Clinical Psychology Review 32 (2012) 81–92

Contents lists available at SciVerse ScienceDirect

Clinical Psychology Review

The effectiveness of long-term psychoanalytic psychotherapy—A meta-analysis of randomized controlled trials

Yolba Smit [a], Marcus J.H. Huibers [b,*], John P.A. Ioannidis [c], Richard van Dyck [d], Willem van Tilburg [d], Arnoud Arntz [b,e]

[a] Leuth, The Netherlands
[b] Department of Clinical Psychological Science, Maastricht University, The Netherlands
[c] Health Research and Policy, and Statistics, Stanford Prevention Research Center, Stanford University School of Medicine, Stanford, CA, USA
[d] Department of Psychiatry, VU University Hospital Amsterdam, The Netherlands
[e] The Netherlands Institute for Advanced Study in the Humanities and Social Sciences, Royal Academy of Netherlands Sciences, Wassenaar, The Netherlands

ARTICLE INFO

Article history:
Received 28 April 2011
Received in revised form 12 October 2011
Accepted 20 November 2011
Available online 4 December 2011

Keywords:
Long-term psychoanalytic psychotherapy
Psychoanalysis
Meta-analysis

ABSTRACT

The effectiveness of psychoanalysis and long-term psychoanalytic psychotherapy (LTPP) is debated. We evaluated the effectiveness of LTPP, compared to other treatments or no treatment, in patients with clearly defined metal disorders. We selected randomised or quasi-randomised controlled trials on LTPP. Two authors independently identified trials for inclusion. Eleven trials were eligible. The risk difference for recovery (primary outcome) at the longest available follow-up was 0.00 (95% CI: −0.17 to 0.17; p = 0.96; I-squared: 58%). The combined Hedges' g, at the longest follow-up for each study, were: for target problems: −0.05 (95% CI −0.55 to 0.46; p = 0.86; I-squared = 88%); general psychiatric symptoms: 0.69 (95% CI −0.19 to 1.57; p = 0.13; I-squared = 96%); personality pathology: 0.17 (95% CI: −0.25 to 0.59; p = 0.42; I-squared = 41%); social functioning: 0.20 (95% CI −0.10 to 0.50; p = 0.19; I-squared = 53%); overall effectiveness: 0.33 (95% CI −0.31 to 0.96; p = 0.32; I-squared = 94%); and quality of life: −0.37 (95% CI: −0.78 to 0.04; p = 0.08; I-squared = 55%). A subgroup analysis of the domain target problem showed that LTPP did significantly better when compared to control treatments without a specialized psychotherapy component, but not when compared to various specialized psychotherapy control treatments. An exploratory meta-regression indicated that there might be a relation between the difference in treatment intensity between the intervention and control group (session ratio) and effect size. We came to conclude that the recovery rate of various mental disorders was equal after LTPP or various control treatments, including treatment as usual. The effect sizes of the individual trials varied substantially in direction and magnitude. In contrast to previous reviews, we found the evidence for the effectiveness of LTPP to be limited and at best conflicting.

Contents

* Corresponding author.
E-mail address: m.huibers@maastrichtuniversity.nl (M.J.H. Huibers).

0272-7358/$ – see front matter © 2011 Elsevier Ltd. All rights reserved.
doi:10.1016/j.cpr.2011.11.003

1. Introduction

Psychoanalysis is known for its strong theoretical background that expanded over the decades, but the empirical evidence for the psychoanalytic theory and its practice has been limited and fragmented. Thus the effectiveness of psychoanalysis and psychoanalytic therapy has been debated. Freud himself contributed to the discussion in 1937: *´One has the impression that one ought not to be surprised if it should turn out in the end that the difference between a person who has not been analyzed and the behavior of a person after he has been analyzed is not so thorough-going as we aim at making it and as we expect and maintain it to be´* (Freud, 1961).

The debate continues today. While the effectiveness of other, mainly short-term, forms of psychotherapy—such as cognitive behaviour therapy, interpersonal psychotherapy and short-term psychoanalytical psychotherapy—has been scrutinised in a large number of controlled trials, controlled research that focuses on long-term, psychoanalytically rooted therapies is sparse. These long-term therapies can be classified as either classical psychoanalysis, or long-term psychoanalytical psychotherapy (LTPP). Psychoanalysis in its classical form is a therapy that stretches over many years, with four to five weekly sessions, in a setting where the patient lies on a couch and the therapist sits behind that couch, thereby conveying an abstinent position. LTPP is an adapted form of classical psychoanalysis, with several different schools existing within LTPP practice, which may lead to differences in LTPP practice. Therapy sessions usually have a frequency of once or twice a week, and the therapist and patient sit while facing each other, as is customary in most forms of psychotherapy. In this meta-analysis, we focus solely on LTPP. We did search extensively for randomised controlled trials on psychoanalysis, but did not find any.

Two recent meta-analyses have previously evaluated the accumulated evidence on LTPP (De Maat, De Jonghe, Schoevers, & Dekker, 2009; Leichsenring & Rabung, 2008; Leichsenring & Rabung, 2011). They concluded that LTPP is an effective treatment for mental disorders, with large standardized overall effect sizes of 0.94 (95% CI not reported) (De Maat et al., 2009) and 1.8 (95% CI 0.7–3.4) (Leichsenring & Rabung, 2008). However, one may wonder whether these conclusions are valid. A major objection is that both meta-analyses synthesized data from within-group differences (i.e. pre-post change), instead of between-group differences. To reliably assess the effectiveness of any treatment, it is necessary to evaluate its outcomes compared to a control group. The change in severity or intensity of a mental disorder over time cannot be attributed solely to the treatment that took place during that time, unless the treatment is controlled for. This is especially so with long-term treatments where the course of symptoms may change (more or less) spontaneously over time, even in personality disorders that were previously thought to be stable and incurable, such as borderline personality disorder. In an update of their meta-analysis, Leichsenring and Rabung acknowledge the necessity to assess between-group differences (Leichsenring & Rabung, 2011). In this update, the overall effectiveness of LTPP was 0.54 (95% CI 0.41–0.67). However, out of the ten included studies, one study did not use randomisation or quasi-randomisation. By contrast, our objective is to systematically review the effectiveness of LTPP in the recovery of patients with a clearly defined mental disorder, as examined in randomised or quasi-randomised controlled trials (RCTs). In addition, we wanted to know if LTPP led to patients' recovery more frequently than control treatments.

2. Methods

2.1. Eligibility

Randomised or quasi-randomised (e.g. randomisation by date of birth, alternation) controlled trials were eligible. Further inclusion criteria were: participants with any clearly defined mental disorder except for schizophrenia; long-term psychoanalytically based psychotherapy as an intervention; a control treatment that differed substantially from the intervention treatment (either a different type of treatment and/or a short-term treatment). Studies in schizophrenia patients were excluded, because LTPP is currently not used in these patients. In order to define the spectrum of LTPP, we asked all Dutch societies of psychoanalytical professionals to give us an overview of the psychotherapies they considered to be psychoanalytical in nature. Three societies sent us a joint overview of such therapies which we used to define LTPP. We defined long-term psychotherapy as having at least 40 sessions and continuing for at least one year. The meta-analyses of Leichsenring and De Maat define long-term as at least 50 sessions (De Maat et al., 2009; Leichsenring & Rabung, 2008). However, LTPP with a once-a-week frequency may result in a total of less than 50 sessions in a year, allowing for patients' and therapists' vacations and missed sessions. Our definition is in line with a Cochrane review on the effectiveness of short-term psychodynamic psychotherapies, which defined short-term as less than 40 sessions on average (Abbass, Hancock, Henderson, & Kisely, 2006). In addition, the Dutch societies of psychoanalytical professionals also defined long-term psychotherapy as more than 40 sessions.

2.2. Information sources and search

Medline, Embase, PsycINFO, the ACP Journal Club, the Cochrane Database of Systematic Reviews, the Cochrane Central Register of Controlled Trials, the Cochrane Methodology Register, the Database of Abstracts of Reviews of Effects (DARE), the National Health

Services Health Technology Assessment Database (NHS HTA) and the NHS Economic Evaluation Database (NHS EED) were searched through text words and indexing terms. Search filters were used to detect studies on LTPP (psychoanaly*, psychodynam*) and controlled studies (clinical trial, meta-analysis, randomized controlled trial, controlled clinical trial, evaluation studies, random*, trial, control*). Searches were performed in January 2009 and updated in July 2011 and were not limited by time-period, language or in any other way. References of meta-analyses, reviews and selected articles were scanned for additional relevant studies, and experts in the field were contacted for information on ongoing or unpublished studies.

2.3. Study selection

Two reviewers (MH and YS) independently selected studies for inclusion. When the researchers disagreed, consensus was reached through discussion.

2.4. Data collection process

Data were abstracted by one researcher (YS) and checked by a second researcher (MH). Outcomes assessed by independent assessors were chosen, if available. Intention-to-treat (ITT) data were used, whenever available. Authors were contacted if the reported data were insufficient or unclear.

2.5. Data items

We sought and extracted data for the following variables: date and place of the trial, sample population, inclusion criteria (including the disorders treated), exclusion criteria, intervention treatment, control treatment, mean number of sessions, mean number of sessions in completers, description of treatment intensity, treatment coherence (use of manuals and supervision, analysis of adherence, therapists' training), co-interventions (policy on co-interventions, description thereof), drop-outs, cross-over, adverse events, and all data relevant to our primary and secondary outcomes (see below). We used the mean number of sessions attended, and the described intensity of treatment to calculate a proxy of the session ratio (number of sessions in the intervention group vs. number of sessions in the control group). A session ratio of 1 would indicate that participants in the intervention group would have received an equal number of sessions as participants in the control group.

2.6. Risk of bias in individual studies

We used (a) the Maastricht–Amsterdam Criteria List (van Tulder, Assendelft, Koes, & Bouter, 1997) and (b) eight criteria proposed by Cuijpers et al. (2009). For the sake of relevance, we amended the criterion for question m2 of the Maastricht Amsterdam criteria list: 'was a long-term follow-up measurement performed?'. We answered this question with 'yes' if there was an outcome assessment more than 2 years after randomisation, instead of more than 6 months after randomisation.

2.7. Summary measures

We meta-analysed: (a) the between-group difference in recovery between treatment groups (primary outcome) and (b) between-group differences in change for the domains target problems, general psychiatric symptoms, personality pathology, social functioning, overall effectiveness and quality of life (secondary outcomes). Overall effectiveness was calculated as the unweighted mean effect size of all available outcomes in a study. Target problems were defined as the problem the treatment was primarily focusing at, and included recovery. In a study of the treatment of depression for example, a measure of depression severity would be a measure of the target problem. We included overall effectiveness for the sake of comparison with previous meta-analyses. However, we do not consider this a very useful outcome, because it is an unweighted mixture of all available outcomes and thus it cannot be clinically interpreted.

2.8. Synthesis of results

We calculated the difference in recovery rate between treatment groups (primary outcome). We calculated standardized differences in means for the outcome scores of intervention and control groups (secondary outcomes). If needed, signs were reversed so a higher score reflected improvement. We used Hedges' g as the metric of choice. Cohen's d tends to overestimate the effect size. A correction factor is used to convert Hedges' g to Cohen's d. This correction factor is very close to 1 unless the number of participants is very small (<10), so the difference is usually trivial (Borenstein, Hedges, Higgins, & Rothstein, 2009). When more than two outcome measures were available for one domain we used the mean effect size and the mean variance, first calculating Hedges' g for each effect size. To calculate the effect size of overall effectiveness we used the mean effect size of all available outcomes in a study. For the meta-analyses we used the longest available follow-up data because LTPP should bring about change that is stable in the long run. If no follow-up data were available we used end-of treatment data. For the sake of convenience we refer to all time points as follow-up, with the time span referring to the time between the start of treatment and the measurement of outcome.

All analyses used the random effects model because we expected that the data could be heterogeneous, given the large diversity in populations (mental disorders) and control treatments (Higgins, Thompson, & Spiegelhalter, 2009; Lau, Ioannidis, & Schmid, 1997). The LTPP group was named the intervention group and a non-LTPP group was named the control group. When more than two intervention groups were available in one study we selected the data from the outpatient individual LTPP intervention group for the main analysis. When more than one control group was available we made a selection for the main analysis based on the following sequence: (1) evidence-based treatment (for example: cognitive behaviour therapy, interpersonal psychotherapy and short-term psychoanalytical psychotherapy) for the condition under study; (2) short-term psychoanalytical psychotherapy (STPP); (3) structured, non-evidence based treatment with the most similar treatment intensity; (4) structured, non-evidence based treatment with the most similar treatment mode (individual or group therapy, outpatient or inpatient therapy); (5) treatment as usual (TAU) or other non-structured treatment. For three selected studies more than one control group was available. We chose the (1) cognitive orientation treatment and nutritional counselling control group over the nutritional counselling control group (Bachar, Latzer, Kreitler, & Berry, 1999); (2) the cognitive-analytic therapy control group over the family therapy group, and over the low contact routine treatment group (Dare, Eisler, Russell, Treasure, & Dodge, 2001); and (3) the STPP control group over the solution-focused therapy control group (Knekt et al., 2008).

Heterogeneity was assessed with the chi-square Q statistic, and the I-squared metric with its 95% confidence intervals (Ioannidis, Patsopoulos, & Evangelou, 2007). Heterogeneity is considered statistically significant for $p<0.10$ and values of I2 above 75% suggest very large heterogeneity beyond chance.

Comprehensive Meta-analysis Version 2 (Biostat Inc.) software package was used for all meta-analyses. Stata Release 10 (StataCorp LP). STATA™10.0 (StataCorp, College Station) was used to calculate 95% CI for frequencies and I-squared (Higgins & Thompson, 2002) and to calculate the significance of differences in baseline factors of individual studies using the *t*-test. Therapies refer to individual outpatient therapy, unless specified otherwise.

2.9. Risk of bias across studies

We constructed funnel plots and applied the Duval and Tweedie's trim and fill test. Inferences should be cautious given the limited number of studies (Lau, Ioannidis, Terrin, Schmid, & Olkin, 2006). We considered the internal validity score to be a covariate that might explain part of the heterogeneity in effect sizes (Cuijpers et al., 2009; Leichsenring & Rabung, 2008) and explored this by random effects meta-regression.

2.10. Additional analysis

We intended subgroup-analyses (per type of disorder, type of treatment and control etc.) and meta-regression to explore heterogeneity, but the small number of studies precluded many of the intended subgroup analyses. We explored the few subgroup analyses that were possible. We considered the proxy session ratio (number of sessions in the intervention group/number of sessions in the control group) to be a covariate that might explain part of the heterogeneity in effect sizes and explored this by meta-regression.

3. Results

3.1. Study selection—excluded studies

We screened a total of 4121 records, of which 4109 were excluded. Fig. 1 gives an overview of the selection process. The main reasons for exclusion were that a short-term treatment was involved, or that a study was not controlled. In one study the exact treatments received by the intervention and control group were unclear, as was the

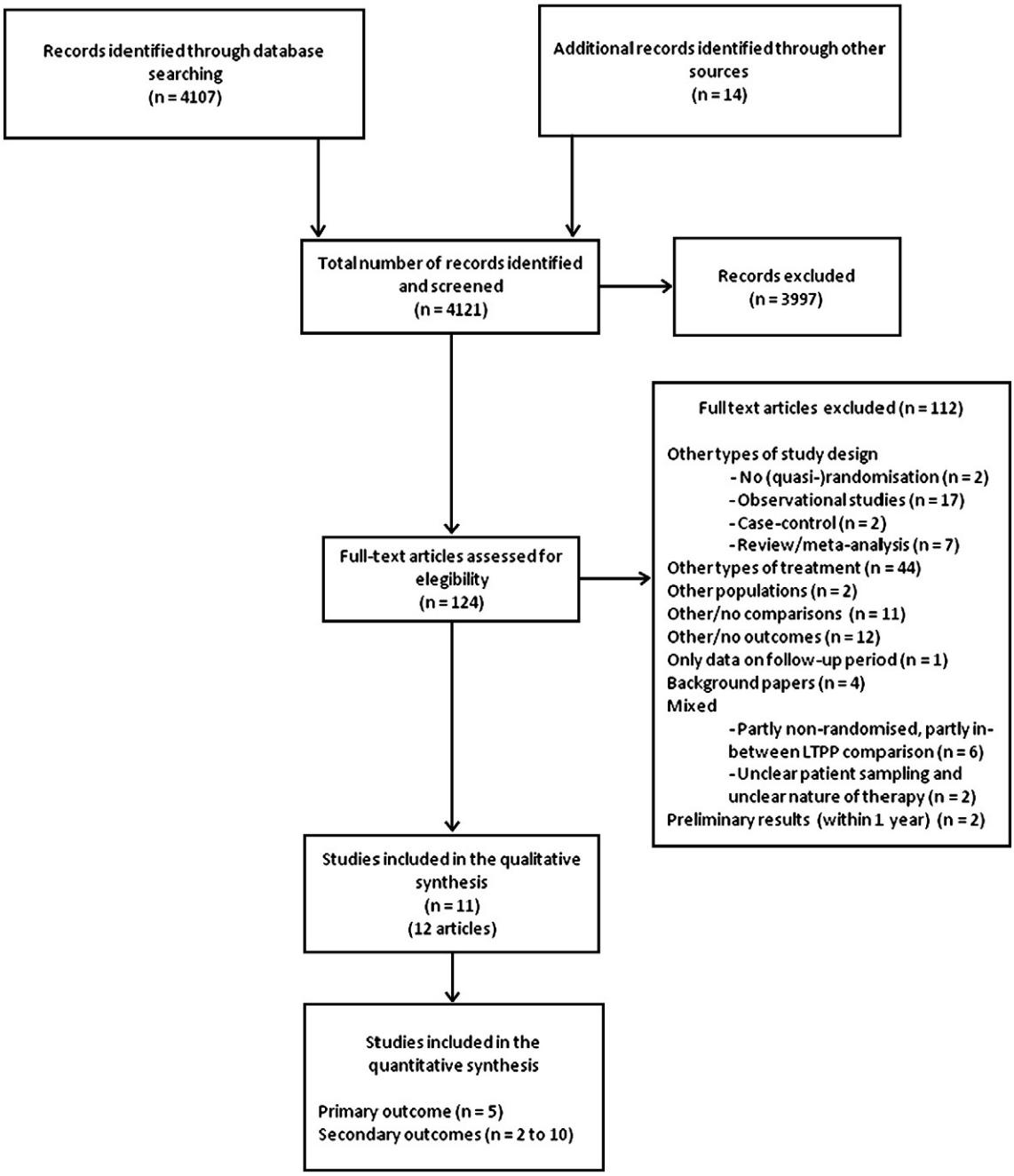

Fig. 1. Flowchart of study search and selection.

sampling of patients (Klar, 2005). Eight controlled studies on LTPP were excluded: in one study controls were sampled in retrospect (Korner, Gerull, Meares, & Stevenson, 2006). In one study the non-LTPP control group was sampled from a completely different population than the two LTPP groups (Chiesa, Fonagy, & Holmes, 2006). The two LTPP groups were recruited from patients referred to a tertiary care facility for inpatient treatment, whereas the control group was recruited from among the caseload of all the senior psychiatrists in a certain district, which makes these two patient groups incomparable as to the severity of their problems. In three other studies (described in five articles) a comparison was made between different forms of LTPP (Hoglend et al., 2006, 2008; Mintz, O'B'rien, & Luborsky, 1976; O'Brien et al., 1972; Vinnars, Barber, Noren, Gallop, & Weinryb, 2005); and three further studies reported data in such a way that we could not meta-analyse them (Clarkin, Levy, Lenzenweger, & Kernberg, 2007; Munroe-Blum & Marziali, 1995; Piper, Debbane, Bienvenu, & Garant, 1984). Of these three latter studies, one study reported the elevation (intercept) of the individual trajectory and the rate of change (slope) of the individual trajectory, but not the post-means and SDs (Clarkin et al., 2007). The second study reported no between-group differences but only the mean scores for all trial participants (Munroe-Blum & Marziali, 1995). In the third study standard deviations or other data to calculate the effect size were missing, so we were not able to meta-analyse the data (Piper et al., 1984). Authors were mailed to obtain the necessary data but we received no reply, or data were no longer available. Notably, we received no answer from Clarkin et al. whom did supply data to Leichsenring and Rabung (2008, 2011).

3.2. Study selection—included studies

Eleven randomised, controlled trials on LTPP were included (described in twelve articles) (Bachar et al., 1999; Bateman & Fonagy, 1999, 2009; Bressi, Porcellana, Marinaccio, Nocito, & Magri, 2010; Dare et al., 2001; Giesen-Bloo et al., 2006; Gregory et al., 2008; Knekt et al., 2008; Linehan et al., 2006; McMain et al., 2009; Svartberg, Stiles, & Seltzer, 2004; van Asselt et al., 2008). In particular, the inclusion of two studies (Bateman & Fonagy, 1999; Linehan et al., 2006) was discussed extensively by us. First, we questioned the type of intervention examined by Bateman and Fonagy (1999): could the reported outcomes be attributed to the psychoanalytical part of the therapy? The intervention (mentalisation-based therapy with partial hospitalisation) was an amalgamate of therapies conducted in an inpatient setting and included: '*1) once-weekly individual psychoanalytic psychotherapy, 2) thrice-weekly group analytic psychotherapy (1 h each), 3) once-a-week expressive therapy oriented toward psychodrama techniques (1 h), and 4) a weekly community meeting (1 h), all spread over 5 days*' (Bateman & Fonagy, 1999). In addition, all therapies were carried out by psychiatrically trained nurses from the hospital's team, who had no formal psychotherapy qualifications. Adherence to therapy was monitored, but by whom and how exactly was not described. We decided to include the Bateman and Fonagy (1999) study but to run a sensitivity analysis without it to check the robustness of our findings. In the second study the control group consisted of community treatment by experts, given by 25 therapists (Linehan et al., 2006). 21 out of 25 (84%) of the therapists described their methods as psychoanalytic or psychodynamic. Three others described themselves as interpersonal therapists, and one therapist described himself as humanistic/client centred [author's reply]. There was a weekly clinical supervision group available at which the therapists could attend. This group met at the Seattle Psychoanalytic Society and Institute and was led by its training director (Linehan et al., 2006). We decided to include the Linehan study; its control group was labelled the intervention group (and vice versa) in our report. To check the robustness of our findings we also ran sensitivity analyses without the Linehan study.

3.3. Study characteristics

Table 1 gives an overview of the main characteristics of the eleven included studies. One study concerned cluster C personality disorder patients, two studies concerned patients with eating disorders, two studies were on anxiety and/or mood disorder patients and six studies concerned patients with a borderline personality disorder. Most studies were small, except for the Knekt and McMain studies which included over 100 patients and 90 patients per treatment group respectively. In three studies there were twice as many non-completers in the intervention groups, compared to the control groups (Giesen-Bloo et al., 2006; Knekt et al., 2008; Linehan et al., 2006). Four studies reported outcomes at follow-up (Bachar et al., 1999; Giesen-Bloo et al., 2006; Linehan et al., 2006; Svartberg et al., 2004); the other seven studies reported outcomes at the end of treatment at the latest. Three studies did not use intention-to-treat data in their analysis (Bachar et al., 1999; Bateman & Fonagy, 1999; Svartberg et al., 2004).

3.4. Risk of bias within studies

We assessed the quality of each study using the Maastricht Amsterdam criteria (van Tulder et al., 1997) and eight criteria proposed by Cuijpers et al. (2009). The inter-rater agreement was 80% overall (77% for the Maastricht Amsterdam criteria and 87% for the criteria used by Cuijpers et al.). Consensus on the quality rating was achieved through discussion and the final ratings for the criteria used by Cuijpers et al. are depicted in Table 2. The quality of selected studies was variable. Most importantly, the quality of internal validity was relatively low with the 'best' studies (Bateman & Fonagy, 2009; Bressi et al., 2010) scoring 7 out of 9 points at the Maastricht Amsterdam criteria score for internal validity (a maximum score of 10 is impossible as the blinding of care providers is not possible), and three studies (Bachar et al., 1999; Bateman & Fonagy, 1999; Svartberg et al., 2004) scoring zero points for internal validity according to the criteria proposed by Cuijpers et al. All studies used randomisation to allocate treatment. Only five out of eleven studies (Bateman & Fonagy, 2009; Bressi et al., 2010; Giesen-Bloo et al., 2006; Knekt et al., 2008; Linehan et al., 2006) described an adequate concealment of treatment allocation. Notably, only three out of eleven studies (Bateman & Fonagy, 2009; Gregory et al., 2008; Linehan et al., 2006) explicitly described the blinding of outcome assessors. Co-interventions, adverse events and compliance were not monitored systematically in most studies. In two studies the treatment groups were not similar regarding the most important prognostic factors. This imbalance was in favour of the LTPP group for one study (Bressi et al., 2010) and in favour of the control group in the other study (Giesen-Bloo et al., 2006). Notably, in one study the baseline scores of the SCL-90-R phobic anxiety score differed significantly between the two groups (with the more favourable mean score in the LTPP group) (Bressi et al., 2010). Anxiety disorders were one of the target problems in this trial and the SCL-90-R was used as an outcome measure. In the other study the LTPP-group reported twice as many suicide planning, steps or attempts (Giesen-Bloo et al., 2006). Controlling for this baseline difference did not affect treatment differences (Giesen-Bloo & Arntz, 2007).

3.5. Synthesis of results: recovery (primary outcome)

Six studies (described in seven articles) gave information on the number of patients that recovered (Bachar et al., 1999; Dare et al., 2001; Giesen-Bloo et al., 2006; Gregory et al., 2008; Knekt et al., 2008; Svartberg et al., 2004; van Asselt et al., 2008). Van Asselt et al. gave recovery rates at 4 years for the Giesen-Bloo study. All six studies gave data on recovery from the targeted disorder, except the Gregory and Knekt studies. Gregory examined patients with a

Table 1
Main characteristics of included studies.

Study	Disorder	Included patients (non-completers)		LTPP intervention	Control intervention [$]	Proxy session ratio [†]	Assessment at	In previous meta-analyses
		LTPP group	Control group [$]					
Bachar (Bachar et al., 1999)	Eating disorder	17 (3)	17 (5)	Self psychological treatment	Cognitive orientation treatment	1.0	Pre- and post-treatment 1 year post-treatment	Leichsenring and Rabung (2008); Leichsenring and Rabung (2011)
Bateman (Bateman & Fonagy, 1999)	Borderline personality disorder	22 (3)	22 (3)	Mentalisation based therapy with partial hospitalisation	General psychiatric outpatient care with community support from mental health nurses, and periods of partial hospital and inpatient treatment as necessary; no specialist psychotherapy	1.8	Pre–treatment 1 year post-randomisation Post–treatment	Leichsenring and Rabung (2008); Leichsenring and Rabung (2011)
Bateman (Bateman & Fonagy, 2009)	Borderline personality disorder	71 (19)	63 (16)	Mentalisation based therapy	Structured clinical management outpatient approach, including advocacy, support, problem oriented activities and case management	NA	Pre-treatment 6 months post-randomisation 1 year post-randomisation Post-treatment	Leichsenring and Rabung (2011)
Bressi (Bressi et al., 2010)	Anxiety or depressive disorder	30 (6)	30 (6)	Psychodynamic psychotherapy	Drug treatment that included an SSRI/SNRI, combined with clinical interviews by the patient's treating psychiatrist, evaluating clinical state, compliance, adverse effects and medication adjustments	NA	Pre- and post-treatment	No
Dare (Dare et al., 2001)	Anorexia	21 (9)	22 (9)	Focal psychoanalytic psychotherapy	Cognitive analytic therapy	2.3	Pre- and post-treatment	Leichsenring and Rabung (2008); Leichsenring and Rabung (2011)
Giesen–Bloo (Giesen–Bloo et al., 2006; van Asselt et al., 2008)	Borderline personality disorder	42 (21) *	44 (11) *	Transference focussed therapy	Schema-focussed therapy	1.2	Pre-treatment 1 year post-randomisation 2 years post-randomisation Post-treatment 4 years post-randomisation	De Maat et al. (2009)
Gregory (Gregory, Chlebowski, Kang, Remen, & Soderberg, 2006)	Borderline personality disorder + alcohol use disorder	15 (5)	15 (6)	Dynamic deconstructive psychotherapy	Remain in current treatment and/or be referred to an alcohol rehabilitation centre; allowed to keep current psychotherapist, if any	0.6	Pre- and post-treatment	Leichsenring and Rabung (2008); Leichsenring and Rabung (2011))
Knekt (Knekt et al., 2008)	Anxiety and/or mood disorder	128 (46) *	101 (13) *	Psychodynamic psychotherapy	Short-term psychodynamic psychotherapy	5.0	Pre-treatment 1 year post-randomisation Post-treatment	Leichsenring and Rabung (2008); Leichsenring and Rabung, 2011)
Linehan (Linehan et al., 2006)	Borderline personality disorder and current and past suicidal behaviour	49 (21) *	52 (10) *	Expert treatment in the community (84% of therapists described their methods as psychoanalytic or psychodynamic)	Dialectical behaviour therapy	0.4	Pre- and post-treatment 1 year post-treatment	No
McMain (McMain et al., 2009)	Borderline personality disorder	90 (34)	90 (35)	Psychodynamic approach drawn from Gunderson	Dialectical behaviour therapy	0.4	Post-treatment	No
Svartberg (Svartberg et al., 2004)	Cluster C personality disorder	26 (1) **	26 (1) **	Dynamic psychotherapy	Cognitive therapy	1.0	Pre- and post-treatment 1 year post-treatment	Leichsenring and Rabung (2008); Leichsenring and Rabung (2011)

Abbreviations: NA: not available; LTPP: long term psychoanalytical psychotherapy; SNRI: serotonin-norepinephrine reuptake inhibitor; SSRI: selective serotonin re-uptake inhibitor.

[$] For the control treatment that was included in the meta-analysis.

[†] Approximation sessions in intervention group/approximation sessions in control group.

* Statistically significant difference in the proportion of non-completers.

** One patient out of a total of 51 randomised patients dropped out but it was unclear from which group.

Table 2
Quality criteria according to criteria used by Cuijpers et al. (2009).

Study / Criterion	Bachar	Bateman 1999	Bateman 2009	Bressi	Dare	Giesen-Bloo	Gregory	Knekt	Linehan	McMain	Svartberg
1. Patients diagnosed using diagnostic system	N	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
2. Use of treatment manual	N	N	Y	Y	N	Y	Y	N	N	Y	Y
3. Therapist trained for intervention under study	Y	N	Y	Y	N	Y	Y	Y	Y	Y	Y
4. Treatment integrity checked (supervision or analysis adherence)	Y	Y	Y	Y	Y	Y	Y	N	Y	Y	Y
5. Intention-to-treat analysis included	N	N	Y	Y	Y	Y	Y	Y	Y	Y	N
6. Adequate statistical power and n = > 50	N	N	Y	N	N	Y	N	Y	Y	Y	N
7. Randomization by independent person or computer	N	N	Y	Y	N	Y	N	Y	Y	Y	N
8. Outcome assessors blinded	N	N	Y	N	N	N	Y	N	Y	N	N
Total Yes (8 items)	2	2	8	6	3	7	6	5	7	7	4
–Total psychotherapy (items 1–4)	2	2	4	4	2	4	4	2	3	4	4
–Total internal validity (items 5–8)	0	0	4	2	1	3	2	3	4	3	0

Abbreviations: N: no; Y: yes.

borderline disorder and alcohol misuse, but recovery data were only available for alcohol misuse. The patients in Knekt's study were mixed in terms of diagnosis. 84.7% of patients had a mood disorder and 43.6% of patients had an anxiety disorder at baseline. Because all patients had to have at least one Axis I disorder and these two disorders were the only ones for which frequencies were given, we assumed that all participants had a mood disorder or an anxiety disorder at baseline. Thus 23.3% of participants had to have both disorders. Because of this overlap we could not treat anxiety disorder patients as being an independent group from mood disorder patients. We thus took the average recovery rate from anxiety disorder and for recovery from mood disorder (both in patients who had the disorder at baseline). Bachar was left out of the analysis on the longest follow-up available, because no data were reported at 2 year follow-up. It was only stated that there was no significant difference between treatment groups at that time. Because at 1 year follow-up a significant difference between groups was reported, we considered this to be a biassed reporting of outcome. We did include the 1-year follow-up data in a sensitivity analysis. For the Dare study we selected the control group that received cognitive analytical therapy as the comparison group (and not the family therapy group or the TAU group). In the Knekt study—that compared LTPP with STPP and a short-term non-evidence based control treatment—we selected the STPP group as the control group, and combined the two available recovery outcomes for this study (recovery from mood disorder and recovery from anxiety disorder).

The meta-analysed recovery difference at the longest available follow-up for each study was 0.00 (95% CI: −0.17 to 0.17; p = 0.96; p for heterogeneity = 0.05, I-squared: 58% (95%CI: 0–85%)) (Fig. 2) (Dare et al., 2001; Gregory et al., 2008; Knekt et al., 2008; Svartberg et al., 2004; van Asselt et al., 2008). These findings were robust and the summary effect remained close to zero and non-significant when we (a) removed each study arbitrarily; (b) used different control groups for the Dare and Knekt studies; (c) did include the Bachar data at 1 year; or (d) used the separate outcomes for the Knekt studies. The meta-analysed recovery difference at 1 year post-randomisation was 0.04 (95% CI: −0.15 to 0.24; p = 0.66; p for heterogeneity = 0.029, I-squared = 63%) (Bachar et al., 1999; Dare et al., 2001; Gregory et al., 2008; Knekt et al., 2008; Svartberg et al., 2004).

3.6. Synthesis of results: target problems, general psychiatric problems, personality pathology, social functioning, overall effectiveness and quality of life (secondary outcomes)

Table 3 and Figs. 3–8 give an overview of the meta-analysed between-group effect sizes at the longest available follow-up (or post-treatment if no follow-up data were available). The estimated effect sizes varied strongly, with negative meta-analysed effect sizes for the domains target problems and quality of life, and positive effect sizes for the domains general psychiatric problems, personality pathology, social functioning and overall effectiveness. However, none of the effect sizes was statistically significant and heterogeneity was large for all estimates. The findings were robust when we (a) removed each study arbitrarily; (b) used different control groups for the Dare and Knekt studies; or (c) did include the Bachar data at 1 year. A subgroup analysis for studies in borderline personality disorder patients indicated a negative but non-significant effect size for the domain target problems (Fig. 9). A subgroup analysis for

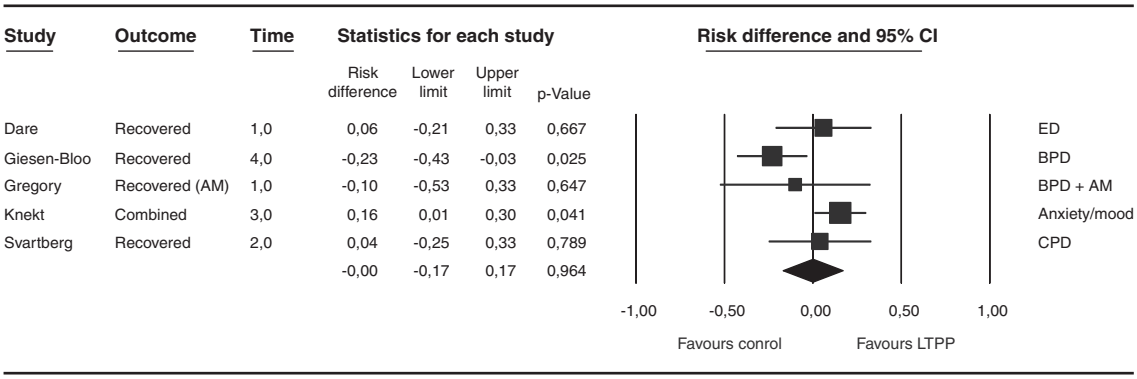

Fig. 2. Meta-analysed difference in recovery at the longest available follow-up. Abbreviations: AM: alcohol misuse; BPD: borderline personality disorder; CPD: cluster C personality disorder; CI: confidence interval; LTPP: long-term psychoanalytical psychotherapy. The outcome for Gregory is recovery from alcohol misuse. The outcome for Knekt is a combination of the outcomes: recovery from mood disorder and recovery from anxiety disorder. The recovery rate at 4 years for the Giesen-Bloo study is taken from van Asselt et al. (2008).

Table 3
Main meta-analysed between-group effect sizes, at the longest available follow-up or end-of-treatment.

Outcomes	Hedges' g (95% CI)	p-value	n	I-squared (95% CI)	p-value heterogeneity	References
Recovery	0.00 (−0.17, 0.17)*	0.96	5	58% (0–85%)	0.05	Dare et al. (2001); Gregory et al. (2008); Knekt et al. (2008); Svartberg et al. (2004); van Asselt et al. (2008))
Target problems	−0.05 (−0.55, 0.46)	0.86	9	88% (80–93%)	<0.01	Bateman and Fonagy (2009); Bressi et al. (2010); Dare et al. (2001); Giesen-Bloo et al. (2006); Gregory et al. (2008); Knekt et al. (2008); Linehan et al. (2006); McMain et al. (2009); Svartberg et al. (2004); van Asselt et al. (2008)
General psychiatric symptoms	0.69 (−0.19, 1.57)	0.13	8	96% (94–97%)	<0.01	Bateman and Fonagy (1999); Bateman and Fonagy (2009); Bressi et al. (2010); Gregory et al. (2008); Knekt et al. (2008); Linehan et al. (2006); McMain et al. (2009); Svartberg et al. (2004)
Personality pathology	0.17 (−0.25, 0.59)	0.42	2	41% **	0.19	Bateman and Fonagy (1999); McMain et al. (2009)
Social functioning	0.20 (−0.10, 0.50)	0.19	5	53% (0–83%)	0.07	(Bateman and Fonagy (2009); Bressi et al. (2010); Gregory et al. (2008); McMain et al. (2009); Svartberg et al. (2004)
Overall effectiveness	0.33 (−0.31, 0.96)	0.32	10	94% (90–96%)	<0.01	Bateman and Fonagy (1999); Bateman and Fonagy (2009); Bressi et al. (2010); Dare et al. (2001); Giesen-Bloo et al. (2006); Gregory et al. (2008); Knekt et al. (2008); Linehan et al. (2006); McMain et al. (2009); Svartberg et al. (2004); van Asselt et al. (2008)
Quality of life	−0.37 (−0.78, 0.04)	0.08	2	55% **	0.14	Giesen-Bloo et al. (2006); McMain et al. (2009)

Abbreviations: CI: confidence interval; n: number of studies included in the analysis.
* Meta-analysed risk difference for recovery.
** 95% CI for I-squared cannot be calculated for a meta-analysis of two studies.

studies that compared LTPP against a straw-man comparator (a comparator without specialized psychotherapy) indicated that LTPP did significantly better in the domain target problems than such comparators, but not than specialized psychotherapy treatments (Fig. 10). (For the primary outcome, recovery, there were too few studies to perform any meaningful subgroup analysis).

3.7. Risk of bias across studies

Using Duval and Tweedie's trim and fill test we did not find formal evidence of bias for the outcomes recovery, target problems, symptoms or social functioning. However, because of the small number of studies we feel we cannot draw a robust conclusion on small-study effects and bias in our review. The internal validity score did not explain effect size on any of the outcomes. We could not delineate a clear 'best category' of studies, to perform a subgroup analysis on. Subgroup analyses for studies with or without adequate concealment of treatment allocation, and for studies with or without blinded assessors showed no significantly different results.

3.8. Meta-regression

In exploratory random effects meta-regression, the session ratio was associated with the recovery difference(B = 0.07; 95%CI: 0.01 to 0.13; p = 0.03) and with the effect sizes of general psychiatric symptoms (B = 0.66; 95%CI: 0.57 to 0.76;p<0.001) and overall effectiveness (B = 0.52; 95%CI: 0.23 to 0.82; p<0.001), but not the effect size of target problems (B = 0.14; 95%CI: −0.25 to 0.53; p = 0.49) or social functioning (B = 0.66; 95%CI: −0.44 to 1.76; p = 0.24). We put the session ratio at '1' for the two studies for which a session ratio could not be calculated (Bateman & Fonagy, 2009; Bressi et al., 2010).

4. Discussion

4.1. Summary of evidence

The recovery rate of various mental disorders was equal after LTPP or various control treatments, including treatments without a specialized psychotherapy component. Similarly, no statistically

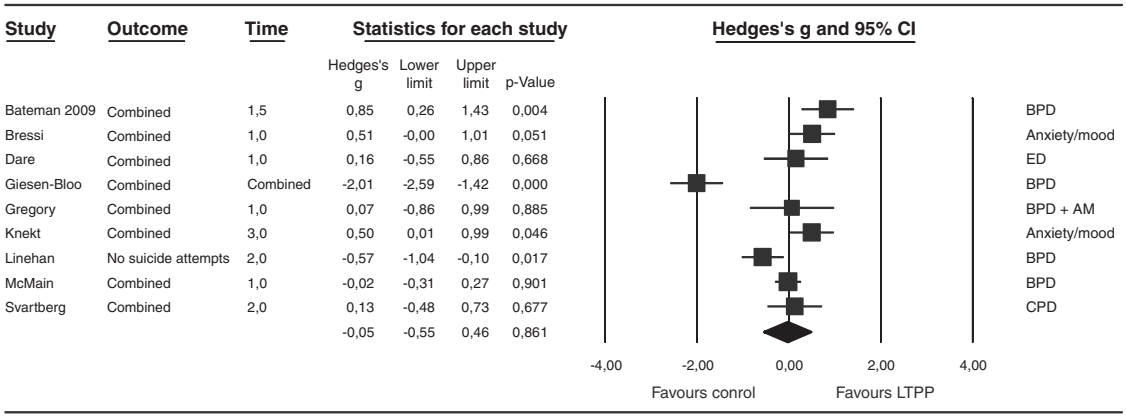

Fig. 3. Meta-analysed Hedges' g for target problems at the longest available follow-up. Abbreviations: AM: alcohol misuse; BPD: borderline personality disorder; CPD: cluster C personality disorder; CI: confidence interval; ED: eating disorder; LTPP: long-term psychoanalytical psychotherapy. The outcome target problems for Giesen-Bloo is a combination of recovery at 4 years (van Asselt et al., 2008) and the Borderline Personality Disorder Severity Index at 3 years.

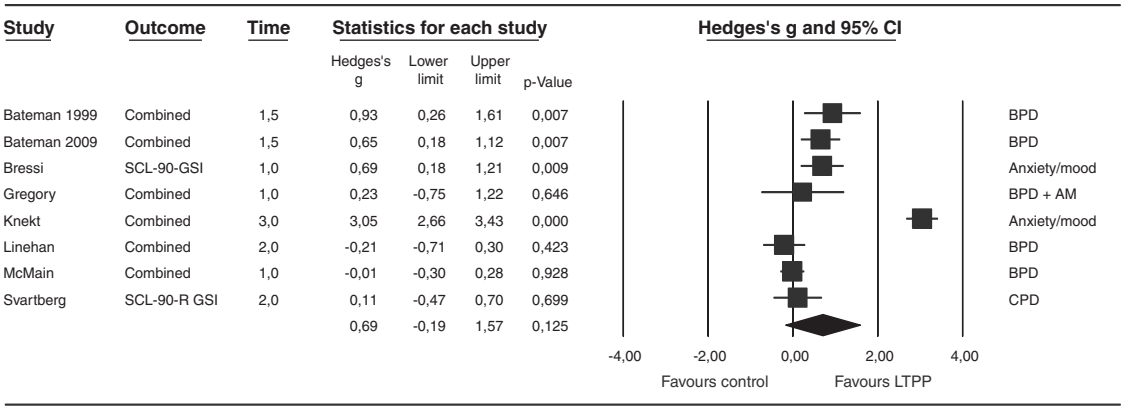

Fig. 4. Meta-analysed Hedges' g for general psychiatric symptoms at the longest available follow-up. Abbreviations: AM: alcohol misuse; BPD: borderline personality disorder; CPD: cluster C personality disorder; CI: confidence interval; ED: eating disorder; LTPP: long-term psychoanalytical psychotherapy.

significant differences were found for the domains target problems, general psychiatric problems, personality pathology, social functioning, overall effectiveness or quality of life. The variation in direction and magnitude of effect indicated that the observed effects in the included studies were highly variable. This makes the evidence on whether LTPP has effect on the recovery from various mental disorders conflicting. With only eleven studies available—and only five available studies on the primary outcome, recovery—the possibilities for meaningful subgroup analyses were very limited. The effects of LTPP for specific mental health disorders and/or against specific control treatments were represented typically by the results of single trials (for the primary outcome recovery) and thus should be interpreted very cautiously.

4.2. Limitations

The overall quality of studies was variable. Short post-treatment follow-up, or only reporting outcomes at the end of treatment, seems curious for trials of a long-term treatment modality. Control conditions were heterogeneous and frequently of low quality, e.g. without a specialized psychotherapy component. If anything, this suggests that LTPP is often compared against relatively ineffective "straw man" comparators (Ioannidis, 2008). LTPP comparisons to specialized non-psychodynamic treatments, like dialectical behaviour therapy and schema-focused therapy, suggest that LTPP might not be particularly effective. Any comparison with STPP is also complicated, as these studies do not inform us about the causes of a difference in effect size, apart from treatment duration. For example, differences might be purely attention and intensity effects, not related to psychoanalytic therapy *per se*. The effect sizes of individual studies varied substantially in direction and magnitude. Differences in disorders and populations, intervention and control treatments, outcome assessment instruments, settings etc. could explain a large part of this heterogeneity. Unfortunately, with so few data points available, both false-positive and false-negative findings can be expected. In exploratory meta-regression analyses we found some indication that effect size might be predicted by the proxy ratio of sessions across groups. Hopefully future studies will further explore this. If a relationship between effect size and session ratio exists, it would be of special interest to examine the effect size when the proxy session ratio equals 1 (indicating the same number of sessions in the intervention group and the control group).

Treatment confounders were present in all studies and included medication and other forms of therapy. It seems practically impossible to control for the use of additional or alternative treatments in an outpatient setting, and possibly more so in an inpatient setting as the Bateman study testifies. Pharmacotherapy cannot be excluded in some disorders, but should at least be monitored. Additional psychosocial treatment may be prohibited in some settings though even then its use should be monitored. Unfortunately most studies do not report treatment confounders in a systematic way. Treatment interaction may be a source of heterogeneity in some of the combined estimates.

We cannot directly compare our results for the primary outcome, recovery rate, against the two previous meta-analyses, because these meta-analyses did not examine the recovery rate. For comparative reasons we also meta-analysed overall effectiveness as an outcome. Our effect size of 0.33 (95% CI: −0.31 to 0.96; p = 0.32) strongly contrasts with previously reported pre-post change effect sizes of 0.94 (95% CI not reported) (De Maat et al., 2009) and 1.8 (95% CI: 0.7–3.4) (Leichsenring & Rabung, 2008). In an update of their review, Leichsenring and Rabung report a between-group overall effect size of 0.54 (95% CI: 0.41 to 0.67) (Leichsenring & Rabung, 2011). The difference between d = 0.33 and d = 0.54 might be explained by the inclusion of different studies and/or different choices in the selection of outcomes or control groups. Leichsenring and Rabung included

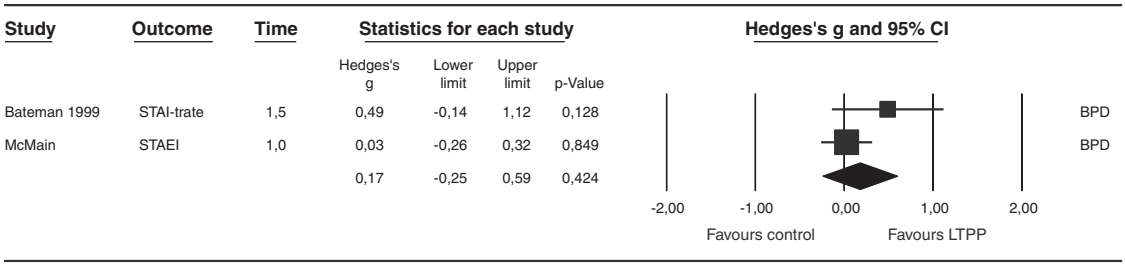

Fig. 5. Meta-analysed Hedges' g for personality pathology at the longest available follow-up. Abbreviations: BPD: borderline personality disorder; CI: confidence interval; LTPP: long-term psychoanalytical psychotherapy; STAI: State-Trait Anxiety Inventory; STAEI: State-Trait Anger Expression Inventory.

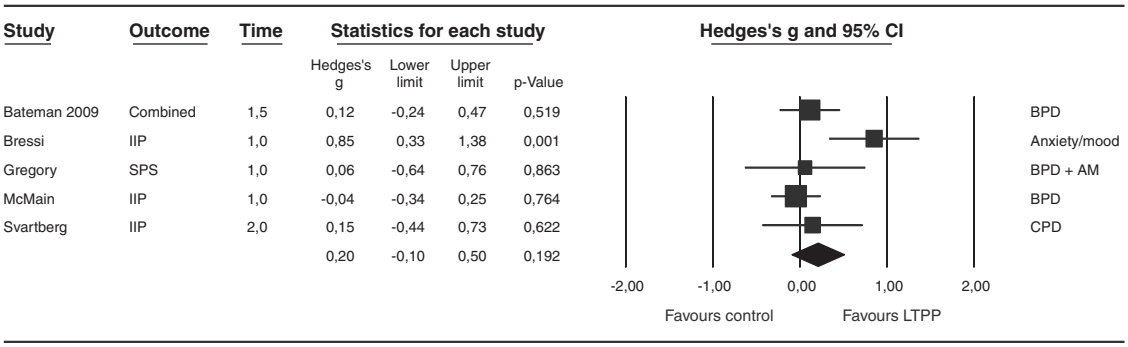

Fig. 6. Meta-analysed Hedges' g for social functioning at the longest available follow-up. Abbreviations: AM: alcohol misuse; BPD: borderline personality disorder; CI: confidence interval; CPD: cluster C personality disorder; IP: inventory of interpersonal problems; LTPP: long-term psychoanalytical psychotherapy; SPS: social provisions scale.

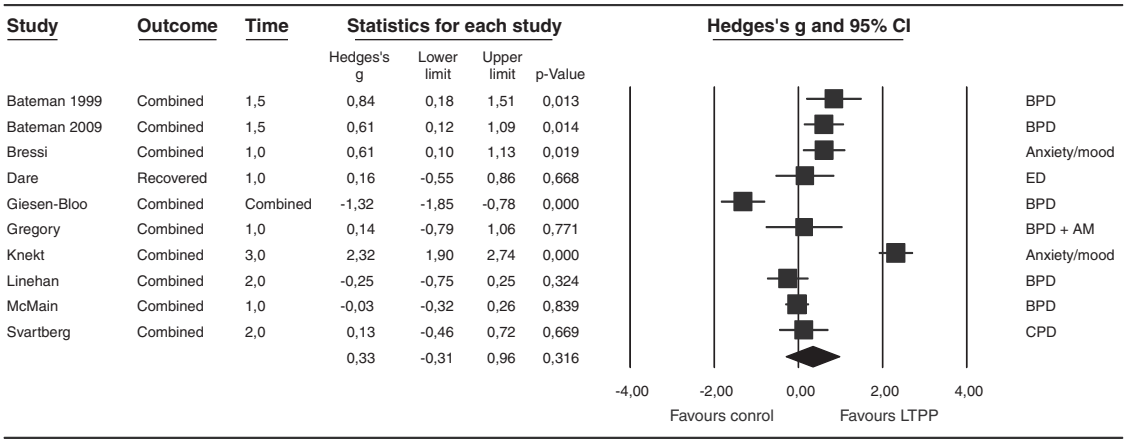

Fig. 7. Meta-analysed Hedges' g for overall effectiveness at the longest available follow-up. Abbreviations: AM: alcohol misuse; BPD: borderline personality disorder; CI: confidence interval; CPD: cluster C personality disorder; ED: eating disorder; LTPP: long-term psychoanalytical psychotherapy. The outcome overall effectiveness for Giesen-Bloo is a combination of recovery at 4 years (van Asselt et al., 2008) and the other available outcomes at 3 years.

three trials that we excluded because (a) we did not receive additional information that would enable us to meta-analyse the data (Clarkin et al., 2007); (b) the study was published as an abstract which did not mention randomisation, control groups or between-group effect sizes (Huber & Klug, 2006); and (c) did not use randomisation or quasi-randomisation (Korner et al., 2006). We included four RCTs that were not included by Leichsenring and Rabung (Bressi et al., 2010; Giesen-Bloo et al., 2006; Linehan et al., 2006; McMain et al., 2009). One RCT was published after the last search date of Leichsenring and Rabung (Bressi et al., 2010); two trials that used LTPP as the control intervention were not identified by them (Linehan et al., 2006; McMain et al., 2009); and one trial was excluded because some participants (in both treatment arms) had not finished treatment completely (Giesen-Bloo et al., 2006). We want to stress the importance of including studies that compared LTPP to other intensive specialized psychotherapies, in addition to studies that used less potent comparison treatments.

The contrast between the effect sizes reported in meta-analysed controlled studies vs. meta-analysed uncontrolled studies underscores also the importance of including controlled studies only, and to examine between-group differences instead of within-group differences\pre-post change. Without control, effect sizes of LTPP cannot be interpreted independently of time effects (including ageing), other non-therapy-specific effects and simple regression-to-the-mean for the scores of recruited patients. This is illustrated by a similar, albeit even greater,

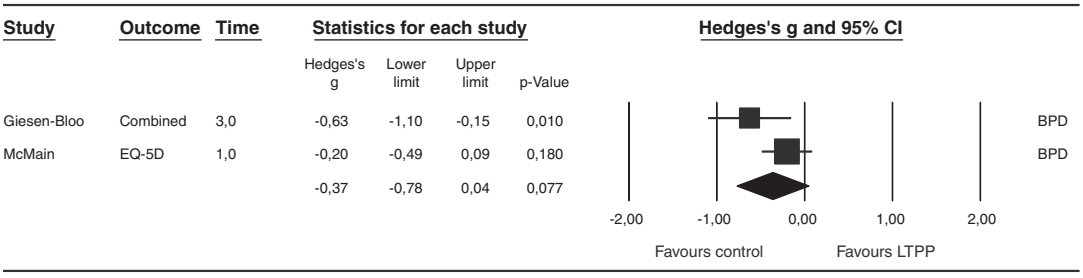

Fig. 8. Meta-analysed Hedges' g for quality of life at the longest available follow-up. Abbreviations: BPD: borderline personality disorder; CI: confidence interval; LTPP: long-term psychoanalytical psychotherapy.

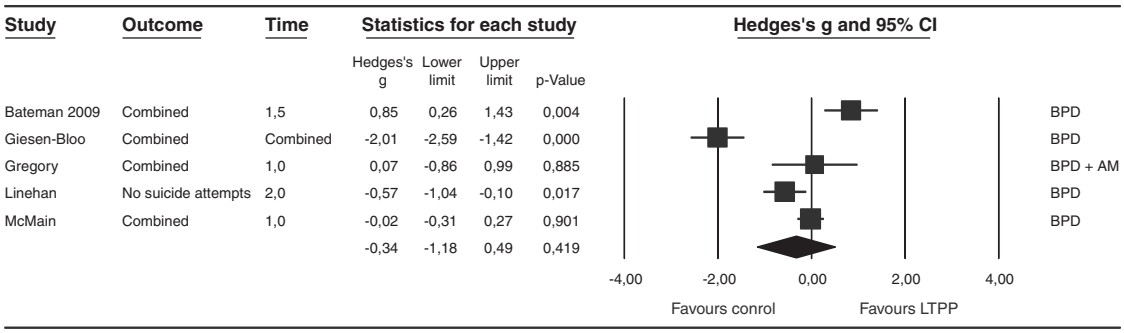

Fig. 9. Meta-analysed Hedges' g for target problems at the longest available follow-up in borderline personality disorder patients. Abbreviations: AM: alcohol misuse: CI: confidence interval; BPD: borderline personality disorder; LTPP: long-term psychoanalytical psychotherapy. The outcome target problems for Giesen-Bloo are a combination of recovery at 4 years (van Asselt et al., 2008) and the Borderline Personality Disorder Severity Index at 3 years.

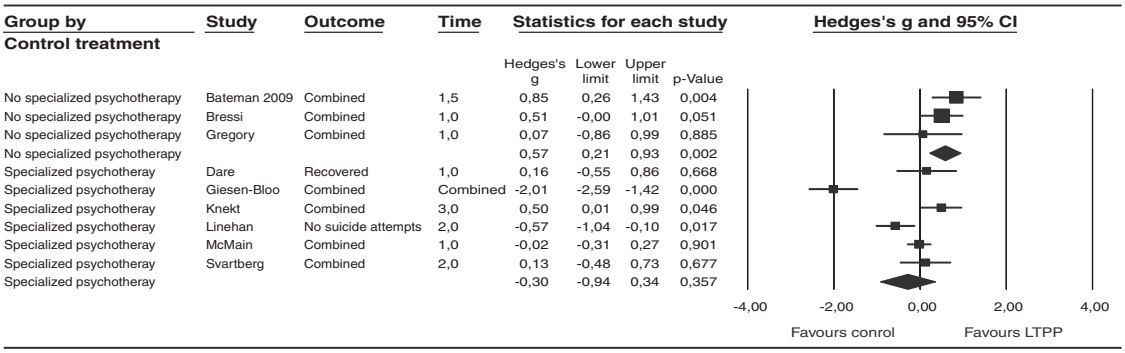

Fig. 10. Meta-analysed Hedges' g for target problems sub grouped by whether LTPP was compared to a specialized psychotherapy or not. Abbreviations: CI: confidence interval; LTPP: long-term psychoanalytical psychotherapy. The outcome target problems for Giesen-Bloo are a combination of recovery at 4 years (van Asselt et al., 2008) and the Borderline Personality Disorder Severity Index at 3 years.

contrast in effect sizes reported in a recent meta-analysis on STPP for depression, in which both within-group and between-group differences were reported. Here, changes in depression level were large in STPP-treated patients, with an effect size of d = 1.34 (pre-treatment to post-treatment changes). However, compared to other psychotherapies, a small negative effect size (d = −0.30) was found directly post-treatment (Driessen et al., 2010). Reported effect sizes can be misleading unless placed in the appropriate context.

5. Conclusions

Acknowledging the caveat that the evidence comes from heterogeneous populations and control treatments, our findings contradict the previously published large effect sizes for LTPP (De Maat et al., 2009; Leichsenring & Rabung, 2008). This shows that the effectiveness of any treatment must be examined by controlled studies. Future studies should compare LTPP to other highly specialised treatments that are equally intensive, like state-of-the art cognitive behaviour therapy in case of eating disorders, or schema-focused therapy or dialectical behaviour therapy in case of borderline personality disorder. Furthermore, studies should focus on populations for which LTPP is frequently used in regular practice and use valid outcomes that allow a meaningful interpretation.

Author contributions

All authors contributed to the conception and design of the study, drafting of the manuscript, and gave final approval of the version to be published. YS and MH performed the analysis of data. YS, MH, and AA were involved in the interpretation of data.

Declaration of interest

No conflicts of interest to be declared.

Acknowledgements

This review was funded in part by the Dutch Health Care Insurance Board (CVZ).

References

Abbass, A. A., Hancock, J. T., Henderson, J., & Kisely, S. (2006). Short-term psychodynamic psychotherapies for common mental disorders. *Cochrane Database of Systematic Reviews*, CD004687.

Bachar, E., Latzer, Y., Kreitler, S., & Berry, E. M. (1999). Empirical comparison of two psychological therapies. Self psychology and cognitive orientation in the treatment of anorexia and bulimia. *The Journal of Psychotherapy Practice and Research*, *8*, 115–128.

Bateman, A., & Fonagy, P. (1999). Effectiveness of partial hospitalization in the treatment of borderline personality disorder: A randomized controlled trial. *The American Journal of Psychiatry*, *156*, 1563–1569.

Bateman, A., & Fonagy, P. (2009). Randomized controlled trial of outpatient mentalization-based treatment versus structured clinical management for borderline personality disorder. *The American Journal of Psychiatry*, *166*, 1355–1364.

Borenstein, M., Hedges, L., Higgins, J. P., & Rothstein, H. (2009). *Introduction to meta-analysis.* Chichester: John Wiley & Sons Ltd..

Bressi, C., Porcellana, M., Marinaccio, P. M., Nocito, E. P., & Magri, L. (2010). Short-term psychodynamic psychotherapy versus treatment as usual for depressive and anxiety disorders: A randomized clinical trial of efficacy. *The Journal of Nervous and Mental Disease*, *198*, 647–652.

Chiesa, M., Fonagy, P., & Holmes, J. (2006). Six-year follow-up of three treatment programs to personality disorder. *Journal of Personality Disorders*, *20*, 493–509.

Clarkin, J. F., Levy, K. N., Lenzenweger, M. F., & Kernberg, O. F. (2007). Evaluating three treatments for borderline personality disorder: A multiwave study. *The American Journal of Psychiatry*, *164*, 922–928.

Cuijpers, P., van, S. A., Bohlmeijer, E., Hollon, S. D., & Andersson, G. (2009). The effects of psychotherapy for adult depression are overestimated: A meta-analysis of study quality and effect size. *Psychological Medicine*, 1–13.

Dare, C., Eisler, I., Russell, G., Treasure, J., & Dodge, L. (2001). Psychological therapies for adults with anorexia nervosa: Randomised controlled trial of out-patient treatments. *The British Journal of Psychiatry*, *178*, 216–221.

De Maat, S., De Jonghe, F., Schoevers, R., & Dekker, J. (2009). The effectiveness of long-term psychoanalytic therapy: A systematic review of empirical studies. *Harvard Review of Psychiatry*, *17*, 1–23.

Driessen, E., Cuijpers, P., de Maat, S. C., Abbass, A. A., de, J. F., & Dekker, J. J. (2010). The efficacy of short-term psychodynamic psychotherapy for depression: A meta-analysis. *Clinical Psychology Review*, *30*, 25–36.

Freud, S. (1961). Analysis terminable and interminable. In J. Strachey (Ed.), *The standard edition of the complete works of Sigmund Freud*. London: Hogarth Press.

Giesen-Bloo, J., & Arntz, A. (2007). In reply. *Archives of General Psychiatry*, *64*, 610–611.

Giesen-Bloo, J., van, D. R., Spinhoven, P., van, T. W., Dirksen, C., van, A. T., et al. (2006). Outpatient psychotherapy for borderline personality disorder: Randomized trial of schema-focused therapy vs transference-focused psychotherapy. *Archives of General Psychiatry*, *63*, 649–658.

Gregory, R. J., Chlebowski, S., Kang, D., Remen, A., & Soderberg, M. (2006). Psychodynamic therapy for borderline personality disorder and co-occurring alcohol use disorders: A newly designed ongoing study. *Journal of the American Psychoanalytic Association*, *54*, 1331–1334.

Gregory, R. J., Chlebowski, S., Kang, D., Remen, A. L., Soderberg, M., Stepkovitz, J., et al. (2008). A controlled trial of psychodynamic psychotherapy for co-occurring borderline personality disorder and alcohol use disorder. *Psychology and Psychotherapy: Theory, Research and Practice*, *45*, 21–48.

Higgins, J. P., & Thompson, S. G. (2002). Quantifying heterogeneity in a meta-analysis. *Statistics in Medicine*, *21*, 1539–1558.

Higgins, J. P., Thompson, S. G., & Spiegelhalter, D. J. (2009). A re-evaluation of random-effects meta-analysis. *Journal of the Royal Statistical Society: Series A: Statistics Society*, *172*, 137–159.

Hoglend, P., Amlo, S., Marble, A., Bogwald, K. P., Sorbye, O., Sjaastad, M. C., et al. (2006). Analysis of the patient-therapist relationship in dynamic psychotherapy: An experimental study of transference interpretations. *The American Journal of Psychiatry*, *163*, 1739–1746.

Hoglend, P., Bogwald, K. P., Amlo, S., Marble, A., Ulberg, R., Sjaastad, M. C., et al. (2008). Transference interpretations in dynamic psychotherapy: Do they really yield sustained effects? *The American Journal of Psychiatry*, *165*, 763–771.

Huber, D., & Klug, G. (2006). Munich Psychotherapy Study (MPS): The effectiveness of psychoanalytic longterm psychotherapy for depression. *Society for Psychotherapy Research, ed. Book of abstracts. From research to practice* (pp. 154). Ulm, Germany: Ulmer Textbank.

Ioannidis, J. P. (2008). Perfect study, poor evidence: Interpretation of biases preceding study design. *Seminars in Hematology*, *45*, 160–166.

Ioannidis, J. P., Patsopoulos, N. A., & Evangelou, E. (2007). Uncertainty in heterogeneity estimates in meta-analyses. *BMJ*, *335*, 914–916.

Klar, F. J. (2005). Effectiveness of indivdual psychological psychoanalytical psychotherapy. *Z Individualpsychol*, *30*, 28–50.

Knekt, P., Lindfors, O., Harkanen, T., Valikoski, M., Virtala, E., Laaksonen, M. A., et al. (2008). 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. *Psychological Medicine*, *38*, 689–703.

Korner, A., Gerull, F., Meares, R., & Stevenson, J. (2006). Borderline personality disorder treated with the conversational model: A replication study. *Comprehensive Psychiatry*, *47*, 406–411.

Lau, J., Ioannidis, J. P., & Schmid, C. H. (1997). Quantitative synthesis in systematic reviews. *Annals of Internal Medicine*, *127*, 820–826.

Lau, J., Ioannidis, J. P., Terrin, N., Schmid, C. H., & Olkin, I. (2006). The case of the misleading funnel plot. *BMJ*, *333*, 597–600.

Leichsenring, F., & Rabung, S. (2008). Effectiveness of long-term psychodynamic psychotherapy: A meta-analysis. *Journal of the American Medical Association*, *300*, 1551–1565.

Leichsenring, F., & Rabung, S. (2011). Long-term psychodynamic psychotherapy in complex mental disorders: Update of a meta-analysis. *The British Journal of Psychiatry*, *199*, 15–22.

Linehan, M. M., Comtois, K. A., Murray, A. M., Brown, M. Z., Gallop, R. J., Heard, H. L., et al. (2006). Two-year randomized controlled trial and follow-up of dialectical behavior therapy vs therapy by experts for suicidal behaviors and borderline personality disorder. *Archives of General Psychiatry*, *63*, 757–766.

McMain, S. F., Links, P. S., Gnam, W. H., Guimond, T., Cardish, R. J., Korman, L., et al. (2009). A randomized trial of dialectical behavior therapy versus general psychiatric management for borderline personality disorder. *The American Journal of Psychiatry*, *166*, 1365–1374.

Mintz, J., O'B'rien, C. P., & Luborsky, L. (1976). Predicting the outcome of psychotherapy for schizophrenics. Relative contributions of patient, therapist, and treatment characteristics. *Archives of General Psychiatry*, *33*, 1183–1186.

Munroe-Blum, H., & Marziali, E. (1995). A controlled trial of short-term group treatment for borderline personality disorder. *Journal of Personality Disorders*, *9*, 190–198.

O'Brien, C. P., Hamm, K. B., Ray, B. A., Pierce, J. F., Luborsky, L., & Mintz, J. (1972). Group vs individual psychotherapy with schizophrenics. A controlled outcome study. *Archives of General Psychiatry*, *27*, 474–478.

Piper, W. E., Debbane, E. G., Bienvenu, J. P., & Garant, J. (1984). A comparative study of four forms of psychotherapy. *Journal of Consulting and Clinical Psychology*, *52*, 268–279.

Svartberg, M., Stiles, T. C., & Seltzer, M. H. (2004). Randomized, controlled trial of the effectiveness of short-term dynamic psychotherapy and cognitive therapy for cluster C personality disorders. *The American Journal of Psychiatry*, *161*, 810–817.

van Asselt, A. D., Dirksen, C. D., Arntz, A., Giesen-Bloo, J. H., van, D. R., Spinhoven, P., et al. (2008). Out-patient psychotherapy for borderline personality disorder: Cost-effectiveness of schema-focused therapy v. transference-focused psychotherapy. *The British Journal of Psychiatry*, *192*, 450–457.

van Tulder, M. W., Assendelft, W. J., Koes, B. W., & Bouter, L. M. (1997). Method guidelines for systematic reviews in the Cochrane Collaboration Back Review Group for Spinal Disorders. *Spine*, *22*, 2323–2330.

Vinnars, B., Barber, J. P., Noren, K., Gallop, R., & Weinryb, R. M. (2005). Manualized supportive-expressive psychotherapy versus nonmanualized community-delivered psychodynamic therapy for patients with personality disorders: Bridging efficacy and effectiveness. *The American Journal of Psychiatry*, *162*, 1933–1940.