CHAPTER 8

Non-pharmacological interventions for somatoform disorders and medically unexplained physical symptoms (MUPS) in adults, a systematic review.

Nikki Claassen - van Dessel
Madelon den Boeft
Johannes C. van der Wouden
Maria Kleinstäuber
Stephanie S. Leone
Berend Terluin
Mattijs E Numans
Henriëtte E. van der Horst
Harm van Marwijk

Published in: Cochrane Database of Systematic Reviews 2014;11:CD011142.

NB: This chapter is a concise version of the full Cochrane review article.
Non-pharmacologic interventions
Abstract

Background
Various psychological and physical therapies have been developed to treat somatoform disorders and MUPS. Although there are several reviews on non-pharmacological interventions for somatoform disorders and MUPS, a complete overview of the whole spectrum is missing.

Objectives
To assess the effects of non-pharmacological interventions for somatoform disorders and MUPS in adults.

Search methods
We searched the Cochrane Depression, Anxiety and Neurosis Review Group’s Specialised Register (CCDANCTR) to November 2013. This register includes relevant randomised controlled trials (RCTs) from The Cochrane Library, EMBASE, MEDLINE, and PsycINFO. We ran an additional search on the Cochrane Central Register of Controlled Trials and a cited reference search on the Web of Science. We also searched grey literature, conference proceedings, international trial registers, and relevant systematic reviews.

Selection criteria
We included RCTs and cluster randomised controlled trials which involved adults primarily diagnosed with a somatoform disorder or an alternative diagnostic concept of MUPS, who were assigned to a non-pharmacological intervention compared with usual care, waiting list controls, attention or psychological placebo, enhanced care, or another psychological or physical therapy intervention, alone or in combination.

Data collection and analysis
We pooled data from studies addressing the same comparison using standardised mean differences (SMD) or risk ratios (RR) and a random-effects model. Primary outcomes were severity of somatic symptoms and acceptability of treatment.

Main results
We included 21 studies with 2658 randomised participants. All studies assessed the effectiveness of some form of psychological therapy. We found no studies that included physical therapy. Fourteen studies evaluated forms of cognitive behavioural therapy (CBT); the remainder evaluated behaviour therapies, third-wave CBT (mindfulness), psychodynamic therapies, and integrative therapy. Fifteen included studies compared the studied psychological therapy with usual care or a waiting list. Five studies compared the intervention to enhanced or structured care. Only one study compared cognitive behavioural therapy with behaviour therapy.
Non-pharmacologic interventions

For all studies comparing some form of psychological therapy with usual care or a waiting list that could be included in the meta-analysis, the psychological therapy resulted in less severe symptoms at end of treatment (SMD -0.34; 95% confidence interval (CI) -0.53 to -0.16; 10 studies, 1081 analysed participants). This effect was considered small to medium; heterogeneity was moderate and overall quality of the evidence was low. Compared with usual care, psychological therapies resulted in a 7% higher proportion of drop-outs during treatment (RR acceptability 0.93; 95% CI 0.88 to 0.99; 14 studies, 1644 participants; moderate-quality evidence). Removing one outlier study reduced the difference to 5%. Results for the subgroup of studies comparing CBT with usual care were similar to those in the whole group.

Five studies (624 analysed participants) assessed symptom severity comparing some psychological therapy with enhanced care, and found no clear evidence of a difference at end of treatment (pooled SMD -0.19; 95% CI -0.43 to 0.04; considerable heterogeneity; low-quality evidence). Five studies (679 participants) showed that psychological therapies were somewhat less acceptable in terms of drop-outs than enhanced care (RR 0.93; 95% CI 0.87 to 1.00; moderate-quality evidence). In all included studies adverse events were seldomly reported.

**Conclusions**

When all psychological therapies included this review were combined they were superior to usual care or waiting list in terms of reduction of symptom severity, but effect sizes were small. As a single treatment, only CBT has been adequately studied to allow tentative conclusions for practice to be drawn. Compared with usual care or waiting list conditions, CBT reduced somatic symptoms, with a small effect and substantial differences in effects between CBT studies. Compared with enhanced or structured care, psychological therapies generally were not more effective for most of the outcomes. Compared with enhanced care, CBT was not more effective. The number of studies investigating various treatment modalities (other than CBT) needs to be increased; this is especially relevant for studies concerning physical therapies.
**Non-pharmacologic interventions**

**Background**

**Description of the condition**

Medically unexplained physical symptoms (MUPS) are physical symptoms for which adequate evidence of an underlying pathophysiological process cannot be identified after appropriate examination and investigation. MUPS are common in all healthcare settings. Up to one-third of all people presenting with physical symptoms have MUPS (Kirmayer 2004). The presence of MUPS is the key feature of conditions known as somatoform disorders. The Diagnostic and Statistical Manual of Mental Disorders (DSM; APA 2000) and International Classification of Diseases (ICD; WHO 2004) describe four somatoform diagnostic categories that include MUPS as their main indication. These categories are: somatisation disorder, (persistent somatoform) pain disorder, undifferentiated somatoform disorder, and unspecified somatoform disorder. The ICD also describes a fifth category: somatoform autonomic dysfunction disorder. All these disorders are established through a validated psychiatric diagnostic interview. Many different diagnostic revisions of somatoform disorders have been suggested and used in research since the early 2000s. Examples of proposed revised diagnoses include abridged somatisation disorder (Escobar 1998), multisomatoform disorder (Kroenke 1997), bodily distress disorder (Fink 2007), and complex somatic symptom disorder (Dimsdale 2009). These alternative diagnoses have their own diagnostic criteria, mainly based on symptom counts. Finally, in some fields, MUPS are not described as a feature of a specific disorder, but as a health problem in their own right. As a result, the treatment of MUPS in general is also described in literature, for example, in primary care research. Assessing the presence of MUPS is usually based on the combination of a validated somatic symptom scale, the duration of symptoms, and clinical judgement by the physician.

Somatoform disorders and MUPS may lead to functional impairment, high levels of psychological distress, a reduced quality of life, and a troubled doctor-patient relationship (Escobar 1987; Gureje 1997; Ring 2004; Zoccolillo 1986). Furthermore, chronic MUPS may lead to absence from work, fragmented and high utilisation of health care, and the associated high costs for society (Konnopka 2012; Kroenke 1989; Smith 1986).

DSM 5 describes the ‘somatic symptom disorder’ (SSD), which requires explicit cognitive criteria (e.g. excessive and disproportionate thoughts, feelings, and behaviours regarding symptoms) (APA 2013). The diagnosis does not require the somatic symptoms to be medically unexplained, and is therefore not taken into account in this review.

**Description of the intervention**

In previous decades, many pharmacological and non-pharmacological interventions for somatoform disorders and MUPS were developed. Pharmacological interventions are described in a separate Cochrane review (Kleinstäuber 2014) and this review only focuses on non-pharmacological interventions.
Most non-pharmacological interventions for MUPS focus on addressing cognitions, behaviour, coping styles, and functional consequences of symptoms. These interventions include psychological therapies as well as physical therapies. In the paragraph below, we described examples of several frequently studied forms of psychological and physical therapies.

**How the intervention might work**

**Psychological therapies - cognitive behavioural therapy (CBT)**

CBT is based on the cognitive behavioural model (Deary 2007). This model proposes that MUPS are caused by a self-perpetuating multi-factorial cycle, based on the interaction of different factors in several domains, including somatic (physical) aspects, cognitions (thoughts), behaviour, emotions, and environment (Sharpe 1992). CBT focusses on addressing cognitions and behaviours that patients have in interaction with their symptoms.

ReattrIBUTion is a specific form of CBT (Goldberg 1989). This method aims to encourage people to reattribute their MUPS to physiological or psychosocial causes rather than to somatic causes. Problem-solving treatment is another form of CBT that has been used for people with MUPS and somatoform disorders. The aim is to reduce complaints associated with unresolved problems in daily life by enhancing a person’s problem-solving capacities in a step-by-step manner (Malouff 2007).

**Psychological therapies - behavioural therapy**

Behavioural therapy, the second group, aims to constructively change a person’s behaviour towards their symptoms using operant conditioning in which a response in a certain context is followed by a reinforcing stimulus or consequence, thereby increasing the likelihood that the same response will follow in future. Biofeedback therapy is an important behavioural intervention relevant to this review. Other forms of behavioural therapy include relaxation therapy (Loew 2000), and psycho-education (Guerney 1971).

**Other psychological therapies**

A third group of psychological therapies, more aimed at increasing insight, such as:

1. third-wave cognitive behavioural therapy (i.e. the development of a new attitude towards symptoms, based on self regulation of attention and acceptance) (van Ravensteijn 2013)
2. psychodynamic therapies, a form of depth psychology, which focusses on revealing the unconscious content of a person’s psyche in order to alleviate psychological of physical tension (Noyes 2008)
3. humanistic therapies, focusing on self development, growth, and responsibilities. Treatment aims to help individuals recognise their strengths, creativity, and choices in the ‘here and now’
4. integrative therapies, which integrate components from several theoretical schools, which aims to work with the person to identify procedural sequences, chains of events, thoughts and emotions that explain how a target problem is established and maintained.
**Non-pharmacologic interventions**

**Enhanced care**
Another group of therapies offered to people with MUPS is enhanced care. Within these therapies people receive care as usual (mostly by their general practitioners (GP)), enhanced with, for example, participant education or structured counselling moments (Rosendal 2013). Within these therapies, there is no specific treatment agenda or structure; the aim is to offer the person some tools to assist in the recovery process, stimulating self-management.

**Physical therapies - physical activity training**
Several studies have indicated that mental health, including mood, pain thresholds, and sleep, can be improved by low- or moderate-intensity activity (Weyerer 1994). Graded activity training is an operant-conditioning behavioural approach in which physical activity is expanded step by step, based on a predetermined time schedule. It focusses on changing the fear-avoidance behaviour that patients with MUPS may have for particular physical activities (Lindström 1992).

**Other physical therapies**
Other examples of physical therapies for somatoform disorders and MUPS include activation therapy, where physical and behavioural activation is increased in a step-wise fashion, and running therapy, where running is used therapeutically, mainly to influence the level of stress.

**Why it is important to do this review**
Although there are several reviews on non-pharmacological interventions for somatoform disorders and MUPS, a complete overview of the whole spectrum is missing. Some reviews did not include a meta-analysis (Edwards 2010; Sumathipala 2007), while other reviews included only specific treatment types (Kroenke 2000; Nezu 2001), or applied restrictions on diagnostic types of MUPS or on treatment setting (Allen 2002; Edwards 2010; Kleinstäuber 2011; Rosendal 2013). Furthermore, currently there are no reviews that evaluate variations in treatment effects on the basis of diagnosis and severity of symptoms at baseline, setting, or healthcare provider.

**Objectives**
To assess the effects of non-pharmacological interventions for somatoform disorders (specifically somatisation disorder, undifferentiated somatoform disorder, somatoform disorder unspecified, somatoform autonomic dysfunction, pain disorder, and alternative somatoform diagnoses proposed in the literature) and MUPS in adults in comparison with treatment as usual, waiting list controls, attention placebo, psychological placebo, enhanced or structured care, and other psychological or physical therapies.
Methods

Types of studies
We included randomised controlled trials (RCTs) and cluster randomised controlled trials (CRCTs). We also planned to include data from the first phase of cross-over trials, but we identified no such trials that met our inclusion criteria. We excluded quasi-randomised trials (e.g. allocation to the study group by day of the week).

Types of participants
Participant characteristics
Participants had to be at least 18 years old. We applied no maximum age, as the condition can be present at any age. We placed no restriction on gender or culture.

Diagnosis
1. Participants had to meet the criteria for a somatoform disorder according to DSM III (APA 1980), DSM IV-TR (APA 2000), ICD-9 (WHO 1975), or ICD-10 (WHO 2004), or the criteria for one of the alternative somatoform diagnoses proposed in the literature. The primary diagnosis (a somatoform disorder) had to be made on the basis of a structured clinical interview or diagnostic checklists.
2. Participants were characterised with MUPS as their primary problem, on the basis of a validated scale for the assessment of MUPS.

Co-morbidities
As we aimed to summarise interventions for multiple symptoms, we excluded studies and reviews that examined participants diagnosed with only one specific functional syndrome or symptom.

Types of interventions
Experimental interventions
Eligible studies included one or more of the following experimental interventions.
1. Psychological therapies: CBT, behavioural therapy, third-wave CBT, psychodynamic therapies, humanistic therapies, integrative therapies.
2. Physical therapies: physical activity training, other physical therapies.
We excluded interventions based on complementary medicine from this review. In addition, pharmacological interventions and solitary consultation letter interventions were beyond the scope of this review; they were evaluated in other Cochrane reviews (Hoedeman 2010; Kleinstäuber 2013).

Comparator interventions
We accepted the following comparator interventions
1. Normal/usual treatment or waiting list procedures.
2. Attention or psychological placebo.
3. Enhanced or structured care.
4. Other psychological therapies.
5. Other physical therapies.
Non-pharmacologic interventions

**Types of outcome measures**
We included studies that met the inclusion criteria described above regardless of whether they reported on the following outcomes

**Primary outcomes:**
1. Severity/intensity of somatic symptoms.
2. Acceptability.
4. Depression and anxiety.
5. Dysfunctional cognitions, emotions, or behaviours (participant-rated).
6. Adverse events.
7. Treatment response (responder versus non-responder).
9. Health care use.

**Timing of outcome assessment**
We analysed primary and secondary outcomes at the following time points, if available: immediately post treatment; within 12 months after treatment ending; and more than 12 months after treatment ended.

**Search methods for identification of studies**

*Electronic searches*
The Cochrane Depression, Anxiety and Neurosis Review Group’s Specialized Register (CCDANCTR)
The Cochrane Depression, Anxiety and Neurosis Group (CCDAN) maintains two clinical trials registers at the editorial base in Bristol, UK: a references register and a studies-based register. The CCDANCTR-References Register contains over 36,000 reports of RCTs in depression, anxiety and neurosis.
1. We searched the CCDANCTR (Studies and References Registers) on 29 November 2013 using the following free-text terms: (somatization or somatisation or somatoform or hysteri* or briquet or polysymptom* or multisomatoform or somatizer* or (somatic NEAR symptom*) or (MUPS or “medical* unexplained” or “unexplained medical*” or (unexplained NEAR (symptom* or syndrom*)) or “frequent attend*” or (multiple NEAR (“physical symptom*” or “symptom diagnos*”)) OR neurastheni*). We screened the records retrieved manually for non-pharmacological interventions.
2. We conducted complementary searches on the following bibliographic databases using relevant subject headings and search syntax, appropriate to each resource: Cochrane Central Register of Controlled Trials (CENTRAL) (all years); Web of Science (from 1945 onwards, cited references search only (April 2014)).
3. To identify ongoing trials, we searched the ClinicalTrials.gov register (clinicaltrials.gov), the Current Controlled Trials metaRegister of Controlled Trials - active registers (mRCT; www.controlled-trials.com/mrct/), and the WHO International Clinical Trials Registry Platform Search Portal (www.who.int/trialsearch).
Searching other resources
We searched the following other resources: grey literature, handsearching for conference proceedings (see full review for more information), reference lists, correspondence

Data collection and analysis
Selection of studies
In the first step, two review authors (NvD, MdB) independently screened the titles and abstracts of reports identified from the literature search. We discarded studies that obviously did not fulfil the inclusion criteria at this stage of the screening process. Two review authors (NvD, MdB) retrieved eligible or potentially eligible articles for full-text assessment. We identified and excluded duplicate records and we collated multiple reports that related to the same study so that each study - rather than each report - was the unit of interest in the review. After full-text assessment, the review authors identified studies for inclusion and exclusion. We recorded reasons for exclusion of studies, and resolved disagreements by consensus - if necessary with the involvement of a third review author (JvdW).

Data extraction and management
We used a data collection form, piloted on one study in the review, to extract study characteristics, participants’ characteristics and outcome data. Independently, four review authors (NvD, MdB, HvdW, HvM) extracted study characteristics and outcome data from included studies. If necessary, we contacted the authors of trial reports for clarification or for additional information. We organised data using the most recent version of Review Manager 5 software (RevMan 2012). Two review author (NvD, HvdW) entered data into Review Manager 5 for analysis (RevMan 2012). We double-checked that data had been entered correctly by comparing the data presented in the systematic review with the data in the study reports. A third review author (MdB) spot-checked study characteristics for accuracy against the trial reports.

Main comparisons
Based on the available evidence we present the following comparisons:
1. psychological therapy versus usual care (or waiting list procedures)
2. psychological therapy versus enhanced (or structured) care
3. psychological therapy versus another psychological therapy

Assessment of risk of bias in included studies
Independently, two review authors (NvD, MdB) assessed the risk of bias for each study using the criteria outlined in the Cochrane Handbook for Systematic Reviews of Interventions (Higgins 2011). We resolved any disagreements by discussion or by involving another review author (HvM, HvdW).

Measures of treatment effect
Dichotomous data
For dichotomous outcomes, we used risk ratio (RR) as the summary statistic, together with 95% confidence intervals (CI).
Non-pharmacologic interventions

Continuous data
As different measures were used to assess the same outcome, we pooled data using the standardised mean difference (SMD); we calculated 95% CI. Specific attention was paid to the secondary outcome ‘functional disability and quality of life’, as the direction of scales for these outcomes can differ.

Assessment of heterogeneity
We assessed the groups for clinical similarities including elements such as age, gender, and setting. First, we assessed statistical heterogeneity visually by inspecting forest plots of standardised mean effect sizes and of relative risks. We used the I² statistic as a second test: I² describes the percentage of variability in effect estimates that is due to heterogeneity rather than chance. We used conventions of interpretation defined by Higgins (Higgins 2011). In the case of substantial levels (i.e. where I² = 50% to 90%) and considerable levels (I² = 75% to 100%) of heterogeneity, we explored data further by means of subgroup and sensitivity analyses (see below).

Assessment of reporting biases
We created funnel plots (treatment effect versus standard error of the effect size), if we included at least 10 trials in a meta-analysis, according to the recommendations of the Cochrane Handbook for Systematic Reviews of Interventions (Higgins 2011; Sterne 2011).

Data synthesis
If we found two or more included studies in a comparison category (see Data extraction and management) that used the same outcome construct, we performed a meta-analysis of the results. We expected to find high heterogeneity in non-pharmacological therapy approaches and in symptom severity, duration of symptoms and co-morbidities among the various study populations. Therefore, we analysed dichotomous and continuous treatment effects using a random-effects model.

Subgroup analysis and investigation of heterogeneity
If analysis of heterogeneity indicated significant heterogeneity, we performed subgroup analysis in order to explore whether methodological and clinical differences between the trials had produced systematic differences observed in the treatment outcomes. If we found statistically significant differences between subgroups, we reported the results of the corresponding subgroup meta-analysis.

When available data allowed (at least 10 studies), we performed subgroup analyses based on the following factors (only for the primary outcomes).
1. Severity, based on symptom count, and duration of MUPS at baseline.
2. Diagnosis at baseline.
3. Somatoform disorders versus alternative somatoform diagnoses versus MUPS.
4. Psychiatric and somatic co-morbidity
5. Primary care versus secondary care and tertiary care
6. Treatment as usual versus a waiting list procedure as a control intervention.
**Sensitivity analysis**

We planned sensitivity analyses to evaluate the robustness of the conclusions in conjunction with decisions made during the review process. In case of sufficient data (i.e. at least 10 studies), we planned these analyses to examine the effects of the following options, restricted to the primary outcome.

1. Exclusion of CRCTs.
2. Exclusion of studies with unclear allocation concealment
3. Exclusion of studies with unclear methods of sequence generation.
4. Exclusion of trials where missing data have been imputed.
5. Exclusion of studies with a drop-out rate higher than 20%.

**Results**

**Description of studies**

**Results of the search**

Searches were conducted up to November 2013 (CCDAN registers) and April 2014 (cited reference searches). Figure 1 shows the flow diagram of the study selection.

![Study flow diagram](image)
### Literature database searches

In the search of the CDANCTR-Studies and CCDANCTR-Reference Register (from now on referred to as CCDAN database), we found 929 abstracts after de-duplication. We excluded 842 records, based on the title and abstract, leaving 82 references (65 studies) selected for full-text retrieval. After reading the full-text, we judged 27 studies (49 articles) eligible for inclusion in this review. We excluded 35 studies (38 articles) and six articles are still awaiting assessment, for example, due to unavailability of a full-text article or difficulties in contacting authors.

The search of CENTRAL database found 995 records. After removing duplicates from the CCDAN search, there were 568 new references. After title and abstract screening, we excluded 560 references, and selected eight articles for full-text reading. After full-text reading, we excluded five articles, and judged three articles eligible for inclusion; however, all three articles described studies already included in the review (e.g. long-term follow-up results) (Gili 2014; Schröder 2013; Zonneveld 2012a). As the Schröder article reported a more detailed trial methodology and higher number of participants, we decided to use this article as the main reference of this study (Schröder 2013) instead of Zaby 2008, which was retrieved from the CCDAN search.

We performed a cited reference search on the Web of Science, for citations to primary reports of all studies expected to be included in this review. When handsearching the retrieved articles, we identified three additional relevant references. After full-text reading, we included one new study (Burton 2012), and excluded one article due to randomisation method (Rembold 2011). One article (Chernyak 2014) described an already included study (Sattel 2012).

### Grey literature

We performed searches for grey literature but found no new articles. We screened the conference proceedings and found no new articles.

### Systematic reviews

We found 14 reviews about (specific) non-pharmacological interventions for somatoform disorders or MUPS. After title screening in the reference lists of the reviews, we selected seven additional articles for screening of abstract. After abstract reading, we excluded four articles, and selected three articles for full-text reading. However, none of these three articles were eligible due to lack of randomisation or inappropriate selection method (Hellmann 1990; Lupke 1996; Tschuschke 2007).

### Trial registers

We performed ongoing trial searches in the databases of www.clinicaltrials.gov, www.controlled-trials.com, and www.who.int/trialsearch. We found six potentially eligible ongoing trials. As full details of the design and study results were not available, we could not included these studies in the review.
Contacting authors
We tried to contact 10 trial authors for missing information regarding the eligibility of studies; four responded and provided the desired information (Lidbeck 1997; Gottschalk 2011; Pols 2008; Steel 2011). We contacted authors of 20 of the included studies for additional information regarding study design and outcomes, of which 10 provided requested data.

Included studies
We included 21 studies, reported in 43 publications, in this review. All included studies concerned psychological interventions.

Design
Twenty of the included 21 studies had a parallel-group, individually randomised design (RCT). One study had a cluster-randomised design (Schaefert 2013). We found not trials with a cross-over design. The total number of randomised participants was 2658, a mean number of 127 per study (range 32 to 328).

Setting
Eight studies recruited participants in primary care only (Burton 2012; Escobar 2007a; Moreno 2013; Schaefert 2013; Schilte 2001; Sumathipala 2000; Sumathipala 2008; Van Ravesteijn 2013a). Only two studies recruited in secondary care (e.g. outpatient clinics) (Sattel 2012; Speckens 1995), and one study recruited inpatients in hospitals (Schweickhardt 2007). Seven studies recruited via medical settings as well as the open population (e.g. through advertisements) (Allen 2006a; Kashner 1995; Katsamanis 2011; Kolk 2004; Martin 2007; Schröder 2013; Zonneveld 2012a). Three studies recruited via primary care as well as secondary care (Fjorback 2013a; Lidbeck 1997; Schröder 2012).

In one study, treatment was performed in group sessions by GPs in primary care who were trained in the specific psychological technique, combined with a psychosomatic specialist (Schaefert 2013). In six other studies, treatment took place at a department of psychiatry or psychology (Allen 2006a; Escobar 2007a; Fjorback 2013a; Katsamanis 2011; Kolk 2004; Van Ravesteijn 2013a). Another six studies treated participants in other outpatient clinics (Kashner 1995; Lidbeck 1997; Schröder 2012; Sumathipala 2000; Sumathipala 2008; Zonneveld 2012a). Five studies treated participants in specific outpatient symptom clinics or outpatient clinics for psychosomatics (Burton 2012; Martin 2007; Sattel 2012; Schröder 2013; Speckens 1995). One study treated participants as inpatients (they were admitted to a ward) (Schweickhardt 2007). One study treated participants at home (Schilte 2001a). Finally, in one study the treatment setting was unknown (Moreno 2013).

Participants
Most studies recruited more women than men. Only one study reported more men in the intervention group (56%, Speckens 1995). The mean age was 43 years in all included studies, ranging from
35 years (Kolk 2004; Sumathipala 2008) to 49 years (Martin 2007; Schaefert 2013). Diagnostic criteria and inclusion criteria varied widely between studies. Fourteen studies used standardised diagnostic interviews (such as CIDI or SCID) to establish the diagnosis, the other seven studies used standardised questionnaires (such as SOMS or PHQ-15). Exclusion criteria varied between studies, but often included dementia, severe psychopathology such as psychosis, active suicidal thoughts, alcohol dependence, pregnancy, and current psychological therapy. Eleven studies reported severity of symptoms at baseline in terms of number of symptoms. This number varied widely, ranging from a lifetime number of seven symptoms (Martin 2007), to a current number of 32 symptoms (Schröder 2012). Only nine studies reported exact duration of symptoms, all of them reported a duration of at least several years, ranging from 3.5 years (Sumathipala 2008) to 25 years (Allen 2006a).

**Interventions**

As described in the ‘Types of interventions’-section, we aimed to select studies investigating psychological therapies, as well as studies on physical therapies. However, we found no studies on physical therapies that were eligible for inclusion. All 21 included studies evaluated a form of psychological therapy. Fourteen studies described certain forms of CBT. Two studies evaluated behaviour therapies (Katsamanis 2011; Schilte 2001a). Two studies described third-wave CBT (mindfulness) (Fjorback 2013a; Van Ravesteijn 2013a), and two studies described psychodynamic therapies (Sattel 2012; Schaefert 2013). In the study of Kolk et al., participants received CBT, client-centred or eclectic therapy, depending on the therapist the participant was assigned to (Kolk 2004); we classified this as integrative therapy.

The duration of treatment ranged from one day (one single session) (Martin 2007) to nine months (Schaefert 2013), most often between one and three months. The mean number of sessions varied among studies and ranged from one session (Martin 2007) to 13 sessions (Zonneveld 2012a). Reported duration of follow-up varied between two weeks (Schweickhardt 2007) and 24 months (Schilte 2001a).

**Comparisons**

Fifteen studies compared an intervention to usual treatment or a waiting list. None of the included studies described a placebo comparator intervention, but five included studies compared an intervention with enhanced or structured care (i.e. more than just usual care or a waiting list condition) (Fjorback 2013a; Sattel 2012; Schröder 2012; Speckens 1995; Sumathipala 2000). One study used compared two psychological therapies (Schröder 2013). We found no studies that compared psychological interventions with physical therapies.

In one study, GPs in both study arms were trained in diagnosis and management of medically unexplained symptoms (Schaefert 2013). In addition, the GPs in the intervention group conducted group sessions for people with MUPS, together with a psychosomatic specialist.
In six studies, in both study arms a consultation letter was sent to the primary care physician after baseline assessment, in addition to the planned psychological therapy (Allen 2006a; Escobar 2007a; Fjorback 2013a; Katsamanis 2011; Moreno 2013; Schröder 2012). This was not a reason for exclusion, and we categorised these studies according to the main comparison.

**Scale used to measure outcomes**

Severity or intensity of somatic symptoms (or both) was most often measured using the PHQ-15 (Kroenke 2002, five studies), the subscale ‘Somatisation’ of the SCL-90R (Derogatis 1986, five studies), the SOMS (Rief 1997, three studies), the BSI (Derogatis 1983, three studies), and the Clinical Global Impression Scale for Somatoform Disorders (CGI-SD; APA 2000, three studies). Acceptability was calculated from provided data about the total number of randomised participants in the study groups and the total number of participants who completed assessments at end of treatment (also in the control group). For anxiety and depressive symptoms, studies mostly used Hospital Anxiety and Depression Scale (HADS) subscales (Zigmond 1983, five studies), SCL-90R subscales (Derogatis 1986; three studies), and PHQ-9 (Kroenke 2001, three studies). Dysfunctional cognitions, emotions, and behaviours were mostly operationalised as health anxiety, and measured by the WI (Pilowsky 1967, seven studies). Withdrawals due to adverse events were incidentally registered by the authors. Treatment response was mainly registered using the clinician-rated CGI (Guy 1976, three studies). Functional disability was mainly measured using the SF-36 (Ware 1992, 10 studies) and the Short Form 12 Questionnaire (SF-12; a shortened version of the SF-36, three studies). No standardised questionnaire was used to measure healthcare use, and mostly consultation counts (as counted by physicians or participants) were reported.

**Risk of bias in included studies**

We classified the methodological quality of the 21 studies according to The Cochrane Collaboration’s tool for assessing the risk of bias. A risk of bias summary graph (figure 2) is presented.
Non-pharmacologic interventions

Figure 2.
Risk of bias summary:
judgements about each risk of
bias item for each included study.
Allocation (selection bias)
Sequence generation
While all studies specified that participants were randomly allocated to conditions (or GP practices randomised to treatment or control conditions), there were two studied that did not describe how sequence generation was performed (Kashner 1995; Schröder 2013). Therefore, we rated them as ‘uncertain’.

Allocation concealment
For five studies, it was unclear who performed allocation, or whether the person allocating participants to the trial groups was independent. Therefore, we rated these studies ‘unclear’ (Allen 2006a; Escobar 2007a; Kashner 1995; Katsamanis 2011; Schröder 2013). We rated the remaining 16 studies ‘low risk of bias’ as there was an adequate description of the person performing allocation or the relation to the researchers and therapists.

Blinding (performance bias and detection bias)
Blinding of participants and personnel
In 18 studies, blinding of participants and personnel was not possible, due to the nature of the interventions (e.g. psychological group therapy versus waiting list). As this may have influenced the judgement, we rated almost all studies ‘high risk’. We rated two studies ‘unclear’ because one of the two groups (participants or personnel) was blinded and the other was not (Sumathipala 2000; Van Ravesteijn 2013a). One study did not describe blinding of personnel (Schröder 2013), and, therefore, we rated it ‘unclear’.

Blinding of outcome assessment
In 19 studies, blinding of outcome assessment was not possible as most outcomes were participant reported. In one study, outcomes were assessed by blinded interviewers, but they did this together with the participants (who were not blinded) (Escobar 2007a). We rated this study ‘unclear’. One study mainly used clinician-rated instruments (Moreno 2013). The outcome assessor was blinded, but, as there also were a few participant report instruments (and participants were not blinded), we rated this study ‘unclear’.

Incomplete outcome data (attrition bias)
All studies reported follow-up rates; nine (43%) studies reported a loss to follow-up of 20% or less. We rated these studies ‘low risk’. We rated one study ‘unclear’, because it had a high loss to follow-up, but corrected for this statistically by multiple imputation (Fjorback 2013a). The remaining 11 studies reported high loss to follow-up (greater than 20%) and, therefore, we rated them ‘high risk’.

Selective reporting (reporting bias)
Seventeen studies reported all intended outcomes and, therefore, we rated them ‘low risk’. For one study, a protocol was lacking, therefore it was impossible to evaluate the possibility of selective outcome reporting (Fjorback 2013a). We rated this study ‘unclear’. We rated the remaining three
Non-pharmacologic interventions

studies ‘high risk’. In Kashner 1995, the outcome ‘days in bed’ was described as assessed, but was not reported in the article. In Moreno 2013, healthcare use and CGI were mentioned as outcomes in the protocol, but they were not reported. Schilte 2001a performed follow-up measurements at six, 12, and 24 months, but only reported outcomes of the last follow-up moment.

Treatment fidelity
Sixteen studies used a treatment manual or protocol for studied treatments. We rated them ‘low risk’. Three studies did not apply a structured intervention according to a protocol (Burton 2012; Kolk 2004; Schilte 2001a), therefore, we rated them ‘high risk’. The two remaining studies did provide information about a form of structure in treatment, but did not mention a protocol or manual for this. We rated them ‘unclear’.

Researcher allegiance
In 18 studies, researchers did not report to have a preference for one of the treatment modalities. In the studies of Burton 2012, Lidbeck 1997, and Schaefert 2013, an author was also (one of) the therapist(s), which may have caused some bias. Therefore, we rated these studies ‘unclear’.

Other potential sources of bias
We included two multiple intervention studies (Moreno 2013; Schröder 2013). In the first study, data were presented for each groups to which participants were randomised, so no other potential sources of bias were found (rating: ‘low risk’). In the second study, participants were randomised for CBT or progressive muscle relaxation (PMR) using random sequences. When both groups were full, newly included participants were allocated to the waiting list group. In a later stage, these participants were included in both intervention groups. As participants were their own controls due to this method, we decided to exclude data from the waiting list group from analysis. For this reason, we rated this study ‘unclear’.

One of the studies was a CRCT (Schaefert 2013). GPs were randomised, after which individuals were recruited. We considered the randomisation method and statistical analysis appropriate for the study design.

In the studies of Schilte 2001a and Katsamanis 2011, we found considerable baseline imbalances. In the study of Schweickhardt 2007, a high percentage (29%) of participants from the control group became involved in psychotherapy. This may have influenced the results, although this study provided data for only one outcome (acceptability) and the effects were in the same order of magnitude as in other studies. We rated these three studies ‘unclear’.

Effects of interventions
For the description of the results, we stratified the comparisons as per the categories of therapies presented in ‘Types of interventions’. Most studies provided data for some of the outcomes. One study did not provide any data that were suitable for meta-analysis, because the authors only
reported change scores (Schilte 2001a). Across all comparisons, outcomes, and time points, we created 44 forest plots. Below, we present the results of the meta-analyses. We also give attention to the subgroups that included a considerable proportion of the studies contributing to the overall comparisons: CBT versus usual care or waiting list and CBT versus enhanced or structured care.

1. Psychological therapy versus usual care or waiting list

Fifteen studies, with 1805 randomised participants, compared some form of psychological therapy with usual care or waiting list controls. They addressed the following psychological therapies:

1. CBT versus usual care or waiting list: 10 studies, 1037 randomised participants (Allen 2006a; Burton 2012; Escobar 2007a; Kashner 1995; Lidbeck 1997; Martin 2007; Moreno 2013; Schweickhardt 2007; Sumathipala 2000; Zonneveld 2012a);
2. Behavioural therapy versus usual care or waiting list: two studies, 209 randomised participants (Katsamanis 2011; Schilte 2001a);
3. Third-wave CBT versus usual care or waiting list: one study, 125 randomised participants (Van Ravesteijn 2013a);
4. Psychodynamic therapy versus usual care or waiting list: one study, 328 randomised participants (Schaefert 2013);
5. Integrative therapies versus usual care or waiting list: one study, 106 randomised participants (Kolk 2004).

In four of these studies, this was combined with a consultation letter sent to the primary care physician after baseline assessment, in both treatment arms (Allen 2006a; Escobar 2007a; Katsamanis 2011; Moreno 2013).

Below, we described results for the whole group and for the subgroup of studies that compared CBT with usual treatment.

Primary outcomes

1.1 Severity of somatic symptoms

Combining all studies that compared some psychological therapy with usual care or waiting list, psychological therapies were significantly more effective at end of treatment, though the effect was small (SMD -0.34; 95% CI -0.53 to -0.16; 10 studies, 1081 analysed participants). Heterogeneity was moderate ($I^2 = 49\%$), and the overall quality of the evidence was low. Compared with usual care, the subgroup of studies that used CBT were also significantly more effective in reducing severity of symptoms at end of treatment (SMD -0.37; 95% CI -0.69 to -0.05; 6 studies, 593 participants, random-effects model). Heterogeneity was substantial ($I^2 = 70\%$), and the overall quality of the evidence was low. The point estimates of all but one of the studies favoured the CBT group. The two studies with the smallest effects offered low-intensity CBT (Burton 2012; Martin 2007). A post-hoc analysis without these two studies provided an SMD of -0.58 (a moderate effect size) (95% CI -0.77 to -0.38) and reduced heterogeneity ($I^2 = 0\%$).
At follow-up, measurements within one year of follow-up, the effect of psychological therapies remained significant (SMD -0.24; 95% CI -0.37 to -0.11; 7 studies, 950 participants; I² = 0%). The same was the case for the subgroup of CBT studies (SMD -0.29; 95% CI -0.49 to -0.09; 4 studies, 496 participants). Heterogeneity was low (I² = 17%).

Only two studies (all of CBT) with 228 participants provided data for this severity of symptoms beyond one year of follow-up (SMD -0.52; 95% CI -0.80 to -0.24). Heterogeneity was low (I² = 0%).

**Figure 3:** Psychological therapies versus usual care or waiting list controls.

### 1.2 Acceptability

Compared with usual care, psychological therapies resulted in a higher proportion of drop-outs (RR acceptability 0.93; 95% CI 0.88 to 0.99 favouring usual care; 14 studies, 1644 participants). Heterogeneity was moderate (I² = 70%). For the studies comparing CBT with usual care, results were of the same magnitude but no longer statistically significant (RR acceptability 0.93; 95% CI 0.85 to 1.01 favouring usual care; 10 studies, 1037 participants). Heterogeneity was considerable (I² = 78%). The overall quality of the evidence for this outcome was moderate.
Secondary outcomes

1.3 Severity of anxiety or depressive symptoms (or both)
For participant-rated anxiety symptoms, there was no significant difference at end of treatment (SMD 0.06; 95% CI -0.20 to 0.32; 4 studies, 270 participants). For the studies comparing CBT with usual care, results were similar (SMD 0.07; 95% CI -0.22 to 0.37; 3 studies, 185 participants). Within one year of follow-up only two studies were available (SMD 0.18; 95% CI -0.22 to 0.58; 134 participants).

For clinician-rated anxiety symptoms at end of treatment, there was a statistically significant difference at end of treatment in favour psychological therapies (SMD -0.40; 95% CI -0.63 to -0.17; 3 studies, 320 participants). Within and beyond one year of follow-up, differences remained statistically significant (within one year: SMD -0.66; 95% CI -1.15 to -0.18, 2 studies both CBT, 251 participants; beyond one year: SMD -0.91; 95% CI -1.26 to -0.55; 1 study, 156 participants). For participant-rated depressive symptoms, there was no significant difference at end of treatment (SMD -0.03; 95% CI -0.22 to 0.16; 6 studies, 661 participants). Similar results were found for the studies that compared CBT with usual care (SMD 0.09; 95% CI -0.13 to 0.31; 4 studies, 325 participants), and for outcomes after not more than one year of follow-up (SMD 0.04; 95% CI -0.34 to 0.42; four studies, 535 participants).

For clinician-rated depressive symptoms, there was a statistically significant difference at end of treatment in favour of psychological therapies (SMD -0.25; 95% CI -0.48 to -0.02; 3 studies, 316 participants). Within one year of follow-up, the difference was no longer statistically significant (SMD -0.55; 95% CI -1.17 to 0.07; 2 studies, 251 participants). Only one study reported on this outcome beyond one year after treatment (SMD -0.81; 95% CI -1.16 to -0.46; 156 participants).

1.4 Dysfunctional cognitions, emotions, and behaviours
Three studies, two of which compared CBT with usual care, with 440 participants, reported on dysfunctional cognitions, emotions, and behaviours. At end of treatment, there was no significant difference between the two groups (SMD -0.11; 95% CI -0.37 to 0.16). The quality of the evidence was moderate. At follow-up within one year, differences remained non-significant (SMD -0.16; 95% CI -0.38 to 0.07).

1.5 Adverse events
Only three studies, all comparing CBT with usual care, reported on adverse events during the treatment period. One study could not be included in the meta-analysis, because no adverse events were found in both groups. The pooled result of the other two studies also showed no significant differences between both conditions (RR 1.31; 95% CI 0.47 to 3.66; 445 participants; I² = 0%).
1.6 Treatment response (clinician rated)
All four studies addressing clinician-rated treatment response comparing CBT with usual care. At end of treatment, results strongly favoured the treatment group (RR 3.30; 95% CI 2.08 to 5.21; 4 studies, 391 participants; $I^2 = 19\%$). We considered the quality of the evidence to be low for this outcome. Three studies provided data for clinician-rated treatment response within one year after end of treatment, still in favour of the treatment group (RR 2.53; 95% CI 1.25 to 5.10; 332 participants; $I^2 = 59\%$). At longer follow up (greater than one year after treatment) only two studies reported outcomes, highly favouring the treatment group (RR 10.31; 95% CI 2.95 to 36.02; 240 participants).

1.7 Functional disability and quality of life
Seven studies, of which four addressing CBT reported on functional disability and quality of life, using a variety of instruments. At the end of treatment, a statistically significant effect was found favouring the psychological therapies (SMD 0.17; 95% CI 0.03 to 0.32; 7 studies, 730 participants; $I^2 = 0\%$). We judged the evidence to be moderate. At follow-up within one year after treatment, differences were similar but no longer significant (less than one year: SMD 0.16; 95% CI -0.01 to 0.33; 4 studies, 526 participants; $I^2 = 0\%$). After one year, only one study provided data for functional disability and quality of life.

Four studies compared CBT with usual care. At end of treatment, a non-significant difference was found favouring CBT (SMD 0.15; 95% CI -0.06 to 0.37; 4 studies, 341 participants; $I^2 = 0\%$).

1.8 Healthcare use
Six studies assessed healthcare use, operationalised in different ways, with moderate quality of evidence. During the treatment phase, two studies found a significant difference in the number of participant-initiated doctor visits and medication usage in favour of CBT (SMD -0.68; 95% CI -1.06 to -0.30; 117 participants). In the period less than one year after treatment, perhaps a more relevant timeframe, four studies found no clear evidence of a difference (SMD -0.09; 95% CI -0.31 to 0.12; 532 participants; $I^2 = 20\%$). We judged the quality of the evidence to be moderate. For one of the studies, the effect was in the opposite direction, that is, favouring the control group (Kolk 2004). No study provided data for healthcare use beyond one year after treatment. See footnotes of analyses for details about the way healthcare use was assessed.

2. Psychological therapy versus enhanced or structures care
Five studies with 680 randomised participants compared a certain psychological therapy with enhanced or structured care. They addressed the following treatments:
1. CBT versus enhanced or structured care: three studies, 349 randomised participants (Schröder 2012; Speckens 1995; Sumathipala 2008);
2. third-wave CBT versus enhanced or structured care: one study, 120 randomised participants (Fjorback 2013a);
3. psychodynamic therapy versus enhanced or structured care: one study, 211 randomised participants (Sattel 2012).
In two of these studies, treatment was combined with a consultation letter sent to the primary care physician after baseline assessment, in both treatment arms (Fjorback 2013a; Schröder 2012). Below we describe the main results, sorted by outcomes. Apart from CBT, only one or two trials provided data for each of the three other types of psychological therapy; hence, for each of these other treatment types there was insufficient evidence. Below, we do not describe the results for these subgroups separately. The reader is referred to the combined forest plots for each outcome.

**Primary outcomes**

2.1 Severity of somatic symptoms

Five studies (with 624 analysed participants) assessed severity of somatic symptoms comparing some psychological therapy versus enhanced care (pooled SMD -0.19; 95% CI -0.43 to 0.04; $I^2$ = 53%). We considered the quality of the evidence to be low. Within one year of follow-up, this effect was similar but now statistically significant (SMD -0.21; 95% CI -0.40 to -0.02; 5 studies, 593 participants; $I^2$ = 25%). Only two studies each comparing a different psychological therapy to enhanced care, assessed severity of somatic symptoms beyond one year after treatment (SMD -0.32; 95% CI -0.73 to 0.10; 172 participants). The subgroup of studies comparing CBT with enhanced care showed similar results. Heterogeneity was substantial at the end of treatment ($I^2$ = 62%) and moderate within one year after treatment ($I^2$ = 39%).

2.2 Acceptability

Five studies, with 679 analysed participants, showed that psychological therapies were less acceptable in terms of drop-outs than enhanced care (RR 0.93; 95% CI 0.87 to 1.00). Heterogeneity was moderate ($I^2$ = 36%), and we judged the quality of the evidence to be moderate. The largest subgroup was CBT. Compared with enhanced care, moderate-quality evidence showed that there was no clear difference between CBT and enhanced or structured care (RR 0.91; 95% CI 0.82 to 1.02; 3 studies, 331 participants). Heterogeneity was moderate to considerable ($I^2$ = 50%).

**Secondary outcomes**

2.3 Severity of anxiety or depressive symptoms (or both)

Five studies assessed severity of anxiety or depressive symptoms (or both) at end of treatment (SMD -0.14; 95% CI -0.30 to 0.02; 624 analysed participants; $I^2$ = 0%), showing no clear difference. Similar results were found within one year after treatment (SMD -0.13; 95% CI -0.29 to 0.03; 5 studies, 593 participants) and beyond one year after treatment (SMD -0.26; 95% CI -0.55 to 0.03; 2 studies, 184 participants).

The studies investigating CBT showed no significant difference in level of anxiety and depressive symptoms between CBT and enhanced care at end of treatment (SMD -0.17; 95% CI -0.40 to 0.05; 3 studies. 307 participants) and within one year after treatment (SMD -0.17; 95% CI -0.40 to 0.06; 3 studies, 289 participants). Heterogeneity was low ($I^2$ = 0% at end of treatment and within one year after treatment). Only one CBT study reported on severity of anxiety or depressive symptoms (or both) beyond one year after treatment.
Non-pharmacologic interventions

2.4 Dysfunctional cognitions, emotions, and behaviours
Four studies, with 499 analysed participants, provided data for dysfunctional cognitions, emotions, and behaviours at end of treatment, showing no clear evidence of a difference between psychological therapy and enhanced care (SMD -0.09; 95% CI -0.29 to 0.10; I² = 14%). We judged quality of the evidence to be moderate. At follow-up within one year after treatment, the difference was statistically significant (P value = 0.05), favouring the psychological therapy over enhanced care (SMD -0.24; 95% CI -0.49 to 0.00; 4 studies, 477 participants; I² = 42%). Beyond one year of follow-up, only two studies reported on dysfunctional cognitions, emotions, and behaviours and showed no significant difference (SMD -0.58; 95% CI -1.27 to 0.11; 2 studies, 184 participants; I² = 82%).

The two studies comparing CBT with enhanced care showed no clear evidence of a difference in dysfunctional cognitions, emotions, and behaviours at end of treatment (SMD -0.28; 95% CI -0.57 to 0.01; 2 studies, 182 participants). Heterogeneity was low (I² = 0%). However, within one year after treatment, levels of dysfunctional cognitions, emotions, and behaviours were significantly lower for CBT (SMD -0.45; 95% CI -0.83 to -0.07; 2 studies, 173 participants), though more heterogeneous (I² = 37%). This effect was even more significant at more than one year after treatment, although this comparison only included one study (SMD -0.94; 95% CI -1.36 to -0.51; 94 participants) (Schröder 2012).

2.5 Adverse events
None of the studies comparing psychological therapy versus enhanced or structures care reported information about adverse events.

2.6 Treatment response
None of the included studies comparing psychological therapy versus enhanced or structures care reported about treatment response using a standardised method as described in ‘Secondary outcomes’. One study reported about treatment response, but for this outcome measure the SF-36 (Ware 1992) was used (Schröder 2012). In this review, we used the outcomes of this questionnaire in the analyses of functional disability. Another study reported about participants’ perceived change in symptoms (Speckens 1995). At all measurement moments after baseline, participants were asked if their symptoms were "recovered", “improved”, “the same”, or “worse” since the previous measurement, using a non-standardised questionnaire. At the end of treatment, 32 (82%) participants in the intervention group declared that symptoms were improved or recovered versus 24 (64%) participants in the control group. Six months after treatment, 27 (73%) participants of intervention group reported recovery or improvement versus 23 (59%) participants of the control group.

2.7 Functional disability and quality of life
At end of treatment, four studies with 497 analysed participants reporting on functional disability and quality of life, found no significant difference (SMD 0.13; 95% CI -0.05 to 0.30; I² = 0%). We considered the quality of the evidence to be moderate. Within one year of follow-up, there was a
Chapter 8

small effect in favour of psychological therapies (SMD 0.20; 95% CI 0.02 to 0.38; 5 studies, 727 participants; I² = 0%). Only two studies reported on functional disability and quality of life beyond one year of follow-up and there was no clear evidence of a difference between the interventions (SMD 0.22; 95% CI -0.16 to 0.60; 2 studies, 184 participants).

For the studies comparing CBT with enhanced care, at end of treatment, moderate-quality evidence showed no significant difference in terms of level of function/quality of life, with a large CI but homogeneous population (SMD 0.21; 95% CI -0.08 to 0.51; 2 studies, 182 participants; I² = 0%). There was a small but significant difference in favour of CBT within one year after treatment (SMD 0.30; 95% CI 0.00 to 0.60; 2 studies, 173 participants). At this time point, heterogeneity was low (I² = 0%). After one year of follow-up, only one study provided data. In this study, CBT resulted in a significantly higher level of function compared with enhanced care (SMD 0.42; 95% CI 0.01 to 0.83; 94 participants).

2.8 Healthcare use

Only two studies provided usable data for this analysis and quality of the evidence was low (Sumathipala 2008; Sattel 2012). There were no significant differences healthcare use between psychological therapies and enhanced care, neither at end of treatment, nor within one year after end of treatment. See footnotes of analyses for details about the way healthcare use was assessed.

3. Psychological therapy versus other psychological therapy

Only one included study addressed psychological therapy versus other psychological therapy (Schröder 2013, 173 randomised participants). The study compared CBT with PMR therapy. The study also included a waiting list group, but we excluded this group from analyses as participants in the waiting list group were not randomly assigned.

Primary outcomes

3.1 Severity of somatic symptoms

No significant difference was found for severity of somatic symptoms between CBT and PMR at end of treatment (SMD 0.10; 95% CI -0.33 to 0.53; 84 participants).

3.2 Acceptability

There was no significant difference in drop-out rates between CBT and PMR during treatment (SMD 0.98; 95% CI 0.83 to 1.15; 90 participants).

Secondary outcomes

3.3 Severity of anxiety or depressive symptoms (or both) at end of treatment

There was no significant difference in level of depression and anxiety between CBT and PMR at end of treatment (SMD 0.01; 95% CI -0.42 to 0.44; 84 participants).
Non-pharmacologic interventions

3.4 Dysfunctional cognitions, emotions, and behaviours
The study did not report about dysfunctional cognitions, emotions, and behaviours.

3.5 Adverse events
The study comparing CBT with PMR did not report about adverse events.

3.6 Treatment response
The study comparing CBT with PMR did not report about treatment response.

3.7 Functional disability and quality of life
There was no significant difference in level of function between CBT and PMR at end of treatment (SMD 0.28; 95% CI -0.15 to 0.71; 84 participants).

3.8 Healthcare use
The study comparing CBT with PMR did not report about healthcare use.

Discussion

*Summary of main results*

**Psychological therapy versus usual care**

Fifteen studies compared some form of psychological therapy with usual care or a waiting list. Combining 10 of these studies, the psychological therapy was significantly more effective on symptom severity at end of treatment, though the effect was small. Heterogeneity was considerable and the overall quality of the evidence was low. Six of the 10 studies compared CBT with usual care; for this subgroup it was also apparent that CBT was more effective in reducing severity of symptoms at the end of treatment. The treatment effect of psychological therapies as a whole was also noted within one year of follow-up (seven studies). After one year, the evidence was limited to two studies (both CBT), but still in favour of the psychological therapy. Results for treatment response, one of our secondary outcomes, supported the findings for symptom severity, with moderate-quality evidence.

Regarding the other primary outcome, acceptability, we found a 7% difference in drop-outs, favouring the usual care group. The quality of the evidence was moderate. After we removed an apparent outlier, the result was smaller (5%), but still statistically significant. There was no significant difference in drop-out rates between CBT and usual care.

For participant-rated symptoms of depression and anxiety, there was no significant difference at the end of treatment or at follow-up. Only three studies reported. There was also no clear evidence of a difference in adverse effects and dysfunctional cognitions, emotions, and behaviours. There was a small difference in functional disability at the end of treatment favouring psychological therapies. This effect was not apparent during follow-up. Two studies (both on CBT) found a small difference
in favour of psychological therapies on healthcare use during treatment, four studies found no effect within one year of follow-up. Due to the small number of studies, these results should be considered with caution.

Only two studies compared behavioural therapy with usual care, of which only one provided relevant data (Katsamanis 2011). In this study, there were no significant differences for any of the outcomes. Only one study compared third-wave CBT (mindfulness therapy) with usual care (Van Ravesteijn 2013a). In this study, mindfulness was more acceptable than usual care, but no evidence of differences was found with respect to other outcomes. One study compared a variety of psychological therapies with usual care (therapy depended on the orientation of the 15 participating therapists) (Kolk 2004). In this study, there was no evidence of differences with respect to any of the outcomes.

**Psychological therapy versus enhanced or structured care**

Five studies compared a certain psychological therapy with enhanced or structured care. The quality of the evidence was moderate for most outcomes. At the end of treatment, there was no clear evidence of a difference for symptom severity, but there was a small statistically significant difference within one year after end of treatment. The psychological therapy groups had a 7% higher drop-out rate than the control groups.

There was no clear evidence of a difference between the groups in terms of severity of anxiety or depressive symptoms or dysfunctional cognitions, emotions, and behaviours at end of treatment. None of the studies in this comparison reported information about adverse events or treatment response in a standardised way. For functional disability and quality of life, there was no clear evidence of a difference at the end of treatment, but there was a small significant difference within one year of follow-up. There were no significant differences in healthcare use between psychological therapies and enhanced care.

Three of the studies compared CBT with enhanced or structured care. For symptom severity, CBT showed similar results as the whole group. There were no differences in drop-out rates. In addition, there were no significant differences in levels of anxiety and depressive symptoms. At the end of treatment, CBT did not result in lower levels of dysfunctional cognitions, emotions, and behaviours, compared with enhanced care. However, within one year of treatment, these levels were lower for CBT (two studies). The level of functional disability at the end of treatment was comparable for CBT and enhanced care. Within and after one year of treatment there was a small difference in favour of CBT, although only a few studies were included in these analyses. Only one CBT study reported data about healthcare use and found no evidence of difference.

**Psychological therapy versus another psychological therapy**

Only one study compared two forms of psychological therapy (CBT versus PMR). There were no differences between the groups for any of the outcomes.
**Studies**

21 studies could be included in this review. In comparison to other existing reviews about non-pharmacological interventions for MUPS or somatoform disorders (e.g. Kleinstäuber 2011; Kroenke 2007a; Rosendal 2013), this number of eligible studies is quite high. However, only a few studies contributed to most of the outcomes. We believe that the included studies cover a broad spectrum of settings and therapists, and both RCTs and CRCTs were included. A limitation of the included studies was the relatively low number of included participants per study as most studies only included 25 to 75 participants per study arm. With only two exceptions (Sumathipala 2000; Sumathipala 2008), studies were performed in developed countries (Western Europe and USA).

**Participants**

Most studies randomised more women than men. This is in line with existing reviews, as MUPS and somatoform disorders are more common among women. Included studies cover a broad age range. However, as the mean age of participants was in the 30s or 40s in most of the studies, it may be possible that younger and older people were relatively underpresented. The number of symptoms at baseline varied widely, ranging from a lifetime number of symptoms of seven (Martin 2007) to a current number of symptoms of 32 (Schröder 2012). Baseline duration, only reported in nine studies, ranged on average from four to 25 years. This suggests that most of the included participants may have had chronic symptoms at baseline. Included studies also reported high psychiatric co-morbidity rates, percentages of participants with a current co-morbid axis 1 disorder varied between 41% (Zonneveld 2012a) and 92% (Escobar 2007a). Taking these findings together, we can say that a limitation may be that participants of included studies were people with relatively severe forms of somatoform disorders and MUPS.

**Interventions**

Fourteen of the included studies compared CBT with another intervention. As a result, relatively robust conclusions could be drawn about the effectiveness of CBT. The number of studies describing other psychological therapies (such as behavioural therapies, third-wave CBT, or psychodynamic therapies) was too low to draw conclusions about these forms of therapy. Duration and number of treatment sessions varied widely between the included studies. It is especially remarkable that we found no studies on physical therapies (such as running therapy). We believe that there is a clear need for this type of research.

**Outcomes**

A problem of the current review was that, with one exception, there were not enough studies to assess reporting bias with funnel plots. According to recommendations of the Cochrane Handbook for Systematic Reviews of Interventions, there should be at least 10 studies to perform this (Sterne 2011). In future updates of this review, the addition of new studies may enable us to produce funnel plots for more comparisons and outcomes.
Adverse effects were very infrequently reported and various ways of reporting were used. Therefore, it was impossible to extract these in a standardised way in order to include them in our meta-analytical calculations, except for the first comparison (psychological therapy versus usual care or waiting list).

**Quality of the evidence**

In regard to different types of biases most of our included studies showed a low risk. However, a few specific domains were often rated as being at high risk of bias across the studies. Especially for blinding of the outcome measurement, we identified a high risk of bias in most of the included studies. Most studies could not blind the outcome reporters, mostly the participants, due to the nature of the intervention. A high risk of bias in blinding of participants and personnel was found for the same reason. Another study aspect that affected the quality of the evidence was the generally low number of included participants per study. The small number of studies did not allow us to assess the effects in subgroups of participants or interventions. Apart from CBT, all other comparisons between specific therapies and usual care or enhanced care, the number of studies was too small (often only one study).

**Potential biases in the review process**

This review has several methodological strengths. The quality of meta-analyses depends on the robustness of the search methods used. In this review, the electronic search was thorough and large in scale with broad parameters. We evaluated published and unpublished studies. The selection criteria were broad, which led to the selection of a relatively high number of studies. As a result, it seems likely that all or almost all evidence in the searched databases that should have been included was included. However, as we did not search Asian databases, this may have led to a potential bias.

The study was performed according to a pre-published protocol. However, due to the fact that not all choices that had to be made were foreseen, there were also post hoc decisions. Excluding studies that trained GPs to deliver some psychological therapy was one of these decisions. Another post hoc decision was the addition of enhanced or structured care as a comparator. We made decisions very carefully and included achieving consensus between several review authors with specific knowledge in the field. However, some studies were difficult to categorise, as, for example, treatments included elements of different treatment categories. Therefore, allocation of these studies remained slightly arbitrary. Other post hoc decisions were to carry out sensitivity analyses by excluding studies that included consultation letters in both study groups, and by excluding studies with the least intensive interventions.

Although acceptability was a primary outcome of our study, we restricted this to the period from randomisation to the end of treatment. We did not take into account the acceptability of the interventions in the recruitment phase. Participants for whom the intervention or control condition was unattractive probably did not participate. This affects the external validity of study findings.
Agreements and disagreements with other studies or reviews
Several systematic reviews have addressed non-pharmacological treatments for participants with some form of somatoform disorder or MUPS (Allen 2002; Blankenstein 2001 (thesis, chapter 2); Guthrie 1996; Hofmann 2012; Huibers 2007; Kleinstäuber 2011; Koelen 2014; Kroenke 2000; Kroenke 2007a; Looper 2002; Nezu 2001; Rosendal 2013; Sumathipala 2007). As many of the included studies in our review were published after 2005, we focused this discussion on the systematic reviews that were published after 2005.

In general, we can say that the results of this review are in line with results of existing reviews. In most reviews, the majority of included studies concerned CBT in some form, and small effect sizes were found. In other reviews, also limited evidence was found for other forms of psychological therapies. Studies investigating physical therapies for somatoform disorders or MUPS were also hardly reported in other reviews.

Authors’ conclusions

Implications for practice
The overall quality of the evidence provided by 21 randomised controlled trials was low to moderate. All psychological therapies combined were superior to usual care or waiting list condition for symptom severity, our first primary outcome, but effect sizes were small. As a single treatment, only cognitive behavioural therapy (CBT) was adequately studied to allow conclusions for practice. Compared with usual care or waiting list conditions, CBT reduced somatic symptoms, with a small effect and substantial differences in effects between CBT studies. The effects were durable within and after one year of follow-up.

Compared with enhanced or structured care, psychological therapies generally were not more effective for most of the outcomes. CBT was also not superior to enhanced care. The question remains how specific CBT is over structured improvements of care. No major adverse events were reported in the intervention groups, although most studies did not describe adverse events as an explicit outcome measure. Apart from CBT, neither psychological nor other non-pharmacological therapies have been adequately studied.

In daily practice, a substantial percentage of people with medically unexplained physical symptoms (MUPS) may not be willing to accept psychologically oriented treatments. Whether such acceptance is associated with the effect of psychological treatments for the total MUPS population was not clear. Due to the small number of studies, we could not draw conclusions about the effect of characteristics such as a profession and experience of the therapist, about treatment intensity and treatment location, on treatment efficacy.

Further optimisation of CBT to target optimal participant profiles and match treatment providers, treatment characteristics, and participants could improve outcomes. Motivating and preparing
people for CBT is important for this participant group (Timmer 2006). As drop-out rates were not much lower than in control groups, this indicates that when a person has accepted involvement in the treatment, the prospects that the treatment will be completed are good.

**Implications for research**

Based on the findings in this review, we can make several recommendations for future research. The number of studies investigating various treatment modalities other than CBT needs to increase to build a broader and more varied evidence-base for the treatment of somatoform disorders and MUPS. As physical therapies may offer a more acceptable starting point for treatment for these people than psychological approaches, investigating the effectiveness of physical therapies is to be considered.

Most studies in our review focused on chronic manifestations of physical symptoms, often of considerable severity. It is conceivable that interventions were more effective in people with milder symptoms, or of shorter duration, but this needs further testing. Preventing symptoms from become chronic may be a relevant outcome to be added in future studies.

In future research, more attention should be paid to the impact of interventions on risk factors for recurrence and persistence of symptoms in somatoform disorders and MUPS. These factors include anxiety; depression; and dysfunctional cognitions, emotions, and behaviours. Most included studies in this review did not report on all of these factors. Specific attention to the effect of treatment duration and number of treatment sessions is also needed. In the studies included in this review, duration and number of sessions varied widely, and it is yet unclear which treatment intensities are effective for which participants.

Psychological treatments were not superior to enhanced care. It could be argued that an active comparator such as enhanced care underestimates treatment effects. However, as this comparative treatment is probably cheaper than more intensive psychological interventions, it would deserve further study (cf. Rosendal 2013).

Finally, future studies should include more participants, preferably use a uniform set of validated outcome measurements, and extend follow-up assessments beyond one year after treatment.