UiT
THE ARCTIC UNIVERSITY OF NORWAY

Faculty of Health Sciences / Department of Community Medicine

Effect of group psychoeducation for major depressive disorder: a systematic review

Helene Sandberg and Åshild Roaldset

HEL-3950 Master's thesis in Public Health,
April 2019

Supervisor: Professor Rigmor C Berg, Norwegian Institute of Public Health

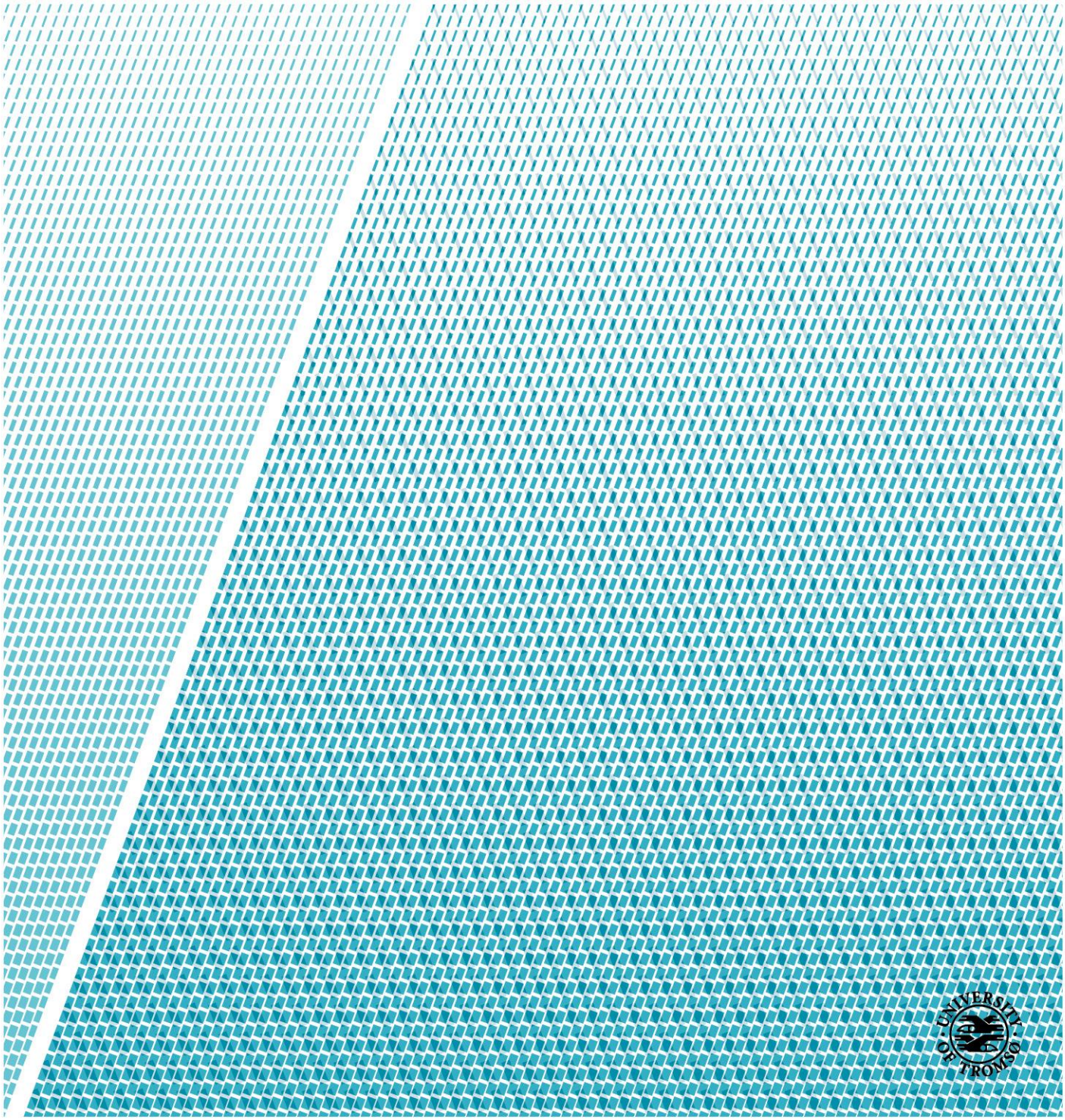


Table of Contents

Acknowledgements

We would like to thank our supervisor Professor /Department Director Rigmor C Berg at the Norwegian Institute of Public Health for her wise advise, enthusiasm and efficiency. She has been generous with her time. We would also like to thank Bror Just Andersen, PhD Psychol, prosjektleder Medisinfrie døgntilbud at Vestre Viken HF for inspiring discussion in the early stages, interest in our project and provision of background literature. We are also grateful to Research Librarian Lien Nguyen for her guidance and execution of the search strategy.

Abstract

Background: Depression is a common mental disorder and a major contributor to the overall global burden of disease. Healthcare systems struggle to provide effective and acceptable treatment to meet the needs of the growing number of patients suffering from depression. Although there are some known, effective medical treatments for depression, far from all of those affected receive such treatments, and there is a corresponding patient- and stakeholder demand for drug-free alternatives to treat depression. Group psychoeducation is a low threshold, drug-free intervention which has proven to be beneficial in the treatment of other mental disorders and which can be adapted to different populations. Use of group psychoeducation for major depressive disorder (MDD) will increase the availability of treatment, if proven to be effective, because it allows for treating several patients in the same session and meets calls for drug-free treatment.

Objective: To systematically review the effectiveness of group psychoeducation for adults with MDD, as sole treatment or in conjunction with treatment as usual (TAU), compared to pharmacological treatment and/or other psychological treatment. Included effect measures are quality of life, depression severity, mortality (suicide), psychosocial functioning, relapse, and compliance.

Methods: The review was planned and described in a PROSPERO (CRD42017077110) registered protocol. The search strategy was executed by a search librarian and it was peer reviewed by another librarian. The search included electronic searches in MEDLINE, Embase, the Cochrane Library, PsycINFO, PubMed, CINAHL, Epistemonikos and a hand search of 29 systematic reviews. The search yielded a total of 4219 records, which were screened independently by two reviewers. We assessed eligible studies for risk of bias using

the Cochrane risk of bias tool for RCTs. We conducted meta-analyses when studies were sufficiently similar in terms of design, population, intervention, and outcomes. Lastly, we evaluated the certainty of the body of evidence using the GRADE approach.

Results: Nine randomized controlled studies (RCTs) with a total of 1249 patients met the inclusion criteria. The meta-analytic results showed that group psychoeducation in conjunction with TAU compared to TAU lead to a reduction in depression at 4-6 weeks, SMD= -0.32 (95% CI: -0.59 to -0.04), and 6 months, SMD= -0.21 (95% CI: -0.38 to -0.04). The effect of psychoeducation in conjunction with TAU was not significant at 12 months follow up, SMD= 0.22 (95% CI:-0.02 to 0.45). Family psychoeducation (groups including patient and caregiver) in conjunction with TAU showed a greater effect on depression than patient group psychoeducation. This was particularly prominent at 3 months follow-up, SMD= -1.21 (95% CI: -1.64 to -0.78). Family psychoeducation in conjunction with TAU also showed greater effect than TAU alone on psychosocial functioning at 3 months follow-up, SMD= 0.98 (95% CI: 0.56 to 1.40). The confidence in the certainty of the evidence varies from high to low. Results for psychosocial functioning was downgraded due to small sample size.

Conclusions: While the current body of research on group psychoeducation shows promise for its effects on depression and psychosocial functioning, further evidence on the short- and long-term effects is needed. Family group psychoeducation seems to give better results than patient group psychoeducation. Robust studies to build a solid evidence on the effect of psychoeducation and knowledge on the effects for different patient groups in various socioeconomic- and cultural settings are necessary, prior to a generalised recommendation on this intervention for patients with major depressive disorder.

List of tables

List of figures

List of abbreviations

BDI	Beck Depression Inventory
CI	Confidence interval
Cochrane CCDAN	Cochrane Common Mental Disorders (research collaboration)
CWDC	Coping with Depression Course
DALYS	Disability-Adjusted Life Year
EQ-5D	Health status measured in five dimensions
FHI	Folkehelseinstituttet (Norwegian Institute of Public Health)
FGPE	Family Group psychoeducation
FS-36	The Short Form (36) Health Survey
HAM-D/HRSD	Hamilton Rating Scale for Depression
HS	Helene Sandberg
ICD-10	International Statistical Classification of Diseases and Related Health Problems
IRBAS	Illness-Related Behaviours and Attitudes Scale
GAF	Global assessment of functioning
GPE	Group Psychoeducation
MAOIs	Monoamine Oxidase Inhibitors
MD	Major Depression
MDD	Major Depressive Disorder
MFPE	Multifamily Psychoeducation
MeSH	Medical Subject Headings
non-RCTs	Non-Randomized Controlled Trials
P	Probability value (quantifying statistical significance)
PE	Psychoeducation
PICO	Patient – Intervention – Comparator – Outcomes
PGPE	Patient group psychoeducation
RB	Rigmor C Berg
RCTs	Randomized Controlled Trials
SD	Standard Deviation
SMD	Standardized Mean Difference

SNRIs	Serotonin and Norepinephrine Reuptake Inhibitors
SSRIs	Selective Serotonin Reuptake Inhibitor
TAU	Treatment as usual
TCAs	Tricyclic Antidepressants
WHO	World Health Organization
ÅR	Åshild Roaldset

1 Introduction

Although many patients with mental health disorders benefit from psychiatric care, a considerable proportion of such patients have limited access, do not want to take drugs, respond poorly or experience adverse effects (Vaaler & Fasmer, 2013). When patients have the capacity to provide informed consent and state that they do not want to be medicated, they should not be forced to, as long as there are alternative, drug-free treatments and care available (Brev fra Helse- og omsorgsdepartementet til de regionale helseforetakene, 2015). Responding to user organisations' calls for the introduction of drug-free treatment alternatives in mental health care, in November 2015, the Norwegian Ministry of Health instructed all the regional health authorities to provide medication-free treatment options (Brev fra Helse- og omsorgsdepartementet til de regionale helseforetakene, 2015). However, according to Norwegian psychiatric user organizations (regjerningen.no, 2015), the supply of non-medical treatment is inadequate. Further, drug-free alternatives challenge the conventional view of what is ethical psychological treatment (Njaa, 2018), and some argue that drug-free programs are "an uninformed measure" that lack evidence (Røssberg, 2017).

Psychoeducation (PE) is a drug-free, psychological treatment that has proven to be beneficial for patients suffering from psychosis and bipolar affective disorder, but there is currently a knowledge gap regarding psychoeducation for major depressive disorder (McFarlane, 2016).

The purpose of this Master thesis in Public Health is to identify and summarize research on the effects of group PE compared with pharmacological and/or other psychological treatment for patients suffering from major depressive disorder. Given that psychoeducation typically is given in conjunction with usual psychological care, it is relevant to examine studies of both PE as a drug-free stand-alone treatment and of PE combined with usual psychological

treatment. The aim of this review is to improve the knowledge base on the effects of PE as a drug-free adjunct treatment and PE as a drug-free sole alternative for patients with major depressive disorder.

1.1 Background

Depression is one of the most common mental disorders with more than 300 million people of all ages affected globally (WHO, 2018). According to the Norwegian Institute of Public Health, about 12-15% of the population at any time are affected by depression (fhi.no, 2015). Depression may become a serious health condition when it is long lasting, or when the intensity is moderate or severe. According to WHO, depression is the leading cause of disease burden globally, and it is on the rise (WHO, 2018). Major depressive disorder (MDD) and dysthymia accounted in 2010 for 2.5% of global Disability Adjusted Life Years (DALYS), a measure of reduced health year, and it is also associated with lower work productivity, suicide and ischemic heart disease (Ferrari, 2013). According to WHO, persons with MDD and schizophrenia have a 40-60% greater chance of dying prematurely compared to the general population. Depression can lead to suicide. Close to 800,000 people die of suicide every year due to all causes (WHO, 2018). Reduction in healthy life years and the associated health issues due to depression has not only an impact on the affected persons and their families but also on the economy worldwide (WHO, 2016).

Disease specific pharmacological treatment for depression is antidepressant medication. According to the National Institute of Mental Health (The National Institute of Mental Health, 2016), the most common antidepressant medications are: Selective Serotonin Reuptake Inhibitor (SSRIs), serotonin and norepinephrine reuptake inhibitors (SNRIs), Bupropion, Tricyclic Antidepressants (TCAs), Tetracyclics Antidepressant, and monoamine oxidase inhibitors (MAOIs).

There are many non-medical interventions for depression disorder. Cochrane Common Mental Disorder lists 87 different psychological therapies (Cochrane CCDAN, n.d.).

Despite the range of treatment options for depression, globally, fewer than half of those affected receive treatment for their depression, and in some countries less than 10% receive treatment (WHO, 2018). Reasons include lack of resources, lack of trained health care providers, inaccurate diagnostic assessment, and social stigma associated with mental disorders (WHO, 2018).

To fulfill the Norwegian government's goals of providing effective and safe medication-free treatment for people suffering from depression, the treatment options on offer to the patients need to be evidence based. If proven to be effective, group psychoeducation (GPE) is a low-cost treatment that can be made widely available. It also has the potential to reduce the social stigma of psychiatric illnesses, because the format of the treatment is a course rather than a therapy, intending to reach people who otherwise may not seek formal treatment (Dowrick, 2001).

The Norwegian Institute of Public Health recently conducted a systematic review aiming to evaluate the effect of different psychosocial therapies without use of antipsychotics for patients with active psychosis. The literature search yielded no relevant studies where psychosocial treatment was given in conjunction with antipsychotic medication (Holte, 2017). Thus, we foresee a two-step approach in evaluation of the effectiveness of GPE. If GPE proves to be effective in conjunction with traditional psychotherapy and/or antidepressant medications, the logical next step would be to conduct research on the effectiveness of drug-free GPE as a sole treatment, for patients who want drug-free alternatives.

1.2 Patient perspective

In 2011, patients and their dependants established the initiative «Fellesaksjonen for medisinfrie behandlingsforløp» in Norway. This stakeholder initiative's sole purpose was to advocate for drug-free alternatives in treatment of psychiatric illnesses. Pressure from «Fellesaksjonen for medisinfrie behandlingsforløp» resulted in the 2015 instruction from the Norwegian ministry of health and care services to all regional health care authorities, demanding provision of drug-free alternatives in psychiatric care, in all regions (Brev fra Helse- og omsorgsdepartementet til de regionale helseforetakene, 2015).

The patients and other stakeholders were, and are still, arguing that medication is often perceived as involuntary treatment (Njaa, 2018) with undesirable side effects. The degree of undesirable side effects can be substantial. In one study (Singh, Liliah & Montagene, 2016), 53.3% of patients on antidepressants reported personality change, and 63.6% of patients reported dependency of their medication, but 88.6% did not feel addicted. Undesirable side effects were reported to be the main cause of non-adherence to treatment.

Respect for the patient's integrity should be of paramount importance to anyone supplying healthcare, therefore research into drug-free alternatives can easily be justified.

1.3 Description of the condition

A depressive episode can be categorized as mild, moderate, or severe depending on the number and severity of symptoms (WHO, 2018). The 10th revision of the International Statistical Classification of Diseases and Related Health Problems (ICD-10) (WHO, 2016) defines different clinical diagnoses of depression. Major depression disorder (MDD) is an episodic disorder with chronic or long-term outcome and increased risk of death (WHO,

2018). During a severe depressive episode, it is unlikely that the person will manage to function well enough to maintain work and social activities (WHO, 2018).

1.4 Description of the intervention

Psychoeducation is defined as a didactic approach aiming to give the participants sound knowledge of the condition and learn how to accept it and cope with it successfully (Ekhtiari, 2017). The intervention can be used for many different conditions. It can include practical tasks, making participants practice skills such as self-assertiveness, communication and problem solving, empowering the participants throughout the program. A psychoeducation program avoids the pathogenetic doctor and patient relationship by considering the patients as participants in the program rather than patients in the psychoeducational setting (Motlova, 2017).

Psychoeducation can be delivered in different formats, to either individuals alone or in groups. Individual psychoeducation may be indicated when an anxious person feels threatened by group situations or wish to stay confidential about the illness (Psychoeducation: Definition, Goals and Methods, 2014). The group format can for some people feel less intimidating than one-to-one sessions and the sharing of experiences in groups will benefit the participants. Support from group members is key to reducing stress and stigma and to be motivated to cope with the disorder (Psychoeducation: Definition, Goals and Methods, 2014).

Lewinsohn, Hoberman, Teri and Hautzinger, (1985) describe group psychoeducation (GPE) as an approach developed from a model considering depression as a product of multiple risk factors acting to transform the emotions, actions and cognitive processes of individuals facing adverse conditions (as cited in Dowrick, 2001). Cognitive behavioral treatment (CBT) techniques and strategies proven to be effective for depression were elements they

incorporated in a structured psychoeducational format for groups (Dowrick, 2001). The most common format is the Coping with Depression Course (CWDC) which was developed by Lewinsohn and Clarke in 1984 (Efthimiou & Psoma, 2012). The aim of psychoeducation is both prevention and treatment, it can be used in combination with drugs and it has been used in both health care and community settings (Dowrick, 2001). Group psychoeducation is implemented in different countries, such as Germany, the Netherlands, and USA (Haringsma, Cuijpers & Spinhoven, 2006). Chile is another example and is one of the few middle-income countries that has implemented a national comprehensive multi-component treatment program for depression including group psychoeducation, and it has proven to have good results (Araya, Alvarado & Minoletti, 2011).

Group psychoeducation will often involve caregivers such as family and friends. People with depression are normally taken care of by a relative in their home. Here, we will use the term caregiver, relative and family interchangeably. Group psychoeducation that includes family member(s) (FPE) emerged from different sources in the late 1970s (McFarlane, 2016). The first model for single-family psychoeducation format was developed by Andersen et al., in 1986 (Fallon, 1984; Miklowitz & Goldstein, 1997) and the multi-family format (MFPE) was introduced in 2002 by McFarlane (MacFarlane, 2016). The format of family psychoeducation can be valuable for the whole family. Information and activities may help family members to better understand the person suffering from the illness, and it may enable them to give the required support, as well as helping the family to get along with one another (Psychoeducation: Definition, Goals and Methods, 2014).

While there are numerous variations of psychoeducational approaches, we have decided to focus on group psychoeducation, with the following characteristics (McFarlane, 2016):

- is provided to groups of patients or groups consisting of one or more patient with their caregivers (family members, friends or other)
- is provided by a health care professional
- includes exercises in practical skills relevant for the patient and family group such as coping strategies, behavior, communication, social interactions, and problem solving
- aims at giving knowledge on depression and treatment of depression
- is specific for patients with depression

Modules in psychoeducation are typically designed to cover these elements:

> "(1) Information transfer (symptomatology of the disturbance, causes, treatment concepts, etc.); (2) Emotional discharge (understanding to promote exchange of experiences with other concerned, contacts, etc.); (3) support of a medical or psychotherapeutic treatment, as cooperation is promoted between the mental health professional and patient (compliance, adherence); and (4) assistance to self-help (e.g. training, so crisis situations are promptly recognised and steps taken to help the patient)" (Kumar & Gupta, 2015).

Psychoeducation can be delivered to the participant in many ways and the mode of delivery may be of importance for the effect of the intervention. Psychoeducation can also be delivered in multifamily groups, in a consultation given by a healthcare professional on one-to-one basis or as a web-based course with no human interaction. We use the term group psychoeducation for an intervention involving both participant and caregiver. However, group psychoeducation can be divided into patient group psychoeducation and family group psychoeducation.

Group psychoeducation may be conducted by health care professionals other than psychologist or psychotherapists (Cuijpers, Munoz, Clarke & Lewinsohn, 2009). The simplicity of psychoeducation allows training which is neither long nor complex in the technique, thus the intervention is not dependent on a specialist (Colomn, 2011). Nurses or other health care professionals who are experts on the disorder, rather than a technique, can conduct the intervention (Colomn, 2011).

The most common format of psychoeducation, the CWDC consists of twelve two-hour sessions over 8 weeks, and a modified version with 8 sessions (Dowrick, 2001). We have included all durations (doses) of group psychoeducation in this review, as long as it is minimum one 60-minute session.

1.5 How the intervention might work

According to Frances Colom (2011);

> "Psychoeducation could be defined as a patient's empowering training targeted at promoting awareness and proactivity, providing tools to manage, cope and live with a chronic condition (i.e. adherence enhancement, early warning sign identification, lifestyle, crisis management, communication), and changing behaviors and attitudes related to the condition. Psychoeducation replaces guilt by responsibility, helplessness by proactive care and denial by awareness" (Colom, 2011).

The core of the psychoeducation intervention is the didactic element, which is meant to provide the participants (patients) with insight into their disease and learn how to accept it and cope with it successfully. The content of the psychoeducation course may vary, but it includes disease-specific knowledge and more general knowledge such as problem-solving skills,

communication skills and healthy living, amongst other. Motlova et al. (2017), points out four active ingredients in psychoeducation:

1. Taking the whole participant into account and building on the participants strengths and resilience
2. Giving the participant emotional support, guidance and sufficient knowledge to accept this reality and adapt to the illness. Including information on how to reduce stress in the household.
3. Psychoeducation must include behavioral interventions, promoting healthy activities such as healthy eating, sufficient sleep, exercise and support from friends.
4. Psychoeducation shall ensure that participant and caregivers have access to reliable sources of information to avoid miseducation that may occur form unreliable internet sources

Psychoeducation focuses on improving the functioning of the patients (Murray-Swank & Dixon, 2011). The therapeutic mechanism may be due to behavior change, leading to interruption of the vicious cycle normally seen in depression, and an improved interaction with the environment with a better function as a result of this (Cuijpers, Munoz, Clarke & Lewinsohn, 2009). Exploring the effect of family psychoeducation on MDD is timely and relevant. Recent research by The Brainstorm Consortium (The Brainstorm Consortium, Anttila et.al. 2018) revealed there is a high degree of heritability and a genetic correlation between common psychiatric diseases such as major depressive disorder, attention deficit hyperactivity disorder, bipolar disorder, and schizophrenia. This means that in a family where one family member is suffering from any of these diseases, there may be other family members at risk of at least one of the associated diseases. A family intervention may therefore be a sensible approach to MDD and perhaps the strict diagnostic criteria of MDD is of lesser

importance for all involved parties. Family psychoeducation programs are already considered an evidence-based intervention in the treatment of schizophrenia and bipolar disorder (Murray-Swank & Dixon, 2005).

In family psychoeducation, caregivers or family members are regarded as important resources for the patient's health recovery. According to Brandy, Kangas & McGill (2016), the aim of FPE is to enhance treatment outcomes by enabling those who are closest to the person, family or other caregiver, to assists in events which may exacerbate the illness. Another key element of family psychoeducation is what Murray-Swank and Dixon (2005) describe as "expressed emotions". Expressed emotions is referring to the level of criticism and emotional over-involvement among caregivers in the patient's environment (Murray-Swank & Dixon, 2005). MDD causes high levels of family burden and expressed emotions (Luciano et al., 2012). A knowledgeable caregiver may have improved coping skills and better ability to withstand the suffering of the depressed family member.

1.6 Why is it important to do this review?

Depression causes significant burden to individuals, families, and society, but access to psychological therapy for depression is limited, and there is a need to strengthen the evidence-base on effective psychological treatments. Psychoeducation is one easily accessible approach that has proven to be beneficial to patients suffering from psychosis and bipolar affective disorder, but there is currently a knowledge gap regarding psychoeducation and major depressive disorder (McFarlane, 2016). Studies have shown that family functioning is important in determining the course of MDD (Keitner et al., 1995). We believe this systematic review on PE involving family members in treatment for MDD can further illuminate the importance of the relative impact on the patient's recovery and ongoing

functioning. There is a clear need to reduce the gap between demand and access to treatment for depression. If group psychoeducation is proven to be effective compared to other available treatment options; it has the potential to play an important role in making treatment accessible to a large number of patients in need, worldwide.

1.7 Review question

Is group psychoeducation as sole therapy or as an adjunct therapy effective for adults with major depressive disorder compared to pharmacological treatment and/or other psychological treatments?

1.8 Objective

The aim is to systematically review the effectiveness of group psychoeducation for adults with major depressive disorder, as sole therapy or in conjunction with treatment as usual, compared to pharmacological treatment and/or other psychological treatment. Included effect measures are quality of life, depression severity, mortality (suicide), psychosocial functioning, relapse, and compliance.

2 Methods

We conducted a systematic review in accordance with the Cochrane Handbook for Systematic Reviews of Interventions, Version 5.1.0 (Higgins & Green, 2011). This chapter describes the methods and choices we made. Our PROSPERO protocol was registered in September 2017 (CRD42017077110), enclosed in appendix 1.

2.1 Search strategy

The search strategy was prepared by us in collaboration with a research librarian from the Norwegian Institute of Public Health, and peer reviewed by another research librarian from the same institution. The search was conducted by the librarian and individually adapted for the following databases:

- MEDLINE (OVID)
- PsycINFO (OVID)
- EMBASE (OVID)
- Cochrane Library (CDSR, HTA, CENTRAL, DARE)
- CINAHL (EBSCO)
- Epistemonikos

The literature search consisted of subject headings and text word-controlled vocabulary, e.g. MeSH in MEDLINE, covering depression and psychoeducation. The search was limited to year 2000 and newer, because of a consensus that the critical elements of family psychoeducation was developed in 1999 (U.S. Department of Health and Human Services, 2009). The search closed in September 2017. The full search strategy is enclosed in appendix 2. In addition to the systematic search in electronic databases, we hand searched the reference lists of systematic reviews and literature reviews to identify any relevant studies not indexed in the databases.

2.2 Inclusion and exclusion criteria

The inclusion criteria for effectiveness studies of group psychoeducation for major depressive disorder compared with pharmacological and/or other psychological treatment are described

by the question's different elements below: Study design - Patient – Intervention – Comparator – Outcomes (SPICO).

2.2.1 Study design

Studies we pre-specified to be eligible for inclusion were randomized controlled trials (RCTs) and non-randomized controlled trials (non-RCTs), controlled before-after studies (CBAs), interrupted time series (ITS) plus prospective and retrospective cohort studies with a control group. We also specified that in the event of identifying several high-quality RCTs and non-RCTs, we would consider not including other study designs. Cluster RCTs analysed on an individual level should be adjusted for intra cluster correlation (ICC).

2.2.2 Population

Eligible participants (patients) were adults with major depressive disorder. Table 1 gives an overview of the included ICD-10 codes included in this systematic review. This list of specific ICD-10 codes to be included in the review were discussed with a psychiatrist, to cover the right diagnostic codes for MDD.

Studies with more than 50% of patients with medical comorbidities were excluded (e.g. anxiety, diabetes, cancer). Other studies excluded were studies with more than 50% of patients with Perinatal Depression, Bipolar Affective Disorder and mental impairment, including dementia. We excluded studies with patients in remission and who thus no longer met the depression criteria. We discussed among us and with a psychiatrist whether the exclusion of patients with anxiety was reasonable, as depression and anxiety often go hand in hand, and concluded that for this review it was prudent to focus on patients with depression only.

Table 1. Included ICD-10 codes

ICD10 code	Clinical description (copied from WHO, 2016, ICD-10 list)

F32.2 Severe depressive episode without psychotic symptoms	An episode of depression in which several of the depressive episode (F32) symptoms are marked and distressing, typically loss of self-esteem and ideas of worthlessness or guilt. Suicidal thoughts and acts are common, and a number of "somatic" symptoms are usually present. Single episode without psychotic symptoms (agitated depression, major depression, vital depression).
F32.3 Severe depressive episode with psychotic symptoms	An episode of depression as described in F32.2, but with the presence of hallucinations, delusions, psychomotor retardation, or stupor so severe that ordinary social activities are impossible; there may be danger to life from suicide, dehydration, or starvation. The hallucinations and delusions may or may not be mood-congruent. Single episodes of: • major depression with psychotic symptoms • psychogenic depressive psychosis • psychotic depression • reactive depressive psychosis
F33.2 Recurrent depressive disorder, current episode severe without psychotic symptoms	A disorder characterized by repeated episodes of depression, the current episode being severe without psychotic symptoms, as in F32.2, and without any history of mania. Endogenous depression without psychotic symptoms Major depression, recurrent without psychotic symptoms Manic-depressive psychosis, depressed type without psychotic symptoms Vital depression, recurrent without psychotic symptoms
F33.3 Recurrent depressive disorder, current episode severe with psychotic symptoms	A disorder characterized by repeated episodes of depression, the current episode being severe with psychotic symptoms, as in F32.3, and with no previous episodes of mania. Endogenous depression with psychotic symptoms Manic-depressive psychosis, depressed type with psychotic symptoms Recurrent severe episodes of: • major depression with psychotic symptoms • psychogenic depressive psychosis • psychotic depression • reactive depressive psychosis
F34.0 Cyclothymia	A persistent instability of mood involving numerous periods of depression and mild elation, none of which is sufficiently severe or prolonged to justify a diagnosis of bipolar affective disorder (F31.-) or recurrent depressive disorder (F33.-). This disorder is frequently found in the relatives of patients with bipolar affective disorder. Some patients with cyclothymia eventually develop bipolar affective disorder. Affective personality disorder Cycloid personality Cyclothymic personality

F34.1 Dysthymia	A chronic depression of mood, lasting at least several years, which is not sufficiently severe, or in which individual episodes are not sufficiently prolonged, to justify a diagnosis of severe, moderate, or mild recurrent depressive disorder (F33.-). Depressive (neurosis, personality disorder) Neurotic depression Persistent anxiety depression Excl.: anxiety depression (mild or not persistent) (F41.2)

2.2.3 Intervention

The intervention was group psychoeducation that had the following characteristics (McFarlane, 2016):

- Is provided to groups of patients or groups consisting of one or more patient with the caregivers (family members, friends or other)
- Is provided by a health care professional
- Includes exercises in practical skills relevant for the patient and family group such as coping strategies, behavior, communication, social interactions, and problem solving
- Aims at giving knowledge on depression and treatment of depression
- Is specific for patients with depression

We included all durations (doses) of group psychoeducation as long as the intervention consisted of minimum one 60-minute session. Group psychoeducation could be given as sole treatment or in conjunction with treatment as usual (TAU).

2.2.4 Comparison

We specified the following comparison conditions:

- Antidepressant medications. We limited medication to be common antidepressants only: SSRIs, SNRIs, Bupropion, Tetracyclic Antidepressant, and MAOIs

- Other psychological intervention (psychological interventions in the Cochrane CCDAN list of 87 interventions were included. See Cochrane CCDAN, n.d.)

2.2.5 Outcomes

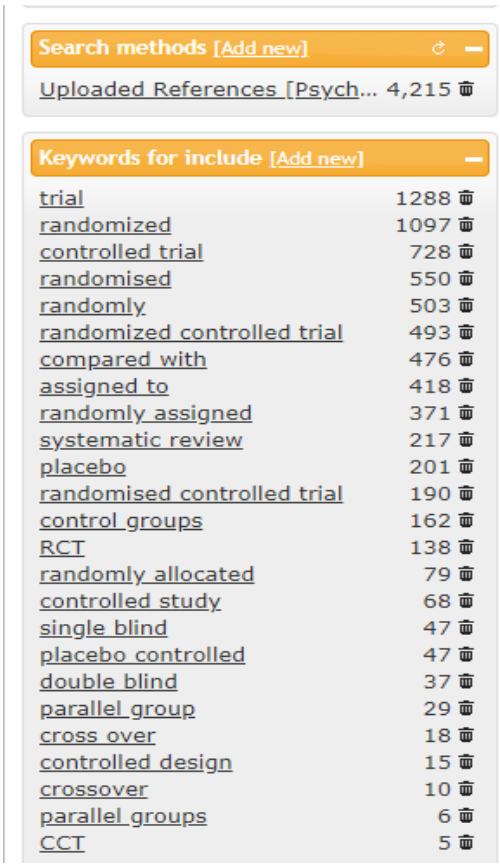
We specified the following primary outcomes: depression severity, quality of life, and mortality (suicide). The secondary outcomes were level of psychosocial functioning, relapse, and treatment adherence.

Measurement tools for the outcomes are described in chapter 3.2.6.

2.3 Selection of literature

Two reviewers (ÅR and HS) screened independently of each other all the abstracts from the literature searches by the use of Rayyan QCRI – a web and mobile app for systematic reviews (Ouzzani, M., Hammady, H., Fedorowicz, Z. & Elmagarmid, A., 2016) (see Figure 1).

Figure 1. Illustration of Rayyan screening tool, used for the 4215 uploaded references



We promoted all relevant abstracts to full text examination, and again independently assessed the studies' relevance relative to the inclusion criteria. The search retrieved both primary studies and systematic reviews. We hand searched the reference lists of the systematic reviews for any relevant primary studies which were not identified through the database searches. Any disagreement about relevance of primary studies was solved by discussion with a third researcher (RB). We present excluded studies read in full text, with the reason for exclusion, in appendix 3. Ongoing studies with a protocol published are described in appendix 4.

2.4 Assessment of methodological quality (Risk of bias assessment)

The two reviewers, HS and ÅR, first independently and then together performed an assessment of risk of bias of each included study. Because we only included RCTs we

assessed the risk of bias for all studies according to the criteria for RCTs in the Cochrane Handbook for Systematic Reviews of Interventions (Higgins & Green, 2011). We evaluated the processes for sequence generation; allocation concealment; blinding of participants and personnel; blinding of outcome assessment; incomplete outcome data; selective outcome reporting; and other sources of bias. The risk of bias for all processes are reported as 'Low Risk', 'Unclear Risk', or 'High Risk'. When there is no cause for concern the procedure is considered to have low risk of bias. When there may be a risk of bias, but there is either insufficient information to assess whether an important risk of bias exists or there is insufficient rationale or evidence that an identified problem will introduce bias, we assigned unclear risk of bias. Procedures with cause for concern is assigned high risk of bias. Any disagreement regarding risk of bias was solved by discussion between the two reviewers and with some guidance from the supervisor. The full risk of bias assessment of included studies is enclosed in appendix 5. If study designs other than RCT were to be included, we would have use the procedures described in the protocol in appendix 1.

2.5 Extraction of data

We created a standard data extraction form. Both reviewers extracted data from all included studies by using the data extraction form. We then compared the extractions to ensure all correct and necessary data were extracted. The information we extracted was: title, authors and other publication details, study design and aim, setting (place and time of recruitment/data collection), sample characteristics (age, gender, etc.), intervention and control characteristics (duration/dose, provider, content, etc.), methods of outcome measurement (instruments/ tools), results/outcomes. The mapping of the outcomes per study is enclosed in appendix 6.

2.6 Data analysis

In statistical analyses we are seeking differences that are not random. A significant result means that we have chosen a level (95%) for how sure we want to be about the results not being random. The significance can be measured by using p-values or confidence intervals (CI).

Our research question is about effect of an intervention. In systematic reviews it is the effect at group level which will be examined in the analysis, meaning it is the mean effect in the group we are considering. The effect value can be calculated in different ways.

Risk ratio (RR) is generally used for dichotomous outcomes. The effect measure is then the ratio between the probability that an event occurs in both groups. If the probability is equal in the two groups, the risk ratio will be 1. A risk ratio of 2 means that the probability that an event occurs is twice as big in the intervention group compared to the control group. For continuous data, the effect measure is mean difference (MD) or standardized mean difference (SMD). We have used 95% confidence intervals (95% CI) to calculate the effect sizes by using the RevMan tool (Review Manager, 2014). MD for continuous outcomes applies when the same measurement scales are used. The effect measure is then the absolute difference between the mean value in the two groups, and it estimates the amount by which the experimental intervention changes the outcome on average compared with the control (Higgins & Green, 2011).

If studies were sufficiently similar with respect to population, intervention, comparison and outcome, we decided we would conduct meta-analyses, using the RevMan tool, and generate forest plots to display the results.

If we pooled studies with continuous outcomes and the studies did not use the same measurement instrument, we would recalculate all results into standardized measures and get the SMD.

There will almost always be some heterogeneity between estimates that are pooled. Heterogeneity can be clinical (differences between the participants, the interventions or outcomes), methodological (differences in study design or in risk of bias) or statistical (Higgins & Green, 2011). We define heterogeneity when there is great variation in results, non-overlapping CIs, P<0,001 and I>50% (Higgins and Green, 2011). We examined causes for and attempted to explain heterogeneity.

Meta-analyses can be conducted with either a random effects model or a fixed effects model. We selected to use a random effects model to combine the effect estimates, rather than a fixed effects model, because we judge there might be systematic differences between the studies while a fixed-effect model would consider each individual study as part of one big study (Higgins & Green, 2011).

For the effect comparisons where there was only one study, or studies could not be pooled, we presented the results narratively in text and tables.

When there are many studies included in a meta-analysis and high heterogeneity, it is possible to conduct sub-group analyses, meaning that studies are grouped to check if there are significant differences between the groups. We pre-specified the following sub-groups or subsets that we would examine:

- Effect of group composition: groups consisting of patients only or patients together with their family members or other caregiver

- Dose effect of the psychoeducation therapy: 12 sessions (of 1-2 hours duration every week) or less were considered low dose, 13-52 sessions (of 1-2 hours duration every week) were considered moderate dose, and 52 sessions and above (1-2 hours duration every week) were considered high dose (Xia, Merinder & Belgamwar, 2011).

In our review, it was possible to elaborate the effect of group psychoeducation with a sub-group analysis investigating family psychoeducation, grouping studies with participants together with their caregivers (family psychoeducation) compared to studies with groups for patients only.

2.7 Assessment of the certainty of the evidence (GRADE)

Grading of Recommendations Assessment, Development and Evaluation (GRADE) methodology was used to consider the certainty of the evidence (The Grade Working Group, 2013). The two reviewers (HS and ÅR) conducted the grading together. We used the standard definitions to assess the certainty of the evidence, with the certainty of the evidence classified as either high, moderate, low or very low. The different levels of certainty of evidence reflects the extent to which we are confident that an estimate of the effect is 'correct'. The certainty of the evidence is rated for each outcome across studies (i.e. for a body of evidence).

Table 2: Grades of certainty of evidence

Grade	Symbol	Definition
High	⊕⊕⊕⊕	We are very confident that the true effect lies close to that of the estimate of the effect
Moderate	⊕⊕⊕◯	We are moderately confident in the effect estimate: The true effect is likely to be close to the estimate of the effect, but there is a possibility that it is substantially different
Low	⊕⊕◯◯	Our confidence in the effect estimate is limited: The true effect may be substantially different from the estimate of the effect
Very low	⊕◯◯◯	We have very little confidence in the effect estimate: The true effect is likely to be substantially different from the estimate of effect

The study design is the starting point for the grade assessment. Randomized controlled trials are considered to start as high and observational studies are considered to start as low. Factors that can lower the certainty of evidence according to The Grade Working Group (2013):

- Methodological study limitations (risk of bias)
- Inconsistency
- Indirectness
- Imprecision
- Publication bias

Factors that can raise the certainty of evidence:

- Large magnitude of effect
- Dose response relationship
- Opposing bias & confounders

The decision to up rate certainty of evidence is only made for observational studies and only when serious limitations in any of the 5 areas reducing the certainty of evidence is absent.

2.8 Changes from the original protocol

To improve precision, we made a small change in the research question. The original question as registered in PROSPERO was: "*Is group psychoeducation effective in improving quality of life in adults with major depressive disorder compared to pharmacological treatment and/or other psychological treatment?"* We amended the research question to: *"Is group psychoeducation, as sole therapy or as adjunct therapy, effective for adults with major depressive disorder compared to pharmacological treatment and/or other psychological treatment?"*

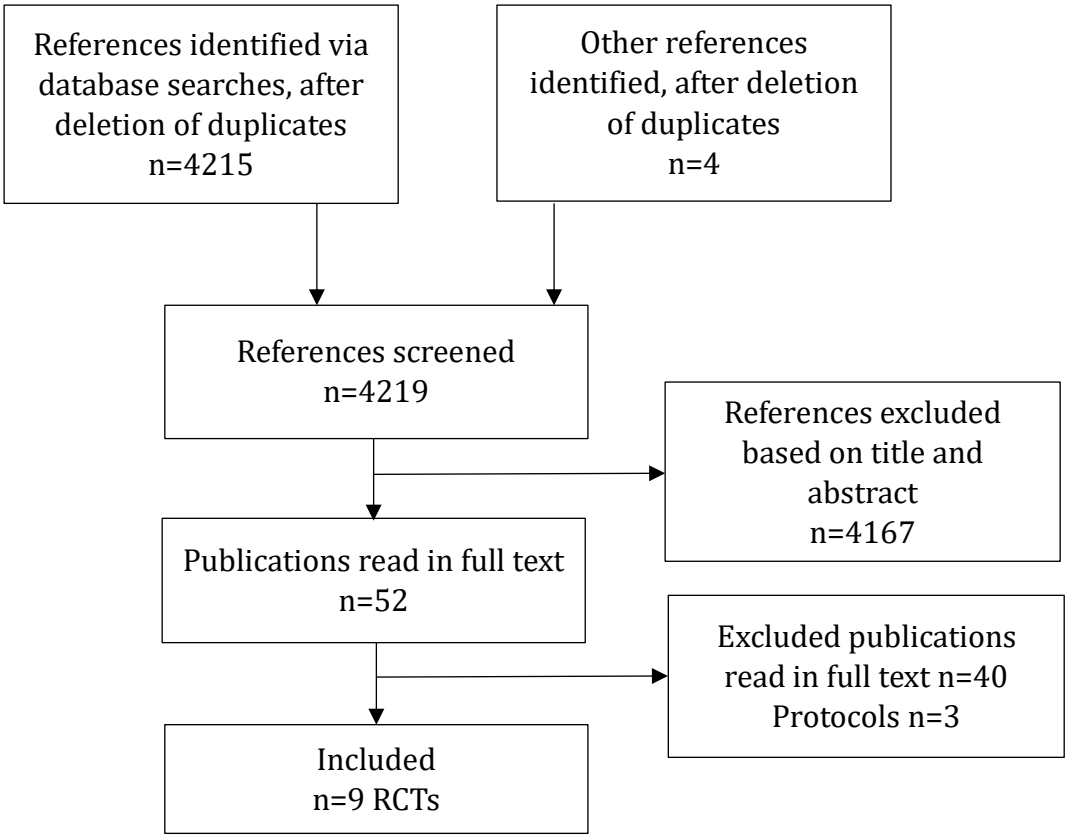
When planning the review, as shown in our PROSPERO registered protocol, we focused on comparing the effect of group psychoeducation with pharmacological interventions and/or other psychological interventions. It rapidly became evident that most patients given psychoeducation continue to receive treatment as usual (TAU), such as psychotherapy and pharmacological treatment. While none of the studies excluded TAU not all patients used all available treatment alternatives. Thus, it was obvious that we would include studies regardless of whether psychoeducation was given as sole treatment or in conjunction with other standard treatment. TAU varies across settings and here it was defined as access to health care professionals and access to pharmacological medication (see table 7 for details on TAU).

3 Results

3.1 Results of the literature search

The search in electronic databases gave 4215 references in total after deletion of duplicates (figure 2). We hand searched 29 systematic reviews for relevant references, which yielded four additional articles, giving a total of 4219 references. We excluded 4167 records based on title and abstract and read 52 full text articles. One author (Conradi, Jonge & Ormel, 2008) was contacted by e-mail to clarify whether the study in question met our inclusion criteria, it did not, and one author (Lara, Navarro & Mondragon, 2003) was contacted in an attempt to retrieve numeric results of the study. Both of these studies are listed among the excluded articles read in full text (appendix 3). We identified three study protocols (see appendix 4). At the end of our selection process, nine RCTs matched our inclusion criteria and other study designs were thus not considered, as per our protocol.

Figure 2. Flowchart illustrating the literature selection process



3.2 Description of included studies and their context

Nine unique studies presented in 10 articles are included in this systematic review. The earliest study is from year 2000 and the most recent study is from 2017. Studies include psychoeducation delivered in three different ways; in groups of patients, as a brochure handed to the patient, and in groups including participants and their caregivers (family psychoeducation). Günadyın & Barlas (2017) had one study arm delivering psychoeducation as a brochure. This intervention did not meet our inclusion criteria, and the participants receiving this intervention are not included in any of the analysis. Dowrick et al (2000) had one study arm receiving problem solving, participants in this arm was also excluded from the analyses.

The included studies were conducted on four different continents (Africa, America, Asia, and Europe) and in 11 different countries (Denmark, Finland, India, Iran, Norway, Republic of Ireland, Spain, South Africa, Turkey, United Kingdom, and USA). One study (Dowrick et al., 2000), was conducted in multiple countries (Finland, Norway, Republic of Ireland, Spain and United Kingdom). The total number of individual research participants in the included studies is 1231. The results of one study (Dalgard, 2006) are provided in two articles (Dalgard, 2004 and Dalgard, 2006), one of which is only published in Norwegian (Dalgard, 2004). This study is referred to as Dalgard 2006 in tables throughout this review. The cultural, sociological and economic contexts vary greatly among the studies and the baseline treatment for depression (treatment as usual) must be interpreted for each study. Chetty and Hoque (2013) described their study as "quasi-experimental" and Sharif, Ashkani and Zoladl (2012) called their study "interventional case control study". However, based on the descriptions, it is clear that both of these studies are RCTs with satisfactory risk of bias. Table 3 gives a brief overview of the included studies.

Table 3. Brief overview of the included studies (N=9)

Study & context	Population	Intervention: GPE	Comparator	Outcome
Aagaard 2017 Denmark, multicentre	N=80 ICD-10 recurrent depression, all severe	8 weekly sessions of 120 minutes Group consisting of patients + 1 session with caregiver TAU	TAU	Decline in BDI Drop-out/non-compliance, psychotropic drugs and social measurement Admission to psychiatric hospital
Casañas 2012 Spain, Barcelona	N= 231 MDD according to ICD-10 Depressive Disorder BDI 10-29 (mild or moderate)	12 weekly sessions, 90 minutes Group consisting of patients TAU	TAU	Remission (BDI <11) Quality of life (EQ-5D)
Chetty 2013 South Africa, southern KwaZulu-Natal	N=30 Depressed South African Indian women according to DSM 4 BDI 10-28 (mild to moderate)	15 weekly sessions, 60-120 minutes Group consisting of patients TAU	TAU	Depressive symptoms (BDI)
Cohen 2010 USA, Long Island	N=35 MDD and Dysthymia, BDI-2 >21	5 weekly sessions of 120 minutes, Patient and caregiver in group TAU	TAU	Depression symptoms (BDI-2, HAM-D) Spouse impact: (FSDS) Change of behavioural and attitude (IRBAS) Overall relationship satisfaction (DAS)
Dalgard 2006 Norway, Oslo	N=155 Unipolar depression according to DSM-4. BDI mean= 21.8/22.9 (moderate depression)	8 weekly sessions, 150 minutes, plus booster sessions 1, 2 and 4 months after the course Group consisting of patients TAU	TAU	Depressive symptoms (BDI)
Dowrick 2000 Finland, Norway, Republic of Ireland, Spain, United Kingdom	N= 425 Depressive episode according to ICD-10	12 sessions, 2 hours over 8 weeks, Group consisting of patients TAU	**Comparator I** TAU **Comparator II** Problem solving TAU	Acceptability of two interventions (withdrawals) Depressive symptoms (BDI) Quality of life (SF-36) Caseness (FS-36)

Study & context	Population	Intervention	Comparator	Outcome
Günadyın 2017 Turkey	N=135 Unipolar depression, DSM-IV BDI score 17-30	5 weekly sessions of 45-60 minutes Group consisting of patients TAU	**Comparator I** TAU **Comparator II** Psychoeducation delivered as a brochure TAU	Depressive symptoms (BDI)
Kumar 2015 India	N=80 Unipolar depression, MDD to dysthymia, according to ICD-10	Sessions on week 0, week 2, week 4 and week 8 Group consisting of caregiver and patients TAU	TAU	Depressive symptoms (HAM-D) Quality of life (PGWBI) Psychosocial functioning (GAF)
Sharif 2012 Iran	N=60 Major depressive disorder according to ICD-10 criteria	6 weekly sessions of 90 minutes duration. Group consisting of 6 patients TAU	TAU	Quality of life (FS-36 completed by researcher)

Legend: N= number of participants, TAU= treatment as usual, ICD-10= 10th International Statistical Classification of Diseases and Related Health Problems.

3.2.1 Study setting

The studies were conducted in a variety of settings (table 4). All of the studies were conducted in urban areas except for the multicentre study by Dowrick et al. (2000), that reported urban and rural settings, without any further information. Patients were recruited as primary patients (Aagaard, Foldanger, Makki, Hansen & Muller-Nielsen, 2017, Casañas, Catalan, del Val, Real, Valero & Casas, 2012, Chetty & Hoque, 2013) as referral patients (Günadyın & Barlas, 2017; Kumar & Gupta, 2015, Sharif & Ashkani & Zoladl, 2012), by using mass media (Dalgard, 2006 and Cohen, O'Leary & Foran, 2010) or by a process described as a "two stage community survey" (Dowrick et al., 2000).

Two studies were pre-registered in Clinical Trials.gov (Casañas et al., 2012 and Dalgard, 2006). However, one of these (Dalgard, 2006), was not pre-registered, but registered just before publication of the article providing the results. Informed consent was reported to be obtained in six studies (Aagaard et al., 2017; Casañas et al., 2012; Chetty & Hoque, 2013;

Günadyın & Barlas, 2017; Kumar & Gupta, 2015; Sharif & Ashkani & Zoladl, 2012) and ethical approval by a board is reported in four studies (Aagaard et al., 2017, Casañas et al., 2012, Chetty & Hoque, 2013, Sharif & Ashkani & Zoladl, 2012). In the studies by Cohen, O'Leary & Foran, 2010, Dalgard, 2006, and Dowrick et al., 2000, there was no information about informed consent.

Table 4. Overview of the setting, recruitment and ethical approval in the included studies

Study Country, City	**Setting**	**Ethical approval/registration**
Aagaard 2017 Denmark	Multicentre study conducted in 4 different Community Mental Health-Centres Urban	Written informed consent obtained from patient Approval from the Danish Data Protection Agency and the Scientific Ethic Committee.
Casañas 2012 Spain, Barcelona	Participants recruited at Primary Care Centres Urban	Informed consent was obtained from all participants Ethical approval by the Gol Guriana Foundation. Clinical Trials.gov NCT00841737
Chetty 2013 South Africa, southern KwaZulu-Natal	Participants recruited at Urban-community-psychiatric-clinic Urban	Informed consent was obtained from all participants Ethical permission was obtained from the Research, Publication and Ethics Committee of the Durban University of Technology
Cohen 2010 USA, Long Island	Participants recruited using TV, radio announcements, flyers and pamphlets in local medical clinics Urban	No information available
Dalgard 2006 Norway, Oslo	Participants recruited using newspaper advertisement Urban	No information available Clinical Trials.gov NCT00319540
Dowrick 2000 Finland, Norway, Republic of Ireland, Spain, United Kingdom	Participants recruited "…by a two stage community survey…" Urban and rural	No information available
Günadyın 2017 Turkey, Istanbul	"The research universe was composed of entire patients applied to psychiatric polyclinic of a state hospital in a month and diagnosed with depression…" Urban	Informed consent was obtained from all participants Ethical approval was obtained according to The Code of Ethics of the World Medical Association
Kumar 2015 India, New Delhi	Recruitment: "The study was conducted at the Department of Psychiatry of	Written informed consent was obtained from all participants

	Vardhman Mahavir Medical College & Safdarjung Hospital…" Urban	
Sharif 2012 Iran, Shiraz	Recruitment: "They were admitted in the psychiatric units of hospitals affiliated to Shiraz University of Medical Sciences at the time of study" Urban	Written informed consent was obtained from all participants Approval by the Ethics Committee of Shiraz University of Medical Sciences was obtained

3.2.2. Study population

Worldwide, there are more women than men who suffer from depression (WHO, 2018). This is seen in the samples in this review. The study by Kumar & Gupta (2015) is the only study where the majority of the patients are male. In this study the majority of caregiver are also male. Two studies (Chetty & Hoque, 2013 and Cohen, O'Leary & Foran, 2010) included female participants only (table 5).

With regard to age, the mean ages across the studies was 33-54. One study (Kumar & Gupta, 2015) included participants from 15 years to 59 years. This study was conducted in New Delhi, India. The average age in this particular study was 36.17 years (SD ±11). While we specified that we would only include populations 18 years or older, the number of participants under 18 in this study is negligible, thus we decided to include the study. We also believe that the age at which the society defines the commencement of adulthood varies between cultures and that Indian adolescent are considered adults at an earlier age than in many other societies. This may be expressed in the high rate of child marriages in India, where 27% of 20 to 24-year olds was married before the age of 18, in the years between 2010 and 2017 (UNICEF, 2018). The study was included in this systematic review because it seemingly had few, if any, underage participants and if any underage participants was included, they had probably already started their adult life, in many ways.

All participants suffered from MDD. The severity of depression of the participants (mild, moderate, severe) varied among the studies, see table 5 for details.

Table 5. Overview of the population in the included studies

Study	No of participants % female Comment	Age	Diagnostic criteria
Aagaard 2017	N=80 71% females	Mean age=48	ICD-10 recurrent depression
Casañas 2012	N=231 89% females	Age> 20 Mean age=53.8	Major Depressive Disorder (MDD) according to ICD-10 Depressive Disorder BDI 10-29 (mild or moderate).
Chetty 2013	N=30 100% females Indian South African women	Age range 31-60 years Mean age=45.2	Depressed according to DSM 4 BDI 10-28 (mild to moderate)
Cohen 2010	N=35 100% females (100% male caregiver)	Mean age=43.74	MDD and Dysthymia BDI-2 >21
Dalgard 2006	N=155 76% females	Mean age=50.3	Unipolar depression according to DSM-4 BDI mean= 21.8/22.9 (moderate depression) Depressive symptoms
Dowrick 2000	N=425 65% females	Age 18-65 years	Depressive episode according to ICD-10
Günadyın 2017	N=153 92% female	Age range 18-65 years	Unipolar depression, DSM-IV BDI score 17-30
Kumar 2015	N=80 38.8% female (57% male caregiver)	Age range 15-59 years Mean age= 33 years	Unipolar depression, MDD to dysthymia, according to ICD-10 criteria
Sharif 2012	N=60 54.9% female	Age >18 years	MDD according to ICD-10 criteria

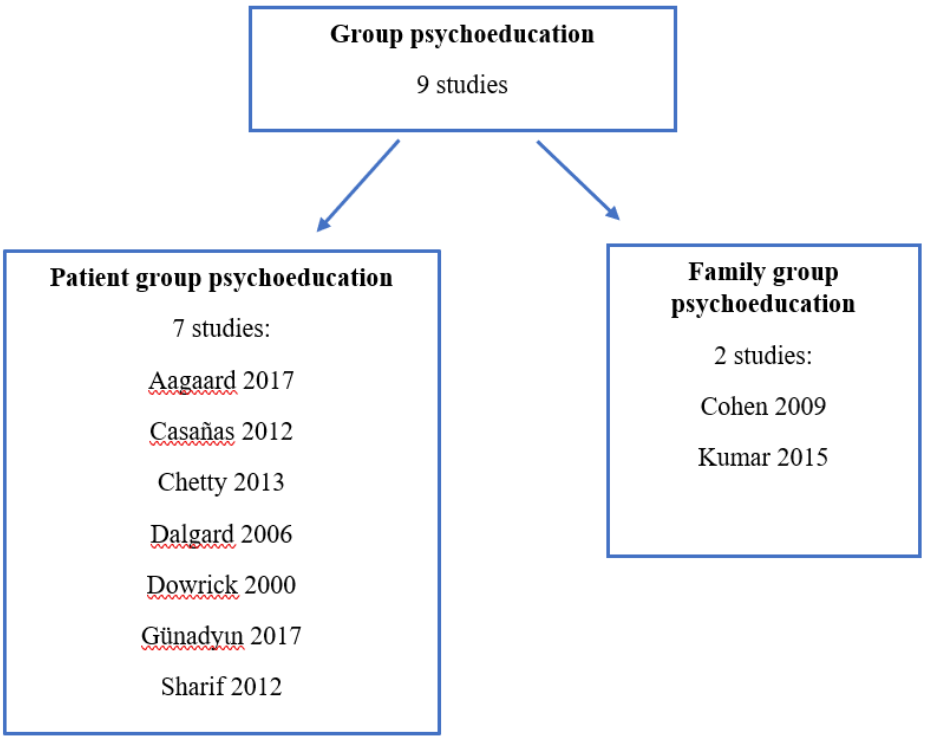
3.2.3 Intervention

The psychoeducation intervention in the included studies all have a didactic element, providing knowledge on depression, medication, and various other aspects. Several of the didactic programmes are based on a pre-defined syllabus such as “Group Psychoeducation Programme” (Aagaard et al., 2017), “Women’s Workbook and Facilitator`s Manual” (Chetty & Hoque, 2013), “Coping with depression course” (Dowrick et al., 2000) and “Continuity Enhancement Therapy for Antidepressant (CETA)” (Günadyın & Barlas, 2017).

Eight of nine studies provided a low dose of intervention, that is, less than 12 sessions of minimum 1 to 2 hours, according to our definition. The ‘dose’ of psychoeducation given to the intervention group varied from 4 sessions over 8 weeks, giving the lowest dose (Kumar & Gupta, 2015), to 15 weekly sessions of 1-2 hours, representing the highest dose (Chetty & Hoque, 2013). Thus, this latter study (Chetty & Hoque, 2013) had a moderate dose of intervention. One study (Dalgard, 2006) had booster sessions: one, two and four months after the end of the psychoeducation course, giving a total dose of 11 sessions. Time of the final assessment varied between 10 weeks and 52 months after randomization.

Two studies included caregivers in the group psychoeducation (Cohen, O’Leary & Foran, 2010 and Kumar & Gupta, 2015) and we therefore regarded these interventions as family group psychoeducation. While, Aagaard et al. (2017) had caregiver included in one session, but we did not categorize this study as family group psychoeducation because the majority of the intervention was provided in a patient group setting. Figure 3 gives an overview of the type of psychoeducation provided in the included studies

Figure 3. Overview of types of psychoeducation provided in the included studies



Nurses seem to play a prominent role in the delivery of GPE interventions. In four of the included studies, nurses conducted the intervention alone or in combination with other professionals (Casañas et al., 2012, Chetty & Hoque, 2013, Dalgard, 2006 and Dowrick et al., 2000). A clinician was the facilitator in one study (Cohen, O'Leary & Foran, 2010) and the researcher in another (Kumar & Gupta, 2015). Two studies do not provide information on the facilitator of the intervention (Günadyın & Barlas, 2017 and Sharif & Ashkani & Zoladl, 2012). Since the studies was conducted in a hospital setting assume that the facilitator of the intervention had relevant professional background. Table 6 gives an overview of the intervention in the included studies.

Table 6. Overview of the intervention in the included studies

Study	Description of intervention	Dose	Facilitator
Aagaard 2017	**Group Psychoeducation Programme** (PEP) covering depressive disease, psychopharmacological treatment, depression and anxiety, impact of depression on family and work, and coping competence to the disease. Homework was considered an important component of the programme- Group composition: 6-8 patients and two conductors A caregiver was present in one session	8 weekly sessions of 120 minutes	Highly experienced group therapist or therapists under training
Casañas 2012	**Group Psychoeducation** including education on health education, diet, physical exercise, sleep, pharmacological treatment, breathing techniques, problem solving, behavioral activation, cognitive behavioral perspective on depression, self-esteem, self-image, pleasant activities, social skills and assertiveness. Group composition: 8-12 patients and 2 nurses	12 weekly sessions, 90 minutes	Primary care centre nurse
Chetty 2013	**Nurse-facilitated-cognitive-group (FCG)** intervention based on Verona Gordon`s (1988) Women's Workbook and Facilitator`s Manual. Goal- setting, depression, self-worth, relationships, assertiveness, conflict-management, stress, nutrition and exercise are topics that were covered. Group composition: 5-15 patients and 2 psychiatric nurses	15 weekly sessions, 60-120 minutes	Advanced psychiatric nurse
Cohen 2010	**Family psychoeducation** for couples where the women was depressed Group composition; patient and caregiver in group with a clinician	5 weekly sessions of 120 minutes	Clinician
Dalgard 2006	**Group Psychoeducation** didactic course aiming at promoting positive thinking, pleasant activities, social skills and social	8 weekly sessions of 2.5 hours, plus booster sessions 1,2 and 4 months after the course	Professionals mainly nurses with background in

	support. Homework was considered an important component of the programme Group composition: 8-10 patients and 2 professionals		psychiatry and primary healthcare
Dowrick 2000	**Intervention I: Group psychoeducation.** A modified version of "coping with depression course" with emphasis on social support was provided to the allocated participants Group composition: patients and facilitator **Intervention II: Problem-Solving treatment** was given to the patient on a one to one basis usually in the patient's home	**Intervention I:** 12 sessions of 2 hours, over 8 weeks **Intervention II:** 6 individual sessions providing less than 4 hours total therapy time	Facilitator with qualifications in psychology, nursing, or allied health professions
Günadyın 2017	**Intervention I: Group psychoeducation** based on CETA including sessions on relationship, attitude, depression recognition, compliance, side effects, cognitive- behavioral techniques were also employed in the course. A large bulk of the course is focused on compliance of medication. Group composition: 8 patients and conductor **Intervention II: Brochure psychoeducation,** CETA program delivered on an individual basis as a brochure intervention	**Intervention I:** 5 weekly sessions of 45-60 minutes **Intervention II:** Dose not specified	Psychiatric nurse
Kumar 2015	**Family psychoeducation** covering cause of depression, trigger factors, stigma, pharmacotherapy, suicide, expressed emotion, caregiver burden outcome, quality of life, problem-solving, forming an action plan, and more. Group composition: patient, caregiver and researcher	4 sessions held on week 0, week 2, week 4 and week 8	Researcher
Sharif 2012	**Group psychoeducation** covering depression, signs and symptoms, medication, treatment, side effect of medication, self-esteem, assertiveness, negative thought patterns, rational thinking, social skills training and relaxation. Group composition: 6 patients and conductor	6 weekly sessions of 90 minutes duration	Not specified

3.2.4 Control and treatment as usual

The nine included studies have various descriptions of the basic healthcare provided to all patients, whether they were in the intervention or the control group, and many of the studies have especially scarce information about care provided to the control group. However, our reading and interpretation of the studies are that all participants had access to basic healthcare for MDD, such as access to healthcare professionals and access to antidepressant medication. It would be considered unethical to conduct a study where participants diagnosed with MDD were deprived of all traditional treatments for depression. Hence, all participants seem to have had access to basic level of care for their MDD. This can be considered treatment as usual (TAU). While we did not apply the term "treatment as usual" in our review protocol, the treatments nonetheless fit our description of included comparisons. We note however that the description of the basic level of care must be interpreted from the description given in each publication and from the social context of each study (see table 7). In this systematic review, we label the basic level of care, provided to all patients, “treatment as usual” and define this term as having access to antidepressant medication and access to a healthcare professional. In some of the included studies, "treatment as usual" (TAU) also includes access to psychotherapy and other available treatments.

Table 7. Overview of level of baseline healthcare provided to all patients

Study	Description of baseline treatment of depression provided to all patients (TAU)	Interpretation of level of baseline care (TAU)
Aagaard 2017	“…2-year outpatient follow-up at a CMHC”	The study was conducted in Denmark. The Danish population has readily access to high quality healthcare, such as general practitioner and antidepressant

	52.4% of the intervention group and 36.8 % of the control group intervention received traditional antidepressant medication. Some of the participants used mood stabilising medication (19.0% / 31.6%), antipsychotic medication (14.3% / 18.4%), anxiolytic and hypnotic medication (11.9% / 13.2 %)	medication. We assume all participants had access to excellent healthcare.
Casañas 2012	"Members of the control group received usual treatment (visits to GP and nurses). There was no pattern of visits established; the patients could go to the primary care centre" 55% the intervention group and 45% of the control group received antidepressant medication. Some of the participants also used hypnotic medication (4.8%), anxiolytics medication (54.3%), blood pressure medication (30.3%) and alternative treatment (22.1%).	The study was conducted in Barcelona, Spain. The participants had readily access to a general practitioner or a nurse and access to antidepressant medication.
Chetty 2013	"…participants were assessed by a psychiatric nurse on a monthly basis or by psychiatric clinic doctor if a script needed to be reviewed or if the patient had problems requiring the doctor's attention and had to collect antidepressant medication from the clinic pharmacy." 100% of participants used antidepressant medication	The participants were recruited from a public, community, psychiatric clinic, serviced by the KwaZulu-Natal Provincial Health Services in South Africa. All participants had regular access to a psychiatric nurse or a doctor and all received antidepressant medication.
Cohen 2010	"Couples assigned to the waiting list group were assessed approximately 5 weeks after entry into the study and again three months later" "Female partners were not restricted from receiving concurrent treatments for their depression as long as they had been in individual psychotherapy for a minimum of 12 weeks or taking a stable dose of psychotropic medication for a minimum of 8 weeks"	The study is set in Long Island, USA and the majority of the spouses are employed we therefore assume that the participants had access to healthcare professionals and antidepressant medication.
Dalgard 2006	"The control group as well as the intervention group were free to continue eventual ongoing treatment (i.e. "treatment as usual")." 44.4% of participants in the intervention group and 42.7 % in the control group used	The study was conducted in Norway. The Norwegian population have readily access to high quality healthcare, such as general practitioner and antidepressant medication.

	medication at time of screening. 24% of the intervention group and 12% in the control group received psychotherapy at time of screening.	
Dowrick 2000	"…controls receiving no treatment." "…(26%) reported concurrently taking antidepressants. There were no significant differences in diagnosis or antidepressant receipt between the study sites or intervention arm."	The study was conducted with participants in several countries (Finland, Norway, Republic of Ireland, Spain, United Kingdom) all of which have high quality healthcare readily available to the population. We assume all participants had access antidepressant medication and healthcare professionals
Günadyın 2017	"… the patients who met the inclusion criteria were divided into experimental ((psychoeducation (n=49), brochure (n=51)) and control (medication (n=53)) group. The psychoeducation group (experimental group1) both received group psychoeducation (usual care+CETA; Continuity Enhancement Therapy for Antidepressants) in five sessions and antidepressant treatment. The brochure group (experimental group 2) both received CETA with a brochure and antidepressant treatment. The medication (control group) group did not receive any psychoeducation about depression, antidepressants and only continued antidepressant treatments that were given by their psychiatrist" "…..... treatment plan have included that antidepressants and have the antidepressant treatment for the first time." 100% of participants used antidepressant medication	The participants in this study are recruited through a state hospital psychiatric polyclinic in Istanbul. Ethical consideration was obtained from a multicentre research ethics committee and from a public hospital, in accordance with Declaration of Helsinki. We therefore assume that the term "usual care" involves access to healthcare professionals.
Kumar 2015	"As the study was designed to evaluate the role of psychoeducation, the treating clinician was absolutely free to continue the treatment of his/her own choice. Both groups also received the routine unstructured counselling."	The study was conducted in Safdarjung Hospital, New Delhi, India The participants in this study had access to antidepressant medication and healthcare professionals.
Sharif 2012	"They were admitted to Shiraz University of Medical Sciences at the time of study and were on antidepressant medication." "The control group did not receive the intervention" 100% of participants used antidepressant medication	The participants in this study were recruited from two hospitals in Shiraz, Iran and the study were approved by Ethics Committee of Shiraz University of Medical Sciences. We therefore assume that the hospital provided access to healthcare professionals for all participants. Antidepressant medication was a part of the inclusion criteria at screening.

3.2.5 Comparisons and number of studies for each comparison

Table 8 gives an overview of comparisons included and number of studies in each comparison. Our protocol outlines that we should do sub-analyses for FGPE and for dose of psychoeducation if possible. Sub-group analyses were performed to examine the effect of FGPE, but it was not possible to do sub-group analyses on dose, because all but one study provided a low dose of the intervention. The following three meta-analyses were performed on the reported outcomes from the included studies;

- Effect of GPE and TAU vs TAU
    - Subgroup analysis: Effect of FGPE in conjunction with TAU vs patient group psychoeducation in conjunction with TAU
- Effect of FGPE in conjunction with TAU versus TAU

The comparisons are described and displayed in table 8.

Table 8. Comparisons and studies included

Main comparison	**Number of studies** (Study)
Group psychoeducation and TAU vs TAU	9 (Aagaard 2017, Casañas 2012, Chetty 2013, Cohen 2009, Dalgard 2006, Dowrick 2000, Günadyın 2017, Kumar 2015, Sharif 2012)
Sub-group comparison	**Number of studies** (Study)
Family group psychoeducation and TAU vs patient group psycho education and TAU	4 (Casañas 2012,Cohen 2009, Günadyın 2017, Kumar 2015,)

3.2.6 Reported outcomes

Primary outcomes: Quality of life was reported in three studies; (Casañas et al., 2012, Dowrick et al., 2000 and Sharif, Nourian, Ashkani & Zoladl, 2012), using 2 different tools (EQ-5D and FS-36). Depression severity was reported in 8 studies using 2 different tools (BDI and HAM-D). None of the nine studies included had any patients lost to suicide.

Secondary outcomes: Level of psychosocial functioning was reported in 2 studies, (Cohen, O'Leary & Foran, 2010 and Kumar & Gupta, 2015) using two different tools (IRBAS, GAF). Relapse (readmittance) was reported in one study (Aagaard et al., 2017). Adherence to treatment was indirectly addressed in two studies. It was reported as drop-outs by Aagaard et al. (2017) who concluded that belonging to the control group was a significant contributor to the drop outs / non-compliance rate. Cohen, O'Leary and Foran (2010), reported on change in compliance of medication and concluded there was no difference in medication adherence between the two groups. Neither of these two studies directly looked at treatment adherence, thus there will be no further discussions on this outcome.

Several tools are available to assess depression, quality of life and level of psychosocial functioning (Wei, McGrath, Hayden, Kutcher, 2016). Table 9 presents the tools used for measuring the outcomes of interest in this review.

Sharif, Nourian, Ashkani & Zoladl (2012) reported the FS-36 outcome in an unrecognisable format with some results being out of scale. In this study the self-report tool used to measure quality of life, (FS-36) was completed by a researcher and not by the participant. We contacted an expert at https://www.optum.com for advice on interpretation of these results (Bjørner, J. B. personal communication July 2018) and was advised to exclude the results from further analysis, which we did.

FS-36 results from Dowrick et al. (2000) was reported in several subgroups without a summarised total effect and could hence not be included in a meta-analysis with EQ-5D results from Casañas et al. (2012).

Table 9. Description of tools used for measuring the outcomes in the included studies

Tool	Study	Outcome measured	Description of tool *
Beck Depression Inventory (BDI)	Aagaard 2017 Casañas 2012 Chetty 2013 Cohen 2010 Dalgard 2006 Dowrick 2000 Günadyın 2017	Level of depression	21-item self-report tool measuring depressive symptoms. Several versions of BDI are available (Cohen et al., 2009). BDI-II: Sensitivity: 81% Specificity: 92% (Psych Congress network BDI-II, 2018)
Hamilton Rating Scale for Depression (HAM-D)	Cohen 2010 Kumar 2015	Level of depression	21-item reporting scale for measuring depression. Should be administered by a clinician experienced in psychology. Several versions of the HAM-D are available. Sensitivity: 86.4% Specificity: 92.2% (Strik, J.J., Honig, A., Lousberg, R. & Denollet J. 2001)
EQ-5D	Casañas 2012	Quality of life (Health Status)	Self-report tool (Herdman et al., 2011). Health status is measured using 5 dimensions; mobility, self-care, usual activities, pain/discomfort, and anxiety/depression
The Short Form (36) Health Survey (FS-36)	Dowrick 2000 Sharif 2012	Quality of life (Health status)	36- item self-report tool. Several versions of the FS-36 are available.
Illness-Related Behaviors and Attitudes Scale (IRBAS)	Cohen 2010	Psychosocial functioning	Self-report tool. Several versions of the IRBAS are available (Cohen et al., 2009).
Global assessment of functioning (GAF)	Kumar 2015	Psychosocial functioning	Clinician administered assessment (Gold 2014)

Legend: *Tools used to measure outcomes outside our interest are not included in the table.

3.2.7 Risk of bias (RoB) assessment of included studies

The nine included RCTs were assessed for risk of bias in accordance with our protocol.

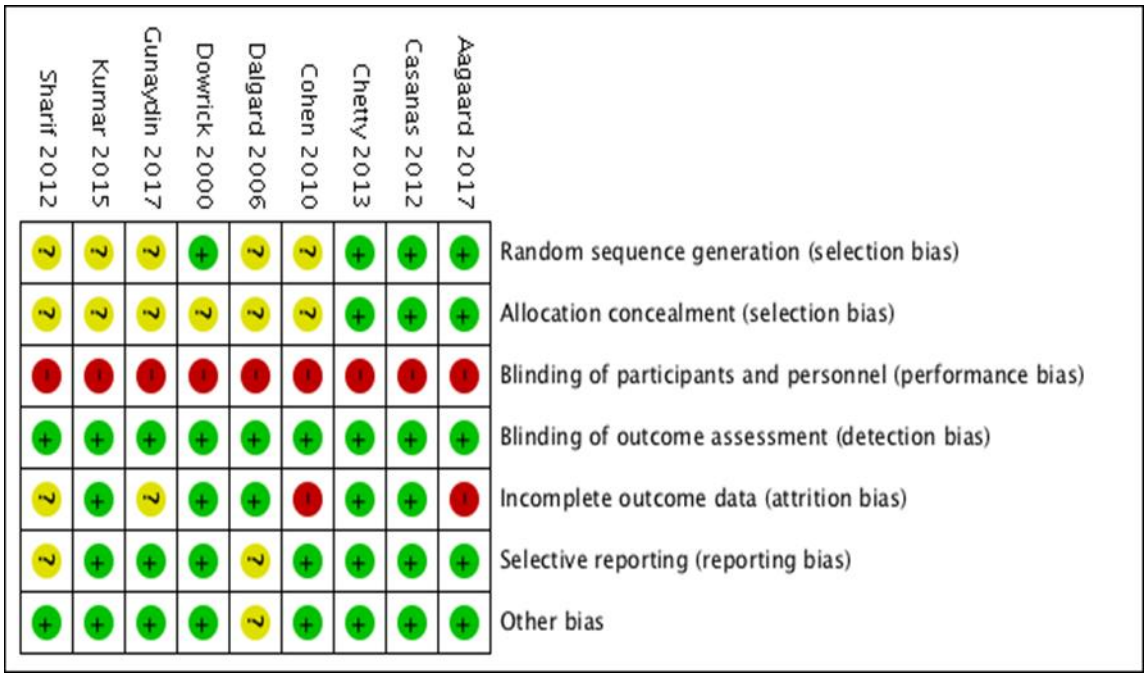
Blinding of participant and personnel is challenging for the intervention studied. We therefore

graded lack of blinding down for all of the 9 studies, under the category “assessment of blinding of participant and personnel” during the RoB assessment.

Five of the studies (Dalgard (2006), Cohen, O’Leary & Foran, 2010, Günadyın & Barlas, 2017, Kumar & Gupta, 2015 and Sharif & Ashkani & Zoladl, 2012) were considered to have unclear risk of bias in the randomization process. Four of these used a sub-optimal method and one had poor description of the randomization process. We attempted, without success, to contact the author of one of these five studies (Günadyın & Barlas, 2017) to get more information on the randomization process as the article provides scarce information: “Randomization methods were employed to achieve homogeneity among the groups”.

The result of the risk of bias assessment is displayed in table 10 and generated using the RevMan tool (Review Manager, 2014). The complete risk of bias assessment is found in appendix 5.

Table 10. Risk of bias assessment of the included studies



3.2.8 Summary of findings

Three meta-analyses and forest plots were created for the outcomes depression and psychosocial functioning. One of these was a subgroup analysis looking into family group psychoeducation versus patient group psychoeducation. The quality of evidence gained from the two main meta-analyses were evaluated by using The Grading of Recommendations Assessment (GRADE) online software developed by the GRADE working group to conduct systematic, transparent and pragmatic grading of strength of evidence from meta-analyses (Guyatt G. H. et al. 2008). The summary of findings from the GRADE evaluation is presented here and the evidence tables can be found in appendix 8.

Table 11. Summary of findings for the comparisons

Group psychoeducation in conjunction with treatment as usual versus treatment as usual
Population: Adults with Major Depressive Disorder
Intervention: Group psychoeducation in conjunction with treatment as usual
Comparison: treatment as usual

Outcomes	Illustrative comparative effect (95% CI)		Relativ effect (95% CI) [1]	Number of participants (studies)	Quality of the evidence (GRADE)	Comments
	Control group	**Intervention group**				
Depression symptoms severity at 4-6 weeks	Range 17.61-26.29	Range 15.62-18.38	**SMD 0.32 SD lower** (0.59 lower to 0.04 lower) in the intervention group	204 (3 RCTs)	⊕⊕⊕◯ MODERATE[2]	BDI and HAM-D
Depression symptoms severity at 3 months	Range 14.71-26.42	Range 8.43-17.53	**SMD 0.61 SD lower (**1.14 lower to 0.09 lower) in intervention group	432 (4 RCTs)	⊕⊕⊕◯ MODERATE[3]	BDI and HAM-D
Depression symptoms severity at 6 months	Range 14.97-18.3	Range 11.18-17.5	**SMD 0.21 SD lower** (0.38 lower to 0.04 lower) in intervention group	756 (5 RCTs)	⊕⊕⊕⊕ HIGH	BDI and HAM-D
Depression symptoms severity at 12 months	Range 12.6-16.0	Range 14.6-18.8	**SMD 0.22 SD higher** (0.02 lower to 0.45 higher) in intervention group	283 (2 RCTs)	⊕⊕⊕◯ MODERATE[4]	BDI
Quality of life 3, 6 ,9 and 12 months	Range 55.54-57.69 Range 34.05-70.39	Range 57.7-59.2 Range 38.78-68.31	No significant difference	528 (2 RCTs)	⊕⊕⊕⊕ HIGH	EQ-5D SF-36
Relapse 2 years before and 2 years after inclusion	No. of admissions: 38	Nr. of admissions: 42	No significant difference	80 (1 RCT)	⊕⊕⊕◯ MODERATE[5]	
Psychosocial functioning at 4-5 weeks	Range 33.45-62.0	Range 40.0-72.0	**SMD 1.07 SD higher** (0.65 higher to 1.48 higher) in intervention group	48 (2 RCTs)	⊕⊕◯◯ LOW[6]	IRBAS GAF
Psychosocial functioning at 12 weeks	Range 33.83-76.1	Range 42.4-84.0	**SMD 0.98 SD higher** (0.56 higher to 1.40 higher) in intervention group	46 (2 RCTs)	⊕⊕◯◯ LOW[7]	IRBAS GAF

Explanations:
1) RR: relative risk; CI: Confidens interval, SMD, SD
2) Certainty of evidence for Depression symptoms at 4-6 is downgraded due to imprecision.
3) Certainty of evidence for Depression symptoms at 3 months is downgraded due to inconsistency
4) Certainty of evidence for Depression symptoms at 12 months is downgraded due to imprecision
5) Certainty of evidence for Relapse is downgraded due to imprecision
6) Psychosocial functioning at 4-5 weeks is downgraded due to serious imprecision
7) Psychosocial functioning at 12 weeks is downgraded due to serious imprecision

3.3 Effect of group psychoeducation and TAU versus TAU

Reported outcomes for the comparison group PE and TAU versus TAU were depression, quality of life, relapse, and psychosocial functioning. Sub-group analysis was made for family group psychoeducation versus patient group psychoeducation.

3.3.1 Effect on depression

Eight studies measured effect of GPE on depression. BDI was used to evaluate effect of the intervention in 7 of the studies and HAM-D was used in two studies. One study (Cohen, O'Leary & Foran, 2010) reported both BDI and HAM-D. The HAM-D result was chosen for further analysis.

Chetty & Hoque (2013) did not report standard deviation and the results could therefore not be included in the meta-analysis. Seven studies (Aagaard et al., 2017, Casañas et al., 2012, Cohen, O'Leary & Foran, 2010, Dalgard, 2006, Dowrick et al., 2000, Günadyın & Barlas, 2017 and Kumar & Gupta, 2015) were considered to be sufficiently similar and therefore included in a meta-analysis, giving a total of 969 pooled patients. Meta-analysis was made for follow-up after 4-6 weeks, 3 months, 6 months and 12 months. Only one study, (Aagaard et al., 2017), had follow-up of patients beyond 12 months, assessing the patients at 18 months and 24 months. The results are presented in table 12.

Table 12. Results of GPE and TAU versus TAU on MDD

Study Yr **N** **Tool**	**4-6 weeks** Mean score (SD)		**8 weeks** Mean score (SD)		**3 months** Mean score (SD)		**6 months** Mean score (SD)		**9 months** Mean score (SD)		**12 months** Mean score (SD)		**18 months** Mean score (SD)		**24 months** Mean score (SD)	
Aagaard 2017							Interv.	Control			Interv.	Control	Interv.	Control	Interv.	Control
N=80							17.5	17.5			18.8	16.0	14.6	15.5	14.7	17.3
BDI							(12.6)	(12.4)			(13.6)	(11.6)	(12.0)	(12.2)	(12.6)	(11.0)
Casañas 2012 *					Interv.	Contro	Interv.	Control	Interv	Control						
N=231 BDI II					15.42	17.54	15.37	16.51	15.09	16.35						
					(7.53)	(7.18)	(8.74)	(7.60)	(8.62)	(7.84)						
Chetty 2013 **	Interv.	Control			Interv.	Control										
N=30	17.9	20.7			14.6	21										
BDI																
Cohen 2010	Interv.	Control			Interv.	Control										
N=35	18.38	26.29			13.60	26.42										
HAM-D	(10.77)	(10.55)			(11.43)	(12.25)										
Dalgard 2006							Interv.	Control								
N=155							14.1	18.1								
BDI							(9.5)	(9.6)								
Dowrick 2000							Interv.	Control			Interv.	Control				
N=297							14.26	14.97			14.60	12.60				
BDI							(9.71)	(10.23)			(8.75)	(9.50)				
Günadyın 2017	Interv.	Control			Interv.	Control	Interv.	Control								
N=102	18.00	19.73			17.53	18.59	11.18	16.16								
BDI	(11.50)	(10.46)			(12.01)	(13.72)	(10.36)	(12.95)								
Kumar 2015	Interv.	Control	Interv.	Control	Interv.	Control										
N=80	15,62	17.61	12.72	16.21	8.43	14.71										
HAM-D	(5.25)	(4.92)	(5.10)	(4.82)	(5.90)	(3.4)										

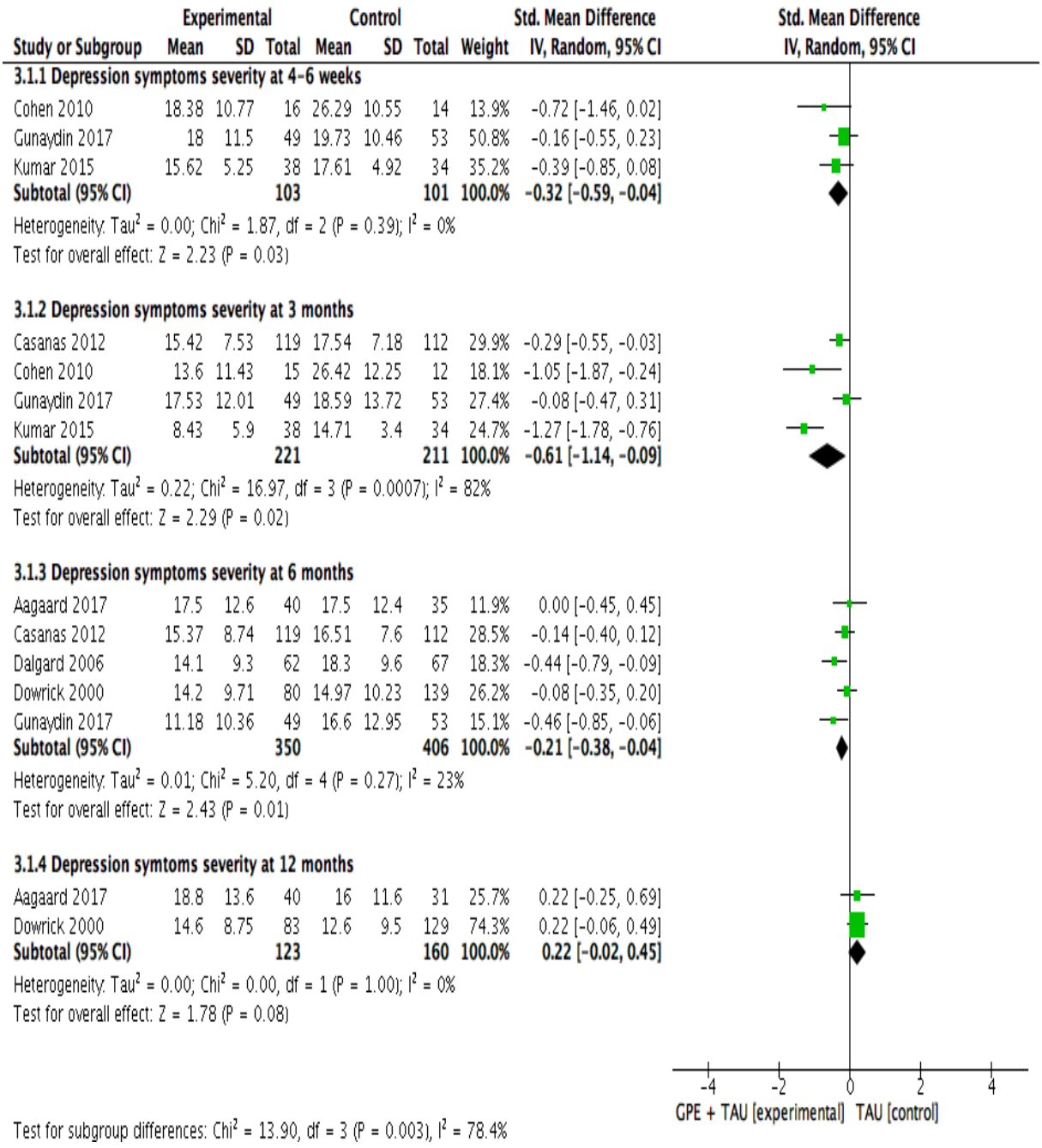
Legend: Interv. = intervention group, Control= control group, SD=standard deviation, N= no of participants

* Split results, reporting effect on mild and moderate depression are available, but not included.

** Standard deviation not reported.

A forest plot displaying the effect of GPE on depression at various follow-up times (subtotals at 4-6 weeks, 3 months, 6 months and 12 months) after randomization was made using the RevMan tool (Review Manager, 2014) and displayed in figure 4.

Figure 4. Meta-analysis of group psychoeducation and TAU versus TAU on depression



The effect of GPE on reduction of depression after 4-6 weeks is (n=204): SMD = -0.32 (CI: -0.59 to -0.04), with no statistical evidence of heterogeneity (I^2=0%, P=0.39), and the result is significant at 5% level (P=0.03). None of the three individual studies give a significant result as the confidence intervals are crossing the line of no effect, but the pooled result shows a significant effect with a small effect size, favouring GPE in conjunction with TAU over TAU on reduction of depression.

The effect of GPE on reduction of depression after 3 months is (n=432): SMD= -0.61 (CI: -1.14 to -0.09), with statistical evidence of heterogeneity (I^2 = 82% and P=0.0007). The effect is significant at 5% level (P=0.02), but the high level of heterogeneity indicates systematic differences among the studies. Three out of four studies show some reduction in level of depression in the group receiving psychoeducation. One study (Günadyın & Barlas, 2017) has no significant result as the confidence interval is crossing the line of no effect. The two studies with highest positive effect of the intervention (Cohen, O'Leary & Foran, 2010 and Kumar & Gupta, 2015) involves a caregiver and these studies therefore belong to the FGPE sub-group.

The effect of GPE on depression after 6 months small (n= 756): SMD= -0.21 (CI: -0.38 to -0.04). The combined effect of the five studies is significant at 5% level (P=0.01), in favour of GPE. Three of the studies are crossing the line of no effect (Aagaard et al.,2017, Casañas et al., 2012, and Dowrick et al., 2000). There is no statistical evidence of heterogeneity (I^2=23%, P=0.27).

Only two studies have measurements from patient assessment after 12 months (Aagaard et al., 2017 and Dowrick et al., 2000). The combined effect after 12 months is small (n= 283):

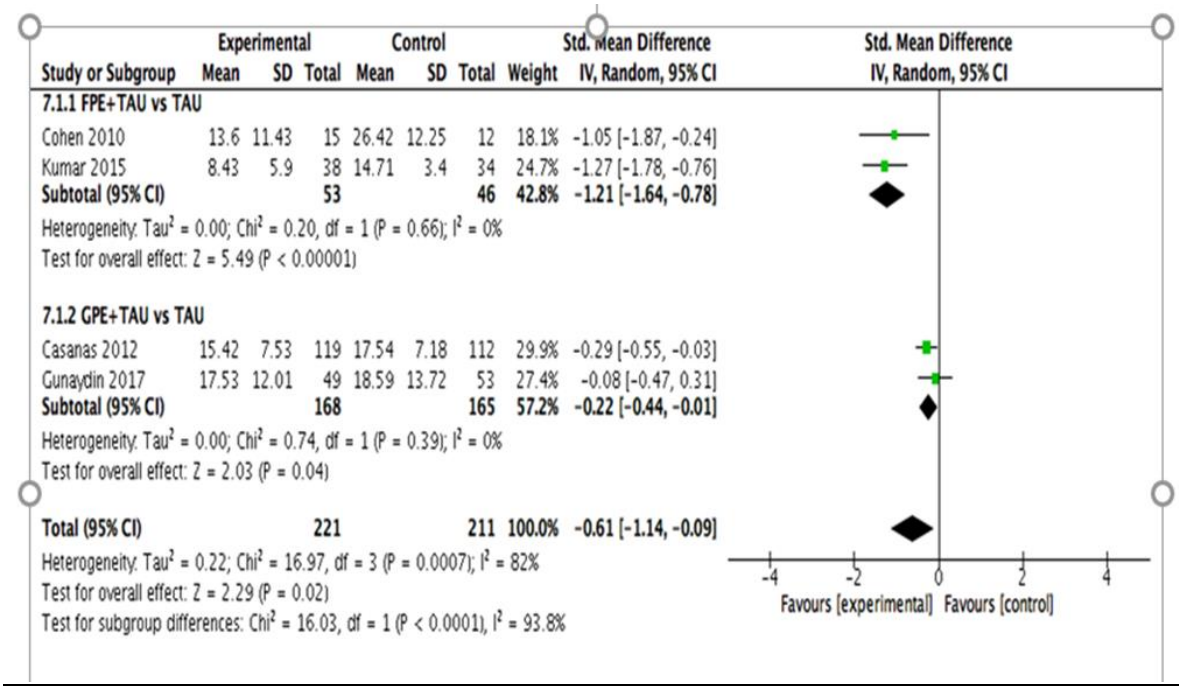
SMD= 0.22 (CI: -0.02 to 0.45) and crossing the line of no effect and hence not significant at 5% level (P=0.08). There is no statistical evidence of heterogeneity (I^2=0%, P=1.00).

We graded the certainty of the body of evidence for the comparison GPE in conjunction with TAU versus TAU to be of moderate quality except for at 6 months where the body of evidence is graded high. The GRADE tables are submitted in appendix 8.

3.3.1.1 Sub-group analysis

We wanted to examine the effect on the patient when including a family member or a caregiver in the psychoeducation program;FGPE. Sub-analysis was made for the effect of FGPE and TAU compared to patient group psychoeducation (PGPE) and TAU on depression. It was possible to make a sub-group meta-analysis, comparing FGPE and PGPE on the effect of depression, at 3 months after randomization. Two studies (Casañas et al., 2012 and Günadyın & Barlas, 2017) measured the effect of PGPE and two studies (Cohen, O' Leary & Foran, 2010 and Kumar & Gupta, 2015) measured the effect of FGPE on depression. The total number of participants in the meta-analysis was 432. The numeric result used in this sub-group analysis is found in table 12. The results are displayed in the forest plot in figure 5.

Figure 5. Sub-group analysis: FGPE versus PGPE analysis.



Effect of FGPE at three months follow up (n= 99) is: SMD= -1.21 (CI: -1.64 to -0.78). There is no statistical evidence of heterogeneity (I^2=0%, P=0.63) and the effect is significant at <1% level (P<0.00001) suggesting a considerable effect of FGPE on reducing level of depression. The test for sub-group difference (Chi^2 =16.03, P<0.0001) shows that there is a significant difference in effect between FGPE and PGPE.

Effect of PGPE at three months is small (n=333), SMD= -0.22 (CI: -0.44 to -0.01), with no statistical evidence of heterogeneity (I^2=0% and P=0.39). The effect is significant at 5% level (P=0.04). One of the two studies (Günadyın & Barlas, 2017) in this sub-group analysis has confidence interval crossing the line of no effect. But the combined result of the two studies suggest a positive effect of GPE when compared with treatment as usual.

3.3.2 Effect on quality of life

Three studies measured the effect of PGPE in conjunction with TAU versus TAU on quality of life (Casañas et al., 2012, Dowrick et al., 2000 and Sharif & Ashkani & Zoladl, 2012). Casañas et al. (2012) used EQ-D5 as measuring tool while the two others used FS-36 to measure quality of life. Sharif, Nourian, Ashkani & Zoladl (2012) reported results in a controversial manner and these results are excluded from this review as discussed in chapter 3.2.6. Dowrick et al. (2000) reported the results in subgroups and it is therefore impossible to include this study in a meta-analysis on the outcome quality of life. The range of the means for the control group was 34.05- 70.38. The range of the means for the PGPE was 38.87-68.31. The standard deviation of the various groups seems high in this study. The results form Dowrick et al. (2000) are found in table 14 and table 15 respectively.

Casañas et al. (2012) found no significant effect of PGPE in conjunction with TAU on quality of life as compared to TAU. Dowrick et al. (2000) found some effect of PGPE on quality of life after 6 months, but no effect after 12 months when compared to treatment as usual. The results from Casañas et al. (2012) is presented in table 13.

*Table 13. Results of PGPE and TAU versus TAU on quality of life (*Casañas et al., 2012)

Casañas 2012 N=231 Tool: EQ-5D	**3 months Mean score (SD)**		**6 months Mean score (SD)**		**9 months Mean score (SD)**	
Group Mean (SD) Difference* (95% CI)	Interv. 59.7 (18.1) 8.97 (12.2 to 5.72)	Control 55.54 (16.36) 2.29 (4.6 to -0.01)	Interv. 57.9 (20.7) 7.09 (10.78 to 3.39)	Control 57.05 (16.97) 3.80 (6.98 to 0.61)	Interv. 59.2 (20.8) 8.46 (11.99 to 4.93)	Control 57.69 (17.325) 4.44 (8.0 to 0.87)
Difference (95% CI between groups (intervention - control)	Difference 4.19 (-0.31 to 8.66) P= 0.067 SES** =0.26		Difference 0.81 (-4.12 to 5.73) P=0.748 SES**=0.05		Difference 1.54 (-3.43 to 6.51) P=0.543 SES**= 0.09	

Legend: SD= standard deviation, CI= confidence interval
*difference was calculated between follow-up measurement and baseline measurement by the authors.
**difference was calculated between intervention group and control group by the authors.

Table 14. Results of PGPE and TAU versus TAU on quality of life (Dowrick et al.,2000)

Treatment v control	**Outcome**	**6 months**		**12 months**	
		Mean (95% CI)	**P value**	**Mean (95% CI)**	**P value**
Depression prevention	SF-36 (mental role)*	12.70 (0.46 to 24.94)	**0.042**	-4.02 (-14.53 to 6.49)	0.454
	SF-36 (social function) **	8.66 (0.07 to 17.25)	**0.048**	2.36 (-6.10 to 10.83)	0.584
	SF-36 (mental health) ***	6.95 (0,76 to 13,14)	**0.028**	-3.25 (-8.47 to 1.97)	0.223

Legend: CI= confidence interval, *after controlling for baseline BDI score, mental role scores and random centre effect. **after controlling for baseline BDI score, social functioning scores and random centre effect. ***after controlling for baseline BDI score, mental health scores and random centre effect.

Table 15. Outcomes for SF-36 (Dowrick et al., 2000)

FS-36	**Baseline Mean (SD)**	**6 months Mean (SD)**	**12 months Mean (SD)**
Mental role:			
-Control	34.05 (38.26)	51.71 (42.70)	63.62 (41.90)
-Intervention	38.87 (38.53)	64.90 (40.70)	61.42 (40.48)
Social function:			
-Control	59.46 (29.23)	64.90 (32.46)	70.39 (30.09)
-Intervention	48.62 (28.23)	68.31 (29.07)	66.89 (27.33)
Mental health:			
-Control	43.51 (17.73)	53.71 (23.58)	60.51 (22.39)
-Intervention	42.98 (16.39)	59.54 (21.41)	57.11 (20.33)

3.3.3 Effect on relapse

One study (Aagaard et al., 2017) reported on admittance to psychiatric hospital 2 years before and two years after the date of inclusion. This outcome can be interpreted as a measurement for relapse. The results, given in table 16 show no significant difference between intervention group and control group on admittance to psychiatric hospital.

Table 16. Admittance to psychiatric hospital 2 years before and 2 years after date of inclusion to the study (Aagaard et al., 2017)

No of admissions	**Intervention**		**Controls**	
	Before	**After**	**Before**	**After**
0	15 (36%)	30 (71%)	13 (34%)	28 (74%)
2	17 (40%)	10 (24%)	15 (39%)	9 (24%)
3	2 (%5)	0 (0%)	2 (5%)	0 (0%)
Total	42	42	38	38
duration, days Mean (SD)	33.5 (42.7)	5.0 (16.2)	47.0 (63.7)	8.5 (19.9)
Median (range)	26 (0-195)	0 (0-82)	15.5 (0-209)	0 (0-97)

3.3.4 Effect on psychosocial functioning

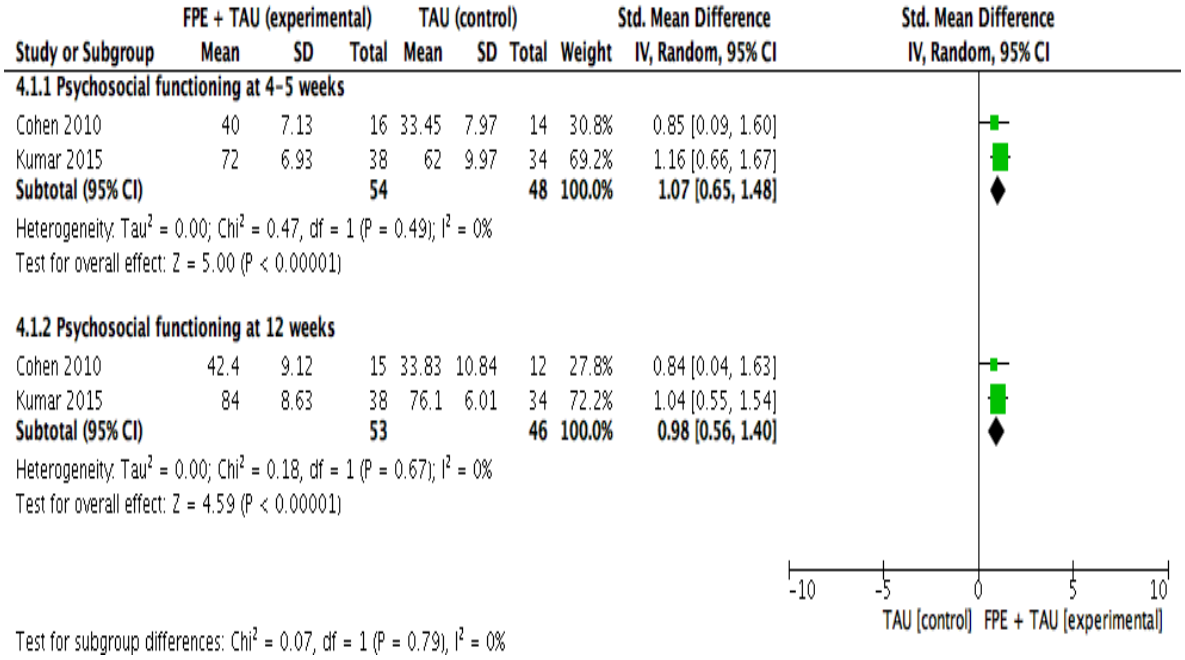
Cohen, O'Leary and Foran (2010) and Kumar and Gupta (2015) examined the effect of FGPE on psychosocial functioning using the measurement tools IRBAS and GAF, respectively. The numeric results are displayed in table 17. The pooled effects after 4 to 6 weeks and after 12 weeks are displayed in a forest plot, figure 6. An *increase* in the level of psychosocial functioning is a desired effect.

Table 17. Results of FGPE and TAU vs TAU on psychosocial functioning

Study Year **No** **Measurement tool**	**4-5 weeks** **Mean** **(SD)**		**8 weeks** **Mean** **(SD)**		**12weeks** **Mean** **(SD)**	
Cohen 2010 N=35 IRBAS	Intervention 40 (7.13)	Control 33.45 (7.97)			Intervention 42.40 (9.12)	Control 33.83 (10.84)
Kumar 2015 N=80 GAF	Intervention 72 (6.93)	Control 62 (9.97)	Intervention 75 (9.20)	Control 67.56 (8.34)	Intervention 84 (8.63)	Control 76.1 (6.01)

Legend: GAF= Global assessment of functioning, IRBAS=Illness-Related Behaviors and Attitudes Scale, SD=standard deviation

Figure 6. Meta-analysis of FGPE and TAU vs TAU on psychosocial functioning



There was a considerable effect of FGPE on psychosocial functioning after 4-5 weeks, (n=102): SMD= 1.07 (CI: 0.65 to 1.48). There is no statistical evidence of heterogeneity (I^2=0%, P=0.49) and the effect is significant at <5% level (P<0.00001).

There was also an effect of FGPE on psychosocial functioning after 3 months (n=99): SMD= 0.98 (CI: 0.56 to 1.40). There is no statistical evidence of heterogeneity (I^2=0%, P=0.67) and the effect is significant at <5% level (P<0.00001).

This is suggestive of a positive effect of FGPE in conjunction with TAU versus TAU on psychosocial functioning after 4-5 weeks and after 3 months. We graded the body of evidence for this comparison to be of low quality, see table 17. There is a serious downgrade of the quality of evidence due to small sample size, at 4-5 weeks (n=48) and at 3 months (n=46).

4 Discussion

4.1 Summary of main results

This review summarizes research on the effect of GPE for adults with major depressive disorder. We preselected six outcomes of interest, including quality of life and outcomes we considered of importance for quality of life such as depression severity, level of psychosocial functioning, relapse, and treatment adherence. Mortality was also prespecified as an outcome of interest. We summarize the main findings below.

4.2 Effect of GPE and TAU versus TAU

4.2.1 Effect on depression

The pooled effects of GPE in conjunction with TAU versus TAU on severity of symptoms of depression at 4-6 weeks and at 6 months after randomization showed a small and statistically significant effect. At 3 months the result showed substantial heterogeneity and the effect was not significant at 12 months post randomization. The high heterogeneity at 3 months cannot be random as the p-value shows significant systematic difference between the studies. Hence, our confidence in the pooled effect on reducing depression, measured at 3 months, is low. It is however interesting to elaborate on the mechanisms accounting for the systematic heterogeneity at 3 months. This will be discussed under the sub-group analysis.

There are five studies contributing to the body of evidence at 6 months post randomization, included in the meta-analysis. The pooled effect of the five studies show a small benefit of GPE in conjunction with TAU as compared to TAU alone. A cautious conclusion in favor of GPE can be made.

The pooled effect of GPE in conjunction with TAU on depression after 12 months did not reach the statistical significance. GPE in conjunction with TAU does not seem to have the desired effect on reducing depression at 12 months after randomization when compared to TAU.

The evidence of the effect of GPE in conjunction with TAU compared to TAU was evaluated by applying GRADE methodology. The GRADE evidence was considered to be of moderate confidence at 4-6 weeks, at 3 months and at 12 months but of high confidence at 6 months. At 4-6 weeks assessment, the quality of evidence was graded down from high to moderate due to serious imprecision; too small sample size. At 3 months assessment, the quality of evidence was downgraded due to serious inconsistency. At 12 months assessment, the small sample size led to downgrading.

When looking at SMD from the various follow-up times compared to the scales of BDI and HAM-D, the size of the change in level of depression seems to be small. Below is an account for the clinical grading of the participants according to the BDI or HAM-D score measured.

Table 17. Level of depression as expressed by the clinical grading tools, BDI, BDI-II and HAM-D.

Tool	**Level 1**	**Level 2**	**Level 3**	**Level 4**	**Level 5**
BDI and BDI-II Assessment score	Minimal depression 0-13	Mild depression 14-19	Moderate depression 20-28	Severe depression 29-63	Not relevant
HAM-D Assessment score	Normal 0-7	Mild depression 8-13	Moderate depression 14-18	Severe depression 19-22	Very severe ≥23

For a researcher it is encouraging to get a significantly lower result in an intervention group, however a small effect of treatment does not necessarily reduce the participants experienced disease burden. The clinical effectiveness and the value for the participant must be considered

the most important outcome. Some of the 9 included studies have also measured other outcomes and concurrent improvement of several outcomes gives a stronger evidence of effect. BDI and HAM-D are self-assessment tools for depression symptoms. This means that the participants in the intervention group may have overrated their self-assessment score. Resulting in an artificially inflated effect of the intervention. We consider the evidence of the positive effect of GPE on depression to be valid, although it is of low magnitude.

4.2.1.1 Sub-group analysis

At three months follow-up the results on depression showed significant hetrogeniety. We elaborated on this by performing a sub-group analysis. The four studies in the 3-month analysis are clustered into two groups. The most encouraging effect is found in the studies performed by Cohen, O'Leary and Foran (2010) and Kumar & Gupta (2015). These two studies showed significant and similar large positive effect of FGPE on depression. Casañas et al. (2012) and Günadyın & Barlas (2017) showed a much smaller and non-significant effect of PGPE. These 4 studies are conducted in different countries (India, Spain and Turkey and USA), in different socioeconomic conditions and different cultures. The two studies contributing to the more effective cluster, have included a caregiver in the intervention and hence measures the effect of FGPE. The two studies showing greatest effect are Cohen, O'Leary & Foran (2010) studying FGPE amongst depressed females including their spouse as caregiver in Long Island, New York, USA and Kumar & Gupta (2015) studying depressed adults in New Delhi, India. In the Kumar & Gupta (2015) study, the majority of the depressed participants and their caregivers were males. In sum, FGPE in conjunction with TAU shows encouraging positive results on reducing depressive symptoms, in studies from two vastly different contexts. It should however be noted that the sample size is small (n=48).

In conclusion, although there are few studies, when assessed at 3 months after randomization, FGPE seems to have a better effect on reducing depressive symptoms than PGPE.

At 6 months after randomization, there are no data recorded for FGPE as both studies examining effect of FGPE stops at 3 months follow-up. Beyond this point we have no information on the effect of FGPE.

It should be noted that the studies measuring FGPE have follow-up assessment between 4 weeks and 3 months, whereas the studies assessing PGPE have follow-up assessment at from 5 weeks to 12 months. Hence this review provides no information on the effect of FGPE after 3 months and very little information on PGPE before 6 months.

The bulk of results on PGPE is reported at 6 months assessment. Aagaard et al. (2017) is the only study having outcome measures (hospital admittance) at 2 years after randomization. Posternak et al. (2006) studied the course of unipolar MDD in patients not receiving somatic treatment and found that there is a median time to recovery of 23 weeks (5-6 months). Furthermore, the results from the study suggests that a high rate of recovery occurs within the first 3-4 months of an episode. The positive effect on depression and psychosocial functioning could perhaps be explained by time as a confounder. PGPE and other intervention may show similar positive effect in this early period, but we lack data to make any conclusions on the effect of PGPE within the first 3 months.

4.2.2 Effect on quality of life

Only three studies reported on quality of life; (Casañas et al., 2012, Dowrick et al., 2000 and Sharif, Nourian, Ashkani & Zoladl, 2012). The results from Sharif, Nourian, Ashkani & Zoladl (2012) will not be discussed, for reasons outlined in chapter 3.3.2.

Dowrick et al. (2000) show some effects on improving quality of life at 6 months, but the positive effect of psychoeducation on quality of life had diminished by 12 months. Casañas et al. (2012) did not find any significant effect on quality of life of GPE in conjunction to TAU versus TAU at 3 months, 6months and 9 months. This corresponds well with the small or even absent effect on depression in these two studies, se chapter 3.3.1.

4.2.3 Effect on relapse

One study (Aagaard et al., 2017) measured relapse, reporting this outcome as admittance to psychiatric hospital. The study found no significant difference between intervention group and control group on admittance to psychiatric hospital 2 years before and 2 years after intervention. This corresponds well with the lack of effect on depression in this study, se chapter 3.3.3.

4.2.4 Effect on psychosocial functioning

The two studies examining the effect of GPE on psychosocial functioning were both investigating the effect of FGPE (Cohen, O'Leary & Foran, 2010 and Kumar & Gupta, 2015). The SMD effect of FGPE in conjunction with TAU on psychosocial functioning were positive and significant both at 4-5 weeks and at 3 months, suggesting a positive effect of FGPE in conjunction with TAU when compared to TAU alone. The positive effects were reported as an increased level of psychosocial functioning. Although the number of participants contributing to the evidence is small, (n=46 at 4-5 weeks and n= 48 at 3 months) the analysis indicates that it is beneficial for the patient to involve their caregiver in psychoeducational treatment. However, we do not have any measurements on psychosocial functioning beyond 3 months. The evidence was evaluated using the GRADE approach, to be of moderate confidence due to small number of participants (imprecision). These results suggest that caretakers may play an important a role on improving psychosocial functioning in patients with major depressive disorder.

4.3 Agreement with other reviews

We have compared our results with other systematic reviews. We found three relevant systematic reviews that have been conducted recently. The three are discussed below.

Effectiveness of psychoeducation for depression: A systematic review

Tursi et al. (2013) conducted a systematic review on the effectiveness of psychoeducation for depression. In this systematic review there were seven studies applying PE for patients in a group setting, three studies of individual PE and three studies of distant /passive PE. Two

studies included families in the intervention. Two studies overlapped with our overview (Dalgard, 2006 and Dowrick et al., 2000). The results reported are suggesting an association between increased knowledge about depression and its treatment and better prognosis in depression. The main conclusion is that there are only a few studies conducted regarding effectiveness on adult patients with MDD. The authors suggest, despite few publications, that psychoeducation is effective in improving the clinical course, treatment adherence, and psychosocial functioning of depressed patients (Tursi et al., 2013). Tursi et al. (2013) conclude that further RCTs on PE for patients with MDD are still needed to better elucidate the effectiveness of PE.

Psychoeducational treatment and prevention of depression: The "coping with depression" course thirty years later. This meta-analysis conducted by Cuijpers, Munoz, Clarke & Lewinsohn (2009) looked at the Coping with Depression (CWD) course as psychoeducational intervention for prevention and treatment of depression. 18 studies were examining treatment of depression. One study was overlapping with our review (Dowrick et al., 2000). The studies differed considerably from each other, ranging from internet interventions without any professional support to minority groups, adolescents to older adults. Results showed an overall mean effect size in reduction of depression symptoms of 0.28 (95% CI: 0.18 to 0.38), with low to moderate heterogeneity (I^2=31.87). They compared the CWD with other psychotherapies (7 studies) and the mean effect size was a non-significant difference in favor of the other psychotherapies – 0.05 (95% CI: -0.25 to -0.16), with moderate heterogeneity (I^2=48.71). Cuijpers, Munoz, Clarke & Lewinsohn (2009) conclude that the effect sizes found were relatively small and that although many studies found clear evidence of efficacy other studies did not find any effect. They argue that more research is needed to examine the differences in patient groups and their benefit from CWD. Another conclusive reflection in

their study is that very few studies have compared the efficacy of CWD to antidepressants and other psychological treatments, and that more research is needed. They also mention that there is unclarity in terms of which elements in the CWD modules that are effective.

"Family matters": A systematic review of the evidence for family psychoeducation for major depressive disorder

This systematic review conducted by Brandy, Kangas & McGill (2016) is reviewing the evidence for family psychoeducation for MDD. This article is reviewing multi-family psychoeducation (MFPE) versus single family intervention (FPE) for MDD, peer-led FPE versus clinical led groups for MDD and FPE for mixed diagnosis versus FPE for MDD only. Nine data samples were included in this review and one study was overlapping with our review (Kumar & Gupta, 2015). Two FPE studies, with groups consisting of patients and caregiver and not only caregivers, other than Kumar & Gupta (2015) reported positive outcomes for the patients, one of them for depressed adolescents. None of the MFPE interventions included the patients. The findings of this review indicate tentative support for FPE for MDD in improving patient functioning and family well-being, based on a small number of international studies.

Altogether it seems the results in the previous comparable systematic reviews are in accordance with our findings for group psychoeducation; there are few RCT studies conducted on GPE for depression available, and the effect sizes found on reducing depression symptoms and psychosocial functioning are small. Although the studies are few, our review contributes with a new dimension in finding a substantial statistically significant difference in effect size for family group psychoeducation versus patient group psychoeducation.

4.4 Certainty of the evidence

The nature of the invention does not allow blinding of personnel and participant. This was taken into consideration in the GRADE assessment. In several of the studies the participants are self-reporting the experienced level of depression, because there is no available objective measure. This lack of blinding may result in an over estimation of the positive effect of the intervention, (Hawthorne effect).

We graded the confidence in the body of evidence for our comparison in regard to all outcomes of interest. We downgraded the confidence primarily for heterogeneity and for small number of events.To sum up, our confidence in the certainty of the evidence varies from low to high. This means that the results where the body of evidence is graded low has to be interpreted with caution.

4.5 Transferability

The nine included studies have been conducted in vastly different socioeconomic and cultural settings. High-, middle-, and low-income countries are all represented amongst our material and four continents are represented. The two studies on family psychoeducation show similar, encouraging results despite the huge difference in socioeconomic and cultural setting under which these studies were conducted. Group psychoeducation, and in particular family group psychoeducation, is a type of intervention that always carries a cultural aspect and it may well be the case that the treatment is not accepted or effective in all cultures or for all groups of patients. It is a strength for the interpretation of the result that two very different studies give similar and positive results for the patient, with consistent positive effect on both level of depression and on psychosocial functioning. However, generalization of these results cannot be easily done due to imprecision caused by few participants. Additional evidence is needed to be able to generalise the findings in this review.

4.6 Ethics

All the included studies are conducted in an ethically acceptable manner. There is no placebo treatment in the studies, meaning there will be no ethical dilemmas regarding the administration of a placebo intervention to ill patients. Six of the studies have reported that they used written informed consent for all the participants and five studies report that the study has acceptance from an ethical committee. All participants seem to have had access to both healthcare professionals and pharmacological treatment.

The ethics of providing GPE as a treatment option is solely dependent on its proven effectiveness because the intervention has no known negative side-effects.

4.7 Strengths and weaknesses

A strength of this review is that all studies included have RCT design, although four of them were somehow unclear about the randomization process. We had enough comparable evidence to conduct three meta-analyses. The review protocol was preregistered in PROSPERO, minor changes in the protocol are described in chapter 2.8.

There are no suicides reported amongst any of the included patients in our studies. Suicide is the ultimate tragical outcome of MDD and research into suicide prevention is of great importance. Our protocol described suicide as an outcome of interest, but in hindsight we experienced that this study design does not give the desired knowledge regarding suicide because the event is too rare.

There were only two studies on FGPE. It would have been desirable to have a few more studies to underpin or reject our findings. More knowledge on the effect of the intervention is also necessary to enable generalizing the results to different cultures and various patient

groups. FGPE show improvement in two of the outcomes (depression and psychosocial functioning). This strengthens the trust in the effectiveness of the intervention.

It is possible that we did not find all relevant studies during the search and screening process, but we consider this to be a minor chance and a minor limitation to this review. It might also be studies that are published after we ended our data search.

Other limitations to our review are that we did not make any restrictions on measurements tools in our inclusion criteria, nor in the manner in which the outcome data were gathered. Neither did we decide on clinical cut-off points for the outcomes, but rather looked at the change of overall baseline levels of depression and psychosocial functioning.

Last but not least, it would have been a strength to the systematic review if we have had a professional background in clinical mental health, previous to performing this review. However, we consulted several clinicians and the project leader for drug free treatment for psychiatric illness at Vestre Viken health authority and believe this is a minor limitation to this review.

4.8 Implication for practice/policy

Bearing in mind the universal shortage of healthcare resources and vulnerability of patients with MDD, PGPE in conjunction with TAU is according to this review not a satisfactory treatment option for patients with MDD. There is however a small effect of GPE on depression at 4-6 weeks, 3 months and 6 months. This indicates that the intervention has benefit to the patients. We have found no evidence supporting any change in admissions to hospital nor evidence to support change in quality of life.

FGPE in conjunction with TAU needs a greater body of evidence to earn a universal recommendation for patients with MDD. If further research confirms a desired effect of FGPE, any potential practical barriers must be removed for successful implementation of FGPE. Knowledge into why some patients drop out and why some patients does not have desired effect is of importance. There is uncertainty regarding whether the patient and their caregiver will opt for attending FGPE in all contexts. It might be challenging for caregivers to commit to attending the whole FGPE program and some cultures may not accept FGPE or achieve desired effect of FGPE. Timing the group sessions may prove difficult, because the participants might need to start treatment at different times. FGPE may not be a desirable treatment for all as some patients may not want to involve their family in their suffering, and some may see the family as part of their problem and would rather seek a therapist on their own. Further investigation on practical barriers and solutions to optimal utilization of FPE programs is recommended.

4.9 Implication for further research

There is a worldwide demand for effective, acceptable and available treatments for depression. FGPE has potential to be of great benefit for depressed patients and solid documentation is necessary before one can recommend the intervention on a large scale.

The substantial knowledge into the shared heritability of common mental disorders, published in Science in June 2018 (The Brainstorm Consortium, Antitila et al., 2018), supports the need for further investigation into family interventions. It also supports the demand for more knowledge on the effect of inviting a loved one (caregiver) into the treatment alliance.

Further research is needed to close the current knowledge gaps, below are some questions that we believe are in need of more research:

- What is the long-term effect of FGPE on MDD?
- Is FGPE beneficial for families who suffer from a familial predisposition to mental health disorders and can these be treated in multi-family groups?
- What are the practical obstacles to implementation of GPE and how can they be overcome?
- What effect does GPE/FGPE have in different societies and for different groups of patients?
- Can FGPE play a role in prevention of MDD and help prevent relapse?

Further knowledge into drug-free alternatives in the treatment of patients with MDD is of great importance in the Norwegian context as all regional health authorities are instructed by the minister of health to provide drug-free treatment option for psychiatric diseases (Brev fra Helse- og omsorgsdepartementet til de regionale helseforetakene, 2015). Treatment offered to the patient must be effective, well tolerated, safe and well documented. FGPE in conjunction with TAU show encouraging results. Research to establish more solid evidence for, or against, this intervention should be performed, preferably in a collaborative manner, including all Norwegian regional health authorities. Use of the intervention prior to further knowledge must be done with caution.

4.10 Conclusion

The effect of GPE in conjunction with treatment as usual without involving caregiver does not seem to have substantial benefits, when compared to treatment as usual and it is therefore unjustifiable to recommend this intervention on a large scale. Further investigation into GPE may provide new evidence.

FGPE involving participants (patients) and their caregiver seem to have beneficial effects on depression and on psychosocial functioning when provided as an adjunct to treatment as

usual. The evidence is based on a very small number of participants and the results for FGPE carry great uncertainties. If further knowledge on FGPE supports effectiveness suggested in this review, this intervention could be an important contributor to treatment of depression worldwide.

5 References

1. Aagaard, J., Foldager, L., Makki, A., Hansen, V. & Muller-Nielsen, K. (2017). The efficacy of psychoeducation on recurrent depression: a randomized trial with a 2-year follow-up. *Nordic Journal of Psychiatry*, 71 (3), 223-229.
2. Anderson, C.M., Gerard, E., Hogarty, G.E. & Reiss, D.J. (1980) Family treatment of adult schizophrenic patients: a psycho-educational approach. *Schizophr Bull*, 6 (3), 490-505.
3. Araya, R., Alvarado, R. & Minoletti, A. (2011). Depression in primary care: Chile's Programme for Screening, Diagnosis and Comprehensive Treatment of Depression. Internatioanl Psychiatry 8 (1), 6-7.
4. Brady, P., Kangas M. & McGill, K. (2017). "Family matters": A systematic review of the evidence for family psychoeducation for major depressive disorder. *Journal of marital and Family Therapy* 43 (2), 245-263. DOI:10.1111/jmft.12204.
5. Brev fra Helse- og omsorgsdepartementet til de regionale helseforetakene. 26.11.2015. Medikamentfri behandling for psykisk syke i alle helseregioner. Retrieved from 2017, March 24th from https://www.regjeringen.no/no/aktuelt/medikamentfri-behandling-for-psykisk-syke-i-alle-helseregioner/id2464240/
6. Berg, R., Øverland, S., Høie, B. Effekten av psykoterapi for mennesker med depressive lidelser, Kunnskapssenteret. Forskningsoversikt 2009. [Electronic version of scientific report]. ISBN : 978-82-8121-255-8, ISSN. Retrieved 2017, August 2nd, from https://www.fhi.no/publ/eldre/effekten-av-psykoterapi-for-mennesker-med-depressive-lidelser/

7. Casañas, R., Catalan, R, del Val, J.L, Real, J., Valero, S. & Casas, M. (2012). Effectiveness of a psycho-educational group program for major depression in primary care: a randomized controlled trial. *BMC Psychiatry*, 230 (12). DOI:10.1186/1471-244X-12-230.
8. Chetty, D. & Hoque, M. (2013). Effectiveness of a nurse facilitated cognitive group intervention among mild to moderately-depressed-women in KwaZulu-Natal, South Africa. African Journal of Psychiatry, 16, 29-34. DOI:10.4314/ajpsy.v16i1.5.
9. Cochrane CCDAN. (n.d.) Topic List: Intervention – Psychological therapies [internet document]. Retrieved 2017, August, 2nd, from http://cmd.cochrane.org/sites/cmd.cochrane.org/files/public/uploads/CCDAN%20topics%20list_psychological%20therapies%20for%20website.pdf
10. Cohen, S., O'Leary, K. D. & Foran, H. (2010). A randomized Clinical Trial of a Brief, Problem-Focused Couple Therapy for Depression. *Behavior Therapy,* 41 (4), 433-446. DOI:10.1016/j.beth.2009.11.004.
11. Colom, F. (2011). Keeping therapies simple: psychoeducation in the prevention of relapse in affective disorders. The British Journal of Psychiatry, 198 (5), 338-340. DOI: 10.1192/bjp.bp.110.090209.
12. Conradi, H.J., de Jonge, P. & Ormel, J. (2008). Cognitive-behavioural therapy v. usual care in recurrent depression. *Br J Psychiatry*, 193 (6). DOI: 10.1192/bjp.bp.107.042937.
13. Cuijpers, P., Munoz, R.F., Clarke, G. N. & Lewinsohn P, M. (2009). Psychoeducational treatment and prevention of depression: The "coping with depression" course thirty years later. *Clinical Psychology Review* 29, 449-458.

14. Dalgard, O.S. (2004). Kurs i mestring av depresjon – en randomisert, kontrollert studie. *Tidsskrift for norsk legeforening*, 124 (23), 3043-3046.
15. Dalgard, O.S. (2006). A randomized controlled trial of a psychoeducational group program for unipolar depression in adults in Norway (NCT00319540). *Clinical Practice and Epidemiology in Mental Health*, 15 (2). Doi:10.1186/1745-0179-2-15.
16. Dowrick, C., Dunn, G., Ayuso-Mateos, J.L., Dalgard, O.S., Page, H., Lehtinen, V.,…Wilkinson, G. (2000). Problem solving treatment and group psychoeducation for depression: multicenter randomized controlled trial. *BMJ*, 321, 1-6.
17. Dowrick, C. (2001). Psychosocial Interventions for Depression in Community Settings. *Z.f. Gesundheitswiss.*, 9 (3), 197-204.
18. Ferrari, A.J., Charlson, F.J., Norman, R.E., Patten, S.B., Freedman, G., Murray, C.J., Vos, T. & Whiteford, H.A. (2013). Burden of Depressive Disorders by Country, Sex, Age, and Year: Findings from the Global Burden of Disease Study 2010. *PLoS Med* 10 (11). DOI: e1001547.
19. Efthimious, K. & Psoma, M. (2012). Lewinsohn's cognitive behavioural group therapy course for depression: structure, application and efficacy. *Encephalos* 49, 60-66.
20. Ekhtiari, H., Rezapour, T., Aupperle, R.L. & Paulus, M.P. (2017). Chapter 10 - Neuroscience-informed psychoeducation for addiction medicine: A neurocognitive perspective. *Progress in Brain Research*, 235, 2017, Pages 239-264.
21. Frisk uten medikamenter (2018). Kompetansesenter for Brukererfaring og Tjenesteutvikling [internet document]. Retrieved 2019, march 2nd, from

http://kbtmidt.no/wp-content/uploads/2018/11/Frisk_uten_medikamenter_en_sammenstilling_og_vurdering_av_de_medikamentfrie_behandlingstilbudene_i_psykisk_helsevern_KBT.pdf

22. Fakta om depresjon (2015). Folkehelseinstituttet. [internet document]. Retrieved 2018, June 24th, from https://www.fhi.no/fp/psykiskhelse/psykiskelidelser/depresjon---faktaark/
23. Gold, L. H. (2014). DSM-5 and the Assessment of Functioning: The World Health Organization Disability Assessment Schedule 2.0 (WHODAS 2.0). *Journal of the American Academy of Psychiatry and the Law*. 42 (2) 173-181.
24. GRADEpro GDT. (2015). GRADEpro Guideline Development Tool [Software]. McMaster University, (developed by Evidence Prime, Inc.). Retrieved, 2018, juli 2nd, from https://gradepro.org/
25. Grunndokument (undated) Fellesaksjonen for medisinfrie behandlingsforløp i psykiske helsetjenester [Internet document] Retrieved 2019, March 2nd, from http://medisinfrietilbud.no/grunndokument/
26. Guyatt, G.H., Oxman, A.D., Vist, G.E., Kunz, R., Falck-Ytter, Y., Alonso-Coello, P., & Schünemann, H.J. (2008). GRADE: an emerging consensus on rating quality of evidence and strength of recommendations. BMJ. 26, 336(7650), 924–926. DOI: 10.1136/bmj.39489.470347.AD
27. Günadyın, N. & Barlas, G.Ü. (2017). Effectiveness of a group and brochure psychoeducation intervention to improve depression level and treatment

continuity among adults with depression in turkey: A controlled study. *Biomedical Research,* 28(4), 1493-1499.

28. Haringsma, R., Engels, G. I., Cuijpers & P., Spinhoven, P. (2006). Effectiveness of the Coping With Depression (CWD) course for older adults provided by the community-based mental health care system in the Netherlands: a randomized controlled field trial. *Int Psychogeriatr.18* (2), 307-25.
29. Herdman, M., Gudex, c., Lloyd, A., Janssen, M.F., Kind, P., Parkin, D., Bonsel, G. & Badia, X. (2011). Development and preliminary testing of the new five-level version of EQ-5D (EQ-5D-5L). *Quality of Life Research.* Dec; 20(10), 1727–1736. DOI: 10.1007/s11136-011-9903-x.
30. Higgins, J.P.T. & Green, S. (2011). Cochrane Handbook for Systematic Reviews of Interventions Version 5.1.0. [Internet publication]. Retrieved 2017, August 2nd from http://handbook.cochrane.org. http://handbook-5-1.cochrane.org
31. Holte, H. H., Austvoll-Dahlgren, A. & Straumann, G. H. (2017). *Psykososial behandling med eller uten bruk av antipsykotika for pasienter med aktiv psykose* (Report 2017). Oslo: Nasjonalt folkehelseinstitutt. Retrieved from: https://www.fhi.no/publ/2017/psykososial-behandling-med-eller-uten-bruk-av-antipsykotika-for-pasienter-m/
32. Iovieno, N., Papakostas, G.I. (2012). Correlation Between Different Levels of Placebo Response Rate and Clinical Trial Outcome in Major Depressive Disorder: A Meta-Analysis*. J Clin Psychiatry* 73(10):1300-1306.
33. Keitner, G.I, Ryan, C.E, Miller, I.W, Kohn, R., Bishop, D.S, Epstein, N.B. (1995). Role of the familiy in recovery and major depression. *Am J Psychiatry*, 152 (7), 1002-8.

34. Kumar, K. & Gupta, M. (2015). Effectiveness of Psycho-educational Intervention in Improving Outcome of Unipolar Depression: Results from a Randomised Clinical Trial. *East Asian Arch Psychiatry*, 25(1), 29-34.
35. Lara, M.A., Navarro, C., Rubi, N.A. & Mondragon, L. (2003). Two levels of intervention in low-income women with depressive symptoms: compliance and program assessment. International Journal of Social Psychiatry, 49 (1), 43-57. DOI: 10.1371/journal.pmed.1001547.
36. Lewinsohn, P. M., Hoberman, H., Teri, L., & Hautzinger, M. (1985). An integrative theory of unipolar depression. In Theoretical issues in behavior therapy. *Reiss S, Bootxin R (Eds.).* New York: Academic Press, 313-359.
37. Luciano, M., Del Vecchio, V., Giacco, D., De Rosa, C., Malangone, C. & Fiorillo, A. (2012). A'familiy affair'? The impact of family psychoeducational interventions on depression. *Expert Review of Neurotherapeutics*, 12 (1), 83-92. DOI:10.1586/ern.11.131.
38. McFarlane, W. R. (2016). Family Interventions for Schizophrenia and the Psychoses: A Review. *Fam. Process*. 55(3), 460–482. DOI:10.1111/famp.12235.
39. Motlova, L., Balon, R., Beresin, E.V, Brenner, A. M., Coverdale, J. H., Guerrero, A. & Weiss, L. R. (2017). Psychoeducation as an Opportunity for Patients, Psychiatrists, and Psychiatric Educators: Why Do We Ignore It?, *Academic Psychiatry,* 41 (4), 447–451.
40. Murray-Swank, A. B. & Dixon, L. (2005). Family Psychoeducation as an Evidence-Based Practice. *CNS Spectrums*, 9 (12), 905-912.
41. Njaa, B. (2018). Fra tvungen medisinering til trygg medisinfrihet. [lectures notes from Dialogkonferanse 20.-21.10 Oslo]. Retrieved from

http://medisinfrietilbud.no/wp-content/uploads/2018/11/PP-Bj%C3%B8rg-Njaa-20.11.2018.pdf .

42. Ouzzani, M., Hammady, H., Fedorowicz, Z., Elmagarmid, A. (2016). Rayyan — a web and mobile app for systematic reviews. *Systematic Reviews* 5:210. DOI: 10.1186/s13643-016-0384-4.
43. Posternak, M.A., Solomon, D.A., Leon, A.C., Mueller, T.I., Shea, M.T., Endicott, J. & Keller, M.B. (2006). The naturalistic course of unipolar major depression in the absence of somatic therapy. *The Journal of Nervous and Mental Disease*. 194(5), 324-9. DOI:10.1097/01.nmd.0000217820.33841.53
44. Psychoeducation: Definition, Goals and Methods. (2014). [Internet publication]. Aipc.net.au.Retrieved2019, March 2nd from (https://www.aipc.net.au/articles/psychoeducation-definition-goals-and-methods/)
45. Review Manager (2014) (RevMan) [Computer program]. Version 5.3. Copenhagen: The Nordic Cochrane Centre, The Cochrane Collaboration.
46. Røssberg J.I., Andreassen O.A., Ilner S.O. Medisinfrie sykehusposter – et kunnskapsløst tiltak. *Tidsskr Nor Legeforen* 2017; 137: 426 – 7.
47. Schünemann, H., Brożek, J., Guyatt, G. & Oxman, A., editors. (2013). GRADE handbook for grading quality of evidence and strength of recommendations [internet handbook]. Retrieved, 2018, juli 2nd, from http://gdt.guidelinedevelopment.org/app/handbook/handbook.html
48. Sharif, F., Nourian, K., Ashkani, H. & Zoladl, M. (2012). The effect of psycho-educational intervention on the life quality of major depressive patients referred

to hospitals affiliated to Shiraz University of Medical Sciences in Shiraz-Iran. *Iranian Journal of Nursing and Midwifery Research,* 17(6), 425-429.

49. Singh, R., Liles, M. A., Montagene, M. (2016). Identifying and describing patient perspectives on long-term antidepressant use*. Cogent Psychology*, 3(1), 1223902. https://doi.org/10.1080/23311908.2016.1223902.
50. Strik, J.J., Honig, A., Lousberg, R. & Denollet J. (2001). Sensitivity and specificity of observer and self-report questionnaires in major and minor depression following myocardial infarction. *Psychosomatics,* 42(5), 423-8.
51. The Brainstorm Consortium. Anttila et.al. (2018). Analysis of shared heritability in common disorders of the brain. *Science,* 360(6395), eaap8757. DOI: 10.1126/science.aap8757.
52. The National Institute of Mental Health. (2016). Mental Health Medication [Fact sheet]. Retrieved 2017, August 2nd, from https://www.nimh.nih.gov/health/topics/mental-health-medications/index.shtml
53. Tursi M.F., Baes, C., Camacho, F.R., Tofoli, S.M., Juruena, M.F. (2013). Effectiveness of psychoeducation for depression: A systematic review. *Australian & New Zealand Journal of Psychiatry* 47(11) 1019-1031. DOI: 10.1177/0004867413491154.
54. UNICEF. (2018). Child marriage [Fact sheet]. Retrieved 2018, 24th of July, from https://data.unicef.org/topic/child-protection/child-marriage/
55. U.S. Department of Health and Human Services. (2009). Family Psychoeducation: The Evidence [Electronic version of scientific report]. (p. 904). Retrieved, 2017,

September 3rd from https://store.samhsa.gov/shin/content/SMA09-4423/TheEvidence-FP.pdf

56. Vaaler A.E., Fasmer O. B., (2013) Antidepressive legemidler – klinisk praksis må endres. *Tidsskr Nor Legeforen* 2013;133: 428-30. Retrieved 2019, March 23rd from https://tidsskriftet.no/2013/02/kronikk/antidepressive-legemidler-klinisk-praksis-ma-endres/kommentarer.
57. Wei, Y., McGrath, P. J., Hayden, j. & Kutcher, S. (2016). Measurement properties of tools measuring mental health knowledge: a systematic review. *BMC Psychiatry.* 16(1), 297. DOI: 10.1186/s12888-016-1012-5.
58. WHO. (2016). Investing in treatment for depression and anxiety leads to fourfold return [Press release]. Retrieved 2018, August 5th, from http://www.who.int/en/news-room/detail/13-
59. WHO. (2016). ICD-10. [Diagnostic Classification]. Retrieved 2017, August 2nd, from http://apps.who.int/classifications/icd10/browse/2016/en#/F32.2.04-2016-investing-in-treatment-for-depression-and-anxiety-leads-to-fourfold-return
60. WHO. (2018). Depression [Fact sheet]. Retrieved 2018, August 3rd, from http://www.who.int/news-room/fact-sheets/detail/depression
61. Xia, J., Merinder, L.B.& Belgamwar, MR. (2011). Psychoeducation for schizophrenia. *Cochrane Database of Systematic Reviews*, 6. Art. No.: CD002831. DOI:10.1002/14651858.CD002831.pub2.
62. Zhao, S., Sampson, S., Xia J., Jayaram, M.B. (2015). Psychoeducation (brief) for people with serious mental illness, *Cochrane Database of Systematic Reviews*, issue 4. Art. No.: CD010823. DOI: 10.1002/14651858.CD010823.pub2.

6 Appendices

6.1 Appendix 1: Protocol published in PROSPERO

NHS National Institute for Health Research

PROSPERO

International prospective register of systematic reviews

Print | PDF

Effect of group psychoeducation for major depressive disorder compared with pharmacological and/or other psychological treatment: a systematic review

Helene Sandberg, Åshild Roaldset

Citation

Helene Sandberg, Åshild Roaldset. Effect of group psychoeducation for major depressive disorder compared with pharmacological and/or other psychological treatment: a systematic review. PROSPERO 2017 CRD42017077110 Available from: http://www.crd.york.ac.uk/PROSPERO/display_record.php?ID=CRD42017077110

Review question

Is group psychoeducation effective in improving quality of life in adults with major depressive disorder compared to pharmacological treatment and/or other psychological treatment?

Searches

The following databases will be searched: MEDLINE, EMBASE, The Cochrane Library (Central), PsycINFO, PubMed, CINAHL. Other databases will be considered in discussion with a search librarian. The reference lists of systematic reviews, literature reviews, and other relevant publications will also be checked manually to identify any relevant studies not covered by the database searches. Databases will be searched from the year 2000, because consensus about the critical elements of family psychoeducation was developed in 1999 (Substance Abuse and Mental Health Services Administration 2009).

Types of study to be included

Types of study to be included: randomized controlled trials (RCTs) and non-randomized controlled trials (non-RCTs), controlled before-and-after studies (CBAs), interrupted time series (ITS), prospective and retrospective cohort studies with a control group are eligible for inclusion. In the event that several high-quality RCTs and non-RCTs are included, we will consider not including other study designs. Cluster RCTs analysed on an individual level must be adjusted for intra cluster correlation (ICC).

Condition or domain being studied

Major depressive disorder (MDD – clinical depression), ICD10, F32.2, F32.3, F33.2, F33.3, F34.0, F34.1 (ICD10, WHO, 2013) is a common mental disorder that occurs in all ages worldwide. MDD is an episodic disorder with a chronic or long-term outcome and increased risk of death. There are more than 300 million people of all ages with depressive disorder globally (WHO 2017). According to WHO (WHO 2017) depression is a leading cause of disease burden and accounted for 4.3% of the global burden of disease in 2010 (Ferrari, 2013). Depression is one of the largest single causes of disability worldwide. MDD and Dysthymia accounted for 2.5% of global disability- adjusted life years (DALYS), a measure of reduced health year, and it is also associated with lower work productivity, suicide and ischemic heart disease (Ferrari, 2013). According to WHO, persons with MDD and schizophrenia have a 40-60% greater chance of dying prematurely compared to the general population. Close to 800,000 people die of suicide every year due to all causes (WHO 2017).

Reduction in healthy life years and the associated health issues due to depression has not only an impact on the affected persons and their families but also the economy worldwide.

There are many non-medical interventions for depression disorder. Cochrane Common Mental Disorder lists 87 different psychological therapies (Cochrane CCDAN n.d.). According to the National Institute of Mental Health, common antidepressant medications (National institute of mental health, Mental Health Medication) are; Selective Serotonin Reuptake Inhibitor (SSRIs), serotonin and norepinephrine reuptake inhibitors (SNRIs), Bupropion, Tricyclic Antidepressants (TCAs), Tetracyclics Antidepressant, and monoamine oxidase inhibitors (MAOIs).

Psychoeducation is a didactic program aiming to give the participants sound knowledge of the condition and learn how to accept it and cope with it successfully. The intervention can be used for many different conditions and will often involve caregivers such as family and friends. The intervention is an independent therapy and is often based on cognitive behavior principles. The psychoeducation program can include practical tasks, making the participants practice skills such as self-assertiveness, communication and problem solving. The participants are empowered through the program and they are not considered to be patients in the psychoeducation setting (Motlova 2017).

The effect of medication and psychological therapies for depression are continuously studied and disputed (Nasjonalt kunnskapssenter for helsetjenesten, 2009, p 16). Psychoeducation has proven to be beneficial to patients suffering from psychosis and bipolar affective disorder, but there is currently a knowledge gap regarding psychoeducation and major depressive disorder (McFarlane 2016).

Participants/population

Patients with MDD, older than 18 years. For inclusion purposes, we will look at study descriptions of participants rather than clinical codes (ICD10). We will exclude studies with more than 50% of patients with medical comorbidities (e.g. cancer, diabetes). Furthermore, we will exclude studies with more than 50% of patients with perinatal depression, bipolar affective disorder and mental impairment, including dementia.

Intervention(s), exposure(s)

" Psychoeducation could be defined as a patient’s empowering training targeted at promoting awareness and proactivity, providing tools to manage, cope and live with a chronic condition (i.e. adherence enhancement, early warning sign identification, lifestyle, crisis management,

communication), and changing behaviors and attitudes related to the condition. Psychoeducation replaces guilt by responsibility, helplessness by proactive care and denial by awareness." (Colom 2011).

We will include group psychoeducation that has the following characteristics (McFarlane 2016):

•is provided to groups of patients or groups consisting of one or more patient with their care givers (family members, friends or other)

•is provided by a health care professional

•includes exercises in practical skills relevant for the patient and family group such as coping strategies, behavior, communication, social interactions and problem solving

•aims at giving knowledge on depression and treatment of depression

•is specific for patients with depression

We will include all durations (doses) of group psychoeducation as long as it is minimum one 60 minute session per week.

Comparator(s)/control

We will include the following comparisons:

1.Group psychoeducation compared with antidepressant medications. We will accept the antidepressant medications listed above.

2.Group psychoeducation plus antidepressant medications (as listed above) compared with antidepressant medications only.

3.Group psychoeducation compared with other psychological intervention (psychological interventions in the Cochrane CCDAN lists of 87 interventions will be included).

4.Group psychoeducation plus other psychological intervention compared to other psychological intervention (as listed above).

5.Group psychoeducation plus antidepressant medications compared to other psychological intervention (as listed above) and/or antidepressant medications (as listed above).

Context

Primary outcome(s)

Patients' quality of life, mortality (suicide), depression severity.

Secondary outcome(s)

Level of psychosocial functioning, relapses, treatment adherence.

Data extraction (selection and coding)

The search result will be screened using Rayyan software (Mourad et al 2016) by two independent reviewers. Each reviewer will screen the abstracts for PICO; followed by full text reading when necessary. Any differences between the two reviewer's evaluation of articles for inclusion, will be discussed and the publication will be inspected, until consensus is achieved.

The following core data will be extracted from all included studies:

• Title, authors and other publication details

• Study design and aim

• Setting (place and time of recruitment/data collection)

• Sample characteristics (age, gender, etc.)

• Intervention characteristics (duration/dose, provider, content, etc)?

• Methods of outcome measurement (instruments/ tools)

• Results related to the outcomes

Risk of bias (quality) assessment

Two reviewers will first independently and then together assess the risk of bias. For the RCTs included we will assess the risk of bias according to the criteria in the Cochrane Handbook for Systematic Reviews of Interventions (Higgins and Green 2011). The following processes will be studied to assess risk of bias: sequence generation; allocation concealment; blinding of participants and personnel; blinding of outcome assessment; incomplete outcome data; selective outcome reporting; and other sources of bias. The quality of evidence for all processes will be reported as 'Low Risk', 'Unclear Risk', or 'High Risk'. When there is no cause for concern the procedure will be considered to have low risk of bias. When there may be a risk of bias, but there is either insufficient information to assess whether an important risk of bias exists or there is insufficient rationale or evidence that an identified problem will introduce bias, we will assign unclear risk of bias. Procedures with cause for concern will be assigned high risk of bias.

For other study-designs than RCTs included, we will use Cochrane EPOC (Effective Practice and Organization of Care Group) checklist. The same assessment procedure as for RCTS will be used. In the event of disagreement between the two reviewers, the supervisor and co-supervisor will be involved to find a solution.

Strategy for data synthesis

We will conduct meta-analyses when possible. Data will be summarized and presented narratively in text and tables for each comparison. For continuous data, mean difference or standardized mean difference and 95% confidence intervals (95% CI) will be used to calculate effect sizes by using the Revman 2014 software. We will analyse dichotomous data as risk ratios and 95% confidence intervals (95% CI). Revman will be used to pool data (meta-analysis) when we have two or more studies reporting the same PICO (Population, Intervention, Comparison, Outcome), and to generate forest plots to display the results. No meta-analyses will be made in the case of diverse methodologies or

unclear therapeutic approach making comparisons difficult. For the primary outcomes we will assess the certainty of the evidence by using GRADE. Heterogeneity will be considered. We define heterogeneity to be when there is great variation in results, non-overlapping CIs, P<0,01 and I > 50% (Higgins and Green 2011). We will examine causes for and attempt to explain heterogeneity.

Analysis of subgroups or subsets

1. Effect of group composition; group consisting of patients only or patients together with their family members or other caregiver

2. Dose effect of the psychoeducation therapy; 12 sessions (of 1-2 hours duration every week) or less will be considered low dose, 13-52 sessions (of 1-2 hours duration every week) will be considered moderate duration, and 52 sessions and above (1-2 hours duration every week) will be considered high dose (Xia et al 2011).

Contact details for further information

Helene Sandberg

sandberg.helene@gmail.com

Organisational affiliation of the review

The Arctic University of Norway (UiT) uit.no

Review team members and their organisational affiliations

Ms Helene Sandberg. UiT
Ms Åshild Roaldset. UiT

Collaborators

Lien Nguyen. Norwegian Institute of Public Health

Anticipated or actual start date

01 September 2017

Anticipated completion date

15 May 2018

Funding sources/sponsors

None

Conflicts of interest

None known

Language

English

Country

Norway

Stage of review

Review_Ongoing

Subject index terms status
Subject indexing assigned by CRD

Subject index terms
Bipolar Disorder; Depressive Disorder, Major; Humans; Psychotherapy

Date of registration in PROSPERO
30 October 2017

Date of publication of this version
15 September 2017

Details of any existing review of the same topic by the same authors

Stage of review at time of this submission

Stage	Started	Completed
Preliminary searches	Yes	No
Piloting of the study selection process	No	No
Formal screening of search results against eligibility criteria	No	No
Data extraction	No	No
Risk of bias (quality) assessment	No	No
Data analysis	No	No

Versions

15 September 2017

6.2 Appendix 2: Search strategy in electronic databases

Database: Ovid MEDLINE(R) Epub Ahead of Print, In-Process & Other Non-Indexed

Citations, Ovid

MEDLINE(R) Daily and Ovid MEDLINE(R) 1946 to Present

Date for search: 19.09.2017

#	Searches	**Results**
1	Mood disorders/	13360
2	Depressive disorder, major/	26064
3	Depressive disorder/	69663
4	Dysthymic disorder/	1135
5	Depression/	101227
6	Cyclothymic Disorder/	636
7	(depress* or dysthym* or ((affective or mood) adj disorder*) or cyclothym*).ti,ab,kf.	427181
8	or/1-7	469371
9	(psychoeducat* or psycho-educat*).mp.	5307
10	((famil* or group?) adj2 intervention*).ti.	1982
11	8 and (9 or 10)	1780
12	Randomized Controlled Trial/	481824
13	Non-randomized controlled trials as topic/	231
14	Controlled Clinical Trial/	96877
15	Controlled Before-After Studies/	284
16	Multicenter Study/	239806
17	Pragmatic Clinical Trial/	674
18	Interrupted Time Series Analysis/	333
19	(random* or trial or intervention? or effect? or impact? or multicenter or multi center or multicentre or multi centre or controlled or control group? or (before adj5 after) or (pre adj5 post) or ((pretest or pre test) and (posttest or post test)) or quasiexperiment* or quasi experiment* or evaluat* or time series or time point? or repeated measur*).ti,ab.	8887916
20	Meta-Analysis/	86965
21	Meta-Analysis as Topic/	16461
22	(((systematic* or literature) adj3 (overview or review* or search*)) or meta-anal* or metaanal* or meta-regression* or umbrella review* or overview of reviews or review of reviews or (evidence* adj2 synth*) or synthesis review*).ti,ab.	453526
23	Review.pt. and (pubmed or medline).ti,ab.	103522
24	or/12-23	9230454

25	11 and 24	1631
26	limit 25 to yr="2000-current"	1505
27	exp Animals/	21982274
28	Humans/	17397204
29	27 not (27 and 28)	4585070
30	(news or editorial or comment).pt.	1197945
31	26 not (29 or 30)	1498
32	remove duplicates from 31	1378

PsycINFO 1806 to September Week 2 2017

Date for search: 19.09.2017

#	Searches	Results
1	affective disorders/	13026
2	major depression/	107206
3	endogenous depression/	1237
4	reactive depression/	298
5	recurrent depression/	735
6	atypical depression/	188
7	dysthymic disorder/	1450
8	"depression (emotion)"/	23856
9	cyclothymic personality/	208
10	(depress* or dysthym* or ((affective or mood) adj disorder*) or cyclothym*).ti,ab,id.	282209
11	or/1-10	286205
12	(psychoeducat* or psycho-educat*).mp.	11217
13	((famil* or group?) adj2 intervention*).ti.	2785
14	11 and (12 or 13)	2153
15	("0400" or "0451" or "1800" or "2100").md. [empirical study/ prospective study/ quantitative study/ treatment outcome/]	2197689
16	Experimental Design/	10593
17	Between Groups Design/	108
18	Quantitative Methods/	2966
19	Quasi Experimental Methods/	143
20	Experiment Controls/	888
21	Pretesting/	236
22	Posttesting/	135
23	Time Series/	1830
24	Repeated Measures/	644
25	(random* or trial or intervention? or effect* or impact? or multicenter or multi center or multicentre or multi centre or controlled or control group? or (before adj5 after) or (pre adj5 post) or ((pretest or pre test) and (posttest or post test)) or quasiexperiment* or quasi experiment* or evaluat* or time series or time point? or repeated measur*).ti,ab.	1913833
26	Meta Analysis/	4048
27	Systematic Review.md.	17378
28	(((systematic* or literature) adj3 (overview or review* or search*)) or meta-anal* or metaanal* or meta-regression* or umbrella review* or overview of reviews or review of reviews or (evidence* adj2 synth*) or synthesis review*).ti,ab.	107679
29	(review and (pubmed or medline)).ti,ab.	12823
30	or/15-29	2968383

31	14 and 30	2024
32	limit 31 to yr="2000-current"	1719
	remove duplicates from 32	1715

Embase 1974 to 2017 September 18

Date for search: 19.09.2017

#	Searches	Results
1	*mood disorder/	7662
2	*depression/	131918
3	*Major affective disorder/	90
4	*Schizo affective psychosis/	2238
5	*Dysthymia/	2225
6	*Endogenous depression/	776
7	*Involutional depression/	147
8	*Treatment resistant depression/	962
9	*cyclothymia/	160
10	*Major depression/	23666
11	(depress* or dysthym* or ((affective or mood) adj disorder*) or cyclothym*).ti,ab,kw	544762
12	or/1-11	568581
13	(psychoeducat* or psycho-educat*).mp.	10750
14	((famil* or group?) adj2 intervention*).ti.	2393
15	12 and (13 or 14)	3254
16	Randomized Controlled Trial/	472907
17	Controlled Clinical Trial/	449115
18	Quasi Experimental Study/	4037
19	Pretest Posttest Control Group Design/	320
20	Time Series Analysis/	20124
21	Experimental Design/	14803
22	Multicenter Study/	165718
23	Pretest Posttest Design/	2143
24	(random* or trial or intervention? or effect* or impact? or multicenter or multi center or multicentre or multi centre or controlled or control group? or (before adj5 after) or (pre adj5 post) or ((pretest or pre test) and (posttest or post test)) or quasiexperiment* or quasi experiment* or evaluat* or time series or time point? or repeated measur*).ti,ab.	11743356
25	Meta Analysis/	133938
26	Systematic Review/	149802
27	(((systematic* or literature) adj3 (overview or review* or search*)) or meta-anal* or metaanal* or meta-regression* or umbrella review* or overview of reviews or review of reviews or (evidence* adj2 synth*) or synthesis review*).ti,ab.	535563
28	(review and (pubmed or medline)).ti,ab.	123202
29	or/16-28	12100334

30	15 and 29	2863
31	limit 30 to yr="2000-current"	2696
32	exp animals/ or exp invertebrate/ or animal experiment/ or animal model/ or animal tissue/ or animal cell/ or nonhuman/	25250688
33	human/ or normal human/ or human cell/	19014640
34	32 not (32 and 33)	6283014
35	(news or editorial or comment).pt.	547024
36	31 not (34 or 35)	2692
37	limit 36 to embase	1757
38	remove duplicates from 37	1658

Cochrane Library (CDSR, DARE, CENTRAL, HTA)

Date for search: 19.09.2017

ID	**Search Hits**	
#1	[mh ^"Mood disorders"]	564
#2	[mh ^"Depressive disorder, major"]	3229
#3	[mh ^"Depressive disorder"]	5216
#4	[mh ^"Dysthymic disorder"]	156
#5	[mh ^Depression]	7307
#6	[mh ^"Cyclothymic Disorder"]	14
#7	(depress* or dysthym* or ((affective or mood) next disorder*) or cyclothym*):ti,ab,kw	55136
#8	{or #1-#7}	55136
#9	(psychoeducat* or psycho-educat*):ti,ab,kw	2136
#10	((famil* or group or groups) near/2 intervention*) .ti.	3678
#11	#8 and (#9 or #10) Publication Year from 2000 to 2017, in Cochrane Reviews (Reviews only) and Trials	1006

#12	(depress* or dysthym* or ((affective or mood) next disorder*) or cyclothym*)	75733
#13	{or #1-#6, #12}	75733
#14	(psychoeducat* or psycho-educat*)	2615
#15	((famil* or group or groups) near/2 intervention*) .ti	3678
#16	#13 and (#14 or #15) Publication Year from 2000 to 2017, in Cochrane Reviews (Protocols only), Other Reviews and Technology Assessments	262
#17	#11 or #16	1268

CINAHL (EBSCO)

Date for search: 19.09.2017

#	Query Limiters/Expanders Last run via	**Results**
S17	S11 AND S16 Limiters - Exclude MEDLINE records; Published Date: 20000101-20170931	215
S16	S12 OR S13 OR S14 OR S15	1,007,977
S15	TI (((systematic* or literature) N2 (overview or review* or search*)) or meta-anal* or metaanal* or meta-regression* or umbrella-review* or "overview of reviews" or "review of reviews" or (evidence* N1 synth*) or synthesis-review*)) OR AB (((systematic* or literature) N2 (overview or review* or search*)) or meta-anal* or metaanal* or meta-regression* or umbrella-review* or "overview of reviews" or "review of reviews" or (evidence* N1 synth*) or synthesis-review*)	90,561
S14	TI ((random* or trial or intervention# or effect* or impact# or multicenter or multi-center or multicentre or multi-centre or controlled or control group# or (before N4 after) or (pre N4 post) or ((pretest or pre-test) and (posttest or post-test)) or quasiexperiment* or quasi-experiment* or evaluat* or time-series or time point# or repeated-measur*) OR AB ((random* or trial or intervention# or effect* or impact# or multicenter or multi-center or multicentre or multi-centre or controlled or control group# or (before N4 after) or (pre N4 post) or ((pretest or pre-test) and (posttest or post-test)) or quasiexperiment* or quasi-experiment* or evaluat* or time-series or time point# or repeated-measur*))	19,921
S13	TI ((random* or trial or intervention# or effect* or impact# or multicenter or multi-center or multicentre or multi-centre or controlled or control group# or (before N4 after) or (pre N4 post) or ((pretest or pre-test) and (posttest or post-test)) or quasiexperiment* or quasi-experiment* or evaluat* or time-series or time point# or repeated-measur*) OR AB ((random* or trial or intervention# or effect* or impact# or multicenter or multi-center or multicentre or multi-centre or controlled or control group# or (before N4 after) or (pre N4 post) or ((pretest or pre-test) and (posttest or post-test)) or	950,412

	quasiexperiment* or quasi-experiment* or evaluat* or time-series or time point# or repeated-measur*))	
S12	(DE "Randomized Controlled Trials" OR DE "Pretests Posttests" OR DE "Control Groups" OR DE "Evaluation Research" OR DE "Quasiexperimental Design" OR DE "Program Validation" OR DE "Program Effectiveness" OR DE "Program Evaluation" OR DE "Outcomes of Treatment")	86,174
S11	S7 AND S10 775	775
S10	S8 OR S9	3,928
S9	TI ((famil* or group or groups) N1 intervention*)	1,094
S8S7	TI ((psychoeducat* or psycho-educat*)) OR AB ((psychoeducat* or psycho-educat*)) OR SU ((psychoeducat* or psycho-educat*))	2,948
S7	S1 OR S2 OR S3 OR S4 OR S5 OR S6	91,034
S6	TI ((depress* or dysthym* or ((affective or mood) W0 disorder*) or mental* ill* or cyclothym*)) OR AB ((depress* or dysthym* or ((affective or mood) W0 disorder*) or mental* ill* or cyclothym*)) OR SU ((depress* or dysthym* or ((affective or mood) W0 disorder*) or mental* ill* or cyclothym*))	91,034
S5	(MH "Affective Disorders, Psychotic")	346
S4	(MH "Cyclothymic Disorder")	9
S3	(MH "Dysthymic Disorder")	169
S2	(MH "Depression")	52,312

S1	(MH "Affective Disorders")	3,036

Epistemonikos

Date for search: 19.09.2017

(title:(psychoeducat* OR psycho-educat* OR group-intervention* OR family-intervention*) OR abstract:(psychoeducat* OR psycho-educat* OR group-intervention* OR family-intervention*)) AND (title:(depress* OR dysthym* OR affective-disorder* OR mood-disorder* OR cyclothym*) OR abstract:(depress* OR dysthym* OR affective-disorder* OR mood-disorder* OR cyclothym*))

6.3 Appendix 3: Excluded studies read in full text

41 excluded studies after full text reading.

Study	Reason for exclusion
Allart-Van Dam, E., Hosman, C. M. H., Hoogduin, C. A. L. & Schaap, C. P.D.R. (2003). The coping with depression course: Short-term outcomes - subclinical depression. Behavior Therapy, 34 (3), 381-396 doi.org/10.1016/S0005-7894(03)80007-2	Excluded depressed patients
Alvarado, R., Rojas, G., Minoletti, A., Alvarado, F. & Domínguez, C. Ruben Alvardo Depression Program in Primary Health Depression Program in Primary Health Care, The Chilean Experience. (2012). International Journal of Mental Health, 41(1), 2012, 38-47. doi.org/10.2753/IMH0020-7411410103	Patients have comorbidity
Bersani, F. S., Biondi, M., Coviello, M, Fagiolini, A., Majorana, M., Minichino, A., Rusconi, A. C., Vergani, L., Vicianza, R. & Coccanari de' Fornari, M. A. (2017). Psychoeducational intervention focused on healthy living improves psychopathological severity and lifestyle quality in psychiatric patients: Preliminary findings from a controlled study. Journal of Mental Health, 26 (3), 271-275, DOI:10.1080/09638237.2017.1294741.	Mixed diagnosis
Brown, J. S. L., Elliott, S.A., Boardman, J., Ferns, J. & Morrison, J. (2004). Meeting the unmet need for depression services with psycho-educational self-confidence workshops: prelimenary report. British Jounal of Psychiatry, 185, 511-515.	Comorbidity (anxiety)
Brown, J. S. L., Elliott, S.A., Boardman, J., Andiappan, M., Landau, S. & Howay, E. (2008). Can the effects of a 1-day CBT psychoeducational workshop on self-confidence be maintained after 2 years? A naturalistic study. Depression and Anxiety, 25, 632-640.	Comorbidity (anxiety)
Canasas, R., Catalan, R., Penades, R., Real, J., Valero, S., Munoz, M.A., Lalucat-Jo, I.L. & Casas, M. (2014). Evaluation of the Effectiveness of a Psychoeducational Intervention in treatment-Naïve Patients with Antidepressant Medication in Primary Care: A Randomized Controlled Trial. The Scientific World Journal. Article ID 718607.	Non-randomized study

Chiesa, A., Mandelli, L., Serretti, A. Alberto Chiesa. (2012). Mindfulness-based cognitive therapy versus psycho-education for patients with major depression who did not achieve remission following antidepressant treatment: a preliminary analysis. J Altern Complement Med., 18(8), 756-60. doi: 10.1089/acm.2011.0407	Wrong comparison
Conradi, H.J., de Jonge, P. & Ormel, J. (2008). Cognitive-behavioural therapy v. usual care in recurrent depression. Br J Psychiatry, 193 (6). doi: 10.1192/bjp.bp.107.042937.	Not group intervention.
Cramer, H., Salisbury, C., Conrad, J., Eldred, J. & Araya R. (2011). Group cognitive behavioural therapy for women with depression:pilot and feasibility study for a randomised controlled trial using mixed methods.. BMC Psychiatry, 11 (82).	Pilot study without results
Delgadillo, J., McMillan, D., Lucock, M., Leach, C., Ali, S.& Gilbody, S. (2014). Early changes, attrition, and dose–response in low intensity psychological interventions. British Journal of Clinical Psychology,53, 114-130. doi.org/10.1111/bjc.12031	A large proportion of the patients have anxiety as comorbidity
Dunn, E., Rogers, E.S., Hutchinson, D.S., Lyass, A., MacDonald, K.L., Wallace, L.R. & Furlong-Norman, K. (2008). Results of an Innovative University-based Recovery Education Program for Adults with Psychiatric Disabilites. Adm Policy Ment Health, 35, 357-369. DOI 10.1007/s10488-008-0176-9	Wrong intervention
Fiorillo, A., Malangone, C., Vecchio, V., Rosa, C., Luciano, M., Giacco, D., Sampogna, G., Gaudio, L. & Maj, M. (2011). The effect of family psychoeducational interventions on patients with depression. European psychiatry [abstracts from the 19th European congress of psychiatry, EPA 2011 MAR 12-15; Vienna, Austria]	Congress abstract
Franchini, L., Bongiorno, F., Spagnolo, C., Florita, M., Santoro, A., Dotoli, D. & Barbini, B. (2006). Smeraldi, E. Psychoeducational group intervention in addition to antidepressant therapy as relapse preventive strategy in unipolar patients. Clinical Neuropsychiatry, 3 (4), 282-285.	Patients in remission
Haringsma, R., Engels, G. I., Cuijpers & P., Spinhoven, P. (2006). Effectiveness of the Coping With Depression (CWD) course for older adults provided by the community-based mental health care system in the Netherlands: a randomized controlled field trial. Int Psychogeriatr.18 (2), 307-25.	Comorbidities

Hundt, N. E., Calleo, J. S., Williams, W. & Cully, J. A. (2016) Does using cognitive-behavioral therapy skills predict improvements in depression. Psychology and psychotherapy, 89, 2, 235-238.	Not RCT
Kellet, S., Clarke, S. & Matthews, L. (2007). Delivering group psychoeducational CBT in Primary Care: Comparing outcomes with individual CBT and individual psychodynamic-interpersonal psychotherapy. British Journal of Clinical Psychology, 46 (2), 211-222.	Comorbidities: anxiety PTSD, OCD, anger, pain
Kiermeir, J., Gassner, L-M., Siebörger, A., Wiethoff K., Ricken H., Stamm, T., Bauer, M., Heinz, A. & Adli, M. (2012). Euthymic Therapy to Reduce Residual Symptoms of Depression and Stengthen Self-Care A Randomised Controlled Trial. German J Psychiatry, 5 (1), 15-22.	Majority of patients in remission
Lara, M.A., Navarro, C., Rubi, N.A. & Mondragon, L. (2003). Two levels of intervention in low-income women with depressive symptoms: compliance and programme assessment. International Journal of Social Psychiatry, 49 (1), 43-57.	Outcome clinical condition not included.
Lucksted, A., Medoff, D:, Burland, J., Stewart, B., Fang, L.J., Brown, C., Jones, A., Lehman, A. & Dixon, L.B. (2013). Sustained outcomes of a peer-taught family education program on mental illness. Acta Psychiatr Scand. 127 (4), 279–286. doi: 10.1111/j.1600-0447.2012.01901.x	Intervention only involving family members
Macrodimitris, S. D., Backs-Dermott B. J., Hamilton, K. E. & Mothersill, K. J. (2010). CBT Basics: A group Approach to teaching fundamental Cognitive-Behavioural Skills. Journal of Cognitive Psychotherapy, 24 (2), 132-146.	Not RCT
Melo-Carillo, A., Van Oudenhove, L. & Lopez-Avila, A. (2012) Depressive symptoms among Mexican medical students: high prevalence and the effect of a group psychoeducation intervention. J Affect Disord., 136 (3). doi: 10.1016/j.jad.2011.10.040.	Wrong diagnosis
Morokuma, I., Shimodera, S., Fujita, H., Hashizume, H., Kamimura, N., Kawamura, A., Nishida, A., Furukawa, T.A., & Inoue, S. (2013). Psychoeducation for major depressive disorders: a randomised controlled trial. Psychiatry Research, 210 (1). doi: 10.1016/j.psychres.2013.05.01	Patients in partial remission
Naismith, S. L., Diamond, K., Carter, P.E., Norrie, L.M., Redoblado-Hodge, M.A., Lewis, S.J. & Hickie, I.B. (2011). Enhancing memory in late-life depression: the effects of a combined psychoeducation and	Wrong Intervention

cognitive training program. Am J Geriatr Psychiatry, 19(3), 240-8. doi: 10.1097/JGP.0b013e3181dba587.	
Pradeep, J., Isaacs, A., Shanbag, D., Selvan, S. & Srinivasan, K. (2014). Enhanced care by community health workers in improving treatment adherence to antidepressant medication in rural women with major depression. Indian J Med Res., 139 (2), 236-45.	Comorbidities (anxiety)
Renner, W. & Berry, J. W. (2011). The ineffectiveness of Group Interventions for female Turkish migrants with recurrent depression. Social Behavior and Personality, **39** (9), 1217-1234.	Comorbidities
Rokke, P.D., Tomhave, J.A. & Jocic, Z. (2000). The role of client choice and target selection in self-management therapy for depression in older adults. Cognitive Therapy and Research. 24 (99). https://doi.org/10.1023/A:1005407125992	Comorbidity
Saleid, G., A., Czaikowski, N.O., Holte, A., Tambs, K. & Aarø, L.E. (2016). Coping With Strain (CWS) course - its effects on depressive symptoms: A four-year longitudinal randomized controlled trial. Scand J Psychol., 57(4). doi: 10.1111/sjop.12289.	Wrong intervention
Saloheimo, H.P., Markowitz, J., Saloheimo, T.H., Laitinen, J.J., Sundell, J., Huttunen, M.H., Aro, T.A., Mikkonen, T.N. & Katila, H.O. (2016). Psychotherapy effectiveness for major depression: A randomized trial in a Finnish community. BMC Psychiatry, 16 (131). doi: 10.1186/s12888-016-0838-1	More than 50% comorbidity (anxiety)
Schimmel-Spreeuw, A., Linssen, A.C.G. & Heeren, T.J.(2000). Coping With Depression and Anxiety: Preliminary Results of a Standardized Course for Elderly Depressed Women. International Psychogeriatrics, 12 (1),77-86.	Comorbidity
Schuster, R., Leitner, I., Carlbring, P. & Laireiter, A-R. (2017). Exploring blended group interventions for depression: Randomised controlled feasibility study of a blended computer- and multimedia-supported psychoeducational group intervention for adults with depressive symptoms. Internet Interventions, 8, 63-71. doi.org/10.1016/j.invent.2017.04.001	Patients only at risk of MDD
Seedat, S., Haskis, A. & Stein, D.J. (2008). Benefits of a consumer psychoeducation: a pilot program in South Africa. Int'L.J. Psychiatry in Medicine, 38 (1), 31-42.	Intervention is newsletters only

Shimodera, S., Furukawa, T. A., Mino, Y., Shimazu, K., Nishida, A., & Inoue, S. (2012). Cost-effectiveness of family psychoeducation to prevent relapse in major depression: Results from a randomized controlled trial. BMC Psychiatry, 12(40). doi.org/10.1186/1471-244X-12-40	Patients in partial remission
Silverman, M. J. (2013). Effects of group songwriting on depression and quality of life in acute psychiatric inpatients: A randomized three group effectiveness study. Nordic Journal of Music Therapy, 22 (2), 131-148 doi.org/10.1080/08098131.2012.709268	Patients had a variety of diagnosis, not only MDD
Solati, K. (2016). Effectiveness of Cognitive-Behavior Group Therapy, Psycho-education Family, and drug Therapy in Reducing and Preventing Recurrence of Symptoms in Patients with Major Depressive Disorder. Journal of Chemical and Pharmaceutical Sciences, 9 (4).	Includes family members only
Stangier, U., Hilling, C., Heidenreich, T., Risch, A. K., Barocka, A., Schlösser, R., Kronfeld, K., Ruckes, C., Berger, H., Röschke, J., Weck, F., Volk, S., Hambrecht, M., Serfling, R., Erkwoh, R., Stirn, A., Sobanski, T., Hautzinger, M. (2013). Maintenance Cognitive-Behavioral Therapy and Manualized Psychoeducation in the Treatment of Recurrent Depression: A Multicenter Prospective Randomized Controlled Trial. Am J Psychiatry, 170 (6). doi: 10.1176/appi.ajp.2013.12060734.	Patients were in remission
Swan, J., Sorrell, E., MacVicar, B., Durham, R. & Matthews, K. (2004). Coping with depression": an open study of the efficacy of a group psychoeducational intervention in chronic, treatment-refractory depression. Journal of Affective Disorders, 82 (1), 125-129. doi.org/10.1016/j.jad.2003.09.002	Wrong study design, not RCT
Tanaka, S., Ishikawa, E., Mochida, A., Kawano, K., Kobayashi, M. (2015). Effects of Early-Stage Group Psychoeducation Programme for Patients with Depression. Occup. Ther. Int., 22, 195-205.	Non randomized study
Thimm J.C. & Antonsen, L. (2014). Effectiveness of cognitive behavioral group therapy for depression in routine practice. BMC Psychiatry, 14 (292). doi.org/10.1186/s12888-014-0292-x	Study design, not RCT
Ward, E.C. & Brown, R. L. (2015). A culturally adapted depression intervention for African American adults experiencing depression: Oh Happy Day. Am J Orthopsychiatry, 85(1), 11-22. doi: 10.1037/ort0000027.	Non-RCT
Zu S., Xiang, Y.T., Liu, J., Zhang, L., Wang, G., Ma, X., Kilbourne, A.M., Ungvari, G.S., Chiu, H.F., Lai, K.Y., Wong, S.Y., Yu, D.S. & Li,	Not group intervention

Z.J. (2014). A comparison of cognitive-behavioral therapy, antidepressants, their combination and standard treatment for Chinese patients with moderate-severe major depressive disorders. Journal of Affective Disorders, 262 (7), 152-154). doi: 10.1016/j.jad.2013.09.022	

6.4 Appendix 4: Ongoing studies

Title	Reference
Effectiveness of the first French Psychoeducational program on unipolar depression: study protocol for a randomized controlled trial	Ducasse, D., Courtet, P., Seneque, M., Genty, C., Picot, M-C., Schwan, R. & Olie, E. (2015). *BMC Psychiatry*, 294 (15), DOI 10.1186/s12888-015-0667-7
Family psychoeducation for major depressive disorder – study protocol for a randomized controlled trial	Timmerby, N., Austin, S.F., Ussing, K., Bech, P. & Csillag,C. (2016). *BMC Trials* 17(1): 427, DOI: 10.1186/s13063-016-1549-0
Evaluating effectiveness and cost-effectiveness of a group psychological intervention using cognitive behavioural strategies for women with common mental disorders in conflict-affected rural Pakistan: study protocol for a randomised controlled trial.	Chiumento, A., Hamdani, S.U., Khan, M.N., Dawson, K., Bryant, R.A., Sijbrandij, M., Nazir, H., Akhtar, P., Masood, A., Wang, D., van Ommeren, M. & Rahman, A. (2017). Chiumento et al. Trials, 18 (190). DOI 10.1186/s13063-017-1905-8.

6.5 Appendix 5: Characteristics of the included studies and risk of bias

6.5.1 Aagaard 2017

Methods	Design: Randomized control trial, multicentre.
Participants	Sample size: 80 Inclusion criteria: ICD-10 recurrent depression (100% severe) Exclusion criteria: No specific exclusion criteria were used Gender: 71% females, 29 % males Age mean: 48 Setting: Outpatients at 4 Community Mental Health Centres (CMHC) in Denmark
Interventions	Intervention: Group Psychoeducation Programme(PEP) based on own 130 pages manual, for groups of patients (6-8), one session including family member. Dose: 120 minutes weekly, 8 sessions + Treatment as usual (TAU), Antidepressant (52,4%) Control group: 2 years outpatient follow by CMHCs, antidepressant (36,8%) Outcome measurement points: 6 months, 18 months, 24 months Therapists: Highly experienced group therapists or therapists under training
Outcomes	Decline in Beck's depression inventory (BDI) and decline in psychiatric inpatient service, Drop-out/non compliance, psychotropic drugs and social measurements
Notes	

Risk of bias table

Bias	**Authors' jugement**	**Support for jugement**
Random sequence generation (selection bias)	Low risk	"Each CMHC had received two boxes with numbered closed envelopes for a sex and center stratified randomization". Page 224.

Allocation concealment (selection bias)	Low risk	Closed envelopes. There were some differences in the groups as follows: Significant differences at 95% level of marital status (64,3% for cases and 36,8% for controls) and absent due to illness for patients attached to the Labor market (88,5% for cases and 61,1 % for controls).
Blinding of participants and personnel (performance bias)	High risk	No blinding of patients nor therapists.
Blinding of outcome assessment (detection bias)	Low risk	BDI is self-report. The patients are not blinded (already assessed as high risk).
Incomplete outcome data (attrition bias)	High risk	3 drop outs from intervention group and 11 from control group (28,9%). Missing BDI data from 11 patients in control group. However reason and numbers are reported; two elderly patients died of ischemic heart disease, three cases and nine controls wished to cease participation. No ITT.
Selective reporting (reporting bias)	Low risk	All expected outcome appear to be reported.
Other bias	Low risk	No other bias suspected.

6.5.2 Casañas 2012

Methods	Design: Randomized controlled trial
Participants	Sample size:231 Recruitment: By General Practitioners and nurses from 12 primary care centers in Barcelona. Inclusion: MDD according to ICD-10 Depressive Disorder; BDI >10 and BDI<30 (mild or moderate depression). Patients older than 20 years. Signed informed consent. Exclusion: Other psychiatric disorder (including substance abuse), suicidal, using 2. mental health services, acute & terminal illness, Inability to speak and understand local language. Sensory or cognitive disabilities, illiteracy, temporary resident on non-provision of consent. Antidepressant not changed during the previous months. Gender: 89 % females, 11 % males

	Age mean: 53,8 Setting: Primary Care (PC), Barcelona, Spain
Interventions	Intervention: Group psychoeducation, 12 groups at different PCCs, consisting of 8-12 participants. The researcher developed a protocol with a program in order to homogenize the study intervention. See program in table 1, p. 4 of article. Dose: 90 minutes weekly, 12 sessions +TAU, Antidepressant (71%) Control group: TAU+ Antidepressant (58%) Therapist: Two nurses Outcome measurement points: 3, 6 and 9 months
Outcomes	Rate of remission; mean BDI score of <11. Quality of life (EQ-5D).
Notes	

Risk of bias table

Bias	**Authors' jugement**	**Support for jugement**
Random sequence generation (selection bias)	Low risk	The participants were randomly allocated to one of two conditions by means of a computer-generated random allocation list.
Allocation concealment (selection bias)	Low risk	"An independent person was responsible for managing the randomization lists in a sealed envelope to the two nurses at each PCC a few days before the intervention began".
Blinding of participants and personnel (performance bias)	High risk	Patients and therapists not blinded.
Blinding of outcome assessment (detection bias)	Low risk	BDI is self-reported. The patients are not blinded (already assessed as high risk).

Incomplete outcome data (attrition bias)	Low risk	The overall drop-outs rate was reported to be 23% after 3 months. The long-term drop outs rate after 9 months was a lot higher (56%). In total 72 patients from intervention group and 58 from control group dropped out during the whole period. The authors writes that there was no statistically difference between the drop outs of the two groups after 3 months follow up. Resons for drop outs were: not contactable by telephone and did not attend the interview with the nurse (42), not interested in the study (1), change of adress (3), referred to a secondary mental health service (2) and other unspecified reasons (6). ITT performed, inputed previous scores for the lost data.
Selective reporting (reporting bias)	Low risk	Study protocol available. All prespecified outcome reported.
Other bias	Low risk	No other bias suspected.

6.5.3 Chetty 2013

Methods	Design: Randomized control trial
Participants	Sample size: 30 Recruitment: A poster displayed 6 weeks prior to the study start, informed female patients with a diagnosis of depression attending the clinic about the proposed study and it purpose. Inclusion criteria: Understand English, be indian, reside in the south of Durban, age between 25-65 years, diagnosed as depressed either by clinical features or by DSM 4 diagnosed by a Medical Officer or a Psychiatric at the clinic. BDI score > 9 and < 29 (mild to moderate depression). Have been on antidepressant medication prescribed at the clinic by the Psychiatrist or the Medical Officer for 3 months or more. Exclusion criteria: Not specified Gender: 100% females Age mean: 45,2 Setting: Urban-community-psychiatric-clinic in South Africa
Interventions	Intervention: Nurse-facilitated-cognitive-group (FCG) intervention followed principles of the cognitive group therapy program, as indicated in the Verona Gordon's (1988) Women's workbook and Facilitator's Manual. Patient groups. Dose: 15 sessions, 60-120 minutes.

	+Usual treatment: Monthly follow up by nurses to collect medications, Antidepressant (66,76% TCA +33,33% TTCA) Control group: Usual treatment = no psychotherapeutic-treatment but use of antidepressant (86,76% TCA and 13,33% TTCA). Monthly meetings with a psychiatric nurse and referred to a doctor in case of complications. Therapist: Nurses Outcomes measurement points: At 6 and 12 weeks after intervention baseline.
Outcomes	Improvement in levels of depression (BDI)
Notes	

Risk of bias table

Bias	**Authors' jugement**	**Support for jugement**
Random sequence generation (selection bias)	Low risk	"30 consenting participants were selected and randomly allocated to the two groups". 30 cards were kept in a hat with coded alphabets (C/G, N/FC/G).The hat with the cards was folded and the participant were picking up a card with a coded message.
Allocation concealment (selection bias)	Low risk	See above.
Blinding of participants and personnel (performance bias)	High risk	Patients and therapists not blinded.
Blinding of outcome assessment (detection bias)	Low risk	BDI is self-reported. The patients are not blinded (already assessed as high risk).
Incomplete outcome data (attrition bias)	Low risk	No losses to follow up, all participants were obliged to attend all sessions.
Selective reporting (reporting bias)	Low risk	All expected outcomes appear to be reported.
Other bias	Low risk	No other bias suspected.

6.5.4 Cohen 2010

Methods	Design: A randomized clinical trial

Participants	Sample size: 35 Recruitment: Newspaper ads, radio, TV announcement, flyers, and pamphlets sent to local clinics in Long Island,US, all of which described a free therapy program for couples struggling with depression. Inclusion: The majority had MDD, and some had Dysthymia. BDI-2> 21 Exclusion: Male caregiver should not be clinical depressed Severely discordant couples were referred to other treatment. Infidelity, domestic violence. Gender: 100% females Age mean: 43,74 Location: Outpatients, Long Island, US.
Interventions	Intervention: Brief, problem-focused couple therapy for depression-a treatment that combined psychoeducational and cognitive-behavioral marital therapy approaches to working with couples in which one partner was depressed. Patient and caregiver in group. Dose: 5 sessions weekly, 120 minutes +TAU, Antidepressant (39%), Individual psychotherapy (17%) Control group: Waiting list (TAU), Antidepressant (94%), Individual psychotherapy (29%) Therapist: Advance clinical doctoral student and first author Outcomes measurement points: at 5 weeks and 3 months
Outcomes	Depression symptom reduction: BDI-2 and HAM-D, spouse impact (FSDS), change of behavioral and attitude (IRBAS) and overall relationship satisfaction
Notes	

Risk of bias table

Bias	**Authors' judgement**	**Support for judgement**
Random sequence generation (selection bias)	Unclear risk	"Eligible couples were randomized to either the treatment or waiting group." No other information provided.

Allocation concealment (selection bias)	Unclear risk	As above.
Blinding of participants and personnel (performance bias)	High risk	Participants are not blinded. The author states that clinicians who provided the treatment did not have access to information about the couple group assignment. Unclear information. Care protocol is standardized and the therapists are evaluated, recordings are taken of the sessions.
Blinding of outcome assessment (detection bias)	Low risk	"A second diagnostician independently rated 25% of these interviews selected randomly at each time point".
Incomplete outcome data (attrition bias)	High risk	Out of 35 in total (18 /17), it was at post assessment a number of 16/14 who completed but only 15/12 completed the follow up after three months. Loss to follow up = 23% (17% vs 29,5%). Reasons for loss not provided. No ITT.
Selective reporting (reporting bias)	Low risk	Blinding of personell.
Other bias	Low risk	No other bias suspected.

6.5.5 Dalgard 2006

Methods	Design: Randomized controlled trial
Participants	Sample size: 155 Recruitment: Advertisement in Oslo newspaper (Aftenposten), failed to recruitment among GPs at primary care. Description about the intervention and expected effect. Inclusion criteria: Adults (>18 years) with unipolar depression according to DSM 4. BDI mean= 21.8/22.9 (moderate depression) Exclusion criteria: Psychosis, sub clinical depression, other psychiatric diagnosis, risk of suicide, preference for other therapy, lack of cognitive skills, other reasons. Gender: 76,1 % females, 23,9 % males. Age mean: 47,3 Setting: Outpatients, Oslo, Norway.

Interventions	Intervention: Group PE course, modified version of the coping with depression course (CWD) from the ODIN study. Additional booster sessions and more theory. Patients group. Dose: 8 weeks, 2,5 hours sessions, plus booster sessions at 1,2 and 4 months + TAU for those who wanted, Antidepressant (44,4%), Individual psychotherapy at inclusion (24%) Therapist: Nurses and students Control group: TAU, Antidepressant (42,7%), Individual psychotherapy at inclusion (12,7%)
Outcomes	Changes of BDI scores, with a change of more than 6 BDI considered as reliable and interesting, and BDI 10 as cut off point for depression Outcome measurement points: 2 and 6 months
Notes	

Risk of bias table

Bias	Authors' jugement	Support for jugement
Random sequence generation (selection bias)	Unclear risk	"Every second person on a list of names (N=155) was assigned to the intervention group, the others to the control group". "The sequence of names on the list was ordered according to time of recruitment". This is not a true randomization method. Unclear whether this creates a bias.
Allocation concealment (selection bias)	Unclear risk	Not described. There is a significant difference in age between the groups. Logistics regression analysis were adjusted for age, sex, marital status, education and baseline BDI.
Blinding of participants and personnel (performance bias)	High risk	Patients and therapists not blinded.
Blinding of outcome assessment (detection bias)	Low risk	BDI is self-reported. Blinding of Data analyzer performed.
Incomplete outcome data (attrition bias)	Low risk	No ITT. Author claims ITT is performed but looking at the table at page 4, the analysis are performed with the lost patients excluded. He also

		states on page 5 that N=26 were excluded from the analysis. Lost to follow up is low, under 20% (17%/17%).
Selective reporting (reporting bias)	Unclear risk	All expected outcomes appear to be reported. However, the protocol was published after the report.
Other bias	Unclear risk	Researcher bias; the activity is developed by the investigator who profited from sales of course materials.

6.5.6 Dowrick 2000

Methods	Design: Randomized controlled trial, multi center.
Participants	Sample size: 425 Recruitment: Two stage community survey Inclusion criteria: 18-65-year-old. Diagnosed with Depressive Disorder, about 71% w/MDD Exclusion criteria: Comorbid psychotic condition, drug or alcohol related disorder and major suicide risk. Gender: 75,3% females, 24,7% males Age mean: Age reported as stratified data Setting: Outpatients, Finland, Ireland, Norway, Spain and UK, rural and urban
Interventions	Intervention: Problem solving treatment (128) and Course on prevention of depression (108). Patient groups. Dose Problem Solving treatment: 6 weeks, individual sessions, less than 4 hours Dose Course prevention of depression: 12 sessions of 2 hour, over 8 weeks, groups of patients only. TAU: Antidepressant not an exclusion criterion, the patients have access to health services. Therapist: Community health worker. Allied health professional. Control group: No treatment/TAU; patients have access to health services
Outcomes	Acceptability of two interventions (withdrawals) Caseness (FS-36) Depression symptoms (BDI)

	Subjective function (SF-36) Outcomes measurement points: 6 and 12 months
Notes	No other bias suspected

Risk of bias table

Bias	**Authors' jugement**	**Support for jugement**
Random sequence generation (selection bias)	Low risk	Allocation schedules were generated by random number tables and administrated by staff not in contact with the participants. No description.
Allocation concealment (selection bias)	Unclear risk	Cases were randomly allocated to one of the three groups.
Blinding of participants and personnel (performance bias)	High risk	Patients and therapists not blinded.
Blinding of outcome assessment (detection bias)	Low risk	BDI & SF-36 are self-reported. The patients are not blinded (already assessed as high risk).
Incomplete outcome data (attrition bias)	Low risk	High number of loss to follow up but equal in both groups; 30% . Those who are lost had severe depression. Inputation of data performed to ensure ITT, p. 321.
Selective reporting (reporting bias)	Low risk	All expected outcomes appear to be reported.
Other bias	Low risk	No other bias suspected.

6.5.7 Günadyın 2017

Methods	Design: Semi-experimental trial
Participants	Sample size: 153 Recruitment: Patients who were in psychiatric policlinic of a state hospital during one month and diagnosed for the first time with MMD. Inclusion and exclusion criteria: Unipolar depression diagnosed based on DSM-4 criteria. BDI between 17-30 (moderate to severe depression). Treatment plan have included that patient should take antidepressant and for the first time. No other diagnosis. Be literate. Age 18-65. Not visually impaired. Not

	previously hospitalised, not using any oral or depot antipsychotic medications. No learning disability, organic brain disease or substance or alcohol abuse. Gender: 92,2% females, 7,8% males. Age mean: Age reported in intervals, no mean reported. Setting: Policlinic, Istanbul, Turkey
Interventions	Intervention: Group PE (Continuity Enhancement Therapy for Antidepressant (CETA), focusing on drug compliance & side effects. 2 Intervention groups: 1. Group PE (CETA) + usual care including antidepressant Dose: 5 weekly sessions, 45-60 minutes Groups consisting of patients 2. Indivual brochure CETA) and antidepressant + TAU Therapist: Psyhiatric nurse Control group: Antidepressant only/TAU
Outcomes	Depression symptoms (BDI) Outcomes measurement points: 1,3 and 6 months
Notes	

Risk of bias table

Bias	Authors' jugement	Support for jugement
Random sequence generation (selection bias)	Unclear risk	The authors states: " Ranomization methods were employed to achieve homogeneity among the groups" without further description of the process.
Allocation concealment (selection bias)	Unclear risk	No description provided.
Blinding of participants and personnel (performance bias)	High risk	Patients and therapists not blinded.

Blinding of outcome assessment (detection bias)	Low risk	BDI is self-reported. The patients are not blinded (already assessed as high risk).
Incomplete outcome data (attrition bias)	Unclear risk	No description provided.
Selective reporting (reporting bias)	Low risk	All pre-specified outcomes are reported.
Other bias	Low risk	No other bias suspected.

6.5.8 Kumar 2015

Methods	Design: Randomized controlled trial
Participants	Sample size: 80 Inclusion criteria patient: Diagnosed with ICD-10 Depressive Disorder (Severe, moderate and mild) Exclusion criteria: Any comorbid physical illness, comorbid psychiatric illness, substance use disorder, bipolar disorder, partial treated or current treatment for depression. Age under 14 and above 60. Exclusion criteria care giver: Age < 18, significant medical or mental disorder, alchohol or other substance abuse disorder. Gender: 61,25% females, 38,75% males. Age 15-59 yeas, mean: 36,17 Setting: Psychiatric clinic, department of Psychiatry of of Vardhman Mahavir Medical College & Safdarjung Hospital, India
Interventions	Intervention: Family Psychoeducation, including caregiver Dose: 4 sessions over 12 weeks TAU including unstructured councelling Therapist: Researcher Control group: TAU including unctructured councelling
Outcomes	Decline in depression; Hamilton Depression Rating Scale (HDRS), Global Assessmentof Functioning (GAF)) and Psychological General Well-Being Index (PGWBI)

	Outcomes measurement points: 4,8,12 weeks
Notes	

Risk of bias table

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	"A total of 80 eligible subjects were recruited from the hospitals psychiatric department and they were randomised alternately into 2 groups". Unclear how the sequence was generated.
Allocation concealment (selection bias)	Unclear risk	No information provided.
Blinding of participants and personnel (performance bias)	High risk	Patients and therapists not blinded.
Blinding of outcome assessment (detection bias)	Low risk	"To minimise bias, outcome measures were rated by a psyciatrist not involved in the psychoeducation". No blinding of assessor.
Incomplete outcome data (attrition bias)	Low risk	Only 8 patients were lost to follow up: 5% from intervention group and 15% from control group.
Selective reporting (reporting bias)	Low risk	All expected outcomes appear to be reported.
Other bias	Low risk	No other bias suspected.

6.5.9 Sharif 2012

Methods	Design: Randomsied controlled trial
Participants	Sample size: 60 Recruitment: Admitted to psychiatric units of hospitals and met criteria. Inclusion/exclusion: MDD, age >18 years, not having other mental disoreder, no delusion or hallucination and able to participate in a group, the depression was not due to psysical disease or bipolar disorder. Gender: 55% females, 45% males Age mean: Not reported

	Setting: Inpatients, two hospitals in Shiraz-Iran.
Interventions	Intervention: Group PE Dose: 6 weekly sessions, 90 minutes Groups consisting of patients only. Therapist: Not specified Medication Control group: TAU+ medication.
Outcomes	Health related quality of life (SF-36) Outcomes measurement points: at 10 weeks
Notes	

Risk of bias table

Bias	**Authors' judgement**	**Support for judgement**
Random sequence generation (selection bias)	Unclear risk	" They were randomly assigned into the experimental and control groups after considering the preintervention baseline measurement undertaken by researcher"(p.426). Randomisation process not described.
Allocation concealment (selection bias)	Unclear risk	No information provided.
Blinding of participants and personnel (performance bias)	High risk	Patients and therapists not blinded.
Blinding of outcome assessment (detection bias)	Low risk	No blinding of researcher. The patients are not blinded (already assessed as high risk).
Incomplete outcome data (attrition bias)	Unclear risk	Loss to follow up is not reported; seems all patients completed all sessions.
Selective reporting (reporting bias)	Unclear risk	All expected outcomes appear to be reported. FS-36 is the only outcome measure. As the patients are treated for depression we find it strange that depression symptoms are not measured.

Other bias	Low risk	No other bias suspected.

6.6 Appendix 6: Outcomes overview per study

6.6.1 Aagaard 2017

Use of psychiatric hospital service during 2 years before and 2 years after the date of inclusion:

	Cases		Controls	
Number of admissions:	**Before**	**After**	**Before**	**After**
0	15 (36%)	30 (71%)	13 (34%)	28 (74%)
1	17 (40%)	10 (24%)	15 (39%)	9 (24%)
2	8 (19%)	2 (5%)	8 (21%)	1 (3%)
3	2 (5%)	0	2 (5%)	0
Total	42	42	38	38
Duration, days				
Mean (SD)	33,5 (42,7)	5,0 (16,2)	47,0 (63,7)	8,5 (19,9)
Median (SD)	26 (0-195)	0 (0-82)	15,5 (0-209)	0 (0-97)

Beck's depression inventory (BDI sum scores) at 6 month intervals during 2 years after the date of inclusion of data:

	Cases			Controls		
	N	**Mean**	**SD**	**n**	**Mean**	**SD**
6 months	40	17,5	12,6	35	17,5	12,4
12 months	40	18,8	13,6	31	16,0	11,6
18 months	39	14,6	12,0	31	15,5	12,2
24 months	39	14,7	12,6	27	17,3	11,0

6.6.2 Casañas 2012

Remission of depression in the overall, mild and moderate sample:

Sample	Months	Control n(%)	Intervention n (%)	% difference at each follow-up	IC95%)	P-value
		(n=112)	(n=119)			
Overall	**3**	21 (18,75)	41 (34,45)	15,70	(4,5 to 26,9)	0,003
	6	30 (26,79)	48 (40,34)	13,55	(1,5 to 25,6)	0,014
	9	30 (26,79)	48 (40,34)	13,55	(1,5 to 25,6)	0,014
		(n=37)	(n=48)			
Mild	**3**	15 (31,30)	21 (56,80)	25,50	(5,01 to 46)	0,009
	6	20 (41,70)	22 (59,50)	17,80	(-3,3 to 39)	0,051
	9	18 (37,50)	24 (64,90)	27,40	(6,7 to 48)	0,006
		n=82	N=64			
Moderate	**3**	6 (9,40)	20 (24,40)	15,00	(2,7 to 27,2)	0,007
	6	10 (15,60)	26 (31,70)	16,10	2,2 to 29,9)	0,011
	9	12 (18,80)	24 (29,30)	10,50	(- 3,4 to 24,5)	0,068

Overall, mild and moderate sample. Changes in BDI within and between the intervention and usual care group with missing data replaced using last value carried forward:

			Usual Care group (n=112)			**Intervention group (n=119)**				
Sample	Months	Mean (SD)	Difference (95% CI)	SRM	Mean (SD)	Difference (95% CI)	SRM	Difference	P-value	SES

	3	17,54 (7,18)	2,07 (1,0 to 3,1)	0,36	15,42(7,53)	5,47 (4,19 to 6,76)	0,77	-2,12(-4,03 to -0,214)	0,029	0,29
Overall	6	16,51 (7,60)	3,1 (1,7 to 4,4)	0,43	15,37 (8,74)	5,52 (3,9 to 7,08)	0,64	-1,13 (-3,27 to 0,992)	0,293	0,15
	9	16,35 (7,84)	3,26 (1,9 to 4,5)	0,44	15,09 (8,62)	5,8 (4,3 to 7,26)	0,72	-1,25(-3,39 to 0,886)	0,249	0,16
			Usual care group (n=48)			Intervention group (n= 37)				
	3	13,23 (5,57)	0,85 (-0,56 to 2,2)	0,17	10,38 (4,94)	3,43 (1,81 to 5,04)	0,71	-2,85 (-5,16 to -0,542)	0,016	0,51
Mild	6	13,15 (6,02)	0,93 (- 0,78 to 2,65)	0,15	10,65 (5,46)	3,16 (1,15 to 5, 16)	0,52	-2,50 (-5,015 to 0,200)	0,052	0,42
	9	12,27 (5,78)	1,81 (0,06 to 3,56)	0,30	9,70 (5,93)	4,11 (2,17 to 6,04)	0,70	-2,57 (-5,114 to -0,220)	0,048	0,44
			Usua care group (n= 64)			Intervention group (n= 82)				
	3	20,8 (6,6)	2,99 (1,45 to 4,51)	0,49	17,7 (7,4)	6,40 (4,71 to 8,09)	0,83	-3,08 (-5,41 to -0,762)	0,010	0,47
Moderate	6	19,00 (7,7)	4,74 (2,86 to 6,60)	0,63	17,5 (9,1)	6,60 (4,54 to 8,65)	0,70	-1,53(-4,35 to 1,28)	0,285	0,20
	9	19,4 (7,8)	4,36 (2,35 to 6,36)	0,54	17,50 (8,6)	6,58 (4,63 to 8,50)	0,74	-1,89 (-4,60 to 0,840)	0,174	0,24

SRM= Stanardized response mean

SES: standardized effect size

Overall, mild and moderate sample. Changes in the EQ-5D within and between the intervention and usual care group with missing data replaced using last value carried forward:

			Usual Care group (n=112)			**Intervention group (n=119)**		**Difference (95% CI) between groups (intervention group-usual care group)**		
Sample			Differences (95% CI)	SRM	Mean (SD)	Differences (95% CI)	SRM	Difference	P-value	SES
	3	55,54 (16,36)	2,29 (4,6 to -0,01)	0,19	59,7 (18,1)	8,97(12,20 to 5,72)	0,50	4,19(-0,31 to 8,66)	0,067	0,26
Overal	6	57,05 (16,97)	3,80 (6,98 to 0,61)	0,22	57,9 (20,7)	7,09(10,78 to 3,39)	0,34	0,81(-4,12 to 5,73)	0,748	0,05
	9	57,69 (17,35)	4,44 (8,0 to 0,87)	0,23	59,2 (20,8)	8,46(11,99 to 4,93)	0,43	1,54(-3,43 to 6,51)	0,543	0,09
			Usual care (n=48)			Intervention group(n=37)				
	3	60,71 (16,00)	2,79 (6,17 to -0,59)	0,24	65,7(16,7)	7,89(13,84 to 1,94)	0,44	4,99(-2,11 to 12,09)	0,166	0,31
Mild	6	60,90 (16,49)	2,98 (7,71 to -1,75)	0,18	64 (18,1)	6,14(12,73 to -0,47)	0,31	3,05(-4,43 to 10,53)	0,420	0,18
	9	62,52 (15,02)	4,60 (9,75 to –0,54)	0,25	67,8(20,5)	9,97(16,82 to 3,11)	0,48	5,26 (-2,39 to 12,92)	0,175	0,35
			Usual care group (n=64)			Intervention group (n=82)				
	3	51,67 (15,67)	1,92 (5,15 to -1,31)	0,15	57,04 (18,08)	9,45(13,39 to 5,50)	0,52	5,36(-0,263 to 10,99)	0,062	0,34
Moderate	6	54,17 (16,88)	4,42 (8,82 to 0,02)	0,25	55,11 (21,35)	7,52(12,07 to 2,97)	0,36	0,93(-5,49 to 7,37)	0,774	0,06

	9	54,06 (18,20)	4,31 (89,33 to -0,70)	0,2 1	55,37 (19,8 3)	7,78(11,95 to 3,60)	0,40	1,30(-5,00 to 7,61)	0,684	0,07

6.6.3 Chetty 2013

The mean Beck Depression Inventory scores for the two groups at the scoring sessions

	Intervention group (NFCG)	**Control Group (CG)**	**P value**
Post-test (6 weeks)	17,90	20,70	0,096
Post-test (12 weeks)	14,60	21	<0,001
p-value	<0,001	0,597	

6.6.4 Cohen 2010

Means and Standard Deviations of Outcome Variables for Depressed Women and Partners:

	Posttreat ment (5 weeks)				**3 Months follow-up**			
Outcome-measure	M	SD	t	d	M	SD	T	D
BDI-2(W)								
Treatment	20,34	13,48	0,93	0,34	14,41	10,56	1,60	0,62
Control	25,28	13,86			26,92	17,16		
HAM-D(W)								
Treatment	18,38	10,77	1,92	0,70	13,60	11,43	2,81	1,09
Control	26,29	10,55			26,42	12,25		
FDSD(H)								
Treatment	49,91	16,15	0,17	0,06	42,20	12,87	0,64	0,25

Control	50,91	13,06			48,86	15,26		
IRBAS								
Treatment(Wife)	40,00	7,13	2,04	0,87	42,40	9,12	2,24	0,86
Control(Wife)	33,45	7,97			33,83	10,84		
Treatment(Husband)	40,44	4,57	1,93	1,32	41,80	5,53	2,83	1,28
Control(Husband)	34,06	5,08			35,66	3,89		
DAS								
Treatment (Wife)	100,6	20,52	1,02	0,37	102,07	22,77	1,10	0,43
Control (Wife)	91,87	23,54			92,94	19,77		
Treatment (Husband)	108,7	19,86	1,34	0,42	108,96	16,65	1,64	0,61
Control (Husband)	100,4	14,00			98,06	18,78		

Hieararchical Linear Modeling Results for Effect of Treatment on Change in Outcome Variables:

	B	**SE**	**t**	**p**	**d**
Outcome Measure					
BDI-2	-0,41	0,16	-2,51	<.01	0,54
HAM-D	-0,47	0,14	-3,44	<.001	0,72
FDSD	-0,42	0,18	-2,31	<.05	0,80
IRBAS	0,27	0,11	2,54	<.0,1	0,39
DAS	0,55	0,19	2,83	<.0,1	0,43

6.6.5 Dalgard 2006

Beck Depression Inventory for intervention group and control group at different measurement points:

		2 months		6 months	
		Mean	**SD**	**Mean**	**SD**
Males	**Intervention group**	14,0	7,0	14,4	8,8
	Control group			17,5	7,7
	Significance			Not significant	
Females	**Intervention group**	15,6	7,8	14,0	9,5
	Control group			18,7	10,3
	Significance			P<0,05	
Total	**Intervention group**	15,2	7,6	14,1	9,3
	Control group			18,3	9,6
	Significance			P<0,05	

6.6.6 Dowrick 2000

Acceptability of problem solving treatment and course on prevention of depression.

Values are numbers of participants unless stated otherwise.

PS= Problem Solving

DP= Depression Prevention

Centre	**Treatment**	**Offered treatment**	**Refused treatment**	**Discontinued treatment**	**Did not attend**	**No (%) who completed treatment**
1	PS	23	0	5	1	17 (74)
2	PS	28	3	4	1	20(71)

3	DP	7	0	2	2	3(43)
4	DP	8	0	1	4	3(38)
5	DP	42	15	5	0	22(52)
6	DP	36	12	5	2	17(47)
7	PS	19	7	0	0	12(63)
8	PS	32	5	7	4	16(50)
8	DP	15	5	1	6	3(20)
9	PS	26	5	6	0	15(58)
	N0 (%) problem solving	128	20(16)	22 (17)	6 (5)	80 (63)
	N0 (%) depression prevention	108	32 (29)	14 (13)	14 (13)	48 (44)

Diagnosis of depressive disorders at 6 and 12 months:

	Proportions not depressed (%)			**Proportions not depressed (%)**		
	6 months			**12 months**		
	Control group	**Treatment group**	**Difference**	**Control group**	**Treatment group**	**Difference**
Centre						
1-PS	8/9(42)	11/19(58)	16	9/17(53)	11/16(69)	16
2-PS	12/20(60)	20/24(83)	23	13/20(65)	13/22(59)	-6
3-DP	3/6(50)	5/5(100)	50	5/7(71)	4/4(100)	29
4-DP	4/6(67)	2/3(67)	0	4/6(67)	3/5(60)	-7

5-DP	9/17(53)	18/31(58)	5	8/16(50)	18/35(51)	-1
6-DP	6/18(33)	12/32(38)	4	10/18(56)	14/31(45)	-10
7-PS	2/6(33)	8/15(53)	20	5/7(71)	11/17(68)	-7
8-PS	10/31(32)	12/22(55)	22	16/25(64)	12/19(63)	-1
8-DP	10/31(32)	7/9(78)	46	16/25(64)	5/8(63)	-2
9-PS	8/18(44)	7/18(39)	-6	8/13(62)	8/15(53)	--8
Total PS	40/94(43)	58/98(59)	17*	50/82(61)	55/89(62)	1**
Total DP	32/78(41)	44/80(55)	14***	43/72(60)	44/83(53)	-7****

*Odds ratio 1,39, number needed to treat 6.
**Odds ratio 1,01.
***Oddsratio 1,34, number needed to treat 7.
**** Odds ratio 0,89.

Logistic regression estimates of treatment effects on diagnosis of depressive disorders. Values are odds ratios (95% confidence intervals), with controls as reference:		
	6 months	**12 months**
Unweighted(complete case analysis)		
Problem solving	0,51(0,27 to 0,97)	0,92(0,48 to 1,77)
Depression prevention	0,50(0,21 to 1,15)	1,02(0,46 to 2,23)
Weighted (to allowe for missing outcomes)		
Problem solving	0,58(0,34 to 1,09)	0,87(0,45 to 1,70)
Depression prevention	0,47(0,20 to 1,12)	1,07(0,46 to 2,48)

Outcomes for Beck depression Inventory and SF-36 at 6 months and 12 months. Values are overall means (SD)		
Beck depression inventory	**6 months**	**12 months**

Controls	14,97(10,23)	12,60(9,50)
Problem Solving	12,48(9,95)	11,15(9,20)
Depression prevention	14,26(9,71)	14,60(8,75)
SF-36		
Mental role:		
Controls	51,71(42,70)	63,62(41,90)
Problem solving	63,91(42,13)	70,53(37,38)
Depression prevention	64,90(40,70)	61,43(40,48)
Social function:		
Controls	64,90(32,46)	70,39(30,09)
Problem solving	73,39(28,81)	75,42(29,28)
Depression prevention	68,31(29,07)	66,89(27,33)
Mental health:		
Controls	53,71(23,58)	60,51(22,39)
Problem solving	60,08(21,09)	62,79(22,00)
Depression prevention	59,54(21,41)	57,11(20,33)

Outcomes for treatment effects at 6 and 12 months					
		6 months		**12 months**	
Treatment v control	**Outcome**	**Mean (95% CI)**	**P value**	**Mean (95% CI)**	**P value**
Problem solving	BDI	-2,63(-495 to -0,32)	0,026	-1,00(-3,31 to 1,31)	0,398
Problem solving	SF-36 (mental role)	12,09(1,17 to 23,01)	0,030	8,31(-2,06 to 18,68)	0,116

Problem solving	SF-36 (social function)	9,57(2,12 to 17,02)	0,012	6,96(-0,74 to 14,59)	0,077
Problem solving	SF-36 (mental health)	7,59(2,26 to 12,92)	0,005	4,14(-0,99 to 9,28)	0,114
Depression prevention	BDI	-1,50(-4,16 to 1,17)	0,272	1,11 (-1,30 to 3,52)	0,901
Depression prevention	SF-36 (mental role)	12,70(0,46 to24,94)	0,042	-4,02(-14,53 to 6,49)	0,454
Depression prevention	SF-36 (social function)	8,66(0,07 to 17,25)	0,048	2,36(-6,10 to 10,83)	0,584
Depression prevention	SF-36 (mental health)	6,95(0,76 to 13,14)	0,028	-3,25(-8,47 to 1,97)	0,223

6.6.7 Günadyın 2017

Comparison of pre/post-education depression scores of the groups

	After 1 month of education the average depression scores	**After 3 month of education the average depression scores**	**After 6 month of education the average depression scores**
Groups	**(M+/- SD)**	**(M+/- SD)**	**(M+/- SD)**
PE group (n=49)	18,00 +/-11,50	17,53 +/- 12,01	11,18 +/- 10,36
Drug group (n=53)	19,73 +/- 10,46	18,59 +/- 13,72	16,6 +/- 12,95
Brochure group (n=51)	18,94 +/- 9,21	16,19 +/- 9,78	12,25 +/- 10,01
	X^2 =76.770	X^2 = 950.838	X^2 = 26.770
	p= 0,703	p= 0,039	p= 0,001

6.6.8 Kumar 2015

Outcome measures:

	At baseline	At 4 weeks	At 8 weeks	At 12 weeks
HDRS PE	24,23 +/- 3,00	15,62 +/- 5,25	12,72 +/- 5,10	8,43 +/- 5,90
HDRS Controls	22,48 +/- 4,31	17,61 +/- 4,92	16,21 +/- 4,82	14,71 +/- 3,40
GAF PE	62 +/- 4,80	72 +/- 6,93	75 +/- 9,20	84 +/- 8,63
GAF Controls	57 +/- 11,38	62 +/- 9,97	67,56 +/- 8,34	76,1 +/- 6,01
PGWBI PE	27,92 +/- 5,31	52,08 +/- 10,40	72 +/-10,80	87,92 +/- 7,30
PGWBI Controls	27,42 +/- 6,41	47,52 +/- 9,21	55 +/- 9,15	70,17 +/- 9,28

Impact of psychoeducation:

	At 4 weeks	At 8 weeks	At 12 weeks
HDRS PE	-8,61 (6,24-10,99)	-11,51 (9,76-14,10)	-15,80 (11,55-19,61)
HDRS Controls	-4,87 (3,24-7,99)	-6,27 (4,22-9,10)	-7,77 (4,06-11,29)
p-Value	0,003	<0,001	<0,001
GAF PE	10 (6,98-13,02)	13 (10,16-15,43)	22 (13,89-29,11)
GAF Controls	5 (4,98-8,02)	10,56 (8,86-13,24)	19,1 (12,89-23,45)
P Value	0,04	0,20	0,03
PGWBI PE	24,16 (18,64-29,67)	44,08 (42,34-46,63)	60 (55,47-64,68)
PGWBI Controls	20,10 (16,82-28,67)	27,58 (25,60-30,11)	42,75 (37,51-47,11)

P Value	0,09	0,001	0,001

6.6.9 Sharif 2012

Means score of life quality domains in groups (SF-36)

	Case group			Control Group		
	Before M(SD)	After M(SD)= 10,33 weeks	P value	Before M(SD)	After M(SD)=10,33 weeks	P value
Physical function	16,7 (3,6)	24,3 (2,6)	0,001	16,9 (4,1)	20,5 (3,09)	0,001
Role performance limitation due to physical problems	-1,16 (1,3)	-3,5 (0,8)	0,001	-1,8(1,3)	-2,2(0,8)	0,117
Role performance limitation due to psychological problems	-0,16 (0,46)	-2,4(0,56)	0,001	-0,06(0,25)	-0,86(0,68)	0,001
Social performance	3,2(1,1)	7,9(1,4)	0,001	3,8(1,3)	5,5(1,1)	0,001
Physical pain	-5,3(2,4)	-9,2(1,9)	0,001	-6,5(2,04)	-7,5(1,5)	0,008
Psychological health	9,0(2,6)	21,4(3,6)	0,001	9,9(2,3)	15,6(3,1)	0,001
Vitality	8,8(3,1)	16,8(2,3)	0,001	9,0(2,3)	13,0(2,9)	0,001
General health perception	13,4(4,2)	23,4(3,9)	0,001	13,7(3,9)	18,2(4,3)	0,001

Means score of life quality domains between groups (before and after)

	Case M (SD)	**Control M (SD)**	**P(value)**
Physical function	7.6 (3.6)	3.6 (4.4)	0,001

Role performance limitation due to physical problems	-2,3 (1.1)	-0,4(1.4)	0,001
Role performance	-2,2 (0,5)	-0,8 (0,6)	0,001
Limitation due to psychological problems			
Social Performance	4,7 (1,5)	1,7 (1,4)	0,001
Physical pain	-3,9 (2,1)	-1,0 (2,0)	0,001
Psychological health	12,4 (4,4)	5,7 (3,6)	0,001
Hapiness	8,0 (3,5)	4,0 (3,6)	0,001
General health perception	10 (4,4)	4,5 (4,4)	0,001

6.7 Appendix 7: Depression outcomes sorted by time of assessment

Depression outcomes assessed at 4 weeks

Study	Comparison	Tool	I (n)	Ctrl (n)	Baseline Mean (SD) Interv.	Baseline Mean (SD) Ctrls	4 weeks Mean (SD) Interv.	4 weeks Mean (SD) Ctrl.
Kumar 2015	Effectiveness of structured group PE including caregiver + TAU vs TAU	HDRS	40/38	40/ 34	24,23 (3)	22,48 (4,31)	15,62 (5,25)	17,61 (4,92)
Gunaydin 2017	Efficacy of group PE (CETA) + usual care vs. antidepressant only	BDI	49/49	53/ 53	26,53(8,01)	25,23 (8,66)	18,00 (11,50)	19,73 (10,46

Depression outcomes assessed at 5 weeks

Study	Comparison	Tool	I (n)	Ctrls (n)	Baseline Mean (SD) Interv.	Baseline Mean (SD) Ctrl	5 weeks Mean (SD) Interv.	5 weeks Mean (SD) Ctrl
Cohen 2010	Effectiveness of couple PE vs. Waiting list	BDI-2	18/16	17/14	31,38 (9,32)	30,16 (11,13)	20,34 (13,48)	25,28 (13,86)
Cohen 2010	Effectiveness of couple PE vs. Waiting list	HAM-D(W)	18/16	17/14	26,89(6,76)	28,53(6,93)	18,38 (10,77)	26,29 (10,55)

Depression outcomes assessed at 6 weeks

Study	Comparison	Tool	I (n)	Ctrl (n)	Baseline Mean (SD) Interv.	Baseline Mean (SD) Ctrl	6 weeks Mean (SD) Intervention	6 weeks Mean (SD) Ctrl
Chetty 2013	Nurse facilitated cognitive group intervention (NFCG) vs usual treatment	BDI	15	15	22,4	20,2	17,9 SD not measured	20,7 SD not measured

Depression outcomes assessed at 8 weeks

Study	Comparison	Tool	I (n)	Ctrls (n)	Baseline Mean (SD) Interv.	Baseline Mean (SD) Ctrl	8 weeks Mean (SD) Interv.	8 weeks Mean (SD) Ctrl
Kumar 2015	Effectiveness of structured group PE including caregiver + TAU vs TAU	HDRS	40/38	40/ 34	24,23 (3)	22,48 (4,31)	12,72 (5,10)	16,21 (4,82)

Depression outcomes assessed at 3 months

Study	Comparison	Tool	I (n)	Ctrls (n)	Baseline Mean (SD) Interv.	Baseline Mean (SD) Ctrl	3 mths Mean (SD) Interv.	3 mths Mean (SD) Ctrl
Casanas 2012	Effectiveness of a GPE vs. TAU	BDI	119 /119	112 /112	20,90 (5,68)	19,62 (5,79)	15,42 (7,53)	17,54 (7,18)
Chetty 2013	Nurse facilitated cognitive group intervention (NFCG) vs usual treatment	BDI	15	15	22,4	20,2	14,6 SD not reported	21 SD not reported

Cohen 2010	Effectiveness of couple PE vs. Waiting list	BDI-2	18/15	17/12	31,38 (9,32)	30,16 (11,13)	14,1 (10,56)	26,92 (17,16)
Cohen 2010	Effectiveness of couple PE vs. Waiting list	HAM-D(W)	18/15	17/12	26,89(6,76)	28,53(6,93)	13,60 (11,43)	26,42 (12,25)
Gunaydin 2017	Efficacy of group PE (CETA) + usual care vs. antidepressant only	BDI	49/49	53/53	26,53(8,01)	25.23(8,66)	17,53(12,01)	18,59(13,72)
Kumar 2015	Effectiveness of structured group PE including caregiver + TAU vs TAU	HDRS	40/38	40/34	24,23 (3)	22,48 (4,31)	8,43 (5,90)	14,71 (3,40)

Depression outcomes assessed at 6 months

Study	Comparison	Tool	I (n)	Ctrls (n)	Baseline Mean (SD) Interv.	Baseline Mean (SD) Ctrl	6 mths Mean (SD) Interv.	6 mths Mean (SD) Ctrl
Aagaard 2017	Effect of supplementary GPE to TAU vs. TAU	BDI	42/40	38/35	21,5 (14,1)	24,1 (12,3)	17,5 (12,6)	17,5 (12,4)
Casanas 2012	Effectiveness of a GPE vs. TAU	BDI	119 /119*	112 /112*	20,90 (5,68)	19,62 (5,79)	15,37 (8,74)	16,51 (7,60)
Dalgard 2006	Effectiveness of CWD course + TAU vs. TAU	BDI	81/62	74/67	21,8 (7,9)	22,7 (8,2)	14,1 (9,3)	18,3 (9,6)
Dowrick 2000	Group psychoeducation vs. TAU	BDI	108/80	189/139	22,41 (9,08)	22,51 (8,01) (10,23)	14,26 (9,71)	14,97 (10,23)
Gunaydin 2017	Efficacy of group PE (CETA) + usual care vs. Antidepressant only	BDI	49/49	53/53	26,53(8,01)	25.23(8,66)	11,18(10,36)	16,16(12,95)

Depression outcomes assessed at 9 months

Study	Comparison	Tool	I (n)	Ctrls (n)	Baseline Mean (SD) Interv.	Baseline Mean (SD) Ctrls	9 mths Mean (SD) Interv.	9mths Mean (SD) Ctrl
Casanas 2012	Effectiveness of a GPE vs. TAU	BDI	119 /119	112 /112	20,90 (5,68)	19,62 (5,79)	15,09 (8,62)	16,35 (7,84)

Depression outcomes assessed at 12 months

Study	Comparison	Tool	I (n)	Ctrls (n)	Baseline Mean (SD) Interv.	Baseline Mean (SD) Ctrls	12 mths Mean (SD) Interv.	12 mths Mean (SD) Ctrl
Aagaard 2017	Effect of supplementary GPE to TAU versus TAU	BDI	42/40	38/31	21,5 (14,1)	24,1 (12,3)	18,8 (13,6)	16,0 (11,6)
Dowrick 2000	Group psychoeducation vs. TAU	BDI	108/83	189/129	22,41 (9,08)	22,51(8,01)	14,60 (8,75)	12,60 (9,50)

Depression outcomes assessed at 18 months

Study	Comparison	Tool	I (n)	Ctrls (n)	Baseline Mean (SD) Interv.	Baseline Mean (SD) Ctrls	18 mths Mean (SD) Interv.	18 mths Mean (SD) Ctrl
Aagaard 2017	Effect of supplementary GPE to TAU versus TAU	BDI	42/39	38/31	21,5 (14,1)	24,1 (12,3)	14,6 (12,0)	15,5 (12,2)

Depression outcomes assessed at 24 months

Study	Comparison	Tool	I (n)	Ctrls (n)	Baseline Mean (SD) Interv.	Basline Mean (SD) Ctrls	24 mths Mean (SD) Interv.	24 mths Mean (SD) Ctrl
Aagaard 2017	Effect of supplementary GPE to TAU versus TAU	BDI	42/39	38/27	21,5 (14,1)	24,1 (12,3)	14,7 (12,6)	17,3 (11,0)

6.8 Appendix 8: GRADE evidence profiles

The effect of GPE and TAU versus TAU for MDD on depression

Bibliography: Aagaard et al., 2017, Casañas et al., 2012, Cohen, O' Leary & Foran, 2010, Dalgard, 2006, Dowrick et al., 2000, Günadyın & Barlas, 2017, Kumar & Gupta 2015.

Certainty assessment							Summary of findings				
Nº of participants (studies) Follow-up	Risk of bias	Inconsistency	Indirectness	Imprecision	Publication bias	Overall certainty of evidence	Study event		Relative effect (95% CI)	Anticipated absolute effects	
							With TAU	With GPE and TAU		Risk with TAU	Risk difference with GPE and TAU
Effect of GPE and TAU compared to TAU for MDD at 4-6 weeks follow-up **Bibliography:** Cohen, O' Leary & Foran, 2010, Günadyın & Barlas, 2017, Kumar & Gupta 2015.											
204 (3 RCTs)	not serious	not serious	not serious	serious [a]	none	⊕⊕⊕◯ MODERATE	101	103	-	-	SMD **0.32 SD lower** (0.59 lower to 0.04 lower)
Effect of GPE and TAU compared to TAU for MDD at 3 months follow-up **Bibliography:** Casañas et al., 2012, Cohen, O' Leary & Foran, 2010, Günadyın & Barlas, 2017, Kumar & Gupta 2015.											
432 (4 RCTs)	not serious	serious [b]	not serious	not serious	none	⊕⊕⊕◯ MODERATE	211	221	-	-	SMD **0.61 SD lower** (1.14 lower to 0.09 lower)
Effect of GPE and TAU compared to TAU for MDD at 6 months follow-up **Bibliography:** Aagaard et al., 2017, Casañas et al., 2012, Dalgard, 2006, Dowrick et al., 2000, Günadyın & Barlas, 2017											

The effect of GPE and TAU versus TAU for MDD on depression

Bibliography: Aagaard et al., 2017, Casañas et al., 2012, Cohen, O' Leary & Foran, 2010, Dalgard, 2006, Dowrick et al., 2000, Günadyın & Barlas, 2017, Kumar & Gupta 2015.

Certainty assessment							Summary of findings				
756 (5 RCTs)	not serious	not serious	not serious	not serious	none	⊕⊕⊕⊕ HIGH	406	350	-	-	SMD **0.21 SD lower** (0.38 lower to 0.04 lower)
Effect of GPE and TAU compared to TAU for MDD at 12 months follow-up											
283 (2 RCTs)	not serious	not serious	not serious	serious [a]	none	⊕⊕⊕◯ MODERATE	160	123	-	-	SMD **0.22 SD higher** (0.02 lower to 0.45 higher)

Explanations: CI: Confidence interval, SMD: Standardised mean difference, a. Less than 400 participants, b. Heterogeneity: I2= 82% P=0.0007

The effect of FGPE and TAU vs TAU for MDD on psychosocial functioning

Bibliography: Cohen, O' Leary & Foran, 2010, Kumar & Gupta, 2015

Certainty assessment							Summary of findings				
№ of participants (studies) Follow-up	Risk of bias	Inconsistency	Indirectness	Imprecision	Publication bias	Overall certainty of evidence	Study event rates (%)		Relative effect (95% CI)	Anticipated absolute effects	
							With TAU	With FPE and TAU		Risk with TAU	Risk difference with FPE and TAU
Psychosocial functioning at 4-5 weeks follow-up											
102 (2 RCTs)	not serious	not serious	not serious	serious [a]	none	⊕⊕◯◯ LOW	48	54	-	-	SMD **1.07 SD higher** (0.65 higher to 1.48 higher)
Psychosocial functioning at 3 months follow-up											
99 (2 RCTs)	not serious	not serious	not serious	serious [a]	none	⊕⊕◯◯ LOW	46	53	-	-	SMD **0.98 SD higher** (0.56 higher to 1.4 higher)

Legend: CI: Confidence interval; SMD: Standardised mean difference a: less than 400 participants

The effect of GPE and TAU vs TAU for MDD on quality of life **Bibliography:** Casañas et al., 2012, Dowrick et al., 2000.												
Certainty assessment							**Summary of findings**					
Nº of participants (studies) Follow-up	Risk of bias	Inconsistency	Indirectness	Imprecision	Publication bias	Overall certainty of evidence	Study event		Relative effect (95% CI)	Anticipated absolute effects		
							With TAU	With GPE and TAU		Risk with TAU	Risk difference with GPE and TAU	
The effect of GPE and TAU vs TAU for MDD on quality of life												
528 (2RCTs)	Not serious	Not serious	Not serious	Not serious	none	⊕⊕⊕⊕ HIGH	301	227	[a]	[a]	No difference between groups [a]	

Explanations: a. the result is reported differently for the two studies but none of the results show significant difference between the intervention group and the control group.

The effect of GPE and TAU vs TAU for MDD on relapse **Bibliography:** Aagaard et al., 2017												
Certainty assessment							**Summary of findings**					
Nº of participants (studies) Follow-up	Risk of bias	Inconsistency	Indirectness	Imprecision	Publication bias	Overall certainty of evidence	Study event		Relative effect (95% CI)	Anticipated absolute effects		
							With TAU	With GPE and TAU		Risk with TAU	Risk difference with GPE and TAU	
The effect of GPE and TAU vs TAU for MDD on quality of life												
80 (1RCTs)	Not serious	Not serious	Not serious	Serious[a]	none	⊕⊕⊕◯ MODERATE	38	42			No difference between groups	

Explanations: a: less than 400 participants