VU Research Portal

The effectiveness of long-term psychoanalytic therapy: A systematic review of empirical studies.

de Maat, S.; de Jonghe, F.; Schoevers, R.A.; Dekker, J.J.M.

Link to publication in VU Research Portal

The Effectiveness of Long-Term Psychoanalytic Therapy: A Systematic Review of Empirical Studies

Saskia de Maat, PhD, Frans de Jonghe, MD, PhD, Robert Schoevers, MD, PhD, and Jack Dekker, PhD

Background: There is a gap in the research literature on the effectiveness of long-term psychoanalytic therapies (LPT). *Aim*: To present a systematic review of studies dealing with LPT effectiveness and published from 1970 onward. *Methods*: A systematic literature search for studies dealing with the effectiveness of individual LPT in ambulatory, adult patients. Data about the overall effectiveness of LPT, its impact on symptom reduction, and its effect on personality changes were pooled both at treatment termination and at follow-up, using effect sizes (ESs) and success rates. *Results*: We found 27 studies ($n = 5063$). Psychotherapy yielded large mean ESs (0.78 at termination; 0.94 at follow-up) and high mean overall success rates (64% at termination; 55% at follow-up) in moderate/mixed pathology. The mean ES was larger for symptom reduction (1.03) than for personality change (0.54). In severe pathology, the results were similar. Psychoanalysis achieved large mean ESs (0.87 at termination; 1.18 at follow-up) and high mean overall success rates (71% at termination; 54% at follow-up) in moderate pathology. The mean ES for symptom reduction was larger (1.38) than for personality change (0.76). *Conclusion*: Our data suggest that LPT is effective treatment for a large range of pathologies, with moderate to large effects. (HARV REV PSYCHIATRY 2009;17:1–23.)

Keywords: effectiveness, psychoanalysis, psychoanalytic therapy, psychotherapy, review

The effectiveness of long-term psychoanalytic therapies (LPT)—consisting of psychoanalysis "proper" and long-term psychoanalytic psychotherapy—has been studied from the very beginning, including by Coriat,[1] Fenichel,[2] Kessel and Hyman,[3] Jones,[4] Alexander,[5] Knight,[6] Schjelderup,[7] Orgel,[8]

© 2009 President and Fellows of Harvard College

Knapp and colleagues,[9] Graham,[10] Klein,[11] Cremerius,[12] Bieber and colleagues,[13] Hamburg and colleagues,[14] and Feldman.[15] These studies can be seen as the first attempts to provide empirical evidence of LPT's effectiveness, but they do not meet contemporary criteria of scientific research. In recent decades, efforts have been made to bring LPT effectiveness research more into line with current standards of evidence-based medicine.

Of the recent reviews and overviews in this field, several are especially noteworthy. Bachrach and colleagues[16] performed an extensive review of literature dealing with the effectiveness of psychoanalysis. They discussed at length the methodology, design, and results of six systematic studies, including a total of 550 patients. The authors found success rates (meaning substantial therapeutic benefit) in the 60% to 90% range, and significant effect sizes (ESs). Doidge[17] conducted an overview of the empirical evidence for psychoanalytic psychotherapies and psychoanalysis. He concluded: "There is considerable experimental and clinical data for efficacy of a range of individual psychoanalytic psychotherapies, from short-term psychoanalytic psychotherapy to

1

long-term psychoanalytic psychotherapy and psychoanalysis." Fonagy's Open Door Review[18] described the study designs and outcomes of scores of LPT studies. Its importance is not easily overrated. It covered both process and outcome studies in a wide variety of patient groups. Probably because of the broad range of these studies, Fonagy did not pool the data or aim to reach general conclusions. The clinical heterogeneity of the studies would certainly have precluded the latter. Fonagy evaluated each study separately and listed the strengths and weaknesses of the designs, but the methodological quality of the studies was not systematically assessed or expressed as a quality score. Doidge[19] evaluated nine quantitative studies of psychoanalysis and concluded that "for the properly chosen patient, psychoanalysis is effective in terms of symptom relief, character changes, and conflict resolution." The review of Leichsenring[20] dealt with short-term psychodynamic psychotherapy and long-term psychoanalytic therapy. He identified 22 randomized, controlled trials (RCTs), all of which dealt with short-term psychotherapy. In another article, Leichsenring[21] reviewed 4 studies of long-term psychoanalytic therapy that met the highest quality requirements for naturalistic effectiveness studies. He concluded that LPT was more effective than shorter forms of psychodynamic psychotherapy and that it yielded ESs that significantly exceeded the effects of untreated or low-dose-treated comparison groups. The strength of his study was that it included only the highest level of evidence for effectiveness studies. The data from individual studies were not pooled. A recent meta-analysis of Leichsenring and Rabung[22] addressed the effectiveness of long-term psychodynamic psychotherapy in complex mental disorders. It is an important, extensive, and well-performed meta-analysis that includes 23 studies involving 1,053 patients. The authors concluded that long-term psychodynamic psychotherapy showed significantly higher outcomes on several measures than shorter forms of psychotherapy and that pre/post ESs of long-term psychodynamic psychotherapy fell in a range between 0.78 and 1.98, reflecting, by definition, a large effect. The meta-analysis concentrated on psychotherapy and did not address psychoanalysis.

In short, the more recent reviews are of good quality, but many of them do not pool the data (e.g., calculate mean ESs or mean success rates) from individual studies. Reviews including psychoanalysis are scarce. Although some reviews explicitly address study quality, they do not use systematic quality scores.

The aim of this article is to present a systematic review of LPT studies (psychotherapy or psychoanalysis) published between 1970 and 2007. To enhance clinical homogeneity, we focus on individual, ambulatory LPT with adult patients having "regular" indications for psychoanalytic therapy. Whenever possible, we divide our results according to treatment (psychoanalysis or psychotherapy), assessment source (patient or therapist), outcome (symptom reduction or personality change), and severity of pathology (moderate/mixed or severe [personality disorders]). We perform a systematic assessment of study quality using an explicit quality criterion. Furthermore, we attempt to pool the data about the effectiveness of LPT. Our pooling is restricted to assessing success rates and ESs in individual studies and to calculating their weighted means among comparable studies. Finally, we compare the results of studies meeting our quality criterion with the results of lower-quality studies.

LONG-TERM PSYCHOANALYTIC THERAPIES

We define "long-term" as therapy consisting of at least 50 sessions and lasting at least one year. Psychoanalysis is always long-term; psychotherapy can be short-term or long-term. LPT therefore includes psychoanalysis and long-term psychoanalytic, including psychodynamic, psychotherapy.

The shared feature of these treatments is that they are rooted in psychoanalytic theories. We differentiate between the two treatments according to principles generally accepted by psychoanalytic therapists. Two easily identifiable aspects, therapy setting and session frequency, discriminate between these types of LPT. In psychoanalysis "proper," the patient lies on a couch, and there are at least three sessions a week. In psychotherapy, patients sit across from the therapist, and there are no more than two sessions a week. Interpretation is the hallmark of psychoanalysis; psychotherapy moves on a continuum between the poles of interpretation and support.

The ultimate goals of LPT are symptom reduction, prevention of recurrence, better social functioning, higher quality of life, and higher life satisfaction, preferably for a long time after treatment termination. These goals are by no means specific to LPT. The distinctive feature of LPT lies in its intermediate goals, which focus on bringing about changes in some aspects of the patient's personality. These changes, meant to be lasting, should enable patients to meet the problems of living more successfully and to make better use of their personal potential and the opportunities afforded by their lives. In other words, people's vulnerabilities are reduced, and their strengths and resources are developed. The aim of LPT is to stimulate this development. Successful LPT does not, by definition, make patients happier; it enhances the possibility that patients will be happy when there is reason to be. It also enhances the possibility of there being such reasons, since people are (up to a certain point) the creators of their own circumstances and of their attitudes toward those circumstances. In psychoanalytic terms, the changes in personality are described as: "structural change," "personality change," "personality

reconstruction or construction," or the development of a "cohesive," "adult," "integrated" self. Freud[23] put it "simply" when stating that psychosynthesis is the ultimate goal of psychoanalysis ("So vollzieht sich bei dem analytisch Behandelten die Psychosynthese").

METHODS

Literature Search Strategy

A literature search was conducted in Pubmed, Embase, the Cochrane Database of Systematic Reviews, the Cochrane Central Register of Controlled Trials, PsycLit, and the ACP Journal Club. The following subject headings were applied: long-term psychotherapy, psychoanalysis, psychoanalytic psychotherapy, psychodynamic therapy. The time limit was between 1970 and 2007 (May), with no limits set on language. Cross-references in the retrieved publications were tracked down. The Open Door Review[18] was an important source of studies regarding psychoanalytic therapies. There was no systematic effort to find additional unpublished data.

The following inclusion criteria were applied:

1. The studies were required to be "outcome-intervention studies." The outcomes had to be measured in terms of symptom reduction or personality change. Issues such as process variables or therapeutic variables were excluded from this review.
2. The studies were required to be RCT or cohort studies. Case studies or case series were excluded, but surveys (such as consumer reports or questionnaires among therapists) were included.
3. The studies were required to deal with individual, ambulatory psychoanalytic therapies with adult patients (18 to 65 years of age). Studies dealing with children or the elderly, as well as studies conducted in clinical[24] or day-care settings were excluded.
4. The studies were required to include patients with regular indications for psychoanalytic therapies (i.e., DSM diagnoses [Axis I or II] or otherwise specified symptoms or personality problems). We excluded psychotherapeutic studies for psychotic disorders such as schizophrenia;[25] studies in connection with somatic disorders such as juvenile-onset insuline-dependent diabetes;[26] studies with an exclusive focus on eating disorders[27] or on comparatively rare diagnoses such as Munchhausen syndrome by proxy; and studies of learning difficulties in children.[28]
5. Treatment was required to last a minimum of one year and to involve at least 50 sessions.

Identification of Relevant Publications and Quality Assessment

Two reviewers checked the search results (SdM and FdJ) and requested all potentially suitable studies. They conducted independent checks of all identified studies. Any disagreements about whether a study should be included were resolved by discussion with a third member of staff (JD). The reviewers evaluated the quality of the studies independently using a self-developed Research Quality Score (RQS; see appendix). The list has not yet been validated beyond the consensus between the two raters (SdM and FdJ) who evaluated the quality of the studies. Its criteria, which are based on those postulated by the Cochrane Collaboration[29] and by Leichsenring,[21] deal with study design, patients included, interventions, outcome data, statistics, dropout, and follow-up.

We defined half of the maximum RQS score to be the minimum for selection as a study with acceptable research quality. Since this cut-off score was ultimately arbitrary, we also evaluated the studies falling below the cutoff score and compared their results with those of the studies meeting our criterion. We distinguish between three design types, each with their own criteria (which overlap extensively), maximum scores, and cutoff scores. A type 1 study is a cohort study without matched control groups, although it may include multiple cohorts. A type 2 study is a cohort study with a matched control group. A matched control group is considered to be either an untreated, matched group with characteristics highly similar to those of the intervention group, or a matched group treated with a therapy that is already considered evidence based. We classify studies with both a psychotherapy group and a psychoanalysis group, even if matched, as type 1 studies: a cohort study consisting of two cohorts. The two groups cannot act as controls for each other, as the indications for "psychoanalysis" versus "psychotherapy"—which take into account considerations such as ego strength, affect tolerance, and capacity for reflection and insight—make the two patient groups inherently different. A type 3 study is a RCT. It goes without saying that a type 3 study ranks higher than a type 2 study, which, in turn, ranks higher than a type 1 study. By the same token, type 3 studies must achieve a higher cutoff score than type 2 studies, and type 2 studies a higher cutoff score than type 1 studies. Nevertheless, each study can be, within the limits of its type, well designed and well performed—that is, achieve the cutoff score appropriate for that research type. In addition, the presence or absence of a follow-up phase is considered a separate criterion from our RQS; otherwise, a good study without follow-up could be discarded, whereas a relatively flawed study with follow-up could be accepted.

Effectiveness Data and Calculations of Results

We first summarized the findings of studies meeting our quality criterion. The primary outcome data were pre/post and pre-follow-up ESs,[30] which we considered to be the most "robust" outcome measures. First, they require pre- and posttreatment assessments using a specific questionnaire, and second, their standardization allows for comparison. Consequently, since any particular ES expresses a particular degree of relative change, the amount of change in different studies—on different scales in different populations—can be compared. ESs less than 0.5 are considered small; from 0.5 to 0.8, medium; and greater than 0.8, large.[30] For each study that presented more than one ES, a mean ES for the study was calculated. The ESs of individual studies were used, in turn, as the basis for calculating an overall mean ES, with the individual study ESs weighted to reflect the study's sample size. As some outcome measures may be more sensitive to change than others,[31,32] we also calculated separate mean ESs for scales concerning symptoms (e.g., the 90-item Symptom Checklist–90–Revised [SCL-90-R])[33] and for scales considered to measure aspects of personality (e.g., the Sense of Coherence Scale [SOC][34] and Social Adjustment Scale [SAS]).[35]

Success rates were adopted as the secondary outcome measure. We consider measures of success rates to be less "robust" since they consist of more general assessments such as appraisals from patients and therapists of therapy success. Scales and criteria varied per study. We defined success as "at least moderate improvement." Obviously, the definition and interpretation of success varies, depending on the scale used. Therefore we mention the scales used in each individual study. Some scales are more global (such as the five- or seven-step Likert scales used in the Clinical Global Assessment [CGI])[36] and do not always differentiate between symptoms or personality changes. However, some studies used symptom scales like the SCL-90-R to measure success rates, and other studies used scales relating to personality aspects (e.g., capacity to enjoy, interpersonal capacities, or the capacity to deal with conflicts and interpersonal relationships). Whenever possible, we divided the results into general/symptom-related assessments and personality-related assessments, and then divided these results, in turn, into groups reflecting either therapist or patient assessments. For each table, we also calculated overall success rates, with the success rates of all studies pooled and weighted to take into account the sample size in each study.

Breakdown of the Results

We broke down our results using five types of differentiation:

1. psychoanalysis "proper" and psychoanalytic therapy (referred to hereinafter as "psychotherapy")
2. "therapist assessment" and "patient assessment"
3. pre/post outcomes and pre-follow-up outcomes, with "post" defined as "at treatment termination," and follow-up as a given period after treatment termination
4. results relating to symptom reduction (e.g., SCL-90-R) and results relating to personality change (e.g., SOC or Inventory of Interpersonal Relationships [IIP])[37]
5. "moderate/mixed" pathology and "severe" pathology, with "moderate/mixed" pathology reflecting patients meeting the full range of "regular" indications for ambulatory LPT as indicated by the extensive survey of Doidge and colleagues,[38,39] and "severe" pathology being reserved for patients in studies that included *only* patients with more severe personality disorders such as borderline personality disorder

In the study just mentioned, the authors conducted a survey of 510 analysts in Australia, Canada, and the United States.[38,39] The results showed that 71% of the patients had a DSM-III-R personality disorder, 68% a major depressive disorder, and 57% an anxiety disorder. Only 3% of the patients did not present a DSM-III-R diagnosis. In addition, comorbidity was high, with the modus of patients presenting three disorders. Although these clinical pictures are complicated diagnostically, we labeled pathology in patients treated in ambulatory psychoanalytic practices to be "moderate/mixed," thereby conforming our usage to the findings of Doidge and colleagues.[38,39]

RESULTS

A total of 742 studies was initially found. Based on screening of titles and abstracts, 672 studies were excluded in the first selection round. Reasons for exclusion were as follows: patients samples based on a somatic illness and not a psychiatric illness; case studies; theoretical articles; epidemiological studies; studies dealing with training of professional therapists; studies dealing with alcohol and drugs; patient samples concerning only geriatric or psychotic patients; rare diagnoses such as Munchausen by proxy; and articles dealing with methodological issues. The remaining 73 articles were retrieved, and a further exclusion of 46 studies followed in the second round. Reasons for exclusion in this round were as follows: process-outcome studies; studies focusing on children; reviews; studies on short-term therapies; studies focusing on therapist variables; and studies on inpatients. Finally, 27 outcome studies were retrieved concerning LPT in regular ambulatory patients. Table 1 presents an overview of those 27 studies. Tables 2 to 4 show the results of the 19 studies that met the quality criterion. Table 5

presents the results of the remaining 8 studies. Finally, Table 6 summarizes all findings.

Included Studies

As can be seen in Table 1, we found 27 relevant publications meeting our inclusion and exclusion criteria. These studies included 1 RCT, 5 surveys, and 21 cohort studies. Of the cohort studies, 16 were prospective, and 5 retrospective. Four surveys[45,51,52,54] were replications of the study by Seligman.[76] The survey by Stehle[67] was based on inquiries among therapists. In all, the studies covered 5,063 patients (3,632 in psychotherapy, and 1,431 in psychoanalysis). The mean number of sessions was approximately 500 (mean = of 3.6 years, with approximately 140 sessions/year) in psychoanalysis and 150 (mean = 2.5 years, with approximately 60 sessions/year) in psychotherapy. The large majority of the studies looked at patients with "moderate/mixed" pathology; only five studies[48,59,60,68,69] were considered to address "severe pathology (only severe personality disorders)." In the study of Keller,[54] we categorized the psychoanalysis condition ($n = 84$) under the heading of "psychotherapy" because the mean frequency of the sessions was 1.6 per week. For similar reasons the study of Stehle[67] was categorized under psychotherapy, although some of the patients received psychoanalysis.

Study Quality

A total of eight studies (30%)[43–47,51,52,66,67] did not meet the quality criterion (see appendix). Six of these studies were retrospective studies, and five of them were surveys (sent to either therapists or patients). The fact that the studies were retrospective accounted for many of the low quality scores. Other factors with a negative effect on the RQS of studies were as follows: not stating inclusion and exclusion criteria for patient selection, not describing the characteristics of the patient sample or the treatment applied, no intention-to-treat analyses, no independent or blind assessments, not defining or describing dropouts, and moderate quality of assessment instruments (e.g, only a five-step Likert scale addressing overall therapy success).

Of the remaining 19 studies, which met the quality criterion, only 1 was an RCT.[48] Two made an attempt at a matched control group,[20,53] and the other 16 were single- or multiple-cohort studies. Examples of cohort studies meeting most of the quality criteria were the studies of Brockman and colleagues,[41] Grande and colleagues,[49] and Leichsenring.[20] The studies were well designed and prospective; patient samples were clearly described; treatments were described; treatment adherence was controlled for (although not in Brockman and colleagues);[41] assessment instruments were well qualified; there were mul-

tiple assessors (although none of the studies included blind assessment); follow-up results were included; and the Leichsenring[20] study attempted to compare the sample to a control group.

LPT Effect Sizes

Tables 2.1 to 2.3 present the ESs for psychoanalysis and psychotherapy in moderate/mixed studies and in those including only patients with severe pathology.

As can be seen in Table 2.1, the majority of the studies of psychotherapy for moderate/mixed pathology reported large ESs, both at treatment termination (65%) and at follow-up (69%). The weighted mean ESs were large at both termination (0.78) and follow-up (0.94). We calculated mean ESs separately for symptom reduction and personality change (scales marked S and P, respectively, in the table); scales that provided only global assessments were not included in the table (and are marked with dashes). All ESs at treatment termination, as well as at follow-up, were included in this calculation of the mean ESs. The weighted mean ESs were 1.03 ($n = 572$; SD = 0.59) for symptom reduction and 0.54 ($n = 599$; SD = 0.33) for personality change.

As can be seen in Table 2.2, approximately half of the studies dealing with psychotherapy in severe pathology reported large ESs, both at treatment termination (46%) and follow-up (56%). The weighted mean ESs were large at both termination (0.94) and follow-up (1.02). However, the latter was a result from only one study. The majority of the studies dealt with personality aspects. Leaving out the one study with no results concerning personality change and using all other ESs, the overall ES for personality change was 1.11 ($n = 67$; SD = 0.02).

As can be seen in Table 2.3, which regards psychoanalysis for moderate/mixed pathology, 50% of the studies reported large ESs at treatment termination, whereas 80% of the studies reported large ESs at follow-up. The weighted mean ESs were large both at termination (0.87) and at follow-up (1.18). We made separate calculations of mean ESs for symptom reduction and for personality change. All ESs at treatment termination and at follow-up were included in these calculations. The weighted mean ESs were 1.38 ($n = 150$; SD = 0.27) for symptom reduction and 0.76 ($n = 186$; SD = 0.27) for personality change. Since the study of Weber[72] presented ESs for only part of the psychoanalysis group (range, 0.40–0.50), we did not include them in the table. We did not find ES data for psychoanalysis for severe pathology.

LPT Success Rates

Symptom-related or general assessments. Tables 3.1 to 3.4 present the success rates for psychoanalysis and psychotherapy

TABLE 1. Studies Included in Review

Author	Study type & Research Quality Score	Treatments (n)[a]	Diagnosis	Duration[b]	Outcome source[c]	Measure of effectiveness	Assessment points
Brockman et al. (2003)[40] Brockman et al. (2002)[41]	Cohort, prospective RQS: 38	Psychotherapy (31)	Moderate pathology (DSM-III-R depression and anxiety disorders)	3 years	Patient: + Therapist: + Observer: + Other: −	Symptoms: SCL-90-R, GSI Personality: IIP Goal attainment: GAS	Pre: + During: + Post: + Follow-up: +
Clarkin et al. (2001)[42]	Cohort, prospective RQS: 41	Psychotherapy (23)	Severe pathology (borderline personality disorder) plus more than one Axis I disorder	1 year	Patient: − Therapist: + Observer: + Other: −	Symptoms: Parasuicidal history interview, treatment history interview Undifferentiated: GAF	Pre: + During: − Post: + Follow-up: −
Dührssen (1986)[43]	Cohort, prospective RQS: 24	Psychotherapy (60) Psychoanalysis (30)	Moderate pathology (unspecified)	Unspecified	Patient: + Therapist: − Observer: − Other: +	Symptoms: FPI, GT, GBB Personality: GT	Pre: + During: − Post: + Follow-up: −
Erle et al. (2003)[44]	Cohort, retrospective (1973–77) RQS: 18	Psychoanalysis (161)	Moderate pathology (unspecified)	5–6 years	Patient: − Therapist: + Observer: − Other: -	Undifferentiated: CGI	Pre: − During: − Post: + Follow-up: −
Erle et al. (2003)[44]	Cohort, prospective (1984–89) RQS: 28	Psychoanalysis (92)	Moderate pathology (unspecified)	5 years	Patient: − Therapist: + Observer: − Other: −	Undifferentiated: CGI	Pre: + During: + Post: + Follow-up: −
Freedman et al. (1999)[45]	Survey, retrospective RQS: 18	Psychotherapy (99)	Moderate pathology (personality disorder, 12%; depression, 58%; anxiety, 17%)	1 year	Patient: + Therapist: − Observer: − Other: −	Treatment satisfaction: Effectiveness Questionnaire	Pre: − During: + Post: − Follow-up: +
Friedman et al. (1998)[46] Friedman et al. (2005)[47]	Cohort, retrospective RQS: 21	Psychotherapy (575)	Moderate pathology (at least 1 Axis I disorder, 88%; at least 1 Axis II disorder, 59%)	1 year	Patient: − Therapist: + Observer: − Other: −	Symptoms: structured interview Personality: structured interview Undifferentiated: GAF	Pre: + During: + Post: − Follow-up: +

Study	Design/RQS	Treatment (n)	Pathology	Duration	Rater results	Outcome measures	Time results
Giesen-Bloo et al. (2006)[48]	RCT RQS: 77	Psychotherapy (42)	Severe pathology (BPDSI-IV borderline personality disorder)	3 years	Patient: – Therapist: + Observer: + Other: –	Symptoms and personality: psychopathology, personality pathology; Personality: BPDSI-IV and others; Life satisfaction: EuroQol, WHOQOL	Pre: + During: + Post: + Follow-up: +
Grande et al. (2006)[49]	Cohort, prospective RQS: 48	Psychotherapy (27)	Moderate pathology (psychoneurosis, 55%; personality disorder, 31%)	2 years	Patient: + Therapist: + Observer: + Other: +	Symptoms: SCL-90-R, GSI, PSKB-Se; Personality: IIP; Social functioning: GAF [SOFAS]	Pre: + During: – Post: + Follow-up: +
Rudolf et al. (2004)[50]		Psychoanalysis (31)		2.5 years			
Hartmann et al. (2004)[51]	Survey, retrospective RQS: 16	Psychotherapy (448) Psychoanalysis (263)	Moderate pathology (unspecified)	2 years	Patient: + Therapist: – Observer: – Other: –	Treatment satisfaction: Effectiveness Questionnaire	Pre: – During: + Post: – Follow-up: +
Heinzel et al. (1998)[52]	Survey, retrospective RQS: 22	Psychotherapy (633)	Moderate pathology (unspecified)	3.4 years	Patient: + Therapist: – Observer: – Other: –	Symptoms and social functioning: 5-point rating scale	Pre: + During: – Post: + Follow-up: +
Holm-Hadulla et al. (1997)[53]	Cohort, prospective RQS: 38	Psychotherapy (37)	Moderate pathology (personality disorder, 27%; dysthymic disorder, 32%; adjustment disorder, 35%)	1.5 years	Patient: + Therapist: + Observer: – Other: –	Symptoms: CGI; Personality: CGI; Symptoms and social functioning (global): 21-item questionnaire	Pre: + During: – Post: + Follow-up: –
Keller et al. (1998)[54]	Survey, retrospective RQS: 29	Psychotherapy (111)	Moderate pathology (personality disorder, 17%; depression, 46%)	2.7 years	Patient: + Therapist: + Observer: + Other: –	Symptoms: SCL-90-R, BSS; Personality: GT	Pre: – During: – Post: – Follow-up: +
Leichsenring et al. (2005)[55]	Cohort, prospective RQS: 44	Psychoanalysis (36)	Moderate pathology (personality disorder, 69%; affective disorder, 69%; phobic disorder, 50%)	3 years	Patient: + Therapist: + Observer: – Other: –	General well-being: VEV; Symptoms: CGI, SCL-90-R; Personality: CGI, IIP; Social functioning: CGI; Life satisfaction: CGI, FLZ	Pre: + During: + Post: + Follow-up: +

TABLE 1. Studies Included in Review (*Continued*)

Author	Study type & Research Quality Score	Treatments (n)[a]	Diagnosis	Duration[b]	Outcome source[c]	Measure of effectiveness	Assessment points
Leuzinger-Bohleber et al. (2001, 2003)[56,57]	Cohort, retrospective RQS: 31	Psychotherapy (194) Psychoanalysis (207)	Moderate pathology (personality disorder, 51%; affective disorder, 27%)	4 years 4 years	Patient: + Therapist: + Observer: + Other: −	Symptoms: SCL-90, BSS, GAF (SOFAS) Personality: SOC Life satisfaction: IRES Treatment satisfaction (measured on a 5-point Likert scale)	Pre: − During: − Post: − Follow-up: +
Luborsky et al. (2001)[58]	Cohort, prospective RQS: 30	Psychoanalysis (17)	Moderate pathology (unspecified)	3 years	Patient: + Therapist: + Observer: + Other: −	Symptoms: HSRS Undifferentiated: GAF Treatment satisfaction: SSI	Pre: + During: + Post: + Follow-up: +
Monsen et al. (1995, 1995)[59,60]	Cohort, prospective RQS: 49	Psychotherapy (25)	Severe pathology, as determined through SCID (personality disorder, 96%; depression, 52%; anxiety, 20%)	2 years	Patient: + Therapist: + Observer: + Other: −	Symptoms: HSRS, SCL-90-R Personality: MMPI, affect consciousness measured by semistructured interview Social functioning: level of social adjustment	Pre: + During: − Post: + Follow-up: +
Von Rad et al. (1998)[61] Heuft et al. (1996)[62] Kordy et al. (1988)[63]	Cohort, prospective RQS:42	Psychotherapy (33) Psychoanalysis (36)	Moderate pathology (psychiatric disorder, 58%)	3.7 years 2.2 years	Patient: + Therapist: + Observer: + Other: −	Symptoms: CGI Personality: GT, CGI Social functioning, treatment satisfaction: FKBS	Pre: + During: + Post: + Follow-up: +
Rudolf et al. (1994)[64]	Cohort, prospective RQS: 32	Psychotherapy (56) Psychoanalysis (44)	Moderate pathology (psychoneurosis, 54%; personality disorder, 20%; psychosomatic, 13%)	1.5 years 1.7 years	Patient: + Therapist: + Observer: − Other: −	Symptoms: global impression on a 5-point Likert scale Personality: PSKB Social functioning: PSKB-Se	Pre: + During: − Post: + Follow-up: −
Sandell et al. (2000)[65]	Cohort, prospective RQS: 35	Psychotherapy (331) Psychoanalysis (74)	Moderate pathology (unspecified)	4 years 4.5 years	Patient: + Therapist: − Observer: − Other: −	Symptoms: SCL-90-R Personality: SOC Social functioning: SAS Caseness criterion: on 3 scales, in worst 10%	Pre: + During: + Post: + Follow-up: +

Study	Design	Treatment (n)	Pathology	Duration	Outcome measures	Observers	Time points
Sashin et al. (1975)[66]	Cohort, retrospective (1959–66) RQS: 21	Psychoanalysis (183)	Moderate pathology (hysteric character neurosis [some with depression], 30%; obsessive-compulsive neurosis, 18%)	4.2 years	Symptoms: CGI Personality: CGI Social functioning: CGI	Patient: − Therapist: + Observer: − Other: −	Pre: + During: − Post: − Follow-up: +
Stehle et al. (2004)[67]	Survey, retrospective RQS: 22	Psychotherapy (581)	Moderate pathology (neurosis, 60%; personality disorder, 31%)	3 years	Symptoms: CGI Personality: CGI	Patient: − Therapist: + Observer: − Other: −	Pre: − During: − Post: + Follow-up: −
Stevenson & Meares (1992)[68]	Cohort, prospective RQS: 48	Psychotherapy (48)	Severe pathology (borderline personality disorder)	1 year	Symptoms: Cornell Index	Patient: + Therapist: − Observer: + Other: +	Pre: + During: − Post: − Follow-up: +
Wallerstein (1986)[69] Kernberg (1972)[70] Harty (1976)[71]	Cohort, prospective RQS: 54	Psychotherapy (20) Psychoanalysis (22)	Severe pathology (majority: borderline personality disorder)	Unspecified.	Symptoms: HSRS Personality: CGI	Patient: − Therapist: + Observer: − Other: −	Pre: + During: − Post: + Follow-up: −
Weber et al. (1985)[72] Bachrach et al. (1985)[73]	Cohort, retrospective RQS: 27	Psychotherapy (138) Psychoanalysis (235)	Moderate pathology (unspecified)	0.75 year 2.5 years	Symptoms: CGI Personality: CGI, ego strength scale Social functioning: social relations scale, work gratification scale	Patient: − Therapist: + Observer: − Other: −	Pre: + During: − Post: + Follow-up: −
Weber et al. (1985)[74] Bachrach et al. (1985)[73]	Cohort, prospective RQS: 35	Psychotherapy (29) Psychoanalysis (36)	Moderate pathology (unspecified)	2 years 3 years	Symptoms: CGI Personality: CGI, ego strength scale Social functioning: social relations scale, work gratification scale	Patient: − Therapist: + Observer: − Other: −	Pre: + During: - Post: + Follow-up: −
Wilczek et al. (2004)[75]	Cohort, prospective RQS: 39	Psychotherapy (55)	Moderate pathology (DSM-III-R Axis I, 53%; DSM-III-R Axis II, 11%)	3 years	Symptoms: DSM-III-R, CPRS-S-A Personality: KAPP, KSP	Patient: + Therapist: − Observer: + Other: −	Pre: + During: − Post: − Follow-up: +

BPDSI-IV, Borderline Personality Disorder Structured Interview; BSS, Beschweden-Schwere-Score; CGI, Clinical Global Impression; CPRS-S-A, Comprehensive Psychopathological Self-Rating Scale for Affective Syndromes; DSM, *Diagnostic and Statistical Manual of Mental Disorders*; EuroQol, European Quality of Life thermometer; FKBS, Fragebogen zu Konfliktbewältigungsstrategien; FLZ, Fragen zur Lebenszufriedenheit; FPI, Freiburger Persönlichkeitsinventar; GAF, Global Assessment of Functioning; GAS, Goal Attainment Scale; GBB, Giessener Beschwerde Bogen; GSI, General Severity Index; GT, Giessen Test; HSRS, Health-Sickness Rating Scale; IIP, Inventory of Interpersonal Problems; IRES, Life satisfaction scale; KAPP, Karolinska Psychodynamic Profile; KSP, Karolinska Scales of Personality; MMPI, Minnesota Multiphasic Personality Inventory; PSKB[-Sel], Psychischer und Sozial-Kommunikativer Befund[–Selbstbeurteilung]; SAS, Social Adjustment Scale; SCID, Structural Clinical Interview; SCL-90-R, Symptom Checklist–90–Revised; SOC, Sense of Coherence Scale; SOFAS, Social and Occupational Functioning Assessment Scale; SSI, Combination of "success," "satisfaction," and "improvement"; VEV, Veränderungsfragebogen des Erlebens und Verhaltens; WHOQOL, World Health Organization Quality of Life assessment.

aFor psychoanalytic psychotherapy, there were one to two sessions per week; for psychoanalysis, three to five sessions.
bApproximation mean duration of therapy.
cThe "observers" (below) were independent.

TABLE 2.1. Psychotherapy Effect Sizes for Studies with Moderate/Mixed Pathology

Study	n	Scale: P/S[b]	Pre/post ES[a]			Pre-follow-up ES[a]			Mean ES	
			Small	Medium	Large	Small	Medium	Large	Post	Follow-up
Brockman et al. (2003)[40]	31	SCL-90-R (GSI): S			1.37				1.28	
		IIP (relationships): P								
Brockman et al. (2002)[41c]					1.19					
Grande et al. (2006)[49]	27	SCL-90-R (GSI): S			0.92			1.04	0.80	0.95 (1 yr)
		IIP (relationships): P		0.67				0.85		
Holm-Hadulla et al. (1997)[53]	37	Global judgment by patient: −			1.13				1.05	
		Global judgment by therapist: −			0.96					
Keller et al. (1998)[54]	111	BSS: S						2.10		2.10 (6 yrs)
Leichsenring et al. (2005)[55]	36	SCL-90-R (GSI): S			1.34			1.38	1.64	1.88 (1 yr)
		IIP (relationships): P			1.28			1.85		
		Life satisfaction: P			1.55			1.81		
		GAS: −			2.39			2.48		
Sandell et al. (2000)[65]	331	SCL-90-R: S					0.60			0.46 (3 yrs)
		SOC: P				0.34				
		SAS: P				0.44				
Weber et al. (1985)[72]	138	Ego strength: P		0.59						
		Social relations: P	0.43						0.51	
		Work gratification: P		0.50						
Weber et al. (1985)[74]	29	Ego strength: P	0.11						0.11	
Wilczek et al. (2004)[75]	36	KAPP: P				0.30				0.82 (0.5 yr)
		KSP: P				0.40				
		GAF: −						0.87		
		CPRS (anxiety): S						0.99		
		CPRS (depression): S						1.23		
		CPRS (obsessive-compulsive): S						1.10		
Total	**776**									
Percentage of ES			**14%**	**21%**	**65%**	**25%**	**6%**	**69%**		
Weighted Mean ES (n)									**0.78 (298)**	**0.94 (541)**
Minimum									**0.11**	**0.46**
Maximum									**1.64**	**2.10**
Standard deviation									**0.45**	**0.69**
Mean duration follow-up										**3.2 yrs**

BSS, Beschwerden-Schwere-Score; CPRS, Comprehensive Psychopathological Self-Rating Scale for Affective Syndromes; ES, effect size; GAS, Goal Attainment Scale; GSI, General Severity Index; IIP, Inventory of Interpersonal Problems; KAPP, Karolinska Psychodynamic Profile; KSP, Karolinska Scales of Personality; SAS, Social Adjustment Scale; SCL-90-R, Symptom Checklist–90–Revised; SOC, Sense of Coherence Scale; yr, year.

[a] Small < 0.5; 0.5 < medium < 0.8; large > 0.8.

[b] P = scale mostly related to personality aspects; S = scale mostly related to symptoms; "−" denotes global outcome.

[c] After 3.5 years in therapy.

TABLE 2.2. Psychotherapy Effect Sizes for Studies Regarding Only Severe Personality Disorders (Severe Pathology)

Study	n	Scale: P/S[b]	Pre/post ES[a]			Pre-follow-up ES[a]			Mean ES	
			Small	Medium	Large	Small	Medium	Large	Post	Follow-up
Clarkin et al. (2001)[42]	23	Self-injurious behavior (no. of incidents): S	0.15						0.33	
		Medical risk of incidents: −	0.37							
		Physical condition after incidents: −	0.46							
Giesen-Bloo et al. (2006)[48c]	42	BPDSI: P			1.85				1.12	
		EuroQol: P		0.64						
		WHOQOL: P			1.16					
		Psychopathology and personality pathology: P			0.84					
Monsen et al. (1995, 1995)[59,60]	25	MMPI (depression): P		0.67			0.67		1.20	1.02 (5.2 yrs)
		MMPI (anxiety): P		0.78			0.78			
		MMPI (ego resilience): P						1.45		
		MMPI (ego strength): P						1.15		
		MMPI (social withdrawal): P			0.92			0.92		
		Affect consciousness: P			2.46			2.30		
		Morey PD scale (MMPI) (passive-aggressive, borderline, antisocial, compulsive, paranoid): P						0.91		
		Morey PD scale (MMPI) (avoidant, dependent, schizoid): P					0.54			
		Morey PD scale (MMPI) (histrionic, narcissistic): P				0.49				
Total	**90**									
Percentage of ES			**27%**	**27%**	**46%**	**11%**	**33%**	**56%**		
Weighted mean ES (n)									**0.94 (90)**	**1.02 (25)**
Minimum									**0.33**	**1.02**
Maximum									**1.20**	**1.02**
Standard deviation									**0.36**	**0.00**
Mean duration follow-up										**5.2 yrs**

BPDSI-IV, Borderline Personality Disorder Structured Interview; ES, effect size; EuroQol, European Quality of Life thermometer; MMPI, Minnesota Multiphasic Personality Inventory; PD, personality disorder; WHOQOL, World Health Organization Quality of Life assessment; yr, year.
[a] Small < 0.5; 0.5 < medium < 0.8; large > 0.8.
[b] P = scale mostly related to personality aspects; S = scale mostly related to symptoms; "−" denotes global outcome.
[c] Three years in treatment.

in moderate/mixed pathology on the basis of therapist and patient opinions about symptom reduction and overall health state. As can be seen in Tables 3.1 and 3.2, for psychotherapy in moderate/mixed pathology, success rates at treatment termination ranged between 30% and 95% (weighted mean = 69%) according to therapists, and between 33% and 93% (weighted mean = 59%) according to patients. A lack of data meant that follow-up success rates could not be calculated.

Turning to the success of psychotherapy for severe pathology, Monsen and colleagues[59,60] reported a reduction of DSM Axis I disorders of 75% at treatment termination and of 83% at follow-up, as rated by independent observers (mean = 79%). In addition, 76% of the patients in the study did not meet the caseness criterion of the SCL-90-R's Global Severity Index at follow-up.

As can be seen in Tables 3.3 and 3.4, for psychoanalysis in moderate/mixed pathology, success rates at treatment

TABLE 2.3. Psychoanalysis Effect Sizes for Studies with Moderate/Mixed Pathology

Study	n	Scale: P/S[b]	Pre/post[a]			Pre-follow-up			Mean ES	
			Small	Medium	Large	Small	Medium	Large	Post	Follow-up
Grande et al. (2006)[49]	32	SCL-90-R (GSI): S			1.54			1.58	1.38	1.44 (1 yr)
		IIP (relationships): P			1.22			1.30		
Rudolf et al. (1994)[64]	44	PSKB-Se (somatic anxiety): S			1.36					0.90
		PSKB-Se (depression): S			0.94					
		PSKB-Se (somatic complaints): S		0.80						
		PSKB-Se (social anxiety): S		0.78						
		PSKB-Se (regressive attachment): P		0.62						
Sandell et al. (2000)[65]	74	SCL-90-R: S						1.55		1.06 (3 yrs)
		SOC: P						1.18		
		SAS: P				0.45				
Weber et al. (1985)[74]	36	Ego strength: P	0.39						0.39	
Total	**186**									
Percentage of ES			**12.5%**	**37.5%**	**50%**	**20%**	**0%**	**80%**		
Weighted mean ES (n)									**0.87 (112)**	**1.18 (106)**
Minimum									**0.39**	**1.06**
Maximum									**1.38**	**1.44**
Standard deviation									**0.41**	**0.17**
Mean duration follow-up										**2.6 yrs**

ES, effect size; GSI (General Severity Index); PSKB-Se, Psychischer und Sozial-Kommunikativer Befund–Selbstbeurteilung; SAS, Social Adjustment Scale; SCL-90-R, Symptom Checklist–90–Revised; SOC, Sense of Coherence Scale; yr, year.
[a] Small < 0.5; 0.5 < medium < 0.8; large > 0.8.
[b] P = scale mostly related to personality aspects; S = scale mostly related to symptoms; "−" denotes global outcome.

TABLE 3.1. Therapist Opinions of Psychotherapy Success[a] Rates (Symptoms/General Assessments)

Study	n	Scale	Post	Follow-up
Holm-Hadulla et al.(1997)[53]	37	CGI (symptoms)	89%	NA
Keller et al. (1998)[54]	111	Global assessment of overall health state	95%	NA
Leichsenring et al. (2005)[55]	36	CGI	84%	NA
Leuzinger-Bohleber et al. (2001)[56]	194	CGI	67%	NA
Von Rad et al. (1998)[61]	33	CGI	30%	55% (3.5 yrs)[b]
Rudolf et al. (1994)[64]	56	PSKB-Se	44%	NA
Weber et al. (1985)[72]	138	CGI	61%	NA
Weber et al. (1985)[74]	29	CGI	66%	NA
Total	**634**			
Mean %			**69%**	
Range			**30%–95%**	

CGI, Clinical Global Impression; NA, not available; PSKB-Se, Psychischer und Sozial-Kommunikativer Befund–Selbstbeurteilung; yr, year.
[a] Success = at least moderate improvement on the scale concerned
[b] Rating by independent observers.

TABLE 3.2. Patient Opinions of Psychotherapy Success[a] Rates (Symptoms/General Assessments)

Study	n	Scale	Post	Follow-up
Brockman et al. (2003)[40]	31	SCL-90-R (GSI)	60%	NA
Brockman et al. (2002)[41]				
Grande et al. (2006)[49]	27	SCL-90-R (GSI)	56%	56% (1 yr)
Holm-Hadulla et al. (1997)[53]	37	CGI (symptoms)	85%	NA
Keller et al. (1998)[54]	111	Global assessment of overall health state	93%	NA
Leichsenring et al. (2005)[55]	36	SCL-90-R (GSI)	77%	NA
Leuzinger-Bohleber et al. (2001)[56]	194	CGI	64%	NA
Rudolf et al. (1994)[64]	56	PSKB-Se	55%	NA
Sandell et al. (2000)[65]	222[b]	No longer meeting caseness criterion[b]	33%	NA
Total	**714**			
Mean			**59%**	
Range			**33%–93%**	

CGI, Clinical Global Impression; GSI, Global Severity Index; NA, not available; PSKB-Se, Psychischer und Sozial-Kommunikativer Befund–Selbstbeurteilung; SAS, Social Adjustment Scale; SCL-90-R, Symptom Checklist–90–Revised; SOC, Sense of Coherence Scale; yr, year.

[a] Success = at least moderate improvement on the scale concerned.

[b] The n here is the number of actual cases in the study, with cases falling in the lowest 10% on the SCL, SOC, and SAS.

termination ranged between 57% and 96% (weighted mean = 70%) according to therapists, and between 66% and 81% (weighted mean = 73%) according to patients. Lack of data precluded the calculation of follow-up success rates.

Personality-related assessments. Tables 4.1 to 4.3 present the success rates of psychoanalysis and psychotherapy for moderate/mixed pathology, based on therapist and patient assessments of personality-related changes.

In psychotherapy for moderate/mixed pathology (Tables 4.1 and 4.2), the success rates stated by therapists at treatment termination (weighted mean = 57%) were comparable to those stated by patients (weighted mean = 59%).

With regard to psychotherapy for severe pathology, Giesen-Bloo and colleagues[48] reported a success rate of 43%

at treatment termination—which was the percentage of patients who achieved the reliable change criterion for borderline personality disorder (BPD) measured by the BPD Index–IV. At follow-up, Stevenson and Meares[68] found a reduction of 30% in patients meeting the criteria for BPD. In the study of Monsen and colleagues,[59,60] 78% of all patients with Axis II diagnoses no longer met the criteria for that diagnosis at treatment termination, and 74% no longer met the criteria at follow-up. The weighted mean of the success rates in these three studies was 51% (according to independent observers).

There were only two studies of psychoanalysis (Table 4.3), each representing patient opinions, and resulting in a weighted mean of 61%. The studies dealt with moderate/mixed pathology. There were no studies dealing with

TABLE 3.3. Therapist Opinions of Psychoanalysis Success[a] Rates (Symptoms/General Assessments)

Study	n	Scale	Post	Follow-up
Erle et al. (2003)[44]	92	CGI	66%	NA
Leuzinger-Bohleber et al. (2001)[56]	207	CGI	60%	NA
Luborsky (2001)[58]	17	SSI	65%	NA
Von Rad et al. (1998)[61]	23	CGI	57%	30% (3.5 yrs)
Rudolf et al. (1994)[64]	44	PSKB-Se	64%	NA
Weber et al. (1985)[72]	235	CGI	67%	NA
Weber et al. (1985)[74]	36	CGI	96%	NA
Total	**610**			
Mean			**70%**	
Range			**57%–96%**	

CGI, Clinical Global Impression; NA, not available; PSKB-Se, Psychischer und Sozial-Kommunikativer Befund–Selbstbeurteilung; SSI, Combination of "success," "satisfaction," and "improvement"; yr, year.

[a] Success = at least moderate improvement on the scale concerned.

TABLE 3.4. Patient Opinions of Psychoanalysis Success[a] Rates (Symptoms/General Assessments)

Study	n	Scale	Post	Follow-up
Grande et al. (2006)[49]	32	SCL-90-R (GSI)	81%	72% (1 year)
Leuzinger-Bohleber et al. (2001)[56]	207	CGI	73%	NA
Rudolf et al. (1994)[64]	44	PSKB-Se	75%	NA
Sandell et al. (2000)[65]	65[b]	No longer meeting caseness criterion[b]	66%	NA
Total	**348**			
Weighted mean			**73%**	
Range			**66%–81%**	

[a] Success = at least moderate improvement on the scale concerned.

[b] The n here is the number of actual cases in the study, with cases falling in the lowest 10% on the SCL-90, SOC, and SAS.

therapist assessments of personality changes in moderate/mixed pathology, and no studies regarding severe pathology.

Overall Success Rates. We calculated overall success rates by pooling all rates. In psychotherapy for moderate/mixed pathology, the success rates were 64% at termination and 55% at follow-up. In psychotherapy for severe pathology (only four studies available),[48,59,60,68,69] the overall success rates were 61% both at termination and at follow-up.

In psychoanalysis for moderate/mixed pathology, the pre/post success rate was 71%, and the pre-follow-up success rate was 54%. In psychoanalysis for severe pathology,

TABLE 4.1. Therapist Opinions of Psychotherapy Success[a] Rates (Personality-Related Assessments)

Study	n	Scale	Post	Mean
Holm-Hadulla et al. (1997)[53]	37	CGI (conflicts)	60%	55%
		CGI (reality)	49%	
Leichsenring et al. (2005)[55]	36	Interpersonal capacities	53%	60%
		Capacity to work	53%	
		Capacity to enjoy	53%	
		Capacity to deal with conflict	79%	
Total	**73**			
Mean				**57%**
Range			**49%–79%**	

CGI, Clinical Global Impression.

[a] Success = at least moderate improvement on the scale concerned.

TABLE 4.2. Patient Opinions of Psychotherapy Success[a] Rates (Personality-Related Assessments)

Study	n	Scale	Post	Mean
Brockman et al. (2003)[40] Brockman (2002)[41]	31	IIP	32%	32%
Grande et al. (2006)[49]	27	IIP	50%	50%
Holm-Hadulla et al. (1997)[53]	37	CGI (conflicts)	84%	75%
		CGI (reality)	65%	
Keller et al. (1998)[54]	111	Emotional condition	94%	72%
		Physical condition	66%	
		Global health	51%	
		Satisfaction job	74%	
		Satisfaction partner	75%	
Rudolf et al. (1994)[64]	56	Somatic problems	60%	40%
		Mental problems	41%	
		Relational problems	20%	
Total	**262**			
Mean				**59%**
Range			**20%–94%**	

CGI, Clinical Global Impression; IIP, Inventory of Interpersonal Problems.

[a] Success = at least moderate improvement on the scale concerned.

only one study was found,[69] with a 59% success rate at treatment termination.

Outcomes of Lower-Quality Studies

Tables 5.1 and 5.2 present the outcomes from studies not meeting the quality criterion.

In psychotherapy, pre/post success rates ranged from 55% to 88%, with a mean of 69%. In psychoanalysis, the range was 41% to 75%, with a mean of 67%. Freedman and colleagues[45] and Hartman and Zepf[51] used a Mean Effectiveness Score. The authors describe this measure as,

TABLE 4.3. Patient Opinions of Psychoanalysis Success[a] Rates (Personality-Related Assessments)

Study	n	Scale	Post	Mean
Grande et al. (2006)[49]	32	Reliable Change Index IIP	72%	72%
Rudolf et al. (1994)[64]	56	Somatic problems	80%	55%
		Mental problems	44%	
		Relational problems	41%	
Total	**88**			
Mean				**61%**
Range			**41%–80%**	

IIP, Inventory of Interpersonal Problems.

[a] Success = at least moderate improvement the scale concerned.

TABLE 5.1. Outcomes of Excluded Studies: Psychotherapy

A. Changes in Scores

	n	Pre	Post	Follow-up	Pre/post	Pre-follow-up
Heinzel et al. (1998)[52]	633					
Physical health[a]		3.20	2.20	2.13	31%	33%
General well-being[a]		4.32	2.25	2.16	48%	50%
Mental health[a]		4.45	2.26	2.18	49%	51%
Social functioning[a]		3.67	2.41	2.14	34%	42%
Dührssen (1986)[43]	60					
Symptoms (GBB)		73[b]	30[c]		41%	
Personality (CGI)		8[b]	4[c]		50%	
Personality (FPI)		12[b]	4[c]		33%	
Friedman et al. (2005)[47]	551					
GAF (0–100)		60.36	76.96		28%	

B. Success Rates

	n	Scale	Pre/Post	Mean
Stehle (2004)[7d]	581	CGI (general)	88%	81%
		CGI (symptoms)	85%	
		CGI (personality)	69%	
Hartman & Zepf (2004)[51e]	448	CGI (success)	55%	55%
Total	**1029**			
Mean success rate				**69%**
Range			**55%–88%**	

CGI, Clinical Global Impression; FPI, Freiburger Persönlichkeitsinventar; GAF, Global Assessment of Functioning; GBB, Giessener Beschwerde Bogen.

[a] Scale: 1 = very good; 5 = very bad.
[b] Number of pretreatment symptoms.
[c] Number of improved symptoms.
[d] Although 61% of patients received psychoanalysis, only 20% had 3–4 sessions a week; 80% had only 1–2 sessions a week. This study was therefore categorized under psychotherapy.
[e] 7% still in treatment.

in essence, a treatment satisfaction scale, with a maximum of 300 points; 150 points indicated no effectiveness, and below 150 indicated negative effectiveness. Freedman and colleagues[45] found a mean score of 209. Hartman and Zepf[51] found a mean score of 237 in psychotherapy and 244 in psychoanalysis.

Summary of Results

We summarize all the findings of this review in Table 6.

DISCUSSION

We conducted a systematic search of the literature between 1970 and 2007 for outcome-intervention studies of LPT effectiveness in ambulatory, adult patients. We found 27 relevant studies. We classified 8 of them (30%) as "lower quality" studies using a quality criterion based on a self-developed research quality control list. We compiled the results of the 19 "higher quality" studies by pooling their data. First, we calculated overall ESs and success rates. Second, we broke down the results according to symptom reduction/personality change, and therapist/patient opinion.

Psychotherapy

In moderate/mixed pathology, the overall success rates (64% at termination; 55% at follow-up), the percentage of large ESs (65% at termination; 69% at follow-up), and the weighted mean ESs (0.78 at termination; 0.94 at follow-up) indicate substantial effectiveness, which was maintained for years after treatment termination. In severe pathology, taking into consideration that the results mainly related to personality pathology, the percentage of large ESs (46% at termination; 56% at follow-up) and the weighted mean ESs (0.94 at termination; 1.02 at follow-up) were certainly not inferior to those in moderate/mixed pathology. However, the

TABLE 5.2. Outcomes of Excluded Studies: Psychoanalysis

A. Changes in Scores

	n	Pre	Post	Follow-up	Pre/post	Pre-follow-up
Dührssen (1986)[43]	60					
Symptoms (GB)		73[a]	26[b]		36%	
Personality (CGI)		8[a]	4[b]		50%	
Personality (FPI)		12[a]	3[b]		25%	

B. Success Rates

	n	Scale	Pre/post	Mean
Erle & Goldberg (2003)[44]	161	CGI	74%	74%
Sashin et al. (1975)[66]	130	CGI (symptoms [discomfort])	61%	54%
		CGI (sexual adjustment)	51%	
		CGI (restriction of life functioning)	62%	
		CGI (interpersonal relationships)	41%	
		CGI (insight)	46%	
		CGI (work productivity)	43%	
		CGI (global)	75%	
Hartman & Zepf (2004)[51c]	263	CGI	69%	69%
Total	**554**			
Mean success rate				**67%**
Range			**41% –75%**	

CGI, Clinical Global Impression; FPI, Freiburger Persönlichkeitsinventar; GB, Giessener Beschwerde Bogen.

[a] Number of pretreatment symptoms.
[b] Number of improved symptoms.
[c] 67% still in treatment.

results in severe pathology should be interpreted cautiously, as they are based on only three studies, with follow-up data in only one study and with one of the studies presenting as many as nine ESs. These ESs are likely to be highly correlated since they present dimensions from only two separate instruments.

When differentiating between symptom reduction and personality changes, it appears that the effects for moderate/mixed pathology were more pronounced for the former (ES = 1.03) than for the latter (ES = 0.54). Patients and therapists were in general agreement about success rates. They both thought the effect on personality change was similar to somewhat lower (59% and 57%, respectively) than that on symptom reduction (59% and 69%, respectively).

As only a few studies presented data on severe pathology, our findings were mainly suggestive. We found a large weighted ES for personality change (1.11) and no ES information about symptom reduction. As with the results for moderate/mixed pathology, success rates for personality change seemed lower than for symptom reduction.

The recent, sophisticated meta-analysis of Leichsenring & Rabung[22] differs on several points from our review. First, the cutoff date for studies considered was a year later than ours (falling in 2008 rather than 2007), enabling it to consider more recently published studies. Second, the focus was explicitly on complex mental disorders—including, for instance, eating disorders—whereas we left out these more complex disorders. Third, the meta-analysis included both ambulant patients and those receiving care in any form of hospital setting, whereas we focused exclusive on the former. Fourth, the authors' inclusion criteria were one year *or* 50 sessions (versus our one year *and* 50 sessions) and the availability of information on ESs (versus our consideration also of studies providing only response rates). As a result of the different inclusion criteria, Leichsenring and Rabung[22] considered 13 studies not included in our review. However, it is noteworthy that the ESs for psychotherapy found by Leichsenring and Rabung (0.78 to 1.98) were similar to the ESs that we found (0.78 to 0.94). In addition, they also seem to have found lower ESs for personality functioning (0.78) than for symptoms (target problems, 1.54; psychiatric symptoms, 0.91). Finally, Leichsenring and Rabung confirmed our cautious findings that results of psychotherapy were stable at follow-up.

TABLE 6. Summary of Results

A. Overall Effect Sizes and Success Rates

	Psychotherapy			Psychoanalysis		
	Mean effect size	% large effect sizes	Overall success rate	Mean effect size	% large effect sizes	Overall success rate
Mixed pathology	(n = 9)		(n = 11)	(n = 4)		(n = 8)
Termination	0.78 (n = 6)	65%	64% (n = 11)	0.87 (n = 3)	50%	71% (n = 8)
Follow-up	0.94 (n = 5)	69%	55% (n = 2)	1.18 (n = 2)	80%	54% (n = 2)
Severe pathology (only PDs)	(n = 3)		(n = 4)			(n = 1)
Termination	0.94 (n = 3)	46%	61% (n = 3)	NA	NA	59% (n = 1)
Follow-up	1.02 (n = 1)	56%	61% (n = 2)	NA	NA	NA

B. Effect Sizes and Success Rates for Symptom Reduction and Personality Change

	Psychotherapy			Psychoanalysis		
	Mean effect size	Success rate according to patient	Success rate according to therapist	Mean effect size	Successrate according to patient	Success rate according to therapist
Mixed pathology	(n = 9)	(n = 8)	(n = 8)	(n = 4)	(n = 4)	(n = 7)
Symptoms	1.03 (n = 6)	59% (n = 8)	69% (n = 8)	1.38 (n = 3)	73% (n = 4)	70% (n = 7)
Personality	0.54 (n = 7)	59% (n = 5)	57% (n = 2)	0.76 (n = 4)	61% (n = 2)	NA
Severe pathology (only PDs)	(n = 3)	(n = 1)	(n = 3)[a]			
Symptoms	NA	76% (n = 1)	79% (n = 1)[a]	NA	NA	NA
Personality	1.11 (n = 2)	NA	51% (n = 3)[a]	NA	NA	NA

NA, not available; PD, personality disorder.
[a] Ratings made by independent observers.

Psychoanalysis

In moderate/mixed pathology, the overall success rates (71% at termination; 54% at follow-up), the percentage of large ESs (50% at termination; 80% at follow-up), and the weighted mean ESs (0.87 at termination; 1.18 at follow-up) indicated substantial effectiveness, which was maintained for years after treatment termination. We did not find ESs for psychoanalysis in severe pathology, and only one study provided information about the success of psychoanalysis for severe pathology: 59%.[69]

After differentiating between symptom reduction and personality changes, we found that in moderate/mixed pathology, the effects were more prominent for symptom reduction (ES = 1.38) than for personality changes (ES = 0.76) (though these results, based on only four studies, must by interpreted cautiously). Patients and therapists concurred in their opinions about success rates, and the assessment of changes in personality (61% and 51%, respectively) seemed to be lower than the assessment of symptom reduction (73% and 70%, respectively) (though note here, too, that the effects on personality change are based on only three studies).

Comparison with Outcomes of "Lower Quality" Studies

Thirty percent of the retrieved studies did not meet our quality criterion. As only a few of them provided information that could be compared to the results of the "higher quality" studies, it is not possible to draw strong conclusions. The data suggest that mean success rates do not greatly differ from those in higher-quality studies. The mean psychotherapy success rate was 69% in the lower-quality studies and 67% in the higher-quality studies. For psychoanalysis, the success rates were 67% and 67%, respectively. Our findings concur with those of Lipsey and Wilson,[77] who concluded on the basis of 27 meta-analyses that the ESs of studies of varying methodological quality did not differ significantly.

Additional Comments

ESs for symptom reduction seem higher than for personality change. This result accords with the findings in the literature and the findings of clinicians that personality change is harder to achieve than symptom reduction. Rudolf and

colleagues[50,64] found that long-term psychoanalytic therapy performed better than moderate-length psychoanalytic therapy (mean = 60 sessions) in terms of structural changes of personality. The results of a study by Kopta and colleagues[78] showed that improvements in personality structure took longer than symptom reduction. Perry and colleagues[79] estimated the length of treatment needed if patients were no longer to meet the criteria for PDs. They found that 50% of the patients would recover after 1.3 years (or 92 sessions), and 75% by 2.2 years (or 216 sessions). The results of our review indicate similar success rates in personality changes. In our opinion, even a moderate effect on personality characteristics may be more important clinically (related to quality of life, social functioning, and vulnerability to relapse in the long run) than a large effect on symptoms.

In addition, our cutoff point for considering therapy as long term must be taken into account—at least 50 sessions over a year or more—to be sure, far less than average for long-term psychoanalytic therapy. Consequently, in many cases within the studies reviewed here, therapy may have satisfied our criterion as being long term but without having continued long enough to result in (much) personality change.

Another issue is the durability of the results. The general literature contains indications that the effects of short-term therapies are themselves more short lived.[80–82] In contrast, our results suggest that the effects of LPT persist for years after treatment termination.

The comparability of psychotherapy and psychoanalysis remains unresolved; in our opinion, our findings do not allow for valid comparisons of the relative effectiveness of psychotherapy and psychoanalysis. Little information was provided that would allow a comparison of the psychotherapy patient group and the psychoanalysis patient group in terms of all relevant pretreatment characteristics. Since the groups were not randomized, it is likely that the personality traits of patients receiving psychotherapy and those receiving psychoanalysis differed in important respects.

Limitations

Our study presents several limitations. First, our review is almost entirely based on cohort studies. In the hierarchy of empirical evidence, RCTs are the gold standard, although treatment RCTs are fraught with their own problems. It has been extensively argued[21,76,83–85] that it is difficult to conduct RCTs of long-term therapies (be they psychological or somatic). Treatments for borderline personality disorders present somewhat fewer difficulties in this respect. Leichsenring[20] has cogently argued that when studying long-term treatments, it is not possible to maintain a waiting list of untreated patients, to apply manualized treatments, or to create equal treatment conditions. He pointed

out that the personality characteristics of patients who opt for long-term psychoanalytic therapy differ from those of patients preferring other therapies. Randomization, even if nominally feasible, violates the therapy-patient match.

Some argue that cohort studies tend to overestimate treatment effect,[86,87] but other studies contradict that view. Concato and colleagues[88] compared the results of RCTs to those of cohort studies dealing with the same interventions for five different somatic disorders. Analyses of 72 meta-analyses of RCTs, 24 meta-analyses of cohort or case-control studies, and 6 meta-analyses involving both designs showed that the results of the observational studies were remarkably close to those of RCTs. The authors concluded: "The popular belief that only randomised, controlled trials produce trustworthy results and that all observational studies are misleading does a disservice to patient care, clinical investigation, and the education of health care professionals." Shadish and colleagues[89] did not find a significant correlation between the degree of clinical representativeness (i.e., RCTs vs. naturalistic studies) and the size of effects reported in studies of psychotherapy. Benson & Hartz[90] compared the results of observational studies with those of RCTs reported between 1985 and 1998 dealing with two or more treatments or interventions for the same condition. The authors concluded: "We found little evidence that estimates of treatment effects in observational studies reported after 1984 are either consistently larger than or qualitatively different from those obtained in randomized, controlled trials."

A second limitation is that our review pools data from studies that were conducted in different countries and settings, at different times, and for different patients, and that used nonmanualized therapies (i.e., ones that, despite being psychoanalytic, undoubtedly have different features). Our results must therefore be interpreted with caution. For the same reason, comparing pre/post findings with pre-follow-up findings is hazardous since they typically refer to different groups of patients. We have tried to enhance the comparability of the studies included by restricting our search to individual therapy in adult, ambulant patients, by distinguishing between moderate/mixed and severe pathology, and by applying a quality criterion. Nevertheless, the studies included are still obviously characterized by clinical heterogeneity (especially if measured by DSM diagnostic standards). However, according to psychoanalytic diagnostic criteria, the patient group shows more actual homogeneity. After all, all patients were indicated, presumably on solid grounds, for psychoanalytic therapy. The recent *Psychodynamic Diagnostic Manual*[91] proposes conceptual and research foundations for a psychodynamically based classification system for mental health disorders. The manual includes, in connection with DSM categories, dimensions of mental functioning such as capacity for emotion regulation, capacity for relationships and intimacy, quality of internal

experience (levels of self-confidence and self-regard), defensive patterns, self-observing capacities, and capacity for differentiation and integration. These dimensions represent (some of) the diagnostic criteria that psychoanalysts use when establishing whether psychoanalytic therapy is indicated.

Third, most of the studies included in our review treated patients with different diagnoses and with high levels of comorbidity. The studies did not always break down their results according to diagnostic categories. Some studies provided limited demographic and clinical information about the patients. Both of these factors limit the conclusions we can draw about the effectiveness of LPT for specific DSM-specified disorders. Nevertheless, the diversity of patients serves to enhance the representativeness of the patient population to that of general LPT clinical practice—which is precisely the strength of naturalistic effectiveness studies. A related consideration regarding the diagnostic aspects of this review is that our description of studies as dealing with "moderate/mixed" or "severe" pathology could be misleading. Some patients with moderate/mixed pathology suffer from severe disorders, and many of these patients have various comorbid disorders—which means that complex clinical pictures are the standard, not the exception. Many have also tried briefer therapies, apparently to no avail.[38,39]

Fourth, the quality of the studies that met our inclusion criteria varied. Unfortunately, the quality of any review—including this one—depends heavily on the quality of the studies included. In an effort to safeguard quality, we devised a 24-item Research Quality Score inspired by the criteria put forward by Leichsenring[21] and by the Cochrane Collaboration.[29] We defined an admittedly arbitrary cutoff point, and the list has not been validated. The rigorousness of our quality check could therefore, in this respect, be amenable to improvement. Thirty percent of the studies we found in the literature did not meet our quality criterion—which suggests that the RQS is doing some serious work—but the quality of the studies that met our criterion nevertheless varied widely. Flaws included: the lack of pretreatment data (retrospective design); an overgeneral description of the patient population (e.g., stating only that the sample included patients treated in a given period); no mention of diagnoses; no detailed information about the therapies (e.g., saying only that the study dealt with psychoanalysis or psychotherapy, but not stating the mean number of sessions or duration); overgeneral outcome criteria (e.g., using only a five-step Likert scale for "therapy success"); no independent, let alone blind, assessors; no information about dropouts (e.g., data only for completers); and no analyses of confounders. Only two studies tried to provide some information about control groups.[20,53]

Fifth, many studies used simple measures of success rates—often comprising no more than global assessments

of success based on patient and therapist opinions. Some studies did show that global assessment scales provided more differentiation to reflect both the severity of the illness and treatment results,[92] but such scales do not provide the same information as detailed questionnaires; they are simple instruments that often demand considerable clinical competence to generate reliable information. Also reflecting the nature of the studies themselves, success in our review was rated predominantly by therapists and patients, and only seldom by independent observers. There are pros and cons, of course, for the rating provided by each group: patients, clinicians, and independent observers. On the one hand, independent observers are less invested in the treatment and might therefore be less biased in their judgment of results. On the other hand, patients and therapists have much more exposure than independent observers to the relevant field of observation. In addition, since the patient's subjective feelings are integrally connected with the goals of therapy, one may well wonder who should be the best judge on any relevant improvement or success. The literature is not in agreement, however, on the question whether patients and therapists might overestimate therapy success. Harty and colleagues[71] performed an analysis of the findings of the Menninger Foundation's psychotherapy research project and found that both therapists (65%) and patients (54%) rated therapy success higher than independent judges (38%). Other studies have found, though, that self-reports present more modest results than observer ratings.[31,93]

Sixth, the statistical analyses we performed were rather simple. Although ESs are a reasonably robust evaluation of treatment effects, we were limited to a simple calculation of the average of the ESs reported in the individual studies. For lack of the original data, we could not calculate mean ESs based on the original means and standard deviations. Furthermore, some of the pooled results are based on only a few studies, and in some cases the results of only one study are mentioned.

Finally, our decisions to categorize measures as involving symptom reduction versus personality changes may be debated. In particular, success rates for symptom reduction were often based on CGI scales that do not necessarily differentiate between symptom improvement and changes in personality. We classified them as symptom reduction measures.

CONCLUSIONS

LPT effectiveness has been extensively investigated, mainly by cohort studies of varying quality. Our review of the available empirical evidence suggests that LPT is an effective treatment modality, with moderate to large effects on both symptom reduction and personality changes. The effects are

stronger in symptom reduction than in personality change, but moderate changes in personality may be more significant clinically in terms of quality of life and relapse prevention. Patients and therapists do not seem to differ in their opinions about therapy success. The results of LPT seem to be maintained in the years after treatment termination.

Further research with cohort studies will bear more fruit if they incorporate stricter research standards.

Declaration of interest: The authors report no conflicts of interest. The authors alone are responsible for the content and writing of the article.

APPENDIX
Research Quality Score (RQS)

Criteria	Single or multiple cohort study: Points	Cohort study with matched control group:[a] Points	Randomized, controlled trial: Points
Study design			
Randomization	NA	NA	35
Prospective design	10	10	NA
(Matched) control group	NA	5, 10	NA
Patient group			
Clear inclusion/exclusion criteria	0, 1, 2	0, 1, 2	0, 1, 2
Characteristics described	0, 1, 2	0, 1, 2	0, 1, 2
Baseline scores comparable	NA	0, 2, 4	0, 2, 4
Adequate sample size	0, 1, 2	0, 1, 2	0, 1, 2
Intervention			
Clear description of intervention	0, 2, 4	0, 2, 4	0, 2, 4
Therapist experience	0, 1, 2	0, 1, 2	0, 1, 2
Adherence therapist checked	0, 1, 2	0, 1, 2	0, 1, 2
Equal treatment groups	NA	0, 2	0, 2
Outcome data			
Outcome criteria clear and relevant	0, 2, 4	0, 2, 4	0, 2, 4
Quality assessment instruments	0, 2, 4	0, 2, 4	0, 2, 4
Blind or independent assessment	0, 2, 4	0, 2, 4	0, 2, 4
Multiple assessors	0, 2, 4	0, 2, 4	0, 2, 4
Statistics			
Adequate statistical methods	0, 1, 2	0, 1, 2	0, 1, 2
ITT or PP analyses	0, 2, 4	0, 2, 4	0, 2, 4
Confounders analyses	0, 1, 2	0, 1, 2	NA
Dropout			
Dropout defined and acceptable	0, 2, 4	0, 2, 4	0, 2, 4
Dropout comparable	NA	0, 1, 2	0, 1, 2
Maximum RQS	52	70	83
Cutoff score review	26	35	41
Follow-up			
Adequate length	0, 2, 4	0, 2, 4	0, 2, 4
Multiple assessors	0, 1, 2	0, 1, 2	0, 1, 2
Loss to follow-up defined and acceptable	0, 2, 4	NA	NA
Loss to follow-up defined, acceptable, and comparable	NA	0, 2, 4	0, 2, 4
(Maximum score for follow-up)	10	10	10
Maximum RQS	62	80	93
Cutoff score review	31	40	47

ITT, intent to treat; NA, not applicable; PP, per protocol.

[a] A matched control group is either an untreated, matched group (with similar features to those of the intervention group) or a group treated with evidence-based treatment.

REFERENCES

1. Coriat IH. Some statistical results of the psychoanalytic treatment of the psychoneuroses. Psychoanal Rev 1917;4:209–16.

2. Fenichel O. Statistischer Bericht über die therapeutische Tätigkeit 1920–1930. In: Radó S, Fenichel O, Müller-Braunschweig C, eds. Zehn Jahre Berliner Psychoanalytisches Institut Poliklinik und Lehranschalt. Wien: Internationaler Psychoanalytischer Verlag, 1930:S13–9.

3. Kessel L, Hyman HT. The value of psychonalysis as a therapeutic procedure. JAMA 1936;107:326.

4. Jones E. Report of the clinic work: 1926–1936. London: Clinic of Psychoanalysis, 1936.

5. Alexander F. Five year report of the Chicago Institute for Psychoanalysis, 1932–1937. Chicago: Institute for Psychoanalysis, 1937.

6. Knight RP. Evaluation of the results of psychoanalytic therapy. Am J Psychiatry 1941;98:434–46.

7. Schjelderup H. Lasting effects of psychoanalytic treatment. Psychiatry 1955;18:109–33.

8. Orgel S. Effects of psychoanalysis on the course of peptic ulcer. Psychosom Med 1958;20:117–25.

9. Knapp PH, Levin S, McCarter R, Werner H, Zetzel E. Suitability for psychoanalysis: a review of 100 supervised analytic cases. Psychoanal Q 1960;29:459–77.

10. Graham SR. The effects of psychoanalytically oriented psychotherapy on levels of frequency and satisfaction in sexual activity. J Clin Psychol 1960;16:94–5.

11. Klein H. A study of changes occurring in patients during and after psychoanalytic treatment. In: Hoch PH, Zubin J, eds. Current approaches to psychoanalysis. New York: Grune & Stratton, 1960:151–75.

12. Cremerius J. Die Beurteilung des Behandlungserfolges in der Psychotherapie. Berlin, Göttingen, Heidelberg: Springer, 1962.

13. Bieber I. Homosexuality: a psychoanalytic study. New York: Basic, 1962.

14. Hamburg DA, Bibring GL, Fisher C, et al. Report of ad hoc committee on central fact gathering data of the American Psychoanalytic Association. J Am Psychoanal Assoc 1967;15:841–61.

15. Feldman H. Results of psychoanalysis in clinical case assignments. J Am Psychoanal Assoc 1968;17:274–300.

16. Bachrach H, Galatzer-Levy R, Skolnikoff A, Waldron S. On the efficacy of psychoanalysis. J Am Psychoanal Assoc 1991;39:871–916.

17. Doidge N. Empirical evidence for the efficacy of psychoanalytic psychotherapies and psychoanalysis: an overview. Psychoanal Inq 1997:17(suppl):102–50.

18. Fonagy P. An open door review of outcome studies in psychoanalysis. 1999. http://www.ipa.org.uk/R-outcome.htm

19. Doidge N. Is psychoanalysis effective? Econ Neurosci 2001;3:41–7.

20. Leichsenring F. Are psychodynamic and psychoanalytic therapies effective? A review of empirical data. Int J Psychoanal 2005;86:841–68.

21. Leichsenring F. Randomized controlled versus naturalistic studies: a new research agenda. Bull Menninger Clin 2004;68:137–51.

22. Leichsenring, F, Rabung, S. Effectiveness of long-term psychodynamic psychotherapy: a meta-analysis. JAMA 2008;300:1551–65.

23. Freud S. Wege der psychoanalytischen Therapie. Ergänzungsband 1919:243.

24. Bateman A, Fonagy P. The effectiveness of partial hospitalisation in the treatment of borderline personality disorder: a randomized controlled trial. Am J Psychiatry 1919;156:1563–9.

25. Stanton AH, Gunderson JG, Knapp PH, et al. Effects of psychotherapy in schizophrenia: I. Design and implementation of a controlled study. Schizophr Bull 1984;10:520–64.

26. Fonagy P, Moran GS. Studies on the efficacy of child psychoanalysis. J Consult Clin Psychol 1990;58:684–95.

27. Quiroga S. Estudio comparativo sobre eficacia terapéutica en trastornos de la alimentación I. Factores diferenciales y determinantes comunes [Comparative study about therapeutic efficacy in eating disorders I. Differential factors and common determinants]. Paper presented at the IV Encuentro Capítulo Sudamericano de la Society for Psychotherapy Research, Montevideo, September 1998.

28. Heinicke CM. Frequency of psychotherapeutic sessions as a factor affecting the child's developmental status. Psychoanal Study Child 1965;20:42–98.

29. Cochrane Collaboration. Cochrane reviewers' handbook. At http://www.cochrane.org

30. Cohen J. Satistical power analysis for behavioural sciences. Hillsdale, NJ: Erlbaum, 1988.

31. 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;160:1223–32.

32. Rosenthal R. Meta-analytic procedures for social research: applied social research methods. Newbury Park, CA: Sage, 1991.

33. Derogatis LR, Lazarus L. SCL-90-R, Brief Symptom Inventory, and matching clinical rating scales. In: Maruish ME, ed. The use of psychological testing for treatment planning and outcome assessment. Hillsdale, NJ: Erlbaum, 1974:217–48.

34. Antonovsky A. Unraveling the mystery of health. San Francisco, CA: Jossey-Bass, 1987.

35. Weissman M, Bothwell S. Assessment of social adjustment by patient self report. Arch Gen Psychiatry 1976;33:1111–5.

36. Guy W. ECDEU assessment manual for psychopharmacology. Rockville, MD: National Institutes of Health, Psychopharmacology Research Branch, 1976.

37. Horowitz LM, Strauss B, Kordy H. Inventory of interpersonal problems. 2nd ed. Göttingen: Beltz, 2000.

38. Doidge N, Simon B, Brauer L, et al. Psychoanalytic patients in the US, Canada, and Australia: I. DSM-III-R disorders, indications, previous treatment, medications and length of treatment. J Am Psychoanal Assoc 2002;50:575–614.

39. Doidge N, Simon B, Lancee WJ, et al. Psychoanalytic patients in the US, Canada, and Australia: II. A DSM-III-R validation study. J Am Psychoanal Assoc 2002;50:615–27.

40. Brockman J, Schlüter T, Eckert J. Therpieziele, Zieländerungen und Zielerreichung im Verlauf psychoanalytisch orientierter und verhaltenstherapeutischer Langzeittherapie. Psychother Psychosom Med Psychol 2003;53:163–70.

41. Brockman J, Schlüter T, Brodbeck D, Eckert J. Die Effekte psychoanalytisch orientierter und verhaltenstherapeutischer Langzeittherapien. Ein vergleichende Studie aus der Praxis niedergelassener Psychotherapeuten. Psychotherapeut 2002;6:347–55.

42. Clarkin JF, Foelsch PA, Levy KN, Hull JW, Delaney JC, Kernberg OF. The development of a psychodynamic treatment for patients with borderline personality disorder: a preliminary study of behavioural change. J Personal Disord 2001;15:487–95.

43. Dührssen A. Dynamische Psychotherapie, Psychoanalyse und analytische Gruppenpsychotherapie im Vergleich. Z Psychosom Med 1986;32:161–80.

44. Erle JB, Goldberg DA. The course of 253 analyses from selection to outcome. J Am Psychoanal Assoc 2003;51:257–93.

45. Freedman N, Hoffenberg JD, Vorus N, Frosch A. The effectiveness of psychoanalytic psychotherapy: the role of treatment duration, frequency of sessions, and the therapeutic relationship. J Am Psychoanal Assoc 1999;47:741–72.

46. Friedman RC, Bucci W, Christian C, Drucker P, Garrison WB 3rd. Private psychotherapy patients of psychiatrist psychoanalysts. Am J Psychiatry 1998;155:1772–4.

47. Friedman RC, Garrison WB 3rd, Bucci W, Gorman BS. Private psychotherapy patients of senior psychoanalysts: an effectiveness study. J Am Acad Psychoanal Dyn Psychiatry 2005;33:583–610.

48. Giessen-Bloo J, Dyck R van, Spinhoven P, et al. Outpatient psychotherapy for borderline personality disorder. Arch Gen Psychiatry 2006;63:649–58.

49. 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;16:470–85.

50. Rudolf G, Dilg R, Grande T, et al. Effektivität und Effizienz psychoanalytischer Langzeittherapie: die Praxisstudie Analytische Langzeitpsychotherapie [Effectiveness and efficiency of long-term psychoanalytic psychotherapy: the practice study of long-term psychoanalytic psychotherapy]. In: Gerlach A, Springer A, Schlösser A, eds. Psychoanalyse des Glaubens [Psychoanalysis of religious belief]. Giessen: Psychosozial Verlag, 2004.

51. Hartmann S, Zepf S. Einflüsse auf die Symptombesserung in der Psychotherapie bei Patienten mit unterschiedlichen Beschwerdebildern. Psychother Psychosom Med Psychol 2004;54:445–56.

52. Heinzel R, Breyer F, Klein T. Ambulante analytische Einzel- und Gruppenpsychotherapie in ein der bundesweiten katamnestischen Evaluationsstudie. Gruppenpsychotherapie und Gruppendynamik 1998;34:135–52.

53. Holm-Hadulla R, Kiefer L, Sessar W. Zur Effektivität tiefenpsychologisch fundierter Kurz- und Psychotherapien. Psychother Psychosom Med Psychol 1997;47:271–8.

54. Keller W, Westhoff G, Dilg R, Rohner R, Studt HH, and the study group on empirical psychotherapy research in analytical psychology. Efficacy and cost effectiveness aspects of outpatient Jungian psychoanalysis and psychotherapy—a catamnestic study. Berlin: Department of Psychosomatics and Psychotherapy, University Medical Center Benjamin Franklin, Free University of Berlin, 1997. At http://www.uni-saarland.de

55. Leichsenring F, Biskup J, Kreische R, Staats H. The Göttingen study of psychoanalytic therapy: first results. Int J Psychoanal 2005;86:433–55.

56. Leuzinger-Bohleber M, Target M, eds. Outcomes of psychoanalytic treatment: perspectives for therapists and researchers. London: Whurr, 2001.

57. Leuzinger-Bohleber M, Stuhr U, Rüger B, Beutel M. How to study the 'quality of psychoanalytic treatments' and their long-term effects on patients well-being: a representative, multi-perspective follow-up study. Int J Psychoanal 2003;84:263–90.

58. 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;49:217–34.

59. Monsen JT, Odland T, Faugli A, Daae E, Eilertsen DE. Personality disorders changes and stability after intensive psychotherapy focusing on affect consciousness. Psychother Res 1995;5:33–48.

60. Monsen JT, 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;36:256–68.

61. Rad M von, Senf W, Bräutigam W. Psychotherapie und Psychoanalyse in der Krankenversorgung: Ergebnisse des Heidelberger Katamnese-Projektes. Psychother Psychosom Med Psychol 1998;48:88–100.

62. Heuft G, Seibuchler-Engec H, Taschke M, Senf W. Langzeitoutcome ambulanter psychoanalytischer Psychotherapien und Psychoanalysen. Ein textinhaltsanalytische Untersuchung von 53 Katamneseinterviews. Forum der Psychoanalyse 1996;12:342–55.

63. Kordy H, Rad M von, Senf W. Time and its relevance for successful psychotherapy. Psychother Psychosom 1988;49:212–22.

64. Rudolf G, Manz R, Öri C. Ergebnisse psychoanalytischer Therapien. Z Psychosom Med 1994;40:25–40.

65. Sandell R, Blomberg J, Lazar A, Carlsson J, Broberg J, Schubert J. Varieties of long-term outcome in psychoanalysis and long-term psychotherapy. A review of findings in the Stockholm outcome of psychoanalysis and psychotherapy project (STOPP). Int J Psychoanal 2000;81:921–42.

66. Sashin JI, Eldred S, Amerongen ST van. A search for predictive factors in institute supervised cases: a retrospective study of 183 cases from 1959–1966 at the Boston Psychoanalytical Society and Institute. Int J Psychoanal 1975;56:343–59.

67. Stehle S. Psychotherapeutische Berfustätigkeit. Ergebnisse der DGPT-Therapeutenerhebung. In: Gerlach A, Schlösser A, Springer A, eds. Psychoanalyse des Glaubens. Gießen: Psychosozial-Verlag, 2004.

68. Stevenson J, Meares R. An outcome study of psychotherapy for patients with borderline personality disorder. Am J Psychiatry 1992;149:358–62.

69. Wallerstein R. Forty-two lives in treatment: a study of psychoanalysis and psychotherapy. New York: Guilford, 1986.

70. Kernberg O, Burstein ED, Coyne L, Appelbaum A, Horwitz L, Voth H. Psychotherapy and psychoanalysis—the final report of the Menninger Foundation's Psychotherapy Research Project. Bull Menninger Clin 1972;36:1–275.

71. Harty M, Horwitz L. Therapeutic outcome as rated by patients, therapists, and judges. Arch Gen Psychiatry 1976;33:957–61.

72. Weber JJ, Bachrach HM, Solomon M. Factors associated with outcome of psychoanalysis: a report of the Colombia Psychoanalytic Center Research Project (II). Int Rev Psychoanal 1985;12:127–41.

73. Bachrach H, Weber J, Solomon M. Factors associated with the outcome of psychoanalysis (clinical and methodological considerations): report of the Colombian Psychoanalytic Center Research Project (IV). Int Rev Psychoanal 1985;12:379–89.

74. Weber JJ, Bachrach HM, Solomon M. Factors associated with outcome of psychoanalysis: a report of the Colombia Psychoanalytic Center Research Project (III). Int Rev Psychoanal 1985;12:251–62.

75. Wilczek A, Berber, JP Gustavson JP, Asberg M, Weinryb RM. Change after long-term psychoanalytic psychotherapy. J Am Psychoanal Assoc 2004;52:1163–84.

76. Seligman MEP. The effectiveness of psychotherapy. The Consumer Reports Study. Am Psychol 1995;50:965–74.

77. Lipsey MW, Wilson DB. The efficacy of psychological, educational, and behavioral treatment: confirmation from meta-analyses. Am Psychol 1993;48:1181–209.

78. Kopta SM, Howard KI, Lowry JL, Beutler LE. Patterns of symptomatic recovery in psychotherapy. J Consult Clin Psychol 1994;62:1009–16.

79. Perry JC, Banon E, Floriana I. Effectiveness of psychotherapy for personality disorders. Am J Psychiatry 1999;156:1312–21.

80. Gloaguen V, Cottraux J, Cucherat M, Blackburn I. A meta-analysis of the effects of cognitive therapy in depressed patients. J Affect Disord 1998;49:59–72.

81. Hollon SD, DeRubeis RJ, Shelton RC, et al. Prevention of Relapse following cognitive therapy vs medications in moderate to severe depression. Arch Gen Psychiatry 2005;62:417–22.

82. Maat S, Dekker J, Schoevers R, Jonghe F de. Relative efficacy of psychotherapy and pharmacotherapy in the treatment of depression: a meta-analysis. Psychother Res 2006;16:562–72.

83. Wallerstein R. Comment on Gunderson and Gabbert. J Am Psychoanal Assoc 1999;47:728–34.

84. Westen D, Novotny CM, Thomson-Brenner H. The empirical status of empirically supported psychotherapies: assumptions, findings, and reporting in controlled clinical trials. Psychol Bull 2004;130:631–63.

85. Maat S, Dekker J, Schoevers R, Jonghe F de. The effectiveness of long-term psychotherapy: methodological issues. Psychother Res 2007;17:59–65.

86. Sacks H, Chalmers TC, Smith H Jr. Randomized versus historical controls for clinical trials. Am J Med 1982;72:33–40.

87. Kunz R, Oxman AD. The unpredictability paradox: review of empirical comparisons of randomised and non-randomised clinical trials. BMJ 1998;317:1185–90.

88. Concato J, Shah N, Horwitz RI. Randomised controlled trials, observational studies and the hierarchy of research designs. N Engl J Med 2000;342:1887–92.

89. Shadish WR, Matt G, Navarro A, Phillips G. The effects of psychological therapies under clinically representative conditions: a meta-analysis. J Consult Clin Psychol 2000;26:512–29.

90. Benson K, Hartz AJ. A comparison of observational studies and randomized, controlled trials. N Engl J Med 2000;342:1878–86.

91. PDM Task Force. Psychodynamic diagnostic manual. Silver Spring, MD: Alliance of Psychoanalytic Organizations, 2006.

92. Ottoson JO. Experimental studies of the mode of action electroconvulsive treatment. Acta Psychiatr Scand 1960;35(suppl 145):69–97.

93. Lambert MJ, Hatch DR, Kingston MD, Edwards BC. Zung, Beck and Hamilton Rating Scales as measures of treatment outcome: a meta-analytic comparison. J Consult Clin Psychol 1986;54:54–9.