College voor zorgverzekeringen

Subject:	**Psychoanalysis is not included among the provisions insured on the grounds of the Zvw, though long-term psychoanalytical psychotherapy is**
Summary:	CVZ has adopted a standpoint on psychoanalysis (PA) and long-term psychoanalytical psychotherapy (LPPT), partly on the basis of a study commissioned by CVZ. The assessment of these two forms of care was carried out according to the statutory criteria: 'normal provision' and 'established medical science and medical practice'. Based on examination according to these two statutory criteria, CVZ has reached the conclusion that PA does not belong among the provisions insured on the grounds of the Zvw, and LPPT does. This means that treatment by psychoanalysis is no longer included in the basic package. New treatments may not be initiated at the expense of the Zvw. Current psychoanalyses can be continued and completed at the expense of the Zvw.
Type of ruling:	SpZ = Zvw standpoint
Date:	23rd March 2010
Care form:	Medical GGZ

The full ruling appears below.

Standpoint **Psychoanalysis and long-term psychoanalytical psychotherapy**

Summary

Since 2008 mental health care is regulated within the framework of the *Zorgverzekeringswet* (Zvw, Health Insurance Act). Transferring it from the *Algemene Wet Bijzondere Ziektekosten* (AWBZ, Exceptional Medical Expenses Act) took place with a minimum of package and budgetary effects. 'Psychoanalysis' was placed on the package agenda partly as a result of that transfer. After all, the AWBZ system is different from that of the Zvw.

Study CVZ commissioned a literature study into evidence of the efficacy of psychoanalysis (PA) and long-term psychoanalytic psychotherapy (LPPT). The study, which has two parts, one over psychoanalysis and one over long-term psychoanalytic psychotherapy, was completed at the beginning of this year.
This report contains the standpoint adopted by CVZ – based partly on the study results – over psychoanalysis and long-term psychoanalytic psychotherapy.
Assessment took place according to the two statutory criteria: 'normal provision' and 'established medical

science and medical practice'.

Normal provision

PA and LPPT are forms of care normally provided by medical specialists and clinical psychologists. Such care is a medical provision under the Zvw as long as it fulfils the criterion established medical science and medical practice.

Established medical science and medical practice

PA does not comply with established medical science and medical practice because insufficient qualitatively adequate studies on its efficacy in practice can be found.

LPPT does fulfil the established medical science and medical practice criterion because sufficient studies have been found which prove that its effects do not differ from those of other treatments.

Conclusion

On the grounds of an examination of these two statutory criteria, CVZ has concluded that psychoanalysis should not be regarded as an insured provision on the grounds of the Zvw and long-term psychoanalytic psychotherapy should.

Consequences for current psychoanalyses

This means that treatment with psychoanalysis is no longer included in the basic package.
New treatments may not be initiated at the expense of the Zvw. Current treatments can be continued and completed at the expense of the Zvw.

Introduction

Reason for placement on package agenda and transfer

Psychoanalysis was included on CVZ's package agenda for 2007-2008[1] due to signals received. The subject was included in the cluster 'long-term care', whereby the main question involved how the subject was regulated under the AWBZ and the Zvw (content, nature and amount) and how this care can be defined in the future. An important role in this was the transfer of mental health care (GGZ) from the AWBZ to the Zvw as of 1st January 2008.

It appears that psychoanalysis is a treatment for a small target group, involving relatively high costs, the efficacy of which is not clear[2]. Similarly to psychoanalysis, long-term psychoanalytic psychotherapy is also a highly intensive treatment within mental health care (GGZ). Both treatments take longer than one year and involve four to five sessions per week for psychoanalysis and one to two sessions per week for psychoanalytic psychotherapy.

[1] Rapport Pakketagenda 2007-2008, 2007, zie website CVZ
[2] Gezondheidsraad, Doelmatigheid van long-term psychotherapy, 2001

Study into efficacy

CVZ commissioned a literature study into evidence of the efficacy of psychoanalysis and long-term psychoanalytic psychotherapy.
This study was carried out by a study group under the supervision of Ms. Y. Smit, MD, with an MSc in Epidemiology. This study group guarantees both content and methodological expertise. The question posed by the study was: How effective are psychoanalysis and long-term psychoanalytic psychotherapy, where necessary specified according to the psychiatric disorder?
The study, which has two parts, one over long-term psychoanalytic psychotherapy (LPPT) and one over psychoanalysis (PA), was completed at the beginning of 2010.

In this report CVZ has adopted a standpoint, based partly on the study results, over psychoanalysis and long-term psychoanalytic psychotherapy.

Report composition

The report takes the following form. In section 2 CVZ discusses the historic and current position of these care-forms within the package. Section 3 discusses the care-forms, psychoanalysis (PA) and long-term psychoanalytic psychotherapy (LPPT). Section 4 contains the legal and medical assessment.
CVZ subsequently discusses its standpoint over the question of whether PA and LPPT should be classed as insured care (section 5).
The consequences of this standpoint are discussed in section 6.
In section 7 CVZ discusses comments they received on a draft of this standpoint.
Lastly, section 8 indicates the date on which the standpoint was adopted.

Context

Transfer

On 1st January 2008 GGZ was transferred from the AWBZ to the Zvw. The point of departure during that transfer was that package and budgetary effects should be kept to a minimum. This means that all GGZ treatment that up till then had been financed by virtue of the AWBZ was transferred to the Zvw.
Nevertheless, the system of the AWBZ differs from that of the Zvw. Within the framework of the AWBZ, care-providers are defined and the Minister is able to stipulate conditions. Under the Zvw, the stipulation of procedural conditions is left to health insurers and health insurers determine who can provide care.

Arrangement under the AWBZ in 2007

Under the AWBZ, treatment relating to a psychiatric disorder could take place via an institution, a psychiatrist, a neurologist or a psychotherapist. In

principle, psychotherapeutic treatment was limited to a maximum of 50 sessions, though this limit did not apply to psychoanalytic treatment provided via an institution. Furthermore, the AWBZ stipulated that the health insurer must grant prior permission and that the indication was determined and the care provided in accordance with the provisions in the 'Indication and treatment protocol for adults' of the *Nederlands Psychoanalytisch Instituut* (Npi, Dutch Psychoanalytical Institute) in Amsterdam.[3]

In 2004 the right to all psychotherapeutic treatments was limited to a maximum number of sessions. Following discussions with the Lower House, this was changed back for psychoanalytic treatment, though the relevant extra conditions were accepted, partly because the efficacy of psychoanalysis had not been convincingly demonstrated.[4]

Transfer to the Zvw

As mentioned above, the stipulation of such conditions is not in keeping with the Zvw system. Health insurers can indicate who is allowed to provide a given form of care and also stipulate detailed conditions.
The subject 'Psychoanalysis' was placed on the package agenda with a view to the planned transfer and the different system within the Zvw.

Care-forms LPPT and PA

CVZ regards ambulant long-term psychotherapies as including:

- long-term psychoanalytic psychotherapy (LPPT), with a frequency of one to two sessions per week, and
- classic psychoanalysis (PA), with a frequency of four to five sessions per week.

CVZ has commissioned research into the efficacy of these care-forms. Within the framework of the Zvw, an intervention can be included in the insured package if it is care that is "normally provided" and if this care complies with "established medical science and medical practice".
Due to the fact that the Zvw-legislation has become applicable, CVZ felt that these care-forms should be examined according to the Zvw-criteria. This report is the result of that examination.

Psychoanalysis and long-term psychoanalytic psychotherapy

This section discusses psychoanalysis and psychoanalytic psychotherapy. The discussion covers not only the theory and the points of departure involved but also the evidence found in the scientific literature.

[3] Artikel 8 Besluit zorgaanspraken AWBZ en artikel 7 Regeling zorgaanspraken AWBZ, zoals deze luidden vóór 1 januari 2008
[4] Stb. 2004, 257

(Long-term) psychotherapy

Description of (long-term) Psychotherapy

Psychotherapy is treatment by an expert, who has received the relevant training, with the aim of healing or improving mental disorders by affecting symptoms methodically and bringing about changes in the perception and cognition of patients and the way in which they function.
Long-term psychotherapy is defined as a form of psychotherapy lasting longer than one year or involving more than 25 sessions.[5]

Two forms

According to this terminology, long-term psychotherapy encompasses both psychoanalytic face-to-face therapy-forms (1 to 2 sessions per week) and psychoanalysis (on the couch, 4 to 5 sessions per week). The first form can take place both individually and in a group, for adults the second is – by definition – individual.
Psychoanalysis takes a special place within long-term psychotherapy, partly due to its exceptionally long duration. It can vary from 4 years to more than 10 years, with 4 or 5 sessions per week. Though it is shorter and less intensive (a minimum of one year, with 1 or 2 sessions per week), long-term psychoanalytic psychotherapy is based on the same points of departure as psychoanalysis. This report focuses on these two forms of long-term psychotherapy which are based on psychoanalytic theory, i.e., on the one hand psychoanalysis, and on the other hand long-term psychoanalytic psychotherapy.

Psychoanalytic theory

Theory and Freud

Freud developed the points of departure for psychoanalytic theory, as an explanation of a great deal of human behaviour, at the start of the 20th century. The development of the theory and the development of classic psychoanalysis as a treatment method went hand-in-hand. These subsequently underwent animated development, which has led to an adjustment in the original ideas and concepts and to the development of different versions of the theory and its application. This also led to a great deal of controversy. In any case, the classic 'drive' model described by Freud has largely been abandoned.
This report does not contain a detailed and coherent discussion of the concepts, theories and controversies. By way of a summary, the following can be stated about the psychoanalytic theory and working method.

Common points of departure

According to current beliefs (Gabbard 2004, cited by De Maat[6]), there are a number of points of departure common to all variations in the psychoanalytic theory and the working method:

[5] Gezondheidsraad, Doelmatigheid van long-term psychotherapy, 2001
[6] De Maat, On the effectiveness of psychoanalytic therapy, 2007

- Much of our emotional life is subconscious.
- Experiences from one's childhood, together with aptitude, form the adult person.
- Transference, from the patient to the therapist, is a primary source of insight.
- Reverse transference, from the therapist, provides valuable insight into what a patient evokes in others.
- The patient's resistance to the therapeutic process is an important point for attention in the therapy.
- Symptoms and behaviour serve a multitude of goals and are determined by complex and often subconscious forces.
- A psychodynamic therapist supports the patient in achieving the perception of authenticity and uniqueness.

These principles of the general theory (which is broader than the mere clinical application) were taken from De Wolf[7] en Schalkwijk[8]: early childhood development in relation to others is the fundament for the personality, an important part of our life is enacted subconsciously, and it can be regarded as a theory of emotions that makes statements on the cohesion of feelings and cognition.

No 'proof of concept'

One of the problems of the psychoanalytic theory and treatment is the lack of solid "proof of concept". Its application as a therapy is not currently based on irrefutable proof of its working mechanism. Cultural phenomena and phenomena that crop up in therapy suggest a degree of plausibility for the therapy, but that does not provide sufficient arguments at to why it should be effective in the individual treatment of patients with certain disorders. Van Tilburg[9] paid attention to the need of empirical research into key concepts of the theory. But the most fundamental assumption of the psychoanalytic theory, i.e., that the subconscious is essential to understanding behaviour, is no longer open to discussion.

Application as a method of treatment

Classic psychoanalysis

The most radical application of psychoanalytic therapy is in the form of classical psychoanalysis (PA), whereby the patient lies on the therapist's couch 4 or 5 times per week and the therapist, who listens and responds, sits somewhere beyond the patient's field of vision. The process focuses in particular on revealing subconscious aspects of dysfunctional behaviour or dysfunctional perceptions. This process can take many

[7] De Wolf, Inleiding in de psychoanalytic psychotherapy, 2002
[8] Schalkwijk, Dit is psychoanalyse, 2006
[9] Van Tilburg, Psychoanalyse en psychiatrie, Tijdschrift voor Psychiatrie 42 (2000) 9

years. PA is not suited to everyone, as it demands a lot of patients with respect to the investment of time, their capacity for self-reflection and their ability to express themselves.

Long-term psychoanalytic psychotherapy

Another form of therapy that is based on psychoanalytic theory, is long-term psychoanalytic psychotherapy (LPPT). This involves face-to-face contact in a normal consulting room situation during 1 to 2 years. The chosen aims may be limited to the treatment of a number of specific complaints or problem areas. The therapist and patient are at liberty to choose for either an open or a more supportive approach. This means that patients who are less suited to classical PA on the couch can also be accepted for treatment.

For whom is it intended?

The target group cannot easily be delineated in the form of properly defined diagnostic categories as in the DSM-IV[10]. All the more because the DSM-IV is particularly based on visible behaviour and conscious complaints, whilst the psychoanalytic approach focuses more on the hidden causes of a multitude of psychopathological phenomena. In general, the indications for psychotherapy are not determined primarily by symptoms, syndromes or a disorder, but rather by personal characteristics that enable patients to profit from a psychoanalytic approach[11]. Subject to these conditions, LPPT is applied to a multitude of complaints and problems. These are defined by the authoritative *Nederlands Psychoanalytisch instituut* (Npi, Dutch Psychological Institute) as follows: severe suffering over a longer period of time due to psychiatric problems such as problems relating to personality and identity or recurrent anxieties and depression. These are patients who cannot be helped sufficiently with a short-term approach[12].

Need for efficacy research

No efficacy study

In 2001 the *Gezondheidsraad* (GR, Health Council of the Netherlands) had already concluded that long-term psychotherapy (i.e., including PA and LPPT) had been insufficiently studied to be able to draw conclusions relating to its efficacy.
Only sporadic research had been carried out into long-term psychotherapy. Various factors had played a role in this.
It is impossible to assess the efficacy, let alone the cost-effectiveness of long-term psychotherapy without more research data. The GR committee felt that the therapy may be effective for a certain group of patients. Examples are patients with personality disorders and

[10] DSM-IV-TR. Diagnostic and Statistical Manual of Mental Disorders. Washington: American Psychiatric Association, 2000
[11] Zwanikken e.a., Psychiatrie, 1993
[12] Indicatie- en behandelingsprotocol voor volwassenen, 2004

patients with persistent mood disorders.

In order to be able to assess the current place of PA and LPPT in the basic package, CVZ commissioned a study into evidence of their efficacy. How evidence-based are these treatments? The decision was made to investigate PA and LPPT separately. This resulted in, respectively, a review and a meta-analyse by Y. Smit et al., which are included as appendices to this standpoint determination.

Assessment of the Zvw standpoint

Laws and legislation

Statutory framework

The statutory framework within which CVZ investigated long-term psychoanalytic psychotherapy and psychoanalysis is formed by the laws and legislation as laid down in the Zvw and related regulations.

Risk to be insured

- Article 10 of the Zvw stipulates which risks must be included in health insurance.
  Article 10, under a of the Zvw, stipulates that the risk insured by virtue of health insurance includes, among other things, medical care.

Amvb: Bzv

- Article 11, third paragraph of the Zvw, indicates that the content and amount of the insured provisions are regulated in more detail in a governmental decree (AMvB). This decree took the form of the Health Insurance Decision (Bzv).

- Article 2.1, first paragraph of the Bzv rules that the forms of care or services include the care and other services referred to in article 11, first paragraph, under a, of the Zvw, the content and amount of which are defined in articles 2.4 up to and including 2.15 of the Bzv.

Established medical science and medical practice

- Article 2.1, second paragraph of the Bzv rules that the content and amount of the forms of care or services are determined in part by established medical science and medical practice and, where such a standard is lacking, by what is regarded in the professional field concerned as responsible and adequate care and services.

Normally provided

- Article 2.4, first paragraph, opening lines and under a, 2 of the Bzv, defines that medical care as care that is normally provided by G.P.s, medical specialists, clinical psychologists, (…), as well as (…).

Within the framework of these statutory stipulations,

CVZ assesses whether PA and LPPT can be regarded as:

- medical care as normally provided by medical specialists and clinical psychologists (paragraph 3.b.),
- care that complies with established medical science and medical practice, or where such a standard is lacking, with care that is regarded in the professional field concerned as responsible and adequate care (paragraph 3.c.).

Normal provision

CVZ establishes whether the criterion 'normal provision' is fulfilled in accordance with the report 'The meaning of the normal provision criterion and its assessment' dated 17th November 2008[13].

Accepted arsenal of care tools and providing can in a professionally appropriate way

CVZ determines whether care is involved which belongs to the accepted arsenal of care tools of the professional groups referred to in the legislation (psychiatrists and clinical psychologists) and whether the care is being provided in the manner regarded as professionally correct by these professional groups.

Psychoanalysis

During the past century psychoanalysis has been applied to a variable degree on the basis of continually evolving and sometimes diverging theoretical models, in cases with more or less clearly defined complaints, problems and mental disorders.
In 2004 the number of adults in analysis in Dutch GGZ-institutions was 229; about 400 were visiting independent analysts[14]. In 2004 the Npi provided the indication for 18 of the 602 patients. We carried out a systematic search for psychoanalysis in the Dutch GGZ-guidelines. It turned out that psychoanalysis had not been given a place in the guidelines for ADHD, alcohol-related problems, anxiety disorders, eating disorders, depression, personality disorders and schizophrenia. This does not mean that the care-providers are not providing this care. The fact is that psychoanalysis actually focuses more on the type of patient and his/her life-problems than on specific disorder entities or on specific diagnostic categories. Guidelines usually take diagnostic categories as point of departure. See also under 3c, where the target group is defined.

There is a lack of clear standpoints from foreign sources on psychoanalysis. In 2002 Blue Cross/Blue Shield of Massachusetts removed psychoanalysis from the list of rights to provisions, after having included it in 2000[15]. Blue Cross/Blue Shield North Carolina adhere to an

[13] Publicatienummer 268: zie website CVZ
[14] Schalkwijk, Dit is psychoanalyse, 2006
[15] BluecrossBlueshield Massachusetts, behaviour health policy nr 423, revised jan. 2010
[16] BluecrossBlueshield North Carolina, evidence based guideline MHCD2041, revised aug. 2008

"evidence-based guideline" dating from 2008[16], which permits room for psychoanalysis under indication conditions similar to those of the Npi.

Long-term psychoanalytic psychotherapy

There are fewer limitations for long-term psychoanalytic psychotherapy than for psychoanalysis when determining the indication. In 2004 702 patients were treated in Dutch GGZ-institutions as a result of indications determined by the Npi and a lot more were probably treated by independent care-providers.

Special trend within psychotherapy

Care-providers who actually use psychoanalytic methods are almost all members of the professional groups of psychiatrists and clinical psychologists.
A large proportion of the psychiatrists and clinical psychologists use psychoanalytic methods in their work. As many as 250 therapists affiliated with the Dutch association for Psychoanalysis, just as many psychologists as doctors. Although ideas differ in medical and psychological circles regarding the efficacy and cost-effectiveness of psychoanalysis, properly trained professionals feel that the theory and its application were correctly assessed from a professional point of view. The specialised professional group has formulated specific training requirements for providing PA and LPPT[17]. These apply in addition to the training requirements laid down in the BIG Act [Individual Health Care Professions Act]. Within the framework of controlled access to long-term psychoanalytic forms of treatment, the indication and treatment protocol developed by the Dutch Psychoanalytic Institute has been the national standard since 2004.
For many years this care has been reimbursed at the expense of the statutory health insurance in the Netherlands.

Normal provision

PA and LPPT are forms of care normally provided by the said professional groups.

Established medical science and medical practice

EBM-principles

In assessing whether a provision complies with established medical science and medical practice, CVZ adheres to the principles of Evidence-Based Medicine (EBM).
In their report, 'Assessment of established medical science and medical practice', dated 5th November 2007[18], CVZ describes how they check whether care complies with this criterion.

CVZ's point of departure for replying to the question of

[17] Zie Kenniscentrum van het NPI, www.psychoanalytischinstituut.nl: drie Nederlandse beroepsverenigingen verzorgen een opleiding tot psychoanalyticus (Nederlandse Vereniging voor Psychoanalyse, Nederlands Psychoanalytisch Genootschap en Nederlandse Psychoanalytic Groep); voor informatie over de opleiding tot psychoanalytisch psychotherapeut de Nederlandse Vereniging voor Psychoanalytic psychotherapy
[18] Publicatienummer 254: zie website CVZ

whether care complies with established medical science and medical practice is that good quality randomised studies are required in order to be able to draw an unequivocally positive conclusion about interventions. If there are no such studies, then a positive assessment can follow on the basis of studies with a lower level of evidence. However, in that case, there must be thorough grounds demonstrating why there are no randomised studies and why they cannot be demanded.

Study

In order to assess whether PA and LPPT are care forms that comply with established medical science and medical practice, CVZ commissioned a study into the cost-effectiveness of PA and LPPT.
The study has two parts, one into the cost-effectiveness of PA and one into that of LPPT. These studies are enclosed as appendices to this report.
The following is a discussion of the results of the study.

Systematic review of psychoanalysis

No evidence of cost-effectiveness

For an elaborate description and justification of the search strategy and the inclusion criteria, please see the review[19] a well as the meta-analysis of LPPT[20].
The search relating to LPPT also focused on psychoanalysis. As no controlled, (quasi)randomised trials on psychoanalysis were found, the authors also searched for comparative cohort studies. They examined the references of existing systematic reviews and searched in Medline, Embase, PsycINFO, and also searched for all Evidence-Based Medicine (EBM) Reviews. Studies were included if one of the treatment groups was in psychoanalysis, and if that group could be compared with a comparable group that was undergoing a different treatment or receiving no treatment. Patients had to have a definable mental disorder. Two researchers selected the studies independently of one another. They did not find any comparative cohort studies that were suitable for inclusion. Such studies are necessary, as is randomised research in which psychoanalysis is compared with evidence-based treatments. Researchers must clear a variety of methodological hurdles if they want to assess the cost-effectiveness of psychoanalysis. Double-blind research is not possible, and it is difficult to include, randomise and subsequently retain patients in a very long study.
The existence of various comparative studies into the cost-effectiveness of long-term psychoanalytic psychotherapy suggest that such studies must also be possible for psychoanalysis.
For the cost-effectiveness of PA, therefore, the researchers were unable to find any studies that fulfil the quality criteria of Evidence-Based Medicine (EBM).

[19] Bijlage 1
[20] Bijlage 2

Meta-analysis of long-term psychoanalytic psychotherapy

Y. Smit et al. searched through the references of existing systematic reviews and they searched in Medline, Embase, PsycINFO, and all Evidence-Based Medicine (EBM) Reviews – without any time or language restrictions – for randomised trials and quasi-randomised trials.
The intervention had to be psychoanalysis or long-term psychoanalytic psychotherapy. The control group had to be receiving a non-psychoanalytic treatment, and/or a short-term treatment, or no treatment (e.g., still on a waiting list). Patients had to have a clearly defined mental disorder. Two researchers selected the studies independently of one another.
The primary outcome was recovery; secondary outcomes were outcomes in the field of *target problems, general psychiatric symptoms, personality pathology, social functioning* and quality of life. Selected studies were assessed for quality, and summarised in tables and texts. A weighted average was calculated (meta-analysis) for the primary and secondary outcomes. In the end eight randomised, controlled studies were included which compared long-term psychoanalytic psychotherapy (LPPT) with a different treatment.
No controlled, randomised studies were found involving psychoanalysis, nor any comparative studies involving a lower level of evidence.
The eight studies included were carried out on extremely different groups of patients and involved a comparison of LPPT with a range of other therapies, including *treatment as usual*. The eight studies included adult patients with various eating disorders, bi-polar disorders and personality disorders. For details see the meta-analysis[21].

No difference of efficacy with other treatments

The frequency of recovery from the various mental disorders after LPPT did not differ from the frequency of recovery after various other treatments.
The individual studies did differ considerably with regard to Effect Sizes (ES), both in direction and in size. This may be the result of differences in patient groups, in control treatments, in measuring instruments, etc.
The number of studies was too small to examine the differences by means of sub-group analyses or meta-regression.
The eloquence of the findings is limited due to

[21] Bijlage 2

[22] Ook de analyse van de GR (2001) wijst op de beperkingen van ES als uitkomst: *In psychotherapy-onderzoek wordt om de vraag naar de grootte van het effect te beantwoorden veelal de effectgrootte (effect size) berekend, waarmee bedoeld wordt (M1 - M2) / s.d. Daarin zijn M1 en M2 de gemiddelden van uitkomstmaten van de behandelde respectievelijk onbehandelde groep patiënten en is s.d. de gemiddelde standaardafwijking. Voor de s.d. wordt eventueel een gewogen gemiddelde genomen, en in sommige gevallen de s.d. bij de behandelde groep. De s.d. is mede afhankelijk van het aantal en mogelijke selectie van de patiënten. Bij een grotere s.d. verkrijgt men een kleinere effectgrootte, en omgekeerd. De effectgrootte is derhalve een relatieve maat, die bruikbaar is om verschillende onderzoeken te vergelijken, maar niet om de absolute grootte van het effect te beoordelen.*
De effectgrootte geeft echter niet aan of de behandeling voor een bepaalde patiënt een verbetering oplevert die klinische betekenis heeft. Klinische significantie heeft betrekking op de praktische waarde van een interventie voor een individuele patiënt.

methodological limitations and the limited importance of the final outcome effect sizes (ES). See also GR (2001)[22].
In other words, the meta-analysis was incapable of demonstrating that the effect of LPPT differs from that of a different approach.

Conclusion

Psychoanalysis Psychoanalysis does not comply with established medical science and medical practice because insufficient qualitatively adequate studies can be found on its cost-effectiveness in practice. In their search for evidence, the authors of the review sank to the level of comparative cohort studies. Evidence of an even lower level than that of the review could not justifiably lead to the conclusion that PA is effective.
In other words: no evidence was found that, with respect to individual treatment, patients are sufficiently better off with psychoanalysis than without it.

Explanation If comparative studies cannot – or should not – be expected, evidence of a lower level is sometimes sufficient to make cost-effectiveness plausible. For example, this is the case if there is experimental proof of a working mechanism. However, this is not the case for psychoanalysis.

Long-term psychoanalytic psychotherapy Long-term psychoanalytic psychotherapy does comply with established medical science and medical practice because sufficient studies were found showing that the effects do not differ from other treatments.

Standpoint

The current question is whether PA and LPPT are forms of care that belong in the insured package.
PA and LPPT are care that is normally provided by medical specialists and clinical psychologists. This care falls under the Zvw medical care provision in as far it fulfils the criterion of established medical science and medical practice.

No sufficiently qualitative studies can be found on the cost-effectiveness of psychoanalysis in practice. This means that psychoanalysis does not comply with established medical science and medical practice.

Sufficient studies have been found showing that the effects of long-term psychoanalytic psychotherapy do not differ from those of other treatments. This means that LPPT does comply with established medical science and medical practice.

On the grounds of the above, CVZ concludes that PA does not belong to the insured provisions provided on the grounds of the Zvw and LPPT does.

Consequences

Consequences for insured clients

The standpoint has no consequences for insured persons receiving LPPT, but it does for those who are currently receiving PA. CVZ discusses these current treatments below.

Insurers can no longer charge the costs of 'classic' psychoanalysis to the Zvw. As psychotherapy is a specialist care-form, it is financed under the DBC-system. In the future, health insurers should bear in mind that ambulant products for treatment groups which involve a lot of time could involve psychoanalysis.

Current psychoanalysis

The implication of the standpoint is that treatment with psychoanalysis, as defined in this report, is no longer included in the basic package. No new treatments can be started at the expense of the Zvw. Failing to comply with established medical science and medical practice automatically means that the treatment is not (/no longer) included in the package.
The possibility exists that interrupting the treatment and the treatment relationship could have detrimental consequences for the insured client.

In view of the nature of health insurance, CVZ feels that health insurers should be able to continue the reimbursement of current treatment.
Health insurance is a compensatory form of insurance (article 1, under d of the Zvw). This means that the date on which the costs were incurred is decisive for the question of whether treatment is an insured form of care.
The costs are incurred at the moment at which the insured person incurs the costs of medical treatment or at the moment that medical treatment is given. The entire treatment should be regarded as a single assessable incident.
For insured persons who are currently undergoing psychoanalytic treatment, this means that the expense was incurred when psychoanalysis was still an insured care-form and that the treatment costs should be reimbursed up to the moment that treatment is completed. Health insurers should reimburse this treatment until it has been completed and they can also charge the costs to risk adjustment. This is a direct result of the nature of health insurance.

Comments from the parties

Different procedure

Within the framework of package management, the products supplied by CVZ are standpoints and advice. In reaching their advice (package or system advice), CVZ generally consults the parties involved as interested

parties. Where standpoints are involved, within the framework of determining the paragraph on established medical science and medical practice, the rule is that experts on the matter are consulted. As mentioned in the introduction, this standpoint is a result of the package agenda in 2007-2008. At that moment it not clear whether a standpoint or advice would be issued. Furthermore, various parties had already expressed a desire to be involved. CVZ agreed to approach the parties as soon as a draft report had been prepared. This means that, by way of a departure from the usual procedure, CVZ presented a draft of this report, for comments, to:

- *GGZ Nederland*
- Dutch Association of Independent Psychologists & Psychotherapists (NVVP)
- Dutch Association for Psychotherapy (NVP)
- Dutch Institute of Psychologists (NIP)
- Dutch Psychoanalytic Group (NPAG)
- Client platform for psychoanalysis and psychoanalytic psychotherapy (CPPP)
- National Platform GGz (LPGGz)
- Dutch Association for Psychiatry (NVvP)
- Dutch Psychoanalytic Institute (Npi)
- Dutch Psychoanalytic Society (NPG)
- Dutch Association for Psychoanalytic Psychotherapy (NVPP)
- Dutch Association for Psychoanalysis (NVPA)
- *Zorgverzekeraars Nederland* (ZN)

Reactions

A response was received from all parties, except for the National Platform GGz. The replies received are enclosed as appendix 3. The three as yet unpublished articles to which various parties refer are not enclosed.
CVZ will reply to each response individually.
Within the framework of this report, CVZ responds to the most important comments below. These comments were made by most of the parties. That does not apply to ZN's response. CVZ discusses that separately at the end of this reply.

Why not accept a lower level of evidence for psychoanalysis?
The parties pointed out that, due to the nature of the treatment, psychoanalysis is not – or poorly – suited to scientific research by means of an RCT. For this reason that feel that CVZ should have based their assessment on lower evidence. They also refer to the report 'Established medical science and medical practice', which cites examples where decisions can be made based on lower evidence.
In response to this, CVZ comments that the text clearly indicates that the evidence requirements were lower than an RCT, e.g., comparative cohort studies. If comparative studies cannot or should not be demanded,

then evidence of a lower level can sometimes be sufficient to suggest the plausibility of cost-effectiveness. For example, this could be the case if the evidence is based on experimental proof of a working mechanism. However, this does not exist for psychoanalysis.

For the rest, CVZ does not conclude that psychoanalysis is ineffective, but that there is insufficient evidence for its efficacy.

Working mechanism
The parties point out that two criteria apply for replying to the question of whether an intervention complies with established medical science and medical practice. The first criterion they cite is that 'sufficient qualitatively adequate studies on efficacy in practice' can be found and the second is 'a proven working mechanism'. This was not our intention. There is only one criterion, and that is the first. The fact that 'proven working mechanism' was wrongly interpreted as a criterion has led to clarification and an adjustment in the text.

Provisional finance
The parties argue for retaining PA in the basic package and for realising sufficient qualitative research in the near future.
In this connection, CVZ cites the recently published report 'Conditional finance within the framework of a responsible package'[23]. If the Minister accepts this advice, then there may be possibilities for finance. CVZ suggests six criteria for permitting interventions to become eligible for the possibility of conditional finance.

ZN's response
ZN states that this standpoint provides clarity and can easily be implemented in practice.

With regard to the comments on the conclusion on LPPT, CVZ comments that much of the available material is subject to methodological limitations.
The researchers took this into account in their analysis, and CVZ has taken it into account in establishing their standpoint.

Appropriateness is not an aspect of an assessment of established medical science and medical practice and it was therefore not discussed in this report.

Adoption of the standpoint

The Board of Directors approved this standpoint on

[23] Rapport Voorwaardelijke financiering in het kader van een verantwoord pakket van 1 december 2009. Publicatienummer 283. Zie website CVZ.

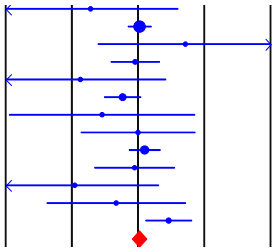
23rd March 2010.

The effectiveness of psychoanalysis - a systematic review of the literature

Final report

January 21, 2010

Yolba Smit MD
Marcus Huibers PhD
John Ioannidis MD PhD
Richard van Dyck MD PhD
Willem van Tilburg MD PhD
Arnoud Arntz PhD



Contact details:
Yolba Smit
Grimmweg 6
47559 Kranenburg
Germany
+49 (0)2826 917262
yolbasmit@yahoo.com

1 Contents

2 Summary

2.a Introduction

For several decades now the effectiveness of psychoanalysis is debated. While the effectiveness of other forms of psychotherapy - such as cognitive behaviour therapy, interpersonal psychotherapy and short-term psycho-analytical psychotherapy - has been scrutinised in controlled trials, research that focuses on psychoanalysis is sparse. We set out to determine the effectiveness of psychoanalysis. In a previous systematic review we did not find any randomised controlled trials or controlled trials comparing psychoanalysis with another treatment or no treatment. Here we search for comparative cohort studies on psychoanalysis, also referred to as studies with a quasi-experimental, quasi-randomised or naturalistic design.

2.b Methods

We tracked references from existing systematic reviews and searched Medline, Embase, PsycINFO, and all Evidence Based Medicine (EBM) Reviews for quasi-randomised studies or controlled cohort studies. Studies would be selected if one of the groups studied were patients in psychoanalysis, and if this group could be compared to a similar group of patients in another type of treatment or not in treatment. Patients had to have a clearly defined mental disorder. Two authors independently identified trials for inclusion.

2.c Results

We were unable to identify any quasi-randomised studies or cohort studies suitable for inclusion in this review.

2.d Discussion

Conductors of studies on psychoanalysis will face several methodological problems. Blinding is not possible, and to include, randomise and retain patients in a study that extents over several years is difficult. The existence of several randomised studies on long-term psychoanalytical therapy shows that such studies can be done, however.

2.e Conclusions

There are no quasi-randomised studies or controlled cohorts comparing psychoanalysis with any other treatment or no treatment in patients with mental disorders. Such studies are required, as are randomised trials comparing psychoanalysis to evidence-based treatments in selected patients.

3 Samenvatting

3.a Introductie

De werkzaamheid van psychoanalyse is al enkele decennia het onderwerp van discussie. Terwijl andere vormen van psychotherapie – zoals cognitieve gedragstherapie, interpersoonlijke psychotherapie en (kortdurende) psychoanalytische psychotherapie – in gecontroleerd onderzoek zijn onderzocht, is onderzoek naar psychoanalyse zeldzaam. Wij onderzochten de werkzaamheid van psychoanalyse. In een eerdere systematische review vonden we geen (gerandomiseerd en) gecontroleerd onderzoek dat psychoanalyse vergeleek met een andere behandeling of met geen behandeling. In deze studie zoeken we naar vergelijkende cohort studies, ook wel quasi-experimentele/-gerandomiseerde studies of naturalistische studies genoemd.

3.b Methodes

We doorzochten de referenties van bestaande systematische reviews en zochten in Medline, Embase, PsycINFO, en alle Evidence Based Medicine (EBM) Reviews naar quasi-experimentele/-gerandomiseerde studies of gecontroleerde cohort studies. Studies zouden geïncludeerd worden als één van de behandelgroepen in psychoanalyse zou zijn, en als die groep vergeleken kon worden met een vergelijkbare groep die een andere behandeling volgde, of niet in behandeling was. Patiënten moesten een definieerbare mentale aandoening hebben. Twee onderzoekers hebben onafhankelijk van elkaar de studies geselecteerd.

3.c Resultaten

We hebben geen quasi-experimentele/-gerandomiseerde studies of gecontroleerde cohort studies kunnen vinden die geschikt waren om te includeren.

3.d Discussie

Er zijn verschillende methodologische hindernissen te nemen voor onderzoekers die de effectiviteit van psychoanalyse willen evalueren. Dubbelblind onderzoek is niet mogelijk, en het includeren, randomiseren en vervolgens vasthouden van patiënten in zeer lang lopend onderzoek is moeilijk. Toch laten verschillende gerandomiseerde studies naar de effectiviteit van langdurige psychoanalytische psychotherapie zien dat het mogelijk is.

3.e Conclusie

Er zijn geen quasi-experimentele/-gerandomiseerde studies of gecontroleerde cohort studies naar de effectiviteit van psychoanalyse. Dergelijke studies zijn nodig, evenals gerandomiseerd onderzoek waarin psychoanalyse met evidence-based behandelingen wordt vergeleken.

4 Introduction

4.a Background

For several decades now the effectiveness of psychoanalysis and long-term psychoanalytic psychotherapy (LTPP) is debated. While the effectiveness of other forms of psychotherapy - such as cognitive behaviour therapy, interpersonal psychotherapy and (short-term) psychoanalytical psychotherapy - has been scrutinised in controlled trials, research that focuses on psychoanalysis and LTPP is sparse.

The Dutch Health Care Insurance Board (CVZ) has funded this meta-analysis to answer the following research questions:

- Is psychoanalysis an effective treatment for mental illness? If so, for which patients or illnesses?
- Is long-term psychoanalytic psychotherapy an effective treatment for mental illness? If so, for which patients or illnesses?

Our work on LTPP is described in a separate report: *The effectiveness of long-term psychoanalytical psychotherapy - a meta-analysis*. Here we searched for studies on psychoanalysis as well. We selected controlled studies only. This included randomised controlled trials and controlled trials where treatment allocation was not at random. However, allocation did have to demonstrate a certain degree of freedom from bias. For example, a study with allocation based on clinical judgement or patient preference would be excluded, while a study with allocation based on time of referral would be included. We did not find any studies on psychoanalysis that met these standards. We then decided to look at comparative cohort studies on psychoanalysis, also referred to as studies with a quasi-experimental or naturalistic design. The present report describes our work and findings.

Definitions (De Maat, 2007;Glass, 2008)and Dutch societies of psychoanalytical professionals:
Cognitive behaviour therapy: a usually short-term psychotherapy focussed on identifying and correcting cognitive patterns that underlie emotional and behaviour symptoms;

Schema-focused therapy and *dialectical behaviour therapy*: focused therapies developed for treatment of borderline personality disorder, usually long-term therapies that combine cognitive, behavioural, interpersonal and experiential techniques;

Interpersonal psychotherapy: time-limited individual therapy developed for the treatment of major depression;

Long-term: consisting of at least 40 sessions and lasting at least one year;

Long-term psychotherapy: any form of psychotherapy that is long-term;

Long-term psychoanalytic psychotherapy (LTPP): therapy rooted in psychoanalytic theories, `A therapy that involves careful attention to the therapist-patient interaction, with carefully timed interpretation of transference and resistance embedded in sophisticated appreciation of the therapists´ contribution to the two-person field(Gunderson, 1999)´. In addition: based on psychoanalytic theory (and not, for instance, on client-centred, interpersonal, or schema-therapy theories). Therapy sessions usually have a frequency of once or twice a week;

Psychoanalysis: as long-term psychoanalytic psychotherapy, with four or more therapeutic sessions weekly. In addition, the patient lies on a couch.

5 Methods

5.a Criteria for considering studies for this review

5.a.1 Types of studies

We wanted to select controlled cohort studies only. One of the cohorts had to be a cohort of patients in psychoanalysis; the other cohort could be patients in a different type of treatment or on a waiting list. Treatment allocation did not need to be at random and could be based on clinical judgement or patient preference. Only studies published in peer-reviewed publications were considered.

5.a.2 Types of participants

Participants had to be adults with a clearly defined mental disorder. Consequently, populations with a primarily behavioural problem were excluded if concurrent mental disorders were not present or not described. E.g. a study with participants with a history of aggression or spouse beating would be excluded if the participants were not diagnosed with a defined mental disorder. In addition, studies with participants with a pure somatic disorder (diabetes, morbus Crohn, irritable bowel disease) were excluded. Studies on patients with schizophrenia were excluded because schizophrenia is very distinct from the mental illnesses for which psychoanalysis is prescribed in the Netherlands at present (Berghout, 2008).

5.a.3 Types of interventions and controls

The intervention had to be psychoanalysis with four or more weekly sessions and the patient lying on a couch. To be considered for inclusion a study should intend to investigate a treatment of at least 160 sessions and continue for at least a year. The mean number of sessions per participant could be less than 160 sessions however, if dropouts or no-shows were included in the calculation of the mean. The control treatment could be any type of treatment or no treatment or waiting list.

5.a.4 Types of outcome measures

We considered the primary outcome the recovery rate, and secondary outcomes were measures of the target problem, of general psychiatric symptoms, personality pathology, social functioning and quality of life (QoL).

5.b Search strategies

We tracked references from existing reviews (De Maat, 2009;Leichsenring, 2008), also looking at studies that were excluded by those reviews, and searched electronic databases from June 2008 onwards. This is the date up to which a recent review by Leichsenring had searched the literature (Leichsenring, 2008). Medline, Embase, PsycINFO, and all Evidence Based Medicine (EBM) Reviews were searched through OVID®. The EBM Reviews are (a) the ACP Journal Club; (b) the Cochrane Database of Systematic Reviews; (c) The Cochrane Central Register of Controlled Trials; (d) the Cochrane Methodology Register; (e) the Database of Abstracts of Reviews of Effects (DARE); (f) the National Health Services Health Technology Assessment Database (NHS HTA); and (g) the NHS Economic Evaluation Database (NHS EED).

We searched using the text word psychoanalysis. The search was restricted to Dutch, English, French and German articles. Experts from our team checked our findings for completeness.

5.c Data collection and analysis

5.c.1 Selection of studies

Two reviewers (MH and YS) independently selected suitable studies for inclusion as detailed below. Where the two reviewers disagreed about the inclusion of a study, disagreements were resolved by consensus of opinion. A third reviewer was consulted if a disagreement could not be resolved. Where resolution was not possible, the author was contacted to obtain more information and clarification. The titles and abstracts of studies identified by searching electronic databases were assessed to determine if an article was eligible. An article was rejected when the title and abstract contained sufficient information to determine that it did not meet the inclusion criteria. The full papers of all remaining articles were retrieved.

5.c.2 Assessment of methodological quality of included studies

We planned to assess methodological quality by two independent researchers (MH, YS). When the researchers disagreed, consensus was to be reached through discussion. We intended to use eight criteria proposed by Cuijpers et al (Cuijpers, 2009). Detailed information on these criteria can be found in section 9.b in the Annexes. In brief, the criteria used by Cuijpers et al were based on an authoritative review of empirically supported psychotherapies (Chambless, 1998), and on the criteria proposed by the Cochrane Collaboration to assess the methodological validity of a study (Anon. 2009). The criteria based on the review of empirically supported psychotherapies assessed the quality of the treatment delivery, while the criteria proposed by the Cochrane Collaboration assessed more methodological sources of bias.

5.c.3 Data extraction and management

We planned to abstract data by one researcher (YS) and checked by a second researcher (MH). Outcomes assessed by independent assessors and intention to treat (ITT) data would be selected, if available.

5.c.4 Data synthesis

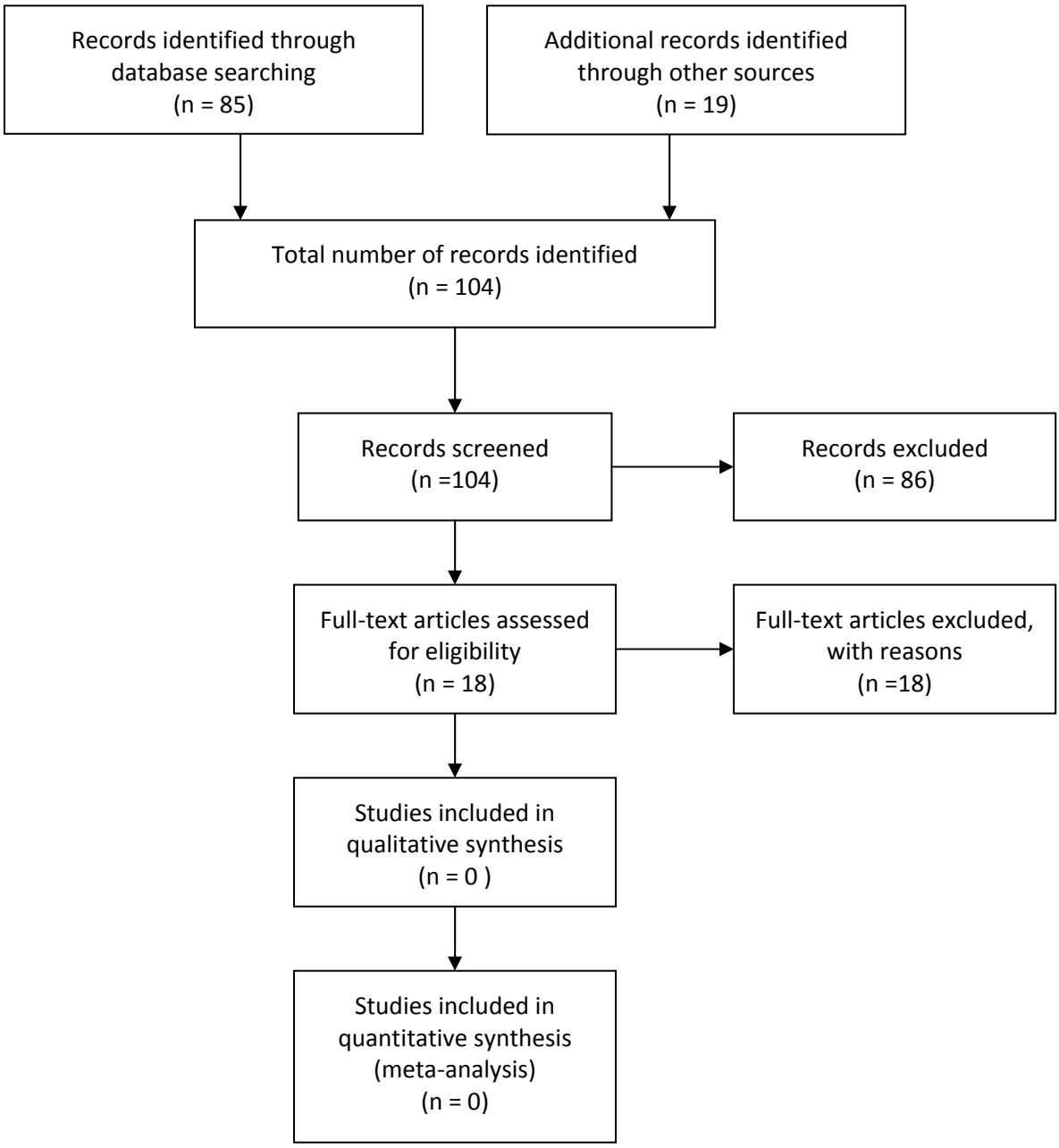
We intended to describe selected studies in tables and text.

6 Results

104 studies were retrieved through database searching and reference tracking. 86 studies were excluded based on the title and/or abstract, because they were:

- Retrospective studies
- Studies with only one group in therapy
- Case studies or case series
- Surveys among analysts or students
- Concerned with long-term psychoanalytical psychotherapy
- Dream content analysis studies
- Studies on theoretical aspects or the concept of psychoanalysis

Figure 1 Flow chart of study search and selection



18 studies were retrieved full text and all were excluded. The two articles by Berghout et al. were found in the conducted electronic databases search from June 2008 onwards (Berghout, 2008;Berghout, 2009); all other articles were found through reference tracking of existing reviews. Most articles were excluded because outcomes were not useful or not presented per treatment type. Some articles did not include a cohort of patients that received psychoanalysis. Detailed reasons for excluding articles are given in Table 1.

Table 1 Studies excluded on full text review

	Reason for exclusion	
1	Narrative review	(Bachrach, 1991)
2	Examines which patients are clinical cases	(Berghout, 2008)
3	Outcomes are not presented per treatment type	(Berghout, 2009)
4	Examines patient-by-treatment interaction	(Blatt, 2004) Menninger Psychotherapy Research Project
5	Not a true cohort study: the observations from different patients are combined, thus the pre-treatment data come from different patients as the post-treatment data do	(Blomberg, 2001;Sandell, 2000) STOPP (Stockholm outcome of psychotherapy and psychoanalysis study)
6	No valid outcomes concerned with recovery, target problems, symptoms or social functioning. Therapists rate a ´clinical global impression' and a ´health-sickness rating scale´, which reflect the therapists´ opinion of success	(Burstein, 1972) Menninger Psychotherapy Research Project
7	Compares patients in psychoanalysis, psychodynamic psychotherapy and group therapy only at the start of therapy (between-group comparison). Then compares hospital admissions before the start of therapy with hospital admissions after therapy ended (within-group comparison)	(Duhrssen, 1986)
8	Retrospective survey	(Dührssen, 1998)
9	Compares a cohort in psychoanalytical psychotherapy to a cohort in psychodynamic psychotherapy	(Grande, 2006) Heidelberg-Berlin study
10	Compares the views of patients, therapists and research judges as to the degree of which treatment goals had been reached	(Harty, 1976) Menninger Psychotherapy Research Project
11	No valid outcomes concerned with recovery, target problems, symptoms or social functioning	(Heuft, 1996;Kordy, 1983;Kordy, 1988;von, 1998) Heidelberg Catamnesis Project
12	Not a peer-reviewed publication (book chapter). Presenting preliminary results only (half a year after treatment start) of a trial of psychoanalysis vs. psychodynamic psychotherapy	(Huber, 2001)
13	Only patients in psychoanalysis	(Leichsenring, 2005)
14	Retrospective sample. The outcomes are all either patients satisfaction or therapists´ satisfaction with the results of treatment. Numerical data are rarely given (outcomes are mainly presented as figures) and the group in psychoanalysis cannot be separated from the group in psychodynamic psychotherapy	(Leuzinger-Bohleber, 2001)
15	Compares a cohort in psychoanalytical psychotherapy to a cohort in psychodynamic psychotherapy.	(Puschner, 2004;Puschner, 2007)
16	Compares a cohort in psychoanalytical psychotherapy to a cohort in psychodynamic psychotherapy	(Rudolf, 1994;Rudolf, 1999)
17	Narrative review	(Schulman, 1990)
18	Outcomes are not presented per treatment type	(Vaughan, 2000)

7 Discussion

Conductors of studies on psychoanalysis will face several methodological problems. Blinding is not possible, and to include, randomise and retain patients in a study that extents over several years is difficult. However, the existence of several randomised studies on long-term psychoanalytical therapy shows that such studies can be done (Bachar, 1999;Bateman, 1999;Dare, 2001;Giesen-Bloo, 2006;Gregory, 2008;Knekt, 2008;Linehan, 2006;Svartberg, 2004).

Beyond (randomised) controlled trials are the quasi-randomised studies or controlled cohort studies we searched for here. Even if we would have found such studies, precise statements on the differential effects of therapies would not have been easy to make. The groups involved might not be comparable, or when groups seemed similar there was no way of ruling out the possibility that natural group formation might have been influenced by factors producing significant outcome differences.

Even further afield are the observational studies where no comparisons with other groups whatsoever are made. It is difficult to examine the effectiveness of any treatment in uncontrolled conditions. Absolute change (before treatment *vs.* after treatment) cannot be interpreted independently from time effects (including aging) and so-called non-specific effects like attention, empathy, expectations, explanations for problems etcetera. Thus, to study the effectiveness of psychoanalysis between-group comparisons are needed. Within-group change is insufficiently precise. Studies on within-group change can play a useful role in preparing researchers to put the right questions forward to be answered in controlled studies. Difficult to conduct as they are in the field of long-tem psychotherapy, controlled studies should be drafted as carefully as possible.

8 Conclusions

We didn't find quasi-randomised studies or controlled cohort studies comparing psychoanalysis with any other treatment or no treatment in patients with mental disorders. Such studies are needed, as are randomised trials comparing psychoanalysis to evidence-based treatments.

9 Annexes

9.a Detailed searches

9.a.1 Medline (in OVID® June 1, 2008 to November 25, 2009)

1. exp psychoanalysis/ (7281)
2. limit 1 to (humans and yr="2008 -Current" and (Dutch or English or French or German) and ("therapy (sensitivity)" or "therapy (specificity)" or "therapy (optimized)" or "costs (sensitivity)" or "costs (specificity)" or "costs (optimized)" or "economics (sensitivity)" or "economics (specificity)" or "economics (optimized)")) (4)

9.a.2 Embase (in OVID® 2008 to November 26, 2009)

1. exp psychoanalysis/ (16209)
2. limit 1 to (human and ("treatment (1 term high sensitivity)" or "treatment (1 term high specificity)" or "treatment (1 term min difference)" or "treatment (2 or more terms high sensitivity)" or "treatment (2 or more terms high specificity)" or "treatment (2 or more terms min difference)") and (Dutch or English or French or German) and yr="2008 -Current" and journal and (adult <18 to 64 years> or aged <65+ years>)) (36)
3. from 2 keep 1-36 (36)

9.a.3 PsycINFO (in OVID® 2008 to November 26, 2009)

1. exp Psychoanalysis/ (40063)
2. limit 1 to (peer reviewed journal and ("treatment (high sensitivity)" or "treatment (high specificity)" or "treatment (min difference)") and ("0430 followup study" or "0450 longitudinal study" or "0451 prospective study" or "0600 field study" or 1800 quantitative study or "2000 treatment outcome/randomized clinical trial") and adulthood <18+ years> and "300 adulthood <age 18 yrs and older>" and "0110 peer-reviewed journal" and journal article and (Dutch or English or French or German) and yr="2008 -Current") (45)
3. from 2 keep 1-45 (45)

9.a.4 OVID® All Evidence Based Medicines Reviews

This database includes:

- DARE (from 1991 onwards)
- NHS EED (from 1995 onwards)
- HTA NHS CRD (from 2001 onwards)
- CMR (Cochrane Methodology Register, from 1995 onwards)
- CCTR (Cochrane Central Register of Controlled Trials, from 1991 onwards)
- Cochrane Database of Systematic Reviews
- ACP Journal Club (from 1991 onwards)

Search:

1. exp psychoanalysis/ (7)
2. limit 1 to yr="2008 -Current" [Limit not valid in DARE; records were retained] (1)
3. from 2 keep 1 (1)

9.b Quality criteria

9.b.1 Quality criteria used by Cuijpers et al

According to the criteria used by Cuijpers et al (Cuijpers, 2009), a study was considered to be of high quality when:

(a) Participants met diagnostic criteria for a mental disorder (as assessed with a personal diagnostic interview and using a diagnostic system such as DSM
(b) The study referred to the use of a treatment manual (either a published manual, or a manual specifically designed for the study)
(c) The therapists who conducted the therapy were trained for the specific therapy, either specifically for that study or as a general training
(d) Treatment integrity was checked during the study (by supervision of the therapists during treatment or by recording of treatment sessions or by systematic screening of protocol adherence by a standardized measurement instrument)
(e) Data were analysed with intention-to-treat analyses, in which all persons who were randomized to the treatment and control conditions initially were included in the analyses
(f) The study had a minimal level of statistical power to find significant effects of the treatment, and included $\geq$50 persons in the comparison between treatment and control groups. This allows the study to find standardized effect sizes of **d**=0.80 and larger, assuming a statistical power of 0.80 and α=0.05. Calculations in Stata (Stata Corp., USA)
(g) The study reported that randomization was conducted by an independent (third) party (this variable was positive if an independent person did the randomization, when a computer program was used to assign patients to conditions, or when sealed envelopes were used)
(h) Assessors of outcome were blinded and did not know to which condition the respondents were assigned to (this was only coded when the effect sizes were based on interviewer-based depression ratings ; when only self-reports were used, it was assumed that this criterion was met)

If a study did not report whether it met the quality criterion is was coded as negative.

10 References

Cochrane handbook for systematic reviews of interventions. Chichester: John Wiley & Sons Ltd; 2009.

Bachar E, Latzer Y, Kreitler S, Berry EM. Empirical comparison of two psychological therapies. Self psychology and cognitive orientation in the treatment of anorexia and bulimia. J Psychother Pract Res 1999;8(2):115-28.

Bachrach HM, Galatzer-Levy R, Skolnikoff A, Waldron S Jr. On the efficacy of psychoanalysis. J Am Psychoanal Assoc 1991;39(4):871-916.

Bateman A, Fonagy P. Effectiveness of partial hospitalization in the treatment of borderline personality disorder: a randomized controlled trial. Am J Psychiatry 1999 Oct;156(10):1563-9.

Berghout CC, Zevalkink J. Identifying clinical cases among patients assigned to psychoanalytic treatment. Bull Menninger Clin 2008;72(3):163-78.

Berghout CC, Berghout CC. Clinical significance of long-term psychoanalytic treatment. [References]. Bulletin of the Menninger Clinic 2009;.73(1).

Blatt SJ, Shahar G. Stability of the patient-by-treatment interaction in the Menninger Psychotherapy Research Project. Bull Menninger Clin 2004;68(1):23-38.

Blomberg J, Lazar A, Sandell R. Long-term outcome of long-term psychoanalytically oriented therapies: first findings of the Stockholm outcome of psychotherapy and psychoanalysis study. Psychother Research 2001;11(4):361-82.

Burstein ED, Coyne L, Kernberg OF, Voth H. Psychotherapy and psychoanalysis. I. The quantitative study of the psychotherapy research project: psychotherapy outcome. Bull Menninger Clin 1972 Jan;36(1):1-85.

Chambless DL, Hollon SD. Defining empirically supported therapies. J Consult Clin Psychol 1998 Feb;66(1):7-18.

Cuijpers P, van SA, Bohlmeijer E, Hollon SD, Andersson G. The effects of psychotherapy for adult depression are overestimated: a meta-analysis of study quality and effect size. Psychol Med 2009 Jun 3;1-13.

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 Mar;178:216-21.

De Maat S, De Jonghe F, Schoevers R, Dekker J. The effectiveness of long-term psychoanalytic therapy: a systematic review of empirical studies (dissertation). Amsterdam: Vrije Universiteit; 2007.

De Maat S, De Jonghe F, Schoevers R, Dekker J. The effectiveness of long-term psychoanalytic therapy: a systematic review of empirical studies. Harv Rev Psychiatry 2009;17(1):1-23.

Duhrssen A. [A comparison of dynamic psychotherapy, psychoanalysis and analytic group psychotherapy]. Z Psychosom Med Psychoanal 1986;32(2):161-80.

Dührssen A, Jorswieck E. Eine empirisch-statistische Untersuchung zur Leistungsfähigkeit psychoanalytischer Behandlung. Z Psychosom Med 1998;44:311-8.

Giesen-Bloo J, van DR, Spinhoven P, van TW, Dirksen C, van AT, et al. Outpatient psychotherapy for borderline personality disorder: randomized trial of schema-focused therapy vs transference-focused psychotherapy. Arch Gen Psychiatry 2006 Jun;63(6):649-58.

Glass RM. Psychodynamic psychotherapy and research evidence: Bambi survives Godzilla? JAMA 2008 Oct 1;300(13):1587-9.

Grande T, Dilg R, Jakobsen T, Keller W, Krawietz B, LANGER M, et al. Differential effects of two forms of psychoanalytic therapy: results of the Heidelberg-Berlin study. Psychother Research 2006;16(4):470-85.

Gregory RJ, Chlebowski S, Kang D, Remen AL, Soderberg M, Stepkovitz J, et al. A controlled trial of psychodynamic psychotherapy for co-occurring borderline personality disorder and alcohol use disorder. Psychother Theory Res Practice Training 2008;45(1):21-48.

Gunderson JG, Gabbard GO. Making the case for psychoanalytic therapies in the current psychiatric environment. J Am Psychoanal Assoc 1999;47(3):679-704.

Harty M, Horwitz L. Therapeutic outcome as rated by patients, therapists, and judges. Arch Gen Psychiatry 1976 Aug;33(8):957-61.

Heuft G, Seibüchler-Engeç H, Tschke M, Senf W. [Long-term outcome of outpatient psychoanalytic psychotherapy and psychoanalysis. Analysis of 53 catamnestic interviews]. Forum Psychoanal 1996;12:342-55.

Huber D, Klug G, Von Rad M. [The Munich processoutcome study: a comparison between psychoanalyses and psychotherapy]. In: Stuhr BM, editor. *Langzeit- Psychotherapie. Perspektiven für Therapeuten und Wissenschaftler*.Stuttgart, Germany: Kohlhammer; 2001. p. 260-70.

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 May;38(5):689-703.

Kordy H, von RM, Senf W. Success and failure in psychotherapy: hypotheses and results from the Heidelberg Follow-Up Project. Psychother Psychosom 1983;40(1-4):211-27.

Kordy H, von RM, Senf W. Time and its relevance for a successful psychotherapy. Psychother Psychosom 1988;49(3-4):212-22.

Leichsenring F, Biskup J, Kreische R, Staats H. The Gottingen study of psychoanalytic therapy: first results. Int J Psychoanal 2005 Apr;86(Pt 2):433-55.

Leichsenring F, Rabung S. Effectiveness of long-term psychodynamic psychotherapy: a meta-analysis. JAMA 2008 Oct 1;300(13):1551-65.

Leuzinger-Bohleber M, Stuhr U, Rüger B, Beutel ME. Langzeitwirkungen von Psychoanalysen und Psychotherapien: eine multiperspektivische, repräsentative Katamnesestudie. Psyche (Stuttg) 2001;55(3):193-276.

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 Jul;63(7):757-66.

Puschner B, Kraft S, Bauer S. Interpersonal problems and outcome in outpatient psychotherapy: findings from a long-term longitudinal study in Germany. J Pers Assess 2004 Dec;83(3):223-34.

Puschner B, Kraft S, Kachele H, Kordy H. Course of improvement over 2 years in psychoanalytic and psychodynamic outpatient psychotherapy. Psychol Psychother 2007 Mar;80(Pt 1):51-68.

Rudolf G, Manz R, Ori C. [Results of psychoanalytic therapy]. Z Psychosom Med Psychoanal 1994;40(1):25-40.

Rudolf G, Manz R, ri C. Outcomes of psychoanalytic therapies. Int Forum Psychoanal 1999;8:125-34.

Sandell R, Blomberg J, Lazar A, Carlsson J, Broberg J, Schubert J. Varieties of long-term outcome among patients in psychoanalysis and long-term psychotherapy. A review of findings in the Stockholm Outcome of Psychoanalysis and Psychotherapy Project (STOPP). Int J Psychoanal 2000 Oct;81 (Pt 5):921-42.

Schulman DG. The investigation of psychoanalytic theory by means of the experimental method. Int J Psychoanal 1990;71:487-98.

Svartberg M, Stiles TC, 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 May;161(5):810-7.

Vaughan SC, Marshall RD, MacKinnon RA, Vaughan R, Mellman L, Roose SP. Can we do psychoanalytic outcome research? A feasibility study. Int J Psychoanal 2000 Jun;81 (Pt 3):513-27.

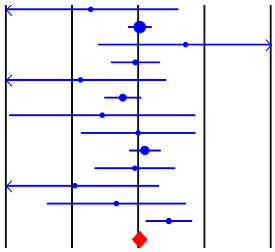
von RM, Senf W, Brautigam W. [Psychotherapy and psychoanalysis in patient management: results of the Heidelberg Catamnesis Project]. Psychother Psychosom Med Psychol 1998 Mar;48(3-4):88-100.

The effectiveness of long-term psychoanalytical psychotherapy - a meta-analysis

Final report

January 21, 2010

Yolba Smit MD
Marcus Huibers PhD
John Ioannidis MD PhD
Richard van Dyck MD PhD
Willem van Tilburg MD PhD
Arnoud Arntz PhD



Contact details:
Yolba Smit
Grimmweg 6
47559 Kranenburg
Germany
+49 (0)2826 917262
yolbasmit@yahoo.com

1 Contents

2 List of tables

3 List of figures

4 Summary

4.a Introduction

For several decades now the effectiveness of psychoanalysis and long-term psychoanalytic psychotherapy (LTPP) is debated. While the effectiveness of other, mostly short-term, forms of psychotherapy - such as cognitive behaviour therapy and interpersonal psychotherapy - has been scrutinised in controlled trials, controlled research that focuses on psychoanalysis and LTPP is sparse. This report describes the effectiveness of LTPP in a systematic review. In contrast to existing reviews, that have used non-controlled data (i.e. within-group differences), we wanted to include controlled data only (i.e. between-group differences).

4.b Methods

We tracked references from existing systematic reviews, and searched Medline, Embase, PsycINFO, and all Evidence Based Medicine (EBM) Reviews without time or language restrictions for randomised controlled trials or controlled trials, published in peer reviewed journals. Studies were selected if patients had a clearly defined mental disorder, and the intervention had to be psychoanalysis or LTPP. The control treatment had to be a non-psychoanalytically based treatment and/or a short-term treatment or no treatment (e.g. waiting list). Two authors independently identified trials for inclusion. The primary outcome was recovery; secondary outcomes were target problems, general psychiatric symptoms, personality pathology, social functioning and quality of life. Selected studies were judged on their methodological quality. Selected studies were described in tables and in text. We performed a meta-analysis of the primary and secondary outcomes.

4.c Results

We included eight controlled trials on LTPP in our meta-analysis (controlled studies on psychoanalysis were not available). These trials compared LTPP in diverse patients to various control treatments, including treatment as usual (TAU). The main meta-analysed results are presented in Table 1. There was substantial to large heterogeneity for most outcomes. The small number of studies precluded any meaningful subgroup analysis. An exploratory meta-regression indicated that there may be a relation between the difference in treatment intensity between the intervention and control group (session ratio) and effect size.

Table 1 Main meta-analysed outcomes across studies (longest available follow-up)

Outcomes	Hedges´ g	95% CI	p-value	n	I-squared
Recovery	0.02	-0.40 to 0.43	0.94	5	52.8%
Target problems	-0.30	-1.05 to 0.46	0.44	6	89.4%
General psychiatric symptoms	0.84	-0.65 to 2.32	0.27	5	97.0%
Personality pathology	*	*	*	1	*
Social functioning	0.11	-0.34 to 0.56	0.63	2	0.0%
Quality of life	*	*	*	1	*
Overall effectiveness	0.29	-0.73 to 1.32	0.58	7	95.4%

* Too few studies to perform analyses. Abbreviations: CI: confidence interval; n: number of studies included in the analysis

4.d Discussion and conclusion

The recovery rate of various mental disorders was equal after LTPP or various control treatments, including TAU and non-evidence based control treatments. The effect sizes of the individual RCTs 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 the small number of

studies precluded a meaningful analysis of subgroups, and severely limited meta-regression. Thus, we consider the meta-regressions we performed exploratory only.

The contrast between the combined effect size for recovery in our meta-analysis (0.02), and the combined effect sizes for overall effectiveness reported in previous meta-analyses by Leichsenring (1.8) and De Maat (0.78) underscores the importance of using a control treatment. Without control treatment, effect sizes are a mixture of time effects (of importance in mental diseases with high remission rates), non-specific effects (such as attention, explanations for symptoms) and the treatment itself.

5 Samenvatting

5.a Introductie

De werkzaamheid van langdurige psychoanalytische psychotherapie en psychoanalyse is al enkele decennia het onderwerp van discussie. Terwijl andere, meestal kortdurende, vormen van psychotherapie – zoals cognitieve gedragstherapie en interpersoonlijke psychotherapie – in gecontroleerd onderzoek zijn onderzocht, is gecontroleerd onderzoek naar langdurige psychoanalytische psychotherapie en psychoanalyse zeldzaam. In dit systematische literatuuronderzoek onderzoeken we de werkzaamheid van psychoanalyse en langdurige psychoanalytische psychotherapie. In tegenstelling tot eerder literatuuronderzoek, waar niet-gecontroleerd onderzoek wordt gebruikt, wilden wij alleen gecontroleerd onderzoek gebruiken.

5.b Methodes

We doorzochten de referenties van bestaande systematische reviews en zochten in Medline, Embase, PsycINFO, en alle Evidence Based Medicine (EBM) Reviews, zonder tijds- of taalrestrictie, naar gerandomiseerde trials en niet-gerandomiseerde trials. De interventie moest psychoanalyse of langdurige psychoanalytische psychotherapie zijn. De controle groep moest een niet-psychoanalytische behandeling krijgen, en/of een kortdurende behandeling, of geen behandeling (bijvoorbeeld op een wachtlijst staan). Patiënten moesten een duidelijk gedefinieerde mentale aandoening hebben. Twee onderzoekers hebben onafhankelijk van elkaar de studies geselecteerd. De primaire uitkomst was herstel; secundaire uitkomsten waren uitkomsten op de domeinen *target problems, general psychiatric symptoms, personality pathology, social functioning* en kwaliteit van leven. Geselecteerde studies werden beoordeeld op kwaliteit, en in tabellen en tekst samengevat. Van de primaire en secundaire uitkomsten werd een gewogen gemiddelde berekend (meta-analyse).

5.c Resultaten

We hebben acht gerandomiseerde, gecontroleerde studies geïncludeerd die lange-termijn psychoanalytische psychotherapie (LTPP) vergeleken met een andere behandeling. (Er zijn geen gecontroleerde, gerandomiseerde studies gevonden over psychoanalyse.) De 8 geïncludeerde studies werden bij zeer verschillende patiëntengroepen uitgevoerd en vergeleken LTPP met een scala aan andere therapieën, inclusief *treatment as usual*. Voor de belangrijkste uitkomst – herstel – was de effectgrootte (*effect size*, ES) 0. Voor de andere uitkomsten waren de ESs variabel in grootte en in richting, maar nooit significant (Tabel 1). Voor de meeste uitkomsten was er sprake van een matige tot grote heterogeniteit. Door het kleine aantal studies was een analyse van subgroepen niet mogelijk. Een verkennende meta-regressie gaf aan dat er mogelijk een verband is tussen het verschil in intensiteit van de behandeling (tussen de interventiegroep en de controlegroep) en de effectgrootte.

Tabel 1 Samenvatting van de belangrijkste uitkomstmaten van de meta-analyse

Uitkomst	Hedges´ g	95% BI	p-waarde	n	I-squared
Herstel	0,02	-0,40 - 0,43	0,94	5	52,8%
Target problems	-0,30	-1,05 - 0,46	0,44	6	89,4%
General psychiatric symptoms	0,84	-0,65 - 2,32	0,27	5	97,0%
Personality pathology	*	*	*	1	*
Social functioning	0,11	-0,34 - 0,56	0,63	2	0,0%
Kwaliteit van leven	*	*	*	1	*
Overall effectiveness	0,29	-0,73 - 1,32	0,58	7	95,4%

* Te weinig studies beschikbaar. Afkortingen: BI: betrouwbaarheidsinterval; n: aantal studies in de analyse.

5.d Discussie en conclusie

De frequentie van herstel van verschillende mentale aandoeningen na LTPP verschilde niet van de frequentie van herstel na diverse andere behandelingen. De individuele studies waren onderling erg variabel wat betreft de ES, zowel in richting als in grootte. Het is mogelijk dat verschillen in patiëntengroepen, in controlebehandelingen, in meetinstrumenten etc. hieraan ten grondslag liggen. Het aantal studies was te klein om met analyses van subgroepen of met meta-regressie de verschillen te onderzoeken.

De ES voor *overall effectiveness* die wij vonden (0.29; 95% CI: -0.73 to 1.32; p=0.58) verschilt van de ESs die door Leichsenring (1.8; 95% CI: 0.7-3.4) en de Maat (0.78; 95% CI niet beschreven; standaard deviatie: 0.45) gerapporteerd werden voor *overall effectiveness*. Dit laat zien dat een ES binnen een groep gemeten, iets heel anders is dan een ES die tussen twee groepen wordt gemeten. Wij zijn er sterk van overtuigd dat alleen de laatste aanpak het effect van therapie kan laten zien. Overigens is het maar de vraag in hoeverre de uitkomst *overall effectiveness* een valide maat is voor het effect van therapie. In deze uitkomstmaat worden alle gevonden uitkomsten in een studie op één hoop gegooid, zonder wegingsfactor. Dat betekent bijvoorbeeld dat het herstel van een eetstoornis even zwaar wordt meegeteld als de score op een vragenlijst naar algemene psychische klachten of de score op een vragenlijst naar persoonlijkheid.

Het contrast tussen de gecombineerde ES voor herstel (0,02) die wij vonden en de gecombineerde ESs voor *overalll effectiveness* die in eerdere meta-analyses door Leichsenring (1,8) en De Maat (0,78) werden gerapporteerd onderstreept het belang van een controlebehandeling. Zonder controlebehandeling zijn ESs een mengeling van het effect van behandeling, het effect van verstrijken van tijd (belangrijk bij mentale aandoeningen waar hoge remissie voorkomt), en van zogenoemde niet-specifieke effecten, zoals aandacht of het verkrijgen van verklaringen voor symptomen of gebeurtenissen.

6 Introduction

6.a Background

For several decades now the effectiveness of psychoanalysis and long-term psychoanalytic psychotherapy (LTPP) is debated. 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 controlled trials, controlled research that focuses on psychoanalysis and LTPP is sparse.

The Dutch Health Care Insurance Board (CVZ) has funded this meta-analysis to answer the following research questions:

- Is psychoanalysis an effective treatment for mental illness? If so, for which patients or illnesses?
- Is long-term psychoanalytic psychotherapy an effective treatment for mental illness? If so, for which patients or illnesses?

First, we assessed the quality of two recent meta-analyses by De Maat and by Leichsenring (De Maat, 2007b;De Maat, 2009;Leichsenring, 2008). We were specifically interested whether we could build upon these existing meta-analyses to answer our research questions, or - if not - how we should proceed to summarise the existing evidence. Five team members (AA, RvD, MH, JI and YS) each assessed the two meta-analyses by using the Quality of Reporting of Meta-analyses (QUORUM) checklist and commented on items if necessary (Moher, 1999;Walker, 1999). The multidisciplinary QUORUM conference aimed to improve the quality of reporting of meta-analyses of randomised controlled trials. One of its products was a 16-item checklist in which the preferred way of reporting a meta-analysis was described (Moher, 1999). QUORUM is now updated in the PRISMA statement (Liberati, 2009). All assessments were compiled into a consensus report which is available in the Annexes (10.a). We will briefly summarise the main points of both meta-analyses here, and then describe the main findings of our assessment.

Definitions (De Maat, 2007b;Glass, 2008) and three Dutch societies of psychoanalytical professionals:
Cognitive behaviour therapy: a usually short-term psychotherapy focussed on identifying and correcting cognitive patterns that underlie emotional and behaviour symptoms;

Schema-focused therapy and *dialectical behaviour therapy*: focused therapies developed for treatment of borderline personality disorder, usually long-term therapies that combine cognitive, behavioural, interpersonal and experiential techniques;

Interpersonal psychotherapy: time-limited individual therapy developed for the treatment of major depression;

Long-term: consisting of at least 40 sessions and lasting at least one year;

Long-term psychotherapy: any form of psychotherapy that is long-term;

Long-term psychoanalytic psychotherapy (LTPP): therapy rooted in psychoanalytic theories, `A therapy that involves careful attention to the therapist-patient interaction, with carefully timed interpretation of transference and resistance embedded in sophisticated appreciation of the therapists´ contribution to the two-person field(Gunderson, 1999)´. In addition: based on psychoanalytic theory (and not, for instance, on client-centred, interpersonal, or schema-therapy theories. Therapy sessions usually have a frequency of once or twice a week;

Psychoanalysis: as long-term psychoanalytic psychotherapy, with four or more therapeutic sessions weekly. In addition, the patient lies on a couch.

6.b De Maat et al.

De Maat conducted a meta-analysis that aimed to examine the effectiveness of long-term psychoanalytic therapies. The authors made a distinction between LTPP and psychoanalysis, and between moderate/mixed pathology and severe pathology – borderline personality disorder (BPD) patients in this study. They conducted a systematic literature review and selected studies on individual LTPP and psychoanalysis in ambulatory, adult patients. 27 studies (16 prospective cohort studies, 10 retrospective cohort studies and 1 RCT) were found. 8/27 studies did not meet the quality criteria of the authors and were not included in the meta-analysis of outcomes, but the pertaining data were presented in tables. They meta-analysed effect sizes (ESs) (Cohen´s d) that were related to pre/post or pre/follow-up changes, and not to between-group differences. Thus, only the LTPP arm of the one included RCT was used in the analyses. For LTPP, the mean ES at therapy termination was 0.78 (standard deviation (SD) 0.45) and at follow-up 0.94 (SD 0.69) in moderate/mixed pathology. For psychoanalysis, the ES at therapy termination was 0.87 (SD 0.41) and 1.18 (SD 0.17) at follow-up. De Maat reported similar results for patients with severe pathology. 95% confidence intervals (95% CI) were not reported. They concluded that LTPP is an effective treatment for a large range of pathologies, with moderate to large effects (De Maat, 2009).

6.c Leichsenring

Leichsenring and Rabung conducted a meta-analysis that aimed to examine the effectiveness of long-term psychoanalytic therapies. The authors were especially interested in the effects of LTPP and psychoanalysis in patients with complex mental disorders, defined as patients with personality disorders, chronic mental disorders, multiple mental disorders or complex depressive or anxiety disorders (i.e. associated with a chronic course and/or multiple disorders). They conducted a systematic search of the literature and included studies on individual LTPP, lasting for at least a year or 50 sessions, with a prospective design and reporting reliable outcome measures. 23 studies (12 prospective cohort studies and 11 RCTs) were included. They calculated ESs (Hedges´ g) for overall effectiveness, target problems, general psychiatric symptoms, personality functioning and social functioning. ESs were related to pre/post or pre/follow-up changes, and not to between-group differences. Thus, only the LTPP arms of the RCTs were used in the analyses. The overall ES at therapy termination was 0.96 (95%CI 0.87 to 1.05). A dummy variable for the type of study (1 for RCTs, 0 for cohort studies) was used to test the correlation between the pre/post ES and the study type. Because this test did not give evidence of a significant correlation, the authors concluded that RCTs give the same results as cohort studies. Of importance, the between-group ESs in RCTs was still not used but only the pre/post ESs in the LTPP arms of RCTs.

Leichsenring then used within-group ESs from the two arms of eight trials that compared LTPP with other forms of therapy. They tested the correlation between the within-group ES and the type of treatment, using a dummy variable (0 for other types of psychotherapy vs. 1 for LTPP). Then, this correlation was used as a measure of between-group effect sizes: *´the point biserial correlation between the within-group ES and treatment condition was significant for overall outcome (r_p=0.60; 95% CI 0.25-0.81; P=0.005, n=20)[...] Thus, LTPP yielded significantly larger pre-treatment/post-treatment ES in overall effectiveness (0.96 vs. 0.47) […] than did other forms of psychotherapy applied in the comparison groups´*. The correlation between within-group ESs and the form of therapy for complex mental disorders was r_p=0.68; 95% CI 0.35-0.86; P=0.002, n=18). n=18 indicates that the ESs of 18 arms of RCTs were used. The authors stated that *´The between-group effect sizes of r_p=0.68 […] are equivalent to Cohen d=1.8 (95% CI 0.7-3.4) […] respectively. [...]a between group ES of 1.8 [...] indicates that after treatment with LTPP, patients on average were better off than 96% of the patients in the comparison groups´*. They concluded that LTPP is an effective treatment for complex mental disorders (Leichsenring, 2008).

6.d Consensus of our assessment of De Maat and Leichsenring´s meta-analyses

We considered the use of non-controlled data - or in the case of Leichsenring: the reduction of controlled data to observational data by separating the arms of RCTS – to be the most serious flaw of both meta-analyses. It is difficult to examine the effectiveness of any treatment in uncontrolled conditions. How can the effects of a treatment be separated from the effects of an alternative treatment? Furthermore, absolute change (before *vs.* after) cannot be interpreted independently from time effects (including aging) and so-called non-specific effects like attention, empathy, expectations, explanations for problems etcetera. Time is an important factor in the course of many mental disorders. E.g. the remission rates for BPD in one study were 40% at two years to 88% at ten years (Zanarini, 2006). In a Dutch population study, 50% of the persons with a new episode of a depressive disorder was recovered within 3 months (Vollebergh, 2003).

In addition, Leichsenring´s analysis seems misleading. Several experts have commented to this extent in a series of letters to the editor (Beck, 2009;Kriston, 2009;Thombs, 2009). First, Leichsenring refers to the estimated ESs as ESs from meta-analysed RCTs. But separating the arms of the trials, and using the within-group pre/post ESs, reduces these RCTs in to observational (cohort) studies. Secondly, a roundabout way is used to estimate a between-group effect size, in which the original intervention groups seem to be separated from their original control groups. The within-group ES of all intervention groups are averaged and compared to the average within-group ES of the control groups. These two averaged within-group ES are then correlated to a dummy variable for the type of therapy, and the correlation coefficient is transformed into a between-group ESs. In an author´s reply Leichsenring stated that he chose this method because the number of RCTs of LTPP was rather small. He also reports a between-group ES *´assessed in the conventional way´* for overall outcome of 0.65 (Hedges´ g)(p=0.03) (no 95% CI reported) (Leichsenring, 2009). Thirdly, each ES should be based on the comparison of a group with its own control to avoid a problem known as Simpson´s paradox. When the unit of the trial is not maintained, bias and confounding by study can be introduced (Borenstein, 2009;Egger, 2001).

Thus, our main conclusion was that the existing meta-analyses could not be used to answer our research questions. A meta-analysis of RCTs – with statistically combining between-group ES – is required to examine the effectiveness of LTPP. This report first describes the exact methods we used, then synthesizes the results of the systematic literature search and meta-analyses the outcome data. In the last sections we discuss our findings and present an overall conclusion.

7 Methods

7.a Criteria for considering studies for this review

7.a.1 Types of studies

We selected controlled studies only. This included randomised controlled trials and controlled trials where treatment allocation was not at random. However, allocation did have to demonstrate a certain degree of freedom from bias. For example, a study with allocation based on clinical judgement or patient preference would be excluded, while a study with allocation based on time of referral would be included. Only studies published in peer-reviewed publications were selected.

7.a.2 Types of participants

Participants had to be adults with a clearly defined mental disorder. Consequently, populations with a primarily behavioural problem were excluded if concurrent mental disorders were not present or not described. E.g. a study with participants with a history of aggression or spouse beating would be excluded, if the participants were not diagnosed with a defined mental disorder. In addition, studies with participants with a pure somatic disorder (diabetes, morbus Crohn, irritable bowel disease) were excluded. Studies on patients with schizophrenia were excluded because schizophrenia is very distinct from the mental illnesses for which LTPP is prescribed in the Netherlands at present (Berghout, 2008).

7.a.3 Types of interventions and controls

The intervention had to be psychoanalysis or long-term psychoanalytically based psychotherapy. 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 (see text box below). If a study designated the index therapy as a mixed form, we considered it to be a non-psychoanalytical therapy. This was the case with cognitive-analytical therapy, that combined elements of cognitive therapy and brief, focused psychodynamic psychotherapy (Dare, 2001). However, we did include studies in which non-mixed LTPP was given in adjunction to other forms of therapy (see also further on in this section, when we discuss the extent to which the control treatment needed to differentiate from the intervention).

Main schools/treatments in the field of LTPP (overview by three Dutch societies of psychoanalytical professionals)
- Ego psychology (S. Freud)
- Object relations theory (Klein, Mahler, A. Freud)
- 'Builders of bridges´ between the ego psychology and object relations theory (Fairbairn, Winnicott, Kernberg and Yeomans)
- Primary love theory (Balint)
- Self psychology (Kohut)
- Attachment theory (Bowlby, M. Ainsworth, M. Main)
- Interpersonal school (Ferenczi, Sullivan)
- Intersubjectivity (Stern, Mitchell, Renik and Hoffman)
- Mentalisation Based Treatment (Fonagy, Bateman and Target)

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, 2009;Leichsenring, 2008). However, in our opinion it is likely that psychotherapy with a once-a-week frequency will result in a total of less than 50 sessions in a year, allowing for patients´ and therapists´ vacations etcetera. 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, 2006). In addition, the Dutch

societies of psychoanalytical professionals defined long-term psychotherapy as more than 40 sessions as well.

To be considered for inclusion a study should intend to investigate a treatment of at least 40 sessions and continue for at least a year. The mean number of sessions per participant could be less than 40 sessions however, if dropouts or no-shows were included in the calculation of the mean.

The control treatment had to differ substantially from the intervention treatment. It either had to be a different type of treatment (a non-psychoanalytically based treatment) and/or a short-term treatment. The comparison made in the study had to give information that helped answer our main research question: is long-term psychoanalytic psychotherapy/psychoanalysis an effective treatment for mental illness? Thus, control treatments could be either treatment as usual (TAU), wait-list conditions, cognitive therapy, short-term psychoanalytical psychotherapy or non-evidence-based non-psychoanalytically based other forms of therapy. A comparison between inpatient *vs.* outpatient LTPP or individual *vs.* group LTPP, for example, was excluded. However, we included studies where LTPP was given in adjunction to other treatments, even if not controlled for in the control condition, despite the problem that effects of LTPP and the other elements of the package could not be disentangled (e.g., the Bateman study (Bateman, 1999)). However, we tested the sensitivity of our results by repeating the analyses with such studies excluded. The small number of RCT's involving LTPP was the mean reason to take such studies into account.

7.a.4 Types of outcome measures

We considered the primary outcome the recovery rate and secondary outcomes were measures of the target problem, of general psychiatric symptoms, personality pathology, social functioning and quality of life (QoL).

7.b Search strategies

Medline, Embase, PsycINFO, and all Evidence Based Medicine (EBM) Reviews were searched through OVID®. The EBM Reviews are (a) the ACP Journal Club; (b) the Cochrane Database of Systematic Reviews; (c) The Cochrane Central Register of Controlled Trials; (d) the Cochrane Methodology Register; (e) the Database of Abstracts of Reviews of Effects (DARE); (f) the National Health Services Health Technology Assessment Database (NHS HTA); and (g) the NHS Economic Evaluation Database (NHS EED).

We searched using text words and indexing terms. Words and terms were generated according to the PICO method (population, intervention, control, and outcome):

***P**: populations with a clearly defined mental disorder;*
***I**: psychoanalysis, psychoanalytical psychotherapy, psychodynamic psychotherapy;*
***C**: no control was specified but the study type was; randomised controlled trial, controlled trial, trial, evaluation study, meta-analysis;*
***O**: outcomes were not specified.*

Searches were not limited by a time-period or a language or in any other way. Search details can be found in 10.b the Annexes. 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. The function ´find similar´ was used to look for studies similar to those in the final selection. References of an existing international guideline were scanned for additional studies (Canceil, 2004).

7.c Data collection and analysis

7.c.1 Selection of studies

Two reviewers (MH and YS) independently selected suitable studies for inclusion as detailed below. Where the two reviewers disagreed about the inclusion of a study, disagreements were resolved by consensus of opinion. A third reviewer was consulted if a disagreement could not be resolved. Where resolution was not possible the author was contacted to obtain more information and clarification. The titles and abstracts of studies identified by searching electronic databases were assessed to determine if an article was eligible. An article was rejected when the title and abstract contained sufficient information to determine that it did not meet the inclusion criteria. The full papers of all remaining articles were retrieved.

7.c.2 Assessment of methodological quality of included studies

Selected studies were judged on their methodological quality by two independent researchers (MH, YS). When the researchers disagreed, consensus was reached through discussion.

We used (a) the Maastricht-Amsterdam Criteria List (van Tulder, 1997) and (b) eight criteria proposed by Cuijpers et al (Cuijpers, 2009). Detailed information on these criteria can be found in section 10.c in the Annexes. In brief, the criteria lists used are characterised by:

(a) The Maastricht-Amsterdam Criteria List was originally designed to rate research in the field of muscular-skeletal disorders but is considered to produce disease non-specific quality ratings and has been used in a Cochrane review on the treatment of mental disorders (Henken, 2007)

(b) The criteria used by Cuijpers et al were based on an authoritative review of empirically supported psychotherapies (Chambless, 1998), and on the criteria proposed by the Cochrane Collaboration to assess the methodological validity of a study (Anon. 2009). The criteria based on the review of empirically supported psychotherapies assessed the quality of the treatment delivery, while the criteria proposed by the Cochrane Collaboration assessed more methodological sources of bias

We changed 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. We felt that this reflected the long-term nature of the treatments under study in a better way.

7.c.3 Data extraction and management

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, when available

7.c.4 Data analysis

We did a meta-analysis of (a) recovery rates of participants (primary outcome) and (b)effect sizes for the effect on target problems, general psychiatric symptoms, personality pathology, social functioning and for overall effectiveness (secondary outcomes). We also wanted to combine QoL measures, but only one study reported those. Target problems were defined as the problem the treatment was primarily focusing at, and included recovery. E.g., in a study of the treatment of depression, some measure of depression severity would be a measure of the target problem. "Target problem" should not be confused with "target complaint". The latter refers to a problem the patient wants to be helped with, is unstandardised and does not relate to any form of psychopathology. For one person the target complaint could be relationship trouble, for another person it could be nightmares, or postponing decisions, etc.

We included overall effectiveness for the sake of comparison with previous meta-analyses. However, we do not consider this a useful outcome. As an unweighted mixture of all available outcomes it cannot be interpreted and is possibly invalid. When more than two outcome measures were available we used the mean effect size, calculating Hedges' g for each effect size. 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 (say less than 10), so the difference is usually trivial (Borenstein, 2009). To calculate the effect size of overall effectiveness we used the mean ES of all available outcomes in a study.

When we use the term follow-up we refer to the time between baseline assessment and any further assessment, which may be during treatment, at the end of treatment (post-treatment) or sometime after the end of treatment. For the meta-analyses we used the longest available follow-up because LTPP should bring about change that is stable in the long run. A difference in treatment effectiveness should be easier to detect at a longer follow-up.

Subgroup analyses for study quality characteristics, type and severity of mental disorders, different types of therapy (individual *vs.* group therapy, outpatient *vs.* inpatient therapy) and therapy intensity were performed. All analyses used the random effects model because we assumed the data to be heterogeneous, amongst others because of diversity in populations (mental disorders) and control treatments. We explored heterogeneity using meta-regression (method of moments). We considered the proxy session ratio (number of sessions in the intervention group/number of sessions in the control group) and the internal validity score to be covariates that might explain part of the heterogeneity in effect sizes (Cuijpers, 2009;Leichsenring, 2008). The rationale for the random effects model is further explained in a recent paper by Higgins et al. (Higgins, 2009).

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:

- Evidence based treatment for the condition under study
- STPP
- Structured, non-evidence based treatment with the most similar treatment intensity
- Structured, non-evidence based treatment with the most similar treatment mode (individual or group therapy, outpatient or inpatient therapy)
- TAU or other non-structured treatment

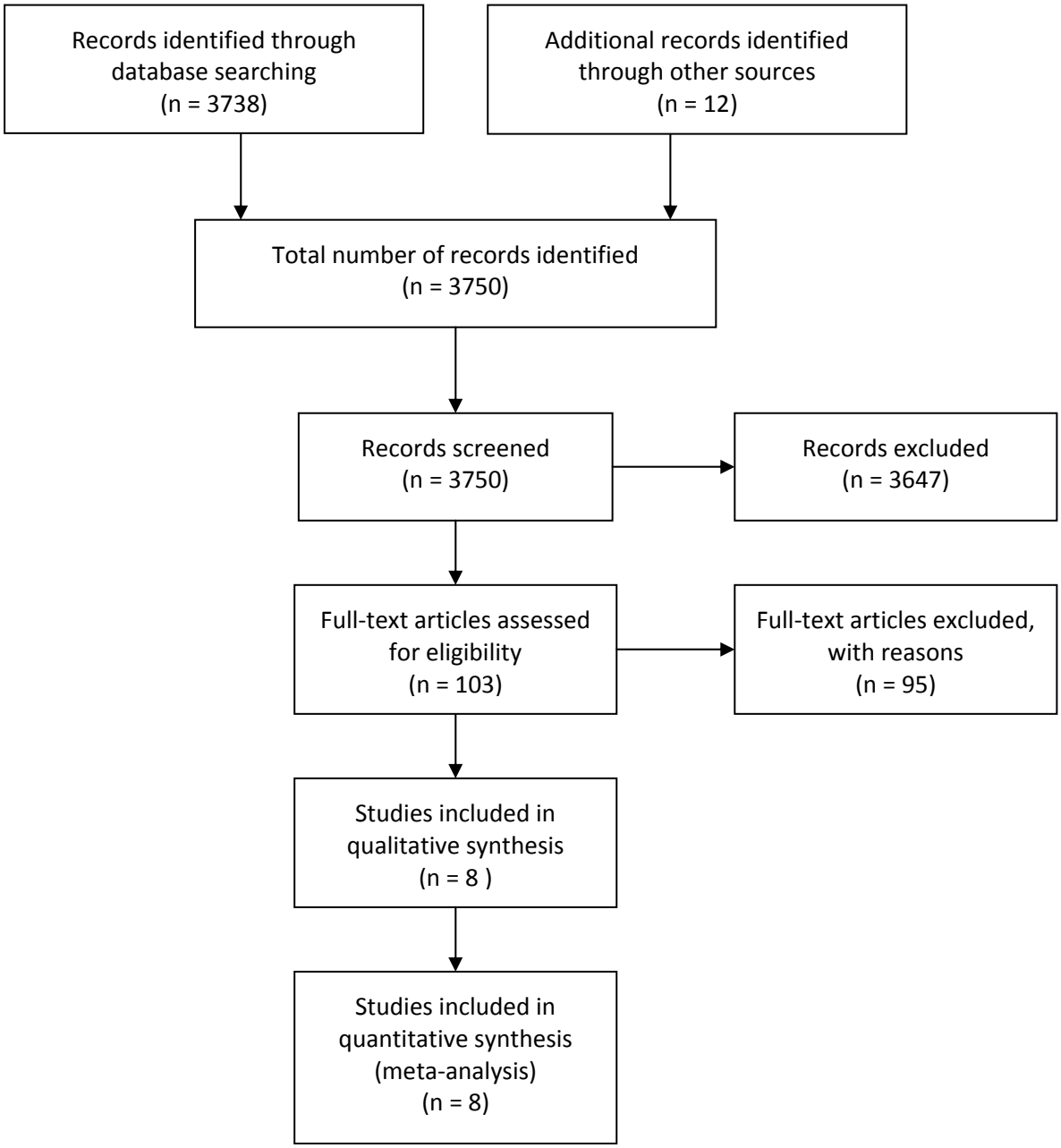
We examined the data for publication bias by visual inspection of funnel plots and by Duval and Tweedie´s trim and fill test. Comprehensive Meta-analysis Version 2 (Biostat Inc) software package was used for all meta-analyses and meta-regression. Stata Release 10 (StataCorp LP). STATA™10.0 (StataCorp, College Station) was used to calculate 95% CI for frequencies when needed. Therapies refer to individual outpatient therapy, unless specified otherwise.

8 Results

We screened a total of 3750 studies, of which 8 were included (Figure 1). The two independent reviewers (MH and YS) agreed on the inclusion and exclusion of all studies, with the exception of seven. These seven studies were discussed to reach consensus:

- Through discussion consensus was reached to exclude four studies (Chiesa, Høglend, O´Brien and Vinnars). The reasoning behind the exclusion is described in 8.a.1
- Further information was requested on three studies, by contacting the corresponding authors by e-mail:
    - The Linehan study: the 25 therapists that were assigned to the control patients were described as *´"eclectic but nonbehavioral" or "mostly psychodynamic*". It was also stated that their "clinical supervision group met at the Seattle Psychoanalytic Society and Institute and was led by its training director" (Linehan, 2006). We felt unsure whether we could categorise this as LTPP. We contacted the authors and received information that 21/25 (84%) of the therapists described their methods as psychoanalytic or psychodynamic, 3/25 described themselves as interpersonal therapists, and 1/25 as humanistic/client centred. We decided to include this study
    - The Korner study. It was unclear how exactly study participants were sampled, but it seemed participants in the intervention group were sampled in retrospect, and on the basis of having completed 12 months of treatment (Korner, 2006). The author confirmed this and this study was excluded
    - The Winston study, who´s title suggests that it concerned short-term psychotherapy. However, the mean number of sessions in the two treatment groups was 40.3 sessions (Winston, 1994). We mailed the author asking for the intended frequency and duration of the therapies included in his study but we did not receive a reply. We then decided to exclude this study because it seemed most likely this was not a study on LTPP as per our definition (at least 40 sessions and a duration of at least 1 year). The article does not state a predetermined duration of the therapies. However, it stated that ´these therapies lasted approximately 40 weeks´. Thus is seems likely that our criterion of a duration of at least 1 year is not met

Figure 1 Flow chart of study search and selection



8.a Description of studies

8.a.1 Excluded studies

95 Articles were excluded on full-text review (Table 2). The two main reasons for exclusion were that a study was either (1) a short-term treatment or (2) uncontrolled. Nevertheless, four *controlled* studies (Chiesa, Høglend, O´Brien and Vinnars) were excluded because the study design did not yield information on the main research question: is LTPP an effective treatment? These four studies made a comparison between two specific forms of LTPP and not between LTPP and another form of psychotherapy or STPP. Thus, they yielded no information on the effectiveness of LTPP as such:

a) Chiesa studied three treatment groups: (1) an inpatient LTPP; (2) a mixed inpatient/outpatient LTPP; and (3) TAU. The TAU group was excluded from any comparison because it was recruited from a very different sample than the two LTPP groups. The two LTPP groups were recruited from personality disorder patients referred to a tertiary care facility for inpatient treatment, whereas the TAU group was recruited from among the caseload of all the senior psychiatrists in a certain district. We excluded the comparison between the inpatient LTPP *vs.* the mixed inpatient/outpatient LTPP because this comparison would not yield any information on the effectiveness of LTPP as such (Chiesa, 2006)
b) Høglend et al. compared LTPP with a moderate level of transference interpretations *vs.* LTPP with no level of transference interpretation (Hoglend, 2006;Hoglend, 2008)
c) O´Brien et al. compared group *vs.* individual psychotherapy, with therapists that "*took a dynamic approach to therapy and were eclectic in orientation*" (Mintz, 1976;O'Brien, 1972)
d) Vinnars´ study intended to compare manualised supportive-expressive psychotherapy *vs.* a non-manualised community-delivered psychodynamic therapy. We considered to include this study as a comparison between LTPP (the supportive expressive therapy group) and STPP (the community delivered psychodynamic therapy group). However, both groups ended up receiving the same number of therapy sessions (26.2 *vs.* 28.0 sessions on average). Community-delivered psychodynamic therapy, as delivered in Vinnars´ study, was close to 50% more intensive than the treatment given to the personality disorder patients in centers that did not participate in the study (Vinnars, 2005). Thus this study was in fact a comparison between two forms of LTPP

Two controlled studies were excluded because we could not combine the data in a way that would reflect the between-group differences (Clarkin, 2007;Munroe-Blum, 1995). One study reported only the elevation (intercept) of the individual trajectory and the rate of change (slope) of the individual trajectory was reported, but not the post-means and SDs (Clarkin, 2007). The author was mailed to obtain the needed data. We received no reply. The other study reported no between-group differences but only the mean scores for all trial participants. In addition, it was unclear whether the psychodynamic psychotherapy treatment was *intended* to be a long-term treatment, though the authors stated that it was an open-ended treatment and 12 months post-randomisation was labeled as end of treatment (Munroe-Blum, 1995). Two other RCTs were excluded because the outcome data were not reported in sufficient detail to be able to meta-analyse the data (Karon, 1972;Piper, 1984).

Lastly, one controlled study was excluded because the exact treatment and control were unclear. It compared LTPP with a wait-list control group in a patient sample of collaborating psychoanalysts who contributed patients. The sampling of patients was unclear, as was the duration the control group spent on the waiting list. In addition, the exact type of therapy received was unclear (Klar, 2005).

Table 2 Studies excluded on full text review

	Reason for exclusion	
1	Describes the concept and proceedings of a health economy evaluation of an extended psychosomatic rehabilitation programme for outpatients. No outcome data described and no (more recent) article with outcome data found	(Albrecht, 2000)
2	8 months of therapy reported on at present, follow-up is continued. The study is a randomised controlled trial that compares day hospital psychotherapy with outpatient individual psychotherapy for patients with personality disorders. Control therapist mainly adhered to psychoanalytic/psychodynamic theories, although cognitive and systemic elements were present. So it is not clear if this study could have been used if it had a longer follow-up.	(Arnevik, 2009)
3	Data on the post treatment period (from 1.5 to 3 years)	(Bateman, 2001)
4	On the background of mentalisation based therapy	
5	A descriptive comparison of the conceptual models of mentalisation based therapy and cognitive analytical therapy	(Bateman, 2007)
6	Data on the post treatment period (from 3 to 8 years)	(Bateman, 2008)
7	See Chiesa 2000	(Beecham, 2006)
8	Short-term treatment (6 weeks)	(Berry, 1989)
9	Short-term treatment (up to 30 sessions)	(Bolz, 1981)
10	Proposes how to conduct new research, with a small pilot-study	(Briffault, 2007)
11	Short –term treatment (an average of five months of psychotherapy ´of the kind the vast majority of clinical patients were getting´)	(Brill, 1966)
12	Prospective study of two cohorts, one in psychoanalytical therapy and one in cognitive behaviour therapy. Participants were consecutive new patients of either a psychoanalytically oriented therapist or a cognitive behavioural therapist. Treatment allocation is not specified. It is mentioned that patients could be referred by health care workers or be self-referrals.	(Brockmann, 2006)
13	Short-term treatment (mean 18 sessions)	(Brom, 1989)
14	Short term treatment (10 weeks)	(Burnand, 2002)
15	Short-term treatment (up to 30 sessions)	(Burzig, 1981)
16	Short-term treatment (up to 8 sessions)	(Chabrol, 2002)
17	Describes the study the authors intend to undertake	(Chiesa, 1999)
18	Compared inpatient LTPP *vs.* mixed inpatient/outpatient LTPP *vs.* TAU. The TAU group was excluded because it was recruited from a very different sample than the two LTPP groups. The comparison between the two LTPP groups was excluded because it yielded no information on LTPP effectiveness	(Chiesa, 2000)
19	See Chiesa 2000	(Chiesa, 2002)
20	See Chiesa 2000	(Chiesa, 2003)
21	See Chiesa 2000	(Chiesa, 2004)
22	See Chiesa 2000	(Chiesa, 2006)
23	Description of concept of transference focussed therapy, exploring studies and preliminary trial results	(Clarkin, 2005)
24	Data could not be combined in a way that would reflect the between-group differences	(Clarkin, 2007)
25	Short-term treatment (10 weeks)	(Cooper, 2003)
26	Short-term treatment (up to 9 months)	(Crits-Christoph, 1997)
27	Short-term treatment (up to 16 sessions)	(Crits-Christoph, 1996)
28	Retrospective comparison of two matched cohorts (one in psychoanalysis and one on a waiting list)	(Dührssen, 1998)
29	No comparison of treatments (100 sessions of client-centred group therapy). Moreover client-centred group therapy is not a psychoanalytic treatment.	(Eckert, 2000)
30	A nested, controlled trial of long-term psychodynamic group therapy *vs.* treatment as usual. However, treatment allocation was biased. Participants all came from a study called the Treatment Enhancement Program for Bipolar Disorders. Participants in the intervention group were referred to the intervention group by psychiatrists or psychiatric nurses. The grounds on which they were referred are not described. Controls were matched to the intervention group.	(Gonzalez, 2007)
31	Prospective observational study of two matched cohorts, one in psychodynamic psychotherapy and one in psychoanalysis. Treatment allocation was done by individual therapists who gave both forms of therapy ´ The therapists were directed to opt for one of these treatment forms before the onset of therapy and to adhere to this option throughout´	(Grande, 2006)
32	No form of psychoanalytical based therapy is examined	(Grawe, 1990)
33	Matched case-control design. Investigates the effect of psychoanalysis on death from cancer or coronary heart disease	(Grossarth-Maticek, 1990)

	Reason for exclusion	
34	No treatment comparisons. Sample of patients in a psychoanalytical outpatient clinic who are treated with psychoanalysis or psychoanalytically oriented treatment. Treatment allocation is unclear. This article compares the individual therapy goals at the beginning of treatment with the outcome of therapy as measured by the text content of a catamnestic interview	(Heuft, 1996)
35	Short-term therapy (11 weeks)	(Hoffart, 1990)
36	Cohort study. Patients with somatoform disorders and patients with social-medically relevant problems, who had been on sick leave for more than six months. No comparisons between therapies made.	(Hoffmann, 2007)
37	Randomised trial of LTPP with a moderate level of transference interpretations *vs.* LTPP with no level of transference interpretation	(Hoglend, 2006)
38	Randomised trial of LTPP with a moderate level of transference interpretations *vs.* LTPP with no level of transference interpretation	(Hoglend, 2008)
39	Not a peer-reviewed publication (book chapter). Presenting preliminary results only (half a year after treatment start) of a trial of psychoanalysis *vs.* psychodynamic psychotherapy	(Huber, 2001)
40	Not a peer-reviewed publication (abstract in a book). Concerning a cohort of patients in psychoanalysis	(Huber, 2006)
41	Cohort study of a group of inpatients in psychodynamic psychotherapy	(Huber, 2009)
42	Meta-analysis of four German studies we reviewed in the selection process	(Jakobsen, 2007)
43	Short-term treatment (46 sessions over a preset 9 month period) (Steuer 1984 publishes on the same study). In addition, patients were recruited through mass media appeals.	(Jarvik, 1982)
44	Short-term treatment (3 months)	(Jäger, 1997)
45	No usable outcome data reported	(Karon, 1972)
46	Study that compared LTPP with a waiting list control group in a patient sample of collaborating psychoanalysts who contributed patients. The sampling of patients is unclear and the duration that the control group spent on the waiting list was not reported. In addition, the exact type of therapy received was unclear	(Klar, 2005)
47	Participants in the intervention group were sampled in retrospect, and on the basis of having completed 12 months of treatment	(Korner, 2006)
48	Population had no clearly defined mental disorder: participants were male patients from a general and psychiatric hospital who had assaulted during the previous 6 months	(Lanza, 2002)
49	Short-term treatment (six months)	(Lanza, 1995)
50	Uses the patients from Clarkin´s trial to examine changes in attachment organization and reflective functioning, which are not themselves indices of psychopathology. This study is more a test of theoretically assumed mechanisms of change	(Levy, 2006)
51	Concerns affect-focused body psychotherapy, which is especially suited for physiotherapists to work with. It is a ´psychodynamic body therapy, primarily based on psychomotor physiotherapy and affect consciousness treatment, within a general psychodynamic frame of reference´. We consider it to be more a form of psychomotor physiotherapy than a form of psychotherapy	(Levy Berg, 2009)
52	Retrospective sample of a cohort in psychoanalysis or psychodynamic psychotherapy	(Leuzinger-Bohleber, 2001)
53	Controlled study of psychoanalytical therapy. Outcome reported at 4 months. In addition: ´the duration …is somewhere between that of brief and long-term psychotherapy and lasts an average of 10 months (range 3 – 24 months)´	(Manos, 1984)
54	The exact form of psychotherapy was unclear: ´psychotherapy, given for an average of not less than two hours a week, supervised by a psychoanalyst experienced in the treatment of schizophrenic patients´. Treatment duration was unclear: `Treatment of all forms was continued up to a maximum of one year, unless the patient was released from hospital before that´	(May, 1965)
55	See Piper 1984	(McCallum, 1990)
56	Control treatment has psychodynamic traits but is not a true psychodynamic treatment. The author´s refer to is as a ´dynamically informed psychotherapy´	(McMain, 2009)
57	Short-term treatment (up to 30 sessions)	(Meyer, 1981)
58	Short-term treatment (up to 30 sessions in 9 months)	(Miklowitz, 2007)
59	Short-term treatment (up to 24 sessions in 12 weeks)	(Milrod, 2000)
60	See Milrod 2000	(Milrod, 2001)
61	See Milrod 2000	(Milrod, 2007)
62	See O´Brien	(Mintz, 1976)
63	Outcome measures were not reported for each trial arm separately	(Munroe-Blum, 1995)

	Reason for exclusion	
64	Group vs. individual psychotherapy, with therapists that "took a dynamic approach to therapy and were eclectic in orientation"	(O'Brien, 1972)
65	Short-term treatment (intended number of treatment sessions 30, which was seldom reached)	(Ojehagen, 1992)
66	Not a peer-reviewed publication (abstract of a presentation and a power point presentation)	(Petrak, 2007)
67	No usable outcome data reported	(Piper, 1984)
68	Examines patients´ characteristics as a predictor of treatment success. Patients are from the included trial of Piper 1984 (Piper, 1984)	(Piper, 1994)
69	Prospective cohort study that compares psychodynamically oriented psychotherapy, cognitive behavioral therapy, and analytic psychotherapy in a cohort of patients that applied for subsidized outpatient psychotherapy. Treatment allocation was based on patients preferences backed by clinical judgment.	(Puschner, 2004;Puschner, 2007)
70	Description of a pilot study on the outcome of psychoanalysis. No outcome data presented	(Rascon, 2005)
71	Prospective cohort study of two cohorts of patients, one in psychoanalysis and one in psychodynamic psychotherapy	(Rudolf, 1999)
72	Prospective cohort in psychodynamic psychotherapy	(Rudolf, 2004)
73	Short-term treatment (2-4 months inpatient treatment)	(Sachsse, 2006)
74	No true pre-treatment assessment, but a retrospective assessment based on non-standardised, written referrals. Observational study labelled as ´naturalistic experimental design´. Treatment allocation was determined by time on a waiting list.	(Sandell, 1997;Sandell, 1999;Sandell, 2000)
75	Short-term treatment (15 weekly sessions)	(Sandahl, 1998)
76	See Klar 2005	(Schleussner, 2005)
77	Partly retrospective (pre-treatment) comparison of two cohorts: one in psychoanalytical oriented treatment and one in cognitive behaviour oriented treatment	(Schulz, 1999)
78	Short-term therapy (five months)	(Shaffer, 1997)
79	Long-term treatment was not a predetermined research question as ´keep patients in treatment for the minimum duration of 6 months´	(Siassi, 1979)
80	Short-term treatment (´counsellors were asked to try to keep the number of sessions to six to 12 sessions´)	(Simpson, 2000)
81	See Simpson 2000	(Simpson, 2003)
82	Retrospective cohort study	(Sohlberg, 1987)
83	Concerned with identifying prognostic markers of treatment outcome. Based on the included trial of Giesen-Bloo et al.	(Spinhoven, 2008)
84	Investigated the quality and development of the therapeutic alliance as a mediator of change. Based on the included trial of Giesen-Bloo et al.	(Spinhoven, 2007)
85	Short-term treatment (46 sessions over a preset 9 month period). In addition, patients were recruited through mass media appeals.	(Steuer, 1984)
86	Prospective comparison of three cohorts: cognitive-behavioural therapy, person-centred therapy and psychodynamic therapy. Treatment allocation was not described. Patients were a sample of patients who attended one of 58 NHS sites delivering counselling and psychotherapy services	(Stiles, 2006)
87	Observational study that compares patients receiving treatment with either cognitive-behavioural, person-centred, or psychodynamic therapies in primary-care routine practice	(Stiles, 2008)
88	Short-term treatment (up to 30 sessions)	(Stuhr, 1981)
89	Short-term treatment (10 sessions)	(Taylor, 1993)
90	Short-term treatment (14 weeks)	(Teusch, 1997)
91	Studies patient sex as moderator of effects in the Høglend trial	(Ulberg, 2009)
92	A comparison of two forms of LTPP (see also 8.a.1)	(Vinnars, 2005)
93	See Vinnars 2005	(Vinnars, 2007)
94	Short-term treatment (20 sessions)	(Walsh, 1997)
95	Short-term treatment	(Winston, 1994)

8.a.2 Included studies

Eight controlled studies were included. Table 3 gives an overview of these studies and the pertaining published articles, corrections and correspondence. Hereafter we will refer to each study by the first author of the main publication (Table 3, second column). All studies were RCTs.

The inclusion of two studies (Bateman and Linehan) was discussed extensively in our team. First, we questioned the type of intervention the Bateman study examined: could the reported outcomes be attributed to the psychoanalytical ingredients? 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 hour each), 3) once-a-week expressive therapy oriented toward psychodrama techniques (1 hour), and 4) a weekly community meeting (1 hour), all spread over 5 days´* (Bateman, 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 study but to run a sensitivity analysis without it to check the robustness of our findings.

Secondly, in the Linehan study the control group consisted of community treatment by experts, given by 25 therapists (Linehan, 2006). 21/25 (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, 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 ran a sensitivity analysis without the Linehan study.

Table 3 Selected controlled studies with the corresponding scientific articles

	Study	Articles, letters and comments	References
1	Bachar	Bachar 1999	(Bachar, 1999)
2	Bateman	Bateman 1999	(Bateman, 1999)
		Bateman 2003	(Bateman, 2003)
		Stern 2001 (comment)	(Stern, 2001)
3	Dare	Dare 2001	(Dare, 2001)
		Bell 2001 (comment)	(Bell, 2001)
		Okhai 2001 (comment)	(Okhai, 2001)
4	Giesen-Bloo	Giesen-Bloo 2006	(Giesen-Bloo, 2006)
			Correction Arch Gen Psych 2006;63:1008
		Van Asselt 2008	(van Asselt, 2008)
		Grenyer 2007(comment)	(Grenyer, 2007)
		Pearce 2007 (comment)	(Pearce, 2007)
		Yeomans 2007 (comment)	(Yeomans, 2007)
5	Gregory	Gregory 2008	(Gregory, 2008)
6	Knekt	Knekt 2008 a	(Knekt, 2008a)
		Knekt 2008 b	(Knekt, 2008b)
7	Linehan	Linehan 2006	(Linehan, 2006)
			Correction Arch Gen Psych 2007;64(12):1401
8	Svartberg	Svartberg 2004	(Svartberg, 2004)

8.a.3 Settings and participants

Table 4 gives an overview of the setting of the included studies. It includes the sample from which study participants were recruited and the in- and exclusion criteria that were used. Psychiatric disorders were diagnosed according to DSM criteria, unless specified otherwise. Two studies concerned patients with eating disorders, four studies concerned patients with a personality disorder and two studies were on mixed patients.

Table 4 Settings and participants of selected controlled studies

Years of intervention	Country	Sample population	Inclusion criteria	Exclusion criteria
Bachar				
1999 *	Israel	Referrals to eating disorder units of two general hospitals	Eating disorder	Co-morbidity Axis I DSM
Bateman				
1993 - 1996	UK	Referrals to a psychotherapy unit	BPD	Schizophrenia Bipolar disorder Substance misuse Mental impairment Organic brain disorder
Dare				
2001 *	UK	Referrals to an outpatient eating disorder service in a psychiatric teaching hospital	Anorexia Age ≥ 18	Urgent hospital admission required (e.g. suicidal risk, extremely low weight, hypoglycaemia, syncope, electrolyte depletion)
Giesen-Bloo				
1999-2004	NL	Referrals to the study	BPD Age 18-60 Dutch literacy	Psychotic disorders Bipolar disorder Dissociative identity disorder Antisocial PD ADHD Addiction needing clinical detoxification Psychiatric disorders secondary to medical conditions Mental retardation
Gregory				
2004-NA	US	Referrals to the study	BPD and alcohol use disorder Age 18-45	Schizophrenia Schizoaffective disorder Mental retardation Psychiatric disorders secondary to medical conditions Low IQ
Knekt				
1994-	Finland	Referrals to the study	Long-standing (>1 year) disorder causing dysfunction in work ability Anxiety or mood disorders and neurosis to higher-level borderline disorder $ Age 20-45	Psychotic disorders Severe personality disorder Adjustment disorder Substance-related disorder Organic brain disease Severe organic disease Mental retardation Treated with psychotherapy in the previous 2 years

Years of intervention	Country	Sample population	Inclusion criteria	Exclusion criteria
Linehan				
2006 *	US	Referred women	BPD and current and past suicidal behaviour (at least 2 suicide attempts or self-injuries in the past 5 years, with at least 1 in the past 8 weeks) Age 18-45	Schizophrenia Schizoaffective disorder Bipolar disorder Psychotic disorder Mental retardation Seizure disorder requiring medication A mandate to treatment A need for primary treatment for another debilitating condition
Svartberg				
2004 *	Norway	Referrals to the study	Cluster C PD Self-defeating PD Age 18-65	Psychotic disorder, Substance dependence/abuse Eating disorder Organic brain disease Serious physical illness Active suicidal behaviour Refusal to discontinue other active treatment

* Years of intervention not available. Data of first publication of outcomes is given
$ According to Kernberg's classification of personality organization
Abbreviations: ADHD: Attention-deficit/hyperactivity disorder; BPD: borderline personality disorder; DSM: diagnostic and statistic manual of mental disorders; IQ: intelligence quotient; NA: not available; NL: the Netherlands; PD: personality disorder; UK: United Kingdom; US: United States

8.a.4 Interventions and controls

Table 5 gives an overview of the treatments that were examined in our selection of controlled studies. For the sake of convenience and unequivocal statistical analysis, the long-term psychoanalytically based treatment was always labelled the intervention treatment. All interventions were LTPP: no controlled studies on psychoanalysis were found. In Table 5 we give a more elaborate description of the treatments used in each study. In the light-grey highlighted rows we give a label to each treatment, meaning a standardised way of referring to each treatment. Thus, a treatment referred to as ´ focal psychoanalytic psychotherapy´ (Dare) or ´ self psychological treatment´ (Bachar) is labelled as an ´individual outpatient LTPP´ by us.

TAU is the ´normal´ care that is given under non-study conditions in the study´s setting. It usually involves a form of general psychiatric care and general supportive measures. Though TAU might also be classified as a non-evidence based treatment, we did make a distinction because TAU was a less structured and less intense control treatment compared to the various non-evidence based control treatments.

8.a.5 Treatment frequency

Frequency of treatment is described in Table 6. Few studies gave a mean number of sessions for completers and most studies did not give an average number of sessions for all participants. To give an indication of treatment intensity we calculated a proxy variable. This proxy equalled the mean number of sessions in completers, if available. If not available, the proxy was the predetermined treatment frequency, multiplied by the attending rate if possible. This proxy of treatment intensity ranged from 40 sessions/1 year to 302 sessions/1.5 years (Table 6). We then calculated a proxy for the session ratio: the number of sessions in the intervention group/the number of sessions in the control group. This session ratio ranged from 0.4 to 7.9.

The Dare and Linehan studies seemed to concern STPP, with a mean number of 24.9 sessions (Dare) and a median number of 33.0 sessions (Linehan) in the intervention groups. In the Dare study however, this mean was calculated across all participants, including non-completers (Dare, 2001). Since 9/21 participants (43%) in the intervention group were non-completers, and because the intervention treatment was intended to take place weekly for at least 1 year, we did consider the Dare study to be a study on LTPP, and we included it.
Similarly, we did consider the Linehan study to be a study on LTPP. Because of the high drop-out rate in the intervention group (59.2% dropped their first study therapist) we considered it likely that a median number of 33.0 sessions across all participants meant that the intervention was a LTPP, though we could not recalculate the mean number of sessions in completers.

Table 5 Characteristics of intervention and control treatments: assessment of comparisons made

Study	*Disorder (N)* - *Our assessment of the treatments* Intervention treatment	Control treatment
Bachar	*Anorexia (7 vs. 6 patients)* - *Individual outpatient LTPP vs. non-evidence based control treatment*	
	Self psychological treatment	Cognitive orientation treatment
	Bulimia (10 vs. 11 vs. 10) - *Individual outpatient LTPP vs. two different non-evidence based control treatments*	
	Self psychological treatment and nutritional counselling	Cognitive orientation treatment and nutritional counselling
		Nutritional counselling
Bateman	*Personality disorder (22 vs. 22 patients)* - *Mixed individual/ group, and mixed day care/outpatient LTPP vs. non-evidence based control treatment*	
	Mentalisation based therapy with partial hospitalisation	TAU: general psychiatric outpatient care with medication prescribed by the consultant psychiatrist, community support from mental health nurses, and periods of partial hospital and inpatient treatment as necessary but no specialist psychotherapy
Dare	*Anorexia (21 vs. 20. vs. 22 vs. 19 patients)* - *Individual outpatient LTPP vs. three different non-evidence based control treatments*	
	Focal psychoanalytic psychotherapy	Cognitive-analytic therapy
		Family therapy
		Low contact ´routine´ treatment
Giesen-Bloo	*Borderline personality disorder (43 vs. 45 patients)* - *Individual outpatient LTPP vs. individual outpatient CBT*	
	Transference focussed therapy	Schema focussed therapy
Gregory	*Borderline personality disorder with alcohol use disorder (15 vs. 15)* - *Individual outpatient LTPP vs. TAU*	
	Dynamic deconstructive psychotherapy	TAU: remain in current treatment and/or referred to an alcohol rehabilitation centre and given names of psychiatric clinics and therapists in the community who might have openings and provide suitable treatment. Allowed to keep current psychotherapist, if any
Knekt	*Mood or anxiety disorder (128 vs. 101 vs. 97 patients)* - *Individual outpatient LTPP vs. individual outpatient STPP or vs. non-evidence based control treatment*	
	Psychodynamic psychotherapy	Short-term psychodynamic psychotherapy
		Solution-focused therapy
Linehan	*Borderline personality disorder* - *Individual outpatient LTPP vs. DBT*	
	Expert treatment in the community (84% of therapists described their methods as psychoanalytic or psychodynamic)	Dialectical behaviour therapy
Svartberg	*One or more cluster C personality disorders (25 vs. 25 patients* - *Individual outpatient LTPP vs. individual outpatient CT*	
	Dynamic psychotherapy	Cognitive therapy

Abbreviations: C(B)T: cognitive (behavioural) therapy; LTPP long-term psychodynamic psychotherapy; STPP: short-term psychodynamic psychotherapy; TAU: treatment as usual

Table 6 Frequency of treatment: mean number of sessions in participants *vs.* completers; a description of treatment intensity and a proxy of treatment intensity in completers

Study	Mean number of sessions (SD): participants		Mean number of sessions (SD): completers		Description of treatment intensity		Proxy of treatment intensity: completers (number of sessions/time) [a]		Session ratio (sessions in intervention group/ sessions in control group)
	I	C	I	C	I	C	I	C	
Bachar									
- Anorexia group	NA	NA	NA	NA	Weekly sessions for 1 year	Weekly sessions for 1 year	50/1 year	50/1 year	1.0
- Bulimia group (Intervention *vs.* cognitive orientation treatment and nutritional counselling)	NA	NA	NA	NA	Weekly sessions for 1 year + weekly or bi-weekly nutritional counselling for 6 months	Weekly sessions for 1 year+ weekly or bi-weekly nutritional counselling for 6 months	68/1 year	68/1 year	1.0
- Bulimia group (Intervention *vs.* nutritional counselling)	NA	NA	NA	NA	Weekly sessions for 1 year + weekly or bi-weekly nutritional counselling for 6 months	Weekly or bi-weekly nutritional counselling for 6 months	68/1 year	18/1 year	3.8
Bateman									
	NA	NA	NA	NA	18 months of individual and group mentalisation based therapy in a partial hospital setting	Duration and intensity not predefined	302/1.5 years [b]	165/1.5 years [c]	1.8
Dare									
- Intervention *vs.* CAT	24.9 (13.0)	12.9 (70)	NA	NA	1 year weekly sessions	7 months with 20 weekly sessions and thereafter 3 monthly sessions	52/1 year	23/ 1 year	2.3
- Intervention *vs.* FT	24.9 (13.0)	13.6 (8.6)	NA	NA	1 year weekly sessions	1 year of sessions each 1-3 weeks	52/1 year	26/1 year	2.0
- Intervention *vs.* TAU	24.9 (13.0)	10.9 (0.5)	NA	NA	1 year weekly sessions	1 year of low-contact sessions	52/1 year	16/1 year [d]	3.3
Giesen-Bloo									
	NA	NA	231 (NA)	190 (NA)	3 years of twice-weekly sessions	3 years of twice-weekly sessions	231/3 years	190/3 years	1.2
Gregory									

Study	Mean number of sessions (SD): participants		Mean number of sessions (SD): completers		Description of treatment intensity		Proxy of treatment intensity: completers (number of sessions/time) [a]		Session ratio (sessions in intervention group/ sessions in control group)
	I	C	I	C	I	C	I	C	
	NA	NA	3.57 (1.22)[e]	6.11 (7.12) [e]	1 year weekly sessions	Duration and intensity not predefined	42.8/1 year [e]	73.3/1 year [e]	0.6
Linehan									
	[f] 33.0	80.5 [f]	NA	NA	At least 1 session weekly for 1 year	Weekly individual and group session with telephone consultations as needed	33 [f]	80.5 [f]	0.4
Knekt									
- *Vs.* short-term psychodynamic psychotherapy	232 (105)[g]	18.5 (3.4)[g]	NA	NA	2 or 3 sessions weekly for up to 3 years	20 weekly treatment sessions	235 (104)/3 years [h]	46.9 (61.9)/3 years [h]	5.0
- *Vs.* solution-focused therapy	232 (105)[g]	9.8 (3.3)[g]	NA	NA	2 or 3 sessions weekly for up to 3 years	≤ 12 sessions ≤ 8 months	235 (104)/3 years [h]	29.9 (43.9)/3 years [h]	7.9
Svartberg									
	NA	NA	40	40	40 weekly sessions	40 weekly sessions	40/1 year	40/1 year	1.0

[a] The proxy of treatment intensity is based on the predetermined number of sessions participants were to receive, multiplied by the attendance rate when available
[b] Treatment for the partially hospitalized group consisted of 1) once-weekly individual psychoanalytic psychotherapy, 2) thrice weekly group analytic psychotherapy (1 hour each), 3) once-a-week expressive therapy oriented toward psychodrama techniques (1hour), and 4) a weekly community meeting (1 hour), all spread over 5 days. In addition, on a once-per-month basis, subjects had 5) a meeting with the case administrator (1 hour) and 6) medication review by the resident psychiatrist. The average length of stay was 1.45 years. Attendance at the program's psychotherapy sessions was 62%. Thus, participants were offered 6 weekly sessions and 2 monthly sessions. Across 1.45 years this is (6x52x1.45)+(24x1.45)=487 sessions, of which they participated 62%=302 sessions
[c] Treatment for the control group consisted of 1) Regular psychiatric review with a senior psychiatrist when necessary (on average, twice per month); 2) inpatient admission as appropriate (admission rate=90%, average stay=11.6 days), with discharge to non- psychoanalytic psychiatric partial hospitalization focusing on problem solving (72% were partially hospitalized, with an average length of stay of 6 months); followed by 3) outpatient and community follow-up (100%, every-2-week visits by a community psychiatric nurse) as standard aftercare. We assumed that treatment intensity during partial hospitalization was comparable to the intervention group, including the attendance rate. Thus, across 1,5 years there would have been 1) 36 sessions with a senior psychiatrist; 2) 6 months partial hospitalization with 156 sessions of which 62% were attended=104 sessions and 3) on average 1 year of every-2-weeks visits=25. 1)+2)+3)=165 sessions
[d] Calculation based on the mean number of sessions in all participants in this TAU group=12.9 sessions x 19 persons=207 sessions, divided by the number of completers in the TAU group: 207/13=16 sessions
[e] Mean number of paid contact hours per month. The proxy of treatment intensity is calculated by multiplying the mean number of sessions per month by 12
[f] The median number of sessions belonging to the study´s intervention and control treatment is given. The proxy of treatment intensity in completers is this median number of treatment sessions
[g] Mean number of sessions in people that started the assigned therapy
[h] mean number of sessions, including auxiliary therapies, in patients starting with therapy
Abbreviations: C: control group; CAT: cognitive-analytical therapy; FT: family therapy; I: intervention group; NA: not available; SD: standard deviation; TAU: treatment as usual

8.a.6 Treatment coherence

We considered a treatment coherent when it could be – and was – described in a manual or protocol, and when the therapists were supervised and especially trained for the intervention. In 3 studies the intervention treatment could be labelled coherent (Giesen-Bloo, Gregory and Svartberg) and in 2 of these studies the control treatment was coherent as well (Giesen-Bloo and Svartberg) (Table 7).

8.a.7 Co-interventions

Treatment confounders were present in all studies (Table 8). Medication was generally allowed as a co-intervention - and prescribed. One study (Dare) did not monitor the use of psychotropic medication. The Bateman study reported a much higher use of psychotropic medication in the control group (78% of participants used psychotropic medication at 1.5 years) *vs.* the intervention group (38%). We did not attempt to quantify the use of medication as most studies did not report this in sufficient detail; especially information on the type of medication was missing. A typical description would be: ´patients were discouraged from taking psychotropic medication and none was prescribed by the therapists´.

Besides medication, other forms of therapy and support were in frequent use as well, though hardly any study monitored these. Some studies forbade the use of other kinds of therapy (Giesen-Bloo). The Gregory study gave a nice overview that made clear that both the intervention (dynamic deconstructive psychotherapy) and control (TAU) group received additional forms of treatment in high frequency. Three months after the start of the study 4/15 patients in the intervention group and 7/15 patients in the control group used professional group therapy. Self-help groups were used by 2/15 in the intervention group and 5/15 in the control group. In addition, alcohol counselling, separate medication management and case management were used in both the intervention and the control group (Gregory, 2008).

It seems practically impossible to control the use of additional or alternative treatments in an outpatient setting. Refusal to discontinue other active treatment was an exclusion criterion in the Svartberg study. However, some participants did use other treatments. 5/50 participants had used hypnotics and 3/50 had used antidepressants during (part of) the study (Svartberg, 2004).

Table 7 Treatment coherence: manuals, integrity and therapists´ training

Study	Manual		Supervision		Analysis of adherence		Therapists trained for intervention		Kind of therapists	
	I	C	I	C	I	C	I	C	I	C
Bachar										
	N	N	Y	Y	Y	Y	Y	Y	Same experience with eating disorders in both I and C groups: 7 residents in clinical psychology (MSc), 2 psychiatric social workers, 1 psychiatrist	
Bateman										
	Y	N	Y	N	Y	N	N	NA	Psychiatrically trained nurses from the hospital´s team without formal psychotherapy qualifications	Senior psychiatrists and community psychiatric nurses
Dare										
- *Vs.* CAT	N	N	Y	Y	N	N	N	N	1 psychologist, 1 social worker, 1 medical doctor. All experienced in psychodynamic therapy	4 Members of the eating disorder team
- *Vs.* FT	N	N	Y	Y	N	N	N	N	See above	Same therapists as intervention group
- *Vs.* TAU	N	N	Y	Y	N	N	N	N	See above	Trainee psychiatrists in their 2nd/3rd year who rotated each ½ year
Giesen-Bloo										
	Y	Y	Y	Y	Y	Y	Y	Y	3 PhD, 37 MSc, 4 Bachelors with postgraduate training (no between-group difference in training). All with previous experience with BPD (no between-group difference in experience)	
Gregory										
	Y	N	Y	N	Y	N	Y	N	1 Psychiatrist and 5 psychiatry residents in 3rd year of training	NA
Knekt										
- *Vs.* STPP	N	N	N	N	N	N	Y	Y	Mainly psychologists (~80%)with standard training in psychoanalytically oriented psychotherapy and a mean of around 17 years of experience	
- *Vs.* SFT	N	Y	N	N	N	Y	Y	Y	See above	Psychologists, physicians and social workers with a mean of 9 years experience
Linehan										
	N	Y	Y	Y	N	Y	NA	Y	Therapists considered experts in treating difficult clients. 56% had a PhD and 56% had >10 years of clinical experience after terminal degree	75% had a PhD and 25% had >10 years of clinical experience after terminal degree
Svartberg										
	Y	Y	Y	Y	Y	Y	Y	Y	3 Psychiatrists and 5 clinical psychologists with 9.2 years experience on average	6 Clinical psychologists with 11.2 years experience on average

Abbreviations: B: borderline personality disorder; C: control group; CAT: cognitive-analytical therapy; FT: family therapy; I: intervention group; LTPP; long-term psychodynamic psychotherapy; N: No; NA: not available; SFT: solution focused therapy; STPP: short-term psychodynamic psychotherapy; TAU: treatment as usual; Y: Yes

Table 8 Co-interventions: studies´ policy and description of use

Study	Policy on co-interventions	Described co-interventions	
		I	C
Bachar			
	NA	1 Patient on fluoxetine for 5 weeks and 1 patient hospitalised for 5 weeks	1 Patient on fluoxetine for 5 weeks
Bateman			
	Polypharmacy was discouraged	The initial types and dosage of medication were similar in both groups	
		At 1.5 years 38% were still taking medication	At 1.5 years 78% were still taking medication 90% was admitted in to hospital (mean stay 12 days) and 72% was partially hospitalized in psychiatric care (mean stay 6 months) 10/63 patients were hospitalised
Dare			
	NA	2/21 patients were hospitalised	NA
Giesen-Bloo			
	Only medication allowed	Psychotropic medication use similar in both groups	
Gregory			
	Allowed and monitored	Comparable: - Receiving case management - Professional group therapy - Participation in self-help groups - Number of psychotropic medications	
		Not comparable: - The intervention group received more individual psychotherapy or alcohol counselling - The control group received more separate medication management	
Knekt			
	Allowed and monitored	Comparable: - Psychotropic medication	
		Not comparable: - The two control groups received more psychotherapy as a co-intervention - No one in the solution-focussed therapy group was hospitalised *vs.* 5% in the other groups	
Linehan			
	Allowed and monitored	Comparable: - Use and types of psychotropic medication at pre-treatment and during first year of follow-up - Use of therapy outside of the study	
		Not comparable: - The use of psychotropic medication decreased significantly less in the intervention group during treatment - The intervention group had significantly more hospital admissions for psychiatric reasons	
Svartberg			
	NA	2 patients received additional psychotherapy, partly during follow-up 3 patients used antidepressants 3 patients used hypnotics	2 patients received additional psychotherapy, partly during follow-up 1 patient used antidepressants 2 patients used hypnotics

Abbreviations: C: control group; I: intervention group; NA: not available

8.b Quality of selected studies

We assessed the quality of each study using the Maastricht Amsterdam criteria (van Tulder, 1997) and eight criteria proposed by Cuijpers et al (Cuijpers, 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 are depicted in Table 9 and Table 10. The quality of selected studies was variable. Most importantly, the quality of internal validity was low with the ´best´ study (Linehan) scoring 6/9 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, Bateman, Svartberg) scoring zero points for internal validity according to the criteria proposed by Cuijpers et al.

All studies used randomisation to allocate treatment. Only 3/8 studies (Giesen-Bloo, Knekt, Linehan) described an adequate concealment of treatment allocation. Notably, only 2/8 studies (Gregory, Linehan) explicitly described the blinding of outcome assessors. Co-interventions, adverse events nor compliance were monitored systematically in most studies. Besides being a quality criterion, we considered the drop-out rates to be an outcome as well. All these characteristics are described elsewhere in more detail (8.a.5, 8.a.7. and 8.c.1).

In 6/8 (Bachar, Bateman, Dare, Knekt, Linehan, Svartberg) studies we saw a statistically even distribution of important patient characteristics between treatment groups; in two studies (Giesen-Bloo, Gregory) there was an uneven distribution. In the Giesen-Bloo study twice as many patients in the intervention group had had recent suicide plans, steps or attempts compared to the control group (76.2% vs. 38.6%, p=0.007). In the Gregory study 6/15 patients allocated to the control group had an additional diagnosis of bipolar disorder vs. 0/15 in the intervention group. The researchers assessed that ´bipolar disorder displayed no significant interactions with group on any primary or secondary outcome measure at any time interval´ (Gregory, 2008). This statement has to be interpreted with due caution, however. Notwithstanding the statistically even distribution of prognostic factors in the remaining 6/8 studies, we assessed two more studies (Bachar, Bateman) as having dissimilar prognostic factors at baseline. Statistically non-significant differences between treatment groups may hide true differences, if the number of participants in the study is small. Much larger numbers would be needed to check for true differences between groups. Thus we looked at an equilibrium in numbers or percentages. If the ratio of any prognostic factor was in the order of 2 or 0.5 we assessed the groups to have dissimilar prognostic factors.

Table 9 Quality criteria according to criteria used by Cuijpers et al (Cuijpers, 2009)

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

Abbreviations: N: no; Y: yes

Table 10 Quality of studies according to the Maastricht Amsterdam criteria (van Tulder, 1997)

		Study / Criterion	Bachar	Bateman	Dare	Giesen-Bloo	Gregory	Knekt	Linehan	Svartberg
Patient selection										
a		Where the eligibility criteria specified?	Y	Y	Y	Y	Y	Y	Y	Y
b		Treatment allocation								
	1	*Was a method of randomization performed?*	*Y*	*Y*	*Y*	*Y*	*Y*	*Y*	*Y*	*Y*
	2	*Was the treatment allocation concealed?*	*D*	*D*	*D*	*Y*	*D*	*Y*	*Y*	*N*
c		Were the groups similar at baseline regarding the most important prognostic factors?	N	N	Y	N	N	Y	Y	Y
Interventions										
d		Were the index and control interventions explicitly described?	Y	Y	Y	Y	Y	Y	N	Y
e		Was the care provider blinded to the intervention?	N	N	N	N	N	N	N	N
f		Were co-interventions avoided or comparable?	D	D	D	N	N	N	N	Y
g		Was the compliance acceptable in both groups?	D	Y	Y	N	D	Y	N	Y
h		Was the patient blinded to the intervention?	N	N	N	N	N	N	N	N
Outcome measurements										
I		Was the outcome assessor blinded to the intervention?	D	D	N	N	Y	N	Y	D
J		Were the outcome measures relevant?	Y	Y	Y	Y	Y	Y	Y	Y
k		Were adverse events described?	N	Y	Y	Y	Y	N	N	N
L		Was the withdrawal/dropout rate described and acceptable?	Y	Y	N	N	Y	N	Y	Y
m		Timing of follow-up measurements								
	1	*Was a short-term follow-up measurement performed?*	*Y*	*Y*	*Y*	*Y*	*Y*	*Y*	*Y*	*Y*
	2	*Was a long-term follow-up measurement performed?*	*N*	*Y*	*N*	*Y*	*N*	*Y*	*Y*	*Y*
n		Was the timing of the outcome assessment in both groups comparable?	Y	Y	Y	Y	Y	Y	Y	Y
Statistics										
o		Was the sample size for each group described?	Y	Y	Y	Y	Y	Y	Y	Y
p		Did the analysis include an intention to treat analysis?	N	N	Y	Y	Y	Y	Y	N
q		Were point estimates and measures of variability presented for the primary outcome measures?	Y	Y	N	Y	Y	Y	Y	Y
		Total Yes (maximum score 17)	**8**	**10.5**	**10**	**10**	**11**	**11**	**11**	**11.5**
-		Total internal validity (b, e, f, g, h, i, j, l ,n, p, maximum 10)	3.5	4.5	4.5	4	5.5	5	6	5.5
-		Total descriptive criteria (a, c, d, k, m, maximum 5)	2.5	4	4.5	4	3.5	4	3	4
-		Total statistical criteria (o, q, max imum 2)	2	2	1	2	2	2	2	2

Abbreviations: D: don´t know; N: no; Y: yes

8.b.1 Outcomes

Table 11 gives an overview of the outcomes for which data were available. The described outcomes of the intervention *vs.* control treatments are those for the longest follow-up available. In general, a higher score indicated a worse state, except for the QoL measures. There were three studies that extended their follow-up after the end of treatment (Bachar, Linehan, Svartberg). At a follow-up of 2 years Bachar does not report the actual data but stated *´a slight continued improvement occurred during the year following termination of therapy. This improvement was not significant in either of the two groups, nor was there a significant difference between groups´* (Bachar, 1999). We contacted the author to obtain data at the 2 year follow-up, but these data were not available.

Reporting of outcomes was not consistent across time points of follow-up. The Bateman study for example reported different outcome variables at 1.5 and 3 years *vs.* 8 years. Also, only the 3 year analysis was an ITT (Bateman, 1999;Bateman, 2001;Bateman, 2008). In addition, the delineation of time periods was confusing and inconsistent. Outcomes were reported that occurred (a) during the treatment – i.e. from 0 to 1.5 years (Bateman, 1999); (b) at a follow-up of 3 years - i.e. from 1.5 to 3 years (Bateman, 2001); and (c) at a follow-up of 8 years – i.e. from 3 to 8 years (Bateman, 2008). It was not possible to recalculate the total number of events during all 8 years of follow-up (including the treatment period) from the available data, so we only used the data at 1,5 years.

We did not meta-analyse the outcomes of ´suicide attempt´ or ´self mutilation´ in the Bateman study. We considered the effect of partial hospitalisation *vs.* outpatient treatment to have a larger impact than the effect of a psychoanalytically based therapy *vs.* TAU for these outcomes. The SD for the outcome on the Social Adjustment Scale (self report) was missing, so we could not use this in the meta-analysis.

In the Knekt study several measurement scales were used that could be considered measurements of target problems for subgroups of patients. E.g. the Beck Depression Inventory and the Hamilton Depression Rating Scale could be seen as measuring target problems for participants with a mood disorder. Similarly, the Symptom Check List-90-Anxiety subscale and the Hamilton Anxiety Rating Scale measure target problems for participants with an anxiety disorder. However, these instruments were used in all participants and not only in the participants with the pertaining disorder at baseline. Thus we used the below mentioned scales as outcome measures for general psychiatric symptoms.

Table 11 Outcomes of measures of psychopathology in included studies

Study	Domains	Tests	Outcome (intervention *vs.* control)
Bachar	*Reported at 1 year (post-treatment), no data at 2 years, completers analysis*		
	Recovery	Recovered (DSM-IV)	Significantly more recovered
	Target problems	DSM-SS	No significant difference in group means
		EAT 26	No significant difference in group means
	Symptoms	SCL-90 GSI	No significant difference in group means
	Personality	Selves Questionnaire	No significant difference in group means
Bateman	*Reported at 1 and 1.5 (post-treatment) years, completers analysis*		
	Symptoms	BDI	Significantly better scores
		SCL-90-R GSI	No significant difference in group means
		SCL-90-R PSTS	No significant difference in group means
		STAI-state	Significantly better scores
	Personality	STAI-trate	No significant difference in group means
	Social functioning	IIP	Significantly better scores
Dare	*Reported at 1 year (post-treatment), ITT analysis*		
	Recovery	Recovered (DSM-IV)	Significantly more recovered in intervention group *vs.* TAU*
Giesen-Bloo	*Reported at 1, 2 and 3 years (post-treatment) , ITT analysis*		
	Recovery	Recovered (BPD Severity Index)	Significantly less recovered

Study	Domains	Tests	Outcome (intervention *vs.* control)
	Target problems	Reliable clinical improvement (BPD Severity Index)	Significantly less improved
	Quality of Life	EuroQol thermometer	No significant difference in group means
		WHOQoL	No significant difference in group means
	Symptoms and personality composite score	BPD Severity Index, SCL-90, Rosenberg Self-Esteem Scale, Miskimins Self-Goal Discrepancy Scale, YSQ, PDBQ, IPO, DSQ-48.	Significantly less improved
Gregory	*Reported at 1 year (post-treatment), ITT analysis*		
	Recovery	Recovered (ASI criteria)	
	Target problems	Alcohol misuse	No significant difference in group %
		BEST	No significant difference in group means
		Parasuicidal behaviour	No significant difference in group %
	Symptoms	Institutional care	No significant difference in group %
		BDI	No significant difference in group means
		DES	No significant difference in group means
	Social functioning	SPS	No significant difference in group means
Linehan	*Reported at 1 year (post-treatment) and 2 years, ITT analysis*		
	Target problems	Suicide attempt or self injury	Significantly more frequent
		Suicide ideation	No significant difference in group %
	Symptoms	RLI	
		- Mean total item score	No significant difference in group means
		- Survival and coping	No significant difference in group means
		HAMD	No significant difference in group means
		Emergency department visits	
		- For psychiatric reason	Significantly more visits
		- For suicide ideation	Significantly more visits
		Hospital admissions	
		- For psychiatric reason	Significantly more admissions
		- For suicide ideation	Significantly more admissions
Knekt	*Reported at 1 and 3 years (post-treatment), ITT analysis*		
	Recovery	Recovery from Axis I diagnosis	No significant difference in group %
		- Mood disorder	No significant difference in group %
		- Major depressive disorder	No significant difference in group %
		- Anxiety disorder	Significantly more recovered *vs.* 2 control groups
	Symptoms	BDI	Significantly better score *vs.* 2 control groups
		HAMD	Significantly better score *vs.* 2 control groups
		HAMA	Significantly better score *vs.* 1 control group
		SCL-90-Anx	Significantly better score *vs.* 2 control groups
		SCL-90-GSI	Significantly better score *vs.* 1 control group
Svartberg	*Reported at 1 (post-treatment) and two years (completers analysis)*		
	Recovery	Recovered (Millon Clinical Multiaxial Inventory)	No significant difference
	Target problems	Millon Clinical Multiaxial Inventory (total mean raw score)	No significant difference in group means
	Symptoms	SCL-90-R GSI	No significant difference in group means
	Social functioning	IIP	No significant difference in group means

* No comparison between the intervention group and two other non-evidence based control groups are made in the original publication

$ A composite score of several tests

Abbreviations: ASI: addiction severity index; BEST: Borderline Evaluation of Severity over Time; BDI: Beck Depression Inventory; BPD: borderline personality disorder; DES: Dissociative Experiences Scale; DSM-III-R, Diagnostic and Statistical Manual of Mental Disorders (Third Edition Revised); DSM –SS: DSM Symptomatology Scale for Anorexia and Bulimia; DSQ: Defence Style Questionnaire; GAF: Global Assessment of Functioning Scale; GSI: global severity index scale; IIP: Inventory of Interpersonal Problems; IPO: Inventory of Personality Organisation; HAMA: Hamilton Anxiety Rating Scale; HAMD: Hamilton Depression Rating Scale; ITT: intention to treat; PD: personality disorder; PDBQ: Personality Disorder Belief Questionnaire; PSTS: positive symptom total score; QoL: quality of life; RLI: Reasons for Living Inventory; SAS-SR: social adjustment scale self report; SCL-90-R: Symptom Check List-90 Revised; SPS: Social Provisions scale; STAI: (Spielberger) State-Trait Anxiety Inventory; TSIA and TSIAI: severity for all target objectives and most important objective; TSP & TSPI, severity for all target objectives and most important objective; TST & TSTI: severity for all target objectives and most important objective; TAU: treatment as usual; WHO: World Health Organisation; Young Schema Questionnaire

8.c Effect of interventions

8.c.1 Non-completers and adverse events

In Table 12 an overview of the non-completers, cross-overs and adverse events is given. The difference in dropout rates was statistically significant in the Giesen-Bloo, Knekt and Linehan studies. In all three studies there were twice as many non-completers in the intervention groups, compared to the control groups. The differences in proportions of non-completers were:

a) 50.0% (95%CI 34.9-65.1%) *vs.* 25.0% (95%CI 12.2-37.8%) in the Giesen-Bloo study (p=0.017)
b) 36.7% (95%CI 28.4-45.1%) *vs.* 15.5% (95%CI 8.3-22.7%) - in SFT group – and *vs.* 12.9% (95%CI 6.3-19.4%) - in the STPP group – in the Knekt study (p<0.001)
c) 42.9% (95%CI 29.0-56.7%) *vs.* 19.2% (95%CI 8.5-29.9%) in the Linehan study (p=0.01)

Adverse events were not monitored systematically in the selected studies. Although the delineation between an adverse event and a sequel of the target problem is not always clear, this certainly deserves more attention in future trials.

Table 12 Non-completers*, cross-overs and adverse events

Study	Non-completers						Cross-overs	Adverse events
	I	(%)	C	(%)	Timing of dropout	Reason for dropout		
Bachar								
	3/17	17.6	5/17	29.4	<5 w	NA	No	NA
Bateman								
	3/22	13.6	3/22	13.6	<6 m	NA	3/22 C->I	3 suicide attempts in C group (these crossed-over)
Dare								
- vs. CAT	9/21	42.9	9/22	40.9	1/3 <2 m 2/3 >2 m	NA	No	12/84 required hospitalisation 1 death TAU group
- vs. FT			6/22	27.3				
- vs. TAU			6/19	31.6				
Giesen-Bloo								
	21/42 †	50.0	11/44	25.0	I: 1/3 < 4 m 73% in first y C: spread evenly across 3 y	In both groups 50% had no faith in therapy or therapist. The other 50% had various reasons	No	No
Gregory								
	5/15	33.3	6/15	40.0	Spread out evenly across treatment y	NA	No	1 incarcerated in I group 1 death in C group (suicide)
Knekt								
- vs. SFT	47/128 †	36.7	15/97	15.5	I: 20.3% of participants never started LTPP C: 4.1% never started therapy	I: 11.7% of participants objected to type of therapy. Rest had various reasons C: various reasons	No	NA
- vs. STPP			13/101	12.9	C: 3.0% never started therapy	C: various reasons		
Linehan								
	21/49 †	42.9	10/52	19.2	I: 1st therapists dropped at a median of 9.7 w C: 1st therapists dropped at a median of 16.9 w	NA	No	NA
Svartberg								
	1/26**	3.8**	1/26**	3.8**	NA	NA	No	NA

* Non-completers are defined as the participants that discontinued their treatment or were lost to follow-up during the treatment duration. Participants lost to follow-up in the post treatment period are excluded
** One patient out of a total of 51 randomised patients dropped out but it was unclear from which group
Abbreviations: C: control group; CAT: cognitive analytical therapy; CI: confidence interval; F: family therapy; I: intervention group; m: month; n: number; NA: not available; TAU: treatment as usual; w: week; y: year
† Statistically significant difference in proportions

8.c.2 Recovery

Six studies (Bachar, Dare, Giesen-Bloo, Gregory, Knekt, Svartberg) gave information on the number of patients that recovered. The main characteristics and the criteria for recovery pertaining to each study are shown in Table 13. All six studies gave data on recovery from the targeted disorder, except the Gregory and Knekt studies. Gregory examined patients with a 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 Hedges´ g for recovery from anxiety disorder and for recovery from mood disorder (both in patients who had the disorder at baseline).

Four out of six studies reported an ITT analysis and the two studies that did not (Bachar and Svartberg) had non-differential drop-out rates (see also Table 12). The Gregory study was the only study in which blinded outcome assessors were used. Even so, the outcome assessor was able to correctly guess the treatment allocation 67% of the time (50% correct guesses expected by chance alone) (Gregory, 2008).

Table 13 Studies with data on recovered patients

Study, Disorder	Recovery criterion	Instrument(s) used to measure criterion	Type of instrument	Independent assessor	Blinded assessor
Bachar					
Eating disorder	No longer meets DSM-IV diagnostic criteria	NA	NA	D	D
Dare					
Eating disorder	No longer meets DSM-IV diagnostic criteria	NA	NA	D	N
Giesen-Bloo					
BPD	Achieving a BPDSI-IV score <15	BPDSI	Semi structured interview	Y	N
Gregory					
BPD+alcohol misuse	No longer meets alcohol misuse criteria ASI	ASI	Structured interview	Y	Y
Knekt					
Mixed	No longer meets criteria for DSM-IV mood disorder	Semi structured interview	Semi structured interview	Y*	N
	No longer meets criteria for DSM-IV anxiety disorder	Semi structured interview	Semi structured interview	Y*	N
Svartberg					
Cluster C personality disorder	MCMI<74	MCMI	Questionnaire	D	D

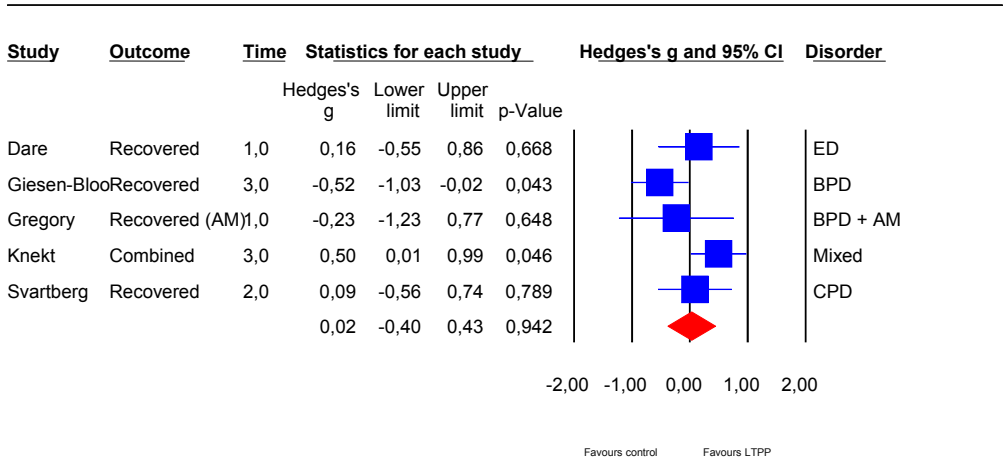
* The interviews were carried out by experienced clinical raters at a separate location, so it is likely these raters were independent

Abbreviations: ASI: Addiction Severity Index; BPD: borderline personality disorder; BPDSI-IV: Borderline Personality Disorder Severity Index, fourth version; D: don´t know; DSM-IV: Diagnostic and Statistic Manual of Mental Disorder, fourth version; MCMI: Million Clinical Multiaxial Inventory; N: no; NA: not available; R: Randomised Clinical Trial; TAU: treatment as usual; Y: yes

To perform meta-analysis on the recovery rates, we combined the bulimia and anorexia patients in the Bachar study and left out the control group that only received nutritional counselling. Bachar was left out of the analysis on the longest follow-up available, because no data were reported at 2 years follow-up. It was only stated that there was no significant difference between treatment groups at that time. 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 combined Hedges´ g for recovery at the longest available follow-up for each study was 0.02 (95% CI: -0.40 to 0.43; p=0.94;I-squared: 52.8%; n=5). These findings were robust 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 combined Hedges´ g for recovery at a follow-up of 1 year was 0.06 (95% CI: -0.42 to 0.53; p=0.81; I-squared=52.8%;n=5).

Figure 2 Combined Hedges´ g for recovery at longest available follow-up



Heterogeneity				Tau-squared			
Q-value	df (Q)	P-value	I-squared	Tau Squared	Standard Error	Variance	Tau
8,480	4	0,076	52,829	0,114	0,156	0,024	0,337

Abbreviations: AM: alcohol misuse; BPD: borderline personality disorder; CPD: cluster C personality disorder; CI: confidence interval; LTPP: long-term psychodynamic 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

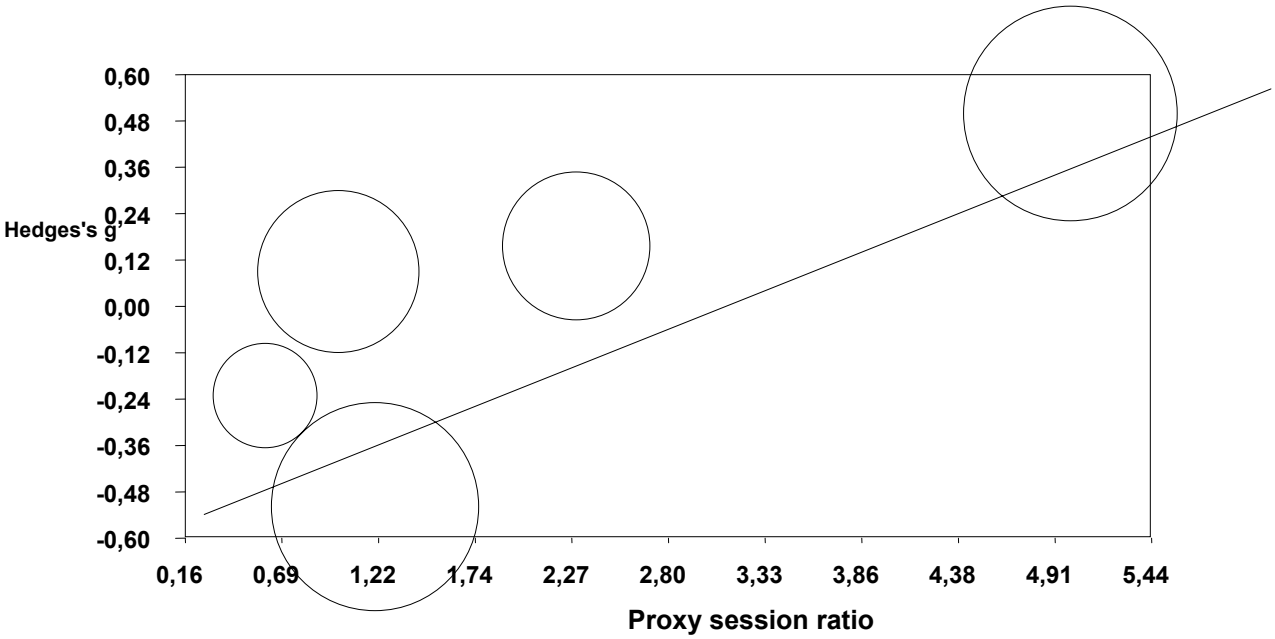
Subgroup and meta-regression analyses

Only one study reported the recovery of patients with BPD. Though the Gregory study was undertaken in BPD patients, only the recovery from concomitant alcohol misuse was assessed. The combined Hedges' g for recovery in eating disorder patients (Bachar and Dare studies, outcomes only available at 1 year) was 0.59 (95% CI -0.38 to 1.56; p=0.23; I-squared=61.2%;n=2).

There was no difference between studies that reported an adequately concealed treatment allocation *vs.* studies where the treatment allocation concealment was not reported or inadequate. Similarly, whether the outcome assessors were or were not blinded to the treatment allocation made no difference (though only one study that reported on recovery used blinded outcome assessors).

Exploratory meta-regression indicated that the proxy session ratio might be a predictor of the effect size of recovery (B=0.19; 95%CI: 0.03 to 0.34; p=0.02;T-squared=0.00;n=5)(Figure 3). The internal validity score did not predict the Hedges´ g for recovery (B=0.23; 95% CI: -0.34 to 0.81; p=0.43;T-squared=0.11;n=5).

Figure 3 Regression of proxy session ratio on Hedges´ g for recovery at the longest available follow-up for each study

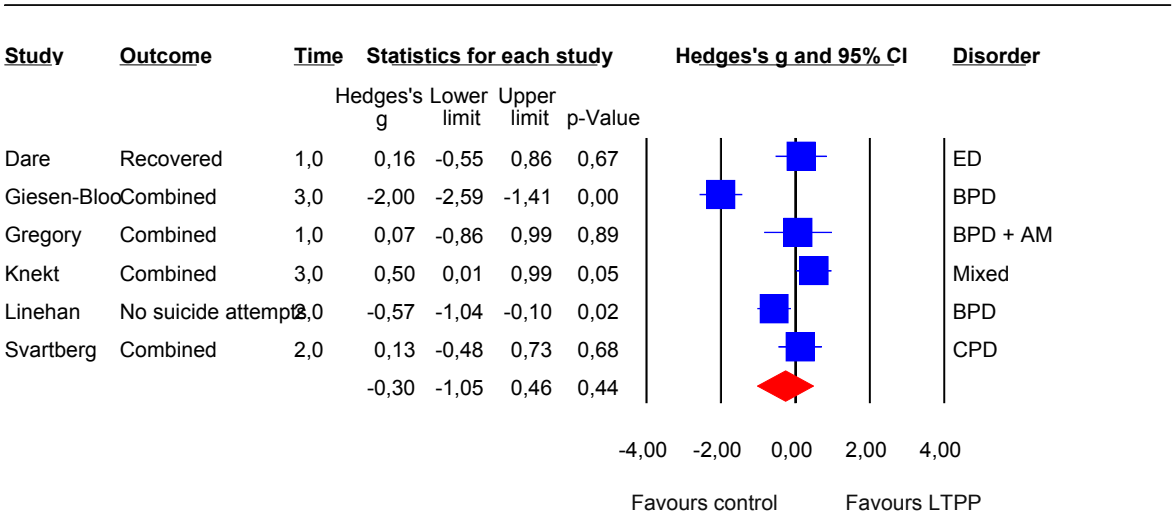


8.c.3 Target problems

We combined the bulimia and anorexia patients in the Bachar study and left out the control group that only received nutritional counselling. Bachar was left out of the analysis on the longest follow-up available, because no data were reported at 2 years follow-up. It was only stated that there was no significant difference between treatment groups at that time. 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 combined effect size (Hedges´ g) for target problems (the mean Hedges´ g for all outcomes on target problems available per study) was -0.30 (95% CI -1.05 to 0.46; p=0.44; I-squared=89.4%; n=6) (at the longest outcome available for each study)(Figure 4). These 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. The combined Hedges´ g for target problems at a follow-up of 1 year was -0.26 (95% CI: -0.96 to 0.44; p=0.47; I-squared=85.6%;n=6).

Figure 4 Combined Hedges´ g for target problems at longest available follow-up



Test of null (2-Tail)		Heterogeneity				Tau-squared			
Z-value	P-value	Q-value	df (Q)	P-value	I-squared	Tau Squared	Standard Error	Variance	Tau
-2,574	0,010	47,298	5	0,000	89,429	0,780	0,582	0,339	0,883
-0,770	0,442								

Abbreviations: AM: alcohol misuse; BPD: borderline personality disorder; CPD: cluster C personality disorder; ED: eating disorder; LTPP: long-term psychoanalytical psychotherapy.

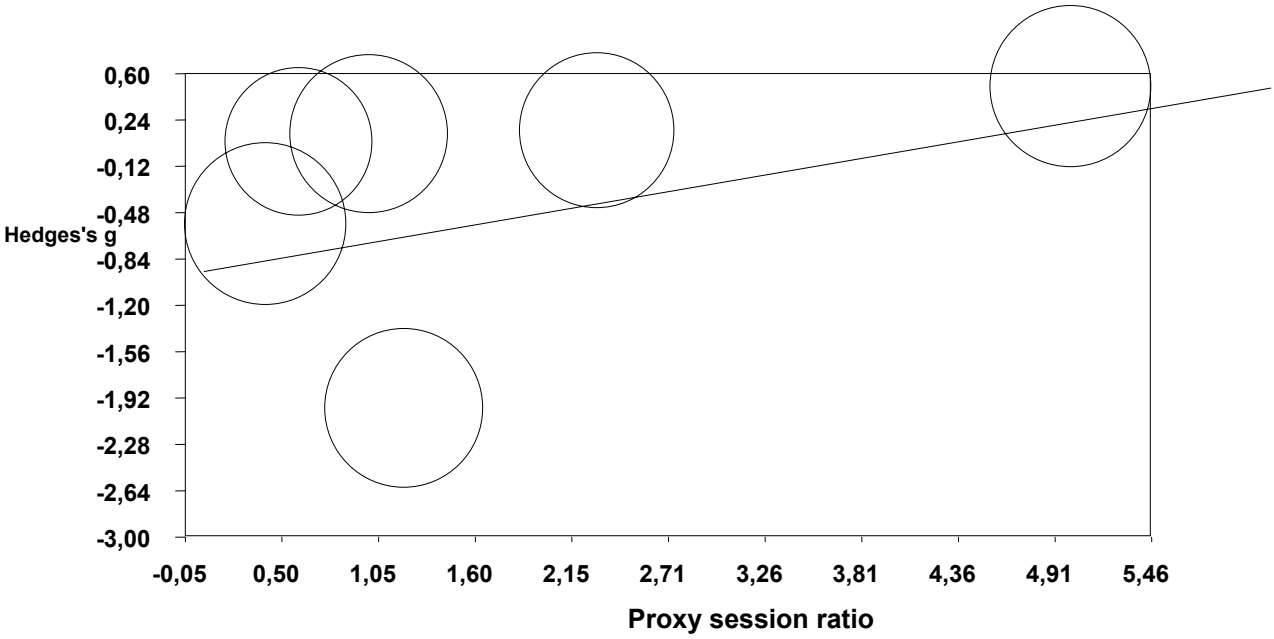
Subgroup and meta-regression analyses

The combined Hedges' g for target problems in BPD patients (Giesen-Bloo, Gregory and Linehan studies) was -0.87 (95% CI -2.00 to 0.27; p=0.14; I-squared=89.6%;n=3) (at the longest available follow-up). The combined Hedges' g for target problems in eating disorder patients (Bachar and Dare studies) was 0.31 (95% CI -0.24 to 0.85; p=0.27; I-squared=0.0%;n=2) (outcomes only available at 1 year).

There was no difference between studies that reported an adequately concealed treatment allocation vs. studies where the treatment allocation concealment was not reported or inadequate. Similarly, whether the outcome assessors were or were not blinded to the treatment allocation made no difference, though only one study that reported on target problems used blinded outcome assessors.

Exploratory meta-regression indicated that the proxy session ratio was not a predictor of the effect size in the domain target problems (B=0.24; 95%CI: -0.23 to 0.70; p=0.32;T-squared:0.76;n=6)(Figure 5). Similarly, the internal validity score did not predict the Hedges´ g for target problems (B=0.72; 95% CI: -0.33 to 1.77; p=0.18;T-squared:0.68 ;n=6).

Figure 5 Regression of the proxy session ratio on Hedges´ g for target problems at the longest available follow-up for each study

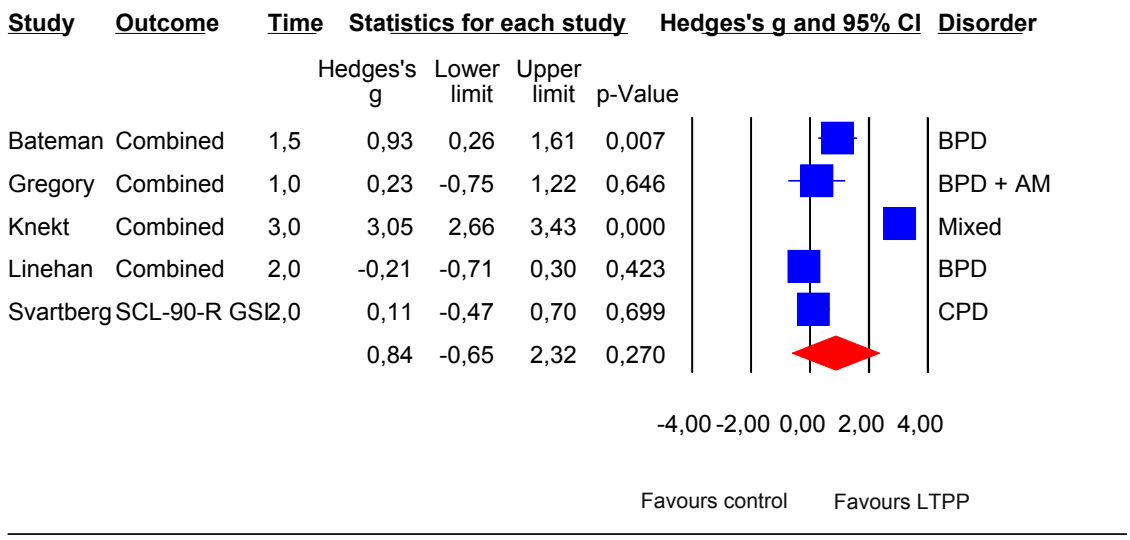


8.c.4 General psychiatric symptoms

We combined the bulimia and anorexia patients in the Bachar study and left out the control group that only received nutritional counselling. Bachar was left out of the analysis on the longest follow-up available, because no data were reported at 2 years follow-up. It was only stated that there was no significant difference between treatment groups at that time.

The combined effect size (Hedges´ g) for general psychiatric symptoms was 0.84 (95% CI -0.65 to 2.32; p=0.27; I-squared=97.0%;n=5) (at the longest outcome available for each study)(Figure 6). These findings were robust when we (a) removed each study arbitrarily; (b) used different control groups for the Knekt study; or (c) did include the Bachar data at 1 year. The combined Hedges´ g for symptoms at a follow-up of 1 year was -0.22 (95% CI: -1.04 to 0.60; p=0.60; I-squared=91.6%;n=5). This ES has overlapping 95%CI with the ES of 0.84 at the longest available follow-up. The differences of effect direction are due to the Knekt study wherein at a follow-up of 1 year the results in the control group were significantly better compared to the LTPP group, and vice versa.

Figure 6 Combined Hedges´ g for general psychiatric symptoms at longest available follow-up



Heterogeneity				Tau-squared			
Q-value	df (Q)	P-value	I-squared	Tau Squared	Standard Error	Variance	Tau
135,505	4	0,000	97,048	2,765	2,236	4,999	1,663

Abbreviations: AM: alcohol misuse; BPD: borderline personality disorder; CPD: cluster C personality disorder; ED: eating disorder; LTPP: long-term psychoanalytical psychotherapy.

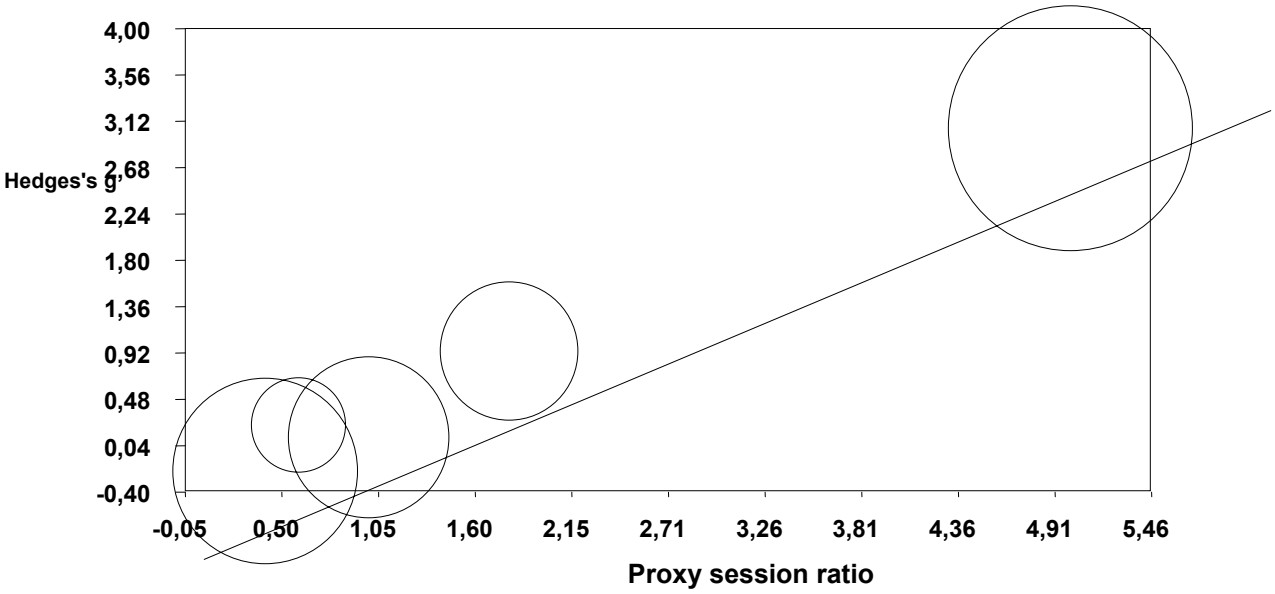
Subgroup and meta-regression analysis

The combined Hedges' g for general psychiatric symptoms in BPD patients (Bateman, Gregory and Linehan studies) was 0.30 (95% CI -0.45 to 1.06; p=0.44; I-squared=74.4%; n=3). Only one study (Bachar) reported symptom scores in eating disorder patients.

There was no difference between studies that reported an adequately concealed treatment allocation *vs.* studies where the treatment allocation concealment was not reported or inadequate. Similarly, whether the outcome assessors were or were not blinded to the treatment allocation made no difference (though only one study that reported on target problems used blinded outcome assessors).

Exploratory meta-regression showed that the proxy session ratio might be a predictor for the effect size in the domain general psychiatric symptoms (B=0.70; 95%CI: 0.58 to 0.82;p<0.00; T-squared=0.00;n=5)(Figure 7). The internal validity score did not predict the Hedges´ g for general psychiatric symptoms (B=-1.22; 95% CI: -3.94 to 1.49; p=0.38; T-squared=2.38;n=5).

Figure 7 Regression of proxy session ratio on Hedges´ g for symptoms at the longest available follow-up for each study



8.c.5 Personality pathology

We could not estimate a combined effect size for personality pathology at the longest available follow-up, as only one study (Bateman) provided this outcome. The combined Hedges´ g for personality pathology at a follow-up of 1 year was 0.22 (95% CI: -0.26 to 0.71; p=0.36; I-squared=0.0%: n=2) (Bachar and Bateman studies).

Subgroup and meta-regression analysis

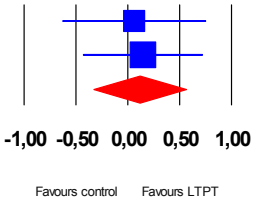
Too few studies were available to perform subgroup or meta-regression analyses.

8.c.6 *Social functioning*

The combined effect size for social functioning was 0.11 (95% CI -0.34 to 0.56; p=0.63; I-squared=0.0%; n=2) (at the longest outcome available for each study)(Figure 8). The combined Hedges´ g for social functioning at a follow-up of 1 year was 0.17 (95% CI: -0.27 to 0.61; p=0.45; I-squared=0.0%;n=2) (Gregory and Svartberg studies).

Figure 8 Combined Hedges´ g for social functioning at longest available follow-up

Study	Outcome	Time	Statistics for each study				Hedges's g and 95% CI
			Hedges's g	Lower limit	Upper limit	p-Value	
Gregory	SPS	1,0	0,06	-0,64	0,76	0,863	
Svartberg	IIP	2,0	0,15	-0,44	0,73	0,622	
			0,11	-0,34	0,56	0,625	



Test of null (2-Tail)		Heterogeneity				Tau-squared			
Z-value	P-value	Q-value	df (Q)	P-value	I-squared	Tau Squared	Standard Error	Variance	Tau
0,489	0,625	0,034	1	0,854	0,000	0,000	0,152	0,023	0,000
0,489	0,625								

Subgroup and meta-regression analysis

There were too few studies available to perform subgroup or meta-regression analyses.

8.c.7 *Quality of life*

Only one study reported quality of life data (Giesen-Bloo).

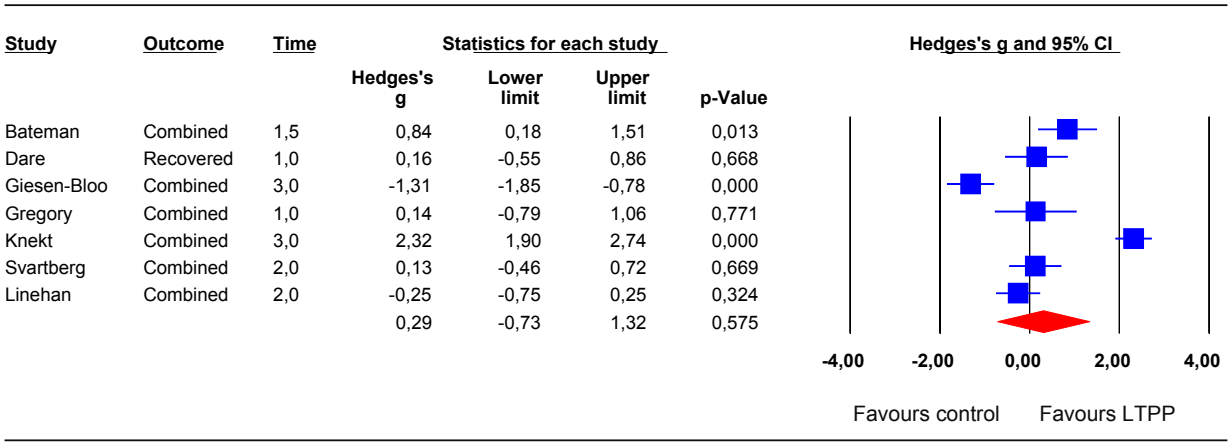
8.c.8 *Overall effectiveness*

We combined the bulimia and anorexia patients in the Bachar study and left out the control group that only received nutritional counselling. Bachar was left out of the analysis on the longest follow-up available, because no data were reported at 2 years follow-up. It was only stated that there was no significant difference between treatment groups at that time. 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 combined effect size (Hedges´ g) for overall effectiveness (the mean Hedges´ g for all outcomes available per study) was 0.29 (95% CI -0.73 to 1.32; p=0.58; I-squared=95.4%;n=7) (at the longest outcome available for each study)(Figure 9Figure 4). These 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. The combined Hedges´

g for overall effectiveness at a follow-up of 1 year was -0.06 (95% CI: -0.53 to 0.41; p=0.79; I-squared=80.9%; n=8).

Figure 9 Combined Hedges´ g for overall effectiveness at longest available follow-up



Study	Outcome	Time	Statistics for each study			
			Hedges's g	Lower limit	Upper limit	p-Value
Bateman	Combined	1,5	0,84	0,18	1,51	0,013
Dare	Recovered	1,0	0,16	-0,55	0,86	0,668
Giesen-Bloo	Combined	3,0	-1,31	-1,85	-0,78	0,000
Gregory	Combined	1,0	0,14	-0,79	1,06	0,771
Knekt	Combined	3,0	2,32	1,90	2,74	0,000
Svartberg	Combined	2,0	0,13	-0,46	0,72	0,669
Linehan	Combined	2,0	-0,25	-0,75	0,25	0,324
			0,29	-0,73	1,32	0,575

Test of null (2-Tail)		Heterogeneity				Tau-squared			
Z-value	P-value	Q-value	df (Q)	P-value	I-squared	Tau Squared	Standard Error	Variance	Tau
4,460	0,000	129,658	6	0,000	95,372	1,813	1,180	1,393	1,346
0,561	0,575								

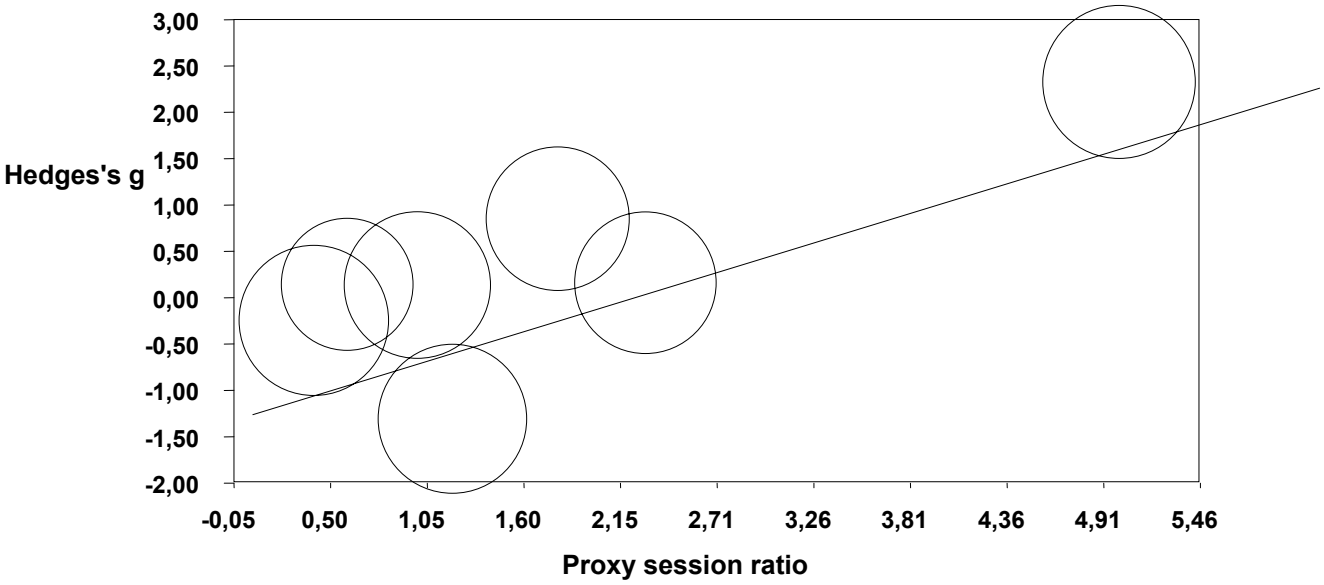
Subgroup and meta-regression analyses

The combined Hedges' g for overall effectiveness in BPD patients was 0.05 (95% CI –0.52 to 0.62; p=0.87; I-squared=67.0%;n=4) (at the longest available follow-up). The combined Hedges' g for overall effectiveness in eating disorder patients (Bachar and Dare studies) was 0.27 (95% CI -0.26 to 0.81; p=0.31; I-squared=0.0%; n=2) (outcomes only available at 1 year).

There was no difference between studies that reported an adequately concealed treatment allocation *vs.* studies where the treatment allocation concealment was not reported or inadequate, for the outcome overall effectiveness. Similarly, whether the outcome assessors were or were not blinded to the treatment allocation made no difference.

Exploratory meta-regression indicated that the proxy session ratio might be a predictor of the effect size for overall effectiveness (B=0.58; 95%CI: 0.21 to 0.96; T-squared=0.48; p<0.00) (Figure 10). The internal validity score did not predict the Hedges´ g for overall effectiveness (B=0.30; 95% CI: -1.38 to 1.99; T-squared=2.12; p=0.73).

Figure 10 Regression of proxy session ratio on Hedges´ g for overall effectiveness at the longest available follow-up for each study



8.c.9 Publication bias

Visual inspection showed an asymmetric funnel plot for the outcome target problems, but not for the outcomes recovery, symptoms or overall effectiveness. Using Duval and Tweedie´s trim and fill test we did not find formal evidence of publication bias for the outcomes recovery, target problems or symptoms. However, because of the small number of studies we feel we cannot draw a formal/statistical conclusion on the existence of publication bias in our review.

9 Discussion

9.a Study quality and assessment of bias

The overall quality of studies was reasonable. All included studies used randomisation to allocate treatment, but only 3/8 studies described an adequate concealment of treatment allocation. Notably, only 2/8 studies explicitly described the blinding of outcome assessors. Subgroup analyses for studies with or without adequate concealment of treatment allocation, and for studies with or without blinded assessors showed no differential results. The internal validity score did not predict effect size on any of the outcomes. We could not delineate a clear ´best category´ of studies, to perform a subgroup analysis on. Short follow-up, or only reporting outcomes at the end of treatment seems curious for a treatment that takes so long.

Sometimes a biased publication of study outcomes occurred. E.g. Bachar et al. described their outcomes at 1 year follow-up at length, while the results at a follow-up of 2 years merited one paragraph only. At a follow-up of one year, the patients with a psychodynamic therapy improved significantly more than patients in a control group. Results at a follow-up of 2 years were available but not reported except as ´in the two psychotherapy groups [*self psychological treatment vs. cognitive orientation treatment, YS*] on all three outcome measures, a slight continued improvement occurred during the year following termination of therapy. This improvement was not significant in either of the two groups, nor was there a significant difference between groups´ (Bachar, 1999).

Control conditions were heterogeneous and mostly of low quality, e.g. non-evidence-based treatments or TAU. 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*. Future studies should compare LTPP to other highly specialised treatments that are equally intensive, like state-of-the art CBT in case of eating disorders, or SFT and DBT in case of BPD.

In the existing meta-analyses almost all included studies were undertaken by researchers with a LTPP background. There is a risk of overoptimistic findings in such studies, because a mix of backgrounds within the research groups is missing. This risk can be counterbalanced by including trials where LTPP is used as a control.

9.b Heterogeneity, interaction and confounding

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 the small number of studies precluded a meaningful analysis of subgroups, and severely limited meta-regression. With so few data points available, both false-positive and false-negative findings can be expected. Thus, we consider the meta-regressions we performed exploratory only. 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 predict the effect size when the proxy session ratio equals 1 (indicating the same number of sessions in the intervention group and the control group).

At first we intended to meta-analyse between-group differences in means, of scores on tests that measure the outcome in target problems. This would have minimised heterogeneity from different outcome assessment tests used. For example, it would be interesting to compare the

difference of mean scores (intervention *vs.* control) in the HAMD (a score test for depression) in depressed patients, after therapy. To do so we would have needed studies that used the same tests in similar patients, which was not the case. The SCL-90 was the only test that was used in several studies, but we felt it was not appropriate to use this single instrument to compare the effectiveness of therapy in the populations it was used on. It is doubtful that the SCL-90 tells us anything specific on the progress made by BPD patients or eating disorder patients.

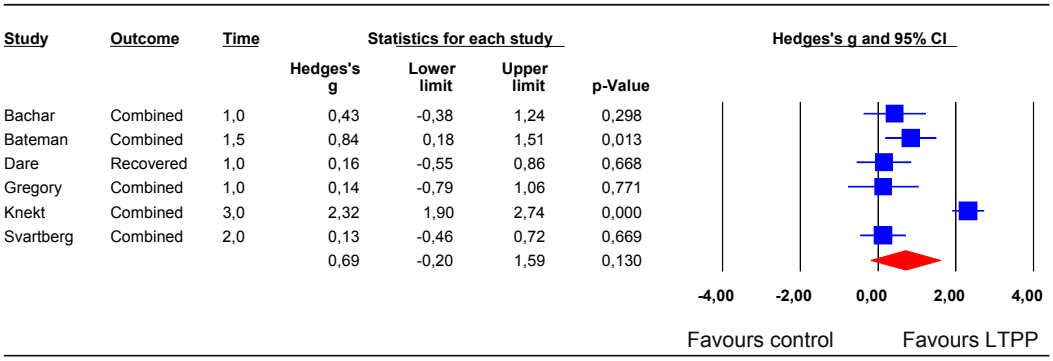
Treatment confounders were present in all studies and included medication and other forms of therapy. It seems practically impossible to control the use of additional or alternative treatments in an outpatient setting, and possibly more so in an inpatient setting as the Bateman study testifies. Pharmacotherapy cannot be excluded in some disorders, but should at least be monitored. Additional psychosocial treatment may be forbidden 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.

9.c Comparison with findings from other meta-analysis

We found a meta-analysed effect size of 0.29 (95% CI: -0.73 to 1.32; p=0.58) for overall effectiveness from various mental disorders in controlled studies that compared LTPP *vs.* a variety of control treatments. This contrasts strongly with the meta-analysed effect size of 1.8 (95% CI: 0.7-3.4) for overall effectiveness, found in the recent meta-analysis by Leichsenring. As we have reasoned in the introduction, Leichsenring seems to reduce the controlled studies he included to an observational level, and his calculation of Hedges´ g is not transparent. In his author´s reply to various letters he states that the effect size *´assessed in the conventional way´* for overall outcome would be 0.65 (Hedges´ g)(p=0.03) (no 95% CI reported) (Leichsenring, 2009). If we calculate an ES for the overall effectiveness without the Giesen-Bloo and Linehan studies (which Leichsenring excluded and did not consider for inclusion respectively), this leads to a similar estimate with a wider confidence interval (Hedges´ g=0.69; 95%CI: -0.20 to 1.59;I-squared=91.3;p=0.13)(Figure 11). Leichsenring may have made different choices in the selection of outcomes or control groups than we have made. In addition, he has included five RCTs that we have excluded (Clarkin, Høglend, Huber, Piper and Vinnars). Unfortunately, we have no further details on the exact calculations Leichsenring used or the studies he included to assess this effect size so we cannot truly compare our findings.

The contrast between the combined effect size for recovery in our meta-analysis (0.02) and the combined effect sizes for overall effectiveness reported by Leichsenring (1.8) and De Maat (0.78) underscores the importance of several points we have discussed previously. First, to including controlled studies only, so as to examine between-group differences instead of within-group differences. Without control, effect sizes of LTPP cannot be interpreted independently of time effects (including aging) and non-specific effects. Second, search for studies that use the intervention of interest as a control treatment. And third, to choose an outcome measure that can be interpreted (recovery) and is not an unweighted mixture of all available outcomes (overall effectiveness).

Figure 11 Combined Hedges´ g for overall effectiveness excluding Giesen-Bloo and Linehan, and including Bachar´s data at one year



Heterogeneity				Tau-squared			
Q-value	df (Q)	P-value	I-squared	Tau Squared	Standard Error	Variance	Tau
57,374	5	0,000	91,285	1,130	0,850	0,723	1,063

9.d Conclusions

Controlled studies on psychoanalysis were not available. The recovery rate of various mental disorders was equal after LTPP or various control treatments, including TAU and non-evidence based control treatments. Similarly, no differences were found for the domains target problems, general psychiatric problems, social functioning or overall effectiveness. Too few studies were available to estimate combined effect sizes for personality pathology and quality of life.

The variation in direction and magnitude of effect indicated that the effect was highly variable. This makes the evidence on whether LTPP has effect on the recovery from various mental disorders conflicting. With only 8 studies available – and 6 available studies on the primary outcome, recovery – the possibilities for subgroup analyses were necessarily very limited. Though we did not find evidence of different effects in BPD or eating disorders we feel that we cannot draw any conclusion from the small number of comparisons we were able to make. We could not perform subgroup analyses for different kinds of control treatments. Our findings refute the previously published large effect sizes for LTPP and show that the effectiveness of any treatment must be examined by controlled studies.

9.e Recommendations for future research

- RCTs on psychoanalysis
- RCTs that compare LTPP with a specialised control treatment (and not with TAU)
- Control treatments should be as comparable as possible in session frequency, therapy duration etc. to be able to measure the effect of the psychoanalytical approach. For example, a partial inpatient LTPP should be compared to a partial inpatient control treatment
- Control treatments should be evidence based, when available
- Studies on LTPP should focus on the disorders that LTPP is used for most frequently. In the Netherlands patients with an indication for LTPP have the following Axis I diagnoses: mood disorder 51% (in particular dysthymia 35%); anxiety disorders 18%; adjustment disorders 8%; mixed Axis I diagnosis 8%. Seventy-four percent of the patients was diagnosed with an additional V-code, of which identity problems and relational problems were most common. Twenty percent of the patients was diagnosed with no Axis I disorder, only a V-code. No data on Axis II diagnoses were available in this recent study (Berghout, 2008)
- A long follow-up is needed to monitor the effect of a long-term treatment
- When events during follow-up are reported, the treatment period and any previous follow-up periods should be included
- Co-interventions and adverse events must be monitored systematically
- RCTs should use power calculations to decide on the number of participants and use ITT analyses
- The choice of outcome assessment instruments should be appropriate and the instruments should be reliable, valid, sensitive to detect change, and be administered by independent assessors/methods
- Cost-effectiveness should also be taken into consideration, given the high costs of LTPP (only two included studies reported cost-effectiveness data)

10 Annexes

10.a Long-term psychotherapy: quality assessment of existing meta-analyses and consequences for an update. Consensus report of expert opinions, April 8 2009

10.a.1 Introduction

Background

For several decades now the effectiveness of psychoanalysis and long-term psychoanalytic psychotherapy (LTPP) are debated. While the effectiveness of other forms of psychotherapy, such as cognitive behaviour therapy and interpersonal psychotherapy, have been scrutinised in controlled trials the research that focuses on psychoanalysis and LTPP is sparse.

The Dutch Health Care Insurance Board (CVZ) funds a systematic review of the literature to answer the following research questions:

- Is psychoanalysis an effective treatment for mental illness? If so, for which patients or illnesses?
- Is long-term psychoanalytic psychotherapy an effective treatment for mental illness? If so, for which patients or illnesses?

Two recent meta-analyses (De Maat 2007 and Leichsenring 2008) meta-analyse the existing evidence on psychoanalysis and LTPP(De Maat, 2007a;de, 2009;Leichsenring, 2008). However, both seem to have their limitations warranting a more firm assessment:

- De Maat´s study involved mainly observational studies (1 randomised trial), including retrospective research. Several clinical trials have been published since
- Leichsenring does not make a distinction between proper psychoanalysis and LTPP
- Leichsenring does not include seven prospective observational studies that were included by De Maat.

Aim

The aim of this quality assessment of the two existing meta-analyses is:

- To assess whether these meta-analyses are in some way useful in answering CVZ´s research questions
- To determine if (parts of) these meta-analyses can be used to build an updated or new meta-analysis on

This consensus report describes the assessment undertaken by experts in epidemiology, psychiatry and psychology. Their findings and recommendations will be used to determine the exact methods of an updated or new systematic literature review into the effectiveness of psychoanalysis and LTPP.

10.a.2 Methods

Assessors

Five experts in relevant fields were asked to assess the meta-analyses of De Maat and Leichsenring. These experts (and their field of expertise) were:

- Professor A. Arntz (psychology)
- Professor R. van Dyck (psychiatry)
- Professor M. Huibers (psychology and epidemiology)
- Professor J. Ioannidis (epidemiology)
- Y. Smit (epidemiology)

Assessment

The assessment was based on the Quality of Reporting of Meta-analyses (QUORUM) checklist (see Annexes) (Moher, 1999;Walker, 1999). The assessors assessed each item on the Quorum

checklist and commented on items if necessary. Assessors compared the meta-analyses of De Maat and Leichsenring, commenting on the weaker or stronger points of either one. Finally, assessors made recommendations for a new/updated systematic review.

Consensus report
One of the experts (Yolba Smit) compiled all assessments in to this report. Comments on Leichsenring in Letters to the Editor, published in JAMA, have been incorporated and are referenced when used (Beck, 2009;Kriston, 2009;Roepke, 2009;Thombs, 2009). A draft version of the consensus report was sent to all experts. Their comments were then incorporated in to the present report.

10.a.3 Quality assessment results

General comments
Neither meta-analysis compares psychoanalysis or LTPP to placebo or any form of control treatment in an adequate way. Eleven randomised controlled trials (RCTs) are included by Leichsenring. Kriston has criticised Leichsenring as follows: *´By calculating point biserial correlations between within-group effect sizes and type of treatment Leichsenring nullified the effect of randomisation. Thus, the findings of the RCTs in his meta-analysis are reduced to an observational level´* (Kriston, 2009). Leichsenring is somewhat cryptic in the methods section about what exactly he has done. From his author´s reply to Kriston´s comment it seems that Kriston is right: *´In response to Dr. Kriston and colleagues, considering treatment groups rather than studies as the unit of analysis can indeed reduce the effect of randomization. It may weaken internal validity but does not necessarily imply serious bias. Observational studies may not systematically overestimate the effects of psychotherapy´* (Leichsenring, 2009).

(Adequate) between-group comparisons would have been much more informative. Now it is unknown whether and how the reported effect sizes compare to no treatment or to alternative treatments. Absolute change (before *vs.* after) cannot be interpreted independent from time effects (including aging) and so-called non-specific effects like attention, empathy, expectations, explanations for problems etcetera.

Furthermore, the two existing meta-analyses:

- Do not give an indication of which patients or which mental illnesses do or do not benefit from psychoanalysis or LTPP. (De Maat gives a global assessment of the pre/post effect size in patients with more severe pathology)
- Tend to over-interpret their findings
- Have analytical problems in the way they combine the data
- Underestimate (or do not estimate at all) the uncertainty involved
- Have problematic quality evaluations
- Publication bias testing is suboptimal or even misleading, with a remarkable discrepancy between the studies retrieved and considered eligible. In the case of Leichsenring a study with an unfavourable outcome for long-term psychodynamic psychotherapy (LTPP) was excluded
- *´It seems unlikely that investigators of small studies (15-30 patients) would attempt to publish negative study results, or that such a study would be accepted for publication. This means that all published studies would have an effect size of at least 0.50 to 0.75, the minimum for statistical significance. This is an artificial floor that guarantees a large effect when these studies are combined´* (Thombs, 2009)

Detailed comments relevant to De Maat
Methods: searches and selection

- Not enough detail to replicate searches exactly
- Not reported whether studies identified by other reviews and meta-analyses, not found by their search, were included

Methods: inclusion and exclusion criteria

- Not entirely clear how it was defined what was included and excluded in psychodynamic, psycho-analytic, and psychoanalysis treatments. E.g. a list of the treatments that were considered eligible would have helped
- The inclusion of retrospective studies is methodologically weak. As this meta-analysis eventually includes 1 RCT, 5 surveys, 5 retrospective cohorts and 16 prospective cohorts it would have been reasonable to exclude retrospective studies completely. The authors rate retrospective studies as second class (by making them low quality studies) but do not exclude them all together. Rather, they make the case that low quality studies too show comparable results, suggesting that we should take them into account

Methods: validity assessment

- It is questionable to combine points in one score, e.g., can a non-RCT design be compensated by other aspects? Quality defects are not exchangeable and/or additive necessarily
- In the validity assessment used, both controlled and uncontrolled studies rate as ´good´
- A quality criterion that uses half the maximal points as a cut-of is flawed or at least arbitrary
- Viewing equal drop-out in conditions as a quality mark is problematic, as treatments might differ in their drop-out rates as a result of their characteristics. Thus, differences in drop-out rates are also an outcome, and equality in an outcome should not be viewed as a quality characteristic

Methods: data abstraction

- No information as to the process used (e.g., completed independently, in duplicate)

Methods: study characteristics

- Sampling methods are ignored (e.g., how patients were recruited; could there be biased recruitment or were patients an unbiased sample of the regular stream of patients?)
- Heterogeneity is ignored, e.g. heterogeneity based on disorder type

Methods: quantitative data synthesis

- Choice for Cohen's d is reasonable, as it is the standard in psychotherapy meta-analyses. But, information as to how d is calculated exactly is missing, e.g. using the standard deviation of the difference score or the standard deviation of baseline or pooled standard deviation of assessments. These different calculations can have a large influence on effect size's d
- Unclear how d is estimated if the descriptive statistics were missing in the original publication
- Unclear how missing data are handled
- Confidence estimates are not assessed
- Heterogeneity is not assessed
- Publication bias is not assessed, which is of concern in this field
- An a priori sensitivity analysis is not reported
- Drop-out rates miss as an outcome measure. (Psychoanalytic treatments can have very high drop-out rates.)
- Some weak outcome measures are used such as CGI and global success rates

Results: study characteristics

- Interactions are not assessed or reported on

Results: quantitative data synthesis

- Agreement on the selection and validity assessment is not reported
- The length of treatment is not taken into account
- Data needed to calculate effect sizes are not reported
- Point estimates are reported but confidence intervals not
- Not reported whether the difference between symptom and personality outcomes is significant

Discussion

- Bias is discussed, but its ability to invalidate its conclusions is dismissed which seems inappropriate

- The overall conclusion that these interventions are effective is too strong. The thin evidence, subject to bias, precludes any conclusion on effectiveness. At best, the conclusion is suggestive
- Information on potential harms or adverse events is missing

Detailed comments relevant to Leichsenring

Methods: searches and selection

- Psychoanalysis is not included
- Not clear how exactly long-term psychodynamic psychotherapy (LTPP) is defined, i.e. which psychotherapies were considered to belong to LTPP.
- The authors include 5 treatments that do not constitute formal psychotherapy as it is generally understood under the designation of ´shorter-term methods of psychotherapy´. These treatments consisted of waitlist control condition, nutritional counselling, standard psychiatric care, low-contact routine treatment and treatment as usual in the community. This is considered as the ´mixing of apples ad oranges´ (Beck, 2009;Roepke, 2009)
- The authors made an effort to identify data from the Internet and through communicating with experts. It's possible that selective reporting of outcomes and biased outcome assessment and analysis may be a more prominent problem in this field rather than classic publication bias

Methods: inclusion and exclusion criteria

- Confusing and potentially biased definition of long-term: exclusion of studies with 40 sessions within 1 year and inclusion of studies with less than 40 sessions over more than 1 year. Moreover, the exclusion of studies in which treatment could continue was violated by the inclusion of at least 2 studies that offered continuation of treatment (Bateman et al., Clarkin et al.). The inclusion criterion that patients should have ended therapy at follow-up only comes up in the discussion. In this way several studies unfavourable for LTPP are excluded via the back door
- It is peculiar that the treatment duration and/or follow-up and possibly outcome assessment differed between trial arms of some trials. A detailed examination of these studies is warranted. Any doubts about the controlled design or the validity of these studies may necessitate exclusion (e.g. Bachar et al. 1999, 12 months psychodynamic psychotherapy *vs.* 12 months cognitive therapy or 6 months nutritional counselling in the control group)

Methods: validity assessment

- The authors modified the Jadad scale to make it appropriate for the type of studies encountered in this field. This is a home-made modification of a scale that is problematic anyhow and therefore it leaves a lot of open questions about how reliable their quality assessment is

Methods: study characteristics

- Patients' characteristics are reported in a limited way
- Clinical heterogeneity assessment is not mentioned
- Mechanism of change measure is an outcome in at least 1 selected study (Clarkin et al., Levy et al.)

Methods: quantitative data synthesis

- Unclear how missing data are handled (intention to treat or completers analysis)
- The use of between-conditions effect size seems erroneous but is at least unclear (Thombs, 2009)
- The use of baseline standard deviations in the estimation of effect sizes is very questionable. A restriction of range at baseline can blow up effect sizes
- Assessments are made for publication bias, but may be misleading in this case: Spearman's correlation and fail-safe N are not the way to go with such data. It is suspected that there is a huge selective analysis and outcome reporting bias in this field. One test that may be useful here is Ioannidis and Trikalinos Clinical Trials 2007 on the evaluation of excess of significant results. It is remarkable that with minute sample sizes, almost all studies give statistically significant results on their own

- Random effects are used in the analysis, which is Ok. However, in this setting it might be even better to use random effects with calculation of the predictive interval for the diverse population effects, not just for the mean population effect (see Higgins et al. JRSS 2009), given the very large diversity of settings, diseases and outcomes. This will give wider confidence(predictive) intervals

Results: study characteristics

- Patients´ age is not described
- Control conditions are not described

Results: quantitative data synthesis

- More simple summary results (mean, standard deviation) would have made the meta-analysis more transparent and easy to control
- An analysis of drop-out proportions is missing

Discussion

- Authors do not acknowledge the weaker points of the analysis. This is especially true for the comment by Thombs and Bassel about the between group difference being larger than the within group difference. This improbable finding suggests a statistical artefact (Thombs, 2009)

Overall

- Impossible to control the results as many German studies were included that cannot be retrieved by regular means and authors refused to send descriptive statistics
- The exclusion of studies in which therapy continues after follow-up is peculiar
- The inclusion of controlled studies were the duration of therapy (and the assessment of outcomes?) differed between arms is questionable
- The fact that results from RCTs and cohort studies appear comparable does not justify throwing them on one pile: why did the authors not keep these results separate?
- Because the authors primarily use post-therapy assessments, the time points at which outcome is assessed shows huge variations, between studies, but also between conditions in one study which seems strange
- The finding that patients are better off after 2 years (long-term therapy) than after 6 months (short-term) is less than surprising and might just as well be the natural course of a disorder. The conclusion that long-term is better than short-term is therefore invalid
- Long-term outcomes of short-term therapies that were available have not been taken in to account
- The conclusion is too strong; the meta-analysis seems to suggest effectiveness by far with large effect sizes, though it is hard to put so much trust in trials of 40 patients each on average. It would not be surprising if no effect is seen in the future, were a large study to be performed

10.a.4 Conclusions

First and foremost: only a calculation of between-group effect sizes can show a true difference between the active treatment and (1) time; (2) non-specific factors; (3) alternative treatments. Therefore it is even possible that having done nothing for these patients would have been better than psychoanalysis or LTPP.

Secondly, the literature in this field is prone to classic publication bias, and it is possible that the selective reporting of outcomes and biased outcome assessment and analysis may be an even more prominent problem.

Third, expectations on what research questions can be answered by examining the literature in this field must be realistic. It seems obvious that the number of studies of good quality is very limited, especially when psychoanalysis is concerned. Therefore, the number of questions that can be answered with the presently available material is also limited. Careful thought should be given to the research priorities. The reliability of a new meta-analysis is upheld only by the foundations on which it is built.

10.a.5 Consequences for an updated or new meta-analysis

We have previously outlined our research proposal for an updated or new meta-analysis on the effectiveness of psychoanalysis and LTPP. From this assessment of the existing meta-analysis it seems clear that we cannot perform a mere update. We need to conduct a new meta-analysis that:

(1) Includes (R)CTs only
(2) Calculates between-group effect sizes

We feel that the levels of evidence need to be separated: the inclusion of cohorts in a new meta-analysis is a repetition of moves that will give similar results as De Maat and Leichsenring presented. Most importantly, the fundamental question whether psychoanalysis and LTPP are effective treatments would not be answered.

In addition to these two main characteristics (and in addition to our research proposal) the new meta-analysis:

(3) Clearly defines/states the types of therapy that are included
(4) Defines long-term as at least 40 sessions within 1 year[1]
(5) Uses the Maastricht-Amsterdam Criteria List and assess agreement
(6) Analyses drop-out rates
(7) Takes the length (dose) of treatment in to account

10.a.6 Annexes

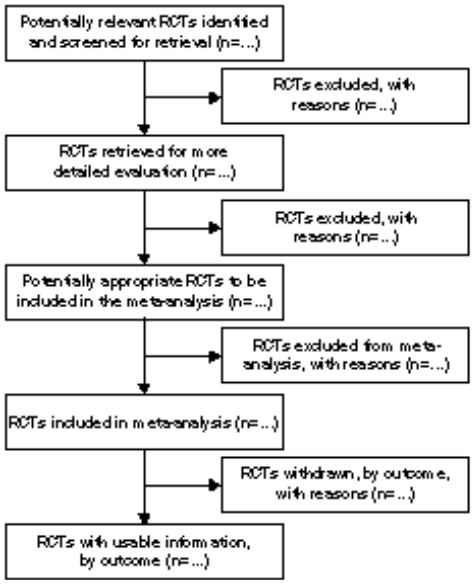
QUORUM checklist (Moher, 1999)

Heading	Subheading	Descriptor	Reported? (Y/N)	Page number
Title		Identify the report as a meta-analysis [or systematic review] of RCTs		
Abstract		Use a structured format		
		Describe		
	Objectives	The clinical question explicitly		
	Data sources	The databases (i.e., list) and other information sources		
	Review methods	The selection criteria (i.e., population, intervention, outcome, and study design); methods for validity assessment, data abstraction, and study characteristics, and quantitative data synthesis in sufficient detail to permit replication		
	Results	Characteristics of the RCTs included and excluded; qualitative and quantitative findings (i.e., point estimates and confidence intervals); and subgroup analyses		
	Conclusion	The main results		
		Describe		
Introduction		The explicit clinical problem, biological rationale for the intervention, and rationale for review		
Methods	Searching	The information sources, in detail (e.g., databases, registers, personal files, expert informants, agencies, hand-searching), and any restrictions (years considered, publication status, language of publication)		
	Selection	The inclusion and exclusion criteria (defining population, intervention, principal outcomes, and study design		
	Validity assessment	The criteria and process used (e.g., masked conditions, quality assessment, and their findings)		
	Data abstraction	The process or processes used (e.g., completed independently, in duplicate)		
	Study characteristics	The type of study design, participants' characteristics, details of intervention, outcome definitions etcetera and how clinical heterogeneity was assessed		
	Quantitative	The principal measures of effect (e.g., relative risk),		

[1] This seems more reasonable than at least 50 sessions as 40 sessions would mean a weekly session (of LTPP), allowing room for vacation and continuous education of the therapist

	data synthesis	method of combining results(statistical testing and confidence intervals), handling of missing data; how statistical heterogeneity was assessed; a rationale for any a-priori sensitivity and subgroup analyses; and any assessment of publication bias
Results	Trial flow	Provide a meta-analysis profile summarising trial flow (see figure)
	Study characteristics	Present descriptive data for each trial (e.g., age, sample size, intervention, dose, duration, follow-up period)
	Quantitative data synthesis	Report agreement on the selection and validity assessment; present simple summary results (for each treatment group in each trial, for each primary outcome); present data needed to calculate effect sizes and confidence intervals in intention-to-treat analyses(e.g. 2x2 tables of counts, means and SDs, proportions)
Discussion		Summarise key findings; discuss clinical inferences based on internal and external validity; interpret the results in light of the totality of available evidence; describe potential biases in the review process (e.g., publication bias); and suggest a future research agenda

Figure 12 Progress through the stages of a meta-analysis for RCTs (Moher, 1999)



10.b Detailed searches

10.b.1 Medline (in OVID® 1950-May week 2 2009)

1. Clinical trial.pt. (451790)
2. Meta-analysis.pt. (21044)
3. Randomized controlled trial.pt. (270142)
4. controlled clinical trial.pt. (79146)
5. Evaluation studies.pt. (117779)
6. or/1-5 (653329)
7. psychoanaly*.ti,ab. (10215)
8. psychodynamic$.ti,ab. (4093)
9. 7 or 8 (13857)
10. 6 and 10 (257)

10.b.2 Embase (in OVID® 1980-May week 1 2009)

1. meta-analysis.ti,ab. (18245)
2. random$.ti,ab. (397432)
3. factorial$.ti,ab. (8294)
4. (crossover$ OR cross over$ OR cross-over$).ti,ab. (39662)
5. placebo$.ti,ab. (110709)
6. (blind$).ti,ab. (139041)
7. trial.ti,ab. (202881)
8. control$.ti,ab. (1502722)
9. or/1-8 (1887659)
10. psychoanaly$.ti,ab. (9028)
11. psychodynam$.ti,ab. (4162)
12. or/10-11 (12523)
13. 9 and 12 (959)

10.b.3 PsycINFO (in OVID® 1806-May week 1 2009)

1. meta-analysis.ti,ab. (7198)
2. random$.ti,ab. (80405)
3. factorial$.ti,ab. (10912)
4. (crossover$ OR cross over$ OR cross-over$).ti,ab. (4529)
5. placebo$.ti,ab. (22095)
6. (blind$).ti,ab. (29240)
7. trial.ti,ab. (42669)
8. control$.ti,ab. (323762)
9. or/1-8 (433688)
10. psychoanaly$.ti,ab. (55811)
11. psychodynam$.ti,ab. (14399)
12. psychodynamic psychotherapy/ (1181)
13. or/10-12 (67492)
14. 9 and 13 (3412)
15. limit 14 to peer reviewed journal (2270)

10.b.4 OVID® All Evidence Based Medicines Reviews

This database includes:

- DARE (from 1991 onwards)
- NHS EED (from 1995 onwards)
- HTA NHS CRD (from 2001 onwards)

- CMR (Cochrane Methodology Register, from 1995 onwards)
- CCTR (Cochrane Central Register of Controlled Trials, from 1991 onwards)
- Cochrane Database of Systematic Reviews
- ACP Journal Club (from 1991 onwards)

Search:
1. psychoanaly*.ti,ab. (62)
2. psychodynamic$.ti,ab. (200)
3. 1 or 2 (252)

10.b.5 www.controlled-trials.com

Searched 21 April 2009 at http://www.controlled-trials.com/mrct/searchform

Searched in both active and archived registers (left hand menu) and in all contributing registers (top menu).

Search term:
psychoanal% OR psychodynamic

10.c Quality criteria

10.c.1 Maastricht Amsterdam criteria list

MAASTRICHT-AMSTERDAM CRITERIA LIST FOR METHODOLOGICAL QUALITY ASSESSMENT

Patient selection	
a. Where the eligibility criteria specified?	
b. Treatment allocation	Yes/No/Don't know
1) Was a method of randomisation performed?	
2) Was the treatment allocation concealed?	Yes/No/Don't know
c. Were the groups similar at baseline regarding the most	Yes/No/Don't know
important prognostic indicators?	Yes/No/Don't know
Interventions	
d. Were the index and control interventions explicitly	
described?	Yes/No/Don't know
e. Was the care provider blinded to the intervention?	Yes/No/Don't know
f. Were co-interventions avoided or comparable?	Yes/No/Don't know
g. Was the compliance acceptable in all groups?	Yes/No/Don't know
h. Was the patient blinded to the intervention?	Yes/No/Don't know
Outcome measurements	
i. Was the outcome assessor blinded to the intervention?	Yes/No/Don't know
j. Were the outcome measures relevant?	Yes/No/Don't know
k. Were adverse events described?	Yes/No/Don't know
l. Was the withdrawal/dropout rate described and	Yes/No/Don't know
acceptable?	
m. Timing follow-up measurements	Yes/No/Don't know
1) Was a short-term follow-up measurement performed?	Yes/No/Don't know
2) Was a long-term follow-up measurement performed?	Yes/No/Don't know
n. Was the timing of the outcome assessment in both groups	
comparable?	
Statistics	Yes/No/Don't know
o. Was the sample size for each group described?	Yes/No/Don't know
p. Did the analyses include an intention-to-treat analysis?	Yes/No/Don't know
q. Were point estimates and measures of variability presented	
for the primary outcome measures?	
Internal validity criteria: b, e, f, g, h, i, j, l, n, p.	
Descriptive criteria: a, c, d, k, m.	
Statistical criteria: o, q.	

a. To score a "yes", the eligibility criteria (e.g. duration of complaints) must be described appropriately.
b. 1. A random (unpredictable) assignment sequence. Methods of allocation using date of birth, date of admission, hospital numbers, or alternation should not be regarded as appropriate.
   2. Assignment generated by an independent person not responsible for determining the eligibility of the patients. This person has no information about the patients included in the trial and has no influence on the assignment sequence or on the decision about eligibility of patients.
c. To receive a "yes", groups must be similar at baseline regarding important prognostic factors (like age, duration of complaints, value of main outcome measures).
d. Adequate description of characteristics like type, modality, application technique, intensity, duration, number, and frequency of sessions for both the experimental intervention and the control condition, so that others could replicate the treatment.
e. The reviewer determines when enough information about the blinding is given in order to score a "yes".

f. Co-interventions should either be avoided in the trial design or comparable between experimental and control groups.
g. The reviewer determines when compliance to the interventions is acceptable, based on the reported intensity, duration, number and frequency of sessions for both the experimental intervention and the control condition.
h. The reviewer determines when enough information about the blinding is given in order to score a "yes".
i. The reviewer determines (per outcome measure) when enough information about the blinding is given in order to score a "yes".
j. The reviewer determines whether the outcome measures were relevant.
k. Each event should be described and correctly attributed to the allocated treatment: if it is explicitly reported that "no adverse events" have occurred, a "yes" should be scored.
l. Participants included in the study who did not complete the observation period or were not included in the analysis must be described. If the percentage of withdrawals and drop-outs does not exceed 20% for short-term follow-up and 30% for long-term follow-up and does not lead to substantial bias, a "yes" is scored. (NB, these percentages are arbitrary, not supported by literature).
m. 1. Outcome assessment at the end of the intervention period.
   2. Outcome assessment more than 2 years after randomisation
n. Timing of outcome assessment should be identical for all study groups and for all important outcome assessments.
o. To be presented for each group at randomization and for most important outcome assessments; NB, this means that, in contrast to previous lists, there is no pre-set cut-off point to determine whether the sample size is sufficient.
p. All randomized patients are reported/analyzed for the most important moments of effect measurement (minus missing values) irrespective of non-compliance and co-interventions.
q. Both point estimates and measures of variability should be presented (to be scored for each important outcome measures separately). Point estimates are: means, medians, modes, etc. Measures of variability are: standard deviations, 95% confidence intervals, etc.

10.c.2 Quality criteria used by Cuijpers et al

According to the criteria used by Cuijpers et al (Cuijpers, 2009), a study was considered to be of high quality when:

(a) Participants met diagnostic criteria for a mental disorder (as assessed with a personal diagnostic interview and using a diagnostic system such as DSM
(b) The study referred to the use of a treatment manual (either a published manual, or a manual specifically designed for the study)
(c) The therapists who conducted the therapy were trained for the specific therapy, either specifically for that study or as a general training
(d) Treatment integrity was checked during the study (by supervision of the therapists during treatment or by recording of treatment sessions or by systematic screening of protocol adherence by a standardized measurement instrument)
(e) Data were analysed with intention-to-treat analyses, in which all persons who were randomized to the treatment and control conditions initially were included in the analyses
(f) The study had a minimal level of statistical power to find significant effects of the treatment, and included $\geq$50 persons in the comparison between treatment and control groups. This allows the study to find standardized effect sizes of **d**=0.80 and larger, assuming a statistical power of 0.80 and α=0.05. Calculations in Stata (Stata Corp., USA)
(g) The study reported that randomization was conducted by an independent (third) party (this variable was positive if an independent person did the randomization, when a computer program was used to assign patients to conditions, or when sealed envelopes were used)

(h) Assessors of outcome were blinded and did not know to which condition the respondents were assigned to (this was only coded when the effect sizes were based on interviewer-based depression ratings ; when only self-reports were used, it was assumed that this criterion was met)

If a study did not report whether it met the quality criterion is was coded as negative.

11 References

Cochrane handbook for systematic reviews of interventions. Chichester: John Wiley & Sons Ltd; 2009.

Abbass AA, Hancock JT, Henderson J, Kisely S. Short-term psychodynamic psychotherapies for common mental disorders. Cochrane Database Syst Rev 2006;(4):CD004687.

Albrecht M, Krauth C, Rieger J, Lamprecht F, Kersting A, Schwartz FW. [Concept for a health care economic evaluation of short- and long-term costs and effectiveness parameters of an expanded ambulatory psychosomatic rehabilitation program]. Gesundheitswesen 2000 Mar;62(3):156-60.

Arnevik E, Wilberg T, Urnes O, Johansen M, Monsen JT, Karterud S. Psychotherapy for personality disorders: short-term day hospital psychotherapy versus outpatient individual therapy - a randomized controlled study. Eur Psychiatry 2009 Mar;24(2):71-8.

Bachar E, Latzer Y, Kreitler S, Berry EM. Empirical comparison of two psychological therapies. Self psychology and cognitive orientation in the treatment of anorexia and bulimia. J Psychother Pract Res 1999;8(2):115-28.

Bateman A, Fonagy P. Effectiveness of partial hospitalization in the treatment of borderline personality disorder: a randomized controlled trial. Am J Psychiatry 1999 Oct;156(10):1563-9.

Bateman A, Fonagy P. Treatment of borderline personality disorder with psychoanalytically oriented partial hospitalization: an 18-month follow-up. Am J Psychiatry 2001 Jan;158(1):36-42.

Bateman A, Fonagy P. Health service utilization costs for borderline personality disorder patients treated with psychoanalytically oriented partial hospitalization versus general psychiatric care. Am J Psychiatry 2003 Jan;160(1):169-71.

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 May;165(5):631-8.

Bateman AW, Ryle A, Fonagy P, Kerr IB. Psychotherapy for borderline personality disorder: mentalization based therapy and cognitive analytic therapy compared. Int Rev Psychiatry 2007 Feb;19(1):51-62.

Beck AT, Bhar SS. Analyzing effectiveness of long-term psychodynamic psychotherapy. JAMA 2009 Mar 4;301(9):931-3.

Beecham J, Sleed M, Knapp M, Chiesa M, Drahorad C. The costs and effectiveness of two psychosocial treatment programmes for personality disorder: a controlled study. Eur Psychiatry 2006 Mar;21(2):102-9.

Bell L. Psychological therapies in anorexia nervosa. Br J Psychiatry 2001 Jul;179:78-9.

Berghout CC, Zevalkink J. Identifying clinical cases among patients assigned to psychoanalytic treatment. Bull Menninger Clin 2008;72(3):163-78.

Berry DM, Abramowitz SI. Educative/support groups and subliminal psychodynamic activation for bulemic college women. Int J Eat Disord 1989;8(1):75-85.

Bolz W, Meyer AE. The Hamburg short psychotherapy comparison experiment: the general setting. Psychother Psychosom 1981;35(2-3):85-95.

Borenstein M, Hedges L, Higgins JP, Rothstein H. Introduction to meta-analysis. Chichester: John Wiley & Sons Ltd; 2009.

Briffault X, Thurin M, Lapeyronnie B, Thurin JM. [New directions for psychotherapy research: an evaluation of a research protocol and a methodological and technical framework proposal]. Encephale 2007 Dec;33(6):911-23.

Brill NQ. Results of psychotherapy. Calif Med 1966 Apr;104(4):249-53.

Brockmann J, Schlüter T, Eckert J. [Long-term effects of long-term psychoanalytic and long-term behavior therapy. A comparative study from the general practices of psychotherapists]. Psychotherapeut 2006;51:15-25.

Brom D, Kleber RJ, Defares PB. Brief psychotherapy for posttraumatic stress disorders. J Consult Clin Psychol 1989 Oct;57(5):607-12.

Burnand Y, Andreoli A, Kolatte E, Venturini A, Rosset N. Psychodynamic psychotherapy and clomipramine in the treatment of major depression. Psychiatr Serv 2002 May;53(5):585-90.

Burzig G, Speidel H, Bolz W, Meyer AE. Our pluridimensional evaluation system for short psychotherapy outcome. Psychother Psychosom 1981;35(2-3):134-7.

Canceil O, Cottreaux J, Fallisard B, Flamant M, Miermont J, Swendsen J, et al. Psychothérapies: trois approches évaluées. Paris: INSERM; 2004.

Chabrol H, Teissedre F, Saint-Jean M, Teisseyre N, Sistac C, Michaud C, et al. [Detection, prevention and treatment of postpartum depression: a controlled study of 859 patients]. Encephale 2002 Jan;28(1):65-70.

Chambless DL, Hollon SD. Defining empirically supported therapies. J Consult Clin Psychol 1998 Feb;66(1):7-18.

Chiesa M, Fonagy P. From the efficacy to the effectiveness model in psychotherapy research: the APP multi-centre project. Psychoanal Psychother 1999;13(3):259-72.

Chiesa M, Fonagy P. Cassel Personality Disorder Study. Methodology and treatment effects. Br J Psychiatry 2000 May;176:485-91.

Chiesa M, Fonagy P. Psychosocial treatment for severe personality disorder. 36-month follow-up. Br J Psychiatry 2003 Oct;183:356-62.

Chiesa M, Fonagy P, Holmes J. Six-year follow-up of three treatment programs to personality disorder. J Pers Disord 2006 Oct;20(5):493-509.

Chiesa M, Fonagy P, Holmes J, Drahorad C. Residential versus community treatment of personality disorders: a comparative study of three treatment programs. Am J Psychiatry 2004 Aug;161(8):1463-70.

Chiesa M, Fonagy P, Holmes J, Drahorad C, Harrison-Hall A. Health service use costs by personality disorder following specialist and nonspecialist treatment: a comparative study. J Pers Disord 2002 Apr;16(2):160-73.

Clarkin JF, Levy KN, Lenzenweger MF, Kernberg OF. Evaluating three treatments for borderline personality disorder: a multiwave study. Am J Psychiatry 2007 Jun;164(6):922-8.

Clarkin JF, Levy KN, Schiavi JM. Transference focused psychotherapy: development of a psychodynamic treatment for severe personality disorders. Clin Neuroscience Res 2005;4:379-86.

Cooper PJ, Murray L, Wilson A, Romaniuk H. Controlled trial of the short- and long-term effect of psychological treatment of post-partum depression. I. Impact on maternal mood. Br J Psychiatry 2003 May;182:412-9.

Crits-Christoph P, Connolly M, Azarian K, Crits-Christoph K, Shappell S. An open trial of brief sypportive-expressive psychotherapy in the treatment of generalized anxiety disorder. Psychother 1996;33(3):418-30.

Crits-Christoph P, Siqueland L, Blaine J, Frank A, Luborsky L, Onken LS, et al. The National Institute on Drug Abuse Collaborative Cocaine Treatment Study. Rationale and methods. Arch Gen Psychiatry 1997 Aug;54(8):721-6.

Cuijpers P, van SA, Bohlmeijer E, Hollon SD, Andersson G. The effects of psychotherapy for adult depression are overestimated: a meta-analysis of study quality and effect size. Psychol Med 2009 Jun 3;1-13.

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 Mar;178:216-21.

De Maat S, De Jonghe F, Schoevers R, Dekker J. The effectiveness of long-term psychoanalytic therapy: a systematic review of empirical studies. Amsterdam: Vrije Universiteit; 2007a.

De Maat S, De Jonghe F, Schoevers R, Dekker J. The effectiveness of long-term psychoanalytic therapy: a systematic review of empirical studies (dissertation). Amsterdam: Vrije Universiteit; 2007b.

De Maat S, De Jonghe F, Schoevers R, Dekker J. The effectiveness of long-term psychoanalytic therapy: a systematic review of empirical studies. Harv Rev Psychiatry 2009;17(1):1-23.

de MS, de JF, Schoevers R, Dekker J. The effectiveness of long-term psychoanalytic therapy: a systematic review of empirical studies. Harv Rev Psychiatry 2009;17(1):1-23.

Dührssen A, Jorswieck E. Eine empirisch-statistische Untersuchung zur Leistungsfähigkeit psychoanalytischer Behandlung. Z Psychosom Med 1998;44:311-8.

Eckert J, Biermann-Ratjen EM, Wuchner M. [Long-term changes in borderline symptoms of patients after client-centered group psychotherapy]. Psychother Psychosom Med Psychol 2000 Mar;50(3-4):140-6.

Egger M, Smith GD. Principles of and procedures for systematic reviews. In: Egger M, Smith GD, Altman D, editors. Systematic reviews in health care. Meta-analysis in context. 2 ed. London: BMJ Publishing Group; 2001. p. 23-42.

Giesen-Bloo J, van DR, Spinhoven P, van TW, Dirksen C, van AT, et al. Outpatient psychotherapy for borderline personality disorder: randomized trial of schema-focused therapy vs transference-focused psychotherapy. Arch Gen Psychiatry 2006 Jun;63(6):649-58.

Glass RM. Psychodynamic psychotherapy and research evidence: Bambi survives Godzilla? JAMA 2008 Oct 1;300(13):1587-9.

Gonzalez JM, Prihoda TJ. A case study of psychodynamic group psychotherapy for bipolar disorder. Am J Psychother 2007;61(4):405-22.

Grande T, Dilg R, Jakobsen T, Keller W, Krawietz B, LANGER M, et al. Differential effects of two forms of psychoanalytic therapy: results of the Heidelberg-Berlin study. Psychother Research 2006;16(4):470-85.

Grawe K, Caspar, f:, Ambühl H. [Differential psychotherapy research: four forms of therapy compared. The Bernese comparative treatment study: basic research questions and design]. Z Klin Psychol 1990;19(4):S294-315.

Gregory RJ, Chlebowski S, Kang D, Remen AL, Soderberg M, Stepkovitz J, et al. A controlled trial of psychodynamic psychotherapy for co-occurring borderline personality disorder and alcohol use disorder. Psychother Theory Res Practice Training 2008;45(1):21-48.

Grenyer BF. Hope for sustaining a positive 3-year therapeutic relationship with patients with borderline personality disorder. Arch Gen Psychiatry 2007 May;64(5):609.

Grossarth-Maticek R, Eysenck HJ. Prophylactic effects of psychoanalysis on cancer-prone and coronary heart disease-prone probands, as compared with control groups and behaviour therapy groups. J Behav Ther Exp Psychiatry 1990 Jun;21(2):91-9.

Gunderson JG, Gabbard GO. Making the case for psychoanalytic therapies in the current psychiatric environment. J Am Psychoanal Assoc 1999;47(3):679-704.

Henken HT, Huibers MJ, Churchill R, Restifo K, Roelofs J. Family therapy for depression. Cochrane Database Syst Rev 2007;(3):CD006728.

Heuft G, Seibüchler-Engeç H, Tschke M, Senf W. [Long-term outcome of outpatient psychoanalytic psychotherapy and psychoanalysis. Analysis of 53 catamnestic interviews]. Forum Psychoanal 1996;12:342-55.

Higgins JP, Thompson SG, Spiegelhalter DJ. A re-evaluation of random-effects meta-analysis. J R Stat Soc Ser A Stat Soc 2009 Jan;172(1):137-59.

Hoffart A, Martinsen EW. Exposure-based integrated vs. pure psychodynamic treatment of agoraphobic inpatients. Psychother 1990;27(2):210-8.

Hoffmann C, Ruf-Ballauf W. [Psychosomatic rehabilitation of in-patients with somatoform disorders and of patients with social-medically relevant problems: results of a two-year follow-up]. Rehabilitation (Stuttg) 2007 Oct;46(5):283-95.

Hoglend P, Amlo S, Marble A, Bogwald KP, Sorbye O, Sjaastad MC, et al. Analysis of the patient-therapist relationship in dynamic psychotherapy: an experimental study of transference interpretations. Am J Psychiatry 2006 Oct;163(10):1739-46.

Hoglend P, Bogwald KP, Amlo S, Marble A, Ulberg R, Sjaastad MC, et al. Transference interpretations in dynamic psychotherapy: do they really yield sustained effects? Am J Psychiatry 2008 Jun;165(6):763-71.

Huber D, Albrecht C, Hautum A, Henrich G, Klug G. [Effectiveness of inpatient psychodynamic psychotherapy: A follow-up study.]. Z Psychosom Med Psychother 2009;55(2):189-99.

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

Huber D, Klug G, Von Rad M. [The Munich processoutcome study: a comparison between psychoanalyses and psychotherapy]. In: Stuhr BM, editor. *Langzeit- Psychotherapie. Perspektiven für Therapeuten und Wissenschaftler*.Stuttgart, Germany: Kohlhammer; 2001. p. 260-70.

Jäger J, von Wietersheim B. [Psychoanalytically oriented psychotherapy of bulimia nervosa]. Psychother Psychosom Med Psychol 1997;47:322-31.

Jakobsen T, Rudolf G, Brockmann J, Eckert J, Huber D, Klug G, et al. [Results of psychoanalytic long-term therapy in specific diagnostic groups: improvement in symptoms and interpersonal relationships]. Z Psychosom Med Psychother 2007;53(2):87-110.

Jarvik LF, Mintz J, Steuer J, Gerner R. Treating geriatric depression: a 26-week interim analysis. J Am Geriatr Soc 1982 Nov;30(11):713-7.

Karon BP, Vandenbos GR. The consequences of psychotherapy for schizophrenic patients. Psychother Theory Res Practice Training 1972;9(2):111-9.

Klar FJ. [Effectiveness of indivdual psychological psychoanalytical psychotherapy]. Z Individualpsychol 2005;30(1):28-50.

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 Maya;38(5):689-703.

Knekt P, Lindfors O, Laaksonen MA, Raitasalo R, Haaramo P, Jarvikoski A. 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 Aprb;107(1-3):95-106.

Korner A, Gerull F, Meares R, Stevenson J. Borderline personality disorder treated with the conversational model: a replication study. Compr Psychiatry 2006 Sep;47(5):406-11.

Kriston L, Holzel L, Harter M. Analyzing effectiveness of long-term psychodynamic psychotherapy. JAMA 2009 Mar 4;301(9):930-1.

Lanza ML, Anderson J, Boisvert CM, LeBlanc A, Fardy M, Steel B. Assaultive behavior intervention in the Veterans Administration: psychodynamic group psychotherapy compared to cognitive behavior therapy. Perspect Psychiatr Care 2002 Jul;38(3):89-97.

Lanza ML, Satz H, Stone J, Kayne HL. Assessing the impact of group treatment for aggressive inpatients. Group 1995;19(4):195-219.

Leichsenring F, Rabung S. Effectiveness of long-term psychodynamic psychotherapy: a meta-analysis. JAMA 2008 Oct 1;300(13):1551-65.

Leichsenring F, Rabung S. Analyzing effectiveness of long-term psychodynamic psychotherapy - Reply. JAMA 2009 Mar 4;301(9):932-3.

Leuzinger-Bohleber M, Stuhr U, Rüger B, Beutel ME. Langzeitwirkungen von Psychoanalysen und Psychotherapien: eine multiperspektivische, repräsentative Katamnesestudie. Psyche (Stuttg) 2001;55(3):193-276.

Levy Berg A, Sandell R, Sandahl C. Affect-focused body psychotherapy in patients with generalized anxiety disorder: evaluation of an integrative method. J Psychother Integration 2009;19(1):67-85.

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 Dec;74(6):1027-40.

Liberati A, Altman DG, Tetzlaff J, Mulrow C, Gotzsche PC, Ioannidis JP, et al. The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate health care interventions: explanation and elaboration. PLoS Med 2009 Jul 21;6(7):e1000100.

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 Jul;63(7):757-66.

Manos N, Vasilopoulou E. Evaluation of psychoanalytic psychotherapy outcome. Acta Psychiatr Scand 1984 Jul;70(1):28-35.

May PR, Tuma AH. Treatment of schizophrenia. An experimental study of five treatment methods. Br J Psychiatry 1965;111(503):510.

McCallum M, Piper WE. A controlled study of effectiveness and patient suitability for short-term group psychotherapy. Int J Group Psychother 1990 Oct;40(4):431-52.

McMain SF, Links PS, Gnam WH, Guimond T, Cardish RJ, Korman L, et al. A randomized trial of dialectical behavior therapy versus general psychiatric management for borderline personality disorder. Am J Psychiatry 2009 Dec;166(12):1365-74.

Meyer AE, Bolz W. The Hamburg short psychotherapy comparison experiment. Foreward and introduction. Psychother Psychosom 1981;35(2-3):81-4.

Miklowitz DJ, Otto MW, Frank E, Reilly-Harrington NA, Wisniewski SR, Kogan JN, et al. Psychosocial treatments for bipolar depression: a 1-year randomized trial from the Systematic Treatment Enhancement Program. Arch Gen Psychiatry 2007 Apr;64(4):419-26.

Milrod B, Busch F, Leon AC, Aronson A, Roiphe J, Rudden M, et al. A pilot open trial of brief psychodynamic psychotherapy for panic disorder. J Psychother Pract Res 2001;10(4):239-45.

Milrod B, Busch F, Leon AC, Shapiro T, Aronson A, Roiphe J, et al. Open trial of psychodynamic psychotherapy for panic disorder: a pilot study. Am J Psychiatry 2000 Nov;157(11):1878-80.

Milrod B, Leon AC, Busch F, Rudden M, Schwalberg M, Clarkin J, et al. A randomized controlled clinical trial of psychoanalytic psychotherapy for panic disorder. Am J Psychiatry 2007 Feb;164(2):265-72.

Mintz J, O'B'rien CP, Luborsky L. Predicting the outcome of psychotherapy for schizophrenics. Relative contributions of patient, therapist, and treatment characteristics. Arch Gen Psychiatry 1976 Oct;33(10):1183-6.

Moher D, Cook DJ, Eastwood S, Olkin I, Rennie D, Stroup DF. Improving the quality of reports of meta-analyses of randomised controlled trials: the QUOROM statement. Quality of Reporting of Meta-analyses. Lancet 1999 Nov 27;354(9193):1896-900.

Munroe-Blum H, Marziali E. A controlled trial of short-term group treatment for borderline personality disorder. J Pers Disord 1995;9(3):190-8.

O'Brien CP, Hamm KB, Ray BA, Pierce JF, Luborsky L, Mintz J. Group vs individual psychotherapy with schizophrenics. A controlled outcome study. Arch Gen Psychiatry 1972 Oct;27(4):474-8.

Ojehagen A, Berglund M, Appel CP, Andersson K, Nilsson B, Skjaerris A, et al. A randomized study of long-term out-patient treatment in alcoholics. Psychiatric treatment versus multimodal behavioural therapy, during 1 versus 2 years of treatment. Alcohol Alcohol 1992 Nov;27(6):649-58.

Okhai MF. Psychological therapies in anorexia nervosa. Br J Psychiatry 2001 Jul;179:78-9.

Pearce S. Knowledge of the effectiveness of treatments for borderline personality disorder is not yet sufficient to justify the lack of a control condition. Arch Gen Psychiatry 2007 May;64(5):609.

Petrak F, Nickel R, Hardt J, Egle UT. Verhaltenstherapie versus psychodynamisch-interaktionelle Therapie bei Patienten mit somatoformen Schmerzen. 1-Jahres-follow-up einer randomisierten, kontrollierten Studie. Verhaltenstherapie 2007;17(Suppl1):19-20.

Piper WE, Debbane EG, Bienvenu JP, Garant J. A comparative study of four forms of psychotherapy. J Consult Clin Psychol 1984 Apr;52(2):268-79.

Piper WE, Joyce AS, Azim HF, Rosie JS. Patient characteristics and success in day treatment. J Nerv Ment Dis 1994 Jul;182(7):381-6.

Puschner B, Kraft S, Bauer S. Interpersonal problems and outcome in outpatient psychotherapy: findings from a long-term longitudinal study in Germany. J Pers Assess 2004 Dec;83(3):223-34.

Puschner B, Kraft S, Kachele H, Kordy H. Course of improvement over 2 years in psychoanalytic and psychodynamic outpatient psychotherapy. Psychol Psychother 2007 Mar;80(Pt 1):51-68.

Rascon SR, Corona PC, Lartigue T, Rios JM, Garza DL. A successful trial utilizing the Leuzinger-Bohleber methodology for evaluation of psychoanalytic treatment: preliminary report. Int J Psychoanal 2005 Oct;86(Pt 5):1425-40.

Roepke S, Renneberg B. Analyzing effectiveness of long-term psychodynamic psychotherapy. JAMA 2009 Mar 4;301(9):931-2.

Rudolf G, Jakobsen T, Micka R, Schumann E. [Results of psychodynamic inpatient psychotherapy in relation to diagnosis]. Z Psychosom Med Psychother 2004;50(1):37-52.

Rudolf G, Manz R, ri C. Outcomes of psychoanalytic therapies. Int Forum Psychoanal 1999;8:125-34.

Sachsse U, Vogel C, Leichsenring F. Results of psychodynamically oriented trauma-focused inpatient treatment for women with complex posttraumatic stress disorder (PTSD) and borderline personality disorder (BPD). Bull Menninger Clin 2006;70(2):125-44.

Sandahl C, Herlitz K, Ahlin G, Rönnberg S. Time-limited group psychotherapy for moderately alcohol dependent patients: a randomized controlled clinical trial. Psychother Research 1998;8(4):361-78.

Sandell R, Blomberg J, Lazar A. When reality doesn´t fit the blueprint: doing research on psychoanalysis and long-term psychotherapy in a public health service program. Psychother Res 1997;7(4):333-44.

Sandell R, Blomberg J, Lazar A. [Repeated long-term follow-up of long-term psychotherapy and psychoanalysis. First findings of the Stockholm Outcome of Psychotherapy (STOP) Project]. Z Psychosom Med Psychother 1999;45(1):43-56.

Sandell R, Blomberg J, Lazar A, Carlsson J, Broberg J, Schubert J. Varieties of long-term outcome among patients in psychoanalysis and long-term psychotherapy. A review of findings in the Stockholm Outcome of Psychoanalysis and Psychotherapy Project (STOPP). Int J Psychoanal 2000 Oct;81 (Pt 5):921-42.

Schleussner D. [Effectiveness of individual psychological psychotherapy - a long term trial regarding psychotherapy research]. Z Individualpsychol 2005;30(1):S51-77.

Schulz H, Lotz-Rambaldi W, Koch U, Jürgensen R, Rüddel H. [One year follow-up of an inpatient treatment in a psychosomatic rehabilitation clinic with either a psychoanalytic or behavior therapy oriented treatment]. Psychother Med Psychol (Stuttg) 1999;49:114-30.

Shaffer HJ, LaSalvia TA, Stein JP. Comparing Hatha yoga with dynamic group psychotherapy for enhancing methadone maintenance treatment: a randomized clinical trial. Altern Ther Health Med 1997 Jul;3(4):57-66.

Siassi I. A comparison of open-ended psychoanalytical-oriented psychotherapy with other therapies. J Clin Psych 1979;40(1):25-32.

Simpson S, Corney R, Fitzgerald P, Beecham J. A randomised controlled trial to evaluate the effectiveness and cost-effectiveness of counselling patients with chronic depression. Health Technol Assess 2000;4(36):1-83.

Simpson S, Corney R, Fitzgerald P, Beecham J. A randomized controlled trial to evaluate the effectiveness and cost-effectiveness of psychodynamic counselling for general practice patients with chronic depression. Psychol Med 2003 Feb;33(2):229-39.

Sohlberg S, Rosmark B, Norring C, Holmgren S. Two year outcome in anorexia nervosa/bulimia: a controlled study of an eating control program combined with psychoanalytically oriented psychotherapy. Int J Eat Disord 1987;6(2):243-55.

Spinhoven P, Giesen-Bloo J, van DR, Arntz A. Can assessors and therapists predict the outcome of long-term psychotherapy in borderline personality disorder? J Clin Psychol 2008 Jun;64(6):667-86.

Spinhoven P, Giesen-Bloo J, van DR, Kooiman K, Arntz A. The therapeutic alliance in schema-focused therapy and transference-focused psychotherapy for borderline personality disorder. J Consult Clin Psychol 2007 Feb;75(1):104-15.

Stern R. Partial hospitalization for borderline personality disorder. Am J Psychiatry 2001 Nov;158(11):1932-3.

Steuer JL, Mintz J, Hammen CL, Hill MA, Jarvik LF, McCarley T, et al. Cognitive-behavioral and psychodynamic group psychotherapy in treatment of geriatric depression. J Consult Clin Psychol 1984 Apr;52(2):180-9.

Stiles WB, Barkham M, Mellor-Clark J, Connell J. Effectiveness of cognitive-behavioural, person-centred, and psychodynamic therapies in UK primary-care routine practice: replication in a larger sample. Psychol Med 2008 May;38(5):677-88.

Stiles WB, Barkham M, Twigg E, Mellor-Clark J, Cooper M. Effectiveness of cognitive-behavioural, person-centred and psychodynamic therapies as practised in UK National Health Service settings. Psychol Med 2006 Apr;36(4):555-66.

Stuhr U, Meyer AE, Bolz W. Outcome results in psychological tests. Psychother Psychosom 1981;35(2-3):138-98.

Svartberg M, Stiles TC, 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 May;161(5):810-7.

Taylor S, McLean P. Outcome profiles in the treatment of unipolar depression. Behav Res Ther 1993 Mar;31(3):325-30.

Teusch L, Bohme H, Gastpar M. The benefit of an insight-oriented and experiential approach on panic and agoraphobia symptoms. Results of a controlled comparison of client-centered therapy alone and in combination with behavioral exposure. Psychother Psychosom 1997;66(6):293-301.

Thombs BD, Bassel M, Jewett LR. Analyzing effectiveness of long-term psychodynamic psychotherapy. JAMA 2009 Mar 4;301(9):930-3.

Ulberg R, Johansson P, Marble A, Hoglend P. Patient sex as moderator of effects of transference interpretation in a randomized controlled study of dynamic psychotherapy. Can J Psychiatry 2009 Feb;54(2):78-86.

van Asselt AD, Dirksen CD, Arntz A, Giesen-Bloo JH, van DR, Spinhoven P, et al. Out-patient psychotherapy for borderline personality disorder: cost-effectiveness of schema-focused therapy v. transference-focused psychotherapy. Br J Psychiatry 2008 Jun;192(6):450-7.

van Tulder MW, Assendelft WJ, Koes BW, Bouter LM. Method guidelines for systematic reviews in the Cochrane Collaboration Back Review Group for Spinal Disorders. Spine 1997 Oct 15;22(20):2323-30.

Vinnars B, Barber JP, Noren K, Gallop R, Weinryb RM. Manualized supportive-expressive psychotherapy versus nonmanualized community-delivered psychodynamic therapy for patients with personality disorders: bridging efficacy and effectiveness. Am J Psychiatry 2005 Oct;162(10):1933-40.

Vinnars B, Barber JP, Noren K, Thormählen B, Gallop R, Lindgren A, et al. Who can benefit from time-limited dynamic psychotherapy? A study of psychiatric outpatients with personality disorders. Clin Psychol Psychother 2007;14(198):210.

Vollebergh W, de Graaf R, ten Have M, Schoemaker C, van Dorsselaer S, Spijker J, et al. [Psychische stoornissen in Nederland. Overzicht van de resultaten van NEMESIS]. Utrecht: Trimbos-instituut; 2003.

Walker A. Meta-style and expert review. Lancet 1999 Nov 27;354(9193):1834-5.

Walsh BT, Wilson GT, Loeb KL, Devlin MJ, Pike KM, Roose SP, et al. Medication and psychotherapy in the treatment of bulimia nervosa. Am J Psychiatry 1997 Apr;154(4):523-31.

Winston A, Laikin M, Pollack J, Samstag LW, McCullough L, Muran JC. Short-term psychotherapy of personality disorders. Am J Psychiatry 1994 Feb;151(2):190-4.

Yeomans F. Questions concerning the randomized trial of schema-focused therapy vs transference-focused psychotherapy. Arch Gen Psychiatry 2007 May;64(5):609-10.

Zanarini MC, Frankenburg FR, Hennen J, Reich DB, Silk KR. Prediction of the 10-year course of borderline personality disorder. Am J Psychiatry 2006 May;163(5):827-32.