inadequate responses to previous treatment and age above 18. Serious medical complication and body mass index (BMI) below 14 were exclusion criteria for admission to the hospital.

From the initial sample of 92, six patients discharged themselves shortly (<2 weeks) after admission. Three (4%) died during the observation time yet from causes unrelated to ED. No one died from suicide. Two (3%) patients could not be reached and four (5%) refused to participate at the 5-year follow-up, leaving us with a sample of 77, i.e. 90% out of the initial 86 patients. These 77 patients constitute the sample of the present study.

At admission mean age was 29.7 years (SD = 7.3) and the mean duration of illness was 13.3 years (SD = 7.1). This was the mean duration from the time when they displayed ED-symptoms. Prior to admission patients had received psychiatric treatment for a mean of 2.9 years (SD = 2.1). Among the patients 45 (58%) had received previous in-patient treatment. As for the diagnostic distribution 13 (17%) had AN, 39 (51%) had BN, and 25 (32%) had Eating Disorders Not Otherwise Specified (EDNOS) according to DSM-IV criteria. Of the 25 EDNOS patients 8 (10%) had subthresold AN with normal weight, and were allocated to the AN group, while 17 (23%) had BN with low symptom frequency and were allocated to the BN group. Furthermore, 57 (74%) had a personality disorder at admission, predominantly avoidant, obsessive-compulsive, and borderline personality disorder. Also, 41 patients (53.2%) reported a history of suicide attempts or self-mutilating behaviors. Table 1 provides detailed socio-demographic information related to diagnostic groups.

Written informed consent was obtained from all participants, and the study was approved by the Regional Committee for Medical Ethics.

Design

Patients were assessed at five points of time, i.e. admission and discharge from the unit, and at 1-, 2- and
5-year follow-up, respectively. In case of missing data the last observations were carried forward. This was considered to be a conservative procedure, as scores generally tended to drop over time. At the 5-year follow-up 35 (45%) patients were assessed in their homes or at offices outside the hospital, the remaining came to the hospital for assessment. The majority of 74 patients (96%) were assessed in a personal interview and the remaining three were interviewed by phone due to long distances.

**Assessments**

**Eating Disorders.** The Eating Disorder Examination interview (EDE) was used to assess ED psychopathology and generate ED diagnoses. The EDE consists of four subscales: “restraint,” “shape concern,” “weight concern,” and “eating concern.” A mean value is calculated on a 0–6 point scale. At the 5-year follow-up the first author conducted the interviews. The Cronbach's alpha for these subscales ranged from 0.91 to 0.97 during all measure points. All interviews were videotaped, and 20 tapes were randomly selected for blind rating by an experienced clinician. Inter-rater reliability was high, with intraclass coefficients (ICC (1,1)) of 0.99 for the mean EDE restraint, eating concern, shape concern, and global EDE and 0.98 for the mean EDE weight concern. There was complete agreement between the two raters on all the 20 interviews about absence or presence of objective binge episodes. For 18 of the EDE-interviews diagnostic ambiguity was resolved by consensus discussions.

Patients completed the Eating Disorder Inventory, EDI. The EDI consists of 64 questions related specifically to eating behaviors, body perception, as well as more general questions about attitudes, behaviors, and emotions. A sum score and eight subscores are calculated, in which the sum score of “weight phobia,” “bulimia,” and “body dissatisfaction” comprises the symptom index, and the sum score of “ineffectiveness,” “perfectionism,” “interpersonal distrust,” “interoceptive awareness,” and “maturity fears” represents the personality index. Cronbach's alpha for these subscales ranged from 0.95 to 0.97 during all measure points.

**General Psychopathology and Interpersonal Problems.** General psychopathology was assessed using the Symptom Check List-25 Revised, SCL-25-R and Inventory of Interpersonal Problems 64, IIP-64.

The SCL-25-R consists of 25 items and derived from the SCL-90. A mean global severity index (GSI) is calculated on a 0–4 scale range. Cronbach's alpha for these subscales ranged from 0.86 to 0.96 during all measure points. We also used a suicidal ideation index (SII) representing the sum score of five SCL-90-items. Cronbach's alpha of these items ranged from 0.80 to 0.90.

The IIP consist of 64 questions with eight subscales assessing interpersonal problems. A mean value on a 0–4 scale is calculated for the total scale and for subscales. Cronbach's alpha values ranged from 0.92 to 0.96.

**Outcome Criteria**

There is no general consensus in the literature with respect to defining outcome in ED. The present study used outcome criteria modified from Dare et al. Full recovery was defined as BMI > 18.5 with no binges or weight compensating behaviors for 12 weeks prior to assessment and/or a mean value >3 at the EDE weight or shape concern subscales, respectively. In contrast to Dare et al. regular menstruation was not included as a criterion of full recovery, as there may be many reason for irregular menstruation unrelated to ED. Moderate recovery was defined as BMI > 17.5, and/or binges and/or weight compensating behaviors like vomiting, laxative and/or diuretic misuse and/or compulsive physical activity less than weekly. Poor recovery was defined as BMI values from 15.0 to 17.5 or weekly binges or weight compensating behaviors as stated above. No recovery was defined as BMI values <15.0, and/or daily binges, and/or daily weight compensating behaviors. To qualify for poor or no recovery, patients also had to satisfy the DSM-IV criteria for an ED. Improvement was defined as a shift to a more favorable recovery status, or in other words to a less severe outcome category.

**Statistical Analyses**

Data were analyzed by SPSS version 13.0 using multivariate measures analyses of variance (MANOVA) and bivariate analyses like t-test, the Wilcoxon Signed Rank Test or the McNemar test. The level of significance was set to <0.05 if not otherwise stated. Interrater reliability was calculated as intraclass correlation coefficients (ICC). To explore the variety of change patterns and subgroups a K-mean cluster analysis was conducted with Euclidian distance measures and running means in the iterations based on EDE-scores. No a priori hypothesis was established for initiating clusters. Successive analyses with increasing numbers of clusters were performed. The criteria for the final solution were based on the value showing the F-ratios for the sums of mean-squares between clusters, the F-ratios for the sums of mean-squares within clusters, the number of patients should be of a certain size in each clusters and the course and outcome of each cluster should be informative from a clinical point of view. This method of partitioning may yield profiles of change patterns. Two limitations in the use of cluster analysis are often cited, i.e. when one fails to provide comparable measures in order to evaluate the validity and utility of the cluster typologies, and the fact that the cluster solutions often are overly dependent on the clustering algorithm and the sample used to generate the typology. This may reduce the predictive validity and clinical utility of the cluster...
TABLE 2. Scores M and (SD) of the difference on EDE, EDI, SCL-25, SII and IIP at admission (adm), discharge (dis) and follow-up for the 77 patients

<table>
<thead>
<tr>
<th>Measure</th>
<th>Adm</th>
<th>Dis</th>
<th>1 Year</th>
<th>2 Years</th>
<th>5 Years</th>
<th>Adm/dis/1/2/5</th>
<th>F-Value</th>
<th>Adm/dis</th>
<th>Dis/1</th>
<th>1/2</th>
<th>2/5</th>
<th>F-Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>EDE</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Global</td>
<td>4.2</td>
<td>3.1</td>
<td>3.2</td>
<td>2.9</td>
<td>2.4</td>
<td>2.8</td>
<td>28.6***</td>
<td>53.9***</td>
<td>0.6</td>
<td>3.8*</td>
<td>9.8***</td>
<td></td>
</tr>
<tr>
<td>Diet. reastment</td>
<td>4.1</td>
<td>2.8</td>
<td>3.0</td>
<td>2.8</td>
<td>2.3</td>
<td>2.1</td>
<td>17.5***</td>
<td>35.1***</td>
<td>1.0</td>
<td>1.7</td>
<td>5.2**</td>
<td></td>
</tr>
<tr>
<td>Eating concern</td>
<td>4.0</td>
<td>2.5</td>
<td>2.5</td>
<td>2.0</td>
<td>1.7</td>
<td>1.7</td>
<td>45.0***</td>
<td>63.3***</td>
<td>0.1</td>
<td>11.3***</td>
<td>3.4</td>
<td></td>
</tr>
<tr>
<td>Shape concern</td>
<td>4.5</td>
<td>3.8</td>
<td>3.8</td>
<td>3.7</td>
<td>3.1</td>
<td>2.7</td>
<td>15.6***</td>
<td>19.5***</td>
<td>0</td>
<td>1.5</td>
<td>13.0***</td>
<td></td>
</tr>
<tr>
<td>Weight concern</td>
<td>4.1</td>
<td>3.1</td>
<td>3.3</td>
<td>3.2</td>
<td>2.7</td>
<td>2.0</td>
<td>13.4***</td>
<td>33.0***</td>
<td>1.5</td>
<td>0.3</td>
<td>9.9***</td>
<td></td>
</tr>
<tr>
<td>EDI</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sum score</td>
<td>90.0</td>
<td>68.7</td>
<td>64.8</td>
<td>58.6</td>
<td>52.5</td>
<td>43.1</td>
<td>41.3***</td>
<td>40.8***</td>
<td>2.1</td>
<td>6.3**</td>
<td>5.7*</td>
<td></td>
</tr>
<tr>
<td>Drive for thinness</td>
<td>14.1</td>
<td>11.2</td>
<td>10.5</td>
<td>9.4</td>
<td>8.2</td>
<td>8.6</td>
<td>19.0***</td>
<td>16.2***</td>
<td>1.1</td>
<td>5.5*</td>
<td>3.2</td>
<td></td>
</tr>
<tr>
<td>Bulimia</td>
<td>9.0</td>
<td>4.2</td>
<td>4.6</td>
<td>3.4</td>
<td>2.9</td>
<td>2.4</td>
<td>39.6***</td>
<td>58.9***</td>
<td>1.1</td>
<td>8.6**</td>
<td>1.3</td>
<td></td>
</tr>
<tr>
<td>Body dissatisfaction</td>
<td>17.8</td>
<td>15.4</td>
<td>14.6</td>
<td>14.3</td>
<td>13.0</td>
<td>10.3</td>
<td>10.3***</td>
<td>9.9**</td>
<td>1.7</td>
<td>0.2</td>
<td>2.8</td>
<td></td>
</tr>
<tr>
<td>Ineffectiveness</td>
<td>14.5</td>
<td>11.4</td>
<td>11.0</td>
<td>10.3</td>
<td>8.7</td>
<td>8.2</td>
<td>15.9***</td>
<td>16.2***</td>
<td>0.5</td>
<td>0.9</td>
<td>6.1***</td>
<td></td>
</tr>
<tr>
<td>Perfectionism</td>
<td>7.6</td>
<td>6.1</td>
<td>5.6</td>
<td>5.3</td>
<td>5.4</td>
<td>4.6</td>
<td>14.0***</td>
<td>24.1***</td>
<td>2.1</td>
<td>1.0</td>
<td>0.1</td>
<td></td>
</tr>
<tr>
<td>Interpersonal distrust</td>
<td>6.5</td>
<td>5.9</td>
<td>5.0</td>
<td>4.5</td>
<td>4.5</td>
<td>4.7</td>
<td>8.4***</td>
<td>4.2*</td>
<td>2.5</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>Interoceptive awareness</td>
<td>14.9</td>
<td>10.4</td>
<td>9.8</td>
<td>7.7</td>
<td>6.1</td>
<td>6.4</td>
<td>42.5***</td>
<td>34.0***</td>
<td>0.7</td>
<td>9.7**</td>
<td>9.8**</td>
<td></td>
</tr>
<tr>
<td>Maturity fears</td>
<td>5.7</td>
<td>4.0</td>
<td>4.1</td>
<td>3.8</td>
<td>3.7</td>
<td>3.2</td>
<td>10.4***</td>
<td>16.4***</td>
<td>0.7</td>
<td>0.1</td>
<td>2.0</td>
<td></td>
</tr>
<tr>
<td>SCL-25</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Global severity index</td>
<td>2.2</td>
<td>1.7</td>
<td>1.7</td>
<td>1.6</td>
<td>1.4</td>
<td>1.4</td>
<td>14.9***</td>
<td>20.8***</td>
<td>0.6</td>
<td>1.5</td>
<td>6.6**</td>
<td></td>
</tr>
<tr>
<td>Suicidal ideation index</td>
<td>2.2</td>
<td>1.8</td>
<td>1.8</td>
<td>1.7</td>
<td>1.3</td>
<td>1.3</td>
<td>8.4***</td>
<td>15.0***</td>
<td>1.0</td>
<td>0.1</td>
<td>1.0</td>
<td></td>
</tr>
<tr>
<td>IIP</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean score</td>
<td>1.8</td>
<td>1.5</td>
<td>1.5</td>
<td>1.3</td>
<td>1.3</td>
<td>1.2</td>
<td>11.5***</td>
<td>17.0***</td>
<td>2.7</td>
<td>5.1*</td>
<td>1.0</td>
<td></td>
</tr>
</tbody>
</table>

Note: EDE, eating disorder examination; EDI, eating disorder inventory; SCL, symptom check list; SII, suicidal ideation index; IIP, inventory of interpersonal problem.

* N = 58.
* * p < 0.05.
** ** p < 0.01.
*** *** p < 0.001.

solutions. Strategies to overcome these limitations are described below.

Results

Clinical Status at the 5-Year Follow-Up

Repeated measures MANOVA for global EDE, total EDI score, and GSI revealed a significant effect of time ($F(12, 65) = 10.78, p < 0.001$) across admission, post-treatment and 1-, 2- and 5-year follow-up (Table 2). The significant effect of time was replicated for subscales on EDE ($F(16, 61) = 10.19, p < 0.001$), EDI ($F(32, 45) = 6.74, p < 0.001$), GSI ($F(4, 73) = 14.92, p < 0.001$), and IIP ($F(32, 25) = 1.96, p < 0.05$).

For the whole series, there was a 77% reduction of binging, 39% of purging, and 40% of laxative misuse from admission to the 5-year follow-up (Table 3). By excluding an extreme case with a total of 700 self-induced vomiting episodes last 28 days at 5-year follow-up, the reduction of purging increased to 60%. A total of 19 had a BMI < 17.5 at admission. The mean value of BMI for this group was 15.6 at admission, 16.3 at discharge, 16.9 at 1-year, 18.9 at 2-year, and 19.4 at 5-year follow-up. Repeated measures ANOVA on weight gain in this sub sample showed significant time effect ($F(4, 15) = 4.6, p < 0.01$) with an average weight increase of 11.0 kg (SD = 11.0).

The number of patients meeting the criteria for BN and AN decreased substantially over the 5 years. At the 5-year follow-up 30 patients (39%) received no DSM-IV ED diagnosis. The number of patients with EDNOS was relatively stable between admission and 5-year, with a slight increase during discharge and 2-year follow-up. Only one patient had no ED diagnosis at discharge and any point of follow-up, 6 (8%) patients had no ED diagnosis at any point of follow-up while 41 (53%) had an ED diagnosis at all measure points.

According to the recovery criteria categories 26 (34%) were completely recovered, 17 (22%) moderately recovered, and 34 (44%) had poor or no recovery. Indeed, there was a pattern of increase in the number of fully recovered from discharge (Table 4). According to the criteria for improvement (moving to a more favorable outcome category), 46 (60%) had improved from admission to the 5-year follow-up. The difference in number of patients belonging to the different outcome categories was only significant from 2- to 5-year (McNemar: $\chi^2 = 21.57, p < 0.001$). The largest contribution of this difference was an increase in the number of recovered patients (12 at 2 year to 26 at 5 year) and a decrease...
Cluster Analysis

The procedure of partitioning into subgroups, according to the K-mean cluster analysis of the EDE scores across assessments, was done according to the cluster solution, yielding the most clinical meaningful solution across the four clusters. The first step in a validation strategy the EDE-scores across assessments were analyzed based on the EDE three-cluster solution, and the same pattern was found. Testing four clusters gave a partitioning consistent with the three-cluster solution, except that the late responders in the four-cluster solution were split into two. The nonresponders remained unchanged when increasing the numbers of clusters. No cluster emerged that involved a relapser. We performed the analysis both with and without using running means, resulting in the same clusters. In addition, the analysis both with and without using running means, resulting in the same clusters. In addition.

TABLE 3. Bulimic symptoms and BMI at admission (adm), discharge (dis), and follow-up for the 77 patients

<table>
<thead>
<tr>
<th>Outcome Measures</th>
<th>Adm</th>
<th>Dis</th>
<th>1 year</th>
<th>2 years</th>
<th>5 years</th>
<th>Pairwise Comparison</th>
</tr>
</thead>
<tbody>
<tr>
<td>No. of patients with objective bulimic episodes last 12 weeks (%)</td>
<td>53 (69)</td>
<td>41 (53)</td>
<td>45 (58)</td>
<td>35 (45)</td>
<td>23 (30)</td>
<td>A vs. D***</td>
</tr>
<tr>
<td>No. of objective bulimic episodes last 28 days (median)</td>
<td>27.1 ± 37.7</td>
<td>8.2 ± 14.7</td>
<td>10.4 ± 17.5</td>
<td>7.5 ± 15.8</td>
<td>6.2 ± 13.0</td>
<td>1 vs. 2**</td>
</tr>
<tr>
<td>No. of self-induced vomiting episodes last 12 weeks (%)</td>
<td>53 (69)</td>
<td>48 (62)</td>
<td>54 (70)</td>
<td>42 (55)</td>
<td>36 (47)</td>
<td>A vs. D***</td>
</tr>
<tr>
<td>No. of self-induced vomiting episodes last 28 days (median)</td>
<td>43.3 ± 60.1</td>
<td>12.3 ± 182.3</td>
<td>20.8 ± 46.1</td>
<td>13.4 ± 24.4</td>
<td>26.4 ± 87.8</td>
<td>0.001</td>
</tr>
<tr>
<td>No. of patients with laxative misuse episodes last 12 weeks (%)</td>
<td>27 (35)</td>
<td>10 (13)</td>
<td>15 (19)</td>
<td>15 (19)</td>
<td>18 (23)</td>
<td></td>
</tr>
<tr>
<td>No. of laxative misuse episodes last 28 days (median)</td>
<td>11.4 ± 27.5</td>
<td>2.6 ± 9.3</td>
<td>5.0 ± 16.6</td>
<td>4.3 ± 14.1</td>
<td>6.9 ± 23.1</td>
<td>A vs. D***</td>
</tr>
<tr>
<td>Mean BMI</td>
<td>20.2 ± 4.7</td>
<td>20.1 ± 4.2</td>
<td>20.5 ± 4.4</td>
<td>21.1 ± 4.7</td>
<td>21.6 ± 4.8</td>
<td></td>
</tr>
</tbody>
</table>

* p < 0.05.  
** p < 0.01.  
*** p < 0.001.  
* For each bulimic symptom, four pairwise comparisons (A vs. D, D vs. 1, 1 vs. 2 and 2 vs. 5) were calculated; those not significant are omitted in the column on the right.

TABLE 4. Distribution (N and %) of eating disorder diagnoses and recovery status at admission, discharge, and follow-up for the 77 patients

<table>
<thead>
<tr>
<th>Diagnosis</th>
<th>Admission</th>
<th>Discharge</th>
<th>1 year</th>
<th>2 years</th>
<th>5 years</th>
</tr>
</thead>
<tbody>
<tr>
<td>Anorexia nervosa</td>
<td>13 (17)</td>
<td>13 (17)</td>
<td>10 (13)</td>
<td>8 (10)</td>
<td>6 (8)</td>
</tr>
<tr>
<td>Bulimia nervosa</td>
<td>39 (51)</td>
<td>14 (18)</td>
<td>22 (29)</td>
<td>12 (16)</td>
<td>14 (18)</td>
</tr>
<tr>
<td>EDNOS</td>
<td>25 (32)</td>
<td>44 (57)</td>
<td>35 (45)</td>
<td>42 (53)</td>
<td>27 (35)</td>
</tr>
<tr>
<td>No ED</td>
<td>6 (8)</td>
<td>10 (13)</td>
<td>15 (20)</td>
<td>10 (13)</td>
<td>15 (20)</td>
</tr>
</tbody>
</table>

Recovery status

- Full recovery — 3 (4) 8 (10) 12 (16) 26 (34) 26 (34)
- Moderate — 5 (6) 21 (27) 15 (19) 21 (27) 17 (22) 17 (22)
- Poor — 21 (27) 33 (43) 25 (33) 25 (33) 11 (14) 11 (14)
- No recovery — 51 (66) 20 (26) 29 (38) 19 (25) 23 (30) 23 (30)

The difference in number of patients across ED diagnoses who improved at 5-year follow-up was not significant (v^2 = 0.39, p = 0.86). About 87% of the patients had some form of comorbid Axis-I disorders at the 5-year follow-up. During the whole 5-year period all patients except one had at some time received psychiatric treatment and 37 patients (48%) had been hospitalized for treatment of their ED. Eighteen (23%) reported at least one episode of suicide attempt and 37 (48%) at least one episode of self-mutilation. The number of patients who were married or cohabitant increased from 28 (36%) at discharge to 36 (47%) at 5-year follow-up. The number of patients fully employed or studying increased from 20 (26%) at discharge to 34 (44%) at 5-year follow-up.
we performed a boot-strapping procedure by ordering the data to be randomly selected (80%), and by rerunning the analyses, similar partitioning as in the original version occurred 8 out of 10 times. Hence, the validation strategies made well-known cluster analysis limitations less likely to affect the outcome findings. Diagnoses at admission for the nonresponder group were AN 33% (N = 7), BN 38% (N = 8), and EDNOS 29% (N = 6). The “nonresponder” group had a mean age of 30.4 (SD = 7.1) duration of illness 13.5 years (SD = 6.2), duration of treatment 3.2 years (SD = 2.2). About 67% (N = 14) of the nonresponders reported self-mutilating and 86% (N = 18) were neither working nor studying.

Conclusion
The history of previous unsuccessful treatment, long duration of illness, and high age at inclusion of the patients in this study picture a patient series of adults with ED usually associated with an unfavorable outcome. Apart from a subset of patients with less or no improvement the present study on 5-years follow-up shows the contrary, i.e. that the course of ED for almost three of four patients was rather favorable. Thus, a highly significant improvement occurred during in-patient treatment and a steady improvement during the follow-up period. Also, no one died from complications of ED or suicide.

There are few previous studies of comparable patient series that merit comparisons with our findings. Dare et al. studied outpatients with a mean age of 26.3 years and AN for a mean duration of 6.3 years, where almost 80% of the patients had received previous ED-treatment. Only 13% recovery was found at the 1-year follow-up. Pike et al. studied AN outpatients with a mean age of 25.2 years with a long duration of illness and previous hospitalization. At their 1-year follow-up only 17% were fully recovered. In a 7.5 years follow-up study of adults with AN treated in ED specialist services the mean age at follow-up was 35.5. One-third of the patients had recovered or had only residual AN.

Considering these findings the results from the present study even more present an optimistic and positive picture of the outcome. Even though our patients were severely ill at admission, their pattern of improvement was rather equivalent to that of ED patients with less severity of illness. For instance, studies by Fichter and coworkers of AN and BN patients showed substantial improvement during inpatient stay, a moderate decline in symptom level during the first 2 years post treatment and further improvement until 12-year follow-up. At 12-year follow-up 28% of the AN patient had a good outcome and 70% of the BN patients showed no major DSM-IV ED. In their studies, the duration of the inpatient treatment was shorter and the mean age was 25 years compared with 30 years in our study. Recovery both in our and Fichter’s study occurred during the whole observation period. Reports on long-term course of AN as well as reviews of outcome studies of AN indicate that there is an increase in remission rates with longer follow-up periods. Moreover, none of our patients died from ED complications or suicide despite a long duration of illness and high rate of previous suicide attempts. This finding is different from other studies. Even in short-time follow-up studies mortality figures usually varies between 2 and 8%, and even higher (15%) in long-term studies of AN. Thus, zero mortality from ED or suicide is
noteworthy as some ED-related mortality would have been expected considering the long follow-up period and high response rate in our study. One may argue that mortality in BN is rather uncommon, yet Fichter and Quadflieg\textsuperscript{10} reported in their 6-year follow-up a 1.2% mortality in their BN series comprising patients with a less severe history and duration of illness than in the present study. On the other hand, an optimistic picture of the outcome is a matter of perspective, and should be balanced against the fact that on average, 87% of the patients had some form of comorbid Axis-I disorders 5 years after admission. A separate paper is underway reporting on this comorbidity in more detail. Also one must consider the fact that only 34% were recovered despite an initial hospitalization of almost 6 months and, that 48% had further hospital treatment for their ED in the 5-year period. Still, one may argue that the patients fared rather well considering their history of ED and their history of past treatment failures. Moreover, one may argue that the high number of EDNOS patients at the onset of the study may confound the outcome findings as one may suspect that patients with EDNOS were already in an improved state at the time of their admission. However, this is hardly likely as the mean EDE scores at admission were almost identical for AN, BN, and EDNOS (data not shown) indicating a similar level of severity. The extensive inpatient treatment program may have contributed to this beneficial outcome. While our data merits no conclusions on this issue, the results may encourage future research exploring the effects of hospital treatment.

The three major change patterns were identified: early, late, and nonresponders. These correspond well with clinical observations and show that cluster analysis may capture basic and varied longitudinal patterns of course and outcome. The early and late responders, comprising close to two-thirds of the sample, showed improvement during the observation time. One should note, however, that at both 2- and 5-year follow-up the mean values of global EDE are within the normal range only for the early responders group (EDE $< 0.80$).\textsuperscript{59} At the 5-year follow-up the late responders group had not reached an average level of EDE within normal range. Patients in this group still had residual symptoms at 5-year follow-up. This finding corroborate results of several other follow-up studies demonstrating that for a considerable number of patients, normal attitudes towards food, weight, and body shape seem difficult to regain.\textsuperscript{2,8}

We identified a group of nonresponders with a high level of symptoms throughout, and hence they may stand the risk of a chronic course. Future studies are under way to determine what characterizes this subgroup. We do not know what kind of treatment approach that optimises outcome in these patients, but it may be the case that they can benefit from treatment approaches more focused on how to increase quality of life with a chronic illness instead of aiming at full recovery.

Strengths of this study include (1) a high response rate, (2) personal interviews at all assessments, (3) a high interrater reliability, and (4) that the design and procedures follow the methodological recommendations in the literature.\textsuperscript{14–16} A limitation is that follow-up assessments were not performed by independent assessors, but by expert clinicians who to some extent had been involved in the treatment. Some bias may thus be introduced. However, this can be an asset more than a flaw. Thus, personal knowledge of the patient may increase data validity and reliability. As ED patients tend to withhold or flavor personal information, disclosure may be easier meeting with an interviewer they know, and who knows the patient. Because patients were notified about who would interview then, this may also have contributed to the high response rate. Some data of our study were collected retrospectively (e.g. previous history of ED or mental disorder, course of illness over the time between assessments). In view of the vast literature on memory and possible biases of retrospective assessments,\textsuperscript{50} this may represent a validity and reliability problem, yet counterbalanced by the use of multiple assessment points. Because of the symptom severity and comorbidity one may argue that the patients in the study are selected, and preventing the results to be generalized to patients with ED in general as the present series represents an atypical subset of patients. However, all diagnostic categories were represented, reflecting the diagnostic distribution commonly seen in clinical practice. Thus, we are rather confident that the results may be valid for adult patients with ED. Also, a clinical picture of symptom severity and comorbidity seems rather common in ordinary clinical settings.\textsuperscript{61} In a similar way Palmer\textsuperscript{62} argue for a general trend that people who are admitted for ED even in general clinical settings present with complicated problems that are difficult to treat. A final word of caution refers to the unequal time intervals between the five measure points that make change patterns difficult to interpret.

To conclude, this study shows a generally positive outcome of ED despite a dismal history of ED and of treatment failures. Still, we identified a subgroup
References


60. Friedman WJ. Memory for the time of past events. Psychol Bull 1993;113:44–66.
