

Kächele H, Kordy H, Richard M (2001) Therapy amount and outcome of inpatient psychodynamic treatment of eating disorders in Germany: Data from a multicenter study. *Psychotherapy Research* 11: 239-257

Therapy Amount and Outcome of Inpatient Psychodynamic Treatment of Eating Disorders in Germany: Data From a Multicenter Study

Horst Kächele, Hans Kordy, Matthias Richard & Research Group TR-EAT
Center for Psychotherapy Research
Stuttgart, Germany

Correspondence to:

Prof. Dr. med. H. Kächele

Forschungsstelle für Psychotherapie

Christian-Belser-Str. 79a

70597 Stuttgart

Abstrakt

Die vorliegende Studie untersucht (1) welche Faktoren die Dauer und Intensität der Psychotherapie für Essstörungen bestimmen und (2) wie diese evtl. in Interaktion mit anderen Faktoren die Effektivität beeinflussen.

In einem naturalistischen Design wurde der Symptomstatus von 1171 Patientinnen über einen Zeitraum von 2,5 Jahren gemessen, nachdem sie in einer der 43 teilnehmenden Fachkliniken zur Behandlung aufgenommen wurden. Die statistische Modellierung erfolgte mit Verfahren der hierarchisch lineare Regression und logistischen Regression. Mögliche Verzerrungen durch konfundierende Variablen wurden mit Hilfe der Propensity-Score-Technik korrigiert.

Die Behandlungsmodalitäten, insbesondere Dauer und Intensität, variierten beträchtlich zwischen und innerhalb der Kliniken, wobei Patientenmerkmale nur eine geringe Rolle spielten. In der 2,5-Jahresuntersuchung zeigte sich für etwa ein Drittel der AN und etwa ein Viertel der BN eine klinisch bedeutsame Symptombesserung. Der Einfluß von Dauer und Intensität auf die Ergebnisse war schwach und nur in Interaktion mit weiteren Patienten- bzw. Klinikmerkmalen zu erkennen. Implikationen zur Behandlungsplanung werden diskutiert.

Abstract

The present study investigates (1) factors that determine length of treatment and (2) the effect of treatment duration, and other factors, on treatment outcome for patients with eating disorders.

It is an observational study. Symptomatic status of 1171 patients was observed for 2.5 years after admission to one of the 43 participating hospitals. Treatment and outcome were modeled by using hierarchical linear models and logistic regression. To control for possible confounding factors, propensity score adjustment was used.

Treatment modalities, especially length and intensity, varied considerably between and within hospitals and were related to patient characteristics to a very small degree. At 2.5 year follow-up, 33% of Anorexia patients and 25% of the Bulimia patients were symptom free. Length of treatment showed weak effect on outcome and only in interaction with other patient characteristics of relevance, while treatment intensity was not related to outcome. Implications for treatment planning are discussed.

Therapy amount and Outcome of Inpatient Psychodynamic Treatment of Eating Disorders in Germany: Data From a Multicenter Study

Anorexia nervosa (AN) and bulimia nervosa (BN) are predominantly female mental disorders arising in adolescence and early adulthood. The frequency of these illnesses, severity of their symptoms, the risk of chronicity, and seriousness of physical complications (Treasure and Smukler, 1995), and the related social costs underline the clinical interest and health political importance of these disorders.

Eating disorder - AN or BN - have been estimated to have a relatively high prevalence (between 1 to 4%). AN has the highest mortality risk of all mental disorder, with 20 % patients dying within 20 years of onset (Crisp et al., 1992). If those patients are considered who have eating disorders symptoms (EDNOS - Eating Disorders Not Otherwise Specified) although they do not fulfill diagnostic criteria for AN and/or BN, the rate of prevalence increases to 5-15% (Herzog et al., 1991).

Since the mid-1970's, the amount of research conducted with regards to eating disorders has increased. Standardization has been achieved in defining both symptoms of illness and diagnostic categories. Several treatments programs have been developed and there efficacy has been investigated. (e.g. Roth & Fonagy, 1996). Although the efficacy of some treatment programs has been supported empirically, especially treatment programs for BN (e.g. Fairburn et al. 1996) these studies show at the same time that much research remains to be done. Empirically validated treatments for BN result only in 50% of cases in long-term recovery (Mitchell et al., 1996). The chances for full recovery in patients with AN is much lower (Crisp et al., 1996).

Only a few consistent predictors of outcome have been consistently (Keel et al., 1999; Wilson et al., 1999; Bulik, 1998; Herzog et al., 1997). Empirical findings have not been significant enough to justify the use of one treatment over another (Herzog & Hartmann, 1997). Most studies were carried out under unique research conditions in specialty treatment centers, often with selected samples (Mitchell et al., 1997). Studies concerned with effectiveness of treatment in a naturalistic setting are sparse, as are studies that investigate

the relations between amount and outcome of treatment (Treasure & Schmidt, 1999; Herzog et al., 1996). Mitchell et al. (1993) reported a higher rate of success at the end of treatment when treatment intensity was increased (two sessions a week instead of one) holding constant the treatment duration. We are not aware of comparative studies of alternative inpatient treatment programs such as those that AN patients often receive (Treasure et al., 1995; McKenzie & Joyce, 1992). Although treatment programs are usually very intensive and involve comprehensive therapeutic resources, data concerning the effect of the composition of the inpatient treatment program, its duration, and intensity on outcome are practically non-existent.

We found only one study in the German-language literature that investigated the effect of treatment intensity on success of outcome (Herzog, Hartmann & Sandholz, 1996). Treatment duration (with the same number of therapy sessions) was not a significant variable. However, once again treatment intensity was an important factor. An increase in treatment intensity allowed for a comparable outcome to be reached within three months as opposed to twelve

Most studies concerned with the effectiveness of psychotherapy on eating disorders have investigated standardized treatment programs with pre-defined treatment duration and intensity. However, this does not reflect the naturalistic process of treatment, where the treatment is not defined at admission and it is possible for treatment to be expended or shortened. Conclusive findings that identify variables predictive of the length or intensity of treatment are largely not available. Thus, the question whether patients with severe symptoms receive the longer and/or more intensive treatments or whether and which other factors are involved can not be answered yet.

The Center for Psychotherapy Research in Stuttgart initiated a multicenter study on the effectiveness of inpatient psychodynamic treatment of eating disorders in the early 1990's. Besides this overall objective, the study aimed to estimate the amount of therapeutic resources that were applied within the various treatment programs in clinical everyday practice and the impact on the short and medium-term course of eating disorders. After the pilot phase, the main study began in 1993. It ended at the end of 1998 with the 2.5 year

follow-up assessment. Forty-three specialty hospitals and departments for psychosomatic medicine and psychotherapy in Germany participated in the data collection.

The present article addresses the main questions investigated by the study (see Kächele, 1992 for the research group TR-EAT): (1) What factors determine the length of inpatient treatment? (2) What is the effectiveness of inpatient psychodynamic therapy; i.e., what rates of recovery or improvement (measured from therapist's as well as patient's perspective) were reached and for which patients? (3) How do treatment duration and intensity contribute to the effectiveness? Can such effects be attributed to specific patient characteristics?

Methods

Study Design

Project TR-EAT is a naturalistic, longitudinal, and observational study. Treatment duration and intensity were not standardized in order to observe the naturalistic clinical course of treatment. Data collection took place from 1993 to 1998. Patients included in this study were at least 18 years and fulfilled diagnostic criteria for AN and/or BN at time of screening.

All patients were assessed over a time of 2.5 years (Figure 1). At therapy admission and discharge, as well as 1 year and 2.5 years after index admission, patients were questioned as to their physical condition, their mental state, and their level of psychosocial functioning using a comprehensive battery of inventories. At the same time, the condition of each patient was clinically evaluated at admission and discharge by their primary therapist. The 2.5 year assessment was conducted by clinical experts and the 1 year assessment was conducted by mail and thus was limited to self-evaluation. The questionnaires covered sociodemographic and historical variables; the battery of psychological inventories included the Symptom Checklist 90-R (SCL 90-R), the Eating Disorder Inventory (EDI), the "Freiburg Personality Inventory" (Freiburger Persönlichkeitsinventar, FPI-R), the „Narcissistic Personality Inventory,, (Narzissmus Fragebogen, NARZ) and the "Parental Care Index" (Familien-Klima-Skalen, FKS). Treatment exposure was measured using weekly documentation of the frequency and duration of all psychotherapeutic contact. In addition, the occupation of the

participant was considered (to estimate therapy cost), and the number of participants in group or family therapy was monitored. The 2.5 year follow-up assessment was completed using a semi-structured interview, the Longitudinal Interval Follow-up Evaluation (LIFE, Keller et al., 1987) adapted for use in present study. Retrospective, longitudinal information on symptomatic disturbance (the so-called "change points") during the post-discharge course of the illness was obtained using LIFE. This information was used to track the course of recovery and relapse between the point of discharge and the 2.5 year follow-up interview (Kordy et al., in press).

 Insert Figure 1 about here

Subjects

Treatment of 1171 patients from 43 sites in Germany were examined. Diagnostic criteria for AN were fulfilled by 355 patients, 647 patients fulfilled criteria for BN, and 169 patients fulfilled diagnostic criteria for both disorders. The majority were female. The proportion of male patients was under 4% (AN: 3.3%, BN: 2.3%). The mean age of the participants was 24.8 years for AN ($SD = 5.6$) and 25.9 years for BN ($SD = 6.3$). The mean duration of illness prior to admission was 8.2 years ($SD = 6.1$) for BN and 5.7 years ($SD = 5.3$) for AN. Anorexic patients were 72.1% of expected body mass index ($SD = 8.4$).

At the 2.5 year follow-up, 879 patients (75.1% of sample) could be contacted. A comprehensive interview was conducted with 781 patients (66.7%). Limited information given by family or doctors was available for 98 patients (8.4%). Of the 292 patients who could not be reached, only 64 declined participation. It was not possible to establish contact with the other 207 patients in spite of repeated attempts to contact them via post and telephone. For various other reasons, no data was available for 11 patients. In the course of our study, 10 patients died: 6 patients through suicide, 1 patient as a result of medical complications related to the illness, and 3 patients of an unknown cause.

Few differences were found between patients who participated in this 2.5 year study and those who did not. Anorexic patients who participated received two more weeks of therapy on average, had a slightly higher body mass index at time of discharge, and were not as motivated to partake in treatment at admission. Patients diagnosed with BN who participated in this study had a slightly higher body mass index (BMI) at index admission and had a longer history of illness on average. They also scored higher on the „phobic anxiety,, scale (SCL-90) and slightly lower on the Morgan-Russell „socio-economic state (E),, scale. Overall, there were few indications of possible bias resulting from non-randomization; one eventual consideration is that the BN sample at the 2.5 year follow-up assessment tended towards chronicity.²

Definitions of Recovery

The diagnostic criteria used to measure outcome were largely based on data concerning symptoms specifically related to eating disorders. This data was obtained through self-report questionnaires and therapist assessments. The three required characteristics for AN (underweight, fear to gain weight, and body perception distortion) and BN (recurrent episodes of binge eating, recurrent compensatory behavior to prevent weight gain, and overconcerns with body and shape) were operationalized according to specific criteria. Outcome was evaluated as successful when at least 2 of the above diagnostic criteria (within each disorder), including underweight for AN and binge episodes for BN, were no longer present at the end of treatment and/or at the follow-up assessment (Table 1). This operationalization orients on diagnostic criteria used in other international studies (Strober et al. 1997, Mitchell et al. 1996, Fichter et al. 1994).

Table 1

Criteria for definition of therapy success

Symptom Domain	Symptom no longer significant at assessment points
Anorexia Nervosa	
Low Body Weight ^t	Weight 85% of expected BMI

Fear of Weight Gain: P^b: EDI "drive for thinness" ≤ 4.5 ,
T^b: \leq "light" ("not at all" to "very strong" in 5 steps)
Body Image Distortion: P^b: estimated vs. real %BMI $\leq 10\%$,
T^b: \leq "light" ("not at all" to "very strong" in 5 steps)

Bulimia Nervosa

Binge eating^a: Less than 2 episodes a week
Weight reducing methods: No vomiting, no use of laxatives and no strict dieting

Weight and shape concern: \leq "sometimes" ("never" to "always" in 5 Steps)^c

^a Main symptom. ^b P = Patient evaluation. T = Therapist evaluation. ^c The question was: "Do you pay a lot of attention to your figure?"

While in most studies underweight is defined by a fixed threshold, usually 17.5 BMI, we use German norm, which allow to take the age into account (Hebebrand et al., 1994; Oehlschlegel, Malewski & Mahon, 1999). Thus the age specific expected weight is set to 100 %. A 25 years old woman with an BMI of 17.5 would be of 84% of expected weight for her (his) age group.

Predictor Variables

The variables included in our study were chosen to cover several central domains and tend to correlate only moderately (Table 2). These variables were investigated in association with treatment duration and outcome. Covered domains included the history of the illness, pre-treatment syndrome status, treatment parameters such as duration and intensity. Hospitals were classified according to the type: specialty eating disorder hospitals, hospitals with a separate eating disorders ward, hospitals with a specialized eating disorders treatment program, and hospitals without treatment settings specifically for patients with eating disorders. Treatment intensity was measured using the hours of therapy from weeks 3 and 4. Hours of therapy were separately measured within two subsets: standard therapy hours (individual and/or group sessions) and extra-therapeutic activity (e.g., ergotherapy, sports therapy, dance therapy, and art therapy. Weeks 3 and 4 were chosen since treatment

intensity at this stage was expected to be most representative of the treatment intensity for the entire program.

Table 2

Overview on predictors investigated

Predictor	Operationalization
<i>Patient Data at Index Admission</i>	
motivation	not motivated - motivated
age	in years
impairment due to Eating Disorders	patient rating, 1= "low" to 5 = "high"
low body weight	weight in % of the mean expected BMI
diagnosis	AN, BN
Eating Disorder Inventory (EDI)	sum of the subscales 1,2 and 3
personality: (FPI)	QS ^a deviations from norm (all scales)
AN and/or BN Symptoms	criteria fulfilled by: none - one – both – all
sick leave	number of days in the past year
multi-impulsivity	substance abuse, aggression, impulsivity (ratings)
Narcissism Inventory	total score
duration of illness	years (logarithmic)
sexual abuse	none known – suspected - present
psychological distress	GSI score (SCL-90-R)
social functioning	Morgan–Russell "socioeconomic state" scale
number of previous treatments	none – one – two – three and more
health care insurer	private – public: BfA, DAK, AOK – others

Table 2 (continued)

<i>Treatment parameters</i>	
duration	short (≤ 11.5 Weeks), long (> 11.5 Weeks)
number of psychotherapy sessions (individual or group)	as documented from weeks 3 and 4
number of sessions of adjunctive therapy (e.g., ergo- or sportstherapy)	as documented from weeks 3 and 4
<i>Type of hospital</i>	
classification of the hospitals	degree of Specialization (see text)
<hr/>	
^a square sum.	

Analyses

The analysis is divided into two main parts: the first is concerned with variables that predict treatment duration, and the second with variables that are predictive of outcome. Associations between the selected predictor variables and inpatient treatment duration (Table 2) and the effect of the main treatment parameters "duration" and "intensity" on outcome rates were investigated. For analyses of outcome, the rate of positive outcome was the dependent variable. To control for possible bias due to confounding variables, estimates were adjusted using the propensity score (Rosenbaum & Rubin, 1983, 1984).

Analysis of Treatment Duration

Treatment duration varies within each clinic and between the various clinics. To test for possible predictive variables, a hierarchical linear modeling was used (HLM; Bryk & Raudenbush, 1992; Goldstein, 1995). This method makes use of a multilevel regression analysis to investigate treatment assignment and treatment outcome. A hierarchy exists between patient and clinic with patients nested within one and only one clinic. On the first level differences within hospitals were modeled, on the second level, differences between hospitals were investigated. Variance in treatment duration within and between hospital was estimated in the first model, which did not include any additional predictor. A sequence of models was fitted by the gradual refinement of this first model: at a second step, patient

characteristics were added as covariates. Finally, at the third step, characteristics of the treatment centers were added.

Analysis of Rate of Recovery and Relapse

The main purpose of this analysis was to test for the effect of treatment duration and intensity on outcome success rates after 2.5 years. Due to the naturalistic design of the study, confounding variables could bias the estimated effect of these predictors on outcome.³ To control for confounding variables, a method developed by Rosenbaum and Rubin (1983) was used. The "propensity score" was developed in the 1980's to remove possible bias in studies conducted without the benefit of randomization and to compare two treatment alternatives. The procedure can be divided into two steps: In the first step, each patient is given a propensity score using logistic regression or a discriminant analysis. The propensity score is the conditional probability of assignment to treatment A or B (here shorter or longer treatment) given the observed covariate information. In the second step, the sample is subclassified into (e.g.) quintiles of the distribution of the propensity score to form subgroups with rather similar propensity scores. As Rosenbaum & Rubin (1983, 1984) showed, the two treatment alternatives within the quintiles balanced with regards to the confounding variables (given that there are no predictor variables that clearly distinguish between the two treatments). Thus an unbiased estimate (or with relatively small bias) of the effect of both treatment alternatives is possible

Logistic regression was used for the statistical analysis of predictors of outcome (Hosmer & Lemeshow, 1989), in which the propensity score method was again applied for removing bias (D'Agostino, 1998). Due to the large number of possible predictor variables, there is a risk of over-specification. Therefore, the proportion of variance explained was corrected by utilizing a bootstrap method. The same regression model was constructed with a random sample of two-thirds of the initial sample and differences between the sampled \underline{R}^2 and the \underline{R}^2 of the entire model were estimated. This procedure was repeated 200 times. The mean of these differences yields an adjustment for the full model \underline{R}^2 that corrects for bias (Efron & Tibshirani, 1993).

Results

The following results are based on a sample of 1112 patients (AN: n=338, BN: n=605, AN&BN: n=169) from 43 hospitals, of which 733 (AN: n=225, BN: n=399, AN&BN: n=109) participated in the 2.5 year follow-up assessment. In order to allow for a consistent interpretation, data from 59 patients were excluded: 20 of the 59 patients received primarily outpatient treatment, the other 39 received less than 4 weeks of inpatient treatment. A treatment program of less than 4 weeks was not offered by any of the clinics, therefore, the treatment of these 59 patients cannot be considered representative for any of the programs and thus are drop outs.

Which patients received which length of treatment?

The mean treatment duration for all 3 subclasses of eating disorders was roughly 11 weeks (AN: Median=11.1; BN: Median=11.4; AN&BN: Median=10.6). Twenty-five percent of patients were treated for 4 to 8 weeks; a further 50% were treated for 9 to 13 weeks. Treatments with a duration of 15 weeks were rare, and those that continue for more than six months were exceptionally rare.

Insert Fig. 2 about here

Figure 2 depicts both sources of variance: the length of each individual boxplot indicates the variance within each clinic, and the differences in length between the boxplots illustrate variance between the various hospitals. Only 3 hospitals had a median treatment duration of 20 weeks or longer (Hospitals 53, 23, 19). A large majority of the hospitals had a median treatment duration of 8 to 12 weeks. A few hospitals showed high variability in treatment duration (Hospitals 53, 19, 5), while in other hospitals the length of treatment varied only slightly (Hospitals 6, 38, 11, 26).

To investigate which variables defined treatment duration, the stepwise linear model was used. The construction of the model required 3 steps. The results are presented in Table 3 and Table 4. Variances within and between hospitals were estimated using a first model without predictors⁵ (Model 1 in Table 3), also known as an "unconditional model" (Bryk &

Raudenbush, 1992). The effect of the patient-related variables on treatment duration is presented in the second column (Model 2). Finally, in the third step, data regarding the specialty status of the treatment offered was included (Model 3). As can be seen in Table 3, patient characteristics accounted for only a small percentage of variance within the hospitals. These patient variables did not explain variance between the hospitals.

Table 3.
Proportion of variance explained of treatment duration within and between hospitals

	Model 1 ^a	Model 2 ^b	Model 3 ^c
	σ^2	σ^2 explained σ^2	σ^2 explained σ^2
	(SE)	(SE)	(SE)
Between	19.19 (4.69)	19.95 – (4.91)	14.75 23.1% (3.69)
Within	20.44 (0.96)	19.78 3.2% (0.94)	19.33 2.3% (0.91)

^awithout predictors ^bpatient characteristics as predictors ^chospital characteristics.

Detailed results for Model 3 can be found in Table 4, which lists all variables with a statistically significant effect. The strongest effect was found in hospitals with a separate eating disorders ward. When compared to a specialty eating disorders hospital, their patients received 7 additional weeks of treatment on average. A few patient characteristics were moderate predictors of outcome. Patients with low treatment motivation at baseline received half a week less treatment on average than patients with higher motivation.

Although the effect of 4.06 for BN diagnosis appears significant, it does not indicate that BN was treated 4 weeks longer than AN on average. Three further significant interactions are involved in interpreting diagnosis effect. For anorectic patients a difference of 1 point on the Global Severity Index, GSI, (SCL-90-R) correlates with an extended treatment length of half a week, while for bulimic patients the same difference on the GSI leads to .24 weeks shorter treatment due to a GSI by BN interaction. A weight gain of 16% BMI, which equals an increase in weight from the upper diagnostic boundary of 17.5 BMI to the

expected body mass index of 20, results in a treatment shortened by 0.1 weeks for anorexic patients. This correlation is even stronger for bulimic patients; a weight gain of 16% BMI results in a treatment shortened by 0.75 weeks. Since bulimic patients have a higher weight at baseline, the effect is even more pronounced. The treatment duration of very overweight BN patients (>130 %BMI) is approximately 2 weeks shorter than that of BN patients with expected body mass index. For anorexic and bulimic patients with 90% of expected body weight, a difference of only 0.56 weeks is found. These examples underscore once more, the role of interaction effects.

Table 4

Selected predictors of treatment duration

Predictor	B	SE
Other vs. specialty hospital	0.42	3.95
Specialized program vs. specialty hospital	-0.33	3.98
Specialized ward vs. specialty hospital	7.29	4.36 ⁺
Motivation to change = No	-0.49	0.16 ^{**}
Psychological distress (GSI, SCL – 90-R)	0.50	0.18 ^{**}
Age	0.013	0.016
Weight (in % of expected BMI)	-0.006	0.004
Diagnosis = BN	4.06	1.32 ^{**}
Weight \times diagnosis (BN)	-0.05	0.01 ^{**}
Psychological distress \times diagnosis (BN)	-0.74	0.23 ^{**}
Age \times diagnosis (BN)	0.05	0.03 [*]

⁺ $p < .1$, ^{*} $p < .05$, ^{**} $p < .01$

For which patients does inpatient psychodynamic therapy provide which rate of recovery, and is treatment duration a predictive factor?

Outcome rates at the end of treatment and at the 2.5 year follow-up assessment are presented in Table 5. Essentially, the present definition of treatment success implies an almost complete lack of symptoms or only symptoms for which immediate further

treatment is not necessary. Positive outcome thus defined (Table 1) was found in a minority of AN patients and patients who met both AN and BN criteria: 11% and 17% respectively at the end of treatment. Self-report evaluations and therapist's evaluation did not differ substantially. Positive outcome rates for BN were markedly higher. However, outcome success rates as assessed in therapist evaluations (45%) differed greatly from self-evaluations (31%). At the 2.5 year follow-up assessment, a significantly greater proportion of patients with AN (36% based on therapists reports and 33% based on patients reports) and with AN and BN (26% versus 21%) were to a large extent symptom free. However, the rate of positive outcome for BN dropped slightly (36% based on therapists reports versus 22% based on patients reports).

Table 5

Rate of success (in %) at discharge and at 2.5 years follow-up classified according to treatment duration

Treatment duration		Discharge				2.5 years Follow-up			
Perspective	≤ 11 Weeks		> 11 Weeks		≤ 11 Weeks		> 11 Weeks		
	Total		OR ^a		Total		OR ^a		
Anorexia Nervosa									
	n=166		n=170		n=112		n=113		
Patients	7.8		13.4	10.6	1.1	34.8	31.0	32.9	
Therapists	9.6		15.1	12.4	1.8	36.6	34.5	35.6	
Bulimia Nervosa									
	n=292		n=303		n=197		n=202		
Patients	33.0		29.0	31.0	0.7 ^b	20.6	24.0	22.2	
Therapists	48.0		43.0	45.4	0.6	35.5	37.1	36.3	
Anorexia & Bulimia Nervosa									
	n=73		n=95		n=44		n=65		
Patients	13.7		15.8	14.9	1.2	27.3	16.9	21.1	
Therapists	13.7		19.0	16.7	1.5	31.8	21.6	25.7	

^a Odds ratios after propensity score adjustment (with the exception of 'Anorexia & Bulimia Nervosa').

^b 1 within the 95% confidence interval.

The propensity score was used to estimate the effect of treatment duration on outcome, thus reducing potential bias due to confounding variables (see Method). This method compares only two alternatives: In this study the effect of short and long treatment was compared. The treatments are thus dichotomized at the median (AN: 11.1 weeks; BN: 11.4 weeks). With the exception of treatment duration and intensity, all variables listed in Table 2 are included in the propensity score adjustment. Figure 3 shows the effect using the example of low body weight in AN. As can be seen, the mean BMI of shorter and longer treated patients is not significantly different within each quintile. The same effect can be applied to the other variables. Overall, the propensity score reduces the bias of confounding variables by 72% for AN and 88% for BN (Liebermann et al., 1996). It is important to note that hospital effects cannot be controlled by this method. As some of the hospitals used only short treatments, it is inevitable that a certain proportion of shorter treatment patients were in different hospitals than those who received longer treatment. Therefore, the potential hospital effect could not be counterbalanced for treatment duration.

 Insert Figure 3 about here

The comparison of the success rate for shorter and longer treatments is demonstrated by the odds ratios⁶ with propensity score adjustments in Table 5. The adjusted odds ratios indicate that a patient-reported difference in outcome at the end of treatment was found only in BN patients. For patients with bulimia, the patient-reported positive outcome rate for shorter treatments was 33% higher than the rate for longer treatments. Patients with anorexia had a therapist-reported success rate that was about 80% higher for longer term versus shorter term cases. However, this difference in success rates was not statistically reliable (despite the considerable sample sizes). Overall, there was no recognizable difference in outcome between shorter and longer treatment at the 2.5 year follow-up assessment. This

analysis does not preclude the possibility that unconsidered variables may interact with treatment duration and intensity in predicting outcomes.

Correlations between eating disorder treatment outcome and predictors

The following analyses aim on the identification of possible predictors of treatment success at the 2.5 year follow-up.⁷ The present analysis made use of a stepwise multivariate logistic regression method⁸. No substantial differences between outcome at the various hospitals could be found. The analysis of the predictive variables of treatment success used a simple logistic regression approach, excluding the „hospital,, factor from the design.

Analyses were conducted within AN and BN groups, separately. The variables listed in Table 2 were investigated as predictors of outcome status. The propensity score and treatment duration, longer-term versus shorter-term, were also included. In order to gain an accurate picture of the possible association between treatment duration and outcome, all predictors were examined in interaction with treatment duration. Only those interactions that contributed significantly to the model were utilized further. The outcome criterion was based on patient reports. The resulting models are described in Table 6. The resulting models for AN and BN differ with regard to the identified predictors as well as to the goodness of fit. It is of particular interest that treatment length was found to interact with different variable sets for the AN analyses when contrasted with the BN analyses.

Table 6

Predictors of therapy success (patient report)

AN ^a			BN ^b		
predictor	<u>B</u>	<u>SE</u>	predictor	<u>B</u>	<u>SE</u>
propensity pcore	1.02	0.80	propensity score	-0.07	1.04
short vs. long treatment	-1.41	0.76 ⁺	short vs. long treatment	3.90	1.52 [*]
age ^c	-1.98	0.68 ^{**}	age	0.03	0.03
personality (FPI "Extreme")	-0.75	0.38 [*]	previous treatments	-0.33	0.13 [*]
weight (% expected BMI)	0.07	0.02 ^{**}	EDI Scales 1+2+3	-0.02	0.01 ⁺
desired body weight (self-report)	0.07	0.02 ^{**}	anorectic symptoms: yes(vs. none)	-1.27	0.44 ^{**}

treatment duration \times age	0.71	0.41 ⁺	multi-impulsivity	-0.53	0.26 [*]
			Morgan-Russell	0.04	0.08
			„socioeconomic state,,		
			treatment duration \times age	-0.10	0.05 [*]
			treatment duration \times		
			Morgan-Russell „socioeconomic	-0.21	0.11 [*]
			state,,		
^a R^2 (without bootstrap) = .31, R^2 (with bootstrap) = 0.25. ^b R^2 (without bootstrap) = .13, R^2 (with bootstrap) = 0.07. ^c logarithmic association. ⁺ $p < .1$, [*] $p < .05$, ^{**} $p < .01$					

Overall, the goodness of fit was moderate for both models. A proportion of 31 % of variance was explained by the model for AN. This was reduced to 25% after conducting a bootstrap adjustment and was thus relatively reliable. The model for BN showed 13 % variance explained. This was reduced to 6.5% after a bootstrap adjustment, thus presenting a low reliability.

Even after controlling for other covariates, treatment intensity had no effect on the outcome for patients with AN. However, treatment length (short vs. long) in interaction with the age variable was a significant predictor. Low body weight and low desired body weight (%BMI) at treatment admission, and significant deviations as measured on the Freiburg personality inventory (FPI) were risk factors associated with a poor outcome. The interaction between treatment duration and age is shown as an example in Figure 4: Overall, high patient age indicates a poorer outcome. The predictive value of this variable was intensified in the outcome of younger patients: In the range from 18 to 20 years, patients had 2 to 4 times higher rates of positive treatment outcome with shorter treatments than patients of the same age with longer treatments. This correlation is reversed for older patient outcomes – those patients who received a longer treatment actually had an increased rate of good outcome. However, this rate as a whole was considerably lower than that of younger patients.

Insert figure 4 about here

The results for BN are quite complex: multi-impulsivity, additional anorectic symptoms, and high number of previous treatments were associated with lower success rates. Treatment success was only slightly lower for patients who scored high on the first three EDI subscales, i.e., those with a more severe eating disorder syndrome. The effect of treatment duration is complicated by interaction effects with age and with the Morgan Russell subscale E ("socioeconomic state"), in which relationships to family and friends are presented. The association between the three variables is shown in figure 5: Patients with difficulty in establishing relationships (low Morgan-Russell scores) had a better prognosis when treated for a longer rather than a shorter period of time. If patients scored at least moderately on psychosocial functioning, the reverse effect resulted in the good outcome increasing with shorter treatment. The advantage of longer treatment for outcome was stronger for younger patients with poor social adjustment than for older patients, whereas for older patients with good psychosocial functioning the advantage of shorter treatment was stronger than for younger patients.

Insert Fig. 5 about here

Discussion

The treatment of eating disorders remains challenging. Although our results are significant, some findings reflect the difficulties that are inherent to the psychodynamic treatment of eating disorders. The success rates at the end of treatment were consistent, to some degree, with rates achieved under research conditions in other studies. Mitchell et. al.'s (1996) overview of the effect of psychotherapy on BN estimated that 40% to 60% of patients were largely free from symptoms at the point of discharge. Fichter et al. (1994) reported that 54% of bulimic patients who received a mean 12-week inpatient behavioral therapy were rated as having a good outcome 2 years after discharge. Our study showed a lower rate of symptom remission at the 2.5 year follow-up assessment. Furthermore, the

differences of the success rates between discharge and 2.5 years follow-up indicate a high rate of relapse and characteristic of a cyclical course of illness with recurring symptomatic phases. The somewhat higher rate of recovery found by Fichter et al. could be related to the behavioral treatment's focus on symptom and relapse management, thus patients may have been better prepared for maintaining treatment gains following hospitalization. However, this hypothesis requires further testing.

Not very many treatment studies of anorexia nervosa have been conducted, and the results of existing studies are difficult to interpret due to various problems, especially treatment as well as study dropout (van Strien et al., 1992; Crisp et al., 1992). In this respect, a solid empirical basis is lacking (Roth & Fonagy, 1996). Recently published results of a long-term study shift attention to outcomes over time: Strober et al. (1997) reported that 21% and 33% of adolescent anorectic patients no longer met the main diagnostic criteria of low body weight and amenorrhoe two and three years after admission, respectively. One percent and 9% of patients were free of all criterion symptoms 2 and 3 years after admission, respectively. Comparisons with such studies can be problematic, as definitions of successful outcome or full recovery are subject to differing criteria. The definition of full recovery by Strober, et. al. (1997) and Fichter & Quadflieg (1997) are consistent with that of the Project TR-EAT in that the criterion of symptom-free status was employed, even when this criterion was operationalized differently.

The interpretations of findings concerning the effect of treatment duration and intensity on outcome are rather challenging. Differences in the length of treatment assigned can only be partly explained through patient characteristics or sociodemographic status at admission. Obviously, hospital-specific resource management factors (largely based on local treatment options, therapist's preferences, and experiences and only marginally based on external, empirical studies) have great effect on determining the length of treatment. The high mean treatment duration in hospitals with a separate eating disorders ward should be critically examined to determine whether higher rates of recovery truly accompany the high resource investment in such settings.

If the syndrome status of the patient does not correlate with differences in the length of treatment assigned, to what extent can a possible increase in rate of recovery justify the implementation of a longer therapy? Success rates were almost equal for short and long treatment when they have been controlled for possible confounding variables. Only a more detailed analysis allows for the identification of subclasses, in which the treatment duration is predictive of treatment outcome.

The pretreatment characteristics of anorexic patients are slightly predictive for treatment outcome. As one would expected, a severely low body weight at admission is related to outcome. Severely underweight AN patients although on average, treated for a longer period had a poorer rate of recovery at the 2.5 year follow-up assessment than patients who had a higher BMI. The longer inpatient treatment duration could be related to the time required to reach target weight. Similar results were found when investigating the effect of age and the number of previous treatments; both variables can be used to indicate the degree of chronicity. Patients who display a higher level of chronicity benefit more from longer treatments, while patients who are not chronically ill benefit more from shorter treatments. However, patients who were around 25 years (the mean age of the sample) have similar predictions of outcome for both longer and shorter treatments. The higher rate of success in younger patients suggests that a long-term treatment outcome of full symptom remission is feasible for these individuals, whereas this target is possibly set too high for older and chronically ill patients. These individuals could target reaching a level of symptomatic that allows them to lead a normal life with reasonable level of life satisfaction.

The model for the prediction of positive outcome with only 6.5 % variance explained is not very impressive, and calls for caution in interpreting the identified predictors especially because of the low stability of the model Severe multi-impulsivity, a high number of previous treatments, and additional anorectic symptoms effect the rate of success adversely. Treatment duration in interaction with psychosocial functioning and seems to be an indicator of outcome: The probability of a good outcome increased for those patients with good social adjustment. This effect was intensified in older patients and suggests that long treatment, i.e., long absence, has a reduced effect for patients with good social

functioning, especially when those patients are married. However, this effect reverses for young patients with difficulties establishing relationships: they do not have these problems in a longer treatment.

The participation of such a large sample of hospitals in Germany is not only an advantage. This also limits possible conclusions. For example, the differences in patient mix between the hospitals limits the interpretation of the found effect of the specialty level of the treatment. The possible bias for this factor alone or in association with the various clinical settings could not be adjusted with the propensity score method (since that would have required that all hospitals apply longer as well as shorter treatments).

Of particular note for the findings regarding treatment length is the range that was used in this study: the duration of treatments investigated run from 5 to approximately 16 weeks. Treatments shorter than 5 weeks were excluded from the analyses as it was uncertain to what extent the duration was intended and what proportion could be attributed to dropouts. Treatments with duration of more than 16 weeks were not found frequently enough to justify inclusion. Furthermore, the possible effect of many variables not included in the model remained unobserved, e.g., parameters for the course of illness not included in this study, further characteristics of the hospitals, or parameters for the treatment program. Other possible factors within the follow-up period, such as the effect of further outpatient treatment or critically decisive life events, were not included in this study – and thus limit generalization.

In spite of these limitations, suggestions for the treatment of eating disorders can be made. The results of this study suggest that inpatient psychodynamic treatment should have a duration of at least 8 to 12 weeks, whereas a longer treatment for older patients over 40 could be beneficial. Further treatment extensions should be based on individual cases as well as on the course of improvement. The decision for further inpatient or outpatient treatment should not be independent of patient status at point of discharge. Rather, new and specified treatment methods should be developed. The results of the medium-term course of eating disorders (Kordy et al., in press) suggest an increased risk of relapse in the period immediately following hospitalization. This could be counteracted by a maintenance

treatment. Booster therapies could be such an approach to reduce the rate of relapse that occurs a few months after discharge. Stepwise care provision could serve as guiding principle for the development of comprehensive treatment strategy (Royal College of Psychiatrists, 1992). Experiences with such strategic approach are positive in those countries that do not separate inpatient and outpatient treatment that strictly as it is tradition in Germany.

References

- Bryk, A.S. & Raudenbush, S.W. (1992). Hierarchical Linear Models. Newbury Park: Sage.
- Bulik, C.M., Sullivan, P.F., Joyce, P.R., Carter, F.A. & McIntosh, V.V. (1998). Predictors of 1-year treatment outcome in bulimia nervosa. Comprehensive Psychiatry, 39(4), 206-214.
- Crisp, A.H., Callender, J.S., Halek, C. & Hsu, L.K. (1992). Long-term mortality in anorexia nervosa. A 20-year follow-up of the St. George's and Aberdeen cohorts. British Journal of Psychiatry, 161, 104-107.
- Crisp, A.H., Joughin, N., Halek, C. & Bowyer, C. (1996). Anorexia nervosa: The wish to change. Hove, East Sussex: Psychology Press.
- D'Agostino, R.B. (1998). Tutorial in biostatistics: propensity score methods for bias reduction in the comparison of a treatment to a non-randomized control group. Statistics in Medicine, 17, 2265-2281
- Efron, B. & Tibshirani, R.J. (1993). An introduction into the bootstrap. New York: Chapman & Hall.
- Fairburn, C.G., Norman, P.A., Welch, S.L., O'Connor, M.E., Doll, H.A., Peveler, R.C. (1995). A prospective study of outcome in bulimia nervosa and the long-term effects of three psychological treatments. Archives of General Psychiatry, 52(4), 304-312.
- Fichter, M.M., Quadflieg, N. & Reif, W. (1994). Course of multi-impulsive bulimia. Psychological Medicine, 24, 591-604.
- Fichter, M.M. & Quadflieg, N. (1997). Six-year course of bulimia nervosa. International Journal of Eating Disorders, 22, 361-384.
- Goldstein, H. (1995). Multilevel Statistical Models. (2. Edition) London: Arnold.
- Hebebrand, J., Hesecker, H., Himmelmann, W., Schäfer, H. & Remschmidt, H. (1994). Altersperzentile für den Body-Mass-Index aus Daten der Nationalen Verzehrstudie einschließlich einer Übersicht zu relevanten Einflußfaktoren. Aktuelle Ernährungsmedizin, 19, 259-265.
- Herzog, D.B., Keller, M.B., Lavori, P.W. & Sacks, N.R. (1991). The course and outcome of bulimia nervosa. Journal of Clinical Psychiatry, 52 (10, suppl.), 4-8.
- Herzog, T., Hartmann, A. & Sandholz, A. (1996). Psychotherapiedauer und Psychotherapiedosis. Die Freiburger prospektive kontrollierte Studie zur Kurz-Psychotherapie der Bulimia Nervosa. In: Hennig, H., Fikentscher, E., Bahrke, U. et al. (Hrsg.). Kurzzeit-Psychotherapie in Theorie und Praxis. Pabst Science Publishers, Lengerich, Vienna, Zagreb.
- Herzog, T. & Hartmann, A. (1997). Psychoanalytisch orientierte Behandlung der Anorexia nervosa. Eine methodenkritische Literaturübersicht unter Verwendung meta-analytischer Verfahren. Psychotherapie, Psychosomatik, Medizinische Psychologie, 47, 263-280.

- Herzog, W., Deter, H.C., Fiehn, W. & Petzold E. (1997). Medical findings and predictors of long-term physical outcome in anorexia nervosa: a prospective, 12-year follow-up study. Psychological Medicine, 27(2), 269-279.
- Hosmer, D. & Lemeshow, S. (1989). Applied Logistic Regression. New York: Wiley & Sons.
- Kächele, H. für die Studiengruppe MZ-ESS. (1999). Eine multizentrische Studie zu Aufwand und Erfolg bei psychodynamischer Therapie von Eßstörungen. Studiendesign und erste Ergebnisse. Psychotherapie, Psychosomatik, Medizinische Psychologie, 49, 100-108.
- Keel, P.K., Mitchell, J.E., Miller, K.B., Davis, T.L. & Crow, S.J. (1999). Long-term outcome of bulimia nervosa. Archives of General Psychiatry, 56, 63-69.
- Keller, M.B., Lavori, P.W., Friedman, B. et al. (1987). The Longitudinal Interval Follow-up Evaluation. A comprehensive method for assessing outcome in prospective longitudinal studies. Archives of General Psychiatry, 44, 540-548.
- Kordy, H., Krämer, B., Palmer, R.L., Papezova, H., Pellet. J., Richard, M., Treasure J. & COST Action B6 (in press). Conceptualisation and rationale for consensus definition in eating disorders: Response, remission, recovery, relapse and recurrence. Journal of Clinical Psychology.
- Liebermann, E., Lang, J.M., Cohen, A., D'Agostino, Jr. A., Datta, S. & Frigoletto, F.D. (1996). Association of epidural analgesia with caesareans in nulliparous women. Obstetrics and Gynecology, 88, 993-1000.
- Littell, R.C., Milliken, G.A., Stroup, W.W. & Wolfinger, R.D. (1996). SAS[®] System for Mixed Models. Cary, NC: SAS Institute Inc., (Kap. 8, 11).
- McKenzie, J., Joyce, P. (1992). Hospitalisation for anorexia nervosa. International Journal of Eating Disorders, 11, 235-241.
- Mitchell, J.E., Pyle, R.L., Pomeroy, C. et al. (1993). Cognitive behavioral group psychotherapy of bulimia nervosa: the importance of logistical variables. International Journal of Eating Disorders, 20, 277-287.
- Mitchell, J.E., Hoberman, H.N., Peterson, C.B., Mussell, M. & Pyle, R.L. (1996). Research on the psychotherapy of bulimia nervosa: half empty or half full. International Journal of Eating Disorders, 20, 219-229.
- Mitchell, J.E., Maki, D.D., Adson, E.E., Ruskin, B.S. & Crow, S. (1997). The selectivity of inclusion and exclusion criteria in bulimia nervosa treatment studies. International Journal of Eating Disorders, 22, 243-252.
- Oehlschlägel, J., Malewski, P. & Mahon, J. (1999). How to define anorectic weight? European Eating Disorders Review, 7(5), 321-333.
- Rosenbaum, P.R. & Rubin, D.B. (1983). The central role of the propensity score in observational studies for causal effects. Biometrika, 70 (1), 41-55.

- Rosenbaum, P.R. & Rubin, D.B. (1984). Reducing bias in observational studies using subclassification on the propensity score. Journal of the American Statistical Association, 79 (387), 516-524.
- Roth, A. & Fonagy, P. (1996). What works for whom? A critical review of psychotherapy research. New York: Guilford Press.
- 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. International Journal of Eating Disorders, 22, 339-360.
- Treasure, J. & Smukler, G.I. (1995). Medical complications of chronic anorexia nervosa. In: Smukler, G.I., Dare, C. & Treasure, J. (eds.). Handbook of Eating Disorders: Theory, Treatment and Research. New York: Wiley & Sons, 197-220.
- Treasure, J., Todd, G. & Smukler, G.I. (1995). The inpatient treatment of eating disorders. Smukler, G.I., Dare, C. & Treasure, J. (eds.). Handbook of Eating Disorders: Theory, Treatment and Research. New York: Wiley & Sons, 275-291.
- Treasure, J. & Schmidt, U. (1999). Invited Article - Beyond effectiveness and efficiency lies quality in services for eating disorders. European Eating Disorders Review, 7, 162-178.
- Van Strien, D.C., Van der Ham & van Engeland, H. (1992). Drop out characteristics in a follow-up study of 90 eating disorder patients. International Journal of Eating Disorders, 12, 341-343.
- Wilson, G.T., Loeb, K.L., Walsh, B.T., Labouvie, E., Petkova, E., Liu, X. & Waternaux, C. (1999). Psychological versus pharmacological treatments of bulimia nervosa: predictors and processes of change. Journal of Consulting and Clinical Psychology, 67(4), 451-459.

Footnotes

1. Detailed tables are available through authors.
2. Chronically ill patients could have received longer rather than shorter treatments. On the assumption that these patients have a poorer prognosis, bias in the estimates of success of shorter and longer treatment programs would have resulted.
3. Since the assumption of homogeneous variance for treatment duration in the hospitals obviously is not adequate (see diagram 2), an estimate of variance *within* each hospital was calculated using a modified method by Littell et al. (1996).
4. AN and BN had a different distribution for body weight at index admission: the median for anorectics was 71.5% BMI and 102.3% BMI for bulimics. The distribution overlapped between 75% and 93% BMI.
5. An odds ratio of 1 indicates that the probability of success are equal for short and long treatments. An odds ratio of 2 indicates a success rate twice as high for long treatments and an odds ratio of 0.5 indicates half as high a success rate when compared to short treatment.
6. This is because the amount of treatment, the duration and intensity, can only be interpreted here as a real predictor and possible causal factor. Decisions as to the length of treatment are assumed to be associated with the outcome and/or the target outcome, although this association is currently not clarified and remained unclarified within this study.
7. We utilized GLIMMIX-Macros (Littell et al., 1996, Chapter 11).

Acknowledgements

The authors are very grateful to the colleagues of Project TR-EAT for providing not only the data , but also many valuable ideas to this paper.

The project was funded by the German Ministry For Education and Research.

Figure Captions

Figure 1. Study design.

Figure 2. Distributions of therapy duration by the hospitals.

Figure 3. Illustration of the effect of propensity adjustment: mean BMI by quintiles (Anorexia Nervosa).

Figure 4. Interaction between treatment duration and age in association with rate of success (Anorexia Nervosa).

Figure 5. Interaction between treatment duration, age, and Morgan-Russell "socio-economic state" scale in association with rate of success (Bulimia Nervosa).