



Systematic review

Acupuncture for depression: An overview of systematic reviews

Tao Gao^{a,b}, Qianhua Zheng^a, Tinghui Hou^a, Yanan Luo^a, Yunzhou Shi^a, Ying Li^{c,*}^a Acupuncture and Tuina School, Chengdu University of Traditional Chinese Medicine, Chengdu, Sichuan, China^b Acupuncture and Tuina School, Shaanxi University of Chinese Medicine, Xi'an, Shaanxi, China^c Graduate school, Chengdu University of Traditional Chinese Medicine, Chengdu, Sichuan, China

ARTICLE INFO

Keywords:

Acupuncture

Depression

Methodological and reporting quality

Systematic review

ABSTRACT

Introduction: As current evidence on the effectiveness of acupuncture for depression is inconsistent, this overview aims to summarize and critically evaluate the evidence of relevant systematic reviews (SRs).

Methods: A comprehensive literature search was conducted using multiple databases for SRs on acupuncture for depression from their inception to December 2018. Two authors independently selected articles, collected data, and assessed the methodological and reporting quality of identified SRs according to revised Assessment of Multiple Systematic Reviews (AMSTAR 2) and Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA), respectively.

Results: Nine SRs were included in this overview. The items of AMSTAR 2 in most SRs were poorly reported, and only one SR achieved a decent overall rating and considered to be of moderate quality assessed by AMSTAR 2. A total of 7 SRs obtained a decent rating by PRISMA. No firm conclusions were made on the effectiveness of acupuncture for depression.

Conclusion: The methodological and reporting quality of SRs on acupuncture for depression was suboptimal. In future studies, more efforts are needed to improve the quality of SRs in this area.

1. Introduction

Depression is a common mental illness in modern society, defined as persistent sadness and loss of interest or pleasure in activities that you normally enjoy, accompanied by an inability to carry out daily activities, for at least two weeks [1]. People with depression generally have several of the following symptoms: depressed mood, decreased interest in people and activities, anxiety, loss of energy, reduced concentration, feelings of guilt and hopelessness, low self-worth, and even thoughts of self-harm or suicide [1]. In addition to mental symptoms, there are somatic symptoms, such as fatigue, sleep disturbance, poor appetite, pain, digestive irregularities, chest tightness, and dizziness. More than 300 million people have been living with depression in 2017, an increase of more than 18% between 2005 and 2015 [1]. The incidence of depression continues to increase worldwide, and the lifetime incidence of major depression disorder (MDD) has been estimated at 16.2% [2]. Depression is among the most common psychiatric disorders in adults and is related to a huge loss of quality of life for patients, substantial social and economic burdens [3–6]. MDD was identified as the second leading cause of disability and a major contributor to suicide according to 2010 Global Burden of Disease study [7]. Depression is the third most common burden of disease worldwide, and is expected to become

the leading burden of disease by 2030 [8]. The financial burden in England is estimated to exceed £9 billion annually, with approximately £370 million accounting for direct costs of treatment [9].

Currently, the worldwide conventional treatment for depression is antidepressant medication [10]. However, lack of effect in some individuals [11], intolerable side effects [12–14], delay in the onset of the antidepressant action [15], relapse [16], and withdrawal reactions [17,18] have hampered the clinical use of antidepressants. Psychotherapy has been shown to be beneficial and efficacious [19,20], and guidelines are increasingly advocating it as a treatment option [21], but it is not uniformly accepted [22]. Faced with the limitations of conventional treatments, depressive patients often seek complementary and alternative medicine, such as acupuncture, to alleviate symptoms and improve life quality [23]. Acupuncture is one of the most important therapies in traditional Chinese medicine, and historically has been used for emotional, psychological and spiritual disorders, such as anxiety, insomnia and depression, in China, Japan and Korea. In recent years, it has become one of the most popular complementary therapies in the west as well.

SRs of high-quality randomized controlled trials (RCTs) are considered to be the best evidence on specific healthcare interventions. Some SRs of RCTs concerning acupuncture described significant

* Corresponding author at: Chengdu University of Traditional Chinese Medicine, No. 37 Shi'er Qiao Rd, Chengdu, Sichuan 610075, China.

E-mail address: liyings@cdutcm.edu.cn (Y. Li).

beneficial effects in the treatment of depression [24,25], whereas others failed to demonstrate the effects. For example, Fan et al. [26] concluded efficacy of combination therapy of acupuncture and medicine remained uncertain due to inadequate study design, and some studies [27,28] suggested no significant difference on symptom relief of depression between acupuncture and sham acupuncture or no treatment. The contradictory results make it difficult to draw conclusions on the effectiveness of acupuncture. It is of significant importance to evaluate the quality of relevant SRs as poor quality of SRs diminishes their value to clinicians, policy makers, and other users. At present, no published report has presented such assessments for the methodological and reporting quality of SRs regarding the effectiveness of acupuncture for depression.

The main purpose of our study is to evaluate the methodological and reporting quality of SRs concerning the effectiveness of acupuncture for depression. Our second goal is to summarize the evidence of effectiveness and safety of acupuncture for depression.

2. Methods

2.1. Search strategy

A systematic literature search for SRs of acupuncture for depression was performed in the following databases: PubMed, Web of Science, Ovid-EMbase, the Cochrane Library, China National Knowledge Infrastructure (CNKI), Chinese Biomedical Literature Database (CBM), China Science and Technology Journal Database (VIP) and Wan Fang Database. All the online databases were scrutinized from their inception dates to December 2018. Search terms used were (acupuncture OR acupuncture therapy OR electroacupuncture) AND (depression OR depressive disorder OR depressive state OR dysthymia OR dysthymic disorder) AND (systematic review OR meta-analysis), with slight modifications for individual searches in each database. The detailed search strategy was available in Appendix A. The reference lists of these papers were also identified for further relevant articles. The language was restricted to English and Chinese. The search was independently conducted by two investigators and discrepancies were resolved by discussion.

2.2. Study selection

The SRs were selected according to the following aspects: (1) the studies included in SRs had to be RCTs or quasi-RCT (q-RCTs); (2) the diagnosis of depression should be based on at least one of the following criteria: depression defined by the Diagnostic and Statistical Manual of Mental Disorders (DSM-III, DSM-IV, or DSM-V), or the International Statistical Classification of Diseases and Related Health Problems (ICD), or the Research Diagnostic Criteria (RDC), or the Chinese Classification of Mental Disorders (CCMD-2 or CCMD-3); (3) the intervention treatment included manual acupuncture or electroacupuncture alone or in combination with other treatment (western medicine, usual care or psychotherapy). Sham acupuncture, no treatment (waiting list control) or other treatments were utilized as control interventions; (4) one or more of the following outcome measures were available: Hamilton Depression Rating scale (HAMD), Self-Rating Depression Scale (SDS), Hamilton Rating Scale for Depression (HRSD), Beck Depression Inventory (BDI), Side Effect Rating Scale (SERS) and efficacy rate; (5) studies with participants diagnosed with post-stroke depression, postpartum depression and other disease-related depression were excluded; (6) studies that presented repeated retrievals or duplicate publications were excluded, and the article with the highest quality was included; (7) review comments, guidelines, editorials were excluded. There was no restriction on gender, age, race, or region.

2.3. Quality assessment

The methodological and reporting quality of included SRs were rated according to AMSTAR 2 and PRISMA tool, respectively. AMSTAR 2 is a measurement tool to assess the methodological quality of SRs, with good validity, reliability and applicability [24]. PRISMA is a currently widely applied reporting guidelines of SRs. AMSTAR 2 is a 16-item questionnaire and PRISMA is a 27-item questionnaire, requiring assessors to answer “Yes”, “No”, or “Partial Yes”. The judgements for each item in the checklist of AMSTAR 2 and PRISMA are as follows: “Yes” is chosen in the condition that the item is answered completely correct and well-documented; “Partial Yes” is selected in the case that the item is answered correct, but with insufficient evidence; “No” is picked when relevant data is absent or evaluation is inappropriate. Two reviewers assessed these domains independently, and disagreements were resolved by discussion and consensus. Any unresolved discrepancies were judged by a third reviewer.

AMSTAR 2 is a recently updated instrument and has been applied in few studies. As an overall score may disguise critical flaws that should diminish confidence in the results of a SR, the AMSTAR 2 expert group suggested that this new tool items not be evaluated to generate an overall score like the original AMSTAR (A Measurement Tool to Assess Systematic Reviews). It is recommended that users adopt the rating process based on identification of critical domains. Seven items are regarded as critical domains which can critically affect the validity of a SR and its conclusions, consisting of item 2 (protocol registered before commencement of SR), item 4 (adequacy of literature search), item 7 (justification for excluding individual studies), item 9 (risk of bias from individual studies being included in the SR), item 11 (appropriate meta-analytical methods), item 13 (consideration of risk of bias when interpreting the results of the SR) and item 15 (assessment of presence and likely impact of publication bias). General standard for evaluation of overall confidence of SR is introduced as follows: SR is rated as high confidence if it has no or one non-critical weakness; moderate confidence if it has more than one non-critical weakness; low confidence if it has one critical flaw, with or without non-critical weaknesses; critically low confidence if it has more than one critical flaw, with or without non-critical weaknesses.

At present, there is another practical method to summarize the results of methodological assessment using AMSTAR 2, that is by calculating the percentage of items with “yes”. Details are as follows: percentage of items with “yes” > 50% is considered to have a decent overall rating; percentage of items with “yes” > 80% is considered to be adequately reported; percentage of items with “yes” < 50% is considered to be poorly reported. The aforementioned methods of calculating the percentage of items with “yes” is also applicable to evaluation of PRISMA.

2.4. Data extraction

The following information for each SR were extracted: title, first author's name, country of first author, year of publication, study type, number of primary studies and patients, searching databases and terms, language limitations, interventions, outcomes, methodological evaluation tool and main conclusions, and whether STRICTA (Standards for Reporting Interventions in Controlled Trials of Acupuncture) or CONSORT (Consolidated Standards of Reporting Trials) statement were indicated in the RCTs of the included SRs. Data were extracted independently by two authors using a specifically designed data extracted form. Any disagreements were resolved by discussion or a third author was consulted if necessary.

2.5. Data analysis

A narrative description of the included studies was provided, and summary statistics were presented with percentages and frequencies. In

In addition, statistical analysis was conducted using SPSS 19.0 Software (SPSS Inc., Chicago, IL, USA). Two-sided $P < 0.05$ indicated a statistical significance. We calculated the inter-rater reliability between the two assessors for AMSTAR 2 and PRISMA using Kappa index. Kappa index of less than 0.4 indicated poor agreement, 0.4 to less than 0.75 suggested fair agreement, and 0.75 or higher reflected excellent agreement.

3. Results

3.1. Selection of the systematic reviews

The initial database search yielded a total of 803 potential relevant articles, including 213 from PubMed, 143 from Web of Science, 322 from Ovid-EMbase, 13 from the Cochrane Library, 14 from CNKI, 17 from CBM, 18 from VIP, and 36 from Wan Fang Database. Two-hundred and eighty-four records were removed for duplicate retrievals, and the titles and abstracts of the remaining 519 records were reviewed. Four-hundred and ninety-four records were excluded for reasons of irrelevance. Twenty-five potentially eligible articles were downloaded for full-text assessment. Sixteen studies were excluded, and the reasons for excluded reports were as follows: the studies included in 7 SRs were not RCTs or q-RCTs, 3 articles provided insufficient data (lack of meta-analysis or quantitative synthesis of results in the SR), 5 studies were not SRs, and 1 article was duplicated report. The detailed information for excluded studies were shown in Appendix B. Finally, 9 studies were included in our study [24–32]. The process of literature retrieval and screening is presented in Fig. 1.

3.2. Characteristics of the systematic reviews

Six SRs [26–31] were published in English-language journals, and the remaining 3 [24,25,32] were in Chinese-language. One Cochrane review [31], 7 peer-reviewed articles [24–30] and 1 dissertation [32] were included in our research. Seven SRs [24–29,32] originate from China, one [30] from Taiwan and another [31] was accomplished by scholars from Australia, reported from 2008 to 2018. The number of

primary studies ranged from 6 to 64, and participants varied from 431 to 7104 in each SR. In regard to searching databases, 6 SRs searched the Cochrane Library [24,27,27,28,29,30,31] and PubMed [24,26,28,30,31,33], 5 [24–26,29,32] searched CNKI and Wan Fang Data, 4 searched EMbase [24,27,28,31], VIP [24–26,32] and Medline [27,25–29,31], and 2 searched Korean databases [28,31]. The study design of the trials included in the SRs were all RCTs, except that in 2 studies were RCTs or q-RCTs [26,32]. As to the intervention measures, 5 studies [24–28] used acupuncture therapy alone, 4 studies [29–32] used acupuncture combined with western medicine. Western medicine was applied as controls in 5 studies [24–26,30,32], and other controls included sham acupuncture, no treatment and psychological therapy. Different primary outcomes were adopted, while the most frequently used was HAMD. Others include efficacy rate, BDI, SDS, SERS, TESS and HRSD. Methodological evaluation tool was adopted in all the studies, 6 of which used Jadad score [34], and 3 of which used Cochrane Handbook. Among the 3 studies using the Cochrane Handbook, the specific assessment by Stub et al. included allocation concealment and blinding [35], and the other two studies [24,31] assessed random sequence generation, allocation concealment, blinding of participants and personnel, blinding of outcome assessment, incomplete outcome data, selective outcome reporting and other bias according to Cochrane Handbook for Systematic Reviews of Interventions 5.1.0 [36]. The items reporting percentage for individual RCT using Consort and STRICTA guidelines were identified in 1 SR [30]. Consort standard guidelines were followed to assess the quality of included RCTs in 1 SR [28]. In all the conclusions of the included SRs, there was no agreement on whether acupuncture was effective. The characteristics details of included SRs are listed in Table 1.

3.3. Quality of the systematic reviews

3.3.1. Methodological quality

A decent overall rating with AMSTAR 2 was achieved in only 1 SR [31], while the items of other 8 SRs were poorly reported. Seven SRs (77.8%) [24,25,27–31] included the components of PICOS (population, intervention, comparator group, outcome and study design) in the

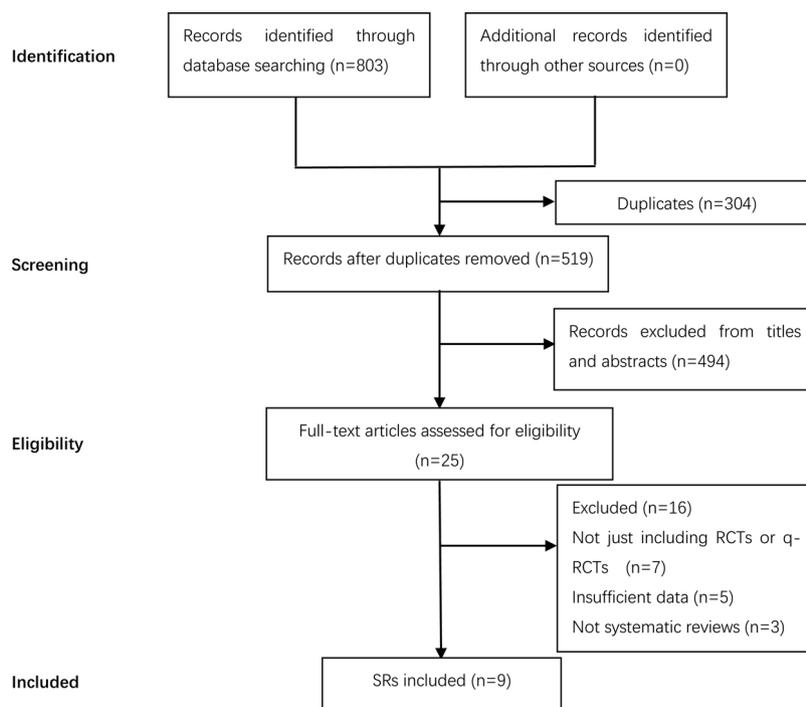


Fig. 1. Flow chart of literature selection.

Table 1
Characteristics of systematic reviews.

Study	Country	No. of primary studies (patients)	Gender	Diagnosis	Study types	Intervention measures		Primary outcome(s)	Methodological evaluation tool	Main conclusions
						Treatment group	Control group			
Wang 2008	China	8(477)	Not mentioned	Major depressive neurosis	RCT	Acupuncture	Sham acupuncture	HAMD, BDI, response rate	Jadad score	Acupuncture was an effective treatment that could significantly reduce the severity of disease in the patients with depression.
Zhang 2016	China	6(431)	No limitations	Primary depression	RCT	SSRIs + electroacupuncture	Electroacupuncture	SDS, HAMD, SPERS, TESS	Jadad score	The combination of electroacupuncture with SSRI in the treatment of depression elicits a faster response onset and better efficacy than SSRI medications.
Chan 2015	Taiwan	13(1046)	Not Mentioned	Depression/ various depressive conditions	RCT	Acupuncture + antidepressant medication	Antidepressant medication	HAMD, response rate	Jadad score	Combined acupuncture and antidepressant treatment is more effective than antidepressants alone in the first 6 weeks of treatment.
Stub 2011	China	26(2173)	Not Mentioned	Depression/ depression neurosis	RCT	Acupuncture/ electroacupuncture/ laser acupuncture	Placebo/ no treatment/ standard medication/ psychotherapies/ other standard care	HRS	Cochrane Handbook	Acupuncture is effective in reducing severity of depression and that TCM- and electro acupuncture may have similar effect as current usual care.
Smith 2018	Australia	64(7104)	No limitations	Depression	RCT	Manual acupuncture/ electroacupuncture/ laser acupuncture/ acupuncture + medication	No treatment/ control acupuncture/ antidepressants/ psychological therapies	BDI, HAMD	Cochrane Handbook	The reduction in severity of depression was less when acupuncture was compared with control acupuncture than when acupuncture was compared with no treatment control. Effects of acupuncture versus medication and psychological therapy are uncertain.
Fan 2010	China	20(2787)	No limitations	Depression/ depressive disorders	RCT q-RCT	Acupuncture alone or combined with western medicine	Western medicine	Efficacy rate	Jadad score	The efficacy of acupuncture was similar to western medicine for treating depression. However, efficacy of combination therapy remains uncertain.
Shen 2014	China	13(884)	No limitations	Depression diagnosed with ICD-10PHC; HAMD ≥ 17	RCT	Acupuncture	Western medicine	Efficacy rate	Jadad score	The efficacy of acupuncture was superior to western medicine for treating depression.
Xiong 2009	China	9(903)	No limitations	Depression neurosis	RCT	Acupuncture	Western medicine	Efficacy rate, HAMD, SDS	Cochrane Handbook	Acupuncture is not inferior to western medicine, and acupuncture is associated with few adverse reactions.
Lv 2014	China	16(1350)	Not mentioned	Depressive disorder	RCT q-RCT	Acupuncture alone or combined with western medicine	Western medicine	Efficacy rate	Jadad score	Acupuncture, electroacupuncture, acupuncture combined with drugs are more effective than drug, but there are differences between the specific groups.

inclusion criteria. Only 1 review (11.11%) [31] reported that a protocol was established prior to its implementation. Merely 1 review (11.11%) [31] provided its explanation on study design for inclusion. Eight reviews (88.89%) [24–30,32] failed to achieve a comprehensive literature search. Majority of the reviews (77.78%) [24–26,28,25–30,32] did not provide a list of excluded studies or justify the exclusions. Only 2 reviews (22.22%) [29,31] described PICOS of the included studies in adequate detail. Four reviews (44.44%) [24,26,27,31] performed duplicate study selection, and 5 reviews (55.56%) [24,25,27,29,30] conducted duplicate data extraction. Eight reviews (88.89%) adopted appropriate statistical methods to combine the study findings. Only 3 reviews (33.33%) [25,26,31] assessed the likelihood of publication bias and its possible impact on the results of the review. When heterogeneity was observed in the results of review, 6 SRs (66.67%) [25–27,29,31,32] investigated source of the heterogeneity and discussed its impact on the results of the review. Three reviews (33.33%) [24,28,31] used a satisfactory technique for assessing the risk of bias (RoB) in randomized trials included in the review, only 1 review [31] (11.11%) assessed its potential impact on the results of meta-analysis, 3 reviews (33.33%) accounted for it when interpreting the results of the review. Five SRs (55.56%) [26,28,28,29,30,31] stated no competing conflict of interest of the SR itself, but none of them reported conflict of interest for individual study included in a SR. Of all the items, only one item on statistical methods to combine results was adequately reported. Nevertheless, there were still eleven items which were poorly reported, including priori protocol, comprehensive search strategy, duplicate study selection, interpretation of study design and excluded studies, detailed characteristics of included studies, conflict of interest of studies included in SR, investigation of publication bias, technique for assessing the RoB in individual studies, assessment of the potential impact of RoB in individual studies on the meta-analysis results as well as discussion of it on the results of review. The details can be found in Table 2 and Fig. 2.

According to the evaluation criteria of the AMSTAR 2, 8 SRs had more than one critical flaws, considered to be of critically low quality, and only 1 SR [31] have 2 non-critical weakness, rated as moderate quality.

The inter-rater reliability was excellent between the two assessors

(Kappa = 0.895). Disagreements were found in some domains after comparison, and more details were shown in Appendix C and D. Finally, Disagreements were settled by discussion and by input from a third author.

3.3.2. Reporting quality

No review fulfilled all 27 items of PRISMA, but a total of 7 reviews (77.78%) [24,26–31] achieved a decent overall rating with PRISMA. None of the reviews provided a highly structured abstract and protocol registration information. Most (88.89%) [24–30,32] mentioned the terms “systematic review” or “meta-analysis” in the title. In the introduction part of review, simply 3 (33.33%) reviews [29,27–31] described the rationale for review well, and only 1 (11.1%) review [31] gave an explicit statement of questions being addressed concerning PICOS. In the methods section, two (22.22%) reviews [24,31] reported carefully and completely defined eligibility criteria, most (88.89%) reviews [24,26–32] described all the information sources, merely one (11.1%) [31] presented full electronic search strategy for at least one major database, more than half stated study selection process (55.56%) [24,26,27,30,31], data extraction methods (66.67%) [24,25,27,29,27–31] and items (55.56%) [25,27,30,27–32], 4 (44.44%) reviews [24,28,30,31] described methods for assessing RoB of individual studies, most (88.89%) reviews [24,27–31] stated the primary outcome measures and methods to combine study results, 5 (55.56%) reviews [24–27,31] specified assessment of RoB across studies that may influence the cumulative evidence, 4 (44.44%) reviews [26,29,31,32] described methods of additional analysis. For the results part, most items of PRISMA checklists were reported appropriately, except for the assessment of RoB across studies. The main findings are summarized in 7 (77.78%) reviews [24,25,27,28,30,27–32]. All the reviews discussed the limitations appropriately and provided a general interpretation of results and implication for future research. Source of funding and role of funders were reported in 3 (33.33%) reviews [26,28,31]. Overall, seven items were adequately reported, including title, information source, summary measures, synthesis of results in the methods section, results of individual studies, synthesis of results, and conclusion. While there were still eleven items poorly reported, including structured abstract, rationale and objectives in the introduction

Table 2
Methodological quality assessment of systematic reviews by AMSTAR 2.

Study	Item 1	Item 2	Item 3	Item 4	Item 5	Item 6	Item 7	Item 8	Item 9	Item 10	Item 11	Item 12	Item 13	Item 14	Item 15	Item 16	of “yes” [n(%)]	Ranking of quality
Wang 2008	Y	N	N	PY	Y	Y	Y	PY	PY	N	Y	N	N	Y	N	N	6(37.50)	Critically low
Zhang 2016	Y	N	N	PY	N	Y	N	Y	PY	N	Y	N	N	Y	N	Y	6(37.50)	Critically low
Chan 2015	Y	N	N	PY	N	Y	N	PY	PY	N	Y	N	N	N	N	Y	4(25.00)	Critically low
Stub 2011	Y	N	N	PY	N	N	N	PY	Y	N	N	N	Y	N	N	Y	4(25.00)	Critically low
Smith 2018	Y	Y	Y	Y	Y	N	Y	Y	Y	N	Y	Y	Y	Y	Y	Y	14(87.50)	Moderate
Fan 2010	N	N	N	PY	Y	N	N	PY	PY	N	Y	N	N	Y	Y	Y	5(31.25)	Critically low
Shen 2014	Y	N	N	PY	N	Y	N	PY	PY	N	Y	N	N	Y	Y	N	5(31.25)	Critically low
Xiong 2009	Y	N	N	PY	Y	Y	N	PY	Y	N	Y	N	Y	N	N	N	6(37.50)	Critically low
Lv 2014	N	N	N	PY	N	N	N	PY	PY	N	Y	N	N	Y	N	N	2(12.50)	Critically low
of “yes” (n)	7	1	1	1	4	5	2	2	3	0	8	1	3	6	3	5		
of “yes” (%)	77.78	11.11	11.11	11.11	44.44	55.56	22.22	22.22	33.33	0	88.89	11.11	33.33	66.67	33.33	55.56		

Y: Yes; PY: Partial Yes; N: No.

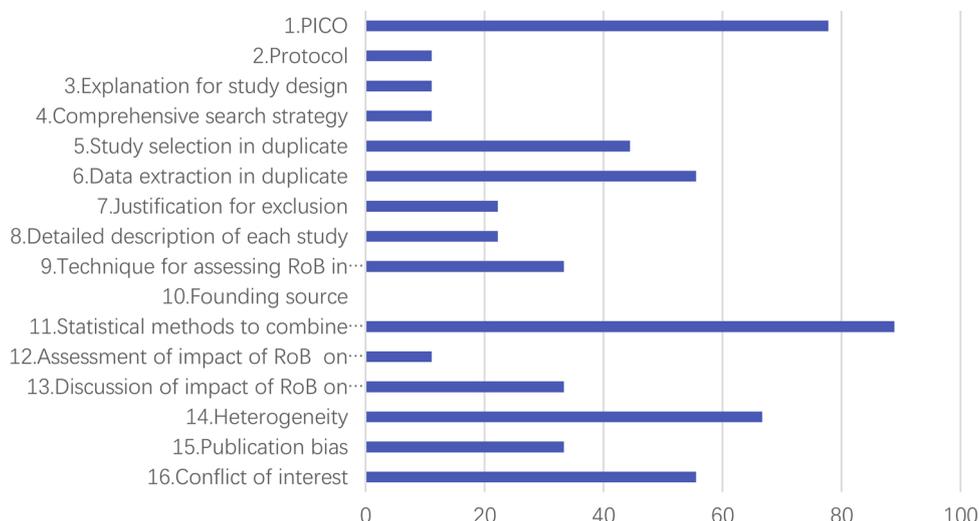


Fig. 2. Percentage of each AMSTAR 2 items with "Yes".

part, protocol and registration, eligibility criteria, search strategy, RoB in individual study, additional analysis in the methods part, RoB within studies in the results part, and funding. The corresponding results can be found in Table 3.

The inter-rater reliability was excellent between the two assessors (Kappa = 0.890). Disagreements were found in some items after comparison, and more details were shown in Appendix E and F. Finally, discrepancies were resolved by discussion and by consulting a third author.

Table 3
Reporting quality assessment of systematic reviews by PRISMA.

Item	Study									
	Wang2008	Zhang2016	Chan2015	Stub2011	Smith2018	Fan2010	Shen2014	Xiong2009	Lv2014	Of Yes [n (%)]
1	Y	Y	Y	Y	N	Y	Y	Y	Y	8(88.89)
2	PY	PY	PY	PY	PY	PY	N	PY	PY	0(0.00)
3	PY	Y	Y	PY	Y	PY	PY	PY	PY	3(33.33)
4	PY	PY	PY	PY	Y	PY	PY	PY	PY	1(11.11)
5	N	N	N	N	N	N	N	N	N	0(0.00)
6	PY	PY	PY	PY	Y	PY	PY	Y	PY	2(22.22)
7	Y	Y	Y	Y	Y	Y	PY	Y	Y	8(88.89)
8	PY	PY	PY	PY	Y	PY	PY	PY	PY	1(11.11)
9	Y	N	Y	N	Y	Y	N	Y	PY	5(55.56)
10	Y	Y	Y	N	Y	N	Y	Y	PY	6(66.67)
11	Y	PY	Y	PY	Y	PY	Y	N	Y	5(55.56)
12	PY	PY	Y	Y	Y	PY	PY	Y	PY	4(44.44)
13	Y	Y	Y	Y	Y	Y	Y	Y	N	8(88.89)
14	Y	Y	Y	Y	Y	Y	Y	Y	N	8(88.89)
15	Y	N	N	N	Y	Y	Y	Y	N	5(55.56)
16	N	Y	N	N	Y	Y	N	N	Y	4(44.44)
17	PY	Y	Y	Y	Y	PY	N	PY	Y	5(55.56)
18	Y	Y	Y	Y	Y	PY	PY	Y	Y	7(77.78)
19	PY	PY	Y	Y	Y	PY	PY	Y	PY	4(44.44)
20	Y	Y	Y	Y	Y	Y	Y	Y	Y	9(100.00)
21	Y	Y	Y	Y	Y	Y	Y	Y	Y	9(100.00)
22	Y	N	N	N	Y	Y	Y	N	Y	5(55.56)
23	N	Y	Y	Y	Y	Y	N	Y	Y	7(77.78)
24	Y	PY	Y	Y	Y	PY	Y	Y	Y	7(77.78)
25	Y	Y	Y	Y	Y	Y	Y	Y	Y	9(100.00)
26	Y	Y	Y	Y	Y	Y	Y	Y	Y	9(100.00)
27	PY	N	N	Y	Y	Y	N	PY	N	3(33.33)
Of Yes [n (%)]	15(55.56)	14(51.85)	17(66.67)	15(55.56)	24(88.89)	14(51.85)	12(44.44)	17(62.96)	13(48.15)	

Y: Yes; PY: Partial Yes; N: No.

3.4. Effects of acupuncture

In present research, five types of comparisons had been made, summarized as follows.

3.4.1. Acupuncture versus no treatment / wait list

One Cochrane SR [31] and 1 non-Cochrane SR [28] compared acupuncture with no treatment or waitlist. The non-Cochrane SR [28] found no positive effect in reduction of depression by acupuncture compared to a waiting list control. While Cochrane SR [31] found

evidence suggesting a reduction in depression for manual acupuncture (SMD -0.56, 95% CI -0.98 to -0.15, five trials, 458 participants, $I^2 = 62\%$) and electro-acupuncture (SMD -1.26, 95% CI -2.10 to -0.43, one trials, 30 participants, $I^2 = 62\%$) compared with control.

3.4.2. Acupuncture versus sham acupuncture

While two SRs reported contradictory results, the one review which was of higher quality [31] provided no evidence of a reduction in the severity of depression in the manual acupuncture group vs. invasive control group (MD -2.97, 95% CI -6.26 to 0.31, seven trials, 418 participants, $I^2 = 81\%$), electro-acupuncture vs. invasive control (MD 0.43, 95% CI -0.54 to 1.4, five trials, 251 participants, $I^2 = 0\%$), or electro-acupuncture vs. non-invasive control (MD 0.17, 95% CI -2.14 to 2.48, two trials, 99 participants, $I^2 = 35\%$). The meta-analysis of another SR [27] demonstrated positive benefits of acupuncture over sham acupuncture (SMD 0.65, 95% CI -1.18 to -0.11, $P = 0.02$).

3.4.3. Acupuncture versus western medicine

Six SRs assessed the efficacy of acupuncture against western medicine. Shen [25] et al. found acupuncture with significantly better outcomes than medication (OR = 1.69, 95% CI 1.13–2.52, $Z = 2.55$, $P < 0.05$). Three SRs [24,26,28] suggested the efficacy of acupuncture was similar to western medicine. Another SR by Lv et al. [32] suggested a great reduction in depression for electro-acupuncture compared with medicine ($X^2 = 2.83$, $P = 0.42$, $I^2 = 0\%$, OR = 1.68, 95% CI 1.10–2.56, $Z = 2.40$, $P = 0.02$), but no reduction for manual acupuncture ($X^2 = 11.67$, $P = 0.07$, $I^2 = 49\%$, OR = 1.15, 95% CI 0.77–1.71, $Z = 0.69$, $P = 0.49$). One Cochrane SR [31] demonstrated superior efficacy of electro-acupuncture to selective serotonin reuptake inhibitors (SSRIs) (SMD -0.47, 95% CI -0.85 to 0.10, five trials, 197 participants, $I^2 = 34\%$), but no better efficacy of acupuncture was observed between manual acupuncture vs. SSRIs (SMD -0.23, 95% CI -0.50 to 0.04, 16 trials, 1570 participants, $I^2 = 84\%$) or tetracyclic antidepressants (TCAs) (SMD -0.28, 95% CI -1.25 to 0.69, three trials, 397 participants, $I^2 = 94\%$) or other antidepressants (SMD -0.22, 95% CI -0.73 to 0.29, one trial, 60 participants, $I^2 = 84\%$), and electro-acupuncture vs. TCAs (SMD -0.20, 95% CI -0.42 to 0.01, five trials, 801 participants, $I^2 = 34\%$) or heterocyclic antidepressants (SMD 0.30, 95% CI -0.20 to 0.80, one trial, 62 participants) or other antidepressants (SMD 0.09, 95% CI -0.52 to 0.70, one trial, 41 participants).

3.4.4. Acupuncture plus medication versus medication alone

The combination therapy of acupuncture and antidepressants was significantly more effective than antidepressants alone in most review results [29–32]. Three SRs [29–31] showed an obvious difference in favor of acupuncture combined with SSRIs. Investigators [31] also provided evidence of a reduction in the severity of depression during treatment from manual acupuncture plus TCAs compared with TCAs alone (SMD -3.06, 95% CI -3.82 to -2.30, one trial, 60 participants, $I^2 = 33\%$). Exceptionally, in one review [26] there was insufficient evidence to prove a superior effect of acupuncture combination therapy over western medicine (OR = 2.46, 95% CI 1.64–3.71, $Z = 4.32$, $P < 0.01$). Analysis from Lv et al. [32] suggested combination therapy took effect at 6 weeks (WMD = 1.97, 95% CI 1.24–3.15, $Z = 2.86$, $P = 0.004$) and 8 weeks (OR = 2.73, 95% CI 1.26–5.90, $Z = 2.56$, $P = 0.01$) during treatment, and was effective regardless of various degrees of depression (rated by HAMD scores).

3.4.5. Acupuncture versus psychological therapy

Among the 9 SRs assessing the efficacy of acupuncture for depression, only the research by Smith et al. [31] compared acupuncture with psychological therapy. The authors found it is unclear whether data show differences between acupuncture and psychological therapy in

the severity of depression (SMD -0.5, 95% CI -1.33 to 0.33). Overall quality of evidence for this comparison was low. In addition, this meta-analysis contains merely 2 trials, one [37] of which showed benefit of acupuncture compared with psychological control (MD -6.49, 95% CI -10.33 to -2.65, 44 participants), while another [38] with larger sample size failed to prove this benefit (MD -0.80, 95% CI -1.92 to -0.32, 453 participants).

3.5. Adverse reactions

A total of 4 SRs [24,29–31] reported adverse events, while the remaining five SRs not. Adverse reactions that occurred with antidepressants consisted mainly of dry mouth, nausea, headache and listlessness. Three studies used scales, including TESS [24,30] and SERS [29] scores, to assess side effects. Analysis from two studies [24,30] suggested great difference between acupuncture and western medicine group, and results from one study [29] also showed acupuncture elicited fewer side effects than did SSRIs alone. There were few documented adverse effects of acupuncture in the literature. Thus, acupuncture could be regarded as a safe treatment modality for depression.

4. Discussion

This overview has provided a summary of therapeutic effects of acupuncture on depression. Nine SRs were identified, and the majority of them were considered to be of critically low methodological quality and low reporting quality evaluated by AMSTAR2 and PRISMA, respectively. Most of them provided insufficient evidence to prove acupuncture is an effective treatment.

SRs, the highest quality in the hierarchy of evidence, are becoming more and more important for evidence-based decision making [39]. This overview shows, between 2008 and 2018, a considerable number of SRs of acupuncture for depression have emerged in the literature, indicating that the interest in this subject has increased greatly. However, poor methodological and reporting quality of SRs diminishes their value. Our study is the first review evaluating the methodological and reporting quality of SRs concerning the effectiveness of acupuncture for depression, to provide useful suggestions on the execution for the future reviewers.

AMSTAR, published in 2007, one of the most widely used instruments, is a practical critical appraisal tool to assess the methodological quality of SRs of RCTs of healthcare interventions [40–42]. AMSTAR 2, updated in 2017, with 16 items, is a major revision of the original AMSTAR. The main modifications contain inclusion of non-randomized studies in SRs, elaboration of the PICO, the way in which RoB was handled during evidence synthesis, discussion of possible causes and significance of heterogeneity, justification of study designs selection. Thus, we think it is reasonable to assume that AMSTAR2 is a comprehensive critical appraisal instrument to evaluate SRs and provides a broad assessment of methodological quality. PRISMA is an update and expansion of the QUOROM (Quality Of Reporting Of Meta-analysis) statement, which was developed in 1996, published in 1999, and conceived as a reporting guidance for authors reporting a meta-analysis of randomized trials [43]. The PRISMA Statement, published in 2009, consists of a 27-item checklist and a four-phase flow diagram, with the aim to help ensure the clarity and transparency of reporting of SRs and reduce the risk of flawed reporting of SRs. Therefore, AMSTAR 2 and PRISMA, as comprehensive and critical tools, have the potency to assess the methodological and reporting quality of SRs of healthcare interventions.

Disappointingly, from the results of quality evaluation by AMSTAR 2 and PRISMA, the methodological and reporting quality of the SRs included in this overview was not satisfactory. The major problems

include protocol registry, search, justification for study types and exclusions, publication bias and conflict of interest. Planning the protocol prior to the SR being undertaken reduce the bias from reviewers, and ensures that all methodological decisions are carefully considered and justified, enhancing the transparency and trustworthiness of the results [44,45]. International registry of SR protocol can be accomplished at the Cochrane Collaboration, PROSPERO and Joanna Briggs Institute. Protocol once completed, reviewers should adhere to the predefined processes, and any significant deviation from the protocol requires justification. The full search strategy for at least one major database should be available. Publications in all relevant languages should be sought and a justification provided when there are language restrictions. In addition, searching should be supplemented by checking specialized registers, consulting experts in the particular field of study, and reviewing the reference lists from the studies found [46]. At least two authors are required to determine eligibility of studies for inclusion. A consensus process should be used when there are disagreements. The review authors should provide a complete list of potentially relevant studies with justification for the exclusion of each. Unjustified exclusion may bias the review findings. Listing excluded studies would ensure comprehensive reporting of SRs and allow readers to examine the excluded studies if they had a different perspective on eligibility criteria. An reasonable explanation for the selection of the study types is also required [45]. We included SRs containing RCTs or q-RCTs, while most authors did not explain their selection of the study design for inclusion. Merely one SR fulfilled this requirement, mainly because this is an new item in AMSTAR 2, which was revised from AMSTAR in 2017, and only 1 review in our review was published after that time. Assessing the publication bias of included studies would inform readers on whether the results were generated from an unbiased sample of related studies. Review authors are required to provide details about PICOS of the included RCTs, which may help explain heterogeneity. Evidence show that meta-analyses with financial ties to a company are associated with favorable conclusions [47,48]. So it is important for authors of both SR and the included trials to state their conflict of interest and funding source to improve transparency and reduce bias. In addition, other domains also need to be reported adequately, namely structured abstract, rationale, objectives, eligibility criteria, technique for assessing the RoB in individual studies, assessment of the potential impact of RoB in individual studies on the results of meta-analysis, discussion of RoB in individual studies on the results of review should be adequately reported. Items like additional analysis in the methods part, RoB within studies in the results part also needs to be paid more attention to. In fact, many of the included SRs were published long time ago, and the earliest was in 2008. At present, the widely accepted methodological and reporting quality assessment tools are AMSTAR 2 and PRISMA, respectively. AMSTAR 2 was updated in 2017, and PRISMA was developed in 2009. The systematic reviews, which were published before the time, were not reported or assessed according to the tools we are using now, so it will inevitably contribute to the low quality of assessment results. Thus, we need to consider the publication time in the evaluation. Moreover, the evaluation include many factors such as outcomes measures and the number of participants included. Therefore, it is wrong to consider these studies of low quality to be completely meaningless. At the same time, we cannot deny the importance of quality assessment. Objectively using the selected scale for quality evaluation, we can only say that these studies are considered to be of low quality with the currently accepted evaluation scales.

Methodological quality is undoubtedly of the utmost importance for a scientific research, and serious methodological flaws render results unreliable. In some cases, methodological evaluation is impossible to implement without good reporting. Only when SRs comply with established methodology, report findings transparently, can their

conclusions be free of bias [49]. Research with excellent reporting quality will definitely point out its methodology defects objectively, while insufficient reporting of a study's methodology and findings prevents critical appraisal of the study. In addition, inadequate reporting of medical research impede the applicability and carries with it an additional risk of inadequate and misleading study results being used by patients and practitioners. Scarce health care resources may be expended on ineffective health care treatments through such inadequate reporting. A transparent and accurate reporting is an essential component of a good research [50].

The conclusions of the SRs were not free of contradictions. Among all the included SRs, only 1 SR [31] achieved a reasonable overall rating with both AMSTAR 2 and PRISMA. Thus, we prefer to be convinced of the conclusions of this review rather than trust the others with critically low methodological quality. They found low quality evidence of acupuncture (manual and electro-) with better outcomes than no treatment, but no evidence of a reduction in the severity of depression in the acupuncture (manual and electro-) group compared with sham acupuncture. They also provided evidence of a reduction in the severity of depression from acupuncture plus medication compared with medication alone. Thus, acupuncture, used as an adjunct to western medicine, was associated with higher rates of clinical efficacy. Comparing acupuncture (manual and electro-) with medicine, the results showed no better outcomes of acupuncture, with the exception of comparison between electro-acupuncture and SSRIs. Investigators provided evidence of a large effect reduction in the severity of depression during treatment from electro-acupuncture plus SSRIs compared with SSRIs alone. Evidence was insufficient to determine the efficacy of acupuncture vs. psychological therapy due to the low quality and small size of the included trials. Present overview suggested low quality evidence for the utility of acupuncture in depression. The low quality of evidence is attributed to the low quality of both RCTs and SRs, as the included SRs are based on the primary data of RCTs, and our overview is a summary of these SRs. Methodological flaws in the primary RCTs, such as inappropriate randomization methods, blinding, allocation concealment, are responsible for the low quality of the evidence. Besides, inadequate reporting of RCTs also hindered researchers from assessing the RoB in these RCTs. Therefore, the conclusions supported by the results from inadequately powered or reported RCTs, may be less than reliable. SRs with low quality are likely to have biased findings. The conclusions of SRs are not equally reliable due to variable quality, and the ones with low quality are likely to have biased findings. To draw a solid conclusion, more efforts are needed to improve the quality of both RCTs and SRs. First of all, well-designed clinical trials are expected to prove the effectiveness of acupuncture for depression, and future trials should follow the STRICTA recommendations for reporting. Secondly, researchers need to meet relevant requirements on methodologies and reporting of SRs, and we recommend that future SRs adhered to the AMSTAR 2 and PRISMA guidelines to produce high quality SRs.

Besides the quality of the SRs we included, the RCTs included in the SRs have an impact on the power of evidence on acupuncture for depression as well. Adequate reporting of RCTs can improve transparency and aid in the interpretation of studies [51]. Without transparent reporting, readers cannot judge the reliability and validity of neither trial findings nor extract information for SRs [52]. There exists a great diversity in acupuncture treatment of depression. The CONSORT statement was developed by a group of scientists and editors to improve the reporting quality of RCTs [52]. STRICTA, as an extension of the CONSORT guidelines for the specific requirements of acupuncture trial studies [53], has been designed to improve the standards for reporting interventions in acupuncture clinical trials. It is intended that use of the CONSORT and STRICTA guidelines will reduce inadequate reporting of acupuncture trials, leading to better data presentation, more strict trial

design, and more robust conclusions. SRs that include reporting on the STRICTA items could supply researchers with a reasonable and transparent interpretation for clinical heterogeneity and provide health professionals with greater confidence when using the related evidence in clinical routine practice. However, in this overview, the characteristics of interventions following CONSORT or STRICTA guidelines were reported in only 2 SRs. There is a need for future SRs and trials in this area to follow the CONSORT and particularly STRICTA guidelines to improve both internal and external validity and hence enhance the quality of acupuncture research.

It is worth noting that HAMD as well as efficacy rate remain the primary outcomes accepted by most researchers for assessing effectiveness of acupuncture for depression. These scales are subjective tools and often evaluate the overall conditions, such as sleeping and gastrointestinal conditions, which will influence the total score. Therefore, it is in urgent need to develop more accurate and objective outcomes that can be widely applied to assess the efficacy of acupuncture to treat depression.

This overview was conducted with rigorous methods. However, it still has several limitations. Firstly, only reviews published in English or Chinese were included. Chinese databases were searched in this overview, as acupuncture originated in China and there have been multitudes of studies published in Chinese. Some Asian countries, such as Japan and Korea, also conduct some research on acupuncture for depression, but we are not accessible to their databases due to resource limits. Thus, we can't be sure that all relevant articles were located. Secondly, there are certain limitations of the tools used to assess the methodological and reporting quality of included SRs. Thirdly, the variable and poor quality of the original RCTs limits the conclusiveness of the SRs. Double blinding is difficult to implement in acupuncture trials, currently it is almost impossible to blind the both the practitioners and participants due to the specialty of the acupuncture procedure, although scientists have already been engaged in developing and designing special device used as a blind control in acupuncture studies [54].

Appendix A

- #1 "Acupuncture" [Mesh]
- #2 "Acupuncture Therapy" [Mesh]
- #3 acupuncture
- #4 electroacupuncture
- #5 #1 OR #2 OR #3 OR #4
- #6 "Depression"[Mesh]
- #7 "Depressive Disorder" [Mesh]
- #8 depression
- #9 depressive disorder
- #10 depressive state
- #11 dysthymia
- #12 dysthymic disorder
- #13 #6 OR #7 OR #8 OR #9 OR #10 OR #11 OR #12
- #14 meta analysis
- #15 systematic review
- #16 #14 OR #15
- #17 #5 AND #13 AND #16

The above search strategy was applicable to PubMed, and other database search strategies was modified according to the above strategy.

5. Conclusion

Acupuncture, either in isolation or as an adjunct to other interventions, has been applied as an effective option for patients with depression in routine clinical practice. The methodological and reporting quality of SRs assessing acupuncture for depression was poor according to AMSTAR 2 and PRISMA guidelines. This overview suggested low quality evidence for the utility of acupuncture in depression. Although we can't draw firm conclusions due to low quality of the SRs and the included RCTs, the available evidence indicated that acupuncture might be effective for depression to some extent. And future research should pay attention to improving the methodological and reporting quality of both SRs and RCTs in this field.

Conflict of interest

None.

Funding

This research was funded by the International Cooperation Program of Science and Technology Department of Sichuan Province, China (No. 2016HH0007). The funder played no role in study design, data collection, analysis, interpretation of data, and writing of the study.

Acknowledgments

We thank all authors in this article. TG conceived and designed the study. TG and YNL searched and identified the articles. TG and THH extracted the data. TG and YZSH assessed the quality of the systematic reviews. TG drafted the manuscript. TG and QHZH analyzed and interpreted the data. YL and QHZH revised the draft. All authors have approved the final version of the manuscript.

B.
See [Table 4](#).

Table 4
Excluded systematic reviews.

References	Reasons for exclusion
Zhang, Z.J., et al., The effectiveness and safety of acupuncture therapy in depressive disorders: systematic review and meta-analysis. <i>J Affect Disord</i> , 124 (2010) 9-21.	Not just including RCTs or Q-RCTs
XU Xu-dong, Wu yu-ju, Meta analysis of effects of electric acupuncture on depression. <i>Journal of Clinical Psychiatry</i> , 18(2) (2008) 111-112.	Not just including RCTs or Q-RCTs
Wang Long, Chi Qingbin, Systematic review of randomized controlled trials of electroacupuncture for depression. <i>Shanghai Journal of Acupuncture and Moxibustion</i> , 27(03) (2008) 36-38.	Not just including RCTs or Q-RCTs
Sun, Y.L., et al., Acupuncture versus western medicine for depression in China: A systematic review. <i>Chinese Journal of Evidence-Based Medicine</i> , 8 (5) (2008) 340-345.	Not just including RCTs or Q-RCTs
BIAN Xing-kun, FU Li-xin, ZUO Xiao-hong, et al., Acupuncture versus medicine for depression in China: a systematic review. <i>Journal of Clinical Acupuncture and Moxibustion</i> , 27(8) (2011) 59-62.	Not just including RCTs or Q-RCTs
Wang, L., et al., Systematic evaluation of therapeutic effect and safety of acupuncture for treatment of depression. <i>Zhongguo zhen jiu</i> , 28 (5) (2008) 381-6.	Not just including RCTs or Q-RCTs
ZHONG Bao-Liang, HUANG Yue-Qin, LI Hui-Juan, The effectiveness and safety of acupuncture for depression: a systematic assessment. <i>Chinese Mental Health Journal</i> , 22(9) (2008) 641-647.	Not just including RCTs or Q-RCTs
Halbreich, U., Systematic reviews of clinical trials of acupuncture as treatment for depression: how systematic and accurate are they? <i>CNS Spectr</i> , 13 (4) (2008) 293-4, 299-300.	Not SR
Wu, J., et al., Acupuncture for depression: a review of clinical applications. <i>Can J Psychiatry</i> , 57 (7) (2012) 397-405.	Not SR
Huang Tao, Acupuncture treatment of depressive syndrome. <i>Journal of TCM</i> , 25(2) (2005) 106-107.	Not SR
Sorbero, M.E., et al., Acupuncture for Major Depressive Disorder: A Systematic Review. <i>Rand Health Q</i> , 5 (4) (2016) 7.	Insufficient data
Leo, R.J. and J.S. Ligot, Jr., A systematic review of randomized controlled trials of acupuncture in the treatment of depression. <i>J Affect Disord</i> , 97 (1-3) (2007) 13-22.	Insufficient data
Mukaiino, Y., et al., The effectiveness of acupuncture for depression—a systematic review of randomised controlled trials. <i>Acupunct Med</i> , 23 (2) (2005) 70-6.	Insufficient data
FAN Ling, FU Wen-Bin, XU Neng-Gui, et al., Meta-analysis of randomized controlled trials on acupuncture and moxibustion on treating depression. <i>Chinese Journal of Gerontology</i> , 30(18) (2010) 2561-2563	Insufficient data
Li, B., et al., Meta analysis and analytic hierarchy process-based intervention strategy research of acupuncture for primary depressive disorder. <i>Zhongguo zhen jiu</i> , 33 (5) (2013) 463-7.	Insufficient data
Fan Ling. The literature evaluation and clinical research of treating depression with acupuncture and moxibustion[D]. Guangzhou University of Chinese Medicine, 2011.	Duplicated SRs

C.
See [Table 5](#).

Table 5
Methodological quality assessment of systematic reviews by AMSTAR 2 (assessor 1).

Study	Item1	Item 2	Item3	Item4	Item5	Item6	Item7	Item 8	Item9	Item 10	Item 11	Item 12	Item 13	Item14	Item15	Item16
Wang 2008	Y	N	N	PY	Y	Y	Y	Y	PY	N	Y	N	N	Y	N	N
Zhang 2016	Y	N	N	PY	N	Y	N	Y	PY	N	Y	N	N	Y	N	Y
Chan 2015	Y	N	N	PY	N	Y	N	PY	PY	N	Y	N	N	N	N	Y
Stub 2011	Y	N	N	PY	N	N	N	Y	Y	N	N	N	Y	N	N	Y
Smith 2018	Y	Y	Y	Y	Y	N	Y	Y	Y	N	Y	Y	Y	Y	Y	Y
Fan 2010	N	N	N	PY	Y	N	N	Y	PY	N	Y	N	N	Y	Y	Y
Shen 2014	Y	N	N	PY	N	Y	N	Y	PY	N	Y	N	N	Y	Y	N
Xiong 2009	Y	N	N	PY	Y	Y	N	Y	Y	N	Y	N	Y	N	N	N
Lv 2014	N	N	N	PY	N	N	N	Y	PY	N	Y	N	N	Y	Y	N

Y: Yes; PY: Partial Yes; N: No.

D.
See [Table 6](#).

Table 6
Methodological quality assessment of systematic reviews by AMSTAR 2 (assessor 2).

Study	Item1	Item 2	Item3	Item4	Item5	Item6	Item7	Item 8	Item9	Item 10	Item 11	Item 12	Item 13	Item14	Item15	Item16
Wang 2008	Y	N	N	PY	Y	Y	Y	PY	PY	N	Y	N	N	Y	N	N
Zhang 2016	Y	N	N	PY	N	Y	N	Y	PY	N	Y	N	N	Y	N	Y
Chan 2015	Y	N	N	PY	N	Y	N	PY	PY	N	Y	N	N	N	N	Y
Stub 2011	Y	N	N	PY	Y	N	N	PY	Y	N	N	N	Y	N	N	Y
Smith 2018	Y	Y	Y	Y	Y	N	Y	Y	Y	N	Y	Y	Y	Y	Y	Y
Fan 2010	N	N	N	PY	Y	N	N	PY	PY	N	Y	N	N	Y	Y	N
Shen 2014	Y	N	N	PY	N	Y	N	PY	PY	N	Y	N	N	Y	Y	N
Xiong 2009	Y	N	N	PY	Y	Y	N	PY	Y	N	Y	N	Y	N	N	N
Lv 2014	N	N	N	PY	N	N	N	PY	PY	N	Y	N	N	Y	N	N

Y: Yes; PY: Partial Yes; N: No.

E.
See [Table 7](#).

Table 7
Reporting quality assessment of systematic reviews by PRISMA (assessor 1).

Item	Study								
	Wang2008	Zhang2016	Chan2015	Stub2011	Smith2018	Fan2010	Shen2014	Xiong2009	Lv2014
1	Y	Y	Y	Y	N	Y	Y	Y	Y
2	PY	PY	PY	PY	PY	PY	N	PY	PY
3	PY	Y	Y	PY	Y	PY	PY	PY	PY
4	PY	PY	PY	PY	Y	PY	PY	PY	PY
5	N	N	N	N	N	N	N	N	N
6	PY	PY	PY	PY	Y	PY	PY	Y	PY
7	Y	Y	Y	Y	Y	Y	PY	Y	Y
8	Y	Y	Y	Y	Y	Y	Y	Y	Y
9	Y	N	Y	N	Y	Y	N	Y	PY
10	PY	PY	PY	N	Y	PY	PY	PY	PY
11	Y	PY	Y	PY	Y	PY	Y	N	Y
12	PY	PY	Y	Y	Y	PY	PY	Y	PY
13	Y	Y	Y	Y	Y	Y	Y	Y	N
14	Y	Y	Y	Y	Y	Y	Y	Y	N
15	Y	N	N	N	Y	Y	Y	Y	N
16	N	Y	N	N	Y	Y	N	N	Y
17	PY	Y	Y	Y	Y	PY	N	PY	Y
18	Y	Y	Y	Y	Y	PY	PY	Y	Y
19	PY	PY	Y	Y	Y	PY	PY	Y	PY
20	Y	Y	Y	Y	Y	Y	Y	Y	Y
21	Y	Y	Y	Y	Y	Y	Y	Y	Y
22	Y	N	N	N	Y	Y	Y	N	Y
23	N	Y	Y	Y	Y	Y	N	Y	Y
24	Y	Y	Y	Y	Y	Y	Y	Y	Y
25	Y	Y	Y	Y	Y	Y	Y	Y	Y
26	Y	Y	Y	Y	Y	Y	Y	Y	Y
27	PY	N	N	Y	Y	Y	N	PY	N

Y, Yes; PY, Partial Yes; N, No.

F.

See Table 8.

Table 8
Reporting quality assessment of systematic reviews by PRISMA (assessor 2).

Item	Study								
	Wang2008	Zhang2016	Chan2015	Stub2011	Smith2018	Fan2010	Shen2014	Xiong2009	Lv2014
1	Y	Y	Y	Y	N	Y	Y	Y	Y
2	PY	PY	PY	PY	PY	PY	N	PY	PY
3	PY	Y	Y	PY	Y	PY	PY	PY	PY
4	PY	PY	PY	PY	Y	PY	PY	PY	PY
5	N	N	N	N	N	N	N	N	N
6	PY	PY	PY	PY	Y	PY	PY	Y	PY
7	Y	Y	Y	Y	Y	Y	PY	Y	Y
8	PY	PY	PY	PY	Y	PY	PY	PY	PY
9	Y	N	Y	N	Y	Y	N	Y	PY
10	Y	Y	Y	N	Y	N	Y	Y	PY
11	Y	PY	Y	PY	Y	PY	Y	N	Y
12	PY	PY	Y	Y	Y	PY	PY	Y	PY
13	Y	Y	Y	Y	Y	Y	Y	Y	N
14	Y	Y	Y	Y	Y	Y	Y	Y	N
15	Y	N	N	N	Y	Y	Y	Y	N
16	PY	Y	N	N	Y	Y	N	N	Y
17	PY	Y	Y	Y	Y	PY	N	PY	Y
18	Y	Y	Y	Y	Y	PY	PY	Y	Y
19	PY	PY	Y	Y	Y	PY	PY	Y	PY
20	Y	Y	Y	Y	Y	Y	Y	Y	Y
21	Y	Y	Y	Y	Y	Y	Y	Y	Y
22	Y	N	N	N	Y	Y	Y	N	Y
23	N	Y	Y	Y	Y	Y	N	Y	Y
24	Y	PY	Y	Y	Y	PY	Y	Y	Y
25	Y	Y	Y	Y	Y	Y	Y	Y	Y
26	Y	Y	Y	Y	Y	Y	Y	Y	Y
27	PY	N	N	Y	Y	Y	N	PY	N

Y, Yes; PY, Partial Yes; N, No.

References

- Health Topics: Depression, World Health Organization, 2016 September 25 www.who.int/topics/depression/en/.
- R.C. Kessler, M. Angermeyer, J.C. Anthony, et al., Lifetime prevalence and age-of-onset distributions of mental disorders in the World Health Organization's World Mental Health survey initiative, *World Psychiatry* 6 (3) (2007) 168–176.
- S.I. Saarni, J. Suvisaari, H. Sintonen, et al., Impact of psychiatric disorders on health-related quality of life: general population survey, *Br. J. Psychiatry* 190 (2007) 326–332.
- C.D. Mathers, D. Loncar, Projections of global mortality and burden of disease from 2002 to 2030, *PLoS Med.* 3 (11) (2006) e442.
- P.E. Greenberg, A.A. Fournier, T. Sisitsky, et al., The economic burden of adults with major depressive disorder in the United States (2005 and 2010), *J. Clin. Psychiatry* 76 (2) (2015) 155–162.
- M. Clark, D. DiBenedetti, V. Perez, Cognitive dysfunction and work productivity in major depressive disorder, *Expert Rev. Pharmacoecon. Outcomes Res.* 16 (4) (2016) 455–463.
- A.J. Ferrari, F.J. Charlson, R.E. Norman, et al., Burden of depressive disorders by country, sex, age, and year: findings from the global burden of disease study 2010, *PLoS Med.* 10 (11) (2013) e1001547.
- P.F. Petter Viksveen, Clare Relton. Homeopathy in the treatment of depression: a systematic review, *Eur. J. Integr. Med.* 22 (2018) 22–36.
- C.M. Thomas, S. Morris, Cost of depression among adults in England in 2000, *Br. J. Psychiatry* 183 (2003) 514–519.
- B. Arroll, C.R. Elley, T. Fishman, et al., Antidepressants versus placebo for depression in primary care, *Cochrane Database Syst. Rev.* 3 (2009).
- S. Ijaz, P. Davies, C.J. Williams, et al., Psychological therapies for treatment-resistant depression in adults, *Cochrane Database Syst. Rev.* 5 (2018).
- J. Read, J. Williams, Adverse effects of antidepressants reported by a large international cohort: emotional blunting, suicidality, and withdrawal effects, *Curr. Drug Saf.* 13 (3) (2018) 176–186.
- R. Gafoor, H.P. Booth, M.C. Gulliford, Antidepressant utilisation and incidence of weight gain during 10 years' follow-up: population based cohort study, *Bmj* 361 (2018) k1951.
- C. Cartwright, K. Gibson, J. Read, et al., Long-term antidepressant use: patient perspectives of benefits and adverse effects, *Patient Prefer. Adherence* 10 (2016) 1401–1407.
- P. Blier, The pharmacology of putative early-onset antidepressant strategies, *Eur. Neuropsychopharmacol.* 13 (2) (2003) 57–66.
- M.P. Hengartner, J. Angst, W. Rossler, Antidepressant Use Prospectively Relates to a Poorer Long-Term Outcome of Depression: Results from a Prospective Community Cohort Study over 30 Years, *Psychother. Psychosom.* 87 (3) (2018) 181–183.
- J. Davies, J. Read, A systematic review into the incidence, severity and duration of antidepressant withdrawal effects: are guidelines evidence-based? *Addict. Behav.* (2018).
- G.A. Fava, A. Gatti, C. Belaise, et al., Withdrawal symptoms after selective serotonin reuptake inhibitor discontinuation: a systematic review, *Psychother. Psychosom.* 84 (2) (2015) 72–81.
- F. Santoft, E. Axelsson, L.G. Ost, et al., Cognitive behaviour therapy for depression in primary care: systematic review and meta-analysis, *Psychol. Med.* (2019) 1–9.
- K. Linde, K. Sigterman, L. Kriston, et al., Effectiveness of psychological treatments for depressive disorders in primary care: systematic review and meta-analysis, *Ann. Fam. Med.* 13 (1) (2015) 56–68.
- A. Ribeiro, J.P. Ribeiro, O. von Doellinger, Depression and psychodynamic psychotherapy, *Braz J Psychiatry* 40 (1) (2018) 105–109.
- S.C. Cook, A.C. Schwartz, N.J. Kaslow, Evidence-based psychotherapy: advantages and challenges, *Neurotherapeutics* 14 (3) (2017) 537–545.
- A.F. Thachil, R. Mohan, D. Bhugra, The evidence base of complementary and alternative therapies in depression, *J. Affect. Disord.* 97 (1–3) (2007) 23–35.
- J. Xiong, Y.H. Du, J.L. Liu, et al., Acupuncture versus western medicine for depression neurosis: a systematic review, [Chinese]. *Chinese Journal of Evidence-Based Medicine* 9 (9) (2009) 969–975.
- Shen Hui, Z.J. Yang Jing, Ni. Yang, Systematic review of randomized controlled trials of acupuncture for depression, *Journal of New Chinese Medicine* 46 (6) (2014) 220–222.
- L. Fan, W. Fu, N. Xu, et al., Meta-analysis of 20 clinical, randomized, controlled trials of acupuncture for depression, *Neural Regen. Res.* 5 (24) (2010) 1862–1869.
- H. Wang, H. Qi, B.S. Wang, et al., Is acupuncture beneficial in depression: a meta-analysis of 8 randomized controlled trials? *J. Affect. Disord.* 111 (2–3) (2008) 125–134.
- T. Stub, T. Alraek, J. Liu, Acupuncture treatment for depression-A systematic review and meta-analysis, *Eur. J. Integr. Med.* 3 (4) (2011) e253–e264.
- Y. Zhang, S.S. Qu, J.P. Zhang, et al., Rapid Onset of the Effects of Combined Selective Serotonin Reuptake Inhibitors and Electroacupuncture on Primary Depression: A Meta-Analysis, *J. Altern. Complement. Med.* 22 (1) (2016) 1–8.
- Y.Y. Chan, W.Y. Lo, S.N. Yang, et al., The benefit of combined acupuncture and antidepressant medication for depression: a systematic review and meta-analysis, *J. Affect. Disord.* 176 (2015) 106–117.

- [31] C.A. Smith, M. Armour, M.S. Lee, et al., Acupuncture for depression, *Cochrane Database Syst. Rev.* 3 (2018) Cd004046.
- [32] L. Zhe, *The Systematic Review Of Acupuncture Treatment Of Depressive Disorder Based On The Literatures Of Recent 30 Years*, Shandong University of TCM, 2014.
- [33] Y. Zhang, S.S. Qu, J.P. Zhang, et al., Rapid onset of the effects of combined selective serotonin reuptake inhibitors and electroacupuncture on primary depression: A meta-Analysis. [German], *Deutsche Zeitschrift fur Akupunktur* 59 (3) (2016) 27–29.
- [34] A.R. Jadad, R.A. Moore, D. Carroll, et al., Assessing the quality of reports of randomized clinical trials: is blinding necessary? *Control. Clin. Trials* 17 (1) (1996) 1–12.
- [35] W. Noguchi, T. Ohno, S. Morita, et al., [Reliability and validity of the functional assessment of chronic illness therapy-spiritual (FACT-Sp) Japanese version in cancer patients—a preliminary study], *Gan To Kagaku Ryoho* 31 (3) (2004) 387–391.
- [36] J.P.T. Higgins, G.S. Cochrane, *Handbook for Systematic Reviews of Interventions Version 5.1.0*. The Cochrane Collaboration (updated March, (2011) <http://www.cochrane-handbook.org>.
- [37] C. Wang, M. Liu, J. Lv, et al., Effect of acupuncture and moxibustion on depressive states of stroke patients' spouses, *Zhongguo Zhen Jiu* 35 (3) (2015) 223–226.
- [38] H. MacPherson, S. Richmond, M. Bland, et al., Acupuncture and counselling for depression in primary care: a randomised controlled trial, *PLoS Med.* 10 (9) (2013).
- [39] L. Manchikanti, R.M. Benyamin, S. Helm, et al., Evidence-based medicine, systematic reviews, and guidelines in interventional pain management: part 3: systematic reviews and meta-analyses of randomized trials, *Pain Physician* 12 (1) (2009) 35–72.
- [40] B.J. Shea, J.M. Grimshaw, G.A. Wells, et al., Development of AMSTAR: a measurement tool to assess the methodological quality of systematic reviews, *BMC Med. Res. Methodol.* 7 (2007) 10.
- [41] B.J. Shea, C. Hamel, G.A. Wells, et al., AMSTAR is a reliable and valid measurement tool to assess the methodological quality of systematic reviews, *J. Clin. Epidemiol.* 62 (10) (2009) 1013–1020.
- [42] B.J. Shea, L.M. Bouter, J. Peterson, et al., External validation of a measurement tool to assess systematic reviews (AMSTAR), *PLoS One* 2 (12) (2007) e1350.
- [43] D. Moher, D.J. Cook, S. Eastwood, et al., Improving the quality of reports of meta-analyses of randomised controlled trials: the QUOROM statement, Quality of Reporting of Meta-analyses. *Lancet* 354 (9193) (1999) 1896–1900.
- [44] D. Moher, L. Shamseer, M. Clarke, et al., Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015 statement, *Syst. Rev.* 4 (2015) 1.
- [45] L.A. Riesenber, E.M. Justice, Conducting a successful systematic review of the literature, part 1, *Nursing* 44 (4) (2014) 13–17.
- [46] B.J. Shea, B.C. Reeves, G. Wells, et al., AMSTAR 2: a critical appraisal tool for systematic reviews that include randomised or non-randomised studies of health-care interventions, or both, *Bmj* (2017) j4008.
- [47] V. Yank, D. Rennie, L.A. Bero, Financial ties and concordance between results and conclusions in meta-analyses: retrospective cohort study, *Bmj* 335 (7631) (2007) 1202–1205.
- [48] A.W. Jorgensen, J. Hilden, P.C. Gotzsche, Cochrane reviews compared with industry supported meta-analyses and other meta-analyses of the same drugs: systematic review, *Bmj* 333 (7572) (2006) 782.
- [49] K. Pussegoda, L. Turner, C. Garrity, et al., Systematic review adherence to methodological or reporting quality, *Syst. Rev.* 6 (1) (2017) 131.
- [50] I. Simera, D.G. Altman, D. Moher, et al., Guidelines for reporting health research: the EQUATOR network's survey of guideline authors, *PLoS Med.* 5 (6) (2008) e139.
- [51] B. Ma, Z.M. Chen, J.K. Xu, et al., Do the CONSORT and STRICTA Checklists Improve the Reporting Quality of Acupuncture and Moxibustion Randomized Controlled Trials Published in Chinese Journals? A Systematic Review and Analysis of Trends, *PLoS One* 11 (1) (2016) e0147244.
- [52] D. Moher, S. Hopewell, K.F. Schulz, et al., CONSORT 2010 explanation and elaboration: updated guidelines for reporting parallel group randomised trials, *Int. J. Surg.* 10 (1) (2012) 28–55.
- [53] H. MacPherson, A. White, M. Cummings, et al., Standards for reporting interventions in controlled trials of acupuncture: the STRICTA recommendations, *J. Altern. Complement. Med.* 8 (1) (2002) 85–89.
- [54] N. Takakura, M. Takayama, A. Kawase, et al., Double blinding with a new placebo needle: a validation study on participant blinding, *Acupunct. Med.* 29 (3) (2011) 203–207.