REVIEW

Effectiveness of Long-term Psychodynamic Psychotherapy
A Meta-analysis

Falk Leichsenring, DSc
Sven Rabung, PhD

Context The place of long-term psychodynamic psychotherapy (LTPP) within psychiatry is controversial. Convincing outcome research for LTPP has been lacking.

Objective To examine the effects of LTPP, especially in complex mental disorders, ie, patients with personality disorders, chronic mental disorders, multiple mental disorders, and complex depressive and anxiety disorders (ie, associated with chronic course and/or multiple mental disorders), by performing a meta-analysis.

Data Sources Studies of LTPP published between January 1, 1960, and May 31, 2008, were identified by a computerized search using MEDLINE, PsycINFO, and Current Contents, supplemented by contact with experts in the field.

Study Selection Only studies that used individual psychodynamic psychotherapy lasting for at least a year, or 50 sessions; had a prospective design; and reported reliable outcome measures were included. Randomized controlled trials (RCTs) and observational studies were considered. Twenty-three studies involving a total of 1053 patients were included (11 RCTs and 12 observational studies).

Data Extraction Information on study characteristics and treatment outcome was extracted by 2 independent raters. Effect sizes were calculated for overall effectiveness, target problems, general psychiatric symptoms, personality functioning, and social functioning. To examine the stability of outcome, effect sizes were calculated separately for end-of-therapy and follow-up assessment.

Results According to comparative analyses of controlled trials, LTPP showed significantly higher outcomes in overall effectiveness, target problems, and personality functioning than shorter forms of psychotherapy. With regard to overall effectiveness, a between-group effect size of 1.8 (95% confidence interval [CI], 0.7-3.4) indicated that after treatment with LTPP patients with complex mental disorders on average were better off than 96% of the patients in the comparison groups ($P=.002$). According to subgroup analyses, LTPP yielded significant, large, and stable within-group effect sizes across various and particularly complex mental disorders (range, 0.78-1.98).

Conclusions There is evidence that LTPP is an effective treatment for complex mental disorders. Further research should address the outcome of LTPP in specific mental disorders and should include cost-effectiveness analyses.

JAMA. 2008;300(13):1551-1565 www.jama.com

THE PLACE OF PSYCHOANALYTIC and psychodynamic treatments within psychiatry is controversial.[1,2] Although some evidence supports the efficacy of short-term psychodynamic psychotherapy (STPP) for specific disorders,[3-7] convincing research on the outcome of long-term psychodynamic psychotherapy (LTPP) has been lacking.[1,2,8] Evidence suggests that short-term psychotherapy is sufficiently effective for most individuals experiencing acute distress.[9] Evidence, however, also indicates that short-term treatments are insufficient for a considerable proportion of patients with complex mental disorders, ie, patients with multiple or chronic mental disorders or personality disorders.[9-11] Some studies suggest that long-term psychotherapy may be helpful for these groups of patients.[9,10,12-16] This is true not only of psychodynamic therapy but also of psychotherapeutic approaches that are usually short-term, such as cognitive-behavioral therapy (CBT).[15,16]

Evidence-based treatments for these patient groups are particularly important. Personality disorders, for example, are quite common in both general and clinical populations. They show a high comorbidity with a wide range of Axis I psychiatric disorders and are significantly associated with functional impairments.[17-19] Furthermore, a high proportion of patients in clinical populations expe-

For editorial comment see p 1587.

Author Affiliations: Department of Psychosomatics and Psychotherapy, University of Giessen, Giessen (Dr Leichsenring); and Department of Medical Psychology, University Medical Center Hamburg-Eppendorf, Hamburg (Dr Rabung), Germany.
Corresponding Author: Falk Leichsenring, DSc, Department of Psychosomatics and Psychotherapy, University of Giessen, Ludwigstrasse 76, 35392 Giessen Germany (falk.leichsenring@psycho.med.uni-giessen.de).

rience not just a single but rather multiple mental disorders.[20,21] Patients with multiple mental disorders report significantly greater deficits in social and occupational functioning.[20,21]

Although some studies suggest that LTPP may be helpful for these patient groups, strong evidence-based support for LTPP has been lacking. No meta-analysis addressing the outcome of LTPP has yet been published, although preliminary data have been reported by Lamb.[22] This article reports the first meta-analysis to our knowledge on the outcome of LTPP.

Experts continue to discuss which type of research design—randomized controlled trials (RCTs) vs effectiveness or observational studies—provides the best evidence that a treatment works.[23-28] Randomized controlled trials are carried out under controlled experimental conditions. As such, their strength lies in the control of factors influencing outcome external to the treatments in question; they thus ensure high internal validity of the study. However, their clinical representativeness (external validity) can be limited by strict experimental control.[26] In contrast, effectiveness studies are carried out under the conditions of clinical practice. Consequently, they ensure clinically representative results (ie, high external validity).[29] However, they cannot control for factors influencing outcome apart from the treatment to the same degree as RCTs, ie, threats to internal validity. Taking these issues into account, this meta-analysis sought to include studies with high internal validity (RCTs) and studies with high clinical representativeness (effectiveness studies) provided that they fulfilled predefined inclusion criteria. Including both types of studies allowed the meta-analysis to test for the effect of the research design on outcome and the generalizability of results.

This meta-analysis of LTPP addresses the following research questions:

1. Is LTPP superior to other (shorter) psychotherapeutic treatments, particularly with regard to complex mental disorders, ie, personality disorders, chronic mental disorders (defined as lasting at least a year), multiple mental disorders, or complex depression and anxiety disorders?

2. How effective is LTPP with regard to overall outcome, target problems, general psychiatric symptoms, personality functioning, and social functioning in patients with various, especially complex mental disorders?

3. What patient, treatment, or research factors contribute significantly to the outcome of LTPP (eg, age, sex, diagnostic subgroups, use of therapy manuals, therapist experience, treatment duration, or concomitant psychotropic medication)?

METHODS

The procedures carried out in this meta-analysis are consistent with recent guidelines for the reporting of meta-analyses.[30,31]

Definition of LTPP

Psychodynamic psychotherapies operate on an interpretive-supportive continuum. An emphasis is placed on more interpretive or supportive interventions depending on the patient's needs.[8,32] Gunderson and Gabbard[8(p685)] defined LTPP as "a therapy that involves careful attention to the therapist-patient interaction, with thoughtfully timed interpretation of transference and resistance embedded in a sophisticated appreciation of the therapist's contribution to the two-person field." There is no generally accepted "standard" duration for LTPP. Lamb[22] compiled more than 20 definitions given by experts in the field. They ranged from a minimum of 3 months to a maximum of 20 years. In this meta-analysis, we included studies that examined psychodynamic psychotherapy lasting for at least a year, or 50 sessions. This criterion is consistent with the definition given by Crits-Christoph and Barber.[33(p456)]

Inclusion Criteria and Selection of Studies

We applied the following inclusion criteria: (1) studies of individual psychodynamic therapy meeting the definition given by Gunderson and Gabbard above[8]; (2) psychodynamic therapy lasting for at least a year, or at least 50 sessions; (3) prospective studies of LTPP including before-and-after or follow-up assessments; (4) use of reliable outcome measures; (5) a clearly described sample of patients with mental disorders; (6) adult patients (≥18 years); (7) sufficient data to allow determination of effect sizes; (8) concomitant (eg, psychopharmacological) treatments were admissible, but studies involving concomitant treatment were evaluated separately in order to compare the results of the combined treatment vs LTPP alone; and (9) both RCTs and observational studies fulfilling the criteria listed above. These criteria are consistent with recent meta-analyses of psychotherapy.[5,10]

We collected studies of LTPP that were published between 1960 and May 2008 based on a computerized search of MEDLINE, PsycINFO, and Current Contents. The following search terms were used: (*psychodynamic* or *dynamic* or *psychoanalytic** or *transference-focused* or *self psychology* or *psychology of self*) and (*therapy* or *psychotherapy* or *treatment*) and (*study* or *studies* or *trial**) and (*outcome* or *result** or *effect** or *change**) and (*psych** or *mental**). In addition, manual searches of articles and textbooks were performed, and we communicated with authors and experts in the field. A flow chart showing the process of study selection is given in FIGURE 1.

Data Extraction

The 2 authors independently extracted the following information from the articles: author names, publication year, psychiatric disorder treated with LTPP, age and sex of patients, duration of LTPP, number of sessions, type of comparison group, sample size in

each group, use of treatment manuals (yes/no), general clinical experience of therapists (years), specific experience with the patient group under study (years), specific training of therapists (yes/no), study design (RCT vs observational), duration of follow-up period, and use of psychotropic medication. Disagreements were resolved by consensus. The raters were not blinded with regard to treatment condition, because evidence suggests that blinding is unnecessary for meta-analyses.[34] Effect sizes were independently assessed by the 2 raters. Interrater reliability was assessed for the outcome domains in question, ie, overall outcome, target problems, general psychiatric symptoms, personality functioning, and social functioning. For all areas, interrater reliability was satisfactory ($r \geq 0.80$).

Assessment of Effect Sizes and Statistical Analysis

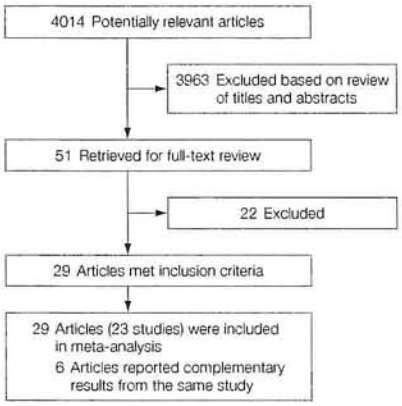
In addition to overall outcome, we assessed effect sizes separately for target problems, general psychiatric symptoms, personality functioning, and social functioning. This procedure was analogous to those in other meta-analyses of psychodynamic therapy.[6,35] As outcome measures of target problems, we included patient ratings of target problems[36] and measures referring to the symptoms specific to the patient group under study (eg, a measure of impulsivity for studies examining borderline personality disorder). For general psychiatric symptoms, both broad measures of psychiatric symptoms such as the Symptom Checklist-90 (SCL-90)[37] and specific measures that do not specifically refer to the disorder under study were included, eg, an anxiety inventory applied to patients with personality disorders. For the assessment of personality functioning, measures of personality characteristics (eg, self-report inventories such as the Defense Style Questionnaire) were included.[38,39] Social functioning was assessed using the Social Adjustment Scale[40] and similar measures.

Whenever a study reported multiple measurements for 1 of the areas of functioning (eg, target problems), we assessed the effect size for each measure separately and calculated the mean effect size of these measures to assess the overall outcome in the respective area of functioning. Overall outcome was assessed by averaging the effect sizes of target problems, general psychiatric symptoms, and personality and social functioning. If a study involved more than 1 form of LTPP, each form was entered separately into this meta-analysis. As a measure of between-group effect size, we used the point biserial correlation r_p as suggested by Cohen and Rosenthal.[41,42] The point biserial correlation also allows us to test for differences between LTPP and other forms of psychotherapy.

As a measure of within-group effect size, the d statistic was calculated for each measure by subtracting the posttreatment mean from the pretreatment mean and dividing the difference by the pretreatment standard deviation of the measure.[42,43] If there was more than 1 treatment group, we calculated a pooled baseline standard deviation as suggested by Hedges and Rosenthal.[41,43] To correct for bias when sample sizes were small, we calculated the Hedges d statistic, an unbiased measure of effect size in small samples (formula 10).[44(p81)] The within-group effect size d gives the difference in the magnitude of change from pretreatment to posttreatment in units of standard deviations. A value of 0.80 is regarded as a large effect.[42,45]

If the data necessary to calculate effect sizes were not published in an article, we asked the study authors for these data. If necessary, signs were reversed so that a positive effect size always indicated improvement. To examine the stability of psychotherapeutic effects, we assessed effect sizes separately for assessments at the termination of therapy and follow-up. If several follow-up assessments were performed, we included only the 1 with the longest follow-up period to study long-term stability of treatment effects. If data pertaining to completers and intent-to-treat samples were reported, the latter were included.

Figure 1. Selection of Trials



See the "Methods" section for study exclusion criteria.

Tests for heterogeneity were carried out using the Q statistic.[44] In case of significant heterogeneity, random-effect models were applied.[46,47] To assess the degree of heterogeneity, we calculated the I^2 index.[48] For control of publication bias, file-drawer analyses were performed.[47,49,50] To test for differences between RCTs and observational studies, point biserial correlations between type of study and outcome were calculated. Only if no significant differences exist, it is appropriate to combine outcome data from RCTs and observational studies.

To compare the effects of LTPP with those of other psychotherapy methods, we performed comparative analyses for the subsample of studies with a control group design. To analyze the effects of LTPP in complex mental disorders, subgroup analyses were carried out for personality disorders, chronic mental disorders, multiple mental disorders, and complex depression and anxiety disorders (the latter being characterized by the chronic course and/or cooccurrence with multiple mental

disorders). For sensitivity analyses, additional subgroup analyses were carried out. Correlation analyses were performed to test the impact of predictor or moderator variables on outcome (eg, concomitant psychotropic medication, use of treatment manuals). Statistical analyses were conducted using SPSS 15.0[51] and MetaWin 2.0.[52] Two-tailed tests of significance were carried out for all analyses. The significance level was defined to be $P=.05$ if not otherwise stated.

Assessment of Study Quality

According to the inclusion criteria described above, only prospective studies of LTPP were included in which reliable outcome measures were used, the patient sample was clearly described, and data to calculate effect sizes were reported. In addition, the quality of studies was assessed by use of the scale proposed by Jadad et al.[53] This scale takes into account if a study was described as randomized, if a study was described as double blind, and if withdrawals and dropouts were described. In psychotherapy research, however, double-blind studies cannot be realized because the patients know or can easily find out which treatment they receive. Thus, all studies of psychotherapy would have to be given a score of 0 points on this item of the Jadad scale. Instead of blinding therapists and patients, the respective requirement in psychotherapy research is that in case of observer-rated outcome measures the ratings were carried out by raters blind to the treatment condition. Additionally, the patient perspective is of particular importance in psychotherapy. For this reason, outcome is often assessed by self-report instruments. We therefore decided to give a score of 1 point on this item if outcome was assessed by blinded raters or by reliable self-report instruments. With this modification, the 3 items of the Jadad scale were independently rated by the 2 authors for all studies included. A satisfactory interrater reliability was achieved for the total score of the scale ($r=0.84$, $P<.001$).

RESULTS

Included Studies

Twenty-three studies met the inclusion criteria (Figure 1).[12-14,39,54-79] For 8 of the studies, we received additional information from the authors.[14,59,65,66,73,75,77,79] The studies are described in TABLE 1. Five studies involved more than 1 LTPP treatment condition.[55,60,62,75,78] Each LTPP condition applied in these studies was entered separately into this meta-analysis.

For 5 studies, some control conditions had to be excluded from this meta-analysis.[65,66,73-75] The observational study comparison groups of the study by Rudolf et al[74] were not included because 1 comparison group did not clearly represent LTPP or STPP due to variability in treatment duration (5 to 200 sessions), and the other condition represented inpatient treatment (Table 1). The comparison group of the study by Huber and Klug[65] was not included because not enough data were available. The low-dose therapy control group of the Sandell et al[75] study was not included, because data to calculate effect sizes were not published for this condition. The data of the short-term psychotherapy groups of the Knekt et al[66] study were not included as control groups because assessments were made at predefined time points that did not represent end of therapy for the short-term treatment group. Of the 4 forms of psychodynamic therapy studied by Piper et al,[73] only the individual long-term and short-term conditions were included. The group treatments were not included due to our inclusion criterion of individual therapy. In all, 8 controlled studies provided the data necessary for comparative analyses of LTPP with other forms of psychotherapy.[12,14,54,59,61,73,77,78] Table 1 and FIGURE 2 indicate 11 studies with 13 LTPP conditions as being RCTS; however, only 8 of these 11 provided data for other forms of psychotherapy.

The results of the studies by Bond and Perry,[39,56] Clarkin et al,[14,57] Knekt et al,[66,67] and Monsen et al[71,72] were reported in 2 journal articles each. Bateman and Fonagy[12,13] presented the data of an 18-month follow-up in a separate article, and Høglend et al[62,63] presented the data of the 1-year and 3-year follow-up in separate articles. We included the data from both articles in our analysis for all of these studies.

For 3 studies,[70,77,79] we received additional information about treatment duration from the authors. In these 3 studies, treatment duration was longer than a year.

In the study by Wilczek et al,[79] not all of the patients under study met the criteria for an Axis I or Axis II diagnosis. To include only individuals with mental disorders, we included only the data from those patients diagnosed with a "character pathology" at intake, according to the Karolinska Psychodynamic Profile as reported by Wilczek et al.[79(p1176)]

In all, 11 RCTs* and 12 observational studies† were included in this meta-analysis. To make the procedures applied in this meta-analysis as transparent as possible, we included the outcome measures used in each study and indicated for which outcome area each measure was included (Table 1). For reasons of space limitations, however, we do not give a reference for each instrument. The reader is referred to the original studies for this information.

The 23 studies included 1053 patients treated with LTPP. For comparative treatments, the number was 257. The 23 included studies cover a wide range of mental disorders (Table 1).

We evaluated the effects of LTPP separately for patients with personality disorders, chronic mental disorders (defined as mental disorders

*References 12, 14, 54, 59, 61, 62, 65, 66, 73, 77, 78.
†References 39, 55, 57, 60, 68-71, 74-76, 79.

Table 1. Studies of Long-term Psychodynamic Psychotherapy

		LTPP Group		Comparison Group			Outcome Measures	
Sources	Mental Disorder	No. of Patients[b]	Duration of Treatment (Follow-up Interval)	No. of Patients	Duration of Treatment (Follow-up Interval)	RCT	Test	Domains
Bachar et al,[54] 1999	Eating disorders	17	40 Sessions; 12 mo	17 CT 10 Control/ nutritional counseling	12 mo CT 6 mo Nutritional counseling	Yes	DSM-SS EAT 26 SCL-90 Selves Q	Target problems Target problems Symptoms Personality
Barber et al,[55] 1997	Avoidant and obsessive-compulsive personality disorders	24 Avoidant personality disorder	52 Sessions			No	WISPI BAI BDI HARS HRSD IIP % Diagnosis[d]	Target problems Symptoms Symptoms Symptoms Symptoms Social functioning
		14 Obsessive-compulsive personality disorder	52 Sessions			No	WISPI BAI BDI HARS HRSD IIP % Diagnosis[d]	Target problems Symptoms Symptoms Symptoms Symptoms Social functioning
Bateman and Fonagy, 1999,[12] 2001[13a]	Borderline personality disorder	19	18 mo	19 Psychiatric TAU Inpatient treatment + Partial hospitalization	11.6 d Inpatient treatment (90% of patients) + 6 mo Partial hospitalization (72% of patients)	Yes	BDI SCL-90-R IIP STAI-state STAI-trait	Symptoms Symptoms Social functioning Symptoms Personality
Bond and Perry, 2004,[39] 2006[56a]	Chronic depression, anxiety, and personality disorders	53	Median, 110 Sessions; Median, 3.0 y; Mean, 3.32 y			No	SCL-90 HRSD GAF DSQ	Symptoms Symptoms Social functioning Personality
Clarkin et al,[57] 2001[a]	Borderline personality disorders	23	12 mo			No	Parasuicide Service utilization	Target problems Social functioning
Clarkin et al,[14] 2007; Levy et al,[58] 2006[a]	Borderline personality disorders	30	12 mo	17 DBT 22 DST	12 mo DBT 12 mo DST	Yes	Aggression scale Anger scale Barrett scale BDI BSI GAF SAS RF Coherence Resolution	Target problems Target problems Target problems Symptoms Symptoms Social functioning Social functioning Personality Personality Personality
Dare et al,[59] 2001	Anorexia nervosa	21	Mean, 24.9 Sessions; 12 mo	22 CAT 22 FT 19 Routine treatment 19 TAU	7 mo CAT 12 mo FT 12 mo TAU	Yes	BMI ABW % Morgan Russel	Target problems Target problems Target problems
Grande et al,[60] 2006	Depressive and anxiety disorders[e]	32 Analytic psychotherapy	Mean, 310 Sessions; Mean, 44.2 mo (12 mo)	2nd LTPP condition		No	SCL-90-R IIP	Symptoms Social functioning
		27 Psychodynamic focal therapy	Mean, 71.1 Sessions; Mean, 24.2 mo (12 mo)	1st LTPP condition		No		
Gregory et al,[51] 2008	Borderline personality disorders	15	12-18 mo	15 TAU	12-18 mo	Yes	BEST BDI DES SPS % Parasuicide, alcohol misuse, institutional care[d]	Target problems Symptoms Symptoms Social functioning

(continued)

Table 1. Studies of Long-term Psychodynamic Psychotherapy (cont)

		LTPP Group		Comparison Group			Outcome Measures	
Sources	Mental Disorder	No. of Patients[b]	Duration of Treatment (Follow-up Interval)	No. of Patients	Duration of Treatment (Follow-up Interval)	RCT	Test	Domains
Høglend et al,[62] 2006, 2008[63]	Depressive, anxiety and personality disorders[e]	52 Transference interpretation	33 Sessions; 12 mo (12 mo, 24 mo)	2nd LTPP condition		Yes	PFS SCL-90-R IIP GAF	Target problems Symptoms Social functioning Social functioning
		48 No transference interpretation	33 Sessions; 12 mo (12mo, 24 mo)	1st LTPP condition		Yes	PFS SCL-90-R IIP GAF	Target problems Symptoms Social functioning Social functioning
Huber and Klug,[65] 2006	Depressive disorders	35	229 Sessions; Mean, 48.8 mo	8 PFT[c]	60.6 Sessions 19.4 mo	Yes	BDI SCL-90-R IIP	Target problems Symptoms Social functioning
Knekt et al,[66,67] 2008	Depressive or anxiety disorders	128	235 Sessions; ≤36 mo	101 STPP[c] 97 SFT[c]	20 STPP sessions 5-6 mo ≤12 SFT sessions ≤8 mo	Yes	BDI HRSD HARS SCL-Anxiety SCL-90-GSI WAI SAS-W PPFS NSLD	Target problems Target problems Target problems Target problems Symptoms Social functioning Social functioning Social functioning Social functioning
Korner et al,[68] 2006	Borderline personality disorder	29	12 mo	31 TAU	12 mo	No	*DSM-III-R* Score GAF	Target problems Social functioning
Leichsenring et al,[69] 2005	Depressive, anxiety, and personality disorders[e]	36	Mean, 253 sessions; Mean, 37.4 mo (12 mo)			No	GAF SCL-90-R FLZ IIP	Target problems Symptoms Personality Social functioning
Luborsky et al,[70] 2001	Heterogeneous disorders	17	>50 Sessions			No	GAF HSRS	Social functioning Social functioning
Monsen et al,[71,72] 1995	Personality disorders	23	Mean, 25.4 mo (60 mo)			No	Affect MMPI [D + Pt + Si] [F + pa + sc]	Target problems Target problems Symptoms Personality
Piper et al,[73] 1984	Heterogeneous disorders 30% Personality disorders	30	Mean, 76 sessions; (6 mo)	27 Individual STPP	Mean, 22 sessions (6 mo)	Yes	TSP TSPI TSIA TSIAI TST TSTI Cornell DA CATT IBSP IBSD SSIAM	Target problems Target problems Target problems Target problems Target problems Target problems Symptoms Symptoms Personality Social functioning Social functioning Social functioning
Rudolf et al,[74] 1994	Depressive, anxiety, and personality disorders[e]	44	Mean, 265 sessions	56 PFT[c] 164 POI[c]	5-200 PFT sessions, 2.6 mo, POI	No	PSKB-SE 1 PSKB-SE 2	Symptoms Personality
Sandell et al,[75] 2000	Heterogeneous disorders	24 Psychoanalysis	Mean, 642 sessions; Mean, 54 mo (12 mo, 24 mo)	27 Low-dose therapies[c]		No	SCL-90-R SOCS SAS	Symptoms Personality Social functioning
		129 LTPP	43 mo LTPP (12 mo, 24 mo)	27 Low-dose therapies[c]		No	SCL-90-R SOCS SAS	Symptoms Personality Social functioning
Stevenson and Meares,[76] 1992	Borderline personality disorders	30	12 mo			No	*DSM-III* Score Cornell Behavior	Target problems Symptoms Social functioning

(continued)

Table 1. Studies of Long-term Psychodynamic Psychotherapy (cont)

		LTPP Group		Comparison Group			Outcome Measures	
Sources	Mental Disorder	No. of Patients[b]	Duration of Treatment (Follow-up Interval)	No. of Patients	Duration of Treatment (Follow-up Interval)	RCT	Test	Domains
Svartberg, et al,[77] 2004	Cluster C personality disorders	25	40 sessions; Mean, 16.9 mo (6, 12, 24 mo)	25 CT	18.3 mo (6, 12, 24 mo)	Yes	Millon SCL-90-R IIP	Target problems Symptoms Social functioning
Vinnars et al,[78] 2005[a]	Personality disorders	80 Manualized psychodynamic therapy	≥12 mo (12 mo, 36 mo)	2nd LTPP condition		Yes	*DSM-IV* score SCL-90-T GAF Change in diagnosis[d]	Target problems Symptoms Social functioning
		76 Community-delivered psychodynamic therapy	≥12 mo (12 mo, 36 mo)	1st LTPP condition		Yes	*DSM-IV* score SCL-90-T GAF Change in diagnosis[d]	Target problems Symptoms Social functioning
Wilczek et al,[79] 2004	Heterogeneous disorders Only "character pathology" patients included	36	Mean, 159 sessions (6 mo)			No	KAPP CPRS-S-A[d] GAF[d]	Target problems

Abbreviations: ABW, average body weight; BAI, Beck Anxiety Inventory; BDI, Beck Depression Inventory; BEST, Borderline Evaluation of Severity Over Time; BMI, body mass index; BSI, Brief Symptom Inventory; CAT, cognitive-analytic therapy; CATT, Cattel's H Scale; CPR-S-A, Self-Rating Scale for Affective Syndromes; CT, cognitive therapy; DA, Depression-Anxiety Subscale of Psychiatric Status Schedule; DBT, dialectic behavioral therapy; DES, Dissociative Experiences Scale; D+Pt+Si, subjective discomfort, anxiety, social introversion subscales of MMPI; *DSM-III-R, Diagnostic and Statistical Manual of Mental Disorders* (Third Edition Revised); *DSM-SS*, DSM Symptomatology Scale for Anorexia and Bulimia; DSQ, Defense Style Questionnaire; DST, dynamic supportive treatment; EAT, Eating Attitudes Test; FLZ, Life Satisfaction Questionnaire; F+pa+sc, F, projection, withdrawal subscales of MMPI; FT, family therapy; GAF, Global Assessment of Functioning Scale; HARS, Hamilton Anxiety Rating Scale; HRSD, Hamilton Rating Scale for Depression; HSRS, Health Sickness Rating Scale; IBSD, Interpersonal Behavior Scale (discrepancy between present and ideal functioning); IBSP, International Behavior Scale (present functioning); IIP, Inventory of Interpersonal Problems; KAPP, Karolinska Psychodynamic Profile; LTPP, long-term psychodynamic psychotherapy; MMPI, Minnesota Multiphasic Personality Inventory; NSLD, number of sick-leave days, PFS, Psychodynamic Functioning Scales; PFT, Psychodynamic Focal Therapy; POI, psychodynamically oriented inpatient treatment; PPFS, Perceived Psychological Functioning Scale; PSKB-SE, psychological and social-communicative state–self-report (Psychischer und Sozialkommunikativer Befund–Selbsteinschätzung); RCT, randomized controlled trial; RF, reflexive Function; SAS, Social Adjustment Scale; SAS-W, Work Subscale of the Social Adjustment Scale; SCL-90-R, Symptom Check List-90 revised; SFT, solution- focused therapy, SOCS, Sense of Coherence Scale; SPS, Social Provisions Scale; SSIAM, Structured and Scaled Interview to Assess Maladjustment; STAI, State-Trait Anxiety Inventory; STPP, short-term psychodynamic psychotherapy; TAU, treatment as usual; TSIA and TSIAI, severity for all target objectives and most important objective, TSP & TSPI, severity for all target objectives and most important objective, TST & TSTI, severity for all target objectives and most important objective; WAI, Work Ability Index, WISPI, Wisconsin Personality Disorders Inventory; % Diagnosis, percentage of patients fulfilling criteria for diagnosis.

[a] LTPP combined with psychotropic medication in some patients of the sample.

[b] Intention to treat samples.

[c] Data of these comparison groups were not included in this meta-analysis.

[d] These outcome measures were not included (no data to calculate effect size d for the respective treatment or patient group).

[e] Predominant diagnoses in sample.

lasting ≥1 year), multiple mental disorders (defined as 2 or more diagnoses of mental disorders), and complex depressive or anxiety disorders.

Treatment manuals or manual-like guidelines were applied in 12 studies.‡

The mean (SD) number of sessions carried out in the 23 studies of LTPP was 151.38 (154.98) and a median of 73.50. The duration of therapy was 94.81 (58.79) weeks and a median of 69.00.

For LTPP the mean (SD) length of follow-up period after treatment was 93.23 (64.93) weeks.

In 16 of the 23 studies, outcome data for LTPP alone without any concomitant psychotropic medication were reported.§ In 7 studies, some patients received concomitant psychotropic medication as needed (ie, because of higher symptom severity or other clinical factors).[12,14,39,57,61,66,78]

Tests for Heterogeneity

To test for heterogeneity of the effects of LTPP, we used the Q statistic.[44,52] To assess the degree of heterogeneity, we calculated the I^2 index.[48] For some outcome analyses, Q yielded a significant result. In the total sample of 23 studies, for example, this was true for overall outcome at the posttest assessment ($Q=53.71$; $P=.002$; $I^2=49\%$). In the 8 comparative studies of LTPP, however, Q was significant for only 2 variables, both for follow-up data for which only 2 studies allowed the calculation of the respective effect sizes (target problems: $Q=11.92$; $P=.001$; $I^2=92\%$; social functioning: $Q=4.53$; $P=.03$; $I^2=78\%$). In the comparative studies, the I^2 index for overall outcome was 0%, for target problems, 45%; for symptoms, 46%; for personality functioning, 60%; and social functioning, 51% at the time of posttest indicating low to medium heterogeneity.[80] For follow-up, the number of studies providing data was too limited to calculate reasonable I^2 statistics. To take heterogeneity between studies into account, we used the random-effects model throughout.

‡References 12, 14, 54, 55, 57, 59, 61, 62, 68, 76-78.

§References 54, 55, 59, 60, 62, 65, 68-71, 73-77, 79.

Tests for Publication Bias

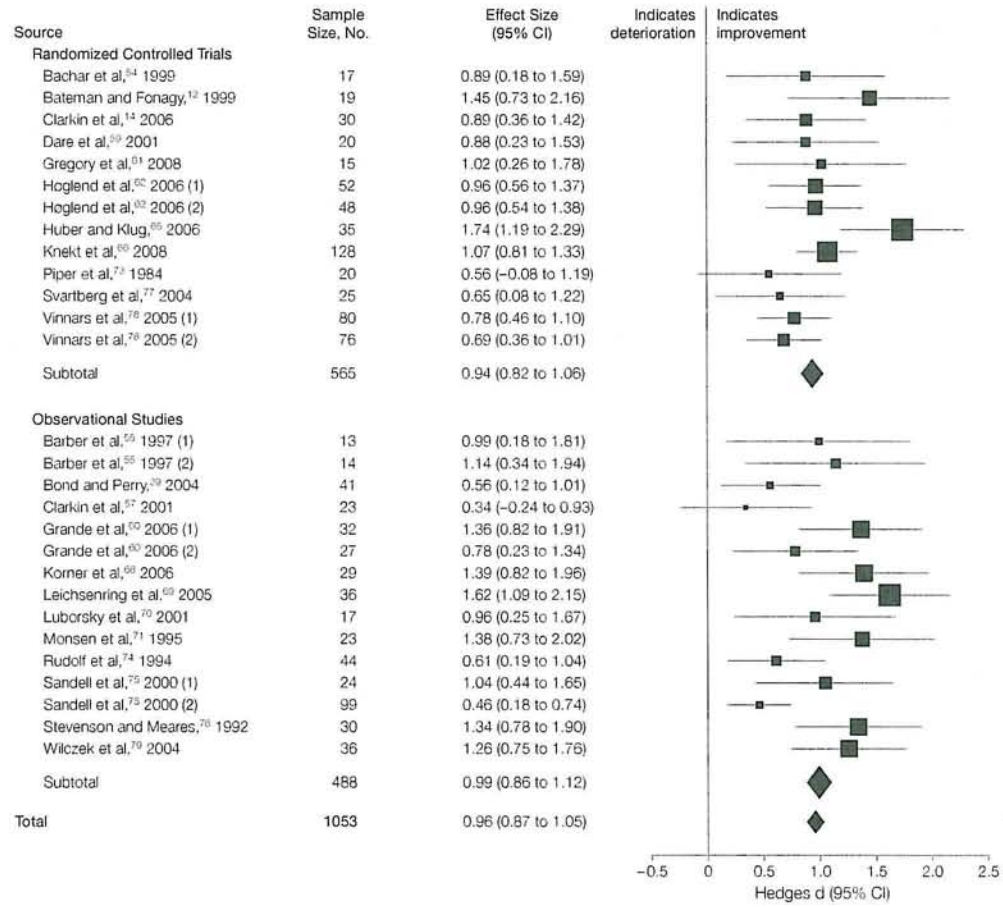
To reduce the file-drawer effect, we tried to identify unpublished studies via the Internet and by contacting researchers. Only 1 additional LTPP study was identified, but it was not included because long-term group therapy was applied.[81] To test for publication bias, we calculated the Spearman rank correlation between effect size and sample size across studies. A significant correlation may indicate a publication bias in which studies with larger effect sizes in 1 direction are more likely to be published.[82] Due to the small number of studies providing follow-up data, we assessed the correlations for only the posttreatment effect sizes. All correlations were nonsignificant ($P>.30$).

As another test for publication bias, we assessed the fail-safe number according to Rosenthal for the posttreatment effect sizes.[49] A fail-safe number is the number of nonsignificant, unpublished, or missing studies that would need to be added to a meta-analysis in order to change the results of the meta-analysis from significance to nonsignificance. In the total sample of studies examining LTPP alone, the fail-safe numbers were 921 for overall outcome, 535 for target problems, 623 for general symptoms, and 358 for social functioning. Due to the smaller number of studies providing data for outcome measures of personality functioning, the fail-safe number was 42 for personality functioning. Even this number is almost twice the number of studies we included. We therefore failed to find any indication of publication bias in this meta-analysis.

Correlation of Quality Ratings With Outcome

To examine the relationship between study quality and outcome of LTPP, the within-group effect sizes for overall outcome, target problems, general symptoms, personality functioning, and so-

Figure 2. Effects of Long-term Psychodynamic Psychotherapy on Overall Outcome



Overall outcome was assessed by averaging the effect sizes of target problems, general psychiatric symptoms, and personality and social functioning. Effect sizes are Hedges d (ie, within-group effect sizes), measured at the beginning and end of therapy. Error bars represent 95% confidence intervals (CIs). Studies are stratified into randomized controlled trials (RCTs) vs observational studies (with or without control groups).

cial functioning were correlated with the total score of the Jadad scale. Due to the small number of studies providing follow-up data, correlations were only calculated for posttreatment effect sizes. For this purpose, the average score of the 2 raters was used. All correlations were nonsignificant ($P>.28$).

Comparison of RCTs With Observational Studies

A forest plot listing the within-group, ie, pretreatment-posttreatment, effect sizes of LTPP on overall outcome for each of the 23 studies is presented in Figure 2. Data are presented separately for RCTs and observational studies. Considering each LTPP condition included in the 23 studies resulted in a total of 13 effect sizes for randomized trials and 15 effect sizes for observational studies.

To test for differences between RCTs and effectiveness (observational) studies, we calculated point biserial correlations in the total sample of 23 studies between the within-group effect size of LTPP at posttest and the type of study design (RCT, 1; observational studies, 0). According to the results, all correlations with outcome measures were nonsignificant ($P>.36$). Observational studies, therefore, did not provide effect sizes significantly different from those of RCTs. There were also no significant differences between the effect sizes of 16 controlled and 7 uncontrolled studies when the 5 studies including observational control groups[55,60,68,74,75] (Table 1) were included ($P>.22$).

In view of these findings, data from RCTs and observational studies were combined in the further analyses of effect sizes of LTPP (see total score in Figure 2).

Effects of LTPP vs Those of Other Psychotherapy Methods

Eight controlled studies provided the data necessary for comparative analyses of LTPP with other forms of psychotherapy.[12,14,54,59,61,68,73,77] These studies included the treatment of personality, eating, and heterogeneous disorders (Table 1). The psychotherapeutic treatments applied in the comparison groups included CBT, cognitive-analytic therapy, dialectical-behavioral therapy, family therapy, supportive therapy, short-term psychodynamic therapy, and psychiatric treatment as usual (Table 1). For the sample of comparative studies, we tested for a correlation between psychotropic medication (0/1) and outcome. Due to the small number of studies providing data for follow-up assessments, tests of significance were carried out only for the posttherapy data, not for the follow-up data. None of the correlations were significant ($P>.13$). For this reason, we included studies of both LTPP alone and LTPP combined with psychotropic medication in the comparative analyses of LTPP vs other methods of psychotherapy.

In the 8 studies included, the mean (SD) duration of LTPP was 53.41 (30.92) weeks and a median of 52 weeks. The mean number of LTPP sessions was 102.57 (135.58), and a median of 49 sessions. In the comparison groups, the mean treatment duration was 39.02 (22.77) weeks, and a median of 52 weeks; the mean number of sessions was 32.58 (27.65), and a median of 22 sessions. Comparing the within-group effects of LTPP with those of the comparison groups will yield information about the possible additional benefit of LTPP. Due to the small number of studies providing data for follow-up assessments, tests of significance were carried out only for the posttherapy data, not for the follow-up data.

We calculated point biserial correlations (r_p) between the within-group effect sizes and type of treatment (LTPP vs other psychotherapies, 1/0) as a measure of between-group effect size as described in the methods section.[41,42] According to Cohen,[42(p82)] a point biserial correlation of 0.371 constitutes a large effect size. The point biserial correlation was also used to test for significance of differences between LTPP and other methods of psychotherapy. As a first step, we compared LTPP with other forms of psychotherapy applied in the comparison groups across the various mental disorders treated in the 8 studies listed above. This comparison included 8 treatment conditions of LTPP and 12 treatment conditions of other psychotherapeutic methods. According to the results, the point biserial correlation between the within-group effect size and treatment condition was significant for overall outcome (r_p=0.60; 95% CI, 0.25-0.81; P=.005, n=20), target problems (r_p=0.49; 95% CI, 0.08-0.76; P=.04, n=18), and personality functioning (r_p=0.76; 95% CI, 0.33-0.93; P=.02, n=9). Thus, LTPP yielded significantly larger pretreatment-posttreatment effect sizes in overall effectiveness (0.96 vs 0.47), target problems (1.16 vs 0.61), and personality functioning (0.90 vs 0.19) than did other forms of psychotherapy applied in the comparison groups. The between-group effect sizes of r_p=0.60, 0.49, and 0.76, respectively, clearly exceed the value of 0.371 and are therefore considered large effects.[42] For social functioning, the between-group effect size was large as well (r_p=0.39; 95% CI, −0.13 to 0.74; P=.19, n=13), but not significant due to the small number of studies reporting this outcome (symptoms: r_p=0.29; 95% CI −0.22 to 0.68; P=.30, n=14).

In the second step of the comparative analysis, we focused on those studies including complex mental disorders that we defined as personality disorders, chronic mental disorders, or multiple mental disorders. For this purpose, 1 study had to be excluded from analysis because the patient sample was not described as having any of these conditions.[73] In order to achieve a sufficient sample size, we did not conduct separate analyses for chronic mental disorders, multiple mental disorders, personality disorders, or complex depression and anxiety disorders. We instead lumped these studies together as studies including "complex mental disorders." In these studies, the point biserial correlation between treatment condition (LTPP vs other psychotherapies) and within-group effect sizes was again

Table 2. Effect Sizes (d) of Long-term Psychodynamic Psychotherapy Alone Across Various Mental Disorders

	No. of Treatment Conditions[a]	d (95% CI)	*P* Value (2-Tailed Test)
Overall effectiveness pretherapy vs posttherapy	20	1.03 (0.84 to 1.22)	<.001
Overall effectiveness pretherapy vs follow-up	8	1.25 (1.00 to 1.49)	<.001
Target problems pretherapy vs posttherapy	14	1.54 (1.20 to 1.87)	<.001
Target problems pretherapy vs follow-up	6	1.98 (1.37 to 2.59)	<.001
Psychiatric symptoms pretherapy vs posttherapy	17	0.91 (0.72 to 1.11)	<.001
Psychiatric symptoms pretherapy vs follow-up	6	1.06 (0.64 to 1.47)	.001
Personality functioning pretherapy vs posttherapy	7	0.78 (0.30 to 1.26)	.005
Personality functioning pretherapy vs follow-up	3	1.02 (−0.99 to 3.03)	
Social functioning pretherapy vs posttherapy	14	0.81 (0.60 to 1.03)	<.001
Social functioning pretherapy vs follow-up	7	0.91 (0.49 to 1.34)	.003

Abbreviations: CI, confidence interval; d, Hedges d; blank cell indicates that no tests of significance were performed due to the small number of studies providing data.
[a] Because some studies included more than 1 form of long-term psychodynamic psychotherapy, the number of treatment conditions in some cases differs from the number of studies.

significant for overall outcome (r_p=0.68; 95% CI, 0.35-0.86; *P*=.002, n=18), target problems (r_p=0.69; 95% CI, 0.34-0.87; *P*=.003, n=16), and personality functioning (r_p=0.96; 95% CI, 0.84-0.99; *P*<.001, n=7). The between-group effect sizes were also large, but not significant for general psychiatric symptoms (r_p=0.40; 95% CI, −0.14 to 0.76; *P*=.20, n=12) and social functioning (r_p=0.45; 95% CI, −0.11 to 0.79; *P*=.17, n=11). The between-group effect sizes of r_p=0.68, 0.69, and 0.96 are equivalent to Cohen d=1.8 (95% CI, 0.7-3.4), 1.9 (95% CI, 0.7-3.5), and 6.9 (95% CI, 3.0-14.6), respectively.[42(p22)] In complex mental disorders, therefore, the differences in effect size between LTPP and other forms of psychotherapy for overall outcome, target problems, and personality functioning were between 1.8 and 6.9 standard deviations. Effect sizes can be transformed into percentiles.[42,83] For example, a between-group effect size of 1.8 (95% CI, 0.7-3.4), as identified in overall outcome, indicates that after treatment with LTPP patients on average were better off than 96% of the patients in the comparison groups.

Comparison of LTPP Alone and LTPP Combined With Psychotropic Medication

In 7 of the total sample of 23 studies, some patients received concomitant psychotropic medication as needed (ie, patients were not randomly assigned to medication but received medication because of higher symptom severity or other clinical factors). Therefore, we again compared the effect sizes of LTPP alone (16 studies)‖ and LTPP combined with psychotropic medication (7 studies)[12,14,39,57,61,66,78] by calculating the point biserial correlation between effect size and treatment condition (LTPP alone vs LTPP combined with psychotropic medication, 0/1). For target problems, the correlation was significant (r_p=−0.45; 95% CI, −0.11 to −0.69; *P*=.05). This means that LTPP combined with psychotropic medication yielded significantly smaller pretreatment-posttreatment effect sizes than LTPP alone in those studies. To avoid bias when estimating the effects of LTPP in specific groups of patients, we decided to include only studies of LTPP alone without concomitant psychotropic medication.

Effects for LTPP Alone Across Various Mental Disorders

As a first step, we assessed the outcome of LTPP alone by evaluating the effect sizes across the various mental disorders treated in the respective studies of LTPP alone.¶ Four of these 16 studies included 2 treatment conditions of LTPP.[55,60,62,75] Thus, 16 studies and 20 treatment conditions of LTPP encompassing 641 patients were evaluated. The within-group effect sizes of LTPP are presented in TABLE 2. According to the results, LTPP yielded significant pretreatment-posttreatment effect sizes that were stable at follow-up for all outcome areas. With the exception of personality functioning (0.78), all effect sizes including those at follow-up were more than 0.80 indicating large effects. For overall outcome, we compared the posttreatment effect sizes of LTPP alone with those at follow-up. The effect sizes significantly increased at follow-up (*t*=3.76, *P*=.007).

Therapy Duration and Effect Sizes

In the studies of LTPP alone, the number of sessions correlated significantly with the outcome for target problems (Spearman r_s=0.62, *P*=.03, n=12) and general psychiatric symptoms (r_s=0.54, *P*=.04, n=15), at posttest time points. The correlations with overall outcome (r_s=0.29, *P*=.25, n=17), changes in personality (r_s=0.43, *P*=.40, n=6), and social functioning (r_s=0.11, *P*=.73, n=12) were not significant. The duration of therapy (weeks) did not show significant correlations with outcome of LTPP alone (*P*>.07). Again, no correlations were calculated for follow-up data due to the small number of studies providing such data.

‖References 54, 55, 59, 60, 62, 65, 68-71, 73-77, 79.

¶References 54, 55, 59, 60, 62, 65, 68-71, 73-77, 79.

Effect Sizes for LTPP Alone in Patients With Personality Disorders

Ten studies included treatments of personality disorders by LTPP (Table 1).# Five studies examined the effects of LTPP alone.[55,68,71,76,77] One study included 2 different groups of personality disorders (avoidant and obsessive-compulsive personality disorder) treated with LTPP.[55] Thus, 5 studies and 6 treatment conditions of LTPP encompassing 134 patients were evaluated with regard to the treatment of personality disorders. According to the results, LTPP alone yielded significant effect sizes for overall outcome, target problems, general psychiatric symptoms, and social functioning at posttest time points (TABLE 3). All these effect sizes were more than 0.80 indicating large effects. Large effect sizes were also observed for personality functioning at posttest and for all outcome areas at follow-up. Due to the small number of studies, however, we performed no tests of significance for these findings (Table 3). Also in the following analyses, no tests of significance were performed for follow-up data because of small sample size.

Effect Sizes for LTPP Alone in Patients With Chronic Mental Disorders

In 7 studies, patients with chronic mental disorders (defined as mental disorders lasting ≥1 year) were treated with LTPP alone.[54,59,60,65,69,74,75] This subsample of studies overlaps in part with the studies of multiple mental disorders and depressive and anxiety disorders described below. Two studies included 2 different treatment conditions of LTPP.[60,75] Thus, the data from 7 studies including 9 LTPP treatment conditions including 334 patients were entered in our meta-analysis. According to the results, LTPP alone yielded significant and large effect sizes for overall outcome, general psychiatric symptoms, personality functioning, and social functioning at posttest time points (Table 3). All effect sizes including those at follow-up were again more than 0.80 indicating large effects in all outcome areas.

Effect Sizes for LTPP Alone in Patients With Multiple Mental Disorders

To assess the outcome of LTPP alone in patients with multiple mental disorders, we separately evaluated those studies in which at least 50% of the patient sample had 2 or more diagnoses of mental disorders. This group of studies overlaps in part with the studies of personality disorders, chronic mental disorders, and depressive and anxiety disorders because these mental disorders are usually highly comorbid.[17-21] This condition was true for 8 studies of LTPP alone.[55,60,62,65,69,71,74,77] Three of

Table 3. Effect Sizes (d) of Long-term Psychodynamic Psychotherapy Alone in Patients With Personality Disorders and Chronic Mental Disorders

	No. of Treatment Conditions[a]	d (95% CI)	*P* Value (2-Tailed Test)
Patients with personality disorders			
Overall effectiveness pretherapy vs posttherapy	6	1.16 (0.82 to 1.50)	<.001
Overall effectiveness pretherapy vs follow-up	2	1.21 (−1.62 to 4.03)	
Target problems pretherapy vs posttherapy	6	1.58 (0.80 to 2.35)	.004
Target problems pretherapy vs follow-up	2	1.65 (−5.90 to 9.19)	
Psychiatric symptoms pretherapy vs posttherapy	5	0.89 (0.49 to 1.29)	.002
Psychiatric symptoms pretherapy vs follow-up	2	0.92 (−1.81 to 3.65)	
Personality functioning pretherapy vs posttherapy	1	0.95 (–)	
Personality functioning pretherapy vs follow-up	1	1.04 (–)	
Social functioning pretherapy vs posttherapy	5	0.82 (0.39 to 1.25)	.007
Social functioning pretherapy vs follow-up	1	1.13 (–)	
Patients with chronic mental disorders			
Overall effectiveness pretherapy vs posttherapy	9	1.05 (0.61 to 1.48)	<.001
Overall effectiveness pretherapy vs follow-up	3	1.36 (0.21 to 2.51)	
Target problems pretherapy vs posttherapy	4	1.70 (0.40 to 3.00)	
Target problems pretherapy vs follow-up	1	2.45 (–)	
Psychiatric symptoms pretherapy vs posttherapy	8	1.05 (0.69 to 1.41)	<.001
Psychiatric symptoms pretherapy vs follow-up	3	1.32 (0.63 to 2.01)	
Personality functioning pretherapy vs posttherapy	5	0.87 (0.18 to 1.56)	.02
Personality functioning pretherapy vs follow-up	1	1.79 (–)	
Social functioning pretherapy vs posttherapy	6	0.88 (0.40 to 1.37)	.004
Social functioning pretherapy vs follow-up	3	1.23 (−0.06 to 2.52)	

Abbreviations: CI, confidence interval; d, Hedges d; blank cell indicates that no tests of significance were performed due to the small number of studies providing data.
[a] Because some studies included more than 1 form of long-term psychodynamic psychotherapy, the number of treatment conditions in some cases differs from the number of studies.

#References 12, 14, 55, 57, 61, 68, 71, 76-78.

these studies included 2 different treatment conditions of LTPP that were evaluated separately.[55,60,62] Thus, 8 studies including 11 LTPP treatment conditions including 349 patients were included in our meta-analysis. According to the results, LTPP yielded significant before-after effect sizes for all outcome domains with the exception of personality functioning. All effect sizes including those at follow-up were more than 0.80 (TABLE 4).

Effect Sizes for LTPP Alone in Patients With Complex Depressive and Anxiety Disorders

In 5 studies of LTPP alone, the majority of patients had complex depressive and anxiety disorders.[60,62,64,69,74] All of these studies included patients with multiple mental disorders. In addition, the patients in 71% of these studies had chronic mental disorders. Thus, the depressive and anxiety disorder subsample of studies included patients with chronic and/or multiple mental disorders (complex mental disorders). Because 2 studies included 2 different treatment conditions of LTPP,[60,62] all 5 studies and 7 LTPP treatment conditions including 274 patients were included in our meta-analysis. According to the results, LTPP alone yielded significant and large effect sizes in overall outcome, general psychiatric symptoms, and social functioning at posttest. All effect sizes, including those at follow-up, were more than 0.80 indicating large effects in all outcome areas (Table 4).

Table 4. Effect Sizes (d) of Long-term Psychodynamic Psychotherapy Alone in Patients With Multiple Mental Disorders or Mainly Complex Depressive and Anxiety Disorders

	No. of Treatment Conditions[a]	d (95% CI)	*P* Value (2-Tailed Test)
Patients with multiple mental disorders			
Overall effectiveness pretherapy vs posttherapy	11	1.09 (0.83 to 1.36)	<.001
Overall effectiveness pretherapy vs follow-up	7	1.28 (1.01 to 1.54)	<.001
Target problems pretherapy vs posttherapy	8	1.62 (1.07 to 2.18)	<.001
Target problems pretherapy vs follow-up	5	1.84 (1.22 to 2.45)	.002
Psychiatric symptoms pretherapy vs posttherapy	11	0.98 (0.76 to 1.21)	<.001
Psychiatric symptoms pretherapy vs follow-up	5	1.18 (0.81 to 1.55)	.001
Personality functioning pretherapy vs posttherapy	3	0.96 (−0.52 to 2.44)	
Personality functioning pretherapy vs follow-up	2	1.43 (−3.32 to 6.18)	
Social functioning pretherapy vs posttherapy	9	0.94 (0.70 to 1.17)	<.001
Social functioning pretherapy vs follow-up	6	1.01 (0.57 to 1.45)	.002
Patients with complex depressive and anxiety disorders			
Overall effectiveness pretherapy vs posttherapy	7	1.13 (0.74 to 1.51)	<.001
Overall effectiveness pretherapy vs follow-up	5	1.30 (0.91 to 1.68)	.001
Target problems pretherapy vs posttherapy	4	1.82 (0.87 to 2.77)	
Target problems pretherapy vs follow-up	3	1.94 (1.01 to 2.88)	
Psychiatric symptoms pretherapy vs posttherapy	7	1.02 (0.70 to 1.34)	<.001
Psychiatric symptoms pretherapy vs follow-up	3	1.32 (0.63 to 2.01)	
Personality functioning pretherapy vs posttherapy	2	0.97 (−6.00 to 7.94)	
Personality functioning pretherapy vs follow-up	1	1.79 (–)	
Social functioning pretherapy vs posttherapy	6	1.02 (0.73 to 1.31)	<.001
Social functioning pretherapy vs follow-up	5	0.99 (0.44 to 1.54)	.009

Abbreviations: CI, confidence interval; d, Hedges d; Blank cell indicates that no tests were performed due to the small number of studies providing data.
[a] Because some studies included more than 1 form of long-term psychodynamic psychotherapy, the number of treatment conditions in some cases differs from the number of studies.

Correlation of Outcome With Specific Patient and Therapist Variables

We examined the effect of the following variables on posttreatment outcome of LTPP (sensitivity analyses): age, sex, diagnostic group (personality disorders, chronic or multiple mental disorders, and depressive and anxiety disorders), general and specific clinical experience of therapists (years), use of treatment manuals (0/1), and specific training in the applied treatment model (0/1). The impact of 10 variables on 10 outcome variables (5 before-after and 5 before follow-up variables) was tested. In order to protect against type I error inflation, we adjusted for multiple testing (0.05/100). All correlations with the outcome of LTPP were nonsignificant ($P > .04$).

COMMENT

A considerable proportion of patients with chronic mental disorders or personality disorders do not benefit sufficiently from short-term psychotherapy.[9,10] However, long-term psychotherapy is associated with higher direct costs than short-term psychotherapy. For this reason, it is important to know whether the benefits of LTPP exceed those of short-term treat-

ments. In this meta-analysis, LTPP was significantly superior to shorter-term methods of psychotherapy with regard to overall outcome, target problems, and personality functioning. Long-term psychodynamic psychotherapy yielded large and stable effect sizes in the treatment of patients with personality disorders, multiple mental disorders, and chronic mental disorders. The effect sizes for overall outcome increased significantly between end of therapy and follow-up.

One limitation of this meta-analysis may be seen in the limited number of studies. The results presented in this meta-analysis, however, were robust. According to the results of sensitivity analyses, they were independent of age, sex, patient subgroups, experience of therapists or use of treatment manuals. We also did not find indications for publication bias. We performed fail-safe number analyses and found that, except for personality functioning, more than 300 studies would need to be added to change the results of the meta-analysis from significance to nonsignificance.

Some of the studies included were carried out in the 1980s and some methodological shortcomings can be expected (eg, problems of randomization, allocation concealment, or observer bias). There was some variance between the included studies with regard to methodological quality as assessed by the scale proposed by Jadad et al.[53] That scale, however, did not show significant correlations with effect sizes of LTPP. This was also true for study design (RCTs vs observational studies). The latter result suggests that the outcome data of the RCTs included in this meta-analysis are representative for clinical practice. On the other hand, the results also show that the data of the observational studies did not systematically overestimate or underestimate the effects of LTPP.[84] Future studies addressing this question should include more specific comparisons of RCTs and observational studies using comparable treatments and diagnostic groups.

Several studies did not meet our inclusion criteria because the majority of patients had not completed their treatment at the time points when effect sizes were assessed by the authors of the original studies. This was true, for example, for the studies by Brockmann et al,[85] Puschner et al,[86] and Giesen-Bloo et al.[87] In the study by Giesen-Bloo, for example, 19 of 42 patients (45%) were still in treatment (LTPP) when outcome was assessed, and only 2 patients had completed LTPP. In the comparison group 27 of 44 patients (61%) were still in treatment, and only 6 patients had completed the treatment. Data from ongoing treatments do not provide valid estimates for treatment outcome at termination or follow-up, eg, if patients received only half of the "dose" of treatment when outcome is assessed.

Whether the effects of psychotherapy improve with longer treatments remains an interesting question. In this meta-analysis, the number of LTPP sessions was significantly correlated with improvements in both target problems and general psychiatric symptoms. These results are consistent with previous findings.[9,10] However, no such correlations were found for the duration of LTPP. The number of sessions and duration of LTPP appear to be different parameters that function differently with regard to the psychotherapeutic process and outcome.

Future research on LTPP, as well as on other approaches when applied as long-term treatment (eg, CBT or interpersonal therapy) should focus on complex mental disorders, such as "double depression" (ie, major depressive disorder plus dysthymic disorder). These studies should compare not only the effects of short-term and long-term psychotherapy but also direct and indirect costs. Some cost-effectiveness studies suggest that LTPP may be a cost-efficient treatment.[88-90]

Author Contributions: Drs Leichsenring and Rabung had full access to all of the data in the study and take responsibility for the integrity of the data and the accuracy of the data analysis.
Study concept and design: Leichsenring, Rabung.
Acquisition of data: Leichsenring, Rabung.
Analysis and interpretation of data: Leichsenring, Rabung.
Drafting of the manuscript: Leichsenring, Rabung.
Critical revision of the manuscript for important intellectual content: Leichsenring, Rabung.
Statistical analysis: Leichsenring, Rabung.
Administrative, technical, or material support: Leichsenring, Rabung.
Financial Disclosures: None reported.
Additional Contributions: We thank Jacques Barber, PhD (Department of Psychiatry, University of Pennsylvania), John Clarkin, PhD (Weill Medical College of Cornell University, New York), Louis Diguer, PhD (Department of Psychology, University of Laval, Québec), Ivan Eisler, PhD (Institute of Psychiatry, King's College London, England), Dorothea Huber, MD, PhD (Department of Psychosomatic Medicine, Technical University Munich, Germany), Paul Knekt, PhD (National Public Health Institute, Helsinki, Finland), William Lamb, PhD (University of California at Berkeley), William Piper, PhD (Department of Psychiatry, University of British Columbia, Vancouver), Sydney Pulver, MD (Department of Psychiatry, University of Pennsylvania), Frank Petrak, DSc (Department of Psychosomatic Medicine, University of Bochum, Germany), Rolf Sandell, PhD (Department of the Behavioural Sciences, Linköping University, Stockholm, Sweden), Martin Svartberg, MD, PhD (Department of Psychiatry, Mount Sinai Hospital, Toronto), and Alexander Wilczek, MD, PhD (Department of Clinical Neuroscience, Karolinska Institute, Stockholm, Sweden) for information about their studies. None of these individuals received compensation for their support.

REFERENCES

1. Gabbard GO, Gunderson JG, Fonagy P. The place of psychoanalytic treatments within psychiatry. *Arch Gen Psychiatry*. 2002;59(6):505-510.
2. Kernberg OF. Psychoanalytic contributions to psychiatry. *Arch Gen Psychiatry*. 2002;59(6):497-498.
3. Fonagy P, Roth A, Higgitt A. Psychodynamic therapies, evidence-based practice and clinical wisdom. *Bull Menninger Clin*. 2005;69(1):1-58.
4. Leichsenring F. Comparative effects of short-term psychodynamic psychotherapy and cognitive-behavioral therapy in depression: a meta-analytic approach. *Clin Psychol Rev*. 2001;21(3):401-419.
5. Leichsenring F, Leibing E. The effectiveness of psychodynamic therapy and cognitive behavior therapy in the treatment of personality disorders: a meta-analysis. *Am J Psychiatry*. 2003;160(7):1223-1232.
6. Leichsenring F, Rabung S, Leibing E. The efficacy of short-term psychodynamic psychotherapy in specific psychiatric disorders: a meta-analysis. *Arch Gen Psychiatry*. 2004;61(12):1208-1216.
7. Leichsenring F. Are psychoanalytic and psychodynamic psychotherapies effective? a review of empirical data. *Int J Psychoanal*. 2005;86(3):841-868.
8. Gunderson JG, Gabbard GO. Making the case for psychoanalytic therapies in the current psychiatric environment. *J Am Psychoanal Assoc*. 1999;47(3):679-704.
9. Kopta SM, Howard KI, Lowry JL, Beutler LE. Patterns of symptomatic recovery in psychotherapy. *J Consult Clin Psychol*. 1994;62(5):1009-1016.
10. Perry JC, Banon E, Ianni F. Effectiveness of psychotherapy for personality disorders. *Am J Psychiatry*. 1999;156(9):1312-1321.
11. Shea MT, Elkin I, Imber SD, et al. Course of depressive symptoms over follow-up: Findings from the National Institute of Mental Health Treatment of Depression Collaborative Research Program. *Arch Gen Psychiatry*. 1992;49(10):782-787.
12. Bateman A, Fonagy P. The effectiveness of par-

tial hospitalization in the treatment of borderline personality disorder, a randomized controlled trial. *Am J Psychiatry*. 1999;156(10):1563-1569.
13. Bateman A, Fonagy P. Treatment of borderline personality disorder with psychoanalytically oriented partial hospitalization: an 18-month follow-up. *Am J Psychiatry*. 2001;158(1):36-42.
14. Clarkin JF, Levy KN, Lenzenweger MF, Kernberg OF. Evaluating three treatments for borderline personality disorder: a multiwave study. *Am J Psychiatry*. 2007;164(6):922-928.
15. Linehan MM, Tutek DA, Heard HL, Armstrong HE. Interpersonal outcome of cognitive behavioral treatment for chronically suicidal borderline patients. *Am J Psychiatry*. 1994;151(12):1771-1776.
16. Linehan MM, Comtois KA, Murray AM, et al. Two-year randomized trial + follow-up of dialectical behavior therapy vs therapy by experts for suicidal behaviors and borderline personality disorder. *Arch Gen Psychiatry*. 2006;63(7):757-766.
17. Lenzenweger M, Lane M, Loranger A, Kessler R. *DSM-IV* Personality disorders in the National Comorbidity Survey Replication. *Biol Psychiatry*. 2007; 62(6):553-564.
18. Kantojärvi L, Veijola J, Laksy K, et al. Co-occurrence of personality disorders with mood, anxiety, and substance abuse disorders in a young adult population. *J Personal Disord*. 2006;20(1):102-112.
19. Grant BF, Hasin DS, Stinson FS, et al. Prevalence correlates and disability of personality disorders in the United States: results from the National Epidemiologic Survey on Alcohol and related conditions. *J Clin Psychiatry*. 2004;65(7):948-958.
20. Olfson M, Fireman B, Weissman MM, et al. Mental disorders and disability among patients in a primary care group practice. *Am J Psychiatry*. 1997; 154(12):1734-1740.
21. Ormel J, VonKorff M, Ustun TB, Pini S, Korten A, Oldehinkel T. Common mental disorders and disability across cultures: results from the WHO Collaborative Study on Psychological Problems in General Health Care. *JAMA*. 1994;272(22):1741-1748.
22. Lamb WK. *A Meta-analysis of Outcome Studies in Long-term Psychodynamic Psychotherapy and Psychoanalysis* [dissertation]. Berkley: University of California; 2005.
23. Benson K, Hartz A. A comparison of observational studies and randomized, controlled trials. *N Engl J Med*. 2000;342(25):1878-1886.
24. Concato J, Shah N, Horwitz R. Randomized, controlled trials, observational studies, and the hierarchy of research designs. *N Engl J Med*. 2000;342(25): 1887-1892.
25. Pocock SJ, Elbourne DR. Randomized trials or observational tribulations? *N Engl J Med*. 2000;342 (25):1907-1909.
26. Rothwell PM. External validity of randomised controlled trials: to whom do the results of this trial apply? *Lancet*. 2005;365(9453):82-93.
27. Seligman ME. The effectiveness of psychotherapy. The Consumer Reports study. *Am Psychol*. 1995;50(12):965-974.
28. Westen D, Novotny CM, Thompson-Brenner H. The empirical status of empirically supported psychotherapies: assumptions, findings, and reporting in controlled clinical trials. *Psychol Bull*. 2004;130(4): 631-663.
29. Shadish WR, Cook TD, Campbell DT. *Experimental and Quasi-Experimental Designs for Generalized Causal Inference*. Boston, MA: Houghton Mifflin Co; 2002.
30. Moher D, Cook DJ, Eastwood S, Olkin I, Rennie D, Stroup DF. Improving the quality of reports of meta-analyses of randomised controlled trials: the QUOROM statement. *Lancet*. 1999;354(9193): 1896-1900.
31. Stroup DF, Berlin JA, Morton SC, et al. Meta-analysis of observational studies in epidemiology: a proposal for reporting. *JAMA*. 2000;283(15):2008-2012.
32. Gabbard GO. *Long-term Psychodynamic Psychotherapy: A Basic Text*. Arlington, VA: American Psychiatric Publishing Inc; 2004.
33. Crits-Christoph P, Barber JP. Long-term psychotherapy. In: Ingram RE, Snyder CR, eds. *Handbook of Psychological Change: Psychotherapy Processes & Practices for the 21st Century*. Hoboken, NJ: John Wiley & Sons; 2000:455-473.
34. Berlin JA. Does blinding of readers affect the results of meta-analyses? *Lancet*. 1997;350(9072): 185-186.
35. Crits-Christoph P. The efficacy of brief dynamic psychotherapy: a meta-analysis. *Am J Psychiatry*. 1992; 149(2):151-158.
36. Battle CC, Imber SD, Hoehn-Saric R, Nash ER, Frank JD. Target complaints as criteria of improvement. *Am J Psychother*. 1966;20(1):184-192.
37. Derogatis L. *The SCL-90 Manual I: Scoring, Administration and Procedures for the SCL-90-R*. Baltimore, MD: Clinical Psychometric Research; 1977.
38. Bond M, Gardner S, Christian J, Sigal J. Empirical study of selfrated defense styles. *Arch Gen Psychiatry*. 1983;40(3):333-338.
39. Bond M, Perry JC. Long-term changes in defense styles with psychodynamic psychotherapy for depressive, anxiety, and personality disorders. *Am J Psychiatry*. 2004;161(9):1665-1671.
40. Weissman MM, Bothwell S. Assessment of social adjustment by patient self-report. *Arch Gen Psychiatry*. 1976;33(9):1111-1115.
41. Rosenthal R. *Meta-analytic Procedures for Social Research: Applied Social Research Methods*. Newbury Park, CA: Sage Publications; 1991.
42. Cohen J. *Statistical Power Analysis for the Behavioral Sciences*. Hillsdale, NJ: Lawrence Erlbaum; 1988.
43. Hedges LV. Distribution theory for Glass's estimator for effect size and related estimators. *J Educ Behav Stat*. 1981;6(2):107-128.
44. Hedges LV, Olkin I. *Statistical Methods for Meta-analysis*. New York, NY: Academic Press; 1985.
45. Kazis LE, Anderson JJ, Meenan RF. Effect sizes for interpreting changes in health status. *Med Care*. 1989; 27(3)(suppl):S178-S189.
46. Hedges LV, Vevea JL. Fixed- and random-effects models in meta-analysis. *Psychol Methods*. 1998;3(44):486-504.
47. Quintana SM, Minami T. Guidelines for meta-analyses of counseling psychology research. *Couns Psychol*. 2006;34(6):839-877.
48. Huedo-Medina TB, Sanchez-Meca J, Botella J, Marin-Martinez F. Assessing hetereogeneity in meta-analysis: Q statistic or I^2 index. *Psychol Methods*. 2006; 11(2):193-206.
49. Rosenthal R. The "file drawer problem" and tolerance for null results. *Psychol Bull*. 1979;86:638-641.
50. Begg CB, Mazumdar M. Operating characteristics of a rank correlation test for publication bias. *Biometrics*. 1994;50(4):1088-1101.
51. SPSS Inc. *SPSS 15.0 Command Syntax Reference*. Chicago, IL: SPSS Inc; 2006.
52. Rosenberg MS, Adams DC, Gurevitch J. *MetaWin. Statistical software for meta-analysis. Version 2.0*. Sunderland, MA: Sinauer Associates; 1999.
53. Jadad AR, Moore RA, Carrol D, et al. Assessing the quality of reports of randomized clinical trials: is blinding necessary? *Control Clin Trials*. 1996;17 (1):1-12.
54. Bachar E, Latzer Y, Kreitler S, Berry EM. Empirical comparison of two psychological therapies: self psychology and cognitive orientation in the treatment of anorexia and bulimia. *J Psychother Pract Res*. 1999; 8(2):115-128.
55. Barber JP, Morse JQ, Krakauer ID, Chittams J, Crits-Christoph K. Change in obsessive-compulsive and avoidant personality disorders following time-limited supportive-expressive therapy. *Psychother*. 1997; 34(2):133-143.
56. Bond M, Perry JC. Psychotropic medication use, personality disorder and improvement in long-term dynamic psychotherapy. *J Nerv Ment Dis*. 2006; 194(1):21-26.
57. Clarkin JF, Foelsch PA, Levy KN, Hull JW, Delaney MSW, Kernberg OF. The development of a psychodynamic treatment for patients with borderline personality disorder: a preliminary study of behavioral change. *J Personal Disord*. 2001;15(6):487-495.
58. Levy KN, Meehan KB, Kelly KM, et al. Change in attachment patterns and reflective function in a randomized control trial of transference-focused psychotherapy for borderline personality disorder. *J Consult Clin Psychol*. 2006;74(6):1027-1040.
59. Dare C, Eisler I, Russell G, Treasure J, Dodge L. Psychological therapies for adults with anorexia nervosa: randomised controlled trial of out-patient treatments. *Br J Psychiatry*. 2001;178:216-221.
60. Grande T, Dilg R, Jakobsen T, et al. Differential effects of two forms of psychoanalytic therapy: results of the Heidelberg-Berlin study. *Psychother Res*. 2006;16(4):470-485.
61. Gregory R, Chlebowski S, Kang D, et al. A controlled trial of psychodynamic psychotherapy for co-occurring borderline personality disorder and alcohol use disorder. *Psychother Theo, Res, Pract Train*. 2008; 45:28-41.
62. Høglend P, Amlo S, Marble A, et al. Analysis of the patient-therapist relationship in dynamic psychotherapy: an experimental study of transference interpretations. *Am J Psychiatry*. 2006;163(10): 1739-1746.
63. Høglend P, Bøgwald KP, Amlo S, et al. Transference interpretations in dynamic psychotherapy: do they really yield sustained effects? *Am J Psychiatry*. 2008; 165(6):763-771.
64. Huber D, Klug G, von Rad M. Die Münchner-Prozess-Outcome Studie: Ein Vergleich zwischen Psychoanalysen und psychodynamischen Psychotherapien unter besonderer Berücksichtigung therapiespezifischer Ergebnisse [The Munich process-outcome study: a comparison between psychoanalyses and psychotherapy]. In: Stuhr BM, ed. *Langzeit- Psychotherapie. Perspektiven für Therapeuten und Wissenschaftler*. Stuttgart, Germany: Kohlhammer; 2001: 260-270.
65. Huber D, Klug G. Munich Psychotherapy Study (MPS): The effectiveness of psychoanalytic longterm psychotherapy for depression. In: Society for Psychotherapy Research, ed. Book of abstracts. From research to practice. Ulm, Germany: Ulmer Textbank; 2006:154.
66. Knekt P, Lindfors O, Harkanen T, et al. Randomized trial on the effectiveness of long-and short-term psychodynamic psychotherapy and solution-focused therapy on psychiatric symptoms during a 3-year follow-up. *Psychol Med*. 2008;38(5):689-703.
67. Knekt P, Lindfors O, Laaksonen MA, et al. Effectiveness of short-term and long-term psychotherapy on work ability and functional capacity—a randomized clinical trial on depressive and anxiety disorders. *J Affect Disord*. 2008;107(1-3):95-106.
68. Korner A, Gerull F, Meares R, Stevenson. Borderline personality disorder treated with the conversational model: a replication study. *Compr Psychiatry*. 2006;47(5):406-411.
69. Leichsenring F, Biskup J, Kreische R, Staats H. The effectiveness of psychoanalytic therapy: first results of the Göttingen study of psychoanalytic and psychodynamic therapy. *Int J Psychoanal*. 2005;86(pt 2): 433-455.
70. Luborsky L, Stuart J, Friedman S, et al. The Penn Psychoanalytic Treatment Collection: a set of complete and recorded psychoanalyses as a research resource. *J Am Psychoanal Assoc*. 2001;49(1): 217-233.

71. Monsen J, Odland T, Faugli A, Daae E, Eilertsen DE. Personality disorders and psychosocial changes after intensive psychotherapy: a prospective follow-up study of an outpatient psychotherapy project, 5 years after end of treatment. *Scand J Psychol*. 1995; 36(3):256-268.
72. Monsen J, Odland T, Faugli A, Daae E, Eilertsen D. Personality disorders: changes and stability after intensive psychotherapy focusing on affect consciousness. *Psychother Res*. 1995;5(1):33-48.
73. Piper WE, Debbane EG, Bienvenu JP, Garant J. A comparative study of four forms of psychotherapy. *J Consult Clin Psychol*. 1984;52(2):268-279.
74. Rudolf G, Manz R, Öri C. Ergebnisse psychoanalytischer Therapien [Outcome of psychoanalytic therapies]. *Z Psychosom Med Psychoanal*. 1994; 40(1):25-40.
75. Sandell R, Blomberg J, Lazar A, Carlsson J, Broberg J, Schubert J. 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). *Int J Psychoanal*. 2000;81:921-942.
76. Stevenson J, Meares R. An outcome study of psychotherapy for patients with borderline personality disorder. *Am J Psychiatry*. 1992;149(3):358-362.
77. Svartberg M, Stiles T, Seltzer MH. Randomized, controlled trial of the effectiveness of short-term dynamic psychotherapy and cognitive therapy for Cluster C personality disorders. *Am J Psychiatry*. 2004; 161(5):810-817.
78. Vinnars B, Barber JP, Noren K, Gallop R, Weinryb RM. Manualized supportive-expressive psychotherapy versus nonmanualized community-delivered psychodynamic therapy for patients with personality disorders: bridging efficacy and effectiveness. *Am J Psychiatry*. 2005;162(10):1933-1940.
79. Wilczek A, Barber JP, Gustavsson JP, Asberg M, Weinryb RM. Change after long-term psychoanalytic psychotherapy. *J Am Psychoanal Assoc*. 2004; 52(4):1163-1184.
80. Higgins JP, Thompson SG, Deeks JJ, Altman DG. Measuring inconsistency in meta-analyses. *BMJ*. 2003; 327(7414):557-560.
81. Petrak F, Nickel R, Hardt J, Egle UT. Verhaltenstherapie versus psychodynamisch-interaktionelle Therapie bei Patienten mit somatoformen Schmerzen: 1-Jahres-Follow-up einer randomisierten, kontrollierten Studie [Cognitive behaviour therapy vs psychodynamic interactional therapy for patients with somatoform pain: 1-year-follow-up of a randomized controlled trial]. *Verhaltenstherapie*. 2007;17(S1): S19-S20.
82. Begg C. Publication bias. In: Cooper H, Hedges LV, eds. *The Handbook of Research Synthesis*. New York, NY: Russel Sage Foundation; 1995:399-409.
83. Roth A, Fonagy P. *What Works for Whom? A Critical Review of Psychotherapy Research*. 2nd ed. New York, NY: Guilford Press; 2005.
84. Shadish WR, Matt GE, Navarro AM, Phillips G. The effects of psychological therapies under clinically representative conditions: a meta-analysis. *Psychol Bull*. 2000;126(4):512-529.
85. Brockmann J, Schlüter T, Eckert J. Langzeitwirkungen psychoanalytischer und verhaltenstherapeutischer Langzeittherapien [Long-term outcome of long-term psychoanalytic and behavioral long-term therapy]. *Psychotherapeut*. 2006;51(1):15-25.
86. Puschner B, Kraft S, Kächele H, Kordy H. Course of improvement over 2 years in psychoanalytic and psychodynamic outpatient psychotherapy. *Psychol Psychother*. 2007;80(1):51-68.
87. Giesen-Bloo J, van Dyck R, Spinhoven P, et al. Outpatient psychotherapy for borderline personality disorder: randomized trial of schema-focused therapy vs transference-focused psychotherapy. *Arch Gen Psychiatry*. 2006;63(6):649-658.
88. Bateman AW, Fonagy P. The development of an attachment-based treatment program for borderline personality disorder. *Bull Menninger Clin*. 2003; 67(3):187-211.
89. de Maat S, Philipszoon F, Schoevers R, Dekker J, de Jonghe F. Costs and benefits of long-term psychoanalytic therapy: changes in health care use and work impairment. *Harv Rev Psychiatry*. 2007;15(6): 289-300.
90. Dührssen A, Jorswieck E. Eine empirisch-statistische Untersuchung zur Leistungsfähigkeit psychoanalytischer Behandlung [An empirical-statistical-study of the effectiveness of psychoanalytic treatment]. *Nervenarzt*. 1965;36(4):166-169.