Complementary therapies for clinical depression: an overview of systematic reviews

Heidemarie Haller, Dennis Anheyer, Holger Cramer, Gustav Dobos

ABSTRACT

Objectives As clinical practice guidelines vary widely in their search strategies and recommendations of complementary and alternative medicine (CAM) for depression, this overview aimed at systematically summarising the level 1 evidence on CAM for patients with a clinical diagnosis of depression.

Methods PubMed, PsycInfo and Central were searched for meta-analyses of randomised controlled clinical trials (RCTs) until 30 June 2018. Outcomes included depression severity, response, remission, relapse and adverse events. The quality of evidence was assessed according to Grades of Recommendation, Assessment, Development, and Evaluation (GRADE) considering the methodological quality of the RCTs and meta-analyses, inconsistency, indirectness, imprecision of the evidence and the potential risk of publication bias.

Results The literature search revealed 26 meta-analyses conducted between 2002 and 2018 on 1–49 RCTs in major, minor and seasonal depression. In patients with mild to moderate major depression, moderate quality evidence suggested the efficacy of St. John's wort towards placebo and its comparative effectiveness towards standard antidepressants for the treatment for depression severity and response rates, while St. John's wort caused significant less adverse events. In patients with recurrent major depression, moderate quality evidence showed that mindfulness-based cognitive therapy was superior to standard antidepressant drug treatment for the prevention of depression relapse. Other CAM evidence was considered as having low or very low quality.

Conclusions The effects of all but two CAM treatments found in studies on clinical depressed patients based on low to very low quality of evidence. The evidence has to be downgraded mostly due to avoidable methodological flaws of both the original RCTs and meta-analyses not follow the Consolidated Standards of Reporting Trials and Preferred Reporting Items for Systematic Reviews and Meta-Analyses guidelines. Further research is needed.

INTRODUCTION

Depression is one of the most prevalent psychiatric disorders, with about 25% of women and 12% of men suffering from at least one depressive episode during their lifetime. According to the criteria for diagnosis recommended by the American Psychiatric Association (APA), depressive disorders can be distinguished by their degree of severity or duration and are also characterised by a high comorbidity and an increase of psychological strain for the affected person. It is evident that a strong comorbid connection to several chronic conditions like addictions, neurodegenerative diseases or different psychiatric diseases exists. This leads depressive disorders as one of the leading causes of disability worldwide.

The most commonly used treatments for depression are antidepressants, psychotherapy or a combination of drugs and psychotherapy. While both treatment strategies (alone and in combination) have been shown to be effective, more recent meta-analyses also found high dropout and low remission rates, as well as clinically significant differences between antidepressant drugs and placebos only for patients at the upper end of the very severely depressed category. This may lead patients to search for alternatives. Increasing mainstream use of complementary and alternative medicine (CAM) support this trend, particularly for different physical conditions with comorbid affective disorders. The NIH defines CAM as therapeutic approaches that are usually not included in conventional Western medicine systems.
with conventional care are considered as complementary, those used instead of conventional care as alternative practices. Types of CAM approaches include natural products, such as herbs and dietary supplements (vitamins, minerals and probiotics) and mind and body practices, such as yoga, chiropractic and osteopathic manipulation, meditation, relaxation, acupuncture, tai chi, qi gong and hypnotherapy. Practices of traditional healers from Europe (naturopathy and homeopathy), Asia (Ayurveda and traditional Chinese medicine) and other continents are also classified as CAM.28 While some complementary therapies have become a promising adjunct in the standard treatment of depression,39 40 others are known for their possible side effects or interactions with standard drugs.30 Recent clinical practice guidelines, in addition, vary widely in their search strategies and resulting recommendations for CAM treatments. While the American College of Physicians (ACP),31 the American Psychiatric Association (APA)32 and the Canadian Network for Mood and Anxiety Treatments (CANMAT) guideline33 provide a more comprehensive overview and critical appraisal of CAM treatments, the Deutsche Gesellschaft für Psychiatrie und Psychotherapie, Psychosomatik und Nervenheilkunde (DGPPN),34 the National Institute for Health and Care Excellence (NICE),35 and the World Federation of Societies of Biological Psychiatry (WFSBP)36 guidelines mainly focus on St. John’s Wort and light therapy. Possible effects and risks of further CAM therapies are not discussed. Thus, the purpose of this overview is to provide a comprehensive search strategy of relevant CAM terms and systematically summarise the existing level 1 evidence for clinical depression as a basis for further guideline recommendations on the efficacy, effectiveness and safety of CAM therapies.

METHODS
This systematic overview of reviews was conducted in accordance with the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines37 38 and the recommendations of the Cochrane Collaboration.39 The protocol was not prospectively registered.

Patient and public involvement
For this overview of reviews, patients or public were not involved.

Inclusion and exclusion criteria
- Types of studies: to be eligible, articles had to be systematic reviews with meta-analyses of randomised controlled trials (RCTs) published in peer-reviewed journals. Conference abstracts or unpublished work were excluded as well as reviews summarising evidence narratively. In cases of including same or similar original studies, only the review with the most recent, most comprehensive search was included. When systematic reviews reported results of RCTs as well as of designs of lower evidence levels, they were considered only if separate meta-analyses for the included RCTs were performed.
- Types of participants: only reviews of patients with a diagnosis of major depression or dysthymia were eligible as well as reviews including patients/general population samples with mild depressive symptoms above a clinical cut-off or seasonal patterns. In contrast, reviews studying depressive symptoms within specific subpopulations of substance-induced or demented patients, secondary depression due to another medical condition (eg, poststroke, cancer or pain patients), bipolar disorders, or females with premenstrual dysphoric disorder or postpartum depression were excluded. Further restrictions regarding the diagnostic criteria or procedures, regarding age, gender, duration of the condition or symptom intensity were not applied.
- Types of interventions: reviews investigating the effectiveness and/or safety of a single, adjunctive or combined CAM treatment were included. For the classification of CAM treatments the definition of the US National Institutes of Health30 was followed. CAM interventions have to be compared against treatment as usual (TAU)/waiting list, placebo/sham or standard medical care.
- Types of outcomes: reviews were eligible if they assessed at least one measure of effectiveness such as severity of depressive symptoms, response rate (generally defined as a 50% decrease in depression scores after a period of up to 12 weeks of treatment),31 remission rate (generally defined as a period of up to 12 weeks during which a patient is asymptomatic or has only few symptoms to a very mild degree),41 relapse rates and/or a measure of safety such as number of adverse events (AEs), drug interactions or numbers needed to harm for study withdrawal due to side effects.

Search strategy
Electronic literature was systematically searched via PubMed, PsycInfo and Central from their inception to 31 January 2018, without restrictions regarding time or language. Search terms for CAM treatments were selected in accordance with Cochrane recommendations (table 1).42 Additional manual search included reference lists of previously published reviews14 29 30 43 and clinical practice guidelines.31-36 Using PubMed Informer,44 the search was updated until 30 June 2018.

Study selection process
To assess eligibility, articles were selected by screening titles and abstracts independently by two authors (HH and DA). Any abstract considered potentially eligible by at least one author was read in full to decide on its eligibility. Disagreements were rechecked with a third author (HC) until consensus was achieved.
Data extraction and quality assessment

Two authors (HH and DA) independently extracted data on the characteristics of the reviews including the type of the intervention, the year of publication, the number and quality of the original RCTs, the total number and age of the participants and effectiveness and safety outcomes. The quality of the included reviews was assessed using the Assessment of the Methodological Quality of Systematic Reviews (AMSTAR) tool. The AMSTAR tool consists of 11 items asking about important methodological quality criteria of systematic reviews such as: a published a priori design, duplicate study selection and data extraction, a comprehensive literature search including grey literature, a list of included and excluded studies, summarised characteristics and quality assessment of included studies, assessment of publication bias, appropriate method of data syntheses and deducing conclusions and a conflict of interests statement. AMSTAR has shown good construct validity and inter-rater reliability. The Intraclass Correlation Coefficient (ICC) of the total AMSTAR score of 11 points was reported as 0.84. For this analysis, the two authors (HH and DA) who independently assessed AMSTAR reached an ICC of 0.96. Disagreements regarding content or quality of the reviews were rechecked with a third author (HC) and resolved by agreement.

Data synthesis

Results were pooled qualitatively by type of the intervention. Outcomes had to be calculated as standard mean differences (SMDs), risk ratios (RRs), HRs or ORs. If meta-analyses displayed mean differences, SMDs were calculated using Review Manager Software (RevMan, V.5.3, The Nordic Cochrane Centre, Copenhagen, Denmark) for better comparability of the results. RevMan was also used to exclude SMDs/RRs/HRs/ORs of selective RCTs that did not fulfil eligibility criteria of this overview. Effect sizes were classified according to Cohen as SMDs of 0.2–0.49=small effect, SMDs of 0.5–0.79=medium effect and SMDs of 0.8 or more=large effect (absolute values) with higher reduction of/improvement in depression scores represented by more negative SMDs or RRs/HRs/ORs of selective RCTs. Statistical heterogeneity between studies was assessed by the χ² test with a p value of ≤0.1 indicating significant heterogeneity. The magnitude of heterogeneity was categorised by the I² statistic with I² of 0%–24%=no heterogeneity, I² of 25%–49%=moderate heterogeneity, I² of 50%–74%=substantial heterogeneity and I² of 75%–100%=considerable heterogeneity.
Quality of evidence

The quality of evidence was assessed according to the Grades of Recommendation, Assessment, Development, and Evaluation (GRADE) approach\textsuperscript{49} individually by two authors (HH and DA). Disagreements were rechecked with a third author (HC) until consensus was achieved. For each outcome, the evidence can be graded as high, moderate, low or very low. Evidence from RCTs is initially assessed as high, but can be downgraded by one level for serious or two levels for very serious limitations of the study quality (of both RCTs and meta-analyses), inconsistency of the results, indirectness of the evidence, imprecision of the results and a potential risk of publication bias (as assessed by the included meta-analyses).\textsuperscript{49}

RESULTS

Study selection

A total of 3582 potentially eligible articles were identified by electronic database search. One additional review was retrieved from manual search,\textsuperscript{50} one from the updated search until June 2018.\textsuperscript{51} After removing duplicates, 2639 articles were excluded by screening of titles and abstracts. The remaining 117 articles were read in full, of which further 91 reviews had to be excluded (figure 1). Reasons for exclusion comprised 55 reviews where newer and/or more comprehensive reviews on higher quality evidence were available.\textsuperscript{50–104} Further 15 reviews have to be excluded as they systematically summarised evidence but did not perform a meta-analysis mostly due to clinical heterogeneity or a limited number of available RCTs.\textsuperscript{105–119} Eight reviews were excluded as they included mixed depressive samples of bipolar or postpartum cases and did not provide (data for) subgroup analyses for patients with the defined depression criteria.\textsuperscript{120–127} Another six reviews contained community samples with non-clinical depression or physically ill patients with comorbid depressive symptoms but displayed no data for patients with a primary diagnosis of depression.\textsuperscript{128–133} Four reviews performed meta-analyses on both RCTs and non-RCTs and did not perform subgroup analyses or extracted sufficient data for post hoc analyses.\textsuperscript{134–137} Three of the reviews analysed standard instead of complementary therapies and were therefore excluded. Finally, 26 meta-analyses could be included and reviewed.\textsuperscript{138–163}

Review characteristics and quality

Characteristics and quality appraisal of the included meta-analyses are summarised in the online supplementary table 1. Meta-analyses were conducted between 2002 and 2018 and included between 1 and 49 RCTs on 40–7104 adult patients. Meta-analyses on children and adolescents were not available or did not meet inclusion criteria. Samples mostly consisted of patients suffering from major depressive disorder\textsuperscript{140–143} 145–151 154 156 157 159 160 but also included patients with mixed diagnoses of non-seasonal depression,\textsuperscript{138 153 162 163} patients with a diagnosis of seasonal depression\textsuperscript{132} and patients with mild to severe symptoms of depression above a clinical cut-off.\textsuperscript{139 141 144 145 151 155 157 158} All but one meta-analysis\textsuperscript{141} reported pooled outcomes based
on common standardised questionnaires or diagnostic interviews. Effects were analysed mostly up to 12 weeks of treatment (short term), except for four meta-analyses that included RCTs with effects reported up to 16, 24 or 32 weeks and further three meta-analyses with long-term analyses equal to or greater than 1 year. The AMSTAR total scores of the included meta-analyses ranged between 4 and 11 points with a median quality of 7 points. The individual AMSTAR ratings are reported in the online supplementary table 2.

**Synthesis of results**

**Acupuncture**

**Manual acupuncture**

A high-quality Cochrane review meta-analysed 49 RCTs in major depressed adults as well as those with clinically relevant symptoms of depression for manual acupuncture. For depression severity, significant effect sizes were found in comparisons with TAU and as in adjunction to standard antidepressants, while acupuncture showed similar effects to invasive sham acupuncture and standard antidepressants (figure 2). The analyses of remission rates did not reveal superiority of acupuncture in the comparisons with TAU, invasive sham, standard antidepressants or in adjunction to standard antidepressants (figure 3). AEs reported in the acupuncture groups were significantly lower than in patients treated with antidepressant drugs. However, most meta-analyses showed significant heterogeneity, a lack of RCTs with low risk of bias and a possibly serious risk of publication bias. Thus, the quality of evidence had to be downgraded to low and very low.

**Electroacupuncture**

For electroacupuncture, the same Cochrane review revealed very low quality of evidence for the comparisons with TAU and invasive sham for both outcomes, severity

<table>
<thead>
<tr>
<th>Quality of evidence</th>
<th>Intervention</th>
<th>Control</th>
<th>Reference</th>
<th>Trials</th>
<th>Participants</th>
<th>Std. Mean Difference (95% CI)</th>
<th>Std. Mean Difference (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Moderate</td>
<td>St. John’s Wort</td>
<td>Placebo</td>
<td>Apaydin, 2016</td>
<td>16</td>
<td>2888 89%</td>
<td>-0.49 [-0.74, -0.23]</td>
<td></td>
</tr>
<tr>
<td>Low</td>
<td>Electroacupuncture</td>
<td>ADM</td>
<td>Smith, 2018</td>
<td>10</td>
<td>995 33%</td>
<td>-0.28 [-0.47, -0.09]</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Manual acupuncture</td>
<td>TAU</td>
<td>Smith, 2018</td>
<td>5</td>
<td>274 53%</td>
<td>-0.94 [-1.16, -0.51]</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Saffron</td>
<td>TAU</td>
<td>Hauensedls, 2013</td>
<td>2</td>
<td>170 0%</td>
<td>-1.06 [-1.45, -0.66]</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Chinese herbs</td>
<td>TAU</td>
<td>Yeung, 2014</td>
<td>4</td>
<td>251 44%</td>
<td>-1.27 [-1.57, -0.97]</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Light therapy</td>
<td>TAU</td>
<td>Martensson, 2015</td>
<td>8</td>
<td>197 82%</td>
<td>0.17 [0.12, 0.24]</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Music therapy</td>
<td>TAU</td>
<td>Zha, 2015</td>
<td>5</td>
<td>244 76%</td>
<td>-0.57 [-1.03, -0.11]</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Faith-adapted CBT</td>
<td>TAU</td>
<td>Anderson, 2015</td>
<td>6</td>
<td>304 82%</td>
<td>-0.46 [-1.17, 0.25]</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Omega-3</td>
<td>Placebo</td>
<td>Appleton, 2015</td>
<td>25</td>
<td>1373 59%</td>
<td>0.30 [0.50, 0.10]</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Probiotics</td>
<td>Placebo</td>
<td>Huang, 2016</td>
<td>21</td>
<td>142 72%</td>
<td>-0.54 [-1.58, 0.48]</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Folate</td>
<td>Placebo</td>
<td>Galizia, 2016</td>
<td>1</td>
<td>73 60%</td>
<td>-0.69 [-1.06, -0.32]</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Vitamin D</td>
<td>Placebo</td>
<td>Shaffer, 2014</td>
<td>2</td>
<td>149 0%</td>
<td>-0.89 [-1.19, -0.60]</td>
<td></td>
</tr>
</tbody>
</table>

**Figure 2** Quality of evidence for depression severity, ADM, antidepressant medication; CBT, cognitive–behavioural therapy; I², heterogeneity; MBCT, mindfulness-based cognitive therapy; MBSR, mindfulness-based stress reduction; n.r., not calculable because of only one included randomised controlled clinical trials; n.r., not reported; SAmE, S-adenosyl methionine; TAU, treatment as usual.


5
(figure 2) and remission (figure 3), because of serious limitations of the quality of the RCTs, imprecision and a high risk of publication bias. For electroacupuncture monotherapy in comparison with standard antidepressants, low-quality evidence homogeneously suggested significant greater effects for severity and similar effects for remission. As an adjunctive treatment to antidepressants, electroacupuncture effectiveness was supported by low quality of evidence showing a significant greater consistent and precise effect for depression severity. Although the mean adjunctive effect can be considered as large, the analysis based on only one RCT with overall low risk of bias and four RCTs of lower methodological quality that missed to include adjunctive sham acupuncture. For remission rates, very low quality of evidence suggested no effects in adjunction to antidepressants. In addition, one RCT showed significant less AEs when electroacupuncture was added to standard antidepressants.

Aromatherapy
The literature search revealed no meta-analysis on aromatherapy. A recent systematic review detected no RCTs in patients with a primary diagnosis of depression. However, two out of five of the reviewed studies on inhalation aromatherapy and five out of eight studies on aromatherapy massage have found significant antidepressive effects in mixed patient samples and healthy adults.116

Biofeedback
No meta-analysis on biofeedback was conducted to date. A recent systematic review revealed only one RCT on the defined inclusion criteria for depression showing some effects in contrast to sham psychotherapy.117

Herbs
St. John’s wort (Hypericum perforatum)
The effectiveness of St. John’s wort was meta-analysed by a Cochrane review of 29 RCTs150 and a more recent, higher quality meta-analysis of 35 RCTs.142 In comparison with placebo, St. John’s wort showed moderate quality evidence of significant greater reductions of depression severity (figure 2) and response rates (figure 4). The evidence had to be downgraded due to significant heterogeneity because of higher effects in studies from German-speaking countries than in those from the USA or other European countries. In contrast, very low quality of evidence suggested no superiority to placebo for remission (figure 3) and relapse rates (figure 5). In
comparison with standard antidepressants, St. John’s wort showed comparable severity reductions, response, remission, and relapse rates. The quality of the evidence of the meta-analyses of severity and relapse rates had to be downgraded to low and very low, respectively. The evidence of the response and remission rates was considered as moderate quality showing the same results in both German and studies from other countries but containing some RCTs with unclear risk of selection bias and detection bias. Moreover, both meta-analyses showed similar AEs of St. John’s wort to placebo but significant less AEs than standard antidepressants.

Saffron (Crocus sativus)
A moderate-quality meta-analysis examined the effectiveness and safety of saffron on depression severity by including five RCTs in adult patients with major depression. It revealed very low quality of evidence for significant greater effects versus placebo and similar effects versus antidepressant medication up to 8 weeks of treatment (figure 2). No serious AEs were reported, but patients receiving saffron tend to report more AEs than those receiving placebo and less AEs than those receiving antidepressant medication. Reasons for downgrading the evidence included no replication of the results (all included RCTs were conducted by the same research group), the small overall sample size and the possibly high risk of publication bias.

Curcumin (Curcuma longa)
For the intake of curcumin, a moderate-quality meta-analysis revealed very low quality of evidence suggesting a small but significant short-term effect of low heterogeneity on depression severity by pooling six RCTs (figure 2). No serious AEs were recorded. Evidence had to be downgraded due to unclear risk of selection, performance, detection, attrition and reporting bias in over the half of the included RCTs, the imprecision of the pooled effect and the possibly high risk of publication bias.

Traditional Chinese herbs
A comprehensive but low-quality systematic review of 296 RCTs of Chinese herbal medicine formulas and single herbs revealed 21 RCTs of mostly unclear to high risk of selection, performance and detection bias and a serious risk of publication bias. Therefore, the evidence supporting the superiority above placebo and the similarity towards standard antidepressants regarding depression severity (figure 2) and response rates (figure 4) was assessed as very low.

Other herbs
For other than the described herbs, no meta-analyses were conducted to date. However, a systematic review found three single RCTs that showed significant improvement in depressive symptoms for Lavandula angustifolia as an adjunctive treatment to standard antidepressant drugs versus antidepressant drugs alone and for Echium amoenum and Rhodiola rosea versus placebo. No serious AEs were reported.

Homoeopathy
No meta-analysis on homoeopathic remedies for depression were conducted yet. A recent systematic review detected no placebo-controlled RCTs in patients with a primary diagnosis of depression.

Hypnosis
No meta-analysis on hypnosis or self-hypnosis techniques met the inclusion criteria of this overview. The only available review on this topic included six RCTs among which only one RCT included adults with mild primary depression. Within the mixed sample of physically ill patients and healthy adults, (self-)hypnosis appeared to be effective in decreasing depressive symptoms.

Light therapy
A high-quality Cochrane review meta-analysed the effects of bright light therapy in adjunction to standard antidepressants versus sham light therapy plus antidepressants on severity and response rates in patients suffering from non-seasonal depression. By pooling 18 RCTs of overall unclear risk of bias, it revealed very low quality of evidence for a significant small but inconsistent and imprecise effect on depression severity (figure 2). A subgroup analysis of two RCTs with low risk of selection bias and detection bias revealed a significant large effect on depression severity but based on one non-peer-reviewed publication and one RCT that also included bipolar patients. Response rates did not significantly differ between

<table>
<thead>
<tr>
<th>Quality of evidence</th>
<th>Intervention</th>
<th>Control</th>
<th>Reference</th>
<th>Trails</th>
<th>Participants</th>
<th>$I^2$</th>
<th>Risk/Odds/Hazard Ratio 95% CI</th>
<th>Risk/Odds/Hazard Ratio 95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Moderate</td>
<td>MBCT</td>
<td>ADM</td>
<td>Kuyken, 2016</td>
<td>4</td>
<td>669</td>
<td>0%</td>
<td>HR: 0.77 [0.60, 0.98]</td>
<td></td>
</tr>
<tr>
<td>Very low</td>
<td>St. John’s W</td>
<td>Placebo</td>
<td>Apaydin, 2016</td>
<td>1</td>
<td>428</td>
<td>n.c.</td>
<td>RR: 0.70 [0.49, 1.02]</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>ADM</td>
<td>Apaydin, 2016</td>
<td>1</td>
<td>241</td>
<td>n.c.</td>
<td>RR: 4.17 [0.47, 33.33]</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Folate</td>
<td>Adjunctive</td>
<td>Almeida, 2015</td>
<td>1</td>
<td>153</td>
<td>n.c.</td>
<td>OR: 0.33 [0.12, 0.94]</td>
<td></td>
</tr>
</tbody>
</table>

Figure 5 Quality of evidence for depression relapse rates. ADM, antidepressant medication; $I^2$, heterogeneity; MBCT, mindfulness-based cognitive therapy; n.c., not calculable because of only one included RCT; RR, risk ratio; TAU, treatment as usual.
groups (figure 4). AEs were reported non-systematically but appeared to be comparable with sham light therapy except for hypomania that occurred more often under verum light therapy.161

For patients with seasonal patterns of depression, a meta-analysis of eight RCTs152 revealed very low quality of evidence for a significant medium effect on depression severity of light monotherapy in comparison with sham light therapy (figure 2). Risk of bias of individual RCTs, heterogeneity and safety were not analysed leading to an overall low quality of the meta-analysis and downgrading of the evidence.

**Massage therapy**
The literature search detected no meta-analysis of *massage therapy* in patients with a primary depression. However, massage therapy appeared to be effective in decreasing depressive symptoms in mixed samples of physically ill patients and healthy adults.139 Future research will show whether these results may be transferable to primary depressed cases.

**Meditative movement therapies**

**Dance therapy**
Short-term effects of improvisatory or structured dance therapy as a combination of movement-based work, interactive group components and insight/expressive methods were meta-analysed by a Cochrane review of high methodological quality.155 It revealed a significant large pooled effect size for depression severity as an adjunctive treatment option to standard medical/psychotherapeutic care based on two RCTs (figure 2). Although the meta-analyses contained no heterogeneity and no imprecise CI, the evidence had to be downgraded to low quality because of mostly unclear/high risk of bias of one of the RCTs as well as the overall small sample size.

**Chinese movement therapies**
A low-quality meta-analysis on Chinese meditative movement therapies revealed 30 RCTs, of which 2 RCTs on Qi Gong and 5 RCTs on Tai Chi met inclusion criteria for patients with mild to severe symptoms of primary depression.151 Very low quality of evidence suggested significant short-term effects for Qi Gong but not for Tai Chi in comparison with TAU. The evidence had to be downgraded due to very serious limitations of the quality of the RCTs and the meta-analysis, significant heterogeneity, imprecision and a possible high risk of publication bias.

**Yoga**
A high-quality meta-analysis of complex yoga interventions for various depressive disorders found 12 RCTs145 of which 5 RCTs of mostly unclear risk of bias met the inclusion criteria of this overview. The pooled short-term effect on depression severity was of large size in comparison with TAU and medium size in comparison with standard exercises (figure 2). However, the evidence was assessed as very low due to serious limitations of quality of the RCTs, significant heterogeneity, imprecision and a possible high risk of publication bias.

A further systematic review of yoga in major depressive disorder, revealed five newer yoga RCTs but did not perform a meta-analysis because of high clinical heterogeneity. Risk of bias was comparably high and evidence mostly conflicting.106

**Mindfulness-based interventions**

**Mindfulness-based cognitive therapy (MBCT)**
A low-quality meta-analysis of mindfulness-based interventions in patients with major depression found four RCTs investigating the effects of MBCT and cognitive–behavioural therapy (CBT) on depression severity.106 It revealed a significant large short-term effect of MBCT in comparison with TAU and similar effects in comparison with CBT (figure 2). However, the quality of the evidence was considered as very low due to the missing risk of bias assessment, inconsistency and imprecision.

A further moderate-quality systematic review on MBCT meta-analysed 9 RCTs on an individual patient data level.149 The sample consisted of patients with recurrent major depression currently in remission. After a period of 60 weeks, MBCT showed a significantly reduced risk of depressive relapse compared with those receiving antidepressant drugs (figure 5). No serious AEs were reported. The evidence was assessed as moderate due to a possible risk of publication bias.

**Mindfulness-based stress reduction (MBSR)**
RCTs of MBSR and MBSR-like interventions were meta-analysed by a recent review144 showing a significant large short-term effect on depression severity in comparison with TAU and enhanced TAU (figure 2). The quality of evidence was assessed as low because of the overall unclear risk of selection and performance bias and significant heterogeneity.

**Music therapy**
Studies on active and receptive music therapy in older patients with a diagnosis of depression were summarised by a recent moderate-quality meta-analysis.163 Out of 19 RCTs, 8 met the inclusion criteria for this overview. Pooled analyses of five of them revealed a significant medium effect size on depression severity against TAU up to 52 weeks, however with bigger short-term than long-term effects, considerable heterogeneity and overall unclear risk of selection, performance and detection bias resulting in very low quality of evidence. Further three RCTs of the same review revealed low-quality evidence for a significant large consistent and precise effect of music therapy as an adjunctive treatment to antidepressants (figure 2).

A newer Cochrane review139 found eight different RCTs showing a significant large pooled effect of music therapy on depression severity against TAU and similar effects as CBT (figure 2). However, both analyses revealed very low quality of evidence due to mostly unclear selection,
performance, detection and reporting bias, significant heterogeneity and imprecision.

Nutrition therapy
No meta-analyses on specific diets for patients with depression were published to date. A systematic review of 11 RCTs on whole-diet interventions in mostly healthy patients with subthreshold physical conditions revealed conflicting evidence on the effectiveness of those interventions for the reduction of depressive symptoms. A further systematic review on fasting in patients with chronic pain and inflammatory diseases included one RCT and seven observational studies, which showed promising short-term but questionable longer term antidepressive effects.

Religious/spiritual interventions
Very low to low quality of evidence was found by a moderate-quality systematic review of nine RCTs on Christian, Muslim and spiritual CBT adaptations. The analyses showed significant greater medium effects on depression severity against TAU and standard CBT (figure 2). Safety data were not reported.

Supplements
Inositol
A low-quality meta-analysis of two RCTs in patients with major depression revealed very low quality evidence for inositol as adjective to standard antidepressants versus placebo in combination to standard antidepressants (figure 2).

Magnesium
No meta-analysis of magnesium supplementation was found. A recent systematic review detected no RCTs in patients with a primary diagnosis of depression.

Omega-3 fatty acids
A high-quality Cochrane review of 26 RCTs found conflicting evidence of the effectiveness of supplementation with omega-3 fatty acids versus placebo in patients with major depression as depression severity significantly improved while response and remission rates did not so (figures 2–4). One additional RCT with a very low sample size showed similar effects of omega-3 fatty acids on severity and response rates in comparison with antidepressant drug treatment (figures 2 and 4). However, all meta-analyses were based on very low quality of evidence because of limitations of the study quality, significant heterogeneity, imprecision and a possibly high risk of publication bias.

Probiotics
The effectiveness of the supplementation with probiotics on depression severity was analysed by a moderate-quality meta-analysis of five RCTs, of which only one RCT with overall low risk of bias was carried out on patients with major depression. The analysis of the RCT revealed a significant medium but imprecise short-term effect in comparison with placebo (figure 2). This led to overall very low quality of evidence for probiotics supplementation.

S-adenosyl methionine (SAMe)
A high-quality Cochrane review of the effectiveness and safety of SAMe supplementation on depression severity revealed two RCTs of low risk of bias that showed no significant pooled effects for SAMe monotherapy versus placebo. One RCT, also of low risk of bias, showed a significant medium short-term effect as adjunctive to standard antidepressant medication, both for depression severity. Further five RCTs, which were rated as having overall unclear risk of bias, showed similar pooled effects of SAMe monotherapy on depression severity compared with standard antidepressant medication (figure 2). Original RCTs reported safety issues insufficiently. For all meta-analyses, the evidence was assessed as low to very low quality because of limitations of the study quality, heterogeneity, imprecision and a possibly high risk of publication bias.

Tryptophan
A moderate-quality Cochrane review found two RCTs investigating the effectiveness and safety of tryptophan supplementation on depression severity. Pooling the effects led to significant greater short-term response rates (figure 4) as well as significant more AEs in the tryptophan group than in the placebo group. The evidence was assessed as very low quality because of an unclear risk of detection, attrition and other bias, imprecision and a possible risk for publication bias.

Vitamins
For Vitamin B6, no meta-analysis was available. A systematic review without meta-analysis revealed two RCTs showing no significant effects when compared with placebo.

Two further moderate to high-quality meta-analyses examined the effects of vitamin B9 (Folate) for major depressive patients. While a Cochrane review calculated a significant medium effect size of folate intake as an adjunctive intervention to standard drug treatment on depression severity, a more recent review revealed non-significant differences on severity and response rates (figures 2 and 3). In contrast, a long-term combined intake of Vitamin B6, B9 and B12 was found to be effective for relapse prevention after remission of symptoms as a result of one RCT (figure 5). However, all comparisons were based on very low quality of evidence mostly due to significant heterogeneity, imprecision and possible high risk of publication bias.

Another moderate-quality meta-analysis revealed evidence of the effectiveness of vitamin D intake on depression severity in comparison with placebo. The analysis of the two included RCTs revealed a significant medium short-term effect in favour of vitamin D in major depressed patients up to 8 weeks (figure 2). The quality of evidence was rated as very low due to limitations of the study quality, missing values of heterogeneity,
imprecision, a high risk of publication bias as well as insufficient reporting of AEs.

**Zinc**

The effectiveness of zinc for major depression was meta-analysed by a low-quality review of three RCTs. It revealed a significant pooled short-term effect of medium size and low heterogeneity when zinc was taken as an adjunctive to standard antidepressant drug treatment (figure 2). However, the available evidence had to be assessed as very low as the meta-analysis did not perform risk of bias assessments and did not report AEs.

**DISCUSSION**

This systematic review provided a comprehensive overview of the evidence of CAM treatments for patients with a diagnosis or clinical symptoms of depression. Moderate-quality evidence suggested the efficacy, comparative effectiveness to standard antidepressants and safety of St. John’s wort on depression severity and response rates. For remission and relapse rates, the evidence was conflicting and of lower quality. Moreover, moderate quality evidence showed that MBCT was superior to standard antidepressant drug treatment for the prevention of depression relapse in patients with recurrent major depression. Low-quality evidence suggested significant greater effects in favour of electroacupuncture in comparison with standard antidepressants alone and in addition to standard antidepressants for depression severity. For remission rates, low quality evidence revealed comparable effects of electroacupuncture and standard antidepressants. Further significant greater effects, which are based on low-quality evidence, were found for MBSR versus TAU, music therapy in conjunction to standard antidepressants, faith-adapted CBT versus CBT, and SAMe versus standard antidepressants. Other treatments such as manual acupuncture, aromatherapy, biofeedback, herbs (crocus sativus, curcuma longa, traditional Chinese herbs, lavandula angustifolia, echium amoenum and rhodiola rosea), homoeopathy, hypnosis, bright light therapy, massage, meditative movement therapies (dance therapy, Qi Gong, Tai Chi and yoga), whole-diet interventions, fasting and supplementation with inositol, magnesium, omega-3 fatty acids, probiotics, tryptophan, B and D vitamins and zinc were based on very low quality of evidence or no level 1 evidence.

The strengths of the review process included the comprehensive literature search based on a structured list of CAM specific topics, which had been operationalised for the Cochrane Collaboration. It therefore included evidence for more than the previously considered CAM approaches and provided systematic information where further high-quality studies are required. In addition, we only included results of RCTs of patients with a diagnosis of depression or clinical relevant depressive symptoms and excluded RCTs on samples with minor or secondary symptoms of depression by newly calculating effect sizes. Finally, we rated AMSTAR and considered the quality of the meta-analyses as well when grading the quality of the evidence.

The conclusions derived from this overview are limited due to possibly missing evidence from newer RCTs, which have not been summarised by the included systematic reviews and meta-analyses. As it was not within the scope of this overview, we did not separately search for individual RCTs. We also did not include meta-analyses on studies of lower evidence levels, which may include bigger samples and may provide additional information about further possible treatment approaches. Moreover, we did not search online registries or conference proceedings for unpublished or ongoing meta-analyses, which may limit the conclusions. Another reason that limits the quality of evidence consists in the unsatisfactory methodological quality of some of the included meta-analyses. Although the methodological quality of the original RCTs might be acceptable, the bad reporting of some meta-analyses led to downgraded evidence. In particular, meta-analyses often missed to search for grey literature, cited excluded studies, adequately assess risk of bias of the original studies and report complete I^2 statistics. As the latter are known to be unstable in meta-analyses with a small numbers of studies, calculating confidence intervals for I^2 should be standard. Moreover, RCTs as well as meta-analyses often missed to systematically report on occurred AEs, which also limits the significance of the conclusions. In RCTs of non-pharmacological interventions, there is always a high or unclear risk of performance bias and possibly high placebo effects. As such, adding credible sham interventions, controlling for patients’ expectancies and performing of intention-to-treat analyses is indispensable. However, meta-analyses mostly did not systematically assess these issues. In meta-analyses of pharmacological interventions, the influence of industrial funding sources was often not adequately analysed. Here, subgroup analyses of studies having received no funding/ non-industrial funding versus those having received industrial funding are needed. Results of meta-analyses that missed to report funding issues completely should be interpreted with caution. In general, it should be noticed that all evidence is based on short-term pooled effects, except for meta-analyses of St. John’s wort, MBCT, music therapy and B and D vitamins that also provided longer term follow-up data.

Clinical recommendations for patients should follow the country-specific clinical practice guidelines considering the quality of evidence, the accessibility of the treatment, costs and the preferences of the patients. While the guidelines agree that clinicians should select between either CBT or second-generation antidepressant drugs for the treatment of major depression, the restricted search strategy of some of the guidelines might limit their recommendations for CAM treatments. For patients who do reject or do not tolerate standard antidepressant drugs, one alternative treatment option may be St. John’s wort. It is also recommended by the
American Psychiatric Association Task Force report and the CANMAT Depression Work Group as being proven sufficiently for the short-term by placebo-controlled and equivalence trials with standard antidepressants for mild to moderate major depression. Particularly for bridging the gap between diagnosis and getting access to psychotherapy and meanwhile reducing/not-worsening depression severity, St. John’s wort may be considered as a possibly better tolerated alternative to standard antidepressant drugs. As St. John’s wort is accessible without prescription and currently not regulated by the US Food and Drug Administration, we agree with the ACP guidelines that it remains difficult for patients to obtain quality-controlled remedies. Moreover, St. John’s wort is associated with numerous herb-to-drug interactions. Therefore, we would recommend clinicians to educate their patients about possible effects, side effects and interactions who in turn should not take St. John’s wort without professional advise. Despite those limitations, we would not discourage a general therapeutic attempt with St. John’s wort, even if we disagree with the NICE guideline in this point. Clinicians may also inform patients with recurrent major depression currently in remission about the superiority of MBCT in comparison with standard antidepressants for relapse prevention. Finally, patients should also be informed that many other CAM treatments might show promising effects but cannot be recommended until further higher quality studies will confirm their effectiveness and safety.

Further research is needed, particularly for interventions that have shown preliminary evidence for reducing secondary symptoms of depression, promising short-term but no longer term effects or insufficient evidence due to low methodological quality of the original RCTs and/or the performed meta-analyses. Reporting of clinical trials and meta-analyses should necessarily follow the Consolidated Standards of Reporting Trials (CONSORT) and PRISMA guidelines, respectively, including rigorous documentation and analysis of AEs. Especially Chinese and Indian trials are still found to be poorly reported and tended to present more positive conclusions than those from western countries. Moreover, seven of the included meta-analyses showed no more than poor methodological quality. All were published in peer-reviewed journals in the past 5 years. One complementary journal is among them, while six of the journals are conventional psychiatric journals with impact factors ranging from 2.419 to 4.369. Thus, particularly, the review process as well as the editorial work need to be improved. Further clinical practice guidelines should extend their search strategies and include standard search terms for CAM. This is also important for CAM therapies that do not show consistent evidence or that are not yet investigated. This information might be equally interesting for physicians as well as for patients to make an informed decision about the treatment for clinical depression.

CONCLUSION

This overview of systematic reviews on CAM treatments for clinical depression aimed to provide a systematic search strategy and evidence base, on which further clinical practice guidelines may build their recommendations. To improve quality of trials and meta-analyses, researchers, reviewers as well as editors are asked to ensure that future articles strictly adhere the CONSORT and PRISMA guidelines.

Contributors HH was responsible for the conception and design of the study, the collection and analysis of the study data and for drafting the manuscript. DA participated in the analysis of the study data and drafting the manuscript. HC participated in the conception and design of the study and the analysis of the study data and critically revised the manuscript. GD participated in the conception and design of the study and critically revised the manuscript. All authors approved the final manuscript.

Funding The authors have not declared a specific grant for this research from any funding agency in the public, commercial or not-for-profit sectors.

Competing interests None declared.

Patient consent for publication Not required.

Provenance and peer review Not commissioned; externally peer reviewed.

Data availability statement All data relevant to the study are included in the article or uploaded as supplementary information.

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, appropriate credit is given, any changes made indicated, and the use is non-commercial. See: http://creativecommons.org/licenses/by-nc/4.0/.

REFERENCES

Open access


59. Cui Y-H, Zheng Y. A meta-analysis on the efficacy and safety of St John’s wart extract in depression therapy in comparison with...


75. Martins JG. Epal but not DHA appears to be responsible for the efficacy of omega-3 long chain polynaturated fatty acid supplementation in depression: evidence from a meta-analysis of randomised controlled trials. *J Clin Endocrinol Metab* 2012;98:525–42.


