

# BMJ Open A critical appraisal of the methodology and quality of evidence of systematic reviews and meta-analyses of traditional Chinese medical nursing interventions: a systematic review of reviews

Ying-Hui Jin,<sup>1</sup> Guo-Hao Wang,<sup>2</sup> Yi-Rong Sun,<sup>1</sup> Qi Li,<sup>3</sup> Chen Zhao,<sup>3</sup> Ge Li,<sup>4</sup> Jin-Hua Si,<sup>5</sup> Yan Li,<sup>1</sup> Cui Lu,<sup>6</sup> Hong-Cai Shang<sup>7</sup>

**To cite:** Jin Y-H, Wang G-H, Sun Y-R, *et al.* A critical appraisal of the methodology and quality of evidence of systematic reviews and meta-analyses of traditional Chinese medical nursing interventions: a systematic review of reviews. *BMJ Open* 2016;**6**:e011514. doi:10.1136/bmjopen-2016-011514

► Prepublication history and additional material is available. To view please visit the journal (<http://dx.doi.org/10.1136/bmjopen-2016-011514>).

Received 17 February 2016  
Revised 7 October 2016  
Accepted 10 October 2016



CrossMark

For numbered affiliations see end of article.

**Correspondence to**  
Professor Hong-Cai Shang;  
[shanghongcai@foxmail.com](mailto:shanghongcai@foxmail.com)

## ABSTRACT

**Objective:** To assess the methodology and quality of evidence of systematic reviews and meta-analyses of traditional Chinese medical nursing (TCMN) interventions in Chinese journals. These interventions include acupressure, massage, Tai Chi, Qi Gong, electroacupuncture and use of Chinese herbal medicines—for example, in enemas, foot massage and compressing the umbilicus.

**Design:** A systematic literature search for systematic reviews and meta-analyses of TCMN interventions was performed. Review characteristics were extracted. The methodological quality and the quality of the evidence were evaluated using the Assessment of Multiple Systematic Reviews (AMSTAR) and Grading of Recommendations Assessment, Development and Evaluation (GRADE) approaches.

**Result:** We included 20 systematic reviews and meta-analyses, and a total of 11 TCMN interventions were assessed in the 20 reviews. The compliance with AMSTAR checklist items ranged from 4.5 to 8 and systematic reviews/meta-analyses were, on average, of medium methodological quality. The quality of the evidence we assessed ranged from very low to moderate; no high-quality evidence was found. The top two causes for downrating confidence in effect estimates among the 31 bodies of evidence assessed were the risk of bias and inconsistency.

**Conclusions:** There is room for improvement in the methodological quality of systematic reviews/meta-analyses of TCMN interventions published in Chinese journals. Greater efforts should be devoted to ensuring a more comprehensive search strategy, clearer specification of the interventions of interest in the eligibility criteria and identification of meaningful outcomes for clinicians and patients (consumers). The overall quality of evidence among reviews remains suboptimal, which raise concerns about their roles in influencing clinical practice. Thus, the conclusions in reviews we assessed must be treated with caution and their roles in influencing clinical practice should be limited. A critical appraisal of systematic reviews/meta-

## Strengths and limitations of this study

- This study is the first attempt to assess the methodology and quality of evidence of systematic reviews and meta-analyses undertaken within traditional Chinese medical nursing (TCMN) published in Chinese journals using the Assessment of Multiple Systematic Reviews (AMSTAR) and Grading of Recommendations Assessment, Development and Evaluation (GRADE) approaches.
- The results show that critical appraisal of systematic reviews/meta-analyses of TCMN interventions published before this review had weaknesses, especially in the use of evidence and decision-making, and suggestions are provided for incorporating improvements into future work.
- The main limitation of this study is that the methodology and quality of the evidence assessments presented were based on the reported published information about the assessment items in the individual systematic reviews and meta-analyses, which may not reflect the actual methodology used.

analyses of TCMN interventions is particularly important to provide sound guidance for TCMN.

## INTRODUCTION

Despite considerable developments in medicine, a large number of people, both in developed and developing countries, turn to complementary and alternative medicine. This includes traditional Chinese medicine (TCM),<sup>1</sup> which is a science nourished by Chinese culture. It is generally delivered by qualified practitioners and has been

practised over thousands of years in China.<sup>2</sup> Traditional Chinese medical nursing (TCMN) is a significant branch of nursing in China, in which TCM nurses use various interventions such as psychological interventions, diet therapy, TCM exercises and medications, acupuncture, massage and cupping.

The holistic philosophy and personalised nature of TCMN concur with the patient-centred approach found in modern nursing elsewhere. In the Chinese Nursing Development Program (2010–2015) it is explicitly pointed out that TCMN should be developed to contribute to the prevention and control of degenerative and chronic diseases and should also be combined with Western medicine nursing techniques.<sup>3</sup> In China, as specialised TCM clinical nursing has developed, TCMN techniques have become more popular, allowing standardised nursing specialties to gradually become established. As a result the level of both TCMN service delivery and scientific research has been significantly improved. A survey of 137 TCM institutions in China showed that 85 TCMN techniques were provided for patients. The 10 most common techniques were moxibustion, cupping therapy, auricular application pressure, TCM fumigation, acupuncture point massage, acupoint sticking, TCM enema, poultices with Chinese medicine, inunction with Chinese medicine and scraping therapy.<sup>4</sup>

Reports of clinical trials assessing the effectiveness of TCMN techniques are greatly needed. Over the past decade, the number of papers reporting trials of TCMN, as well as systematic reviews and meta-analyses based on them, has steadily increased. Systematic reviews and meta-analyses serve a vital role in the development of clinical practice guidelines (CPGs).<sup>5</sup> Assessing and

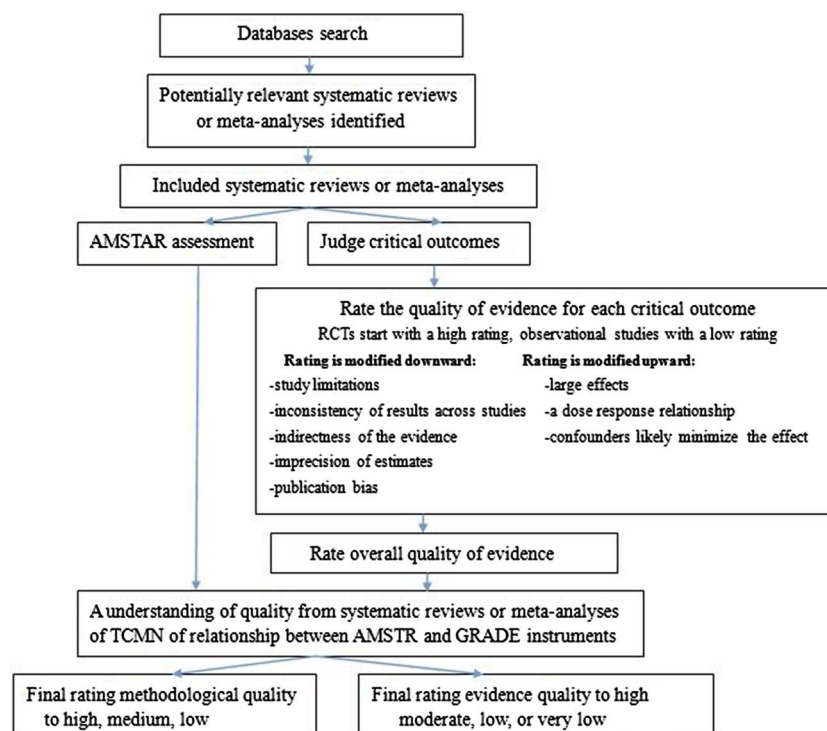
synthesising primary studies of TCMN interventions in systematic reviews and meta-analyses and then developing a CPG for integrated TCM and Western medicine care can promote the sustainable development of TCMN. Although systematic reviews and meta-analyses strive to provide scientifically rigorous, independent and accurate summaries of the scientific evidence for a specific question of interest,<sup>6</sup> the methodological deficiencies which they contain may result in misleading results and overestimation or underestimation of the investigated effects.<sup>7</sup> Even methodologically sound systematic reviews and meta-analyses may provide only indirect or imprecise evidence for the question of interest. For the CPG developers, the quality ratings reflect the extent of our confidence that estimates of an effect are adequate to support a particular decision or recommendation.<sup>8</sup>

A critical appraisal of systematic review and meta-analysis of TCMN can increase nurses' confidence and facilitate efficient application of evidence.<sup>9</sup> In this study, we used widely accepted instruments—namely, the Assessment of Multiple Systematic Reviews (AMSTAR) tool<sup>10 11</sup> and Grades of Recommendations Assessment, Development and Evaluation (GRADE) approach,<sup>12</sup> to critically assess the methodology and quality of evidence of TCMN interventions in Chinese journals and determine their contribution to the development of evidence-based decision-making.

## METHODS

The technology road mapping of this study is presented in figure 1.

**Figure 1** Technology road mapping of this study. AMSTAR, Assessment of Multiple Systematic Reviews; GRADE, Grading of Recommendations Assessment, Development and Evaluation; RCTs, randomised controlled trials; TCMNA, traditional Chinese medical nursing ????.



## Eligibility criteria

We included a study if it met the following criteria: (1) the study design was a systematic review, meta-analysis, or systematic review and meta-analysis; (2) the topic was TCMN care in China; (3) the papers were full-text articles in professional nursing journals or the four professional evidence-based medicine (EBM) journals in China. Articles were excluded (1) if the interventions focused on a broad concept of TCMN (eg, TCM care vs Western medicine care) without subgroup analysis, which meant that the review had a particularly broad scope, reflecting great clinical heterogeneity; (2) the intervention group included non-TCMN interventions (eg, TCMN combined with Western medicine or combined with acupuncture). A flow diagram showing selection of systematic reviews and meta-analyses is presented in figure 2.

## Data sources

A comprehensive literature search was conducted to identify systematic reviews and meta-analyses written in Chinese by searching CNKI (China National Knowledge Infrastructure), VIP (information/Chinese Scientific Journals database), Wanfang, CBM (Chinese Biomedical Literature database) from inception through to April 2016. The search terms 'systematic review' or 'meta-analysis' were used to navigate electronic journals to locate systematic reviews/meta-analyses of TCMN interventions published in nursing and EBM journals. The reference lists of the retrieved review articles were also screened to identify potential studies. If several updates of a study were available, only the most recent version was included (for details of the search strategy see online supplementary file).

## Study selection and data extraction

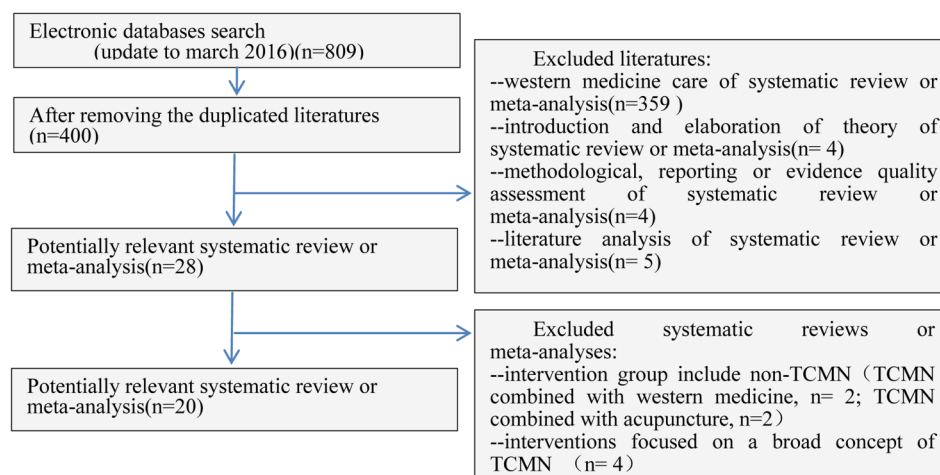
Two reviewers independently screened the abstracts and titles of studies and subsequently reviewed the full-text articles for inclusion; after this, data extraction was performed. We categorised the outcomes of systematic reviews and meta-analyses into the following types:

endpoint, quality of life (QoL), the target event occurred, symptoms, laboratory outcome, composite outcome (synthesis of multiple different outcomes), adverse events and economic evaluations. The risk ratio and 95% CI of dichotomous data, and weighted mean difference or standardised mean difference with 95% CI of continuous data, of the outcome were extracted when possible. In addition, basic characteristics of every review, such as the surname of the first author, year of publication, journal names, intervention and comparison, were extracted. Information related to AMSTAR and GRADE evaluation was also extracted. This included methodological quality of the original studies (allocation concealment, blinding, follow-up and whether or not the research adhered to the intention-to-treat principle), details of interventions and controls used in all included original studies, reporting of outcomes and outcome measures, the pooled estimate and 95% CI for the difference in effect between intervention and control for outcome, total sample of outcome, the extent to which each trial contributed to the estimate of magnitude of effect based on study sample size and number of outcome events, tests of heterogeneity and  $I^2$ , subgroup effects and the method and result of assessment of publication bias.

## Quality assessment

For every systematic review and meta-analysis, quality assessment was carried out by two assessors independently using the AMSTAR tool<sup>10 11</sup> and GRADE approach.<sup>12</sup> To improve standardisation, special training was given and a pre-test was performed. Disagreement between reviewers was resolved by discussion or by a third assessor. Agreement between the two reviewers was determined by the  $\kappa$  statistic with corresponding 95% CI. Different assessors carried out the AMSTAR evaluation and GRADE evaluation to ensure that their judgement was not affected by previous impressions. Appraisers were not allowed to communicate or confer with each other during the appraisal process.

**Figure 2** Flowchart of identified, included and excluded systematic reviews or meta-analyses of traditional Chinese medical nursing (TCMN) interventions.



According to the AMSTAR criteria, a score of 0 or 1 was given for each criterion, with equal weight given to each domain. We judged each item as 'yes (score 1)' when the criterion was explicitly met, 'no (score 0)' when the criterion was explicitly not met, 'cannot answer' when the item was relevant but not described adequately or not reported at all and 'not applicable' when the item was not relevant. When specific domains were not reported in sufficient detail, we gave a score of 0.5 for that domain. The overall score was categorised into three levels: 8–11 was high quality; 4–7 was medium quality and 0–3 was low quality. All assessors reached a more complete and unanimous standard for AMSTAR criteria after careful and full discussion between all authors.

To grade the quality of evidence, the authors identified outcomes that are of key importance to patients and then reviewers applied GRADE to determine the quality of the evidence and considered the five possible reasons to downgrade the evidence or the three possible reasons to upgrade the evidence.<sup>8 9 12</sup> The assessors were conservative in their judgement of downgrading or upgrading. When the systematic review did not provide sufficient information to judge the quality of evidence, the assessor made an attempt to contact authors of individual studies. Finally, the definitions 'high', 'moderate', 'low' and 'very low' were used to grade the quality of evidence.

### Data analysis

We established a database using Microsoft Excel 2007 software to extract data. Information on each included paper was imported into the database for analysis. We used descriptive statistics on the distribution of scores according to AMSTAR items, and summary statistics for the observed AMSTAR scores, for each included systematic review and meta-analysis. A GRADE evidence profile, which included an explicit judgement of each factor that determines the quality of evidence for the outcome of each included systematic review and meta-analysis, was obtained using GRADE profiler 3.6 software.

The AMSTAR instrument and GRADE approach were each applied to assess the methodological and evidence quality based on different criteria and systems, but they do have some similarities. For example, item 3 about a thorough and comprehensive search (eg, searching in international, national, regional and subject-specific databases, the Cochrane Central Register of Controlled Trials (CENTRAL), conference abstracts and other grey literature and ongoing trials) to identify as many relevant studies as possible helps to reduce a high probability of publication bias (GRADE downrating item). The correlation between AMSTAR and GRADE instruments was studied by scatterplot using SPSS V.17.0.

## RESULTS

### Characteristics of included studies

The literature search yielded 809 potentially relevant references; of which, 28 were selected for full-text

review. Finally, 20 studies<sup>13–32</sup> were included in this study. The year of publication<sup>13–32</sup> ranged from 2010 to 2016, with the number of reviews published in 2014 accounting for near a half of these reviews (8/20, 40%). Eleven TCMN interventions were assessed: acupressure, acupoint massage, acupoint stimulation, auricular point therapy, Tai Chi, Qi Gong, electroacupuncture combined with auricular point plaster therapy, Chinese herbal retention enema, inunction with Chinese medicine, foot bath therapy or foot massage with TCM, and compressing the umbilicus with Chinese herbs. None of the studies included observational research. Two included both randomised controlled trials (RCTs) and quasi-RCTs. No systematic review or meta-analysis used indirect comparison. None of the 20 studies used the GRADE approach to summarise evidence. The general characteristics of the assessed systematic reviews and meta-analyses are shown in [table 1](#).

### AMSTAR methodological quality

The two reviewers had satisfactory agreement ( $\kappa=0.87$ ). The methodological quality of all the included reviews is presented in [table 2](#). In summary, compliance with the AMSTAR checklist items ranged from 4.5 to 8 and the majority of systematic reviews and meta-analyses were of medium (16/20, 80.0%) methodological quality.

None of the 20 studies provided a registered protocol. For all 20 studies, study selection and data extraction were conducted, respectively, by two independent reviewers. Most of them (13/20, 65%) adequately described the characteristics of the included trials, but none provided a list of included and excluded studies. The search strategy design was not sufficiently comprehensive in 10 studies (50.0%). The mean number of electronic databases searched in the reviews was 6 (SD 2.2, range 2–11). The most frequently searched databases were PubMed (14/20, 70%) and CNKI (19/20, 95%). Two reviews<sup>14 24</sup> searched only the Chinese databases. Only two studies<sup>19 23</sup> considered the status of publication (eg, grey literature). The literature search in 10 of them was supplemented by consulting textbooks, experts in the particular field of study or by retracing references. No review searched ongoing trials. All the reviews assessed scientific quality of the included studies. The risk of bias tool from the Cochrane handbook criteria (11/20, 55%) and the Jadad scale (6/20, 30%) were the most common criteria for quality assessment of included studies.

The majority of the systematic reviews and meta-analyses used appropriate methods to combine the findings of the studies included. They all stated that a random-effects model was used to combine study data when there was heterogeneity. When substantial heterogeneity was detected, possible explanations were explored in subgroup analyses in six cases. There were no reviews in which meta-regression was applied. Two reviews conducted sensitivity analysis by exchanging the statistical approach for data synthesis (random effects vs



**Table 1** The characteristics of included systematic reviews/meta-analyses

Study	Design	No. of patients	Population	Intervention	Comparison	Outcomes	The range of literature search
Baoxia Chang 2014 <sup>23</sup>	RCT	467	Simple obesity	Electroacupuncture combined with auricular point plaster	Electroacupuncture	Rate of effectiveness, BMI, percentage of body fat, waist circumference	Cochrane Library, PubMed, VIP, CNKI, CBM, references of the included literature, the grey literature
Wangqin Shen 2010 <sup>13</sup>	RCT	327	Phlebitis	Ruyijinhuangsan	Magnesium sulfate by wet compression	Rate of effectiveness	Medline, CBM, CNKI, references of the included literature
Hongying Pu 2011 <sup>15</sup>	RCT	703	Pressure ulcers	Moist exposed burn ointment	Skin disinfection solution or antibiotic ointment	Cure rate Time of cure	PubMed, Embase, Cochrane database, CBM, VIP, references of the included literature
Xuan Zhou 2011 <sup>16</sup>	RCT	1117	Postoperative patients	Acupressure wristbands	Placebo wristband	Incidence of nausea, incidence of vomiting	Medline, CNKI, CBM, VIP, WanFang
Xiaoli Wu 2012 <sup>17</sup>	RCT	860	Diabetic foot ulcers	Foot bath therapy or foot massage with traditional Chinese medicine	Hot foot bath	Effectiveness rate	Medline, CNKI, CBM, VIP
Xilan Zheng 2012 <sup>18</sup>	RCT	551	Pressure ulcers	Traditional Chinese medicine for elimination of necrotic tissues	Skin disinfection solution or antibiotic ointment	Cure rate, cure time, frequency of dressing change	Medline, Embase, PubMed, Cochrane Library, CBM, VIP, CNKI, Wanfang
Na Li 2011 <sup>14</sup>	RCT	712	Phlebitis	External application with aloe vera	Magnesium sulfate by wet compression	Effectiveness rate	CNKI, VIP, Wanfang
Jiaqi Xu 2013 <sup>20</sup>	CCT RCT	610	Pressure ulcer	Resina draconis	Skin disinfection solution or antibiotic ointment	Effectiveness rate, cure time	Cochrane Library, PubMed, Elsevier SDOL, Web of Knowledge, CBM, CNKI, VIP, Wanfang, references of the included literature
Yuan Zhao 2013 <sup>21</sup>	RCT	2796	Elderly individuals	Tai Chi	Regular sport or physical therapy	Rate of falls, time up and go test, functional reach test, Berg balance scale (BBS)	PubMed, Web of Science, Cochrane Library, Embase, CBM, CNKI, VIP, Wanfang database, references of the included literature

Continued

**Table 1** Continued

Study	Design	No. of patients	Population	Intervention	Comparison	Outcomes	The range of literature search
Xiaoyan Wen 2013 <sup>19</sup>	RCT	1735	Viral hepatitis	Chinese herbal retention enema and comprehensive treatment	Comprehensive treatment	Effectiveness rate, liver function index	Cochrane library, PubMed, Embase, VIP, CNKI, CBM, Wanfang, references of the included literatures, SIGLE <a href="http://www.opengrey.eu/search/request">http://www.opengrey.eu/search/request</a>
Pingping Zheng 2013 <sup>22</sup>	RCT	3194	Elders living in home	Tai Chi	Regular sport or physical therapy or blank control	Rate of fall, falls efficacy, time standing on one leg with eyes closed or open, body flexibility	Cochrane Library, Medline, EBSCO, CNKI, Wanfang, references of the included literature
Guohao Wang 2014 <sup>30</sup>	RCT Quasi-RCT	3084	Constipation	Acupoint massage and ventral massage	Routine nursing	Rate of effectiveness, defaecation frequency in the first day, defaecation frequency in the first 2 days, defaecation difficulty rate, defaecation time, dry stool rate, incomplete defaecation feeling rate, laxative provided	CNKI, VIP, CBM, Wanfang, EBSCO, PubMed, Cochrane Library
Jihuan Feng 2014 <sup>25</sup>	RCT	959	Patients with cancer receiving adjuvant chemotherapy	Acupoint massage	Routine nursing	Duration of chemotherapy-induced nausea, vomiting and retching, frequency of chemotherapy-induced nausea, vomiting and retching, severity of chemotherapy-induced nausea, vomiting and retching, antiemetic medication dosage, quality of life	PubMed, Medline, Embase, AMED, Cochrane Library, CBM, CNKI, VIP, Wanfang
Shaoxia Meng 2014 <sup>28</sup>	RCT	383	Diabetics with peripheral neuropathy	Foot massage or massage in foot reflection area or acupoint massage for lower limbs	Routine nursing	Rate of effectiveness, nerve conduction velocity	PubMed, CBM, CNKI, VIP, Wanfang database, references of the included literature
Xijuan Cui 2014 <sup>24</sup>	RCT, Quasi-RCT	860	Postoperative patients having abdominal operation	Acupoint massage for <i>Zusanli</i> point, <i>Zusanli</i> point acupuncture, Chinese medicine application at the <i>Zusanli</i> point	Routine nursing	First bowel sound time, first aerofluxus time, first defaecation time	CNKI, Wanfang

Continued

**Table 1** Continued

Study	Design	No. of patients	Population	Intervention	Comparison	Outcomes	The range of literature search
Zhong Sun 2014 <sup>29</sup>	RCT	524	Patients with primary dysmenorrhoea	Umbilical compression with Chinese herbs	Analgesic drug	Rate of effectiveness, cure rate, comparison of the score and scale of symptoms, rate of prostaglandin excretion, level of haemorheology and oestradiol, A/B, RI and PI	Cochrane Library, PubMed, Proquest, CNKI, VIP, Wanfang
Ye Li 2014 <sup>26</sup>	RCT	888	Women in labour	Acupoint massage for <i>Sanyinjiao, Hegu, Zhiyin, Taichong, Ashi, Shenshu</i> and others	Blank control	Rate of effectiveness of relieving labour pain	CNKI, VIP, CBM, Wanfang, PubMed
Yuanyuan Yang 2015 <sup>31</sup>	RCT	939	Insomnia	Auricular point therapy	Acupuncture or drug therapy	Rate of effectiveness	PubMed, Cochrane library, CBM, CNKI, VIP, Wanfang, references of the included literature

A/B, the ratio of the systolic peak value and the end-diastolic velocity of blood flow; AMED, Allied and Complementary Medicine Database; BMI, body mass index; CBM, Chinese Biomedical Literature database; CNKI, China National Knowledge Infrastructure; EBSCO, Elton Bryson Stephens Company; PI, pulse index; RI, resistance index; SDOL, Science Direct on Line; SIGLE, System for Information on Grey Literature in Europe; VIP, information/Chinese Scientific Journals database.

**Table 2** AMSTAR scores for the methodology of reviews included in this study

Study	Priori design	Data extraction	Comprehensive literature search	Status of publication	List of studies	Characteristics of the included studies	Quality assessment	Forming conclusion	Method for combining	Publication bias	Conflict of interest	Score/rank
Wangqin Shen <sup>13</sup>	1	1	0	0	0	0	1	0	0.5	0	0	4.5/medium
Hongying Pu <sup>15</sup>	1	1	0	0	0	0	1	1	1	1	0	6/medium
Na Li <sup>14</sup>	1	1	0	0	0	1	1	0	1	1	0	6/medium
Xuan Zhou <sup>16</sup>	1	1	0	0	0	0	1	1	0.5	1	0	5.5/medium
Xiaoli Wu <sup>17</sup>	1	1	0	0	0	0	1	1	0.5	1	0	6.5/medium
Xilan Zheng <sup>18</sup>	1	1	0	0	0	0	1	0	0.5	0	1	4.5/medium
Jiaqi Xu <sup>20</sup>	1	1	1	0	0	1	1	1	1	1	0	8/high
Yuan Zhao <sup>21</sup>	1	1	1	0	0	1	1	1	0.5	0	0	6.5/medium
Xiaoyan Wen <sup>19</sup>	1	1	1	1	0	1	1	1	1	0	0	8/high
Pingping Zheng <sup>22</sup>	1	1	1	0	0	1	1	1	0	0	0	6/medium
Guohao Wang <sup>30</sup>	1	1	0	0	0	1	1	1	1	1	0	7/medium
Jihuan Feng <sup>25</sup>	1	1	1	0	0	1	1	1	1	0	0	7/medium
Shaoxia Meng <sup>28</sup>	1	1	1	0	0	1	1	1	0.5	0	0	6.5/medium
Xijuan Cui <sup>24</sup>	1	1	1	0	0	1	1	1	1	1	0	8/high
Zhong Sun <sup>29</sup>	1	1	0	0	0	1	1	1	1	0	0	5/medium
Ye Li <sup>26</sup>	1	1	0	0	0	1	1	1	1	1	0	7/medium
Yuanyuan Yang <sup>31</sup>	1	1	1	0	0	0	1	1	1	0	0	7/medium
Baoxia Chang <sup>23</sup>	1	1	1	1	0	0	1	1	0.5	0	0	7.5/medium
Yue Ma <sup>27</sup>	1	1	0	0	0	1	1	1	1	1	0	7/medium
Weiwei Wu <sup>32</sup>	1	1	1	0	0	1	1	1	1	1	0	8/high
Total	1.00	1.00±0.00	0.50±0.51	0.10±0.31	0.00	0.65±0.49	1.00±0.00	0.85±0.37	0.78±0.30	0.45±0.51	0.05	6.58±1.10
Mean±SD	±0.00				±0.00						±0.22	

AMSTAR, Assessment of Multiple Systematic Reviews.



fixed effects) to determine the robustness of the conclusion. Most of them appropriately used the methodological quality of the included trials in formulating conclusions. None of them conducted evaluation of the quality of the body of evidence. All the included studies drew definitely positive conclusions in favour of TCMN interventions, while all reviewers suggested that there might be some benefits in the interventions. The findings should be interpreted with caution owing to the poor quality of trials or limited trial sample. Ten systematic reviews and meta-analyses (50%) assessed publication bias using funnel plots, and one review<sup>30</sup> used the Egger test. Only one<sup>18</sup> stated any conflict of interest.

### GRADE evidence quality

The two reviewers had a satisfactory agreement ( $\kappa=0.82$ ). None of the 20 studies cited any observational research, so upgraded items were excluded from the assessment of evidence quality. The evidence quality of all the included reviews is presented in [table 3](#).

For outcomes, there were adverse events (1/20, 5.0%) and symptoms (6/20, 30.0%), laboratory outcomes (5/20, 25.0%) and composite outcomes, such as total effectiveness rate (17/20, 85.0%), in the 20 reviews and no review considered the endpoint, economic evaluations or QoL. At the start, we determined the critical outcomes for each review. Judgements about what constitutes a critical outcome may change for different research goals and results. For instance, in a review entitled 'Acupressure wristbands prevent postoperative nausea and vomiting: a meta-analysis',<sup>16</sup> the goal was to evaluate the therapeutic effects on nausea and vomiting, so raters set nausea and vomiting as the critical outcome. Meanwhile, the outcomes of a systematic review of electroacupuncture combined with auricular point plaster therapy for patients with simple obesity<sup>23</sup> were the rate of effectiveness, body mass index (BMI) and waist circumference. The rate of effectiveness is equal to the numbers of patients recovering and the number for whom treatment was markedly effective or effective divided by the total, according to the author's description.

The criteria for recovery, markedly effective, effective and ineffective were stated as follows:

- ▶ Recovery: body weight was in the normal weight range or BMI  $<23 \text{ kg/m}^2$ ;
- ▶ Markedly effective: body weight decreased by no less than 5 kg, or BMI decreased by no less than  $2 \text{ kg/m}^2$ ;
- ▶ Effective: body weight decreased by no less than 2 kg and less than 5 kg, or BMI decreased by no less than  $0.5 \text{ kg/m}^2$  and less than  $2 \text{ kg/m}^2$ ;
- ▶ Ineffective: body weight decreased by  $<2 \text{ kg}$ , or BMI decreased by  $<0.5 \text{ kg/m}^2$ .

Because it was considered that the rate of effectiveness contained far more therapeutic information than BMI and waist circumference, the rate of effectiveness was set as the critical outcome by raters. In total, 31 bodies of evidence in the 20 reviews were assessed for quality.

### Rationale for downgrading

The quality of the evidence we assessed ranged from very low to moderate and no high-quality evidence was found.

The reasons for downrating confidence in effect estimates among the 31 bodies of evidence assessed were the risk of bias (26 times, 83.9%), inconsistency (16 times, 51.6%), indirectness (8 times 25.8%), imprecision (13 times, 41.9%) and publication bias (15 times, 48.4%). The detailed reasons for downgrading due to the risk of bias (80 times in total) included failure to conceal allocation (26 times, 32.5%), failure to blind (23 times, 28.8%), incomplete reporting of random sequence generation in most of the studies included (24 times, 30.0%), use of invalidated outcome measure (0 times, 0.0%), loss to follow-up and failure to adhere to the intention-to-treat principle (3 times, 3.8%), non-RCT included (4 times, 5.0%).

Downgrading for inconsistency was generally due to certain CIs showing little overlap from individual studies and significant heterogeneity. The quality of evidence was downgraded for indirectness nine times in total for the following reasons: substantial differences existed between the interventions (3, 33.3%) or the controls (5 times, 55.6%), or patient-important endpoints were replaced by surrogate endpoints (1 time, 11.1%). In a review evaluating the effectiveness of Tai Chi in preventing falls in the elderly, the authors stated that the scores of the Berg Balance Scale in the Tai Chi group were higher than in the control group and argued that Tai Chi can effectively reduce the risk of falls for elderly people. However, the Berg balance scale is a surrogate outcome for occurrence of fall.

The reasons for downgrading evidence for imprecision (13 times in total) included failure to meet optimal information size criterion (12 times, 92.3%) and wide CIs (1 time, 7.7%).

The quality of evidence was downgraded for publication bias (19 times in all) because of flaws in literature searching (12 times, 63.2%) and funnel plot asymmetry (7 times, 36.8%). The scatterplot showed no correlation between AMSTAR and GRADE instruments ([figure 3](#)).

## DISCUSSION

It is important to assess the methodological and evidence quality of a systematic review/meta-analysis before any conclusions can be reached about clinical decision-making. To the best of our knowledge, this is the first study to assess methodological and evidence quality of systematic reviews and meta-analyses of TCMN interventions in Chinese journals using AMSTAR and GRADE tools.

### Systematic reviews, meta-analyses and primary studies

All systematic reviews and meta-analyses and primary studies lacked important outcomes, which depressed the quality rating of evidence. GRADE specifies that both

**Table 3** GRADE evaluation of the quality of evidence of reviews included in this study

Study	Patients	Intervention	Comparison	Outcome	No. of patients/ studies	Effect RR/ OR/MD (95% CI)	Absolute effect (95%)	Quality	AMSTAR score
Baoxia Chang 2014 <sup>23</sup>	Simple obesity	Electroacupuncture combined with auricular point plaster	Electroacupuncture	Rate of effectiveness	397 (5)	RR=1.18 (1.07 to 1.25)	132 more per 1000 (from 51 more to 183 more)	Moderate (1a, 1c)	7.5
Wangqin Shen 2010 <sup>13</sup>	Phlebitis	Ruyijinhuangsan	Magnesium sulfate by wet compression	Rate of effectiveness	327 (6)	RR=1.32 (1.26 to 1.34)	236 more per 1000 (from 191 more to 250)	Very low (1a, 1b, 1c, 3a, 4a, 5a)	4.5
Hongying Pu 2011 <sup>15</sup>	Pressure ulcers	Moist exposed burn ointment	Skin disinfection solution or antibiotic ointment	Cure rate	432 (11)	RR=2.22 (2.07 to 2.33)	468 more per 1000 (from 410 more to 510 more)	Very low (1a, 1b, 1c, 2, 3a, 5a, 5b)	6
				Time of cure	261 (5)	MD=6.93 (7.7 to 6.15)	MD=6.93 lower (7.7 to 6.15 lower)	Very low (1a, 1b, 1c, 3a, 4a, 5a, 5b)	6
Xuan Zhou 2011 <sup>16</sup>	Postoperative patients	Acupressure wristbands	Placebo wristband	Incidence of nausea	1117 (9)	RR=0.85 (0.72 to 1)	46 fewer per 1000 (from 86 fewer to 0 more)	Moderate (4b,5a)	5.5
				Incidence of vomiting	1117 (9)	RR=0.44 (0.31 to 0.62)	107 fewer per 1000 (from 72 fewer to 131 fewer)	Moderate (4a, 5a)	5.5
Xiaoli Wu 2012 <sup>31 17</sup>	Diabetic foot ulcers	Foot bath therapy or foot massage with traditional Chinese medicine	Hot foot bath	Effective rate	845 (8)	RR=1.44 (1.4 to 1.46)	293 more per 1000 (from 267 more to 307 more)	Low (1a,1b, 2, 5a,5b)	6.5
Xilan Zheng 2012 <sup>18</sup>	Postoperative patients	Traditional Chinese medicine for elimination of necrotic tissues	Skin disinfection solution or antibiotic ointment	Cure rate	551 (9)	RR=1.89 (1.65 to 2.17)	401 more per 1000 (from 293 more to 527 more)	Low (1a, 1b, 1c, 2)	5.5
				Cure time	355 (6)	MD=9.33 (9.9 to 8.76)	MD=9.33 lower (9.9 to 8.76 lower)	Low (1a, 1b, 1c, 2)	5.5
Na Li 2011 <sup>14</sup>	Phlebitis	External application with aloe vera	Magnesium sulfate by wet compression	Effective rate	712 (7)	RR=1.25 (1.2 to 1.27)	192 more per 1000 (from 153 to 206)	Low (1a, 1b, 1c, 5a)	6
Jiaqi Xu 2013 <sup>20</sup>	Pressure ulcer	Resina draconis	Skin disinfection solution or antibiotic ointment	Effective rate	573 (13)	RR=1.2 (1.13 to 1.28)	162 more per 1000 (from 105 more to 227 more)	Moderate (1a, 1b, 1c)	8
Yuan Zhao 2013 <sup>21</sup>	Elderly individuals	Tai Chi	Regular sport or physical therapy	Rate of falls	1443 (4)	RR=0.82 (0.73 to 0.92)	83 fewer per 1000 (from 37 fewer to 124 fewer)	Moderate (3a)	6.5
				BBS	345 (2)	MD=2.45 (1.47 to 0.43)	MD=2.45 higher (1.47 to 3.43 higher)	Low (3a, 3b, 4a)	6.5

Continued

Table 3 Continued

Study	Patients	Intervention	Comparison	Outcome	No. of patients/ studies	Effect RR/ OR/MD (95% CI)	Absolute effect (95%)	Quality	AMSTAR score
Xiaoyan Wen 2013 <sup>19</sup>	Viral hepatitis	Chinese herbal retention enema and comprehensive treatment	Comprehensive treatment	Effective rate: after 2 weeks from cure time	260 (4)	RR=1.51 (1.3 to 1.67)	259 more per 1000 (from 152 more to 340 more)	Low (1a, 1b, 1c, 4a)	8
				Effective rate: after 4 weeks from cure time	333 (5)	OR=4.17 (2.37 to 7.32)	250 more per 1000 (from 173 more to 300 more)	Low (1a, 1b, 1c, 4a)	8
Pingping Zheng 2013 <sup>22</sup>	Elders living in home	Tai Chi	Regular sport or physical therapy or blank control	Rate of fall	2624 (9)	RR=0.85 (0.79 to 0.92)	73 fewer per 1000 (from 39 fewer to 102 fewer)	Moderate (2)	6
Guohao Wang 2014 <sup>30</sup>	Constipation	Acupoint massage and ventral massage	Routine nursing	Rate of effectiveness	2170 (19)	RR=1.93 (1.86 to 2)	396 more per 1000 (from 366 more to 426 more)	Low (1a, 1c, 2, 3a)	7
Jihuan Feng 2014 <sup>25</sup>	Patients with cancer receiving adjuvant chemotherapy	Acupoint massage	Routine nursing	Duration of chemotherapy-induced nausea	942 (7)	MD=1.52 (1.77 to 1.26)	MD=1.52 lower (1.77 to 1.26 lower)	Low (1a, 1b, 1c, 1e, 2)	7
				Frequency of chemotherapy-induced nausea	942 (7)	MD=1.08 (1.32 to 0.83)	MD=1.08 lower (1.32 to 0.83 lower)	Low (1a, 1b, 1c, 1e, 2)	7
				Severity of chemotherapy-induced nausea	942 (7)	MD=1.17 (1.37 to 0.96)	MD=1.17 lower (1.37 to 0.96 lower)	Low (1a, 1b, 1c, 1e, 2)	7
Shaoxia Meng 2014 <sup>28</sup>	Diabetics with peripheral neuropathy	Foot massage or massage in foot reflection area or acupoint massage for lower limbs	Routine nursing	Rate of effectiveness	323 (5)	RR=1.47 (1.29 to 1.68)	297 more per 1000 (from 183 more to 430 more)	Low (1a, 1c, 3a, 4a, 5a)	6.5
Xijuan Cui 2014 <sup>24</sup>	Postoperative patients with abdominal operation*	Acupoint massage for <i>Zusanli</i> , <i>Zusanli</i> point acupuncture, Chinese medicine application at the <i>Zusanli</i> point	Routine nursing	First aerofluxus time: subgroup for <i>Zusanli</i> point acupuncture*	317 (3)	MD=14.52 (15.49 to 13.54)	MD=14.52 lower (15.49 to 13.54 lower)	Low (1a, 1b, 1c, 1f, 2, 4a, 5a)	8
				Subgroup for acupoint massage for <i>Zusanli</i> *	326 (4)	MD=22.7 (25.67 to 19.73)	MD=22.7 lower (25.67 to 19.73 lower)	Low (1a, 1b, 1c, 1f, 2, 4a, 5a)	8
				Subgroup for Chinese medicine application at the <i>Zusanli</i> point*	1048 (6)	MD=18.25 (18.6 to 17.9)	MD=18.25 lower (18.6 to 17.9 lower)	Moderate (1a, 1b, 1c, 1f, 2, 5a)	8
Zhong Sun 2014 <sup>29</sup>	Patients with primary dysmenorrhoea	Umbilical compression with Chinese herbs	Analgesic drug	Rate of effectiveness	496 (5)	RR=1.93 (1.45 to 2.57)	214 more per 1000 (from 104 more to 362 more)	Very low (1a, 1b, 2, 3a, 4a)	5

Continued

Table 3 Continued

Study	Patients	Intervention	Comparison	Outcome	No. of patients/studies	Effect RR/OR/MD (95% CI)	Absolute effect (95%)	Quality	AMSTAR score
Ye Li 2014 <sup>26</sup>	Women in labour	Acupoint massage for <i>Sanyinjiao</i> , <i>Hegu</i> , <i>Zhiyin</i> , <i>Taichong</i> , <i>Ashi</i> , <i>Shenshu</i> and others	Blank control	Rate of effectiveness	766 (6)	RR=1.64 (1.56 to 1.7)	347 more per 1000 (from 304 more to 380 more)	Low (1a, 1b, 1c, 2, 5a, 5b)	7
Yuanyuan Yang 2015 <sup>31</sup>	Insomnia*	Auricular point therapy	Acupuncture or drug therapy	Effective rate after 2 weeks from cure time	606 (3)	RR=1.28 (1.2 to 1.37)	212 more per 1000 (from 152 more to 281 more)	Low (1a, 1b, 1c, 2)	7
		Auricular point therapy	Acupuncture or drug therapy	Effective rate after 4 weeks from cure time	333 (4)	RR=1.25 (1.13 to 1.37)	186 more per 1000 (from 97 more to 276 more)	Low (1a, 1b, 1c, 4a)	7
Yue Ma 2014 <sup>27</sup>	Ileus	Enema and gastrointestinal intubation with traditional Chinese medicine	Blank control	Rate of effectiveness	2821 (27)	RR=1.24 (1.2 to 1.29)	179 more per 1000 (from 149 more to 216 more)	Low (1a, 1b, 1c, 1f, 5b)	8
Weiwei Wu 2016 <sup>32</sup>	Elderly individual	Traditional Chinese exercise	Other intervention or regular nursing	Subgroup for PSQI for Tai Chi	554 (8)	MD=-2.15 (-4.61 to 0.30)	MD=2.15 lower (4.61 lower to 0.3 higher)	Low (1a, 1b, 1c, 2, 5b)	8
				Subgroup for PSQI for Qi Gong	55 (2)	MD=-4.29 (-5.29 to -3.29)	MD=4.29 lower (5.29 lower to 3.29 lower)	Low (1a, 1b, 1c, 4a, 5b)	8

Risk of bias: (1a) failed to conceal allocation; (1b) no blinding used; (1c) incomplete reporting of random sequence generation in most studies included; (1d) use of unvalidated outcome measure; (1e) loss to follow-up and failure to adhere to the intention-to-treat principle; (1f) non-randomised controlled trial was included.

Inconsistency: (2) unexplained heterogeneity or inconsistency of results.

Indirectness: (3a) differences in therapeutic methods between intervention and control groups; (3b) surrogate outcome.

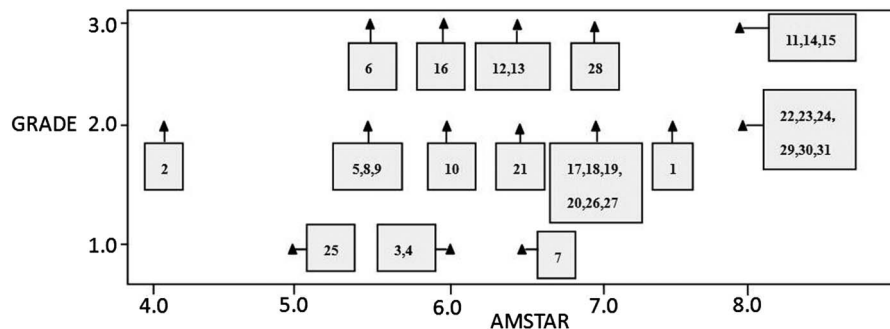
Imprecision: (4a) optimal information size criterion was not met; (4b) wide confidence intervals.

Publication bias; (5a) flaws in literature search; (5b) funnel plot asymmetry.

\*Subgroup.

AMSTAR, Assessment of Multiple Systematic Reviews; GRADE, Grading of Recommendations Assessment, Development and Evaluation; PSQI, psoriasis; PSQI, Pittsburgh Sleep Quality Index; RR, risk ratio; MD, mean difference.

**Figure 3** Scatter plot for exploring correlation between AMSTAR and GRADE instruments. AMSTAR, Assessment of Multiple Systematic Reviews; GRADE, Grading of Recommendations Assessment, Development and Evaluation.



those conducting systematic reviews and those developing practice guidelines should begin by specifying every important outcome of interest.<sup>33</sup> Unfortunately, systematic reviews and meta-analyses in our study usually did not deal with all important outcomes. For instance, a review aiming to verify the effect of foot bath therapy or foot massage with TCM for diabetic foot ulcers did not consider amputation as an outcome, although foot ulcers are a high risk factor for infection, gangrene, amputation and even death among patients with diabetes. We thought the amputation rate might be the preferable long-term outcome to verify the effect of the intervention but this was not considered by the reviewers.<sup>17</sup>

In general, systematic reviews and meta-analyses should include all outcomes that are likely to be meaningful to clinicians, patients, the general public, administrators and policy makers. For example, outcomes may include survival, clinical events, patient-reported outcomes (eg, symptoms or QoL), adverse events, burdens (eg, demands on caregivers, frequency of tests, restrictions on lifestyle) and economic outcomes (eg, cost and resource use). But primary studies typically focused on short-term benefit without considering long-term outcomes, harm or economic outcomes. None of the reviews listed any adverse effects as outcomes of TCMN interventions. These all confirmed that systematic reviews and meta-analyses and primary studies all had shortcomings in research design, which also made it difficult to use the results of systematic reviews and meta-analyses to make appropriate recommendations as these were based on incomplete outcomes.

### Risk of bias

The risk of bias resulted in downgrading in most reviews.

RCTs are critical for assessing and providing valuable evidence about the effectiveness of TCMN interventions. However, the reliability and acceptability of the results of any intervention study depend on the extent to which the studies employ scientific principles and use a valid research design. In this study, we used the Cochrane Collaboration's tool for assessing risk of bias.<sup>11</sup> Most of the reviews were downgraded because of lack of allocation concealment and blinding and lack of details of randomisation in primary studies. A total of 228 RCTs

were included in 20 systematic reviews and meta-analyses, of these the authors noted that 184 (80.7%) were published in a Chinese journal.

This finding is consistent with the results of similar research. Yao *et al* conducted a systematic review using GRADE to assess the quality of evidence of Chinese meta-analyses. They indicated that risk of bias was the most common factor for downgrading evidence in Chinese meta-analyses and emphasised that the inferior quality of evidence in meta-analyses related to TCM studies might be caused by the poor-quality reporting in RCTs.<sup>34</sup> Wu found that more than 90% of RCTs published in core Chinese journals lacked an adequate description of randomisation in 2009 and most trials despite claiming to be RCTs did not fulfil the criteria for a true RCT.<sup>35</sup> Although the number of RCTs in nursing research in China is increasing, the quality of most of them remains unsatisfactory. Xing *et al*<sup>36</sup> published a comprehensive evaluation of 7391 nursing intervention studies published in simplified Chinese between 1979 and 2012. Their results showed that among the 10 characteristics considered in quality evaluations, the lowest ratings were for 'use of a blind method', 'description of loss to follow-up', 'appropriate calculation of sample size' and 'randomised assignment of patients to treatments'.

The usefulness of current systematic reviews and meta-analyses as guidelines is often limited because they rate risk of bias by studies across outcomes rather than by outcome across studies.<sup>37</sup> Authors of systematic reviews and meta-analyses should bear in mind that the importance of sources of bias may vary across outcomes; it means that summarising study limitations must be outcome specific.<sup>37</sup> For example, the assessors downgraded the evidence many times for not using blinding in studies with subjective outcomes, which are much more vulnerable to biased judgements. The above-mentioned review evaluating the effect of foot bath therapy or foot massage for diabetic foot ulcers used rate of effectiveness as the only outcome, which was based on subjective observable judgement of the condition (eg, ulcer area, local swelling and skin colour). Raters categorised this evidence as having serious study limitations on account of lack of blinding during the study.<sup>17</sup> Problems with the design and execution of individual studies of TCMN interventions raises questions



about the validity of their effects and results in downgrading of the quality of evidence.

### Heterogeneity

In systematic reviews and meta-analyses, heterogeneity has not been adequately explored or there is an inappropriate combination of the studies' findings which usually reduces the quality of evidence on the grounds of inconsistency.

The raters decreased the quality of evidence when significant heterogeneity was detected for which the authors had failed to identify a reasonable source or provide an explanation. Although studies brought together in a systematic review will inevitably have some differences, reviewers should look for robust explanations for any significant heterogeneity.<sup>37</sup>

Clinical variation will lead to heterogeneity if the intervention effect is affected by the factors that vary across individual studies—for example, most obviously, patient characteristics or specific interventions. In our study, variability in interventions was the most common reason for heterogeneity—for example, different points for acupoint massage, different medicine for Chinese herbal retention enema or different Chinese herbal prescriptions for umbilical compression. Raters considered the true intervention effect might be different in different studies and reduced the quality of evidence for inconsistency of results, and methodological quality for inappropriate combination of the study's findings.

When heterogeneity cannot readily be explained, incorporating it into a random-effects model is often the only option for reviewers. Reviewers should know that a random-effects model does not 'take account' of the heterogeneity. When the meta-analysis results show that heterogeneity is statistically significant, the most important treatment method is to analyse possible reasons for heterogeneity rather than simply using the random-effects model.

Information about study limitations, imprecision, inconsistency, indirectness and publication bias is necessary for TCMN in order to understand, and have confidence in, the assessment of quality and the estimate of effect size. GRADE provides a framework for assessing outcome quality that encourages transparency and explicit accounting for the judgements made. In this study, the quality of evidence was low in 20 and very low in 4 cases among the 31 bodies of evidence. High-quality evidence is more likely to be associated with a strong recommendation, but it is important to note that sometimes low or very low quality evidence may lead to a strong recommendation. When using the evidence for TCMN interventions, nurses should consider patient values and preferences and resource implications as well as confidence in estimates of the effect of primary outcome used in the GRADE system. In addition, in view of the unsatisfactory methodological and evidence quality of systematic reviews and meta-analyses of TCMN which we included, we considered it important to

present to readers the inbuilt problems of systematic reviews and meta-analyses of TCMN interventions rather than presenting readers only with the available evidence.

### Methodological quality assessment using AMSTAR

Methodological quality assessment using AMSTAR should be a precondition for further evaluation with the GRADE approach.

Systematic reviews/meta-analyses may differ considerably in their methodological quality.<sup>38</sup> Using a rigorous methodology with a clearly formulated research question and a comprehensive search strategy, systematic reviews should provide reproducible results and include all potentially relevant studies, thereby limiting bias and random errors.<sup>33</sup> Systematic reviewers will clearly specify the interventions of interest in their eligibility criteria, ensuring that only directly relevant studies are eligible. However, our study showed that several systematic reviews/meta-analyses included studies inconsistent with their eligibility criteria. For instance, a review aiming to evaluating foot massage for diabetic peripheral neuropathy, included studies with different interventions: foot massage or massage in the foot reflection area or acupoint massage for lower limbs.

In addition, to minimise bias, systematic reviews require a thorough, objective and reproducible search of a range of sources to identify as many relevant studies as possible. In this respect, some systematic reviews/meta-analyses in this study are far from satisfactory. They set up a flawed search strategy—for example, using only free-text searching without performing Mesh (index terms) searching, searching only the Chinese database and failing to identify 'negative' studies. Moreover, in our study, assessors found that high-quality clinical trials do not always exist, especially in TCMN, and non-RCTs were sometimes included in systematic reviews/meta-analyses. The problem with the methodology in some systematic reviews/meta-analyses discussed above resulted in lowering of AMSTAR scores, and downgrading of the quality of evidence. Although GRADE guidelines suggest that GRADE should not be used for systematic reviews/meta-analyses with serious flaws, we did not exclude any study since no review was rated as having a low methodological quality score, although some imperfect methodology does exist in the studies reviewed.

In our study, we were unable to identify any correlation between methodological quality and quality of evidence using a scatterplot. This is understandable because GRADE is much more than a simple rating system. It offers a transparent and structured process for developing and presenting evidence summaries for systematic reviews,<sup>8</sup> surveying some methodological characteristics of the production of systematic reviews/meta-analyses which influence the quality of evidence and also exploring factors resulting in inconsistency or imprecision.



Methodological flaws in the quality of systematic reviews/meta-analyses could severely affect decision-making and the application of evidence. We suggest that methodological quality should be assessed before assessment of the quality of evidence. There is no need to evaluate the evidence quality for a systematic review/meta-analysis for which a low methodological quality score is assigned owing to major flaws.

### Suboptimal reporting

Suboptimal reporting may contribute to an underestimation of methodological quality. Some of the research items for GRADE and AMSTAR rely on transparency in reporting in the systematic reviews/meta-analyses document. Even the most methodologically rigorous process, if not clearly described, will leave assessors or users uncertain about the reliability of the systematic reviews/meta-analyses in question.

Most Chinese journals impose strict limits on word numbers. The Chinese journal editors usually encourage authors to focus on the research results and discussion sections of their manuscripts and shorten the methods section of their papers. Even the *Chinese Journal of Nursing*, a leading domestic journal in China, typically limits the length of articles on nursing intervention studies to no more than four pages.<sup>38</sup> Although we made an attempt to contact developers of the reviews we included, we found this difficult as most authors' contact information was not presented in the published papers. So the results in this study were possibly underestimated owing to a lack of some important information. PRISMA (Preferred Reporting Items of Systematic Reviews and Meta-analyses), which consists of 27-item checklist and a four-phase flow diagram, informs authors of the preferred way to present every part of a report of a systematic review/meta-analysis. We hope that editors of medical journals in China recognise and promote the use of reporting guidelines in their publications and that authors will adhere to them.

### Implications for research and practice

High methodological quality is the basic precondition of systematic reviews for identifying the best available evidence for specific research questions and conducting GRADE evaluation. Authors and editors of systematic reviews/meta-analyses should make every effort to adhere to well-established methodological standards to enhance the impact of their research efforts. But high methodological quality does not fully reflect the quality of a review, the quality of a body of evidence is critical in decision-making. The GRADE approach can provide clinicians and patients with guidance in using results from systematic review/meta-analysis in clinical practice and provide policy makers with a guide to their use in developing health policy.

The overall quality of systematic reviews/meta-analyses of TCMN interventions published in Chinese journals remains suboptimal, especially their risk of bias, which

reduces the quality of evidence for almost all indications, raising concerns about their role in influencing clinical practice. Therefore, their conclusions needs to be treated with caution. Critical appraisal of systematic reviews/meta-analyses of TCMN interventions is particularly important.

### Author affiliations

<sup>1</sup>Evidence-Based Nursing Center, School of Nursing, Tianjin University of Traditional Chinese Medicine, Tianjin, China

<sup>2</sup>Nursing Department, North China University of Science and Technology Affiliated Hospital, TangShan, China

<sup>3</sup>Graduate College, Tianjin University of Traditional Chinese Medicine, Tianjin, China

<sup>4</sup>Public Health Department of Tianjin University of Traditional Chinese Medicine, Tianjin, China

<sup>5</sup>Library of Tianjin University of Traditional Chinese Medicine, Tianjin, China

<sup>6</sup>Emergency Department, Tianjin TEDA hospital, Tianjin, China

<sup>7</sup>Key Laboratory of Chinese Internal Medicine of Ministry of Education and Beijing, Dongzhimen Hospital, Beijing University of Chinese Medicine, Beijing, China

**Contributors** Y-HJ and H-CS designed the study; G-HW, QL, and Y-RS searched the databases for full-text papers; GL, J-HS, and YL extracted and analysed the data; Y-HJ, CZ, and H-CS performed the critical appraisal; Y-HJ and CL wrote the manuscript, H-CS and J-HS reviewed the manuscript.

**Competing interests** None declared.

**Provenance and peer review** Not commissioned; externally peer reviewed.

**Data sharing statement** No additional data are available.

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

### REFERENCES

1. Adib-Hajbaghery M, Hoseinian M. Knowledge, attitude and practice toward complementary and traditional medicine among Kashan health care staff, 2012. *Complement Ther Med* 2014;22:126–32.
2. Chan HY, Chui YY, Chan CW, *et al.* Exploring the influence of Traditional Chinese Medicine on self-care among Chinese cancer patients. *Eur J Oncol Nurs* 2014;18:445–51.
3. Ministry of Health of the People's Republic of China. The development program of nursing in China (2011–2015). *Zhonghua Hu Li Za Zhi* 2012;47:286–8.
4. Zhang SQ, Chen LL, Zhou JM, *et al.* Promoting the subject development using the construction of nursing key specialty of traditional Chinese medicine. *Zhongguo Huli Guanli* 2013;10:4–6.
5. Committee on Standards for Developing Trustworthy Clinical Practice Guidelines, Board on Health Care Services, Institute of Medicine, *et al.* *Clinical Practice Guidelines We Can Trust*. Washington DC: National Academies Press, 2011.
6. Gartlehner G, Sommer I, Evans TS, *et al.* Grades for quality of evidence were associated with distinct likelihoods that treatment effects will remain stable. *J Clin Epidemiol* 2015;68:489–97.
7. Fleming PS, Koletsi D, Seehra J, *et al.* Systematic reviews published in higher impact clinical journals were of higher quality. *J Clin Epidemiol* 2014;67:754–9.
8. Balshem H, Helfand M, Schünemann HJ, *et al.* GRADE guidelines: 3. Rating the quality of evidence. *J Clin Epidemiol* 2011;64:401–6.
9. Guyatt G, Oxman AD, Akl EA, *et al.* GRADE guidelines: 1. Introduction-GRADE evidence profiles and summary of findings tables. *J Clin Epidemiol* 2011;64:383–94.
10. Shea BJ, Hamel C, Wells GA, *et al.* AMSTAR is a reliable and valid measurement tool to assess the methodological quality of systematic reviews. *J Clin Epidemiol* 2009;62:1013–20.
11. Zeng X, Zhang Y, Kwong JS, *et al.* The methodological quality assessment tools for preclinical and clinical studies, systematic

- review and meta-analysis and clinical practice guideline: a systematic review. *J Evid Based Med* 2015;8:2–10.
12. Mustafa RA, Santesso N, Brozek J, *et al.* The GRADE approach is reproducible in assessing the quality of evidence of quantitative evidence syntheses. *J Clin Epidemiol* 2013;66:736–42; quiz 42.e1–5.
  13. Shen W, Qian H, Yu H. Meta-analysis of efficacy on good fortune golden powder to treat patients with phlebitis. *Chin Nurs Res* 2010;24:85–6.
  14. Li N. Meta analysis of the comparison of the effects between external application with Aloe vera and wet dressings with magnesium sulfate in the treatment of phlebitis. *J Qilu Nurs* 2011;17:3–5.
  15. Pu H, Cheng W, He J. The effect of moist exposed burn ointment (MEBO) on pressure ulcers: a meta-analysis. *J Nurs Sci* 2011;26:79–81.
  16. Zhou X, Wang Q. Acupressure wristbands prevent postoperative nausea and vomiting: a meta-analysis. *J Nurs Sci* 2011;26:81–4.
  17. Wu X, Yin L, Ji H. Meta-analysis on effects of clinical nursing of integrated traditional Chinese medicine. *Nurs J Chin People Liberation Army* 2012;29:24–6.
  18. Zheng X, Jiang Z, Hu R. Effect of traditional Chinese medicine in eliminating necrotic tissues and promoting granulation on stage III to IV pressure ulcer: a Meta analysis. *Chin J Pract Nurs* 2012; 28:1–4.
  19. Wen XY, Meng FJ, Jin YH, *et al.* Effectiveness of Chinese herbal retention enema in viral hepatitis patients: a meta-analysis. *Chin J Evid Based Med* 2013;13:339–45.
  20. Xu JQ, He YN, Li J, *et al.* Effectiveness and safety of resina draconis for pressure ulcer: a systematic review. *Chin J Evid Based Med* 2013;13:1236–43.
  21. Zhao Y, Wang Y, Xu XD, *et al.* Effectiveness of Tai Chi in fall prevention and balance function in the elderly: a meta-analysis. *Chin J Evid Based Med* 2013;13:339–45.
  22. Zheng P, Zhang J, Tong L. Meta analysis of the effect of Tai Chi on reducing falls among elders living at home. *Chin J Mod Nurs* 2013;19:1123–7.
  23. Chang B, Meng F. Meta-analysis on effect of electro-acupuncture combined with auricular point plaster therapy for patients with simple obesity. *Chin Nurs Res* 2014;28:884–7.
  24. Cui X, Qiao L, Shan T. Effect of acupuncture using zusanli point on postoperative function of gastrointestinal tract: a systematic review. *Today Nurse* 2014;50–4.
  25. Feng J, Yang G, Jiao L. Effect of acupressure on chemotherapy-induced digestive tract reaction for malignant tumor patients. *Chin J Pract Nurs* 2014;30:51–5.
  26. Li Y, Tang L. Effect of acupoint massage on labor pain relief in women: a meta-analysis of randomized controlled trials. *J Nurs* 2014;21:12–15.
  27. Ma Y, Meng F, Jin Y, *et al.* Chinese herbal enema plus gastrointestinal intubation for ileus: a systematic review. *Chin J Evid Based Med* 2014;14:1254–62.
  28. Meng Z, Zhi C, Li Y. System evaluation of foot massage for treatment of patients with diabetic peripheral neuropathy. *Chin Nurs Res* 2014;28:3187–9.
  29. Sun Z, Wang Q, Li J, *et al.* Systematic review on intervention effect of compressing umbilical with Chinese herbal for primary dysmenorrheal patients. *Chin Nurs Res* 2014;28:506–10.
  30. Wang G, Jin Y, Wang X, *et al.* Prevention and treatment effect of acupressure and ventral massage in constipation: a systematic review. *Chin J Pract Nurs* 2014;30:30–35.
  31. Yang Y, Wang Y, Li W, *et al.* Effectiveness of auricular point insomnia: a meta-analysis. *J Nurs Sci* 2015;30:4–8.
  32. Wu W, Lan X, Kwong W, *et al.* The effects of traditional exercises on sleep quality in older adults: a meta-analysis. *Chin J Nurs* 2016;51.
  33. Guyatt GH, Oxman AD, Kunz R, *et al.* GRADE guidelines: 8. Rating the quality of evidence—indirectness. *J Clin Epidemiol* 2011;64:1303–10.
  34. Yao L, Sun R, Chen YL, *et al.* The quality of evidence in Chinese meta-analyses needs to be improved. *J Clin Epidemiol* 2016;74:73–9.
  35. Wu T, Li Y, Bian Z, *et al.* Randomized trials published in some Chinese journals: how many are randomized? *Trials* 2009;10:46.
  36. Xing W, Fu L, He M, *et al.* A quality evaluation of nursing intervention studies in mainland China: from 1979 to 2012. *Int J Nurs Sci* 2014;1:145–50.
  37. Guyatt GH, Oxman AD, Vist G, *et al.* GRADE guidelines: 4. Rating the quality of evidence—study limitations (risk of bias). *J Clin Epidemiol* 2011;64:407–15.
  38. Remschmidt C, Wichmann O, Harder T. Methodological quality of systematic reviews on influenza vaccination. *Vaccine* 2014;32:1678–84.