Psychology and Psychotherapy: Theory, Research and Practice (2007), 80, 51–68

The British Psychological Society

www.bpsjournals.co.uk

Course of improvement over 2 years in psychoanalytic and psychodynamic outpatient psychotherapy

Bernd Puschner[1], Susanne Kraft[2], Horst Kächele[3] and Hans Kordy[2]*

[1]Department of Psychiatry II, Ulm University, Germany
[2]Center for Psychotherapy Research, University Hospital Heidelberg, Germany
[3]Department of Psychosomatic Medicine and Psychotherapy, Ulm University, Germany

Objective. To assess and predict the level and course of symptomatic improvement in psychoanalytic (PAP) and psychodynamic psychotherapy (PD).

Methods. In a comprehensive longitudinal study, the course of improvement of 116 patients in PAP and of 357 patients in PD was tracked over a period of 2 years and analysed via hierarchical linear models.

Results. At baseline, over 90% of the patients reported substantial psychological, physical or interpersonal distress. In both forms of treatment, the course of improvement could be adequately fitted by a linear model. Symptom distress decreased notably within 2 years, with an especially sharp decline before the first formally scheduled therapy session. No significant differences between forms of treatment as to level or pace of symptom improvement could be observed. Prediction of speed of improvement was poor, with initial symptom distress showing the strongest influence while initial helping alliance had no predictive value. When comparing patients who finished their treatment within the 2-year observation period with those with still ongoing treatments, the former showed quicker symptom improvement.

Discussion. Strategies for the optimal allocation of valuable therapeutic resources should be reconsidered. An adaptive, outcome-oriented allocation strategy of therapeutic resources is proposed.

Since the first edition of the open-door review of psychotherapy outcomes (Fonagy *et al.*, 1999) sponsored by the International Psychoanalytical Association (IPA), a considerable amount of evidence on the outcome of psychoanalytically oriented treatments has

* *Correspondence should be addressed to Hans Kordy, University Hospital Heidelberg, Center for Psychotherapy Research at the Institute for Psychosomatic Cooperative Research and Family Therapy, Bergheimer Str. 54, 69115 Heidelberg, Germany (e-mail: hans.kordy@med.uni-heidelberg.de).*

DOI:10.1348/147608306X107593

accumulated, especially phase IV research on treatments of limited duration and session frequency (Fonagy *et al.*, 2002; Leichsenring, Rabung, & Leibing, 2004). Nevertheless, compared with research on other forms of psychotherapy, it remains strikingly sparse. Sandell *et al.*'s (2000) sceptical judgement ('. . . psychoanalysts have been strangely uninterested in demonstrating the value of their practice in any systematic way that is likely to satisfy the traditional scientific community', p. 922) still seems justified.

The Berlin Psychotherapy Study (Rudolf, 1991; Rudolf, Manz, & Ori, 1994) yielded results indicating more improvement for psychoanalytically than for psychodynamically treated patients on several outcome criteria, even though no differences in improvement of symptomatic impairment were found.

An innovative design was realized in the STOPPP project (Sandell *et al.*, 2000). Patients (74 patients in psychoanalysis, four to five sessions a week; 331 patients in psychoanalytic psychotherapy, two to three sessions a week) in various stages of treatment (before, during or after long-term psychodynamic therapy or psychoanalysis) were assessed annually over a 3-year period. Both groups started treatment at about the same level of symptom distress measured via the Global Severity Index of the SCL-90-R (Derogatis, 1986). At the end of treatment, however, patients in psychoanalysis were somewhat less distressed than the patients in psychotherapy. This difference substantially increased during follow-up. In social functioning, patients in psychoanalysis started rather better off, but both groups showed about the same amount of improvement until follow-up. In a subsample of 156 patients, post-treatment change could be predicted by frequency and duration in interaction, with worse outcome related to higher session frequency and shorter treatment duration (Sandell, Blomberg, & Lazar, 2002).

In sum, there are only a few prospective longitudinal studies of psychoanalysis or psychoanalytic psychotherapy as performed in daily practice. So far, no final conclusions can be drawn as to the course of improvement before, during or after treatment, and very little is known about predictors of the course.

Usually, researchers have collected data at the beginning and end of treatment and at one or more follow-up time points. This limits our understanding of possible differences in patterns of change within treatments between various forms of psychotherapy that usually differ in length (as in the projects cited above). It could be that the speed of improvement is similar in two treatment groups, but that the course of improvement continues for patients with longer treatments. Thus, length of treatment would make the critical difference rather than clinical specifications and their theoretical background. Longitudinal data gathered independently of the actual length of treatment are needed to provide information about whether and when the courses of improvement diverge in various forms of psychotherapy.

Such a longitudinal perspective was chosen for this study. The course of symptomatic improvement in analytic and psychodynamic psychotherapy over a 2-year observation period was estimated. Also, the possible impact of covariates such as SES, initial impairment, form and planned amount of treatment on course of improvement were examined. The two forms of analytically oriented treatment were compared on the following specific dimensions:

(1) Initial impairment in relevant domains (symptom, physical and interpersonal)
(2) Course of symptom improvement
(3) Predictors of course of improvement
(4) Course of improvement after short-/mid-term and long-term psychotherapy
(5) Course of symptomatic impairment after treatment.

Method

Design and data collection

From September 1998 to February 2002, data for the study Transparency and Outcome Orientation in Outpatient Psychotherapy (TRANS-OP) were collected consecutively by the Center for Psychotherapy Stuttgart (CPR, now at University of Heidelberg). By means of a longitudinal design, a sample of $N = 932$ insurees of a major private health insurance company (Deutsche Krankenversicherung, DKV) from all over Germany were recruited when they applied for subsidized outpatient psychotherapy (required for reimbursement if psychotherapy of more than 25 sessions is intended). For the patients, the design of TRANS-OP comprised a total of five measurement points over a 2-year period. Furthermore, therapists were contacted via their patients and requested to contribute their clinical judgement at intake and $1\frac{1}{2}$ years later.

All patients received questionnaires at intake (T1), as well as $1\frac{1}{2}$ and 2 years later (T4 and T5). Intermediate measurement points T2 and T3 were administered randomly at two out of seven possible points in time (4, 8, 16, 26, 40, 52 and 64 weeks from intake). This design was optimized for the application of hierarchical linear models. It provides a rather fine-graded time grid for the sample (a total of 10 measurements over 2 years), while at the same time keeping the burden on the individual patient at an acceptable level. Since we expected more rapid change in the first year, we chose non-equidistant intervals to allow for more frequent assessments in the early treatment phase.

The flow of study participants is displayed in Figure 1. In this paper, data analyses focus on those patients who applied for psychoanalytic (PAP; $N = 144$) or psychodynamic psychotherapy (PD; $N = 472$), while findings on patients who intended to start cognitive-behavioural therapy ($N = 283$) are reported elsewhere.

High response rates confirm the advantages of this measurement plan. Of the 616 subjects, 473 (76.8%) who had been contacted by the CPR returned their first questionnaire (116 or 80.6% in PAP and 357 or 75.6% in PD) and gave their written consent to participate. For these 473 participants, response rates for the succeeding second, third, fourth and fifth patient questionnaires were 91.8, 86.7, 78.9 and 75.5% (not all 473 subjects who participated at T1 were contacted for the succeeding questionnaires since we refrained from sending questionnaires to subjects who refused

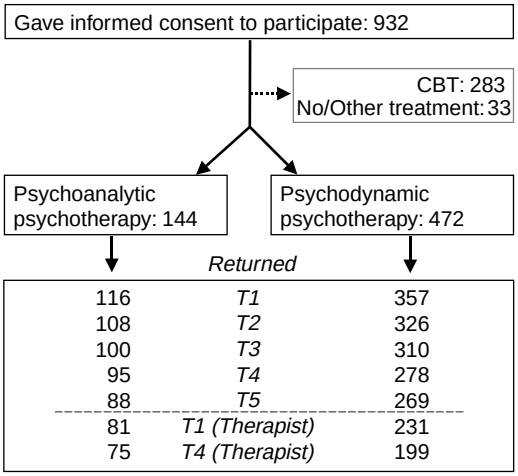

Figure 1. Participant flow. *Note*: Legend indicates number of observations/number of subjects.

to participate at any time of the study). Also, the therapists' readiness to participate was remarkable, i.e. 312 (66.0% of participating patients) sent back their first and 274 (57.9%) their second questionnaire. Neither patients' nor therapists' return rates differed significantly by form of intended treatment.

Data sources and assessment instruments

Measures included assessment of the central problems for which people seek professional help, i.e. psychological and physical symptoms as well as social distress and dissatisfaction with life. In addition, therapists were asked to provide their clinical judgement and the DKV contributed information about health service utilization (Kraft, Puschner, Lambert, & Kordy, 2006).

Patient

Psychological impairment was measured through the German version (Franke, 1995) of Derogatis' (1986) symptom-checklist (SCL-90-R). This is a widely used self-report scale comprising 90 items each on a five-point Likert scale ('not at all' . . . 'very much'). The Global Severity Index (GSI) indicating the mean impairment over all 90 items is used throughout this paper as a global indicator of psychological impairment. Interpersonal problems were assessed by the subscale 'interpersonal relations' of the German version of the Outcome Questionnaire-45 (OQ-45.2; Lambert, Hannöver, Nisslmüller, Richard, & Kordy, 2002) consisting of 11 five-point items. Physical complaints were tapped into by means of the Giessener Beschwerdebogen (GBB-24; Brähler & Scheer, 1995) with 24 five-point items, of which a sum score was calculated as a global indicator of subjective physical impairment. Life satisfaction was measured by the Fragebogen zur Lebenszufriedenheit (FLZ; Fahrenberg, Myrtek, Wilk, & Kreutel, 1986), which investigates eight crucial areas of life satisfaction (health, job, friendships, etc.). A sum score of these seven-point items reflects general satisfaction with life. The quality of the patient–therapist relationship was measured by the German version (Bassler, Potratz, & Krauthauser, 1995) of Alexander and Luborsky's (1986) Helping Alliance Questionnaire. Patients were presented with 11 six-point items. Overall item means yield a global rating of the quality of the helping alliance. Single items assessed therapy motivation ('With regards to this treatment, I am not . . . very motivated', four-point scale), problem duration ('How long do the problems persist which you need help for?', six-point scale) and duration of sick leaves during the year prior to application for treatment (five categories).

Therapist

Patient's impairment from the therapists' point of view was measured by the Impairment Score (IS), which comprises three five-point items (psychological, physical and social problems) (Schepank, 1995) and a sum score. The quality of the patient–therapist relationship from the therapist's point of view was assessed by the Helping Alliance Questionnaire (HAQ; Bassler *et al.*, 1995) which contains nine six-point items (see above). In addition, we took the diagnostic information (coded according to ICD-10; WHO, 1993) from the application forms which therapists are requested to fill out for subsidization of the treatment. For those therapists who participated in the study, we obtained the diagnostic information from their first questionnaire.

Health insurance company

The DKV provided dates of first enquiry of the insuree about reimbursement of outpatient psychotherapy, of therapist's and insuree's application for reimbursement, of approval of reimbursement, and of beginning of treatment (date of first reimbursed session, up to five probatory sessions not included). Furthermore, information on the form of psychotherapy (as indicated on the therapists' application form), as well as on the number of sessions applied for, reimbursed and actually utilized was supplied.

Procedures

While most patients sent back four or even all five questionnaires, owing to the study design and differences in willingness to participate, the number of measurement points varies. Multilevel analysis (Hox, 2002) or hierarchical linear modelling (HLM) (Bryk & Raudenbush, 1987; Raudenbush & Bryk, 2001) is the method of choice for modelling courses of improvement using such unbalanced longitudinal data. HLM makes it possible to use all available data of all participants, even if they missed a number of measurement points and took part at different time intervals, as long as data loss occurred at random. This condition is met in the study (at least for the responders) because the patients were randomly assigned to the intermediate measurement points.

Analyses were carried out using S-PLUS® (version 6.1). First (unconditional), level-1 models for the course of improvement before, during and after treatment were generated. Second (conditional), level-2 models scrutinizing the predictor effects on the course of improvement during and after treatment were calculated. In order to avoid very small group sizes, certain categories of factorial covariates were collapsed: family status into not married, married and separated (widowed, divorced, living separately); educational status into high track (Gymnasium), middle track (Realschule) and low track (Hauptschule) plus others; professional status into university degrees vs. no university degree; duration of sick leaves into up to 1 month vs. more than 1 month; problem duration into up until 2, 10 or 20 years; and amount of initially approved sessions in fewer than 80, 80 or more than 80 for PAP, and fewer than 50, 50 or more than 50 for PD.

Sample

The sample consisted of highly educated people, most of whom held a university degree (see Table 1). Also, note the high percentage of males compared to other studies. Predominant ICD-10 diagnoses included affective (F3) and neurotic disorders (F4), while behavioural syndromes with physical factors and personality disorders were diagnosed to a lesser extent. Subjects did not differ by form of treatment on baseline SES variables and diagnoses, except that people intending to start PAP ($M = 41.81$; $SD = 12.51$ years) were somewhat younger compared with those expected to begin PD ($M = 44.04$; $SD = 11.34$) ($t = -1.7963$, $df = 471$, $p = .07$).

Treatments

Treatment allocation was based on patients' preferences backed by clinical judgement. PAP refers to treatments as they are carried out in day-to-day practice. This includes treatments with between two and four sessions per week and varying length. PAP, as reimbursed by German health insurance companies, is usually restricted to 240 treatment sessions (although exceptions are possible). In contrast, PD is restricted to a total of 80 sessions (again with exceptions possible), with usually one session a week.

Table 1. Socio-economic status and main diagnoses (ICD-10) by form of intended treatment

		PAP		PD	
Variable (N_{PAP}/N_{PD})	Categories	*N*	%	*N*	%
Sex (116/357)	Male	50	43.1	158	44.3
	Female	66	56.9	199	55.7
Age (years) (116/357)	18–29	24	20.7	48	13.4
	30–49	53	45.7	189	52.9
	50 and older	39	33.7	120	33.6
Marital status (115/356)	Single	39	33.9	104	29.2
	Married	47	40.9	160	44.9
	Divorced/living separated/widowed	29	25.2	92	25.7
Educational status (115/355)	Abitur (high track)	93	80.9	257	72.4
	Realschule (middle track)	13	11.3	60	16.9
	Hauptschule (low track)/other	9	7.8	38	10.7
Professional status (115/352)	University degree	71	61.7	206	58.5
	No university degree	44	38.3	146	41.5
ICD-10 diagnoses[a] (86/300)	F30–F39	42	48.8	152	50.7
	F40–F48	30	34.9	123	41.0
	F50–F59	6	6.9	10	3.3
	F60–F69	8	9.3	12	4.0
	Other	0	0	3	1.0

[a]F30–F39, mood (affective) disorders; F40–F48, neurotic, stress-related and somatoform disorders; F50–F59, behavioural syndromes associated with physiological disturbances and physical factors; F60–F69, disorders of adult personality and behaviour.

While PAP aims at resolving the patient's neurotic structure by using processes of transference, countertransference and interpretation of resistance, PD only takes into account these processes while focussing on current neurotic conflicts.

There is no further information available about what treatment was actually carried out. However, every application went through a peer review system where the treatment plan was assessed by an experienced clinician and, thus, clinical standards were ensured.

The number of sessions initially approved for reimbursement differed markedly by form of intended treatment, i.e. subjects whose therapists had applied for PAP mainly were approved 80 sessions (median; mean = 87.52, $SD = 44.66$), while those in PD were mainly approved 50 sessions (median; mean = 40.24, $SD = 15.30$).

Of the 532 treatments (86.3% of entire sample) which actually started (i.e. more than one session reimbursed), 67.7% (PAP, 57.0%; PD, 71.3%; $\chi^2 = 5.1$, $p < .05$) were terminated within the 2-year observation time. Observation time for the course of improvement was 2 years for all subjects, but for the service utilization data – due to consecutive sampling – varied from 2 to more than 4 years (whole study time September 1998 to February 2002). Thus, treatment length could be estimated through Kaplan–Meier survival analysis by censoring cases (106 or 19.9%) which did not terminate treatment within the study time. The estimated median treatment length was 17.7 months (95% CI = 15.9–19.5; 75th percentile = 28.5 months) for the whole sample and differed substantially by form of treatment. Median treatment length was 21.4 months (95% CI = 18.3–24.5) for PAP and 16 months (95% CI = 14.2–17.8) for PD (75th percentiles: PAP = 35.0, PD = 25.9; log rank = 13.3, $df = 1$; $p < .001$).

Results

Initial impairment and helping alliance

For almost 90% in each group, self-reported psychological symptom severity (SCL-90-R GSI) were above the 68th percentile of the normative sample and displayed impairment in this sense (Table 2).

Table 2. Initial impairment and quality of helping alliance

Scale	Treatment	Mean	*SD*	Not impaired (*N*/%)	Impaired (*N*/%)	*N*
Psychological impairment (SCL-90-R GSI)	PAP	1.12	0.55	12/10.5	102/89.5	114
	PD	1.01	0.51	47/13.4	303/86.6	350
	Total	1.04	0.52	59/12.7	405/87.3	464
Interpersonal problems (OQ-45-IR)	PAP	17.88	7.27	27/23.9	86/76.1	113
	PD	15.97	6.66	110/31.4	240/68.6	350
	Total	16.44	6.85	137/29.6	326/70.4	463
Physical complaints (GBB-24)	PAP	26.43	15.29	49/43.0	65/57.0	114
	PD	27.31	14.88	132/37.8	217/62.2	349
	Total	27.09	14.97	181/39.1	282/60.9	463
Life quality (FLZ)	PAP	31.61	8.48	36/33.0	73/67.0	113
	PD	30.05	7.60	128/38.6	204/61.4	350
	Total	30.43	7.85	164/37.2	277/62.8	463
HAQ patient's view (HAQ)	PAP	1.68	0.61			111
	PD	1.82	0.62	*Not standardized*		340
	Total	1.78	0.62			451
Impairment therapist's view (IS)	PAP	6.37	1.63	24/29.6	57/70.4	81
	PD	6.06	1.88	77/34.4	147/65.6	224
	Total	6.14	1.82	101/33.1	204/66.9	305

More than two thirds showed impairment in interpersonal functioning and around 60% reported severe physical impairment. Almost half of the subjects (PAP, 46.9%; PD, 44.8%) displayed impairment in all three and a considerable number (PAP, 34.5%; PD, 35.2%) in two domains, while only some (PAP, 12.4%; PD, 12.2%) showed impairment in just one, and hardly any (PAP, 6.2%; PD, 7.9%) were in the functional range in all three domains at intake. Furthermore, about two thirds of the participants reported serious dissatisfaction with life. Accordingly, about 70% of the participants (of those whose therapists contributed data for the study) were attested severe (i.e. more than five IS-points; cf. Schepank, 1995) overall impairment by their therapists.

Participants rated the quality of the helping alliance rather high with no differences by form of psychotherapy.

Course of improvement

The course of improvement was operationalized as a change in symptom severity on SCL-90-R GSI. The resulting models (mean fixed effects) for the time segments before and in treatment intended as PAP or PD are displayed in Figure 2. Interrupted horizontal

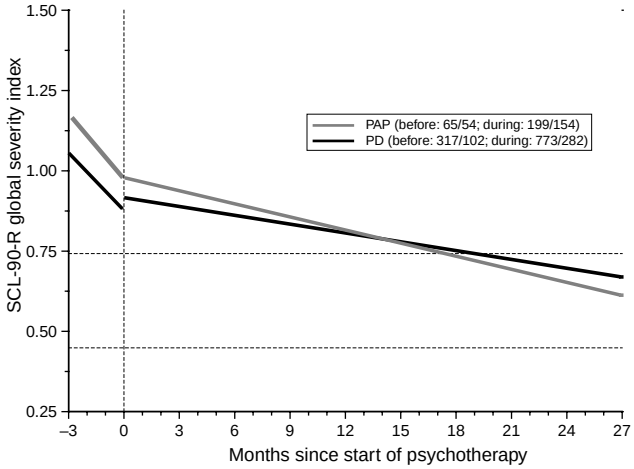

Figure 2. Courses of symptom improvement before and during PAP or PD.

lines indicate the cut-off-points at the 68th (no impairment vs. impairment) and the 95th (impairment vs. severe impairment) percentile of the normative sample.

Figure 2 shows linear fits only. Alternatively, logarithmic fits were tried for the time in treatment and the time in treatment plus the time before treatment. However, differences between the goodness-of-fit indices ('smaller is better'; cf. Pinheiro & Bates, 2000) were marginal for the entire sample, as well as for the subgroups PAP and PD (see Table 3, available at http://www.psyres.de/transop). Thus, following the criterion of simplicity, the linear model was given priority because the more complex logarithmic model did not add substantial gain.

Table 3. Comparison of model fits[a] (linear vs. logarithmic) for entire sample and by form of treatment (available at http://www.psyres.de/transop)

Time segment	Sample *N* (obs./s.s[b])	Fit	*df*	AIC	BIC	logLik
Before and in treatment	PAP	Linear	6	401.11	424.78	−194.55
	382/108	Logarithmic	6	391.16	414.84	−189.58
	PD	Linear	6	1052.23	1081.5	−520.11
	972/308	Logarithmic	6	1037.13	1066.41	−512.56
	Total	Linear	6	1449.99	1481.25	−718.99
	1,354/416	Logarithmic	6	1429.98	1461.24	−708.99
In treatment	PAP	Linear	6	339.51	362.05	−163.76
	317/102	Logarithmic	6	326.94	349.48	−157.47
	PD	Linear	6	864.72	892.59	−426.36
	773/282	Logarithmic	6	846.3	874.18	−417.15
	Total	Linear	6	1200.19	1230.13	−594.09
	1,090/384	Logarithmic	6	1173.07	1203.01	−580.54

[a] Originating from the log-restricted likelihood (logLik), the Akaido information criterion (AIC) and the Bayesian information criterion (BIC) in addition consider number of model parameters and observations. Comparison of GOF indices between models calculated with different samples is not meaningful.

[b] Observations of subjects.

As can be seen in Figure 2, marked improvement – 0.0589 GSI points per month (PAP, 0.0748; PD, 0.0586) – already took place before the start of outpatient psychotherapy (note though that the number of pretreatment observations is rather low). Participants entered treatment with 0.9339 GSI points (PAP, 0.9775; PD, 0.9152) with a monthly reduction in GSI of 0.0106 points (PAP, 0.0135; PD, 0.0091) which for both groups was significantly different from 0 (PAP, $t = -4.87$; $p < .0001$; PD, $t = -5.05$; $p =< .0001$). Even though participants in PAP started a little more impaired, they improved somewhat faster than those in PD, leaving the severely impaired range after about 17 months – approximately 1 month earlier than those in PD. However, neither difference in intercept nor in slope attained statistical significance.

As an addition, the concept of reliable clinical change (Jacobson & Truax, 1991; Kordy & Hannöver, 2000) was applied to assess improvement over 2 years of treatment. Accordingly, of the subjects with data at T1 and T5 ($N = 349$; PAP, $N = 88$; PD, $N = 261$), $N = 219$ (62.8%; PAP = 70.5%; PD = 60.2%) were classified as reliably improved, $N = 89$ (25.5%; PAP = 19.3%; PD = 27.6%) as unchanged, and $N = 41$ (11.7%; PAP = 10.2%; PD = 12.3%) as reliably deteriorated. RCI rates did not differ by form of treatment ($\chi^2 = 3.11$; $df = 2$; $p = .21$). The stricter criterion of clinically significant improvement (reliable improvement and passing cut-off at 68th percentile) was attained by 29.5% of the subjects (PAP = 28.1%; PD = 30.0%).

Predictors of the course of improvement

A comprehensive list of possible predictors was analysed in order to search for differential effects on the course of improvement for the time under treatment for both forms of therapy (available at http://www.psyres.de/transop) (Table 4). Apart from the trivial finding that initial GSI was substantially related to intercept, some other variables predicted baseline symptom distress: educational status in both PAP (low track starting more impaired than middle track) and PD (low track starting more impaired than high track), and periods of sick leave in PD (patients with shorter periods of sick leave were more impaired at intake than those with longer ones). More interestingly, only a few of the predictors in the model affected the speed of symptomatic improvement: in both PAP and PD, initial psychological impairment strongly affected slope (the higher initially impaired, the quicker the improvement), while professional status impinged on slope in PAP to some extent (those with no university degree progressed faster).

Following Liao (1994), the relevance of predictor effects on the course of improvement was further analysed graphically. According to the principle of ceteris paribus, interesting values of the selected predictor were inserted into the model equation while keeping constant (i.e. inserting means of continuous and modes of discrete variables) all other predictors' values. Figure 3 shows variations of course of improvement by initial symptomatic impairment for PAP and PD.

In both forms of treatment, intercept differed strongly by initial GSI, but slope was also affected, i.e. the higher initial impairment, the faster the patients improved.

Short-/mid-term vs. long-term psychotherapy

Comparison of short-/mid-term (pragmatically defined here as those treatments that were terminated during the 2-year observation period) with long-term treatments (those lasting longer than 2 years) yielded the models graphically displayed in Figure 4.

It can be seen for both PAP and PD that patients in short-/mid-term and long-term therapy started at the same level of symptomatic impairment, while termination status

Table 4. HLM coefficients of the comprehensive model for PAP and PD (in treatment) including test for differences by form of psychotherapy (available at http://www.psyres.de/transop)

	PAP (281 observations of 89 s.s)			PD (281 observations of 89 s.s)			Difference	
Variable	Value	*t*	*p*	Value	*t*	*p*	*t*	*p*
Intercept	.294	1.01	.31	.058	.24	.81	−.39	.69
Age (years)	−.004	−1.10	.27	−.001	−.30	.76	.63	.53
Sex (male vs. female)	.054	.78	.44	−.022	−.51	.61	−.80	.42
Family status (married vs. not married)	.069	.89	.38	−.018	−.29	.77	−.84	.40
Family status (married vs. widowed/separated/divorced)	.035	.39	.69	−.013	−.19	.85	−.46	.65
Educational status (low vs. middle track)	.366	2.66	*<.01*	.164	1.83	.07	−.97	.33
Educational status (low vs. high track)	.171	1.25	.22	.196	2.29	*.02*	.11	.91
Professional status (university degree vs. no university)	−.122	−1.52	.13	.034	.61	.54	1.29	.19
Sick leaves (up to 1 vs. more than 1 month)	−.001	−.01	.99	.110	2.02	*.04*	1.01	.31
Problem duration (up to 2 vs. up to 10 years)	.028	.39	.70	−.073	−1.47	.14	−1.25	.21
Problem duration (up to 2 vs. up to 20 years)	.137	1.53	.13	−.037	−.71	.48	−1.56	.12
Therapy motivation	−.095	−1.63	.11	−.012	−.32	.75	1.03	.30
Initial SCL-90-R GSI	.852	9.60	*<.001*	.669	11.77	*<.001*	−1.54	.13
Initial OQ-45 IR	−.005	−1.00	.32	.005	1.32	.19	1.33	.18
Initial GBB-24 total	.002	.63	.53	.001	.58	.56	.06	.95
Initial FLZ total	.006	1.32	.19	.002	.52	.60	−1.08	.28
Initial HAQ total	−.068	−1.26	.21	−.066	−1.76	.08	−.23	.82
Number of sessions initially approved[a]	.036	.45	.65	−.106	−.48	.63	–	–
Number of sessions initially approved[b]	.161	1.69	.09	.004	.09	.93	–	–
Waiting for approval (days)	−.001	−1.65	.10	.000	1.02	.31	1.66	.09
Slope	−.041	−1.07	.29	−.005	−.19	.85	.47	.64
Age (years)	.001	1.14	.25	.000	.71	.48	−.10	.92
Sex (male vs. female)	−.001	−.21	.84	−.003	−.72	.47	−.08	.93
Family status (married vs. not married)	−.013	−1.36	.18	.004	.62	.54	1.35	.18
Family status (married vs. widowed/separated/divorced)	−.010	−.84	.40	.001	.15	.88	.46	.65
Educational status (low vs. middle track)	−.033	−1.77	.08	−.011	−1.17	.24	.98	.33
Educational status (low vs. high track)	−.013	−.73	.46	−.011	−1.22	.22	.01	.99

Table 4. (*Continued*)

Variable	PAP (281 observations of 89 s.s)			PD (281 observations of 89 s.s)			Difference	
	Value	*t*	*p*	Value	*t*	*p*	*t*	*p*
Professional status (university degree vs. no university)	.024	2.36	.02	.003	.49	.62	−1.40	.16
Sick leaves (up to 1 vs. more than 1 month)	−.000	−.01	.99	−.009	−1.72	.09	−.57	.57
Problem duration (up to 2 vs. up to 10 years)	−.004	−.46	.64	.003	.50	.62	.67	.50
Problem duration (up to 2 vs. up to 20 years)	−.009	−.87	.38	.000	.03	.97	.86	.39
Therapy motivation	.012	1.45	.14	−.003	−.66	.51	−1.95	.05
Initial SCL-90-R GSI	−.029	−2.63	<.01	−.012	−2.07	*.04*	1.59	.11
Initial OQ-45 IR	.001	1.09	.28	−.000	−.56	.57	−1.32	.19
Initial GBB-24 total	.000	.47	.64	.000	1.34	.18	.06	.95
Initial FLZ total	−.000	−.52	.61	.000	.89	.37	1.30	.19
Initial HAQ total	−.003	−.48	.63	.002	.39	.69	.83	.41
Number of sessions initially approved[a]	.005	.48	.63	.002	.07	.95	–	–
Number of sessions initially approved[b]	.000	.02	.99	.001	.13	.89	–	–
Waiting for approval (days)	.000	.78	.44	.000	.66	.51	−.34	.74

Random effects (*SD* of intercept/slope/residual): PAP = 0.0622/0.0172/0.2672; PD = .0951/.0092/.2991.
Goodness of fit (AIC/BIC/logLik): PAP = 509.50/662.83/−210.75); PD = 805.03/998.12/−358.51.
$p < .05$ in italics.
[a] < 80 vs. 80 for PAP; < 50 vs. > 50 for PD.
[b] < 80 vs. > 80 for PAP; < 50 vs. 50 for PD.

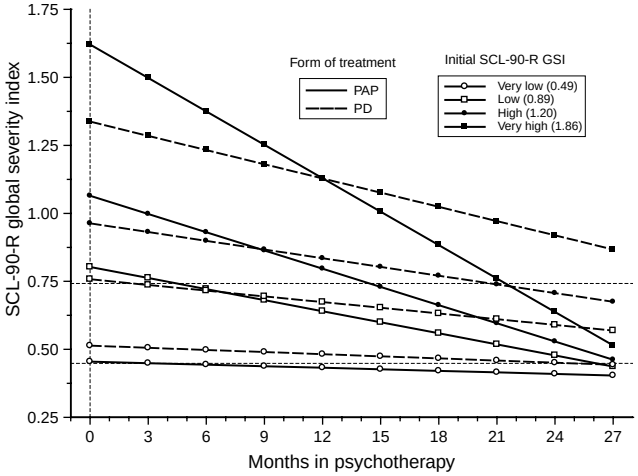

Figure 3. Courses of symptom improvement during PAP or PD by initial symptom severity.

was substantially associated with the pace of improvement. As compared with participants who terminated early, those who stayed in therapy longer than 2 years progressed slower by 0.003 GSI points per month in PAP and by 0.006 points in PD, respectively. The difference reached statistical significance only for subjects in PD (PAP, $t = -0.829$; $p = .408$; PD, $t = -2.891$; $p = .004$). The number of sessions allotted for reimbursement did not contribute significantly to any of the models.

Post-treatment symptomatic impairment

Finally, the relation of course of improvement during treatment and the further development after treatment end was explored. For this purpose, data were available on 232 participants (PAP, 44; PD, 188) who had terminated their therapy during the 2-year observation time. The predictive value of service utilization (duration of treatment and

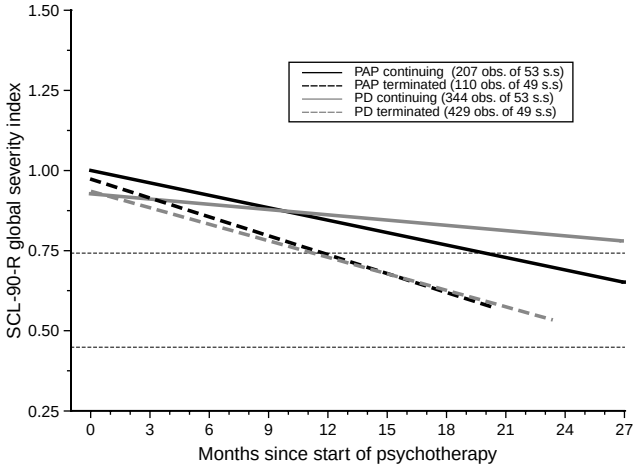

Figure 4. Courses of symptom improvement in PAP or PD over 2 years by termination status.

sessions utilized) and treatment outcome (in-treatment slope and GSI at the end of treatment) for symptomatic impairment (GSI) at the end of the observation period was examined by means of a linear regression model (slope and last in-treatment status were obtained from the random effects coefficients of the unconditional hierarchical linear model using all in-treatment observations; see above). Furthermore, the interaction effect between last in-treatment status and duration of therapy was included in the regression since the time between status at last treatment session and assessment 2 years after study intake varied substantially between participants. Separate models were computed for the two forms of treatment since treatment utilization differed substantially between them.

Table 5 shows that, for participants who had received PAP, symptomatic impairment measured 2 years after intake was not predicted by any of the variables entered into the regression model. This was different for participants who underwent PD. For these, GSI at the last treatment session and the interaction of therapy duration and status showed predictive value.

Table 5. Prediction of psychological impairment 2 years after intake for participants in PAP or PD who terminated treatment before end of observation time (linear regression)

		B	SE	*t*	*p*
PAP[a]	Constant	−0.454	0.47	−0.97	0.34
	Utilized sessions	−0.003	0.00	−1.75	0.09
	Duration of therapy (months)	0.031	0.03	1.14	0.26
	Status at last treatment session (GSI)	0.935	0.57	1.65	0.11
	Slope in treatment (change in GSI/month)	−22.583	20.23	−1.12	0.27
	Duration of therapy × last status	−0.014	0.04	−0.36	0.72
PD[b]	Constant	−0.001	0.24	0.00	1.00
	Utilized sessions	0.002	0.00	0.69	0.49
	Duration of therapy (months)	−0.006	0.01	−0.47	0.64
	Status at last treatment session (GSI)	0.478	0.21	2.29	0.02
	Slope in treatment (change in GSI/month)	−1.113	14.69	−0.08	0.94
	Duration of therapy × last status	0.030	0.01	2.02	0.05

[a]$R^2 = .4018$.
[b]$R^2 = .4394$.

Discussion

The main objective of this study was to investigate the course of improvement of people receiving either analytic or psychodynamic psychotherapy over a 2-year period. These two psychoanalytically oriented treatments cover about two thirds of mid- and long-term psychotherapy – i.e. expected treatment length of more than 25 sessions – reimbursed in the German health service system. As such, this study may be considered an example of service research (National Institute of Mental Health, 1999). Because of its design, this study could show how gains develop over time (during and after treatment) and how progress is related to the form, amount and length of therapy. This knowledge is of special concern in issues of health service provision, which values the optimal allocation of valuable therapeutic resources.

The observation period began when the insuree contacted the health insurance company and asked for the forms for the application of reimbursement. We understand this request as a significant hint that the insuree took psychotherapy into serious consideration at this time. This understanding was supported by the observation that study participants displayed clear psychological, physical and social impairment (see Table 2) at the time when they initiated the application for reimbursement.

There were no significant differences in the course of improvement between PAP and PD, neither in intercept nor in slope. This is surprising as peer reviewers – in agreement with the German psychotherapy guidelines – suggested more therapy sessions for PAP than for PD (medians 80 vs. 50) at first application, and also as treatment length was considerably longer for PAP (medians 21 vs. 16 months). It is possible that reduction of symptomatic impairment does not play the same role for setting up treatment plans for these two forms of psychoanalytically oriented treatments.

According to the resulting models, psychological distress declined quickly during the time before the first session. More than one third of the expected improvement over the full 2-year observation period was achieved during this first phase. This quick improvement might be related to a specific condition of the German service system, i.e. the possibility of up to seven (PAP) or five (PD) so-called probatory sessions before the application for reimbursement. Unfortunately, no information was available about whether and when probatory sessions were actually used. Therefore, it remains open how many such 'preparatory' sessions contributed to this progress. Another possibility, besides mere regression to the mean, is discussed in the research literature as a 'door handle' effect, i.e. prospect on start of possibly long-awaited treatment raises hope and entails swift initial symptom improvement (e.g. Frank, 1971; Lueger, Saunders, Howard, Vessey, & Nunez, 1999).

Despite substantial improvement during the preparatory phase, patients began treatment with considerable psychological distress. The initial status of psychological distress as indicated by the GSI score was substantially above the 95th percentile of the representative population and declined substantially over the following period of up to 2 years. Linear models proved equally suitable to describe the course of improvement as logarithmic models. This finding is in contrast to the still dominating research literature, where negatively accelerating models such as logarithmic, logistic or probit models are proposed (although not actually empirically compared with linear models). Such models suggest 'a law of diminishing return' (Howard, Kopta, Krause, & Orlinsky, 1986), i.e. less additional gain for later sessions, whilst linearity implies that at any time of therapy the same amount of improvement can be expected as return for any additional therapy session. Thus, the linearity found in this project supports the invitation for reconsideration of the dose–effect relationship by Barkham *et al.* (2006); see also Percevic, Lambert, & Kordy (2006). However, one has to keep in mind that all these models are population models. Application for individual patients should be considered carefully.

To make such models more helpful for individual treatment planning, predictors are needed which could help develop more specific models for subpopulations. For this purpose, a long list of predictors was explored. Initial symptomatic impairment showed the most substantial effect on the course of improvement after the first therapy session. Patients who started severely impaired improved quickly, while those with negligible symptoms remained almost unchanged. Furthermore, some of the SES variables, such as lower professional status, indicated somewhat quicker improvement.

Special clinical attention should be directed at the finding that the initial quality of the helping alliance did not substantially affect outcome. In contrast to the research literature (for an overview, see Beutler *et al.*, 2004), helping alliance hardly discernably impinged on initial status and the speed of symptom improvement. This is in agreement with critical appraisals of the practical clinical value of the helping alliance (e.g. Martin, Garske, & Davis, 2000). In the context of service research, a further null finding deserves attention. Neither the form of therapy nor the allotted number of therapy sessions proved a useful predictor for the course of improvement.

Many psychoanalytically oriented practitioners might consider ending psychoanalytic treatment before 2 years have elapsed as premature termination. However, with regard to reduction of psychological distress, patients and therapists appear to behave quite rationally. While 'early' terminators in PAP, as well as in PD, started at about the same level of psychological distress as continuers, the speed of improvement was considerably higher for terminators, especially in PD. Correspondingly, early terminators were substantially better off than continuers at the end of the 2-year observation period. These findings raise doubts on whether termination was actually premature, but rather indicate that the experience of approaching a 'good enough level' (Barkham *et al.*, 2006) allowed termination. On the contrary, one could speculate that continuers actually did need more and longer treatment in order to achieve satisfactory improvement.

Rapid subjective improvement is sometimes understood by clinicians as a hint towards a patient's reluctance to actually get involved in therapy. Concerns are raised about the stability of such quick gains. However, such concerns were not supported by our finding that speed of change during therapy negatively affected the mid-term course of symptomatic impairment after the end of treatment. Rather, the opposite is suggested in so far that – at least for PD – last status of psychological distress in treatment proved predictive for the symptomatic status at the end of the observation period.

This study has important implications for psychotherapy provision. The idea of diminishing returns has been used to justify a strategy of recommending termination or change of treatment in the case of slow or non-response (e.g. Haas, Hill, Lambert, & Morrell, 2002; Lueger *et al.*, 2001). At least under the conditions of the German health service system and for a time span of 2 years, this general strategy is not supported. The finding that linear models adequately reflect the course of improvement suggests that patients should be treated until the desired level of psychological well-being is achieved – and to stop treatment as soon as this level is achieved. This means that quick responders should get shorter and slow responders longer treatment. Such an individually tailored outcome-oriented allocation strategy, described in more detail by Kordy, Haug, and Percevic (2006), is suggested by this study's results for both forms of psychoanalytically oriented therapies.

The practical use of such a strategy would depend on a convincing operationalization of psychological well-being. Even though one might assume that symptomatic improvement is a common goal in any psychotherapeutic treatment, outcome of especially psychoanalytic psychotherapy might be a more delicate matter not measurable by means of symptom checklists.

One specific question that results from this general limitation is whether and how PAP and PD actually differ in practice. There was neither specific training nor any control installed to ensure that the treatments were carried out according to theoretical concepts. Furthermore, no data were gathered about treatment quality. However, these therapies can at least be considered as intended as PAP or PD in the

sense, that – according to the rules of the German health service system – therapists and patients applied for reimbursement of costs for PAP and PD, and the treatment plan was approved through peer review.

This research was carried out under the specific conditions of the German health service system. It remains open whether the reported findings hold true under different service system conditions. Further limitations include: (1) The 2-year observation period might still be too short to judge the stability of the therapeutic gains. (2) Patient recruitment was restricted to privately insured people, which might limit the sample's representativeness for the German health service system since better educated insurees with higher income are over-represented. It might be that – compared with clients of, for example, public health insurers – clients of a private health insurance company are more motivated and able to use health service provision in general (including psychotherapy). (3) As a special bias, it was observed that male and female patients were equally represented, while usually this relation in psychotherapy studies is about 1:2.

A list of other TRANS-OP papers can be found on http://www.psyres.de/transop.

Acknowledgements

This study was funded by the Deutsche Krankenversicherung (DKV), Cologne, Germany. We are grateful to two anonymous reviewers for helpful comments. We also wish to thank Robert J. Lueger at Marquette University, Milwaukee, USA, for final corrections.

References

Alexander, L. B., & Luborsky, L. (1986). The Penn Helping Alliance Scales. In L. S. Greenberg & W. M. Pinsof (Eds.), *The psychotherapeutic process: A research handbook* (pp. 325–366). New York: Guilford Press.

Barkham, M., Connell, J., Stiles, W. B., Miles, J. N. V., Margison, F., Evans, C., & Mellor-Clark, J. (2006). Dose–effect relations and responsive regulation of treatment duration: The good enough level. *Journal of Consulting and Clinical Psychology*, *74*, 160–167.

Bassler, M., Potratz, B., & Krauthauser, H. (1995). Der Helping Alliance Questionnaire (HAQ) von Luborsky: Möglichkeiten zur Evaluation des therapeutischen Prozesses von stationärer Psychotherapie [Luborsky's HAQ: Possibilities to evaluate the therapeutic process in inpatient psychotherapy]. *Psychotherapeut*, *40*, 23–32.

Beutler, L. E., Malik, M., Alimohamed, S., Harwood, T. M., Talebi, H., Noble, S., & Wong, E. (2004). Therapist variables. In M. J. Lambert (Ed.), *Bergin and Garfield's handbook of psychotherapy and behavior change* (pp. 227–306). New York City: Wiley.

Brähler, E., & Scheer, J. W. (1995). *Der Gießener Beschwerdebogen (GBB) (Testmappe mit 2., ergänzter und revidierter Auflage des Handbuchs) [The Gießen complaints questionnaire]*. Bern: Huber.

Bryk, A. S., & Raudenbush, S. W. (1987). Application of hierarchical linear models to assessing change. *Psychological Bulletin*, *101*, 147–158.

Derogatis, L. R. (1986). SCL-90-R: Self Report Symptom Inventory. In Collegium Internationale Psychiatriae Scalarum (Ed.), *Internationale Skalen für Psychiatrie*. Weinheim: Beltz.

Fahrenberg, J., Myrtek, M., Wilk, D., & Kreutel, K. (1986). Multimodale Erfassung der Lebenszufriedenheit: Eine Untersuchung an Herz-Kreislauf-Patienten [Multimodal assessment of life satisfaction: A study of cardio-vascular patients]. *Psychotherapie und Medizinische Psychologie*, *36*, 347–354.

Fonagy, P., Jones, E. E., Kächele, H., Krause, R., Clarkin, J. F., Perron, R. G. A., Gerbev, A., & Allisou, E. (2002). *An open door review of outcome studies in psychoanalysis* (2nd Rev. ed.). London: Research Committee of the International Psychoanalytical Association.

Fonagy, P., Kächele, H., Krause, R., Jones, E., Perron, R., & Lopez, D. (1999). *An open door review of outcome studies in psychoanalysis* (1st ed.). London: Research Committee of the International Psychoanalytical Association.

Frank, J. D. (1971). Therapeutic factors in psychotherapy. *American Journal of Psychotherapy*, *25*, 350–361.

Franke, G. H. (1995). *Die Symptom-Checkliste von Derogatis–Deutsche Version [Derogatis' symptom checklist–German version]*. Göttingen: Beltz Test.

Haas, E., Hill, R. D., Lambert, M. J., & Morrell, B. (2002). Do early responders to psychotherapy maintain treatment gains? *Journal of Clinical Psychology*, *58*, 1157–1172.

Howard, K. I., Kopta, S., Krause, M., & Orlinsky, D. (1986). The dose-effect relationship in psychotherapy. *American Psychologist*, *41*, 159–164.

Hox, J. J. (2002). *Multilevel analysis: Techniques and applications*. Mahwah, NJ: Erlbaum.

Jacobson, N. S., & Truax, P. (1991). Clinical significance: A statistical approach to defining meaningful change in psychotherapy research. *Journal of Consulting Psychology*, *59*, 12–19.

Kordy, H., & Hannöver, W. (2000). Die Evaluation von Psychotherapie und das Konzept der Klinisch bedeutsamen Veränderung [Evaluation of psychotherapy and the concept of clinically significant change]. In A.-R. Laireiter (Ed.), *Diagnostik in der Psychotherapie*. Wien, New York: Springer.

Kordy, H., Haug, S., & Percevic, R. (2006). Patients differ – A plea for individually tailored service allocation. *European Eating Disorders Review*, *14*, 1–7.

Kraft, S., Puschner, B., Lambert, M. J., & Kordy, H. (2006). Medical utilization and treatment outcome in mid- and long-term outpatient psychotherapy. *Psychotherapy Research*, *16*, 241–249.

Lambert, M. J., Hannöver, W., Nisslmüller, K., Richard, M., & Kordy, H. (2002). Fragebogen zum Ergebnis von Psychotherapie: Zur Reliabilität und Validität der deutschen Übersetzung des Outcome Questionnaire 45.2 (OQ-45.2) [Questionnaire on the results of psychotherapy: Reliability and validity of the German translation of the Outcome Questionnaire 45.2 (OQ-45.2)]. *Zeitschrift für Klinische Psychologie*, *1*, 40–47.

Leichsenring, F., Rabung, S., & Leibing, E. (2004). The efficacy of short-term psychodynamic psychotherapy in specific psychiatric disorders: A meta-analysis. *Archives of General Psychiatry*, *61*, 1208–1216.

Liao, T. F. (1994). *Interpreting probability models: Logit, probit, and other generalized linear models*. Thousand Oaks, CA: Sage.

Lueger, R. J., Howard, K. I., Martinovich, Z., Lutz, W., Anderson, E. E., & Grissom, G. (2001). Assesssing treatment progress of individual patients using expected treatment response models. *Journal of Consulting and Clinical Psychology*, *69*, 150–158.

Lueger, R. J., Saunders, S. M., Howard, K. I., Vessey, J. T., & Nunez, P. R. (1999). Entering psychotherapy. In N. E. Miller & K. M. Margruder (Eds.), *Cost-effectiveness of psychotherapy* (pp. 63–71). Oxford: Oxford University Press.

Martin, D. J., Garske, J. P., & Davis, K. D. (2000). Relation of therapeutic alliance with outcome and other variables: A meta-analytic review. *Journal of Consulting and Clinical Psychology*, *68*, 438–450.

National Institute of Mental Health. (1999). *Bridging science and service: A report by the National Advisory Mental Health Council's Clinical Treatment and Services Research workgroup*. Retrieved April 7, 2004, from http://www.nimh.nih.gov/research/bridge.htm.

Percevic, R., Lambert, M. J., & Kordy, H. (2006). What is the predictive value of responses to psychotherapy for its future course? Empirical explorations and consequences for outcome monitoring. *Psychotherapy Research*, *16*, 364–373.

Pinheiro, J. C., & Bates, D. M. (2000). *Mixed-effects models in S and S-PLUS*. New York: Springer.

Raudenbush, S. W., & Bryk, A. S. (2001). *Hierarchical linear models: Applications and data analysis methods* (2nd ed.). Newbury Park, CA: Sage.

Rudolf, G. (1991). Free University of Berlin: Berlin Psychotherapy Study. In L. E. Beutler & M. Crago (Eds.), *Psychotherapy Research. An International Review of Programmatic Studies* (pp. 185–193). Washington: American Psychological Association.

Rudolf, G., Manz, R., & Ori, C. (1994). Ergebnisse der psychoanalytischen Therapien [Outcomes of psychoanalytic therapies]. *Zeitschrift für Psychosomatische Medizin und Psychoanalyse*, *40*, 25–40.

Sandell, R., Blomberg, J., & Lazar, A. (2002). Time matters: On the temporal interactions in long-term follow-up of long-term psychotherapies. *Psychotherapy Research*, *12*, 39–59.

Sandell, R., Blomberg, J., Lazar, A., Carlsson, J., Broberg, J., & Schubert, J. (2000). Varieties of long-term outcome among patients in psychoanalysis and long-term psychotherapy: A review of findings in the Stockholm Outcome of Psychoanalysis and Psychotherapy Project (STOPPP). *International Journal of Psychoanalysis*, *81*, 921–942.

Schepank, H. (1995). *Der Beeinträchtigungs-Schwere-Score: Ein Instrument zur Bestimmung der Schwere einer psychogenen Erkrankung [The Impairment Score: An instrument to determine the severity of mental illness]*. Göttingen: Beltz.

World Health Organisation. (1993). *Internationale Klassifikation psychischer Störungen: ICD-10, Kapitel V (F) [International classification of mental disorders: ICD-10, chapter V (F)]* (2nd ed.). Bern, Göttingen, Toronto, Seattle: Huber.

Received 13 May 2005; revised version received 2 February 2006