

# Recovery and Relapse in Anorexia and Bulimia Nervosa: A 7.5-Year Follow-up Study

DAVID B. HERZOG, M.D., DAVID J. DORER, Ph.D., PAMELA K. KEEL, Ph.D., SHERRIE E. SELWYN, B.A.,  
ELIZABETH R. EKEBLAD, A.B., ANDREA T. FLORES, M.Ed., DARA N. GREENWOOD, B.A.,  
REBECCA A. BURWELL, B.A., AND MARTIN B. KELLER, M.D.

## ABSTRACT

**Objective:** To assess the course and outcome of anorexia nervosa (AN) and bulimia nervosa (BN) at a median of 90 months of follow-up in a large cohort of women with eating disorders. **Method:** A prospective, naturalistic, longitudinal design was used to map the course of AN and BN in 246 women. Follow-up data are presented in terms of full and partial recovery, predictors of time to recovery, and rates and predictors of relapse. **Results:** The full recovery rate of women with BN was significantly higher than that of women with AN, with 74% of those with BN and 33% of those with AN achieving full recovery by a median of 90 months of follow-up. Intake diagnosis of AN was the strongest predictor of worse outcome. No predictors of recovery emerged among bulimic subjects. Eighty-three percent of women with AN and 99% of those with BN achieved partial recovery. Approximately one third of both women with AN and women with BN relapsed after full recovery. No predictors of relapse emerged. **Conclusions:** The findings suggest that the course of AN is characterized by high rates of partial recovery and low rates of full recovery, while the course of BN is characterized by higher rates of both partial and full recovery. *J. Am. Acad. Child Adolesc. Psychiatry*, 1999, 38(7):829–837. **Key Words:** recovery, relapse, anorexia, bulimia.

Anorexia nervosa (AN) and bulimia nervosa (BN) are frequently associated with chronicity. Studies with follow-up periods ranging from 10 to 15 years demonstrate persistent, full-criteria AN in approximately 12% (Eckert et al., 1995) to 14% (Strober et al., 1997) of subjects. A recent meta-analysis of BN outcome studies extending between 5 and 10 years found that approximately 50% of subjects with an initial diagnosis of BN have attained full recovery, while nearly 20% continue to meet full criteria for the disorder (Keel and Mitchell, 1997).

Few studies have investigated the rates of relapse in eating disorders. Compared with AN, BN has been more commonly associated with relapse. Relapse data across BN studies are relatively consistent, indicating that approximately 30% of BN subjects relapse after a recovery (Keel and Mitchell, 1997; Olmsted et al., 1994).

Studies have found several variables to be predictive of eating disorder outcomes. The type of eating disorder diagnosed at intake appears to be one of the most significant of these predictors: a diagnosis of AN has been observed to have a worse outcome than a diagnosis of BN (Herzog et al., 1993). Low body weight has been noted to have significant predictive value for AN (Hebebrand et al., 1997). In addition, duration of illness has emerged as a predictor of outcome in AN in several studies (Herzog et al., 1993; Ratnasuriya et al., 1991). Age at eating disorder onset has also been commonly reported as a predictor of outcome (see Steinhausen et al., 1991, for a review). Predictors of outcome in BN remain poorly understood (Fichter and Quadflieg, 1997; Keel and Mitchell, 1997). No predictors of relapse have been reported in AN. In BN, only age has been reported as a significant predictor of relapse (Olmsted et al., 1994).

Accepted February 8, 1999.

Dr. Herzog is Professor of Psychiatry (Pediatrics), Harvard Medical School, President and Founder of the Harvard Eating Disorders Center, and Director of the Eating Disorders Unit, Massachusetts General Hospital, Boston. Dr. Dorer is biostatistical consultant. Dr. Keel is Assistant Professor of Psychology, Harvard University, Cambridge, MA. Ms. Selwyn, Ms. Ekeblad, Ms. Flores, Ms. Greenwood, and Ms. Burwell are research assistants at the Eating Disorders Unit, Massachusetts General Hospital. Dr. Keller is Professor and Chairman, Department of Psychiatry and Human Behavior, Brown University, Providence, RI.

This study was sponsored in part by grants from the NIMH (5R01 MH 38333 05), the Rubenstein Foundation, and the Harvard Eating Disorders Center. The authors thank Ana Richards and Karin Nussbaum for their assistance.

Reprint requests to Dr. Herzog, Massachusetts General Hospital, Eating Disorders Unit-ACC 725, 15 Parkman Street, Boston, MA 02114.

0890-8567/99/3807-0829©1999 by the American Academy of Child and Adolescent Psychiatry.

Despite an abundance of eating disorder outcome studies, methodological limitations have prevented a cogent characterization of outcome in clinical populations of AN and BN. Few studies have used prospective observation with frequent follow-ups to assess the course and outcome of AN and BN, with most relying on cross-sectional measurements. The cross-sectional design fails to capture the fluctuations in course (such as recoveries and relapses) that occur between follow-up evaluations. In addition, nonstandardized definitions of outcome make comparison of results across studies difficult.

Since 1987, we have been conducting a prospective, naturalistic, longitudinal study of AN and BN. The design of this study advances the existing methodology in several ways. First, we follow a large sample of treatment-seeking women with AN and/or BN. Second, interviews have been conducted every 6 months throughout 11 years of the study. Third, the study uses the modern statistical methods of survival analysis to describe the timing of events such as recovery and relapse. Finally, the definitions of outcome implemented by this study, which are based on the *DSM-IV* criteria and the Psychiatric Status Rating (PSR) scale, are evenly applied to both AN and BN subjects. Thus, these definitions allow for comparisons between AN and BN regarding course and outcome. Such a comparison is critical to understanding eating disorders as a whole, given the shared symptomatology and established crossover between disorders. In this report we will address the following: (1) time to full or partial recovery among anorexic and bulimic subjects over a median of 90 months of follow-up; (2) time to relapse among anorexic and bulimic subjects; and (3) variables that predict time to full recovery, time to partial recovery, and time to relapse.

## METHOD

### Subjects and Recruitment

The original sample consisted of 225 women who sought treatment for an eating disorder at the Massachusetts General Hospital Eating Disorders Unit and at other Boston area eating disorder programs between October 1987 and June 1990. Of the 268 women who met inclusion criteria for the study, 229 (85%) agreed to participate. Four of these women never completed their first follow-up interview. The original sample was composed of 40 women with AN, 95 with BN, and 90 with AN/BN. AN/BN was diagnosed when subjects met *DSM-III-R* criteria for both AN and BN within their intake episode and the anorexic and bulimic episodes were not separated by a period of remission of at least 8 weeks' duration.

To increase our statistical power for analyses of AN and to allow for better comparison of course across the 2 groups (AN and BN), we

recruited an additional 21 women with AN in 1991. A detailed description of the inclusion criteria and recruitment process has been reported elsewhere (see Herzog et al., 1992). Of the 26 women meeting inclusion criteria, 21 (81%) agreed to participate. In accordance with the *DSM-IV* subtyping of AN into restricting and binge-eating/purging categories, the sample was redivided into 3 groups: (1) subjects meeting full criteria for AN who restrict their dietary intake without regularly engaging in bingeing or purging behavior (ANR); (2) subjects meeting full criteria for AN who do *regularly engage (at least once per week)* in binge eating and/or purging behavior (vomiting, diuretic use, laxative use) (ANBP); (3) subjects meeting full criteria for bulimia nervosa (BN). The current sample of 246 women meet *DSM-IV* criteria for ANR ( $n = 51$ ), ANBP ( $n = 85$ ), and BN ( $n = 110$ ).

The reclassification of subjects from *DSM-III-R* diagnoses to *DSM-IV* diagnoses resulted in an additional 15 women identified as BN who had been classified previously as AN/BN. This increase reflects the fact that our initial diagnosis of AN/BN included women with past episodes of AN in addition to an intake diagnosis of BN, whereas the *DSM-IV* diagnosis of ANBP reflects *currently* overlapping diagnoses.

Table 1 summarizes the characteristics of our sample. The mean age at intake was 24.8 years (range 13–45 years). The mean duration of illness was 6.7 years (range 3 months–21 years). We found high levels of current and past psychiatric comorbidity across the eating disorder groups at intake. Nearly all subjects have received treatment over the course of the study (see Herzog and Sacks, 1993, for details). Further demographic characteristics of our sample have been described elsewhere in detail (Herzog et al., 1992, 1993).

### Data Collection Protocol-Procedure

Informed consent was obtained from subjects who met inclusion criteria. These subjects were then contacted by phone and scheduled for intake interviews. The main purpose of the intake interview was to assess each subject's current and lifetime psychiatric history. All intake interviews were conducted in person by trained clinical interviewers. Interviewers were trained by a senior training interviewer using the training program developed by the National Institute of Mental Health-Collaborative Depression Study (Gibbon et al., 1981).

Follow-up interviews occurred every 6 months and were conducted in person whenever possible. When a subject had moved away, the interviews were conducted by telephone. Of the most recent interviews conducted with each subject, 82.9% were completed over the telephone. We have only partial information about 17 subjects who are no longer participating in the study. To date, 7 subjects have died.

### Instruments

Lifetime Axis I psychiatric history was assessed with a modified version of the Schedule for Affective Disorders and Schizophrenia-Lifetime version (SADS-L), which includes the diagnostic criteria for *DSM-III-R* eating disorders (EAT SADS-L). Axis II disorders were assessed with the Structured Interview for *DSM-III* Personality Disorders in all subjects 18 years or older. Each subject's proportion of average body weight (pABW) was calculated using the Metropolitan Life Insurance Company (1983) height and weight norms (Field et al., 1997; Herzog et al., 1993).

Follow-up assessments were completed using the Eating Disorders Longitudinal Interval Follow-up Evaluation (LIFE-EAT-II), a modified version of the LIFE II (Keller et al., 1987). The LIFE-EAT-II, which has been described in detail in a previous article (Herzog et al., 1992), is a semistructured instrument designed for collecting longitudinal data on eating disorders, Research Diagnostic Criteria comor-

**TABLE 1**  
Characteristics of Sample at Intake (*N* = 246)

	Full Sample ( <i>n</i> = 246)	ANR ( <i>n</i> = 51)	ANBP ( <i>n</i> = 85)	BN ( <i>n</i> = 110)	Statistic ( <i>df</i> )	<i>p</i> Value
	Mean (SD)	Mean (SD)	Mean (SD)	Mean (SD)		
Age	24.8 (6.7)	23.9 (8.5)	24.5 (5.9)	25.5 (6.5)	Kruskal- Wallis	.09
Proportion of average body weight	0.89 (0.18)	0.73 (0.09)	0.82 (0.10)	1.03 (0.15)		<.0001
Duration of intake episode	6.7 (6.1)	6.4 (6.7)	7.6 (5.4)	6.1 (6.3)		.004
Age at eating disorder onset	18.1 (5.6)	17.5 (6.1)	16.9 (4.7)	19.4 (5.8)		.0006
	%	%	%	%	$\chi^2$	
Lifetime history of major depression	65.2	64.7	71.3	60.7	(2)	.30
Lifetime history of Axis I disorders	73.2	62.7	78.1	74.1	(2)	.14
Lifetime history of Axis II disorders	28.8	25.5	37.9	23.2	(2)	.06
Lifetime history of substance use disorders	12.4	5.9	16.1	12.5	(2)	.21
Lifetime history of hospitalization	32.8	62.7	39.1	14.3	(2)	<.0001
Lifetime history of suicidality						
None	81.2	84.3	73.6	85.7	(4)	.16
Gesture	7.6	3.9	12.6	5.4	(4)	
Attempt	11.2	11.8	13.8	9.0	(4)	

*Note:* ANR = subjects meeting full criteria for anorexia nervosa who restrict their dietary intake without regularly engaging in bingeing or purging behavior; ANBP = subjects meeting full criteria for anorexia nervosa who regularly engage in binge eating and/or purging behavior; BN = subjects meeting full criteria for bulimia nervosa.

bid psychopathology, treatment participation, and psychosocial functioning. The eating disorders section was derived from the Diagnostic Interview Schedule (Robins et al., 1981).

### Definitions of Outcome

Definitions of outcome consist of 2 components: a state of illness (i.e., level of symptomatology) and a duration (i.e., the minimum length of time a person must remain at that level of symptomatology). Recovery and relapse were assessed using the 6-point Psychiatric Ratings Scale from the LIFE-EAT-II (see Herzog et al., 1993 and 1996, for details). Full recovery is defined as the absence of symptoms or the presence of only residual symptoms for at least 8 consecutive weeks (PSR of 1 or 2). Partial recovery is defined as a reduction of symptoms to less than full criteria for at least 8 consecutive weeks (PSR of 3 or 4). A relapse is defined as the return of full criteria symptoms (PSR of 5 or 6) for at least 8 weeks following a state of full recovery (PSR of 1 or 2).

### Statistical Methods

We used the Kaplan-Meier survival method to estimate the cumulative probability of recovery after intake into the study. The log-rank test was used to compare time to recovery across eating disorder diagnoses. Using the method of proportional hazards (Cox) regression, we examined the extent to which specific eating disorder characteristics and comorbid psychopathology ascertained at the time of intake predicted time to recovery (Cox and Oakes, 1984). The proportional hazards assumption for the final model was tested with the procedure of Therneau (1994) (S-Plus function *cox.zph*). We also analyzed plots of martingale residuals as a diagnostic aid to model building (Fleming and Harrington, 1991).

Nested models were compared using the likelihood ratio statistic *G*-squared (*G*<sup>2</sup>), which follows a  $\chi^2$  distribution. A Cox model pro-

vides a model coefficient ( $\beta_i$ ), hazard multiplier,  $\exp(\beta_i)$ , which we will denote as HM, and confidence interval (CI) for the hazard multiplier. Hazard multipliers greater than 1 increase the hazard, thus shortening the time to event, whereas multipliers less than 1 lengthen the time to event. A multiplier of 1 has no effect on the time to event.

We analyzed the following eating disorder characteristics: intake eating disorder diagnosis (AN versus BN; ANR versus ANBP versus BN); pABW; age at onset of first eating disorder; age at entry into the study; duration of intake episode; history of hospitalization for an eating disorder; history of major depressive disorder; history of drug abuse; history of suicide attempt; history of alcohol abuse; history of Axis I disorders; history of Axis II disorders; maximum bingeing frequency during 8 weeks prior to intake; and maximum purging frequency during 8 weeks prior to intake.

No correction is made for multiple comparisons in the reported *p* values of Cox models. A conservative Bonferroni-type correction for 13 independent comparisons would give a bound on a *p* value at .05 for a nominal value of 0.0038 = 0.05/13. Because of the high correlation between the analysis variables, it is likely that this correction is very conservative.

Since the power of the log-rank test increases with increasing events (Collett, 1994), our analyses have the greatest power for partial recovery (217 events), less power for full recovery (120 events), and the least power for relapse (39 events). Statistical calculations were performed using the S-plus statistical package version 4 for Windows.

## RESULTS

### Full Recovery

Approximately half (49.0%) of the full sample, 33.7% of AN subjects, and 73.8% of BN subjects achieved full

recovery at some point over a median of 90 months of follow-up (Kaplan-Meier estimates). The full recovery rates increased over time, as increasing numbers of subjects sustained an asymptomatic state for at least 8 weeks (Fig. 1). At 2 years of follow-up, 8% of ANR subjects, 13% of ANBP subjects, and 53% of BN subjects had reached full recovery at some point. At 7 years of follow-up, 34% of the ANR subjects, 32% of the ANBP subjects, and 73% of the BN subjects had achieved full recovery at some point.

Figure 1 shows the estimates for full recovery split by eating disorder diagnosis at intake. The ANR and the ANBP curves level out with about one third of subjects recovered at 400 weeks. There is no Kaplan-Meier estimate for a median time to recovery in the 2 AN groups. For the BN group, the median time to full recovery was 83 weeks. The 3 curves are different (log-rank  $\chi^2 = 55.5$ ,  $df = 2$ ,  $p < .0001$ ). The ANR and ANBP curves do not differ significantly (log-rank  $\chi^2 = 0.01$ ,  $df = 1$ ,  $p = .92$ ).

#### Predictors of Full Recovery

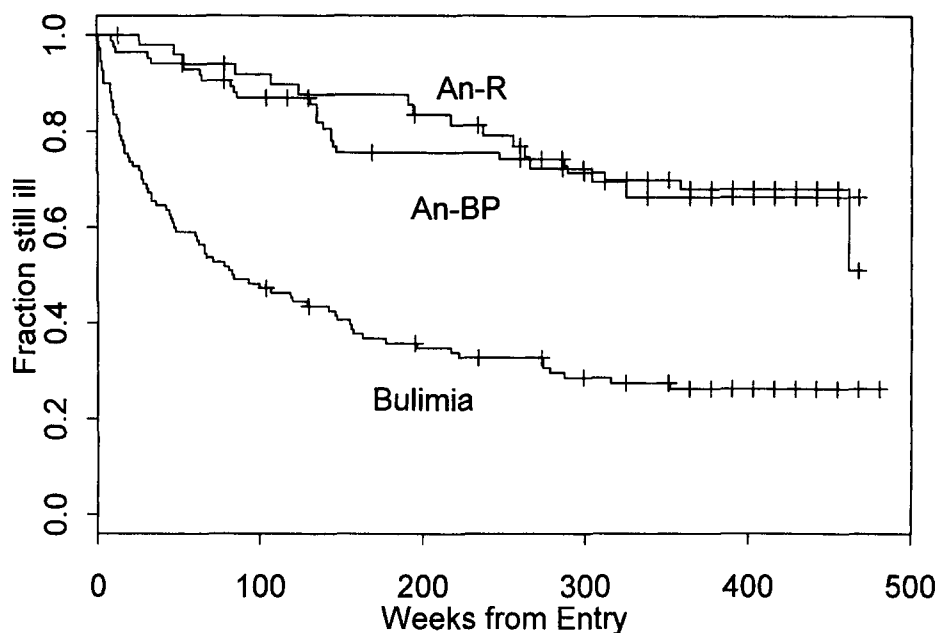
The strongest predictor of full recovery was intake diagnosis (AN versus BN) ( $G^2 = 49.6$ ,  $df = 1$ ,  $p \ll .001$ ,  $HM = 3.77$ ; 95%  $CI = 2.58, 5.53$ ). To simplify the model interpretation, and, because of the very large

effect of AN versus BN, the analyses were conducted for the AN and BN subgroups separately.

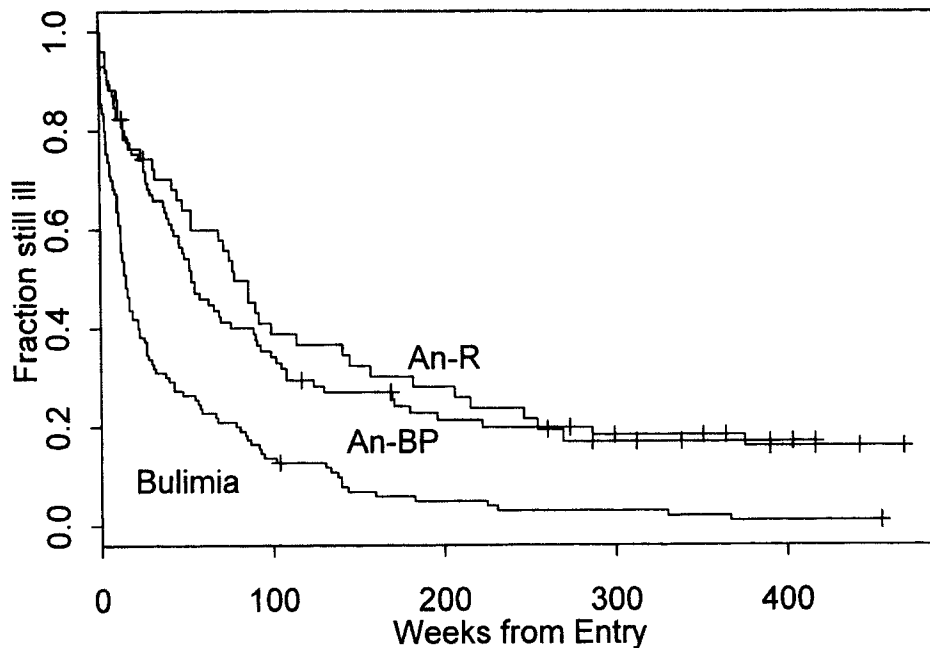
Among the AN group, we found that higher pABW at intake ( $G^2 = 9.41$ ,  $df = 1$ ,  $p = .002$ ) and shorter duration of intake episode ( $G^2 = 11.9$ ,  $df = 1$ ,  $p = .0005$ ) were associated with shorter time to full recovery. This leads to a final model containing duration of intake episode ( $HM = 0.89$ ;  $CI = 0.81, 0.96$ ) and pABW ( $HM = 250.1$ ;  $CI = 6.90, 9,066$ ) for predicting time to full recovery, such that longer intake episodes and lower body weights predict longer times to full recovery. Among the BN group, no factors emerged as significant predictors of full recovery.

#### Partial Recovery

A large majority (91%) of the full sample, 83.7% of AN subjects, and 99.0% of BN subjects achieved partial recovery at some point during a median of 90 months of follow-up (Kaplan-Meier estimates). At 2 years of follow-up, 61% of ANR subjects, 67% of ANBP subjects, and 88% of BN subjects had achieved partial recovery at some point. At 7 years of follow-up, 83% of ANR subjects, 82% of ANBP subjects, and 98% of BN subjects had achieved a partial recovery. Thus, for 17% of ANR subjects, 18% of ANBP subjects, and 1% of BN subjects,



**Fig. 1** Full recovery rates (Psychiatric Status Rating  $< 3$  for 8 consecutive weeks). An-R = subjects meeting full criteria for anorexia nervosa who restrict their dietary intake without regularly engaging in bingeing or purging behavior; An-BP = subjects meeting full criteria for anorexia nervosa who regularly engage in binge eating and/or purging behavior; Bulimia = subjects meeting full criteria for bulimia nervosa.



**Fig. 2** Partial recovery rates (Psychiatric Status Rating = 3 or 4 for 8 weeks or <3 for fewer than 8 consecutive weeks). An-R = subjects meeting full criteria for anorexia nervosa who restrict their dietary intake without regularly engaging in bingeing or purging behavior; An-BP = subjects meeting full criteria for anorexia nervosa who regularly engage in binge eating and/or purging behavior; Bulimia = subjects meeting full criteria for bulimia nervosa.

eating disorder symptomatology has never remitted for at least 8 weeks since intake into the study through year 7.

Figure 2 shows the estimates of partial recovery split according to eating disorder intake diagnosis. The median time to partial recovery is 78 weeks for ANR, 53 weeks for ANBP, and 14 weeks for BN. As with full recovery, the 3 curves are different (log-rank  $\chi^2 = 37.3$ ,  $df = 2$ ,  $p < .0001$ ), and there are no significant differences between the ANR and ANBP curves (log-rank  $\chi^2 = 0.4$ ,  $df = 1$ ,  $p = .53$ ). As with full recovery, predictors of partial recovery were analyzed for AN and BN subjects separately.

#### Predictors of Partial Recovery

In the AN group, longer duration of intake episode, pABW, and a history of hospitalization predicted delayed onset of partial recovery, whereas history of major depression ( $G^2 = 5.5$ ,  $df = 1$ ,  $p = .02$ ) predicted shorter time to partial recovery. History of hospitalization was still significant when it was added to the model with pABW and duration of intake episode ( $G^2 = 4.5$ ,  $df = 1$ ,  $p = .03$ ). Thus, history of hospitalization was significant when controlling for pABW and duration of intake episode. Furthermore, all possible combinations for the sequential addition of these 3 variables lead to signifi-

cant likelihood ratios ( $G^2$ ). Thus, these 3 variables provide independent information on the time to partial recovery. We investigated the three 2-way interactions among these 3 variables: pABW  $\times$  duration of intake episode ( $G^2 = 4.36$ ,  $df = 1$ ,  $p = .037$ ), duration of intake episode  $\times$  history of hospitalization ( $G^2 = 0.003$ ,  $df = 1$ ,  $p = .96$ ), as well as pABW  $\times$  history of hospitalization ( $G^2 = 3.46$ ,  $df = 1$ ,  $p = .062$ ). Duration of intake episode  $\times$  history of hospitalization was excluded from the final model because of a large  $p$  value. Starting with a base model of 5 terms—pABW, duration of intake episode, history of hospitalization, pABW  $\times$  duration of intake episode, and pABW  $\times$  history of hospitalization—we explored the predictive significance of the remaining variables. History of major depression was significant.

The multiplier for duration of intake episode indicates that a longer intake episode results in a longer time to partial recovery. However, there is also a significant interaction between pABW and duration of intake episode in the model for partial recovery (Table 2). Because of this, the size and direction of the effect of duration of intake episode will depend on the level of pABW, as determined by the size of  $\beta_i$ . For values of pABW less than 0.93, longer duration of intake episode corresponds

**TABLE 2**  
Final Model for Predictors of Partial Recovery in Anorexia Nervosa

Coefficient	$\beta_i$	Hazard Multiplier	95% Confidence Interval
Duration of intake episode	-0.46	0.63	0.45, 0.87
pABW	0.47	18.89	0.32, 1,105.56
History of hospitalization	1.69	29.60	1.11, 791.21
History of major depression	0.25	1.64	1.07, 2.51
Duration of intake episode $\times$ pABW	0.50	1.65	1.10, 2.47
pABW $\times$ history of hospitalization	-2.47	0.007	0.0001, 0.44

Note: pABW = proportion of average body weight.

to a longer time to partial recovery. For a 1-year increase in duration of intake episode with pABW fixed at its mean of 0.78, the hazard of recovery decreases by 7%. For pABW fixed at its minimum of 0.48, a 1-year increase in duration of intake episode corresponds to a 20% decrease in hazard. In summary, for the majority of subjects with low pABW (less than 0.93), those with a longer duration of intake episode have a longer time to recovery. For those with high pABW, duration of intake episode has the opposite effect.

The model also has a significant interaction between pABW and history of hospitalization. For pABW values greater than 0.69, subjects with a history of hospitalization have a longer time to partial recovery. For low values of pABW (0.48–0.69), patients with a history of hospitalization achieve partial recovery more rapidly than patients of similar pABW and no history of hospitalization. Thus, among subjects with a very low body weight, those with a history of hospitalization have a better prognosis. Subjects with pABW between 0.69 and 1.03 who also have a history of hospitalization have a poorer prognosis. After controlling for all the other variables, we found that subjects with a history of major depression have a shorter time to partial recovery.

In the BN group, none of the variables predicted time to partial recovery. The most significant variables were increased purging frequency ( $G^2 = 3.5$ ,  $df = 1$ ,  $p = .06$ ), lack of history of suicide attempts ( $G^2 = 3.1$ ,  $df = 1$ ,  $p = .08$ ), and increased bingeing frequency ( $G^2 = 3.1$ ,  $df = 1$ ,  $p = .08$ ), which each predicted delayed onset of partial recovery in BN.

#### Relapse From Full Recovery

Forty percent of AN subjects relapsed after full recovery. Of BN subjects, 35.3% relapsed after full recovery. Figure 3 shows the time to relapse from full recovery, split by intake eating disorder diagnosis. At 100 weeks,

about 30% of fully recovered AN and BN patients have relapsed. The relapse curves level off with 40% relapsed at 200 weeks. After 400 weeks, no new relapses occurred. There is no significant difference between the 3 curves (log-rank  $\chi^2 = 0.0$ ,  $df = 2$ ,  $p = .99$ ). There were no significant predictors of relapse from full recovery in the total sample.

#### DISCUSSION

The majority of women with AN and BN reported periods of substantial symptomatic improvement over the course of 7.5 years. Subjects in the AN group were less likely to achieve full recovery compared with those in the BN group. The large majority of the subjects from each of the 3 diagnostic groups, however, attained a partial recovery by a median of 90 months of follow-up. Rates of full recovery increased slightly between the 5- and 7-year follow-up intervals, suggesting that a small number of AN and BN subjects continue to recover. While our rates of recovery appear to be consistent with the existing literature, it is difficult to make strict comparisons between various findings because definitions of outcome differ according to study. In addition, most other studies report cross-sectional outcomes at a single point in time (Eckert et al., 1995; Fichter and Quadflieg, 1997; Ratnasuriya et al., 1991), whereas our study accounts for recoveries and relapses that occur at any point throughout the follow-up.

Studies on AN have traditionally relied on relatively general definitions of outcome. These definitions are differentiated between good, fair, and poor outcomes and are often based on weight and menstrual functioning (Morgan and Russell, 1975). A review summarizing 67 AN outcome studies with 1 to 30 years of follow-up found that half of the subjects had a good or very good outcome (based on global outcome scores), while approximately 30% of the subjects had an intermediate out-

come, and fewer than 20% remained chronically ill (Steinhausen et al., 1991). A recent review of a decade of AN studies reported a similar breakdown of outcomes for both long-term (at least 4 years) and short-term treatment studies (Steiner and Lock, 1998).

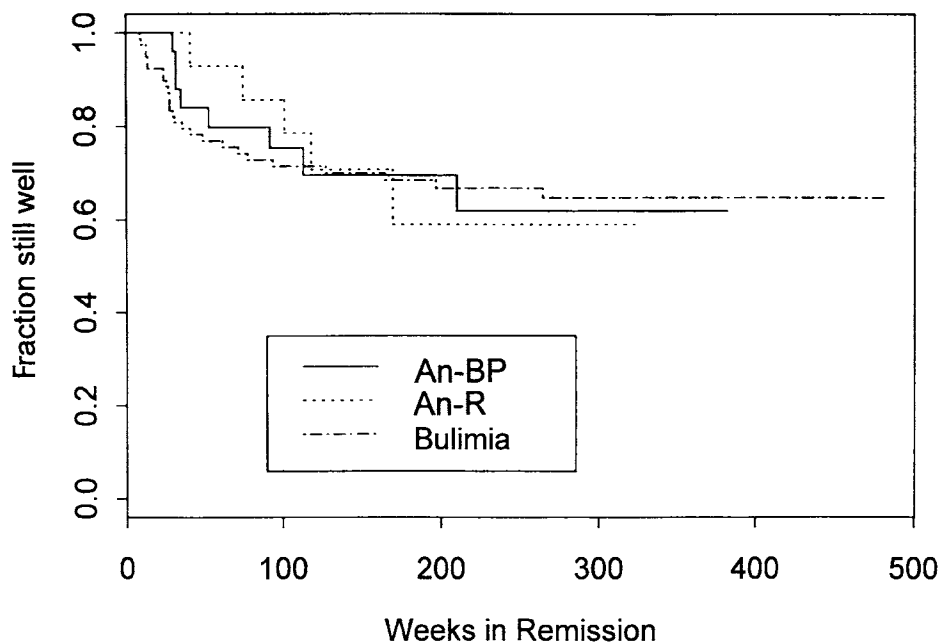
In a 10- to 15-year follow-up study of adolescents with AN, Strober et al. (1997) reported that the full recovery rate was approximately 75%. While we found lower rates of full recovery for AN (approximately 33% at a median of 90 months), our study applied more stringent outcome criteria, such as attainment of 0.95 pABW and regular menstruation. In addition, Strober and colleagues' (1997) sample was younger and had a shorter duration of illness. Furthermore, their sample had agreed to participate in intensive and lengthy inpatient treatment, which may contribute to their higher recovery rate.

Our finding that approximately one sixth of AN subjects continued to meet full criteria for the disorder throughout the study duration is consistent with previous studies (Strober et al., 1997; Sullivan et al., 1998). While symptoms improve in the majority of AN subjects, the chronic nature of AN is further supported by the substantial percentage of subjects who continue to demonstrate eating disorder symptoms, even if they do not meet full criteria for the disorder.

Relapse rates for AN are not frequently documented and are difficult to compare because of varying definitions. Our relapse rates are comparable with those reported by other researchers (Eckert et al., 1995; Norring and Sohlberg, 1993; Strober et al., 1997), although our definition of full recovery is stricter and, consequently, our recovery rates are lower.

Controlled cognitive-behavioral treatment studies of BN report generally high rates of short-term recovery (Fairburn et al., 1995). Fairburn and colleagues' data suggest that these rates may persist over follow-up periods of 5 or more years. Our naturalistic sample reports high rates of recovery as well. Notably, as duration of follow-up increases, controlled treatment outcome studies are influenced by subsequent treatment received and thus may not differ significantly from our naturalistic design. This phenomenon may explain the similarity of results.

Other long-term bulimia studies report full or partial recovery rates that range from 83% to 91% (Collings and King, 1994; Russell et al., 1987). Similarly, our rates of full and partial recovery in BN are on the high end of the range. Our high rates may be explained by (1) our use of survival analyses that account for recoveries occurring at any point throughout the follow-up period; and



**Fig. 3** Relapse rates (Psychiatric Status Rating  $\geq 5$  for 8 consecutive weeks after 8 consecutive weeks at 1 or 2). An-R = subjects meeting full criteria for anorexia nervosa who restrict their dietary intake without regularly engaging in bingeing or purging behavior; An-BP = subjects meeting full criteria for anorexia nervosa who regularly engage in binge eating and/or purging behavior; Bulimia = subjects meeting full criteria for bulimia nervosa.

(2) our definitions of recovery, which may have shorter duration criteria than those used in other studies.

We caution that the substantial rate of recovery in the bulimic cohort must be interpreted within the context of the sample's 35% relapse rate. Within the same BN sample, Field et al. (1997) reported a relapse rate of 63% based on the definition of relapse requiring at least 4 consecutive weeks of either bingeing or purging weekly or bingeing 2 or more times per week. Consistent with our relapse rate, other investigators have found that approximately one third of recovered bulimics relapse in the first 1 to 4 years of follow-up (Keel and Mitchell, 1997; Olmsted et al., 1994).

The strongest predictor of recovery was intake diagnosis; we found that BN subjects recovered more rapidly than AN subjects. Although few studies compare AN and BN outcomes, most BN studies report higher recovery rates and shorter times to recovery than do AN studies. Treatment studies, in particular, highlight the high rate of recovery in BN in the short term. AN patients, however, often experience a protracted recovery. For example, Strober et al. (1997) found that subjects rarely attained recovery during the first 4 years of follow-up.

The duration of intake episode was inversely associated with full recovery among the AN cohort, a finding which is consistent with previous outcome literature (Ratnasuriya et al., 1991). Our finding that low pABW predicts poor outcome was also noted by Hebrebrand et al. (1997). We expected duration of illness and pABW to be significant predictors of outcome because these factors are common indicators of illness severity.

While several predictors such as long duration of illness and low body weight have been reported consistently in AN outcome (Hebrebrand et al., 1997; Ratnasuriya et al., 1991; Steiner and Lock, 1998; Steinhausen et al., 1991), few predictors have been replicated across studies in BN (Keel and Mitchell, 1997). Greater severity of symptoms, longer duration of symptoms, and comorbid Axis I disorders have been associated with worse outcome (see Keel and Mitchell, 1997, for a review). No predictors of recovery in BN emerged in our study.

Predictors of relapse in AN and BN remain ambiguous. We were unable to identify any strong predictors of relapse in the total sample. Similarly, Strober et al. (1997) reported that no variable was predictive of relapse after either a partial or full recovery from AN.

We found that history of major depression was associated with a shorter time to partial recovery in AN. We

also found that for AN subjects with lowest pABW, history of hospitalization was associated with shorter time to partial recovery. This finding may indicate a positive effect of structured inpatient treatment in the most severely ill AN patients, thereby highlighting the need for hospitalization in this population. The association between hospitalization and more favorable outcome in the severely ill AN patient has been reported by other researchers as well (Eckert et al., 1995; Strober et al., 1997).

Our findings must be interpreted within the context of several limitations. First, we only studied subjects who sought treatment for an eating disorder. A second limitation involves 2 aspects of our outcome definitions: (1) full recovery from AN required subjects to achieve 0.95 pABW, which may be too stringent; and (2) the 8-week duration required in our definitions of recovery and relapse may be insufficient. In addition, the current lack of consensus on definitions of outcome variables limits the comparability of our findings. Third, the data are based on self-reported information, and the large majority of the recent interviews were conducted over the telephone, which prevented us from weighing subjects. Nonetheless, recent studies indicate that BN subjects and underweight women report their weights reliably (Doll and Fairburn, 1998; Keel et al., 1999; Rowland, 1990). Moreover, no significant differences in eating disorder outcome have been found between telephone and personal interviews (Keel et al., 1999).

#### Clinical Implications

The chronic course of eating disorders for many patients suggests the need to identify these problems early and to intervene before the disorders become intractable. In addition, the association between longer duration of intake episode and worse outcome may represent greater efficacy of treatment received early in the course of the disorder.

#### Conclusions

Our results support that AN is a chronic disorder. While the health of most of the AN subjects improved over time, relatively few achieved a completely asymptomatic state. Of those who did achieve recovery, a substantial proportion subsequently relapsed. AN subjects at greatest risk are those with lowest weights and longest durations of illness. These patients will require more intensive treatments for lengthy periods of time.

The need for more progressive treatment strategies is also implicated by the substantial rates of relapse in both the AN and the BN subgroups. Our results indicate that course and outcome in AN differs from that in BN, supporting the current nosological classification of eating disorders. Future research regarding patterns and predictors of relapse may contribute to the development of appropriate clinical interventions.

## REFERENCES

- Collett D (1994), *Modeling Survival Data in Medical Research*. London: Chapman and Hall
- Collings S, King M (1994), Ten-year follow-up of 50 patients with bulimia nervosa. *Br J Psychiatry* 164:80–87
- Cox DR, Oakes D (1984), *The Analysis of Survival Data*. London: Chapman and Hall
- Doll HA, Fairburn CG (1998), Heightened accuracy of self-reported weight in bulimia nervosa: a useful cognitive distortion. *Int J Eat Disord* 24:267–273
- Eckert ED, Halmi KA, Marchi P, Grove E, Crosby R (1995), Ten-year follow-up of anorexia nervosa: clinical course and outcome. *Psychol Med* 25:143–156
- Fairburn CG, Norma PA, Welch SL, O'Connor M, Coll HA, Peveler RC (1995), A prospective study of outcome in bulimia nervosa and the long-term effects of three psychological treatments. *Arch Gen Psychiatry* 52:308–312
- Fichter MM, Quadflieg N (1997), Six-year course of bulimia nervosa. *Int J Eat Disord* 22:361–384
- Field AE, Herzog DB, Keller MB, West J, Nussbaum K, Colditz GA (1997), Distinguishing recovery from remission in a cohort of bulimic women: how should asymptomatic periods be described? *J Clin Epidemiol* 50:1339–1345
- Fleming T, Harrington D (1991), *Counting Process and Survival Analysis*. New York: Wiley
- Gibbon M, McDonald-Scott P, Endicott J (1981), Mastering the art of research interviewing: a model training procedure for diagnostic evaluation. *Arch Gen Psychiatry* 38:1259–1262
- Hebrebrand J, Himmelman GW, Herzog W et al. (1997), Prediction of low body weight at long-term follow-up in acute anorexia nervosa by low body weight referral. *Am J Psychiatry* 154:566–569
- Herzog DB, Field AE, Keller MB et al. (1996), Subtyping eating disorders: is it justified? *J Am Acad Child Adolesc Psychiatry* 35:928–936
- Herzog DB, Keller MB, Lavori PW, Kenny GM, Sacks NR (1992), The prevalence of personality disorders in 210 women with eating disorders. *J Clin Psychiatry* 53:147–152
- Herzog DB, Sacks NR (1993), Bulimia nervosa: comparison of treatment responders vs nonresponders. *Psychopharmacol Bull* 29:121–125
- Herzog DB, Sacks NR, Keller MB, Lavori PW, Von Ranson KB, Gray HM (1993), Patterns and predictors of recovery in anorexia nervosa and bulimia nervosa. *J Am Acad Child Adolesc Psychiatry* 32:835–842
- Keel PK, Mitchell JE (1997), Outcome in bulimia nervosa. *Am J Psychiatry* 154:313–321
- Keel PK, Mitchell JE, Miller KB, Davis TL, Crow SJ (1999), Long-term outcome of bulimia nervosa. *Arch Gen Psychiatry* 56:63–69
- Keller MB, Lavori PW, Friedman B et al. (1987), The Longitudinal Interval Follow-up Evaluation: a comprehensive method for assessing outcome in prospective longitudinal studies. *Arch Gen Psychiatry* 44:540–548
- Metropolitan Life Insurance Company (1983), Metropolitan height and weight tables. *Stat Bull Metrop Life Found* 64:2–9
- Morgan HG, Russell GF (1975), Value of family background and clinical features as predictors of long-term outcome in anorexia nervosa: a 4-year follow-up study of 41 patients. *Psychol Med* 5:355–371
- Norring CEA, Sohlberg SS (1993), Outcome, recovery, relapse and mortality across six years in patients with clinical eating disorders. *Acta Psychiatr Scand* 87:437–444
- Olmsted MP, Kaplan AS, Rockert W (1994), Rate and prediction of relapse in bulimia nervosa. *Am J Psychiatry* 151:738–743
- Ratnasuriya RH, Eisler I, Szmukler GI, Russell GF (1991), Anorexia nervosa: outcome and prognostic factors after 20 years. *Br J Psychiatry* 158:495–502
- Robins LN, Helzer JE, Croughan J, Ratcliff KS (1981), National Institute of Mental Health Diagnostic Interview Schedule: its history, characteristics, and validity. *Arch Gen Psychiatry* 38:381–389
- Rowland M (1990), Self-reported weight and height. *Am J Clin Nutr* 52:1125–1133
- Russell GF, Szmukler GI, Dare C, Eisler I (1987), An evaluation of family therapy in anorexia nervosa and bulimia nervosa. *Arch Gen Psychiatry* 44:1047–1056
- Steiner H, Lock J (1998), Anorexia nervosa and bulimia nervosa in children and adolescents: a review of the past 10 years. *J Am Acad Child Adolesc Psychiatry* 37:352–359
- Steinhausen HC, Rauss-Mason C, Seidel R (1991), Follow-up studies of anorexia nervosa: a review of four decades of outcome research. *Psychol Med* 21:447–451
- Strober M, Freeman R, Morrell W (1997), The long-term course of severe anorexia nervosa in adolescents: survival analysis of recovery, relapse and outcome predictors over 10–15 years in a prospective study. *Int J Eat Disord* 22:339–360
- Sullivan PF, Bulik CM, Fear JL, Pickering BA (1998), Outcome of anorexia nervosa: a case-control study. *Am J Psychiatry* 155:939–946
- Therneau TM (1994), *A Package for Survival Analysis in S*. Mayo Foundation, available from StatLib <http://lib.stat.cmu.edu/S/>