The Outcome of Bulimia Nervosa: Findings From One-Quarter Century of Research

Hans-Christoph Steinhausen, M.D., Ph.D., D.M.Sc.
Sandy Weber, Cand.Phil.

Objective: The present review addresses the outcome of bulimia nervosa, effect variables, and prognostic factors.

Method: A total of 79 study series covering 5,653 patients suffering from bulimia nervosa were analyzed with regard to recovery, improvement, chronicity, crossover to another eating disorder, mortality, and comorbid psychiatric disorders at outcome. Forty-nine studies dealt with prognosis only. Final analyses on prognostic factors were based on 4,639 patients.

Results: Joint analyses of data were hampered by a lack of standardized outcome criteria. There were large variations in the outcome parameters across studies. Based on 27 studies with three outcome criteria (recovery, improvement, chronicity), close to 45% of the patients on average showed full recovery of bulimia nervosa, whereas 27% on average improved considerably and nearly 23% on average had a chronic protracted course. Crossover to another eating disorder at the follow-up evaluation in 23 studies amounted to a mean of 22.5%. The crude mortality rate was 0.32%, and other psychiatric disorders at outcome were very common. Among various variables of effect, duration of follow-up had the largest effect size. The data suggest a curvilinear course, with highest recovery rates between 4 and 9 years of follow-up evaluation and reverse peaks for both improvement and chronicity, including rates of crossover to another eating disorder, before 4 years and after 10 years of follow-up evaluation. For most prognostic factors, there was only conflicting evidence.

Conclusions: One-quarter of a century of specific research in bulimia nervosa shows that the disorder still has an unsatisfactory outcome in many patients. More refined interventions may contribute to more favorable outcomes in the future.

(Am J Psychiatry 2009; 166:1331–1341)

The introduction and definition of bulimia nervosa were presented only 30 years ago. In his salient publication, Gerald Russell (1) emphasized the dread of overeating, various compensatory measures, and the morbid fear of gaining weight and getting fat. Within this relatively short period of time, a remarkably large number of outcome studies have been published. Early reviews included one review based on eight studies (2) and another on 24 follow-up studies (3). The latter found a mean recovery rate of 47.5% and mean rates of 26% for both improvement and chronicity. Shortly before the turn of the century, Keel and Mitchell (4) analyzed the course of 56 patient series and found a recovery rate of 50% and chronicity in 20% of the patients. The mortality rate of 0.3% was slightly lower than that reported in the prior review of 24 studies (3). The review by Vaz (5) concentrated on prognostic factors only, and Quadflieg and Fichter (6) reviewed various outcome measures in a total of 33 studies. Finally, a recent, more selective review based on very rigorous inclusion criteria by Berkman et al. (7) provided findings from 13 patient series. Steinhausen’s analysis (8) of the outcome of anorexia nervosa in the second half of the last century served as a model for the present review.

Method

Selection criteria for the inclusion of studies in the present review were 1) data on at least one out of five outcome measures for bulimia nervosa after a minimum follow-up evaluation period of 6 months following the treatment episode and/or 2) data on any prognostic factor of the disorder. The aforementioned reviews consisted of a total of 141 studies. A systematic search with various databases was performed using the terms “bulimia nervosa,” “outcome,” “follow-up,” and “prognosis.” A search using PubMed led to an additional 60 studies, and a search using PsycINFO resulted in 19 studies. Thus, before applying any exclusion criteria, a total of 220 published studies were available. As a result of duplicate or review-type studies, 68 studies were excluded. Furthermore, 26 studies did not include sufficient information on the duration of follow-up evaluation or provided pre-post measures only or were based only on follow-up periods <6 months. In another 21 studies, information on the course of the disorder was insufficient (e.g., no information on sample size or no independent assessment of anorexia and bulimia nervosa patients was provided). Finally, there were 14 studies dealing with prognostic factors only, and another 12 studies presented additional data based on the same patient cohort. Thus, at the end, a total of 79 patient series were entered into the analyses for the outcome of bulimia nervosa in the present review. Findings were published between 1981 and 2007. Data based on reports from a total of 32 studies were extracted by the expert senior author. Decisions on

This article is the subject of a CME course (p. 1437) and is discussed in an editorial by Dr. Kaye (p. 1309).
TABLE 1. Global Outcome Classification for Bulimia Nervosa in 79 Patient Series (N=5,653)

<table>
<thead>
<tr>
<th>Variable</th>
<th>Group Size</th>
<th>Number of Studies</th>
<th>Rate of Outcome (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Three-level classification</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Recovery</td>
<td>1,235</td>
<td>27</td>
<td>Mean 44.9, SD 15.6, Range 5–73</td>
</tr>
<tr>
<td>Improvement</td>
<td>1,235</td>
<td>27</td>
<td>Mean 27.0, SD 12.9, Range 4–67</td>
</tr>
<tr>
<td>Chronicity</td>
<td>1,235</td>
<td>27</td>
<td>Mean 22.6, SD 15.1, Range 2–79</td>
</tr>
<tr>
<td>Crossover&lt;sup&gt;a&lt;/sup&gt;</td>
<td>383</td>
<td>6</td>
<td>Mean 5.1, SD 8.6, Range 0–28</td>
</tr>
<tr>
<td>Two-level classification with supplement parameter</td>
<td>1,284</td>
<td>27</td>
<td>Mean 59.9, SD 12.2, Range 24–78</td>
</tr>
<tr>
<td>Recovery</td>
<td>1,284</td>
<td>19</td>
<td>Mean 59.9, SD 12.2, Range 24–78</td>
</tr>
<tr>
<td>Chronicity</td>
<td>1,284</td>
<td>19</td>
<td>Mean 30.0, SD 17.2, Range 5–76</td>
</tr>
<tr>
<td>Crossover&lt;sup&gt;a&lt;/sup&gt;</td>
<td>891</td>
<td>9</td>
<td>Mean 10.1, SD 8.7, Range 0–27</td>
</tr>
<tr>
<td>Two-level classification without supplement parameter</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Recovery</td>
<td>389</td>
<td>8</td>
<td>Mean 42.3, SD 17.6, Range 15–71</td>
</tr>
<tr>
<td>Improvement</td>
<td>389</td>
<td>8</td>
<td>Mean 41.3, SD 13.7, Range 29–85</td>
</tr>
<tr>
<td>One-level classification</td>
<td>2,745</td>
<td>25</td>
<td></td>
</tr>
<tr>
<td>Recovery</td>
<td>1,639</td>
<td>15</td>
<td>Mean 42.4, SD 10.7, Range 15–75</td>
</tr>
<tr>
<td>Improvement</td>
<td>57</td>
<td>2</td>
<td>Mean 66.6, SD 22.2, Range 45–89</td>
</tr>
<tr>
<td>Chronicity</td>
<td>1,049</td>
<td>8</td>
<td>Mean 50.8, SD 22.8, Range 4–73</td>
</tr>
<tr>
<td>Crossover&lt;sup&gt;a&lt;/sup&gt;</td>
<td>1,031</td>
<td>8</td>
<td>Mean 31.65, SD 10.7, Range 16–44</td>
</tr>
<tr>
<td>Crossover total&lt;sup&gt;a&lt;/sup&gt;</td>
<td>2,305</td>
<td>23</td>
<td>Mean 22.5, SD 12.1, Range 2–44</td>
</tr>
<tr>
<td>Anorexia nervosa</td>
<td></td>
<td></td>
<td>Mean 5.7, SD 9.4, Range 0–27</td>
</tr>
<tr>
<td>Eating disorder not otherwise specified</td>
<td></td>
<td></td>
<td>Mean 16.4, SD 14.4, Range 0–43</td>
</tr>
<tr>
<td>Binge eating disorder</td>
<td></td>
<td></td>
<td>Mean 0.4, SD 1.3, Range 0–7</td>
</tr>
<tr>
<td>Mortality</td>
<td>4,309</td>
<td>76</td>
<td>Mean 0.3, SD 0.6, Range 0–2</td>
</tr>
<tr>
<td>Comorbid mental disorders</td>
<td>5,653</td>
<td>79</td>
<td></td>
</tr>
<tr>
<td>Affective disorders</td>
<td>451</td>
<td>8</td>
<td>Mean 22.5, SD 17.3, Range 7–76</td>
</tr>
<tr>
<td>Neurotic (anxiety) disorders</td>
<td>376</td>
<td>4</td>
<td>Mean 16.2, SD 6.9, Range 2–22</td>
</tr>
<tr>
<td>Drug addiction/alcoholism</td>
<td>501</td>
<td>8</td>
<td>Mean 7.3, SD 5.9, Range 1–15</td>
</tr>
<tr>
<td>Personality disorders</td>
<td>341</td>
<td>5</td>
<td>Mean 15.3, SD 20.7, Range 3–73</td>
</tr>
<tr>
<td>Obsessive-compulsive disorder</td>
<td>89</td>
<td>1</td>
<td>Mean 1.0, SD 1.0, Range 0–2</td>
</tr>
<tr>
<td>Schizophrenia spectrum disorders</td>
<td>80</td>
<td>1</td>
<td>Mean 1.0, SD 1.0, Range 0–2</td>
</tr>
</tbody>
</table>

<sup>a</sup>Data represent crossover to another eating disorder.

the inclusion of the remaining studies were jointly made by both the junior and senior authors.

**Study Characteristics**

The 79 published reports (9–87) were composed of 5,653 patients (group mean size=71.6 [SD=113.4], range=4–884). There were considerable differences in design, group size, methods, duration of follow-up evaluation, and missing data. Diagnostic classification changed considerably over the period in which the studies were conducted. Since the 1990s, there has been an increasing reliance on DSM-IV and ICD-10 criteria. In 46 studies consisting of 2,508 patients, the mean age at onset was 17.2 years (SD=1.7, range=4.3–23.2), and the mean age at follow-up assessment was 28.4 years (SD=4.3, range=16.6–38.0) in 39 studies consisting of 2,478 patients. The mean duration of follow-up evaluation varied between 6 months and 12.5 years (mean=3.2, SD=4.3, range=0.3–884). There were considerable differences in design, group size, methods, duration of follow-up evaluation, and missing data. Diagnostic classification changed considerably over the period in which the studies were conducted. Since the 1990s, there has been an increasing reliance on DSM-IV and ICD-10 criteria.
TABLE 2. Outcome of Bulimia Nervosa in 27 Patient Series by Dropout Rate

<table>
<thead>
<tr>
<th>Variable</th>
<th>1: Low (N=514)</th>
<th>2: High (N=579)</th>
<th>Analysis</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean (%)</td>
<td>SD</td>
<td>Mean (%)</td>
</tr>
<tr>
<td>Recovery</td>
<td>49.8</td>
<td>12.6</td>
<td>36.8</td>
</tr>
<tr>
<td>Improvement</td>
<td>23.7</td>
<td>12.6</td>
<td>30.7</td>
</tr>
<tr>
<td>Chronicity</td>
<td>24.4</td>
<td>12.8</td>
<td>23.4</td>
</tr>
<tr>
<td>Crossover*</td>
<td>0.9</td>
<td>2.5</td>
<td>9.2</td>
</tr>
</tbody>
</table>

*Data represent crossover to another eating disorder.

were calculated by weighting each reported rate with the size of the study group. Data for all studies were converted into individual data for performance of statistical analyses using SPSS 14 (SPSS, Chicago).

All analyses were based on adjusted sample sizes at follow-up assessments rather than actual sample sizes after patient recruitment. Differences between these two figures were considered the dropout rate. The latter was dichotomized into high (≥16% of the original sample) and low (0%–15% of the original sample) categories and served as a first independent effect variable. In accordance with previous analyses (8), the duration of follow-up evaluations was the second independent effect variable and categorized as <4 years, 4 to 9 years, or ≥10 years. Studies with a variable length of course were not considered for analyses of effect. In case there was more than one report based on the same cohort, only the last report with the longest duration of follow-up assessment was considered for the analyses. The third independent variable was represented by the type of intervention. Because of limited data, the aforementioned classification of available information on treatment into five types was restricted to the following three types: nonbehavioral psychotherapy, unspecified medical therapy, and CBT.

Effects of these three variables for treatment type on four outcome measures were analyzed using multivariate analyses of variance. In addition, effect sizes were calculated using partial eta-squared ($\eta^2$) as a measure of association between independent and dependent variables. According to Cohen (107), $\eta^2$=0.01–0.059 represents a small effect, $\eta^2$=0.06–0.13 represents a medium effect, and there is a large effect starting with $\eta^2$=0.14. Furthermore, the frequencies of positive, negative, and insignificant prognostic factors were calculated.

Results

Main Outcome Findings

The main findings on outcome are presented in Table 1. Twenty-seven studies used a three-level classification of global outcome (recovery, improvement, chronicity), supplemented by additional information on rates of crossover to another eating disorder in six studies. According to this procedure, close to 43% of the patients on average showed full recovery from bulimia nervosa, whereas 27% on average improved considerably and nearly 23% on average had a chronic protracted course. In another 27 studies, only two outcome parameters were used. Nineteen of these studies used recovery and chronicity, supplemented by additional information on rates of crossover as an indicator of the course of illness in nine studies. In these studies, recovery increased to almost 60% on average, 30% of patients on average had a chronic course, and 10% on average were marked by crossover to another eating disorder. Eight studies used a two-level classification (recovery and improvement) and found a mean of 42% of patients on average to be recovered and 41% on average to be improved. There was no information on crossover to another eating disorder in addition to a rather substantial amount of missing outcome data. Twenty-five studies used a single outcome criterion only. Among these studies, the mean recovery rate was 42% on average in 15 studies, whereas the mean improvement rate was two-thirds on average in two studies and chronicity was nearly 51% on average in eight studies. Eight of the 25 studies provided additional information on crossover to another eating disorder, which was almost 32% on average.

Detailed information on crossover diagnoses was available in 23 studies. As shown in Table 1, more than one-fifth of the patients fulfilled this criterion, with a majority of 16% on average crossing over to eating disorder not otherwise specified, which in most cases was a subclinical manifestation of bulimia nervosa, and nearly 6% on average developed full anorexia nervosa. A few patients developed binge eating disorder.

Seventy-six studies reported on mortality, and there were 14 deaths among 4,309 patients, leading to a crude mortality rate of 0.32%. For two patients, car accidents were the cause of death, four deaths were the result of suicide, one death was the result of a drug overdose, two deaths were caused by an eating disorder, and no causes of death were given for two subjects.

There was a large list of reported comorbid mental disorders. At follow-up assessments, patients most frequently suffered from affective disorders, followed by neurotic disorders (mostly anxiety disorders). Nonspecific personality disorders and borderline personality disorder were also frequent. Substance use disorders were less frequently seen, and obsessive-compulsive and schizophrenia spectrum disorders were described only in a single study.

Findings From Repeated Follow-Up Assessments

A few studies shed some light on the differential course of bulimia nervosa across time. Fichter and Quadflieg (91) published findings after 2-, 6-, and 12-year follow-up evaluations (34) and found substantial improvement in patients who completed longer follow-up assessment periods. Similarly, in the study by Herzog et al. (42), recovery rates increased between the first follow-up assessment after 2 years and the second follow-up assessment after 7 years.
years, but rates of improvement and chronicity declined correspondingly. Another study on intervention (59) found that the reduction of binge eating episodes and compensatory vomiting and laxative abuse remained constant between the 6- and 12-month follow-up evaluations. A change in outcome rates was observed in two smaller studies. Nevonen and Broberg (64) found a decreased recovery rate and an increased improvement rate between 1 and 2.5 years of follow-up evaluation when comparing individual and group psychotherapy. Similarly, in a very small sample of six patients, Toro et al. (80) observed that all patients were recovered at the first follow-up evaluation, but only four patients remained recovered after 25 years.

**Intervention Effects on Outcome**

There are various intervention studies on bulimia nervosa that include follow-up assessments. Two studies (28, 30) compared three types of interventions and found that both interpersonal therapy and CBT were superior to behavior therapy without cognitive components. Other studies reported that CBT was superior to interpersonal therapy (29, 84) or found no significant differences in the effects of either CBT or the combination of behavior therapy and hypnotherapy (40). In another study (12), more patients became symptom-free by the use of a self-help manual rather than CBT. Physical activity, but not diet counseling, was shown to be superior to CBT in one report (78).

In a study on the optimization of CBT (82), one group of patients used a self-help manual and received additional CBT when necessary. The other patients received behavior therapy only. There were no significant differences between these two approaches. In a study on the effectiveness of either the antidepressant fluoxetine or interpersonal therapy after a first inefficient phase of CBT (62), no significant differences between the two approaches were found.

When comparing the effects of the antidepressant desipramine, CBT, or a combination of both measures (11), CBT or the combined intervention resulted in stronger reduction of symptoms after 24 weeks relative to the pure antidepressant treatment. Adding exposure and response prevention to CBT did not result in improved results (19). When comparing the effects of the antidepressant desipramine, CBT, or a combination of both measures (11), CBT or the combined intervention resulted in stronger reduction of symptoms after 24 weeks relative to the pure antidepressant treatment. Adding exposure and response prevention to CBT did not result in improved results (19).

Another study (59) revealed that relative to a coping with stress program, diet counseling led to a more rapid improvement of eating behavior and to reduction and abstinence of binge eating episodes.

Two studies on the effects of family therapy revealed contradictory findings. One study (74) found that family intervention was inferior to individual psychotherapy, whereas family therapy was found to be superior in another study (39). No significant differences were found in the effectiveness of group versus individual psychotherapy (64). Finally, a positive effect of after-treatment control visits on the course of bulimia nervosa has been documented (58).

**Effect Variables**

As a result of methodological restrictions, the presentation of findings based on effect analyses was limited to the set of 27 studies that used the three-level classification with recovery, improvement, and chronicity, supplemented by the smaller number of studies that provided additional information on rates of diagnoses for crossover to other eating disorders. There were no missing data in this data set.

### TABLE 3. Outcome of Bulimia Nervosa in 21 Patient Series by Duration of Follow-Up Evaluation

<table>
<thead>
<tr>
<th>Variable</th>
<th>1: &lt;4 Years Follow-Up Evaluation (N=603)</th>
<th>2: &gt;4 Years Follow-Up Evaluation (N=146)</th>
<th>3: &gt;10/&gt;20 Years Follow-Up Evaluation (N=217)</th>
<th>Analysis</th>
<th>Post Hoc Test</th>
</tr>
</thead>
<tbody>
<tr>
<td>Recovery</td>
<td>Mean (%) 39.2, SD 14.5</td>
<td>Mean (%) 66.6, SD 11.3</td>
<td>Mean (%) 44.0, SD 4.0</td>
<td>F 286.63, df 2, 963, p &lt;0.001, η² 0.373</td>
<td>2&gt;3&gt;1</td>
</tr>
<tr>
<td>Improvement</td>
<td>Mean (%) 32.2, SD 14.1</td>
<td>Mean (%) 21.6, SD 6.1</td>
<td>Mean (%) 28.4, SD 0.8</td>
<td>F 53.08, df 2, 963, p &lt;0.001, η² 0.099</td>
<td>1&gt;3&gt;2</td>
</tr>
<tr>
<td>Chronicity</td>
<td>Mean (%) 26.9, SD 15.4</td>
<td>Mean (%) 11.9, SD 17.3</td>
<td>Mean (%) 10.6, SD 0.8</td>
<td>F 145.99, df 2, 963, p &lt;0.001, η² 0.233</td>
<td>1&gt;2, 3</td>
</tr>
<tr>
<td>Crossover</td>
<td>Mean (%) 0.8, SD 3.5</td>
<td>Mean (%) 0, SD 0</td>
<td>Mean (%) 17.0, SD 4.0</td>
<td>F 2000.30, df 2, 963, p &lt;0.001, η² 0.806</td>
<td>1&lt;3</td>
</tr>
</tbody>
</table>

* Data represent follow-up assessment periods following the treatment episode.

* Data represent crossover to another eating disorder.

**FIGURE 1. Effect of the Duration of Follow-Up Assessment on the Outcome of Bulimia Nervosa in 23 Patient Series**

* Data represent crossover to other eating disorders.
did not find a significant association between age at onset and outcome. Only in three studies (10, 21, 34) was a young age at onset favorable for the outcome. Various indicators of severity did not show a clear association with the outcome of bulimia nervosa. Only two studies (63, 73) found that frequent binge eating episodes were related to poor outcome. Four studies (25, 63, 65, 105) showed the same association for purging behavior, whereas one study (31) identified laxative abuse as a positive prognostic factor. Predominant vomiting was a negative factor in one study (25) but insignificant in another study (10). Restrictive dieting was not a significant factor in any study. An overemphasis on weight and body shape was a negative prognostic factor in three studies (88, 89, 101), whereas it was insignificant in three other studies (30, 63, 97). Furthermore, no significant associations with outcome were found for low ideal weight (25, 52), maximum and minimum body weight, or body mass index prior to treatment (10, 25, 52, 65). Weight fluctuations were considered to have a positive effect in one study (31) and a negative effect in another study (63).

Additional patient characteristics. The findings on the effect of age were mostly heterogeneous. Older age was

### TABLE 4. Effect of Intervention Methods on Outcome of Bulimia Nervosa in 10 Patient Series

<table>
<thead>
<tr>
<th>Variable</th>
<th>1: Psychotherapy (N=164)</th>
<th>2: Medical Therapy (N=150)</th>
<th>3: Behavior Therapy (N=139)</th>
<th>Analysis</th>
<th>Post Hoc Test</th>
</tr>
</thead>
<tbody>
<tr>
<td>Recovery</td>
<td>62.89 (9.05)</td>
<td>49.21 (2.71)</td>
<td>31.89 (10.00)</td>
<td>F = 576.11, df = 2,450, p &lt; 0.001, η² = 0.719</td>
<td>1&gt;2&gt;3</td>
</tr>
<tr>
<td>Improvement</td>
<td>16.10 (7.98)</td>
<td>26.08 (10.99)</td>
<td>39.96 (12.34)</td>
<td>F = 195.52, df = 2,450, p &lt; 0.001, η² = 0.465</td>
<td>1&lt;2&lt;3</td>
</tr>
<tr>
<td>Chronicity</td>
<td>21.01 (8.90)</td>
<td>22.07 (13.93)</td>
<td>24.58 (13.95)</td>
<td>F = 3.25, df = 2,450, p &lt; 0.001, η² = 0.014</td>
<td>1&lt;3</td>
</tr>
<tr>
<td>Crossover</td>
<td>0.00 (0.00)</td>
<td>2.64 (4.11)</td>
<td>3.57 (6.55)</td>
<td>F = 28.24, df = 2,450, p &lt; 0.001, η² = 0.112</td>
<td>1&lt;2&lt;3</td>
</tr>
</tbody>
</table>

* Data represent crossover to another eating disorder.

**Dropout rate.** The dropout rate had a strong effect on rates of recovery and crossover diagnoses (Table 2). The rate of recovered patients at follow-up evaluations was lower in studies with high dropout rates, and the reverse pattern was true for the rate of crossover. There was a medium effect size for the rate of improvement and no significant effect for chronicity.

**Duration of follow-up evaluation.** The effect of duration of follow-up evaluation was robust among the three effect variables. The effect size for recovery, chronicity, and crossover was strong but medium for improvement (Table 3, Figure 1). Post hoc comparisons indicated that the recovery rate was strongest after 4 years of the course of bulimia nervosa but less robust before 4 years and later at ≥10 years. The rates of improvement and the number of patients with a chronic course of the disorder were highest before 4 years of follow-up evaluation. A considerable number of crossover diagnoses were only seen in patients with long-term follow-up assessment >10 years.

**Type of intervention.** In general, the effects of outcome of intervention in a small sample of 10 patient series were strong. There was a clear gradient of effects for recovery with psychotherapy, which ranked highest, followed by medical therapy and behavior therapy (Table 4, Figure 2). The gradient effect for improvement was highest with behavior therapy, followed by medical therapy and psychotherapy, and there was no real eminent difference in the proportion of patients with chronic illness across intervention types except that it was higher for behavior therapy than nonbehavioral psychotherapy. The effect of treatment type on chronicity was weak. Crossover was observed most frequently with behavior therapy, less frequently with medical therapy, and not at all with psychotherapy.

**Prognostic Factors**

In general, many prognostic factors of bulimia nervosa have been evaluated in the literature. Findings were grouped into various categories as follows:

**Specific characteristics of bulimia nervosa.** Duration of the disorder has been assessed in most studies, finding no effect on the course of illness. In five studies (25, 54, 71, 88, 90), a short duration was beneficial, whereas it was negative in one study (105). Similarly, most of the studies did not find a significant association between age at onset and outcome. Only in three studies (10, 21, 34) was a young age at onset favorable for the outcome. Various indicators of severity did not show a clear association with the outcome of bulimia nervosa. Only two studies (63, 73) found that frequent binge eating episodes were related to poor outcome. Four studies (25, 63, 65, 105) showed the same association for purging behavior, whereas one study (31) identified laxative abuse as a positive prognostic factor. Predominant vomiting was a negative factor in one study (25) but insignificant in another study (10). Restrictive dieting was not a significant factor in any study.

An overemphasis on weight and body shape was a negative prognostic factor in three studies (88, 89, 101), whereas it was insignificant in three other studies (30, 63, 97). Furthermore, no significant associations with outcome were found for low ideal weight (25, 52), maximum and minimum body weight, or body mass index prior to treatment (10, 25, 52, 65). Weight fluctuations were considered to have a positive effect in one study (31) and a negative effect in another study (63).
found to be a negative prognostic factor in one study (73), a positive factor in two studies (25, 65), and nonsignificant in six studies (10, 42, 54, 61, 71, 101). Advanced academic education (24) and having children (21) or giving birth to a child (96) were not found to be significantly associated with follow-up status.

One study addressed the effect of either current or past premorbid anorexia nervosa and found no influence of these conditions on the course of bulimia nervosa (43). Two studies examined coexisting personality disorders and, particularly, the effect of borderline personality disorder. Although Johnson et al. (51) found a negative effect of borderline and nonborderline personality disorders on the course of bulimia nervosa, this was not fully replicated in the study by Steiger and Stotland (75). In the latter study, personality disorders were associated more strongly with general psychiatric symptoms rather than with the course of bulimia nervosa. The differentiation between patients with or without treatment with imipramine (24) did not lead to significant differences in the course of bulimia nervosa.

Introversion (91), perfectionism (91), neuroticism (76), and emotional lability (91) were not significant in terms of prognosis. One study found that low self-esteem had a negative effect on the course of bulimia nervosa (108), but six other studies did not replicate this finding (30, 65, 88, 91, 97, 103). Patients seeking social support during the first month of treatment experienced a more favorable course (94), and adaptive coping contributed to recovery (106).

No association was found between obsessive or various impulsive features and the course of the disorder (25, 52). Patients with multi-impulsive behaviors, including self-injury behaviors and substance use disorders, had a worse course than those without these behaviors (98). Findings on the effect of suicidal behaviors were contradictory, with two studies implying a negative prognostic function (10, 25) and four studies failing to replicate this finding (15, 42, 43, 52). Cue reactivity, in terms of reacting by binge eating after presentations of food, was found not to be a significant prognostic factor in one study (95), but it was a negative predictor in another study (17).

**Patient history.** Only two studies found that a history of anorexia nervosa was an unfavorable prognostic sign (34, 57), and there was no indication that premorbid eating disorders in early infancy had any effect (76). However, a premorbid overweight condition was considered as both an impediment to a favorable course (17, 30, 88) and a nonsignificant factor (34, 76, 91, 97). Two studies found that depressive symptoms before treatment of bulimia nervosa had a negative effect (17, 92). Most studies did not find a significant association between a substance use disorder and the course of bulimia nervosa, with only two studies finding a negative effect (54, 57) and one study finding a positive effect (17).

Among a large list of comorbid mental disorders, none of the various axis I disorders and only a single axis II disorder showed a significant negative association with outcome (25). Only the combination of all comorbidities had a significant negative effect in a single study (34), which was not replicated in another study (91). Furthermore, frequent hospitalizations (25, 52, 101) and treatment of other diagnoses (10, 30, 42, 52, 61, 71, 76) did not imply poor response to treatment of bulimia nervosa.

**Family history and environment.** The search for family history and environment variables with an effect on the course of bulimia nervosa resulted in only a few significant findings. No significant associations were found for affective disorders (25, 30, 52, 71, 90) or eating disorders in patients’ family history (21, 30, 52, 90) or for any parental psychopathology (15, 30, 90). Nonsignificant findings were obtained for parental socioeconomic status (52), sexual abuse (30, 90), and quality of family bonding (30). However, a single study (90) found that physical maltreatment within a controlling family environment lacking support and affection contributed to an unfavorable course of bulimia nervosa.

A family history of obesity (30, 73) and disturbed family relationships (73) had a negative effect on outcome, whereas the findings for substance use disorders were heterogeneous. Four studies revealed nonsignificant findings for substance use disorders (30, 52, 90, 91), one study identified substance use disorders as favorable prognostic factors (21), and another study (15) identified substance use disorders as unfavorable prognostic factors.

**Social Factors**

Positive social adjustment (88), close social relationships (55), and high socioeconomic status (21) were all identified as positive prognostic factors, whereas a high amount of psychosocial distress (72, 73) and low job status (72) were considered negative factors.

**Treatment Factors**

Neither previous outpatient (34, 52, 63) nor hospital treatment (10, 52) had a significant effect on the course of bulimia nervosa. However, poor motivation at treatment onset was an unfavorable prognostic factor in one study (102). A rapid reduction in symptoms during the first 4 weeks of treatment was linked to a positive course (15, 97), but the number of treatment sessions was not significant (35). Continuous dieting and overemphasis on weight and body shape after treatment were clearly linked to a negative course of the disorder (17, 35, 88). Furthermore, continuous treatment after hospital discharge was associated with a poor prognosis in two studies (92, 102) and found to be nonsignificant in another study (76).

**Discussion**

Similar to the previous extended review by Steinhausen (8) on the course and outcome of anorexia nervosa, the present review on the course and outcome of bulimia nervosa was based on the largest database currently available.
and was not confined to descriptive statistics, which was the case in most prior reviews. Using methods of inferential statistics, an attempt was also made to isolate factors that might have influenced the course of bulimia nervosa. Findings were based on a large group of 5,653 patients on whom data were published in little more than one-quarter of a century. In the Steinhausen review, a similar number of patients suffering from anorexia nervosa (N=5,590) was studied within one-half a century. Thus, the intensity of studying the outcome of bulimia nervosa has been very strong. Furthermore, it has to be noted that the present review was based on published data on patients who had been seen predominantly by expert groups. However, as long as there are no reports from other sources, such as community physicians, it is impossible to conclude whether or not the present findings are biased as a result of patient selection.

The discussion of the main outcome findings was severely hampered by a lack of commonly accepted outcome criteria. Different three- and two-level classifications or single criteria for the outcome of bulimia nervosa were presented in the literature. Given the wide acceptance of the distinction among recovery, improvement, and chronicity as classification of the global outcome of diseases in general, and in the previous review by Steinhausen on the course of anorexia nervosa in particular (8), findings based on this classification should be considered those with the highest face value. Within this scheme, findings on mean recovery rates for bulimia nervosa (45%) and anorexia nervosa (46%) (8) were remarkably similar. In addition, the mean improvement rates (bulimia nervosa, 27%, versus anorexia nervosa, 33%) and mean chronicity rates (bulimia nervosa, 23%, versus anorexia nervosa, 20%) were not extremely different in these two largest reviews. However, it needs to be mentioned that, first, these figures represented only a central tendency, and there was a large variation across studies in both reviews. Second, the different criteria of outcome on the course of bulimia nervosa in the studies add to the uncertainty of the data. Thus, studies relying on other schemes of classification resulted in higher or similar mean rates of recovery or higher mean rates of improvement and chronicity. There is no meta-analytical strategy to overcome these different findings because they are rooted in basic differences of the studies.

According to the present review, crossover to other eating disorders in the course of bulimia nervosa is very common. However, as a result of differences in the design of the outcome criteria of the studies, it was difficult to identify precisely the mean rate of crossover diagnoses, which was between a 10% and 32% range, depending on the criteria for the outcome. Obviously, the most common crossover at follow-up evaluations was to eating disorder not otherwise specified, followed by anorexia nervosa, and the least common crossover was to binge eating disorder. However, the low rate of binge eating disorder may be partially the result of underreporting because the term was not yet introduced when many of the older outcome studies were performed.

With a mean crude mortality rate of 0.32%, including a number of deaths not caused by bulimia nervosa, bulimia nervosa was definitely less fatal than anorexia nervosa, which resulted in a mean crude mortality rate of 5% in the review by Steinhausen (8). However, the frequencies of comorbid psychiatric disorders were high for both disorders. Affective and neurotic/anxiety disorders ranked highest, and there was a sizable proportion of patients with personality disorders at follow-up evaluations. Although the crude figures for comorbid disorders were higher for anorexia nervosa, these differences should not be overestimated because many studies did not clearly indicate the criteria for assessment or diagnosis.

Several conclusions may be drawn from the present analyses of three central variables (dropout rate, duration of follow-up evaluation, and type of intervention) with an effect on outcome. First, there is some evidence that, perhaps counter to expectation, patients who dropped out of follow-up studies may have had a more favorable course of bulimia nervosa. Staying in the follow-up cohort may reflect patients’ continuous need for further treatment. Thus, one might argue that the outcome of representative samples of patients without sample loss may be more favorable than delineated from the present data.

Second, duration of follow-up evaluation was the variable with the strongest effect on outcome. The distribution of data did not allow for a more fine-grained analysis, particularly for ≤4 years of follow-up evaluation. The profiles shown in Figure 1 clearly indicate a curvilinear course of bulimia nervosa, particularly if one considers the curve representing crossover diagnoses to be a part of the chronic course of illness. It is also obvious that the mean recovery rate peaked in the 4- to 9-year follow-up interval and declined thereafter, whereas the rates for chronicity and crossover followed the reverse pattern and the rates of improvement remained relatively stable across time. According to these data, the developmental trajectory of bulimia nervosa in the present analyses was rather different from the course of anorexia nervosa in the previous review by Steinhausen (8), which showed a linear relationship indicating better outcome with increasing duration of follow-up evaluation. However, it needs to be taken into consideration that these data were derived from a composition of cross-sectional samples only. The few larger longitudinal bulimia nervosa outcome studies with repeated assessments over extended follow-up periods tended to show a more favorable course with increasing duration of follow-up evaluation (34, 42). However, these findings may not be representative because they were based on patients who had been treated in expert centers.

Third, both the analyses of effect variables and of intervention studies allow for conclusions on the role of treatment for bulimia nervosa. This is in contrast to the
studies on anorexia nervosa, which included a paucity of facts on the effect of interventions on outcome (8). In the entire data set of studies in the present analyses, there were strong effects indicating a clear superiority of psychotherapy over medical therapy and behavior therapy. However, this finding was overshadowed by a lack of clear description of the type and modalities of treatment employed in the various studies. Furthermore, as a result of methodological shortcomings and/or contradictory findings, intervention studies provided only limited evidence regarding which treatments actually may contribute to a positive outcome. Standards of intervention studies, such as controlled randomization or a waiting comparison group analysis, are almost entirely missing in the literature because of problems with clinical practicality. At least for CBT, there is some limited evidence that it may contribute to more favorable outcomes for bulimia nervosa. Clearly, this field is in need of more sophisticated research.

Finally, there was a large number of studies dealing with prognostic factors in various domains. Despite the considerable effort that was invested in this area of research, the majority of these factors did not prove to have any significant effect on the disorder. Only a few of the significant factors were replicated, and many studies resulted in contradictory findings. Information on prediction models, excluded variables, or nonsignificant factors was mostly missing in the reports. Most frequently, prognostic factors were seen in isolation. In the outcome literature on anorexia nervosa, it was shown that favorable, unfavorable, or nonsignificant prognostic factors may be extracted from the various study reports (8). However, multivariate analyses with consideration of a large group of factors in a large sample of patients with anorexia nervosa have clearly shown that very few variables actually have an effect on outcome if their covariation is considered (109). Thus, most of the findings on prognostic factors of bulimia nervosa must be considered as insufficiently controlled. Certainly, this field also is in need of further refinement of research. From the existing literature, a concentration on treatment factors seems most promising. However, one has to consider the likely nature of study data that preclude any delineation of rules for individual prognosis for a given patient.

In conclusion, the present comprehensive review of one-quarter of a century of outcome research shows that bulimia nervosa remains a serious disorder with unsatisfactory recovery and improvement rates and high rates of patients who continue to have chronic eating disorder problems and other comorbid psychiatric disorders over extended periods of their lives. Future research efforts could benefit from more collaborative and prospective studies based on large unselected samples and standardized assessment procedures using a common set of operationalized criteria of outcome. Refined study designs should particularly focus on the role of intervention in the long-term outcome of the disorder so that research might also identify beneficial effects on the course of the individual patient.

References

34. Fichter MM, Quadflieg N: Twelve-year course and outcome of bulimia nervosa. Psychol Med 2004; 34:1395–1406
59. Laesolle RG, Beumont PJ, Butow P, Lennerts W, O’Connor ME, Pirke KM, Touyz SW, Waadt S: A comparison of nutritional man-

60. Maddocks SE, Kaplan AS, Woodside DB, Langdon L, Piran N: Two year follow-up of bulimia nervosa: the importance of abstinence as the criterion of outcome. Int J Eat Disord 1992; 12:133–141


93. Anderson CB, Joyce PR, Carter FA, McIntosh VV, Bulik CM: The effect of cognitive-behavioral therapy for bulimia nervosa on temperament and character as measured by the Temperament and Character Inventory, Compr Psychiatry 2002; 43:182–188


95. Carter FA, Bulik CM, McIntosh VV, Joyce PR: Cue reactivity as a predictor of outcome with bulimia nervosa. Int J Eat Disord 2002; 31:240–250


