

Review article

Long-term psychodynamic psychotherapy in complex mental disorders: update of a meta-analysis

Falk Leichsenring and Sven Rabung

Background

Dose–effect relationship data suggest that short-term psychotherapy is insufficient for many patients with chronic distress or personality disorders (complex mental disorders).

Aims

To examine the comparative efficacy of long-term psychodynamic psychotherapy (LTPP) in complex mental disorders.

Method

We conducted a meta-analysis of controlled trials of LTPP fulfilling the following inclusion criteria: therapy lasting for at least a year or 50 sessions; active comparison conditions; prospective design; reliable and valid outcome measures;

treatments terminated. Ten studies with 971 patients were included.

Results

Between-group effect sizes in favour of LTPP compared with less intensive (lower dose) forms of psychotherapy ranged between 0.44 and 0.68.

Conclusions

Results suggest that LTPP is superior to less intensive forms of psychotherapy in complex mental disorders. Further research on long-term psychotherapy is needed, not only for psychodynamic psychotherapy, but also for other therapies.

Declaration of interest

None.

From both a clinical and a health-economic perspective it is important to distinguish between patients who will benefit sufficiently from short-term psychotherapy and those for whom long-term psychotherapy is required. Data on dose–effect relationships suggest that most patients experiencing acute distress benefit from short-term psychotherapy.¹ Short-term psychotherapy may be defined as a treatment of up to 25 sessions,² applying this definition to the data reported by Kopta *et al*, about 70% of the patients with acute distress recovered after short-term therapy.¹ For patients with chronic distress, about 60% recovered after 25 sessions. For patients with characterological distress, i.e. personality disorders, the data of Kopta *et al* suggest that about 40% recovered after 25 sessions.¹ Perry *et al* estimated the length of treatment necessary for patients with personality disorder to achieve recovery (defined as no longer meeting the full criteria for a personality disorder): according to these estimates, half of such patients would recover after 1.3 years or 92 sessions, and three-quarters after 2.2 years or about 216 sessions.³ Summing up, the majority of patients with acute distress benefit significantly from short-term psychotherapy, whereas for many patients with chronic distress and for the majority of patients with personality disorders, short-term psychotherapy seems not to be sufficient.

Evidence-based treatments for these groups of patients are particularly important. Personality disorders, for example, are not uncommon in both general and clinical populations. They show a high comorbidity with a wide range of Axis I disorders and are significantly associated with functional impairments.^{4–6} Furthermore, personality disorders were found to have a negative prognostic impact on depressive disorders.⁷ For this reason, experts recommend not focusing on the depressive disorder but primarily treating the associated personality disorder.^{7,8} Another population for whom short-term treatment may not be sufficient are those with multiple mental disorders. A high proportion of patients in clinical populations have not just one but several

mental disorders, and such patients report significantly greater deficits in social and occupational functioning.^{9,10}

Some data suggest that long-term psychotherapy may be helpful for these groups of patients.^{1,3,11–14} This is true not only for psychodynamic therapy, but also for psychotherapeutic approaches that are usually short-term, such as cognitive–behavioural therapy (CBT).^{14,15} For long-term psychodynamic psychotherapy (LTPP), however, strong evidence-based support as yet is lacking. In a recent meta-analysis of the effectiveness of LTPP we focused on complex mental disorders which were defined as personality disorders, chronic mental disorders or multiple mental disorders.¹⁶ Twenty-three studies fulfilled the inclusion criteria. Both randomised controlled trials (RCTs) and quasi-experimental observational studies were included, allowing us to test for differences between study type. As the number of controlled studies was small, we calculated within-group effect sizes throughout. Large and stable effect sizes were reported for LTPP in patients with these complex disorders.¹⁶ For the studies including control groups, we compared the within-group effect sizes between the LTPP conditions and the control conditions: effect sizes for LTPP were significantly larger than those in the control conditions.¹⁶ However, comparing within-group effect sizes between treatments uses treatment conditions rather than studies as units of analysis, which may reduce the effect of randomisation.¹⁷ This may weaken internal validity, but it does not necessarily imply that internal validity is severely impaired.¹⁸ In order to address this problem we decided to update this meta-analysis, including new studies where available. For the comparison of LTPP and the control conditions between-group effect sizes were assessed, focusing on complex mental disorders as defined above. Our 2008 meta-analysis was criticised by some authors for addressing an ‘unconventionally broad research question’ by including heterogeneous patient populations and comparison conditions.¹⁷ On the contrary, however, researchers often adopt unnecessarily narrow entry criteria; a broad

perspective on meta-analysis covering different patient populations and settings increases the generalisability and usefulness of results.¹⁹ If results are not homogeneous, subgroup analysis can be used to examine the reasons. In the 2008 meta-analysis we carried out several subgroup analyses for different diagnostic groups.¹⁶ In line with these considerations, our updated meta-analysis focused on complex mental disorders (again defined as personality disorders, chronic mental disorders or multiple mental disorders), addressing the question whether LTPP is superior to shorter or less intensive psychotherapy in treating these disorders.

Method

The procedures followed in our study are consistent with recent guidelines for the reporting of meta-analyses.²⁰

Definition of LTPP

Psychodynamic psychotherapy serves as an umbrella concept encompassing treatments that operate on a continuum of supportive–interpretive psychotherapeutic interventions.^{2,21–23} Interpretive interventions aim to enhance patients' insight into repetitive conflicts sustaining their problems;² supportive interventions aim to strengthen abilities that are temporarily inaccessible to patients owing to acute stress (e.g. traumatic events) or have not been sufficiently developed (e.g. impulse control in borderline personality disorder). The establishment of a helping (or therapeutic) alliance is regarded as an important component of supportive interventions.²² Transference, defined as the repetition of past experiences in present interpersonal relations, constitutes another important dimension of the therapeutic relationship. In psychodynamic psychotherapy, transference is regarded as a primary source of understanding and therapeutic change.^{2,22} The emphasis that psychodynamic psychotherapy puts on the relational aspects of transference is a key technical difference from cognitive–behavioural therapies.²⁴ The use of more supportive or more interpretive (insight-enhancing) interventions depends on the patient's needs. The more severely disturbed a patient is or the more acute the problem, the greater is the need for supportive interventions, whereas an emphasis on interpretive approaches is more suitable for less disturbed patients.²² Psychodynamic psychotherapy can be carried out either as a short-term (time-limited) or as a long-term open-ended treatment. Open-ended psychotherapy in which treatment duration is not fixed *a priori* is not identical to unlimited psychotherapy.²² Short-term treatments are time-limited, usually lasting between 7 and 24 sessions.² There is no generally accepted standard duration for long-term psychodynamic psychotherapy. Lamb compiled more than 20 definitions given by experts in the field,²⁵ ranging 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 1 year or 50 sessions. This criterion is consistent with the definition given by Crits-Christoph & Barber and other experts in the field.²⁶

Inclusion criteria and selection of studies

We applied the following inclusion criteria, consistent with recent meta-analyses of psychotherapy:²⁷

- (a) studies of psychodynamic therapy meeting the definition given above;^{2,21–23}
- (b) psychodynamic therapy lasting for at least 1 year or at least 50 sessions;

- (c) active treatments applied in the control conditions;
- (d) prospective studies of LTPP including pre- and post-treatment or follow-up assessments;
- (e) treatments must have been terminated (no study assessing outcome for ongoing treatments);
- (f) use of reliable and valid outcome measures;
- (g) a clearly described sample of patients with 'complex' disorders (personality disorders, chronic mental disorders or more than one mental disorder);
- (h) adult patients (at least 18 years of age);
- (i) sufficient data to allow determination of between-group effect sizes.

We collected studies of LTPP that were published between January 1960 and April 2010 based on our previous meta-analysis and an updated computerised search of Medline, PsycINFO and Current Contents.¹⁶ 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*) AND (rct* OR control* OR compar*). In addition, articles and textbooks were manually searched, and we communicated with authors and experts in the field.

Data extraction

We independently extracted the following information from the articles: author names, publication year, psychiatric disorder treated with LTPP, age and gender 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 *v.* effectiveness), duration of follow-up period and use of psychotropic medication.¹⁶ Disagreements were resolved by consensus. Rating was done without masking to treatment condition, since evidence suggests that such masking is unnecessary for meta-analyses.²⁸ Effect sizes were independently assessed by two raters. Interrater reliability was assessed for the outcome domains in question: overall outcome, target problems, psychiatric symptoms, personality functioning and social functioning. For all areas interrater reliability was high ($r \geq 0.95$, $P \leq 0.002$).¹⁶

Assessment of effect sizes and statistical analysis

We assessed effect sizes for target problems, psychiatric symptoms, personality functioning, social functioning and overall outcome. As outcome measures of target problems, we included both patient ratings of target problems and measures referring to the symptoms specific to the patient group under study (e.g. measures of depression in treatment studies of major depressive disorder or a measure of impulsivity for studies examining borderline personality disorder).²⁹ For psychiatric symptoms we included both broad measures of psychiatric symptoms such as the Symptom Check List 90 (SCL-90) and specific measures such as measures of depression or anxiety.³⁰ For the assessment of personality functioning, measures of personality characteristics were included (e.g. the Millon Clinical Multiaxial Inventory).³¹ Social functioning was assessed using the Social Adjustment Scale and similar measures.³² Whenever a study reported multiple measures for one of the areas of functioning (e.g. target

psychiatric symptoms), we assessed the effect size for each measure separately and calculated the mean effect size of these measures within each study. In our previous meta-analysis outcome measures were assigned either to target problems or to psychiatric symptoms, personality functioning or social functioning.^{16,27} In a study of depressive disorders, for example, a reduction in depression could be attributed only to target problems, not to psychiatric symptoms. However, this procedure may artificially narrow the data basis for the estimation of actual therapeutic effects in the respective outcome areas. In order to avoid this problem in this meta-analysis, we first assigned each outcome measure to one (and only one) of the three domains of psychiatric symptoms, personality functioning or social functioning. Overall outcome was assessed by averaging the effect sizes of these three areas. To obtain information about changes in target problems, outcome measures referring to criteria specific to the patient group under study (e.g. measures of depression in depressive disorders), which were in the first step of evaluation assigned to one of the aforementioned three areas, were additionally assigned to the domain of target problems. This means that the results for target problems are not independent of the other three areas, but more realistic estimates of therapeutic effects will be achieved. As a measure of between-group effect size for continuous measures, we calculated Hedges' *d* and the associated 95% confidence interval.³³ This measure is a variation of Cohen's *d* which corrects for bias due to small sample sizes.³³ Hedges' *d* was calculated by subtracting the mean pre-treatment to post-treatment or follow-up difference of the control condition from the corresponding difference of LTTP, divided by the pooled pre-treatment standard deviation. This quotient was multiplied by a coefficient *J* correcting for small sample size to obtain Hedges' *d*. If a study included more than one LTTP or comparison group, we used the averaged effect sizes of these groups. We aggregated the effect sizes estimates (Hedges' *d*) across studies, adopting a random effects model which is more appropriate if the aim is to make inferences beyond the observed sample of studies.³⁴ To obtain a mean effect sizes estimate we used MetaWin version 2.0 for Windows.³⁵ If the data necessary to calculate effect sizes were not published in the article, we asked its authors for this information. If necessary, signs were reversed so that a positive effect size always indicated improvement. In order to examine the stability of psychotherapeutic effects, we assessed effect sizes separately for assessments at the termination of therapy and follow-up. If data pertaining to completers and intention-to-treat (ITT) samples were reported, the latter were included. To control for bias related to withdrawal, we additionally carried out ITT analyses. For studies that did not report ITT data we conservatively set the effects for patients who withdrew after randomisation to zero. By this procedure, the effect sizes reported for the completers sample were adjusted for missing ITT data. If a study, for example, reported a pre-post treatment difference of 0.40 for a group of 20 patients who completed the study with 5 patients having withdrawn, we used an adjusted difference of 0.32 ($0.40 \times 20/25$) for the ITT analysis. Tests for heterogeneity were carried out using the *Q* statistic.³³ To assess the degree of heterogeneity, we calculated the *I*² index.³⁶ In cases of significant heterogeneity random effect models are more appropriate.^{34,37} To control for publication bias, tests for asymmetry in funnel plots and 'file drawer' analyses were performed.³⁶⁻³⁹ Statistical analyses were conducted using SPSS version 15.0 and MetaWin version 2.0.^{35,40} Two-tailed tests of significance were carried out for all analyses. The significance level was set to $P = 0.05$ unless otherwise stated. If more is better, outcome should increase with dosage and duration of treatment. For this analysis we used within-group effect sizes which were calculated for each condition by

subtracting the post-treatment mean from the pre-treatment mean and dividing the difference by the pooled pre-treatment standard deviation of the measure.^{41,42} If more than one LTTP condition or more than one control condition was included, we treated them separately in this analysis. Spearman correlations were assessed between within-group effect sizes and both duration of treatment and number of sessions.

Assessment of study quality

According to the inclusion criteria described earlier, we analysed only prospective studies of LTTP 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.*⁴³ This scale takes into account whether a study is described as randomised and double-blind, and whether withdrawals and 'drop-outs' are itemised. In psychotherapy research, however, studies cannot be double-blind because the participants know or can easily find out which treatment they receive.¹⁶ Thus, all studies of psychotherapy would have to be given a score of zero on this item of the Jadad scale. Instead of masking of therapists and patients, the respective requirement in psychotherapy research is that any observer-rated outcome measure is rated by assessors unaware of 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 one point on this item if outcome was assessed by masked raters or by reliable self-report instruments.¹⁶ With this modification, the three items of the Jadad scale were independently rated by us for all studies included; a satisfactory interrater reliability was achieved for the total score of the scale ($r = 0.92$, $P < 0.001$).

Results

Ten studies met the inclusion criteria (Fig. 1).^{11,13,44-51} For three of these studies we received additional information from the authors.^{13,46,51} Levy *et al* reported additional data on outcome for the study by Clarkin *et al.*^{13,52} In contrast to our 2008 meta-analysis, we now included the supportive treatment of the study by Clarkin *et al* as a form of LTTP because of its description by Levy *et al* as a psychodynamic therapy.^{13,52} The study by Korner *et al* used a non-randomised comparison group.⁵⁰ Meta-analytic results, however, have shown that non-randomised comparison group designs yield comparable – if anything, slightly smaller – effect size estimates to randomised designs.⁵³ For this reason we included the study by Korner *et al.*⁵⁰ In an RCT by Knekt *et al* comparing LTTP, short-term psychodynamic psychotherapy and (short-term) solution-focused therapy in long-standing depressive and anxiety disorders, the authors assessed the effects of the short-term treatment groups at predefined time points that did not exactly represent end of therapy for the short-term treatments.^{49,54} Mean duration of treatment was 5.7 months and 7.5 months respectively for these treatments.⁴⁹ To include the study by Knekt *et al* in this meta-analysis, we used the effects of the short-term treatments assessed after 9 months, which is the time point following most closely the end of the short-term treatments. As the effect sizes at 9 months were almost identical to those found at 7 months, no bias was introduced by this procedure. For LTTP we used the outcome assessed after 36 months (end of treatment). In another new RCT, Bateman & Fonagy compared LTTP (mentalisation-based treatment) with a structured clinical management approach in the treatment of patients with borderline personality disorder.⁴⁵ In addition, we received further

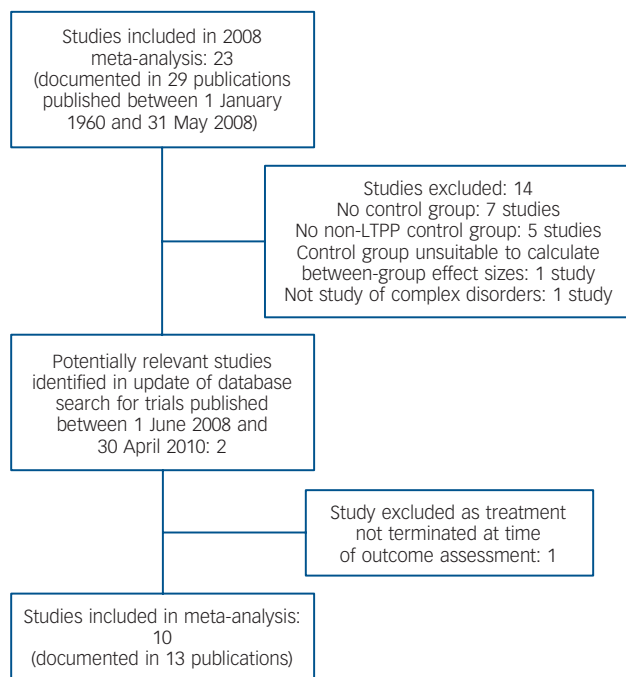


Fig. 1 Selection of trials for update of authors' 2008 meta-analysis of long-term psychodynamic psychotherapy (LTPP).¹⁶

information about another RCT of LTPP which fulfilled the inclusion criteria.⁴⁸ Huber & Klug provided us with data on the comparison groups of their study that were unavailable at the time of our previous meta-analysis.^{16,48} Thus, we included this study in this meta-analysis as another RCT. As both the analytic psychotherapy and the long-term psychodynamic psychotherapy group of that study fulfilled our criterion for LTPP, we included both treatments in this category. The ten studies included are described in online Table DS1.

Tests for publication bias

To reduce the 'file drawer' effect we tried to identify unpublished studies through the internet and by contacting researchers. To test for publication bias we calculated correlations between sample size and between-group effect sizes across studies. A significant correlation may indicate a publication bias in which larger effect sizes in one direction are more likely to be published.⁵⁵ Alternatively, the standard error instead of the sample size can be used to test for publication bias. Owing to the small number of studies providing follow-up data, we assessed these correlations only for the post-treatment between-group effect sizes. Since for comparisons with treatment as usual (TAU) smaller sample sizes (and larger between-group effect sizes) can be expected than for a comparison with a specific form of psychotherapy, we calculated partial correlations in order to control for the type of comparison condition (TAU *v.* specific psychotherapy). According to the results, the mean partial correlation between outcome and sample size was $r_p = 0.05$ (range -0.06 to 0.14 , $P > 0.73$); for outcome and standard error, r_p was 0.16 ($P > 0.46$). As another test for publication bias we assessed the fail-safe number according to Rosenthal: this is the number of non-significant, 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 significant to non-significant.³⁹ An effect size can be regarded as robust if the

fail-safe number exceeds $5K + 10$, where K is the number of studies.⁵⁶ For overall outcome the fail-safe number was 66. As this exceeds 60 ($5K + 10$), the effect can be regarded as robust. Summing up, we did not find any cogent indication of publication bias.

Total number of participants

The ten studies included encompassed 466 patients treated with LTPP and 505 patients receiving comparative treatments.

Therapy duration

For LTPP the mean number of sessions in the ten studies was 120.5 (s.d. = 117.5) and the mean duration of therapy was 78.0 weeks (s.d. = 38.2). For the treatments in the control groups the mean number of sessions was 45.4 (s.d. = 28.1) and the mean duration of therapy was 62.9 weeks (s.d. = 24.0).

Follow-up assessments

Follow-up assessments were carried out in three studies, at intervals of 1–8 years.^{48,51,57}

Mental disorders

The ten controlled studies of complex mental disorders included the treatment of patients with long-standing depressive and anxiety disorders (two studies),^{48,49} cluster C personality disorders (one study),⁵¹ borderline personality disorder (five studies),^{11,13,45,47,50} and eating disorders (two studies).^{44,46} As the number of studies was too small to conduct separate analyses for specific disorders we combined them into one group called 'complex mental disorders'.

Comparison groups

The psychotherapeutic treatments applied in the comparison groups included cognitive (behavioural) therapy (CBT/CT; three groups),^{44,48,51} cognitive analytic therapy (one group),⁴⁶ dialectical behavioural therapy (DBT; one group),¹³ short-term psychodynamic psychotherapy (one group),⁴⁹ solution-focused therapy (one group),⁴⁹ family therapy (one group),⁴⁶ structured clinical management (one group),⁴⁵ and routine psychiatric treatment as usual (four groups).^{11,46,47,50} In addition, one study of eating disorders included nutritional counselling as another control condition.⁴⁴ The authors described this condition as not including psychotherapy. Including nutritional counselling as one of the control conditions of LTPP might lead to underestimating the effects of the control conditions. For this reason we did not include this therapy in the comparison conditions of this meta-analysis. Because of the small number of studies examining one specific comparison treatment, we did not carry out separate analyses for the different comparison conditions (e.g. LTPP *v.* CBT) but combined the treatments into one group called 'less intensive forms of psychotherapy'. According to this procedure the question of whether LTPP yielded a better outcome than less intensive forms of psychotherapy was studied.

Treatment manuals

Treatment manuals or manual-like guidelines for LTPP were applied in all but two studies.^{48,49}

Tests for heterogeneity

We used the Q statistic to test for heterogeneity of between-group effect sizes,^{33,35} and the I^2 index to assess the degree of

Table 1 Comparison of long-term psychodynamic psychotherapy with other forms of psychotherapy: between-group effect sizes

Outcome domain	Number of comparisons	Hedges' <i>d</i>		<i>Q</i>	<i>I</i> ² , %
		<i>d</i> ^a	95% CI ^b		
Overall effectiveness	10	0.54 (0.52)	0.26–0.83	11.72	23
Target problems	9	0.49 (0.48)	0.27–0.71	9.12	12
Psychiatric symptoms	9	0.44 (0.41)	0.15–0.73	11.52	31
Personality functioning	7	0.68 (0.63)	0.31–1.04	5.97	0
Social functioning	8	0.62 (0.59)	0.18–1.06	12.44	44

ITT, intention to treat.
a. Adjusted for ITT sample.
b. Unadjusted *d*.

heterogeneity (Table 1).³⁶ For *Q*, all tests of significance yielded insignificant results ($P \geq 0.09$). The *I*² index for overall outcome, target problems, symptoms, personality functioning and social functioning indicated low to moderate heterogeneity (Table 1).³⁶ For follow-up, the number of studies providing data was too limited to calculate reasonable *Q* and *I*² statistics.

Correlation of quality ratings with outcome

In order to examine the relationship between study quality and outcome the between-group effect sizes were correlated with the total score of the Jadad scale for overall outcome, target problems, general symptoms, personality functioning and social functioning. Owing to the small number of studies providing follow-up data, correlations were only calculated for post-treatment assessment effect sizes. For this purpose, the average quality score of the two raters was used. All correlations were non-significant ($P > 0.14$, r_s -0.13 to 0.53). Although not statistically significant, the Spearman correlation was relatively high for symptoms ($r = 0.53$). Accordingly, studies of higher quality tended to yield larger between-group effect sizes in favour of LTPP for psychiatric symptoms.

Effects of LTPP v. other methods of psychotherapy

Because of the small number of studies providing data for follow-up assessments, between-group effect sizes were only assessed for the post-therapy data, except for some preliminary analyses. Between-group effect sizes (Hedge's *d*) in overall outcome

are presented for each of the ten studies (Fig. 2). The random effects model was applied in order to aggregate effect sizes across studies: the differences in outcome between LTPP and other forms of psychotherapy in complex mental disorders were 0.54, 0.49, 0.44, 0.68 and 0.62 respectively for overall outcome, target problems, psychiatric symptoms, personality functioning and social functioning (Table 1). The ITT analysis yielded similar results (Table 1). According to Cohen these effect sizes can be regarded as medium to large.⁴¹ All between-group effect sizes differed significantly from zero ($P < 0.05$). Effect sizes can be transformed into percentiles:⁴¹ for example, a between-group effect size of 0.54 as identified in overall outcome indicates that after treatment with LTPP, patients on average were better off than 70% of the patients treated in the comparison groups. Only three studies provided data to assess between-group effect sizes for follow-up assessments.^{48,51,57} For this reason the results are only preliminary. For these three studies the between-group effect sizes were 0.55, 0.54, 0.48, 0.76 and 0.37 respectively for overall outcome, target problems, psychiatric symptoms, personality functioning and social functioning. According to these data the differences in favour of LTPP at follow-up are comparable with those at the end of treatment.

Correlations of outcome with dosage and duration

Including all treatment conditions (LTPP and non-LTPP), all outcome variables except for target problems showed significant Spearman correlations with the number of sessions (Table 2). Treatment duration was significantly correlated with improvements

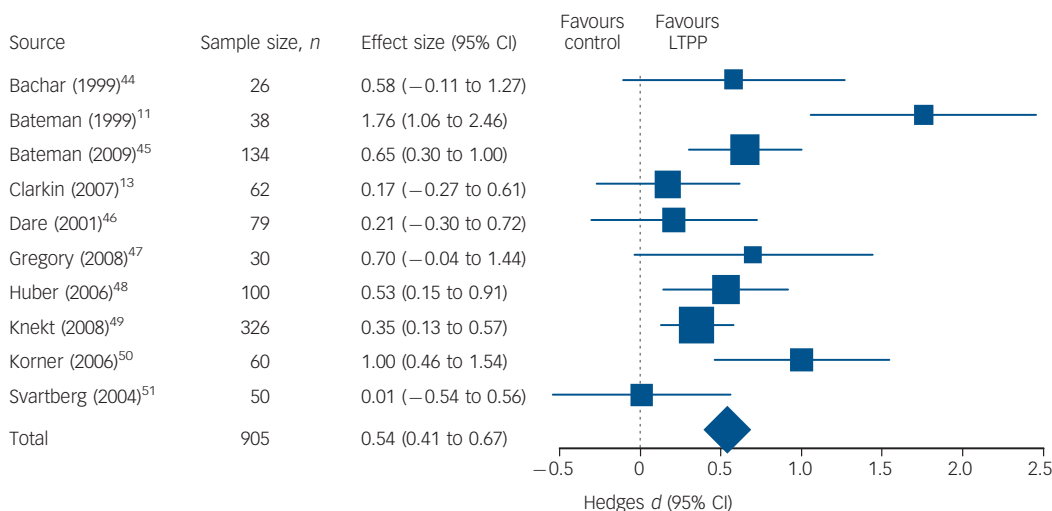


Fig. 2 Comparative effects of long-term psychodynamic psychotherapy (LTPP) on overall outcome (number of patients in analysis sample may differ from intention-to-treat sample).

Table 2 Spearman correlations of outcome (pre–post treatment effect sizes) with duration of therapy and number of treatment sessions

	Overall outcome	Target problems	Psychiatric symptoms	Personality functioning	Social functioning
All treatment conditions					
Duration	0.36*	0.24	0.39*	0.19	0.50*
Sessions	0.54**	0.33	0.37*	0.48*	0.63**
LTPP only					
Duration	0.59*	0.28	0.83**	0.18	0.57*
Sessions	0.68*	0.28	0.67*	0.31	0.79**
Control conditions only					
Duration	−0.10	−0.23	−0.19	0.30	0.37
Sessions	0.02	−0.05	−0.19	− ^a	0.20

LTPP, long-term psychodynamic psychotherapy.
a. Insufficient data to calculate correlations.
* $P < 0.05$, ** $P < 0.01$ (one-tailed).

in overall outcome, psychiatric symptoms and social functioning. The other correlations were of small to medium size but insignificant owing to the small number of conditions. Both the direction and significance of correlations of outcome with duration or dosage of therapy are consistent with the results that showed superiority of LTPP over shorter-term treatments.

In some studies treatment lasted for a year or more but comprised fewer than 50 sessions (online Table DS1). In order to control for the effect of dosage of LTPP, we additionally assessed Spearman correlations between pre–post effect sizes and the number of sessions for the LTPP conditions only (Table 2). Again, all correlations were positive. These correlations were large (> 0.50) and significant for overall outcome, symptoms and social functioning. For target problems and personality functioning, small to medium correlations were found that were insignificant. Thus, the inclusion of studies in the LTPP group in which the number of sessions was less than 50 can be assumed to have reduced the effects of LTPP. In the control conditions only, no significant correlation was found (Table 2).

As a further check regarding the importance of dosage, we assessed between-group effect sizes without those studies in which fewer than 50 sessions were applied in the LTPP conditions (Dare *et al*, Bachar *et al*, Svartberg *et al*).^{44,46,51} For all outcome measures the effect sizes increased after exclusion of the these three studies (overall outcome from 0.54 to 0.66; target problems from 0.49 to 0.55; psychiatric symptoms from 0.44 to 0.55; personality functioning from 0.68 to 0.77; social functioning from 0.62 to 0.72).

Discussion

A considerable proportion of patients with chronic mental disorders or personality disorders do not benefit sufficiently from short-term psychotherapy.^{1,3} Long-term psychotherapy, however, is associated with higher direct costs than short-term psychotherapy. For this reason it is important to know whether the effects of long-term psychotherapy exceed those of short-term treatments. In this meta-analysis, LTPP was superior to less intensive methods of psychotherapy in complex mental disorders. Furthermore, we found positive correlations between outcome and duration or dosage of therapy. Both of these results are consistent with data on dose–effect relations.¹

One limitation of this meta-analysis may be seen in the scarcity of controlled studies. Further studies of LTPP are required to confirm the results and allow for more refined analyses. With a small number of studies it is of particular importance to test for publication bias. For that purpose, we applied several measures. Fail-safe number analysis indicated that for overall outcome, 66

studies would need to be added to this meta-analysis in order to change the results of the meta-analysis from significant to non-significant. Furthermore, we found no significant correlation between outcome and sample size nor with standard error of effect sizes. We also found no significant correlation between outcome and the methodological quality of the studies as assessed using the scale proposed by Jadad *et al*.⁴³ However, the size of some correlations may indicate a systematic relationship, in that studies of higher quality tended to yield larger between-group effect sizes in favour of LTPP. Another limitation can be seen in the small number of studies that reported follow-up assessments. It is of interest to know whether the between-group effect sizes in favour of LTPP are stable beyond the end of treatment. The results of our previous meta-analysis suggest that the effects of LTPP even increase after the end of treatment.¹⁶ When follow-up data from the studies included are available, this question can be addressed directly. As another limitation, not all studies reported ITT analyses. In this meta-analysis, however, we could show that adjusting for missing ITT data did not substantially change the results. Nonetheless, future studies should include ITT analyses whenever possible.

Duration of therapy

There is no generally accepted standard duration for LTPP. We included studies that lasted for at least a year or in which at least 50 sessions were applied. In some studies treatment lasted for a year or more but comprised fewer than 50 sessions; for this reason, some of these studies were included in previous meta-analyses as short-term. This was true, for example, for the study by Svartberg *et al* in which 40 sessions were applied.^{27,51} Apparently, the inclusion of studies depends on the question of research addressed and the specific definition that is used in a meta-analysis. The correlations between dosage and outcome in the LTPP studies reported above suggest that the inclusion of studies in which LTPP lasted for fewer than 50 sessions reduced the treatment effects of LTPP. However, including only studies that fulfilled both the dosage and the duration criteria would have further reduced the already small number of studies. Future meta-analyses of LTPP or of long-term psychotherapy in general should include studies that fulfil both the dosage and the duration criteria. Furthermore, a differentiation between long-term, medium-term and short-term therapy might be useful.

Critical discussion of results

This meta-analysis took several points of critique put forward against our 2008 meta-analysis into account, such as lack of

between-group effect sizes or of ITT analyses, possible publication bias or inclusion of inactive control conditions.^{17,58} According to the results presented here we did not find cogent indication for any systematic bias. The methodological quality both of our meta-analyses and of the studies included is comparable to that of many studies of CBT.⁵⁹

Some controlled studies did not meet the inclusion criteria because the majority of patients had not completed their treatment when the effect sizes were assessed. This was true, for example, for the studies by Brockmann *et al*, Doering *et al*, Giesen-Bloo *et al* and Puschner *et al*.^{15,60–62} In the study by Giesen-Bloo, for example, 19 of 42 patients treated with LTPP (45%) were still in treatment 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.¹⁵ Data from ongoing treatments do not provide reliable estimates for treatment outcome at termination or follow-up, for example if patients had received only half of the 'dose' of treatment when outcome was assessed. By analogy, if one runner enters a 100 m race and a second enters a 10 000 m race, the time taken after 100 m will not be representative of the short-distance speed of the second runner. The runners will adapt their speed to the short or long distance they are going to face. This is true for patients in psychotherapy as well.⁴⁹ Psychotherapy is not a drug that works equally under different conditions, but a psychosocial process.

We compared the effects of LTPP with a group of mixed psychotherapeutic treatments. The control conditions consisted of specific forms of psychotherapy, including established forms such as CBT or DBT, as well as several TAU conditions. Including TAU can be assumed to reduce the mean effect size of the control group; on the other hand, the control conditions included not only short-term psychotherapy but also long-term psychotherapy applied as long as LTPP in the respective studies (e.g. DBT, CBT), in turn increasing the mean effect of the control condition. It is noteworthy that it was on average that duration and the number of treatment sessions applied was higher in the LTPP conditions. Thus, we used the alternative treatments as an unspecific (mixed) control group including both TAU and specific forms of alternative psychotherapy. Consequently, we do not claim that LTPP is superior to any specific form of psychotherapy in complex mental disorders that is carried out equally intensively, rather that it is superior to less intensive forms of psychotherapeutic interventions in general. We expect this to be true for other more intensive approaches of formal psychotherapy as well, for example that higher-dose CBT is superior to lower-dose CBT in borderline personality disorder. For psychodynamic psychotherapy this should also be true. With regard to the hierarchy of evidence, our comparison of LTPP with a mixed control group including TAU and specific psychotherapy is stricter than a comparison with a waiting-list group, placebo therapy or pure TAU, but less strict (and specific) than a comparison with specific or established forms of psychotherapy only.^{63,64}

Future research

Without doubt comparisons of LTPP with specific therapies are desirable, both short-term and long-term. At present, however, not enough studies are available. For CBT or DBT more comparative studies exist. Thus, it would be interesting to compare long-term CBT or DBT with short-term CBT or DBT in specific mental disorders. For some mental disorders for which response rates are not satisfactory, such as social anxiety disorder, experts in the field propose increasing treatment duration.⁶⁵

Falk Leichsenring, DSc, Department of Psychosomatics and Psychotherapy, University of Giessen; **Sven Rabung**, PhD, Department of Medical Psychology, University Medical Centre Hamburg-Eppendorf, Germany

Correspondence: Professor Dr Falk Leichsenring, Department of Psychosomatics and Psychotherapy, University of Giessen, Ludwigstrasse 76, 35392 Giessen, Germany. Email: Falk.Leichsenring@psycho.med.uni-giessen.de

First received 18 May 2010, final revision 6 Nov 2010, accepted 17 Jan 2011

Acknowledgements

We thank Dr Martin Rock (Yeshiva University, New York, USA) for information concerning the study by Clarkin *et al*, and Drs D. Huber and G. Klug and Drs A. Bateman and P. Fonagy for giving us access to their data. We also thank John Clarkin, PhD (Weill Medical College of Cornell University, New York, USA), Ivan Eisler, PhD (Institute of Psychiatry, King's College London, UK), Paul Knekt, PhD (National Public Health Institute, Helsinki, Finland) and Martin Svartberg, MD, PhD (Department of Psychiatry, Mount Sinai Hospital, Toronto, Ontario, Canada) for information about their studies.

References

- 1 Kopta SM, Howard KI, Lowry JL, Beutler LE. Patterns of symptomatic recovery in psychotherapy. *J Consult Clin Psychol* 1994; **62**: 1009–16.
- 2 Gabbard GO. *Long-term Psychodynamic Psychotherapy: A Basic Text*. American Psychiatric Publishing, 2004.
- 3 Perry J, Banon E, Floriana I. Effectiveness of psychotherapy for personality disorders. *Am J Psychiatry* 1999; **156**: 1312–21.
- 4 Grant BF, Hasin DS, Stinson FS, Dawson DA, Chou SP, Ruan WJ, et al. Prevalence correlates and disability of personality disorders in the United States. Results from the International Epidemiologic Survey on Alcohol and Related Conditions. *J Clin Psychiatry* 2004; **65**: 948–58.
- 5 Kantojärvi L, Veijola J, Laksy K, Jokelainen J, Herva A, Karvonen JT, et al. Co-occurrence of personality disorders with mood, anxiety, and substance abuse disorders in a young adult population. *J Pers Disord* 2006; **20**: 102–12.
- 6 Lenzenweger MF, Lane MC, Loranger AW, Kessler RC. Personality disorders in the National Comorbidity Survey Replication. *Biol Psychiatry* 2007; **62**: 553–64.
- 7 Gunderson JG, Morey LC, Stout RL, Skodol AE, Shea MT, McGlashan TH, et al. Major depressive disorder and borderline personality disorder revisited: longitudinal interactions. *J Clin Psychiatry* 2004; **65**: 1049–56.
- 8 Elkin I, Shea MT, Watkins JT, Imber SD, Sotsky SM, Collins JF, et al. National Institute of Mental Health Treatment of Depression Collaborative Research Program. General effectiveness of treatments. *Arch Gen Psychiatry* 1989; **46**: 971–83.
- 9 Olsson M, Fireman B, Weissman MM, Leon AC, Sheehan DV, Kathol RG, et al. Mental disorders and disability among patients in a primary care group practice. *Am J Psychiatry* 1997; **154**: 1734–40.
- 10 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**: 1741–8.
- 11 Bateman A, Fonagy P. The effectiveness of partial hospitalization in the treatment of borderline personality disorder, a randomized controlled trial. *Am J Psychiatry* 1999; **156**: 1563–9.
- 12 Bateman A, Fonagy P. Treatment of borderline personality disorder with psychoanalytically oriented partial hospitalization: an 18-month follow-up. *Am J Psychiatry* 2001; **158**: 36–42.
- 13 Clarkin JF, Levy KN, Lenzenweger MF, Kernberg OF. Evaluating three treatments for borderline personality disorder: a multiwave study. *Am J Psychiatry* 2007; **164**: 922–8.
- 14 Linehan MM, Comtois KA, Murray AM, Brown MZ, Gallop RJ, Heard HL, et al. Two-year randomized controlled trial and follow-up of dialectical behavior therapy vs. therapy by experts for suicidal behaviors and borderline personality disorder. *Arch Gen Psychiatry* 2006; **63**: 757–66.
- 15 Giesen-Bloo J, van Dyck R, Spinhoven P, van Tilburg W, Dirksen C, van Asselt T, et al. Outpatient psychotherapy for borderline personality disorder: randomized trial of schema-focused therapy vs transference-focused psychotherapy. *Arch Gen Psychiatry* 2006; **63**: 649–58.
- 16 Leichsenring F, Rabung S. The effectiveness of long-term psychodynamic psychotherapy: a meta-analysis. *JAMA* 2008; **300**: 1551–64.
- 17 Kriston L, Hölzel L, Härter M. Analyzing effectiveness of long-term psychodynamic psychotherapy [letter]. *JAMA* 2009; **301**: 930–1.
- 18 Leichsenring F, Rabung S. Analyzing effectiveness of long-term psychodynamic psychotherapy [reply]. *JAMA* 2009; **301**: 932–3.

- 19 Gotzsche PC. Why we need a broad perspective on meta-analysis. *BMJ* 2000; **21**: 585–6.
- 20 Moher D, Liberati A, Tetzlaff J, Altman D. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *BMJ* 2009; **339**: b2535.
- 21 Gunderson JG, Gabbard GO. Making the case for psychoanalytic therapies in the current psychiatric environment. *J Am Psychoanal Assoc* 1999; **47**: 679–704.
- 22 Luborsky L. *Principles of Psychoanalytic Psychotherapy. Manual for Supportive-Expressive Treatment*. Basic Books, 1984.
- 23 Wallerstein RS. Psychoanalytic treatments within psychiatry: an expanded view. *Arch Gen Psychiatry* 2002; **59**: 499–500.
- 24 Cutler JL, Goldyne A, Markowitz JC, Devlin MJ, Glick RA. Comparing cognitive behavior therapy, interpersonal psychotherapy, and psychodynamic psychotherapy. *Am J Psychiatry* 2004; **161**: 1567–73.
- 25 Lamb WK. *A meta-analysis of outcome studies in long-term psychodynamic psychotherapy and psychoanalysis*. Dissertation. ProQuest, 2004.
- 26 Crits-Christoph P, Barber JP. Long-term psychotherapy. In *Handbook of Psychological Change: Psychotherapy Processes and Practices for the 21st Century* (eds RE Ingram, CR Snyder): 455–73. Wiley, 2000.
- 27 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**: 1208–16.
- 28 Berlin JA. Does blinding of readers affect the results of meta-analyses? *Lancet* 1997; **350**: 185–6.
- 29 Battle CC, Imber SD, Hoehn-Saric R, Nash ER, Frank JD. Target complaints as criteria of improvement. *Am J Psychother* 1966; **20**: 184–92.
- 30 Derogatis L. *The SCL-90 Manual I: Scoring, Administration and Procedures for the SCL-90-R*. Clinical Psychometric Research, 1977.
- 31 Millon T. *Millon Clinical Multiaxial Inventory* (3rd edn). National Computer Services, 1984.
- 32 Weissman MM, Bothwell S. Assessment of social adjustment by patient self-report. *Arch Gen Psychiatry* 1976; **33**: 1111–5.
- 33 Hedges LV, Olkin I. *Statistical Methods for Meta-analysis*. Academic Press, 1985.
- 34 Hedges LV, Vevea JL. Fixed- and random-effects models in meta-analysis. *Psychol Methods* 1998; **3**: 486–504.
- 35 Rosenberg MS, Adams DC, Gurevitch J. *MetaWin: Statistical Software for Meta-analysis*, version 2.0. Sinauer, 1999.
- 36 Huedo-Medina TB, Sanchez-Meca J, Botella J, Marin-Martinez F. Assessing heterogeneity in meta-analysis. Q statistic or I^2 index. *Psychol Methods* 2006; **11**: 193–206.
- 37 Quintana SM, Minami T. Guidelines for meta-analyses of counseling psychology research. *Couns Psychol* 2006; **34**: 839–77.
- 38 Begg CB, Mazumdar M. Operating characteristics of a rank correlation test for publication bias. *Biometrics* 1994; **50**: 1088–101.
- 39 Rosenthal R. The file drawer problem and tolerance for null results. *Psychol Bull* 1979; **86**: 638–41.
- 40 SPSS. *SPSS 15.0 Command Syntax Reference*. SPSS, 2006.
- 41 Cohen J. *Statistical Power Analysis for the Behavioral Sciences*. Erlbaum, 1988.
- 42 Hedges LV. Distribution theory for Glass's estimator for effect size and related estimators. *J Educ Stat* 1981; **6**: 107–28.
- 43 Jadad AR, Moore RA, Carroll D, Jenkinson C, Reynolds DJ, Gavaghan DJ, et al. Assessing the quality of reports of randomized clinical trials: is blinding necessary? *Control Clin Trials* 1996; **17**: 1–12.
- 44 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**: 115–28.
- 45 Bateman A, Fonagy P. Randomized controlled trial of outpatient mentalization-based treatment versus structured clinical management for borderline personality disorder. *Am J Psychiatry* 2009; **166**: 1355–64.
- 46 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**: 216–21.
- 47 Gregory RJ, Chlebowski S, Kang D, Remen AL, Soderberg MG, Stepkovich J, et al. A controlled trial of psychodynamic psychotherapy for co-occurring borderline personality disorder and alcohol use disorder. *Psychotherapy* 2008; **45**: 28–41.
- 48 Huber D, Klug G. Munich Psychotherapy Study (MPS): the effectiveness of psychoanalytic long-term psychotherapy for depression. In *Book of Abstracts. From Research to Practice* (ed Society for Psychotherapy Research): 154. Ulmer Textbank, 2006.
- 49 Knekt P, Lindfors O, Harkanen T, Valikoski M, Virtala E, Laaksonen MA, 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**: 689–703.
- 50 Korner A, Gerull F, Meares R, Stevenson J. Borderline personality disorder treated with the conversational model: a replication study. *Compr Psychiatry* 2006; **47**: 406–11.
- 51 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**: 810–7.
- 52 Levy KN, Meehan KB, Kelly KM, Reynoso JS, Weber M, Clarkin JF, 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**: 1027–40.
- 53 Lipsey MW, Wilson DB. The efficacy of psychological, educational, and behavioral treatment. Confirmation from meta-analysis. *Am Psychol* 1993; **48**: 1181–209.
- 54 Knekt P, Lindfors O, Laaksonen MA, Raitasalo R, Haaramo P, Järviskoski A, 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**: 95–106.
- 55 Begg CB. Publication bias. In *The Handbook of Research Synthesis* (eds H Cooper, LV Hedges): 399–409. Russell Sage Foundation, 1995.
- 56 Rosenthal R. *Meta-analytic Procedures for Social Research: Applied Social Research Methods*. Sage, 1991.
- 57 Bateman A, Fonagy P. 8-year follow-up of patients treated for borderline personality disorder: mentalization-based treatment versus treatment as usual. *Am J Psychiatry* 2008; **165**: 631–8.
- 58 Beck AT, Bhar SS. Analyzing effectiveness of long-term psychodynamic psychotherapy [letter]. *JAMA* 2009; **301**: 931.
- 59 Leichsenring F, Rabung S. Double standards in psychotherapy research. *Psychother Psychosom* 2011; **80**: 48–57.
- 60 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**: 15–25.
- 61 Puschner B, Kraft S, Kächele H, Kordy H. Course of improvement over 2 years in psychoanalytic and psychodynamic outpatient psychotherapy. *Psychol Psychother* 2007; **80**: 51–68.
- 62 Doering S, Hörz S, Rentrop M, Fischer-Kern M, Schuster P, Benecke C, et al. Transference-focused psychotherapy v. treatment by community psychotherapists for borderline personality disorder: randomised controlled trial. *Br J Psychiatry* 2010; **196**: 389–95.
- 63 Chambles DL, Hollon SD. Defining empirically supported therapies. *J Consult Clin Psychol* 1998; **66**: 7–18.
- 64 Gabbard GO, Gunderson JG, Fonagy P. The place of psychoanalytic treatments within psychiatry. *Arch Gen Psychiatry* 2002; **59**: 505–10.
- 65 Zaider T, Heimberg RG. Non-pharmacologic treatments for social anxiety disorder. *Acta Psychiatr Scand Suppl* 2003; **417**: 72–84.

