Pharmacological and non-pharmacological treatments for major depressive disorder: review of systematic reviews

Gerald Gartlehner,1,2 Gernot Wagner,1 Nina Matyas,1 Viktoria Titscher,1 Judith Greimel,3 Linda Lux,2 Bradley N Gaynes,4 Meera Viswanathan,2 Sheila Patel,2 Kathleen N Lohr2

ABSTRACT

Objectives This study aims to summarise the evidence on more than 140 pharmacological and non-pharmacological treatment options for major depressive disorder (MDD) and to evaluate the confidence that patients and clinicians can have in the underlying science about their effects.

Design This is a review of systematic reviews.

Data sources This study used MEDLINE, Embase, Cochrane Library, PsycINFO and Epistemonikos from 2011 up to February 2017 for systematic reviews of randomised controlled trials in adult patients with acute-phase MDD.

Methods We dually reviewed abstracts and full-text articles, rated the risk of bias of eligible systematic reviews and graded the strength of evidence.

Results Nineteen systematic reviews provided data on 28 comparisons of interest. For general efficacy, only second-generation antidepressants were supported with high strength evidence, presenting small beneficial treatment effects (standardised mean difference: −0.35; 95% CI −0.31 to −0.38), and a statistically significantly higher rate of discontinuation because of adverse events than patients on placebo (relative risk (RR) 1.88; 95% CI 1.0 to 3.28). Only cognitive behavioural therapy is supported by reliable evidence (moderate strength of evidence) to produce responses to treatment similar to those of second-generation antidepressants (45.5% vs 44.2%; RR 1.10; 95% CI 0.93 to 1.30). All remaining comparisons of non-pharmacological treatments with second-generation antidepressants either led to inconclusive results or had substantial methodological shortcomings (low or insufficient strength of evidence).

Conclusions In contrast to pharmacological treatments, the majority of non-pharmacological interventions for treating patients with MDD are not evidence based. For patients with strong preferences against pharmacological treatments, clinicians should focus on therapies that have been compared directly with antidepressants.

Strengths and limitations of this study

► This is the first review of systematic reviews assessing the benefits and harms of more than 140 pharmacological and non-pharmacological treatments for major depressive disorder.
► We used rigorous systematic review and novel graphical methods to summarise treatment effects and present the strength of the underlying evidence.
► Like any review of systematic reviews, we could draw conclusions only about interventions that had been assessed by systematic reviews.
► We did not take combination or augmentation strategies of antidepressants with non-pharmacological interventions into consideration, but in clinical practice, this is a common treatment strategy.

INTRODUCTION

Major depressive disorder (MDD)1 is the most prevalent and disabling form of depression, affecting more than 30 million Europeans per year.2 In the USA, the estimated lifetime prevalence of MDD is 16%.3 In addition to its burden of disease, MDD exerts a negative impact on physical health4–7 and adherence to medical treatment.8–9

Second-generation antidepressants (eg, selective serotonin reuptake inhibitors or selective serotonin norepinephrine reuptake inhibitors) are the most commonly used treatments for acute MDD.10 Most evidence-based guidelines recommend these medications as a first-step therapy.11–12

Nevertheless, patients with depression may prefer non-pharmacological options because antidepressant therapies also come with considerable risks for harms. Up to 63% of patients on second-generation antidepressants experience adverse events; between 7% and 15% of patients discontinue treatment because of adverse events.13 Concerns about the ‘addictiveness’ of antidepressants are also a common reason for patients’ scepticism.
about prescription medications, women and ethnic minorities, in particular, often prefer non-pharmacological options as first-step treatments of depression. Antidepressants also have a substantially higher treatment-specific stigma than, for example, herbal remedies.

Such scepticism toward antidepressants reflects a general trend toward ‘natural treatments’ throughout medicine. In 2012, an estimated 59 million persons in the USA spent US$30.2 billion in out-of-pocket expenses on some type of complementary health approach. In a survey of psychiatric patients, more than half of patients with self-reported depressive disorders used complementary and alternative medicine (CAM) therapies.

Non-pharmacological treatment options for depression are vast. The Cochrane Depression and Neurosis Group lists 87 psychological interventions, a comprehensive summary from an Australian patient advocacy group catalogued 56 CAM interventions for the treatment of depression (beyondblue. A guide to what works for depression (http://resources.beyondblue.org.au/prism/file?token=BL/0556)).

Because of the multitude of non-pharmacological options, for clinicians, the great challenge is how to balance patients’ interest in alternatives to medications with the professional responsibility to choose treatments that are supported by scientific evidence.

The goal of this project was to provide an overview of the general efficacy and risk of harms of pharmacological and non-pharmacological interventions for treating patients with MDD. Furthermore, we strove to compare benefits and harms of non-pharmacological interventions with second-generation antidepressants as the most common treatments for acute-phase MDD.

METHODS
A review of systematic reviews is designed to compile evidence from multiple systematic reviews of interventions into one accessible, usable document. We registered the protocol in PROSPERO (International Prospective Register of Systematic Reviews; registration number: 42016035580).

Populations, interventions, comparators, outcomes, timing and settings
Table 1 presents eligibility criteria for populations, interventions, comparators, outcomes, timing and settings of systematic reviews and meta-analyses. In this table, the term ‘articles’ refers to any systematic reviews or meta-analyses of randomised controlled trials (RCTs) published in peer-reviewed journals or other sources. We limited the publication period to 2011 or later because methods research indicates that more than 50% of systematic reviews are outdated 5.5 years after publication.

For eligible psychological interventions, we used the Cochrane Depression and Neurosis Group classification. For CAM, we were interested in any intervention that the non-profit patient advocacy group beyondblue listed as a ‘non-medical’ intervention for treating depressed patients. Online supplementary file 1 lists the 87 eligible psychological interventions and the 56 eligible CAM interventions.

Literature searches
To identify relevant systematic reviews or meta-analyses, we searched MEDLINE (via PubMed), Embase, the Cochrane Library, PsycINFO and Epistemonikos. We used both index terms (eg, Medical Subject Headings, Emtree) and free-text keywords to search for MDD. We limited the electronic searches to ‘human,’ ‘English, German or Italian language,’ ‘adults’ and systematic reviews or meta-analyses. We searched sources from 1 January 2011 to 20 February 2017.

We imported all citations into an electronic database (EndNote X.6.0.1). The search strategies and yields of the searches are given in online supplementary file 2.

Screening process
We developed and pilot-tested review forms using the eligibility criteria in table 1. In a two-stage review process, two persons independently reviewed abstracts and full-text articles. We resolved discrepancies by consensus or by consulting a third senior investigator. For each comparison and outcome, we chose a single systematic review providing the best available evidence. If more than one systematic review on the same intervention met eligibility criteria, we chose the review with (1) the lowest risk of bias, (2) the most recent search date and (3) the most comprehensive scope. For each eligible systematic review, we determined whether RCTs included in it also met our inclusion criteria (see table 1).

Data abstraction
We designed and used a structured form to ensure consistency of data abstraction. If all studies in a systematic review met our eligibility criteria, we extracted summary estimates from meta-analyses. If one or more studies did not meet our eligibility criteria, we extracted data from individual studies. For example, when systematic reviews included mixed populations with different depressive disorders, we retrieved individual publications on patients with MDD. When data were unclear or contradictory, we contacted review authors for clarification. A second senior reviewer evaluated the completeness and accuracy of the data abstraction.

Risk of bias assessment
To assess methodological limitations (risk of bias) of eligible systematic reviews, we used the AMSTAR (Assessing Methodological quality of Systematic Reviews) tool. Two independent reviewers assigned ratings for study limitations. They resolved any disagreements by consensus or by consulting a third independent party. For the risk of bias of individual studies in a systematic review, we relied on the ratings of the original reviews’ authors. AMSTAR ratings of included studies are given in online supplementary file 3.
Our aim was to depict the magnitude of beneficial and harmful treatment effects and the confidence that patients and clinicians can have in the underlying science about these effects. We used effect estimates of systematic reviews if all included RCTs met our eligibility criteria. In instances where individual RCTs of eligible systematic reviews did not meet our eligibility criteria (eg, because they used treatment as usual as a control group), we recalculated quantitative analyses removing ineligible studies.

For general efficacy, we were interested in the improvement of depressive symptoms. We present standardised mean differences because methods of assessments differed substantially across systematic reviews. A
standardised mean difference of 0 indicates that both groups had similar improvements; effects of −0.5 or −1 indicate that 69% or 84% of patients in the intervention group, respectively, had greater reductions on depression scores than the average patient in the control group. For the risk of harms, we present overall discontinuation rates and discontinuation rates because of adverse events.

For the comparative efficacy of non-pharmacological treatments with second-generation antidepressants, we used relative risk (RR) of response to treatment (as defined by the authors but most commonly presented as a 50% reduction of symptoms on a depression rating scale). If necessary, we recalculated RR so that a value below 1 would represent fewer responses of patients using non-pharmacological treatments and a value greater than 1 would represent more responses. We present treatment effects also as absolute risk reductions or increases (differences in numbers of patients who respond to treatment, per 1000 treated patients) with the related 95% CIs.

Quantitative analyses
As described above, in instances where individual RCTs of eligible systematic reviews did not meet our eligibility criteria, we recalculated quantitative analyses removing ineligible studies. To summarise data quantitatively, we followed established guidance. For all analyses, we used both random-effects and fixed-effects models. We report results of random-effects analyses (DerSimonian & Laird). In general, the findings from the random-effects and fixed-effects analyses were similar. We assessed statistical heterogeneity between studies by calculating the \( I^2 \) statistic and Cochran’s q. We used the \( I^2 \) statistic (the proportion of variation in study estimates attributable to heterogeneity) to estimate the magnitude of heterogeneity. We examined potential sources of heterogeneity using sensitivity analyses and assessed publication bias with funnel plots and Kendall’s tests.

For general efficacy, we estimated standardised mean differences using Hedges’ \( g \). If systematic reviews presented effect sizes as Cohen’s \( d \), we used a correction factor \( J \) to convert to Hedges’ \( g \): \( J = 1 - \frac{3}{4df - 1} \), where df stands for ‘degrees of freedom’.

If systematic reviews presented effect estimates of general efficacy as dichotomous outcomes, we calculated \( \log OR \) and converted them first to Cohen’s \( d \), then to Hedges’ \( g \) using the correction factor presented above. For each estimate, we calculated variances and CIs.

For all statistical calculations, we used Microsoft Excel (V.2010, Microsoft, Redmond, WA, USA) or Review Manager 5.3 (V.5.3. Copenhagen, The Cochrane Collaboration, 2014).

Strength of the evidence
We graded the strength of evidence based on guidance for AHRQ Evidence-based Practice Centres on the use of GRADE (Grading of Recommendations Assessment, Development and Evaluation) Working Group. Strength of evidence can take four grades: high, moderate, low or insufficient. We considered grades of high or moderate strength as reliable evidence.

RESULTS
Searches detected 2532 citations; 19 systematic reviews met our eligibility criteria and provided the most recent summaries of evidence on 28 comparisons of interest. Thirty-one additional systematic reviews formally met eligibility criteria, but their content was superseded by at least one of the 19 reviews mentioned above (online supplementary file 4). figure 1 presents the flow of the literature; table 2 presents characteristics of included reviews.

For the majority of non-pharmacological treatments, we did not find any systematically appraised evidence.

In the following sections, we first provide an overview of treatment effects of non-pharmacological and common pharmacological treatments compared with inactive interventions.

We then present results on the comparative benefits and harms of non-pharmacological interventions and second-generation antidepressants.

Non-pharmacological and pharmacological treatments compared with inactive interventions

Benefits of treatments
Sixteen systematic reviews provided data on 17 comparisons with inactive interventions (placebo, sham interventions or waiting list). figure 2 provides an overview of treatment effects of non-pharmacological and common pharmacological treatments for MDD when compared with inactive interventions using standardised mean differences. The four commonly used pharmacological interventions in the figure are agomelatine, alprazolam, second-generation antidepressants and tricyclic antidepressants.

The comparisons in the figure are ordered by the strength of evidence grades and then alphabetically by the name of the intervention. figure 2 also presents the numbers of trials and the total number of subjects in those trials; thus, the size of the circles reflects the numbers of participants (on a logarithmic scale). Online supplementary file 5 provides detailed strength of evidence ratings.

The only treatments for acute-phase MDD with high strength of evidence were second-generation antidepressants (figure 2). Within this class, the medications rendered modest treatment effects (−0.35; 95% CI −0.31 to −0.38). Although the dataset included 24 unpublished studies, treatment effects might still be inflated because several method studies indicate that publication bias is a serious problem in this drug class.

Reviews on some psychological interventions (third wave cognitive behavioural therapy (CBT) and psychodynamic therapies) reported large treatment effects (third wave CBT: −0.97; 95% CI −1.9 to −0.43; psychodynamic therapies: −2.02; 95% CI −2.1 to −1.9). The dataset included 28 comparisons with inactive interventions. AHRQ Evidence-based Practice Centres on the use of pharmacological interventions in the figure are agomelatine, alprazolam, second-generation antidepressants and tricyclic antidepressants.

The comparisons in the figure are ordered by the strength of evidence grades and then alphabetically by the name of the intervention. figure 2 also presents the numbers of trials and the total number of subjects in those trials; thus, the size of the circles reflects the numbers of participants (on a logarithmic scale). Online supplementary file 5 provides detailed strength of evidence ratings.

The only treatments for acute-phase MDD with high strength of evidence were second-generation antidepressants (figure 2). Within this class, the medications rendered modest treatment effects (−0.35; 95% CI −0.31 to −0.38). Although the dataset included 24 unpublished studies, treatment effects might still be inflated because several method studies indicate that publication bias is a serious problem in this drug class.
these two psychological interventions used waiting lists as control interventions. Patients on waiting lists usually do not experience beneficial placebo effects, which can lead to artificially large treatment effects when active interventions are compared with waiting list controls. Placebo effects in psychiatric populations can be substantial; for example, on average, 35%–40% of patients in double-blinded trials of antidepressants achieved a response (usually defined as a 50% reduction of symptoms) to placebo treatment.53

For many of the therapies in figure 2, the types of inactive comparators varied and involved different magnitudes of placebo effects. Consequently, comparisons of treatment effects across different interventions have to be made cautiously.

Risk of harms
Information on overall discontinuation and discontinuation because of adverse events was scarce. figure 3 depicts the absolute risk reductions or increases for overall discontinuation and discontinuation because of adverse events—namely, the bars showing the 95% CIs of either fewer or more discontinuations per 1000 patients.

Only patients on second-generation antidepressants had a statistically significantly higher rate of discontinuation because of adverse events than patients on placebo (4.5% vs 2.6%; RR 1.88, 95% CI 1.07 to 3.28). Most comparisons were of low or insufficient strength of evidence, indicating little certainty in the available effect estimates (details in online supplementary file 5).

Non-pharmacological treatments compared with second-generation antidepressants
Three systematic reviews provided data on response to treatment for 11 non-pharmacological interventions (four psychological, six CAM and exercise) compared with second-generation antidepressants for the treatment of acute-phase MDD.30 31 44 We used response to treatment as defined by authors of the reviews; in most cases, this was a 50% reduction of symptoms as measured on a depression rating scale (eg, Hamilton Depression Rating Scale). figure 4 depicts the absolute risk reductions or increases for response to treatment per 1000 patients. As in the other figures, the comparisons are ordered by the strength of evidence grades and then alphabetically by the name of the intervention. These estimates are based

---

**Figure 1** Flow diagram of review of systematic reviews of treatments for major depressive disorder in adults.
## Table 2  Characteristics of included systematic reviews

<table>
<thead>
<tr>
<th>Review</th>
<th>Risk of bias</th>
<th>Years covered by searches</th>
<th>Eligible study designs</th>
<th>Population</th>
<th>Intervention</th>
<th>Control</th>
<th>K relevant studies, N analysed</th>
</tr>
</thead>
<tbody>
<tr>
<td>Abbass 2014</td>
<td>Low</td>
<td>NR to July 2012</td>
<td>RCTs</td>
<td>Adults, ≥18 years of age, with common mental disorders, allowed comorbid medical or psychiatric disorders (relevant study of African American women, 20–50 years of age, with depression)</td>
<td>Psychodynamic therapies (short term)</td>
<td>Inactive treatment (waitlist)</td>
<td>Reduction: K=1, n=20</td>
</tr>
<tr>
<td>Al-Karawi 2016</td>
<td>Medium</td>
<td>NR to December 2015</td>
<td>RCTs</td>
<td>Patients with non-seasonal depression diagnosed by standardised depression scales</td>
<td>Bright light therapy</td>
<td>Inactive treatment (placebo device and pill-placebo)</td>
<td>Reduction: K=1, n=62 Discontinuation (overall): K=1, n=62 Discontinuation (adverse events): K=1, n=62</td>
</tr>
<tr>
<td>Apaydin 2016</td>
<td>Medium</td>
<td>January 2007 to November 2014</td>
<td>RCTs</td>
<td>Adults, ≥18 years of age, with a diagnosis of MDD</td>
<td>St John’s wort</td>
<td>Inactive treatment (pill-placebo)</td>
<td>Reduction: K=16, n=2888</td>
</tr>
<tr>
<td>Appleton 2015</td>
<td>Low</td>
<td>All years to May 2015 (except CINAHL, to September 2013)</td>
<td>RCTs, cross-over and cluster RCTs</td>
<td>Adults, ≥18 years of age, with a primary diagnosis of MDD or unipolar depressive disorder, allowed comorbid conditions</td>
<td>Omega-3 fatty acids (n-3PUFAs)</td>
<td>Inactive treatment (pill-placebo)</td>
<td>Reduction: K=6, n=308 Discontinuation (overall): K=7, n=446</td>
</tr>
<tr>
<td>Cujipers 2014</td>
<td>Medium</td>
<td>1966 to January 2012</td>
<td>RCTs</td>
<td>Adults diagnosed with a depressive disorder, allowed comorbid medical or psychiatric disorders</td>
<td>Humanistic therapy (supportive therapy)</td>
<td>Inactive treatment (pill-placebo)</td>
<td>Reduction: K=1, n=101</td>
</tr>
<tr>
<td>Ekers 2014</td>
<td>High</td>
<td>1966 to January 2013</td>
<td>RCTs</td>
<td>Adults, ≥16 years of age, with a primary diagnosis of depression</td>
<td>Third Wave CBT (behavioural activation therapy)</td>
<td>Inactive treatment (waitlist, placebo)</td>
<td>Reduction: K=9, n=338</td>
</tr>
<tr>
<td>Furukawa 2017</td>
<td>Medium</td>
<td>NR to January 2015</td>
<td>RCTs</td>
<td>Adults with MDD, diagnosed according to DSM or ICD-10</td>
<td>CBT</td>
<td>Inactive treatment (pill-placebo)</td>
<td>Reduction: K=5, n=509</td>
</tr>
<tr>
<td>Galizia 2016</td>
<td>Medium</td>
<td>NR to February 2016</td>
<td>RCTs</td>
<td>Adults, aged 18–80 years with a diagnosis of major depression</td>
<td>SAMe</td>
<td>Inactive treatment (pill-placebo)</td>
<td>Reduction: K=2, n=142 Discontinuation (overall): K=2, n=142 Discontinuation (adverse events): K=1, n=124</td>
</tr>
</tbody>
</table>

Continued
<table>
<thead>
<tr>
<th>Review</th>
<th>Risk of bias</th>
<th>Years covered by searches</th>
<th>Eligible study designs</th>
<th>Population</th>
<th>Intervention</th>
<th>Control</th>
<th>K relevant studies, N analysed</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gartlehner 2015</td>
<td>Medium</td>
<td>January 1990 to September 2015</td>
<td>RCTs, allowed non-randomised studies for harms</td>
<td>Adults, ≥19 years of age, with MDD during initial treatment attempt or second treatment attempt among those who did not achieve remission after treatment with an SGA</td>
<td>Acupuncture</td>
<td>SGA</td>
<td>Response; K=93 (NWMA), n=173</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>CBT</td>
<td>SGA</td>
<td>Response; K=5, n=660</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Exercise</td>
<td>SGA</td>
<td>Response; K=90 (NWMA), n=0</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Integrative therapy (interpersonal psychotherapy)</td>
<td>SGA</td>
<td>Response; K=1, n=318</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Omega-3 fatty acids</td>
<td>SGA</td>
<td>Response; K=92 (NWMA), n=40</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>SAMe</td>
<td>SGA</td>
<td>Response; K=90 (NWMA), n=0</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>St John’s wort</td>
<td>SGA</td>
<td>Response; K=9, n=1517</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Third Wave CBT (Behavioural activation)</td>
<td>SGA</td>
<td>Response; K=2, n=243</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>SGA</td>
<td>Inactive treatment (pill-placebo)</td>
<td>Reduction; K=62, n=13759</td>
</tr>
<tr>
<td>Josefsson 2014</td>
<td>High</td>
<td>NR to April 2012</td>
<td>RCTs</td>
<td>Adults, ≥18 years of age, with depression or depressive symptoms</td>
<td>Exercise (aerobic or non-aerobic exercise, as monotherapy or with usual care, excluding eastern meditative practices)</td>
<td>Inactive treatment (no treatment, placebo)</td>
<td>Reduction; K=11, n=368</td>
</tr>
<tr>
<td>Jun 2014</td>
<td>Medium</td>
<td>NR to February 2014</td>
<td>RCTs, quasi-RCTs</td>
<td>Individuals of any age and either sex with depression, allowed comorbid diseases</td>
<td>Gan Mai Da Zao (decoction or modified decoction)</td>
<td>SGA</td>
<td>Response; K=3, n=148</td>
</tr>
</tbody>
</table>

Continued
<table>
<thead>
<tr>
<th>Review</th>
<th>Risk of bias</th>
<th>Years covered by searches</th>
<th>Eligible study designs</th>
<th>Population</th>
<th>Intervention</th>
<th>Control</th>
<th>K relevant studies, N analysed</th>
</tr>
</thead>
<tbody>
<tr>
<td>Linde 2015&lt;sup&gt;34&lt;/sup&gt;</td>
<td>Medium</td>
<td>NR to December 2013</td>
<td>RCTs</td>
<td>Adults with prevalent or incident unipolar depressive disorder</td>
<td>St John’s wort</td>
<td>Inactive treatment (pill-placebo)</td>
<td>Discontinuation (overall): K=4, n=619</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Discontinuation (adverse events): K=3, n=522</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>TCA</td>
<td>Inactive treatment (pill-placebo)</td>
<td>Discontinuation (overall): K=4, n=484</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Discontinuation (adverse events): K=3, n=421</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>SGA</td>
<td>Inactive treatment (pill-placebo)</td>
<td>Discontinuation (overall): K=5, n=1195</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Discontinuation (adverse events): K=6, n=1572</td>
<td></td>
</tr>
<tr>
<td>Liu 2015&lt;sup&gt;17&lt;/sup&gt;</td>
<td>High</td>
<td>NR to February 2014</td>
<td>RCTs</td>
<td>Older adults, mean age ≥60 years, with depressive symptoms, and allowed comorbidities</td>
<td>Tai Chi, Qigong</td>
<td>Inactive treatment (newspaper reading or reading and discussion group, health education)</td>
<td>Reduction: K=3, n=193</td>
</tr>
<tr>
<td>Okumura, 2014&lt;sup&gt;43&lt;/sup&gt;</td>
<td>High</td>
<td>1994 to June 2013</td>
<td>RCTs, cluster RCTs, quasi-RCTs</td>
<td>Adults, ≥18 years of age, with depression (elevated depressive symptoms, depressive disorders or minor depression), allowed comorbid physical illness</td>
<td>CBT (group CBT, mindfulness-based cognitive therapy)</td>
<td>Inactive treatment (waitlist, pill-placebo)</td>
<td>Discontinuation (overall): K=7, n=834</td>
</tr>
<tr>
<td>Sorbero 2015&lt;sup&gt;33&lt;/sup&gt;</td>
<td>Medium</td>
<td>NR to January 2015</td>
<td>RCTs</td>
<td>Adults, ≥18 years of age, with a clinical diagnosis of MDD at enrolment or formerly depressed if primary outcome of study was depression relapse or recurrence</td>
<td>Acupuncture (specific, needle or electroacupuncture)</td>
<td>Inactive treatment (non-specific acupuncture)</td>
<td>Reduction: K=3, n=168</td>
</tr>
<tr>
<td>Taylor 2014&lt;sup&gt;43&lt;/sup&gt;</td>
<td>Medium</td>
<td>NR to March 2013</td>
<td>RCTs</td>
<td>Adults with depression</td>
<td>Agomelatine</td>
<td>Inactive treatment (pill-placebo)</td>
<td>Reduction: K=12, n=3855</td>
</tr>
</tbody>
</table>
### Table 2: Continued

<table>
<thead>
<tr>
<th>Review</th>
<th>Risk of bias</th>
<th>Years covered by searches</th>
<th>Eligible study designs</th>
<th>Population</th>
<th>Intervention</th>
<th>Control</th>
<th>K relevant studies, N analysed</th>
</tr>
</thead>
<tbody>
<tr>
<td>Undurraga 2012</td>
<td>High</td>
<td>1980 to August 2011</td>
<td>RCTs</td>
<td>Adults in an acute, apparently unipolar MDD episode or with ≤10% identified cases of bipolar depression or diagnoses other than MDD</td>
<td>TCA</td>
<td>Inactive treatment (pill-placebo)</td>
<td>Reduction: K=21, n=3094</td>
</tr>
<tr>
<td>Van Marwijk 2012</td>
<td>Low</td>
<td>All years to February 2012</td>
<td>RCTs</td>
<td>Adults, ≥18 years of age, with a primary diagnosis of MDD, a depressive episode, or if considered depressed and eligible for antidepressant treatment by a clinician</td>
<td>Alprazolam</td>
<td>Inactive treatment (pill-placebo)</td>
<td>Reduction: K=5, n=603</td>
</tr>
<tr>
<td>Yeung 2014</td>
<td>Medium</td>
<td>NR to May 2013</td>
<td>RCTs, quasi-RCTs</td>
<td>Individuals diagnosed with depression</td>
<td>Chinese herbal medicine</td>
<td>SGA</td>
<td>Response: K=5, n=1360</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Saffron</td>
<td>SGA</td>
<td>Reduction: K=2, n=171</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Saffron</td>
<td>SGA</td>
<td>Response: K=1, n=38</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Saffron</td>
<td>SGA</td>
<td>Reduction: K=2, n=80</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Saffron</td>
<td>SGA</td>
<td>Discontinuation (overall): K=2, n=80</td>
</tr>
</tbody>
</table>

CBT, cognitive behavioural therapy; K, number of studies that were eligible for review of reviews; MDD, major depressive disorder; N, number of participants in eligible studies; NR, not reported; NWMA, network meta-analysis; n-3PUFA, n-3 polyunsaturated fatty acid; RCT, randomised control trial; SGA, second-generation antidepressant; TCA, tricyclic antidepressants.

Open Access

on meta-analyses or, if meta-analyses were not feasible, on results from the largest and most reliable trial. Online supplementary file 5 provides detailed information on our ratings of strength of evidence domains.

**Psychological interventions**

One systematic review reported on the efficacy of four psychological treatments relative to second-generation antidepressants (figure 4); these included CBT, integrative therapies, psychodynamic therapies and third wave CBT. Integrative therapies also had response rates similar to those for antidepressants (low strength of evidence). Patients treated with third wave CBT had significantly higher response rates than those on antidepressants, but the strength of evidence was insufficient because of the small sample size and under-dosing of antidepressants in the available trial. No evidence on response was available for psychodynamic therapies, but the available evidence indicated remission rates similar to those for second-generation antidepressants.

**CAM interventions**

Three systematic reviews reported on comparisons with second-generation antidepressants for seven (of 56 interventions) that were rated as moderately well supported by evidence. Of these, musculoskeletal interventions were the only group of interventions that did not differ from placebo in any report. The CAM interventions overall were found to be substantially more effective than placebo but not as strongly supported as the meta-analytic evidence for the antidepressants.

**Figure 2** Overview of the strength of evidence of non-pharmacological and pharmacological interventions compared with inactive interventions for the treatment of adult major depressive disorder. CBT, cognitive behavioural therapy; SAMe, S-adenosyl-L-methionine; SGA, second-generation antidepressants; SMD, standardised mean difference; TCA, tricyclic antidepressants.

**Figure 3** Absolute risk reductions or increases of overall discontinuation or discontinuation because of adverse events comparing non-pharmacological interventions with inactive interventions. CBT, cognitive behavioural therapy; SAMe, S-adenosyl-L-methionine; SGA, second-generation antidepressants; TCA, tricyclic antidepressants.
A network meta-analysis of antidepressants for major depressive disorder. 3 Number of participants in trials that directly compared intervention with second-generation antidepressants. 4 Number of trials in network meta-analysis that contributed to the effect estimate. CBT, cognitive behavioural therapy; RR, relative risk; SAMe, S-adenosyl-L-methionine; SGA, second-generation antidepressants.

Some comparisons had wide CIs (eg, acupuncture, Gan Mai Da Zao, omega-3 fatty acids, S-adenosyl-L-methionine (SAMe), St John’s wort and saffron) rendering inconclusive findings about the comparative efficacy of treatments. Other comparisons had more precise results (eg, Chinese herbal medicine or St John’s wort) but severe methodological shortcomings. For example, several trials of St John’s wort had serious design or analytic limitations. Therefore, the lack of statistical significance of most comparisons should not be interpreted as equivalence of treatment effects.

Comparative harms

The discontinuation of treatment because of adverse events were generally lower for patients treated with non-pharmacological interventions than for those receiving second-generation antidepressants, although differences did not always reach statistical significance. Patients on St John’s wort had a statistically significantly lower rate of discontinuation because of adverse events (3.8% vs 6.8%; RR 0.59; 95% CI 0.38 to 0.89). 44 Patients on any psychological treatment had a numerically lower risk for discontinuation of treatment because of adverse events (2.1% vs 7.1%; RR 0.37; 95% CI 0.12 to 1.12). 44 Likewise, patients who used physical exercise discontinued treatment because of adverse events less often than those treated with antidepressants (0% vs 6%; RR 0.15; 95% CI 0.01 to 2.86), but the difference did not reach statistical significance. 44 Little evidence on treatment discontinuation was available for most CAM interventions, particularly for Chinese herbal medicine or saffron. 30 31

DISCUSSION

Out of more than 140 interventions of interest, our review identified only five treatments for which the general efficacy for acute-phase MDD is supported by reliable evidence (ie, evidence graded as high or moderate strength of evidence). Among those, CBT is the only psychological intervention and St John’s wort is the only CAM intervention. For the vast majority of non-pharmacological interventions, either no systematic review evidence was available or the certainty of the evidence was severely limited. When compared with second-generation antidepressants, only CBT had similar efficacy based on moderate strength evidence. Overall, our analyses highlighted a lack of robust evidence for the majority of non-pharmacological treatments. To our knowledge, our study was the first review of systematic reviews assessing more than 140 interventions for treating adults with MDD. It provides a unique synthesis of the available, systematically appraised evidence on these treatment options, beyond the individual reviews on depression therapies that have been published over the past decade.

Our study does have several limitations, however. First and most importantly, like any review of systematic reviews, we could draw conclusions only about interventions that had been assessed by systematic reviews. Conceivably, RCTs are available for some interventions that have never been evaluated systematically in a review. Therefore, the absence of systematic reviews cannot be equated with an absence of RCTs. In addition, eligibility criteria of these reviews sometimes included only a subset of available studies (eg, studies conducted in primary care settings). Such reviews do not provide a picture of the totality of the evidence but sometimes

were the only ones that were available on a specific comparison of interest. Second, reviews of systematic reviews rely on results from other investigators. Although most of the reviews had few problems in methods, conceivably these authors did miss some RCTs. Likewise, we relied on the risk of bias appraisals of RCTs that authors of included systematic reviews had done. Most reviews used two independent reviewers to rate risk of bias; double checking their ratings was beyond the scope of our study. Third, reporting of characteristics of populations, interventions, comparators and outcomes in included systematic reviews was often suboptimal. Frequently, we could not tell with certainty whether included populations were exclusively adult patients with acute-phase MDD; sometimes, we could not determine the exact control interventions that authors had combined in their meta-analyses. We did not take several meta-analyses into consideration that combined studies with inactive treatments and treatment as usual as control interventions. Because treatment as usual cannot be viewed as ‘inactive,’ we believe that such meta-analyses will lead to biased results. Fourth, as in any literature review, the reliability of our results is directly related to number of available studies and their quality. Some of the systematic reviews included only few studies with few events. The strength of evidence grades reflect these concerns and the certainty of our results; for most cases, these grades were low or insufficient. Such low strength of evidence indicates that future studies might have a substantial impact on the effect estimates reported in our review. Furthermore, we had no way to assess how meta-biases such as reporting biases or funding biases could have affected our findings. Finally, we did not take combination or augmentation strategies of antidepressants with non-pharmacological interventions into consideration, but in clinical practice, this is a common treatment strategy.

We believe that our results may have important clinical implications. They provide patients and clinicians with solid and up-to-date information about which treatment options have (or have not) been evaluated in rigorous systematic reviews. For patients with strong preferences against pharmacological treatment, clinicians can offer therapies that have been compared directly with antidepressants. CBT, for example, is a well-supported, first-step alternative to pharmacological treatment of MDD. Other psychological or CAM interventions might be equally effective, or nearly so, but the evidence base is less reliable. The majority of psychological and CAM interventions, however, are not evidence based; given better alternatives, clinicians should probably advise against them. Such shared and informed decision making might enhance treatment adherence and could ultimately improve treatment outcomes for patients with MDD. This is especially important because treatment continuity is one of the main challenges in treating such patients.

Our findings also highlight key areas of future research needs. Subsequent trials need to address gaps in our current knowledge about the efficacy of non-pharmacological interventions and about the comparative benefits and harms of pharmacological and non-pharmacological treatments for MDD. In particular, major research gaps pertain to information about the comparative risk of harms and patient-relevant outcomes such as functional capacity and quality of life. For patients and clinicians alike, balancing benefits and harms based on objective information is crucial. Lack of information about harms can lead to a biased knowledge base and the potential for decisions that cause more harm than good. Future studies should assess benefits and harms with standardised measures to allow for more direct comparisons across studies.

In the end, even in the absence of clearly informative evidence, clinicians and patients need to make decisions. They can discuss what is known and what is not known about the available options to treat MDD, and our work provides a way to start those conversations. For patients with strong preferences against pharmacological treatments, clinicians should focus on therapies that have been compared directly with antidepressants. This review provides a framework to guide discussion of the potential options.

Acknowledgements We would like to thank Monika Kypelova from Danube University Krems and Loraine Monroe from RTI International for administrative support. We are also grateful to Irma Klerings from Danube University Krems for the literature searches and Joshua Green from RTI International for help with data abstraction.

Contributors GG, KL and MV developed the concept of the study; GG, JJ, GW, NM and VT conducted the literature review; GW, NM and VT abstracted data and conducted statistical analyses; MV and LL rated the risk of bias of included systematic reviews; GG, GW and NM graded the strength of evidence; BG provided clinical expertise throughout the study; GG and KL wrote the first draft of the manuscript; all authors reviewed the manuscript and provided comments and revisions.

Funding The paper was supported by internal funds from RTI International, Research Triangle Park, North Carolina.

Competing interests None declared.

Provenance and peer review Not commissioned; externally peer reviewed.

Data sharing statement The datasets used for meta-analyses are available from the corresponding author on reasonable request.

Open Access This is an Open Access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited and the use is non-commercial. See: http://creativecommons.org/licenses/by-nc/4.0/

© Article author(s) (or their employer(s) unless otherwise stated in the text of the article) 2017. All rights reserved. No commercial use is permitted unless otherwise expressly granted.

REFERENCES


