
This is the author’s final accepted version.

There may be differences between this version and the published version. You are advised to consult the publisher’s version if you wish to cite from it.

[http://eprints.gla.ac.uk/148733/](http://eprints.gla.ac.uk/148733/)

Deposited on: 18 April 2018

Enlighten – Research publications by members of the University of Glasgow
[http://eprints.gla.ac.uk](http://eprints.gla.ac.uk)
Methological quality of systematic reviews and meta-analyses on acupuncture for stroke: a review of reviews

CHEN Xin-lin¹, MO Chuan-wei¹, LU Li-ya², GAO Ri-yang¹, XU Qian¹, WU Min-feng³, ZHOU Qian-yi³, HU Yue¹, ZHOU Xuan¹, and LI Xian-tao¹

1. College of Basic Medical Science, Guangzhou University of Chinese Medicine, Guangzhou (510006), China;
2. Public Health, Institute of Health and Wellbeing, University of Glasgow, Glasgow, Scotland (G12 8RZ), UK;
3. First Clinical College, Guangzhou University of Chinese Medicine, Guangzhou (510405), China

Correspondence to: Prof. LI Xian-tao, Tel: 86-20-39358629, E-mail: lxt150@126.com

*Supported by the National Natural Science Foundation of China (No. 81403296); the Outstanding Youth Foundation of Guangdong Province Colleges and Universities (No. YQ2015041); Guangdong High Level Universities Program of Guangzhou University of Chinese Medicine
ABSTRACT

Objective: To assess the methodological quality of systematic reviews and meta-analyses regarding acupuncture intervention for stroke and the primary studies within them.

Methods: Two researchers searched PubMed, Cumulative index to Nursing and Allied Health Literature, Embase, ISI Web of Knowledge, Cochrane, Allied and Complementary Medicine, Ovid Medline, Chinese Biomedical Literature Database, China National Knowledge Infrastructure, Wanfang and Traditional Chinese Medical Database to identify systematic reviews and meta-analyses about acupuncture for stroke published from the inception to December 2016. Review characteristics and the criteria for assessing the primary studies within reviews were extracted. The methodological quality of the reviews was assessed using adapted Oxman and Guyatt Scale. The methodological quality of primary studies was also assessed.

Results: Thirty-two eligible reviews were identified, 15 in English and 17 in Chinese. The English reviews were scored higher than the Chinese reviews (P=0.025), especially in criteria for avoiding bias and the scope of search. All reviews used the quality criteria to evaluate the methodological quality of primary studies, but some criteria were not comprehensive. The primary studies, in particular the Chinese reviews, had problems with randomization, allocation concealment, blinding, dropouts and withdrawals, intent to treat analysis and adverse events.

Conclusions: Important methodological flaws were found in Chinese systematic reviews and primary studies. It was necessary to improve the methodological quality and reporting quality of both the systematic reviews published in China and primary studies on acupuncture for stroke.

KEYWORDS
systematic review, meta-analyses, stroke, acupuncture, review of review, methodological quality
INTRODUCTION

Stroke is the second most common cause of death and disability worldwide.\(^{(1)}\) More than two thirds of global stroke occurs in developing countries.\(^{(2)}\) In China, stroke is also the second leading cause of death, with a similar incidence and prevalence but more intracerebral haemorrhage than that in developed countries.\(^{(3,4)}\) The management in a stroke care unit, intravenous tissue plasminogen activator within 3 h, aspirin within 48 h of stroke onset, and decompressive surgery for supratentorial malignant hemispheric cerebral infarction are proven to be effective in the treatment of stroke patients.\(^{(1)}\) Acupuncture, a major method of Chinese medicine (CM), has been used as a treatment for many diseases for over 3,000 years, and can be applied to treat post-stroke rehabilitation.\(^{(5,6)}\) Some researches reported that acupuncture could improve neuragenesis by promoting the proliferation and differentiation of neural stem cells in brain tissues.\(^{(7,8)}\)

Some reviews have been used to evaluate the effectiveness of acupuncture for stroke patients,\(^{(9-11)}\) but these reviews still have some problems. Firstly, most of these reviews focused mainly on the effectiveness of a particular acupuncture therapy. The methodological quality or risk of bias of primary studies within them was not fully evaluated. Secondly, the Preferred Reporting Items for Systematic Reviews and Meta-analyses (PRISMA statement), which have been developed and recommended to evaluate the systematic reviews,\(^{(12,13)}\) were not used frequently. The poor reporting of systematic reviews seriously affected the conclusions. Different scopes of search or different criteria of inclusion and exclusion will result in discordance among the conclusions.\(^{(14,15)}\)

A review of reviews about stroke rehabilitation therapy was performed to assess the methodological issues of systematic reviews. The methodological flaws of the reviews,
in particular non-Cochrane reviews, were identified.\(^{(16)}\) The methodological quality of systematic reviews about acupuncture for stroke patients has not been evaluated so far. It is not clear whether these reviews also suffer methodological flaws. Therefore, our main aim was to assess the methodological flaws of systematic reviews in the field of acupuncture for stroke. The reporting quality of primary studies had an influence on the reviews. We also assessed the methodological quality of primary studies within them.

**METHODS**

**Data source**

An electronic search of the published review articles on stroke patients receiving acupuncture was performed using PubMed, Cumulative index to Nursing and Allied Health Literature (CINAHL), Embase, ISI Web of Knowledge, Cochrane, Allied and Complementary Medicine, Ovid Medline, Chinese Biomedical Database, China National Knowledge Infrastructure (CNKI), Wanfang and Traditional Chinese Medical Database. The following keywords were used (stroke, apoplexy, cerebrovascular attack, cerebrovascular accident, cerebrovascular\(^*\), cerebral infarction OR cerebral hemorrhage) AND (acupuncture, needling, moxibustion OR acu\(^*\)) AND (systematic review, meta-analyses OR meta-analysis). Only studies undertaken on humans and published from inception to December 31, 2016 in English or Chinese were included.

**Inclusion and exclusion criteria**

Inclusion was restricted to systematic reviews or meta-analyses that: reported the effectiveness based on randomized controlled trials (RCTs), quasi-RCTs, or clinical controlled trails (CCTs); compared different types of acupunctures (including moxibustion and needling), compared acupuncture with placebo acupuncture, sham acupuncture or no acupuncture, or compared acupuncture plus conventional therapies
(physical rehabilitation, conventional medication, or Chinese herbal medicine) with placebo acupuncture, sham acupuncture or plus conventional therapies (physical rehabilitation, conventional medication, or CM); assessed methodological quality. Gray reviews and unpublished reviews of letters, government documents, research reports, and conference proceedings were also searched. The corresponding authors were contacted to share their reviews if full text was unavailable. The reviews were excluded if they only compared other types of conventional therapies other than acupuncture or compared different diagnostic tests in stroke. Reviews on the association between polymorphism and stroke were excluded. Commentary, conference articles, these articles, narrative or descriptive reviews were also excluded.

**Literature screening and extraction**

Two researchers (Wu MF and Zhou QY) screened and assessed all titles and abstracts independently and made a selective decision. Disagreements on final inclusion were settled by discussion and consensus with another author (Chen XL).

The following information was extracted from each of the eligible studies: author, year of publication, inclusion criteria, type of stroke, intervention, sample size, number of primary studies, outcomes and results. Reasons for the exclusion of reviews were also recorded. The methodological quality of the reviews and primary studies within them were also collected by the two researchers.

**Quality assessment**

The methodological quality of systematic reviews was assessed using an adapted scoring approach based on the 9-item Oxman and Guyatt Scale. The scale consisted of 9 items: search methods used to find evidence (Item 1), scope of search (Item 2), criteria for inclusion and exclusion (Item 3), criteria for avoiding bias (Item 4), reporting of criteria for assessing the validity of primary studies, assessing the
validity of primary studies using appropriate criteria (Item 6), methods of combining the data (Item 7 and 8), and conclusions supported by the data or analysis reported in the review (Item 9). The score of each item ranged from 0 point to 2 points, resulting in a maximum score of 18 points. A score 10 was considered as low overall quality, between 11 to 13 as moderate overall quality, and 14 as high overall quality.\(^{(17,18)}\)

The methodological quality of primary studies within the reviews including randomization, allocation concealment, baseline comparability, blinding (patients, health care providers, data collectors and outcome assessors), dropouts and withdrawals, intention-to-treat (ITT) analysis, adverse events, and contamination and cointervention, was also assessed using frequencies and percentages.

**Statistical analysis**

All analyses were performed using SPSS 22.0 software (SPSS Inc., Chicago, IL, USA). The data were expressed as mean and standard deviation (\(\bar{x} \pm s\)). The proportion of the methodological quality (binary data) in the primary studies was calculated. The two independent-samples \(t\)-test was used to compare the scores between English reviews and Chinese reviews. \(P<0.05\) was considered statistically significant.

**RESULTS**

**Reviews characteristics**

A total of 923 articles were initially enrolled and screened. Among them, 845 articles were excluded as duplicates or failure in meet the inclusion criteria. The remaining 78 articles were searched for the full text. A systematic review about acupuncture for hypoxic ischemic encephalopathy in neonates was excluded because no trials satisfied the authors’ predefined inclusion criteria.\(^{(19)}\) In total, 32 systematic reviews met the
criteria and were included for further data extraction (Figure 1).\textsuperscript{(9-11,20-48)}

Of the 32 eligible reviews, 15 in English \textsuperscript{(9,10,20-32)} and 17 reviews were in Chinese (Appendix 1).\textsuperscript{(11,33-48)} Only 2 reviews were Cochrane systematic papers,\textsuperscript{(10,31)} and others were traditional peer-reviewed papers. Twelve reviews have been published since 2010. Fourteen reviews included all stages of stroke such as acute, sub-acute, chronic, and included all types of stroke such as ischemic or hemorrhagic stroke. Eight reviews were about post-stroke hemiplegic (PSH). The treatment group and control group differed within each review. The two groups usually could be classified into 2 types of comparisons: acupuncture (needling or moxibustion) or acupuncture plus conventional therapies vs. shame acupuncture, placebo acupuncture, no acupuncture or conventional therapies; acupuncture (needling or moxibustion) or acupuncture plus conventional therapies vs. another acupuncture or another acupuncture plus conventional therapies. Conventional therapies usually contained rehabilitation, conventional medications, and CM.

The number of eligible primary studies varied from 5 to 56. All the reviews were predominantly based on RCTs, quasi-RCTs, or CCTs. The number of patients evaluated in the reviews varied from 368 to 5650. The median of treatment group patients was 668 (range: 193 to 3156); while the median of control group patients was 617 (range: 175 to 2494).

**Methodological quality of systematic reviews**

Twenty-three (92.0\%) systematic reviews were scored 14 based on the adapted Oxman and Guyatt scale (Appendix 2). The lowest score was 13.\textsuperscript{(9,47)} Two Cochrane systematic reviews were scored 18,\textsuperscript{(10,31)} which was related to the method of standardized reporting and enough word allowances.

The score of English reviews were higher than the Chinese reviews (P=0.025), with
total mean scores 16.73 and 15.29, respectively (Table 1). The scores of item 1, item 2 and item 4 between English reviews and Chinese reviews were also significantly different (P<0.05). In Chinese reviews, item 4 "Was bias in the selection of articles avoided?" was scored 0.88 averagely, vs. 1.73 in English reviews. The mean scores of Item 1 in English reviews were 2.00, vs. 1.71 in Chinese reviews. The mean scores of Item 2 in English reviews were 1.93, vs. 1.59 in Chinese reviews. The scores of other items between English reviews and Chinese reviews were not significantly different (P>0.05).

All the reviews used the quality assessment criteria to evaluate the methodological quality of primary studies. The criteria included Cochrane risk of bias tool,\(^{(49)}\) Physiotherapy Evidence Database (PEDro) scale,\(^{(50)}\) Jadad scale,\(^{(51)}\) Ac-specific standards for Reporting Interventions in Controlled Trials of Acupuncture (STRICTA) criteria,\(^{(52)}\) and the criteria drafted by the authors. The criteria drafted by the authors were used in 10 reviews,\(^{(10,24,26,27,29,31,41,46-48)}\) which usually contained randomization, concealment, blinded, dropout, and ITT analysis. Fourteen reviews used Cochrane risk of bias tool,\(^{(11,20,22,23,25,28,30,33,36,37,42-45)}\) and 2 reviews also used PEDro.\(^{(28,43)}\) Jadad scale was used in 8 reviews,\(^{(9,21,32,33,35,38-40)}\) STRICTA criteria was also used to assess in the review conducted by Wu, et al.\(^{(29)}\)

**Methodological quality of primary studies**

Baseline comparability, and contamination or cointervention were more likely to achieve in primary studies in the primary studies of English and Chinese reviews (Table 2). Allocation concealment, blinded, and reporting adverse events were assessed lowly in the primary studies. Blinding of health care providers was assessed only in one primary study. The proportion of patients blinded, assessors blinded, dropouts and withdrawals, ITT analysis, and reporting adverse events in English reviews were higher than those in Chinese reviews. For example, the proportion of dropouts and withdrawals in English reviews was 56.9%, while that in Chinese reviews was 6.9%.
DISCUSSION

Our results suggested that systematic reviews on acupuncture for stroke published in English journals were better conducted than those in Chinese journals. The English reviews took more measures to avoid the bias in selection of primary studies, and searched more comprehensive and wide database than in Chinese reviews. Many Chinese reviews did not report who reviewed the primary studies or who extracted the data.\(^{39,45,47}\) Many also reported that only one researcher conducted the search and extracted the data and that the researcher was not blinded to identify features of studies and outcomes.\(^{38,46}\) Some Chinese reviews only included the primary studies in the Chinese databases without searching the databases abroad such as PubMed, CINAHL, ISI Web of Knowledge, and Cochrane.\(^{38-40,47,48}\) This resulted in publication bias.

The systematic reviews had other methodological flaws, which were not assessed in the Oxman and Guyatt Scale, such as the choice of the stroke patients, and the choice of control method. The methods of acupuncture treatment and their effectiveness might differ due to different disease stages. The effect sizes could differ if the researchers chose different control groups. (1) Nine reviews focused on all stages of stroke patients.\(^{9,11,23,31,37,38,40,45,47}\) One of them reported that the death rate and dependent rate in acupuncture group were not significantly different compared with the control group.\(^{31}\) While two other reviews recruited primary studies on acute stroke patients and showed that these rates were lower in the acupuncture group.\(^{10,46}\) (2) Some reviews usually enrolled the studies about acupuncture vs. placebo acupuncture, acupuncture vs. sham acupuncture, acupuncture vs. no acupuncture, and the two groups both adding the same conventional therapies (rehabilitation, conventional medication, or CM). Some reviews just pooled all the data from different control groups together. The rationale for
this was not always adequately addressed due to the confounding variables.

The methodological quality assessment of primary studies was an important part in systematic reviews. It was commonly referred to a judgment on the internal validity and risk of bias, and reflected the confidence that the estimates of the effect were correct\(^{(53)}\) Different reviews used different criteria to assess the risk of bias, such as Jadad, PEDro and Cochrane risk of bias tool. Future quality assessments should endeavor to be more comprehensive in assessing the risk of bias in primary studies, which should contain randomization, concealment, blinded, dropout, ITT analysis, baseline comparability, reporting adverse events, contamination or cointervention, and sample size calculation.

The primary studies within the reviews did not achieve some of the criteria for methodological quality. All the primary studies enrolled in the reviews were RCTs, quasi-RCTs, or CCTs, but these studies were not free of important flaws. The description of how to conduct randomization was only reported in about 40% of primary studies. Other primary studies did not report the randomization procedure, or only reported by using the term "randomization".

Allocation concealment was not evaluated in most of the primary studies. Schulz, et al\(^{(54)}\) thought that inadequate concealed treatment allocation tended to exaggerate treatment effects. Blinding of patients and assessors, baseline comparability, dropouts and withdrawals, ITT analysis, and adverse events were also not properly undertaken or described. To improve the quality of primary studies, the reporting should follow the Consolidated Standards for Reporting Trials (CONSORT),\(^{(55)}\) or STRICTA criteria.\(^{(52)}\) The primary studies in English reviews achieved more quality criteria than those in Chinese reviews, because they enrolled more RCTs and quasi-RCTs with high quality.

There were several methodological limitations in our studies. Firstly, only systematic
reviews that assessed methodological quality were considered and any narrative or descriptive reviews were subsequently excluded. Secondly, we focused on the evaluation of biases about internal validity, such as randomization, concealment and blinding. We did not evaluate the adequacy of sample size, validity of the outcomes selected, or defined important clinical differences, although these could be important sources of error in primary studies.

In conclusion, this study identified some methodological flaws in both systematic reviews and primary studies within the reviews on stroke patients receiving acupuncture. The primary studies in these systematic reviews did not meet some of the quality criteria in terms of randomization, allocation concealment, blinding, dropouts and withdrawals, ITT analysis and adverse events. It is therefore necessary to improve the methodological quality and reporting of both systematic reviews and the primary studies about acupuncture for stroke patients.

**CONFLICT OF INTEREST**

The authors have declared that no competing interests exist.

**AUTHOR CONTRIBUTIONS**

Chen XL designed the study, interpreted the results and wrote the manuscript. Mo CW and Lu LY interpreted the results and helped to write the manuscript. Gao RY and Xu Q performed the statistical analyses and contributed to the discussion. Wu MF and Zhou QY participated in study selection and extracted the data. Hu Y and Zhou X modified the manuscript. Li XT participated in the study design and wrote the manuscript.
ACKNOWLEDGEMENT

We thank Dr. Michael Fleming from University of Glasgow for his assistance in editing the manuscript.

ELECTRIONIC SUPPLEMENTARY MATERIAL

Supplementary material (Appendix 1 and 2) is available in the online version of this article at http://dx.doi.org/10.1007/s11655-017-2763-7.

REFERENCES

Figure 1. Flower chart of search strategy for systematic review
Table 1. Mean scores between different types of reviews ($\bar{x} \pm s$)

<table>
<thead>
<tr>
<th>Item</th>
<th>English reviews (n=15)</th>
<th>Chinese reviews (n=17)</th>
<th>t</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2.00±0.00</td>
<td>1.71±0.47</td>
<td>2.42</td>
<td>0.022</td>
</tr>
<tr>
<td>2</td>
<td>1.93±0.26</td>
<td>1.59±0.51</td>
<td>2.37</td>
<td>0.024</td>
</tr>
<tr>
<td>3</td>
<td>2.00±0.00</td>
<td>1.88±0.33</td>
<td>1.37</td>
<td>0.181</td>
</tr>
<tr>
<td>4</td>
<td>1.73±0.40</td>
<td>0.86±0.60</td>
<td>4.40</td>
<td>&lt;0.01</td>
</tr>
<tr>
<td>5</td>
<td>1.93±0.26</td>
<td>1.82±0.39</td>
<td>0.92</td>
<td>0.365</td>
</tr>
<tr>
<td>6</td>
<td>1.87±0.35</td>
<td>1.82±0.39</td>
<td>0.33</td>
<td>0.747</td>
</tr>
<tr>
<td>7</td>
<td>1.60±0.74</td>
<td>1.88±0.33</td>
<td>-1.43</td>
<td>0.164</td>
</tr>
<tr>
<td>8</td>
<td>1.86±0.36</td>
<td>1.76±0.44</td>
<td>0.63</td>
<td>0.533</td>
</tr>
<tr>
<td>9</td>
<td>1.98±0.26</td>
<td>1.94±0.24</td>
<td>-0.09</td>
<td>0.930</td>
</tr>
<tr>
<td>Total score</td>
<td>16.73±1.87</td>
<td>15.29±1.61</td>
<td>2.34</td>
<td>0.026</td>
</tr>
</tbody>
</table>

Table 2. Number and proportion of primary studies achieving the quality criteria

<table>
<thead>
<tr>
<th>Items</th>
<th>English reviews</th>
<th>Chinese reviews</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Eligible article (n)</td>
<td>High score in primary studies (all)</td>
</tr>
<tr>
<td>------------------------------</td>
<td>------------------------</td>
<td>----------------------------------</td>
</tr>
<tr>
<td>Randomization</td>
<td>15</td>
<td>103 (234)</td>
</tr>
<tr>
<td>Allocation concealment</td>
<td>14</td>
<td>38 (225)</td>
</tr>
<tr>
<td>Baseline comparability</td>
<td>7</td>
<td>66 (76)</td>
</tr>
<tr>
<td>Patients blinded</td>
<td>15</td>
<td>46 (234)</td>
</tr>
<tr>
<td>Health care providers blinded</td>
<td>13</td>
<td>1 (192)</td>
</tr>
<tr>
<td>Assessors blinded</td>
<td>12</td>
<td>42 (174)</td>
</tr>
<tr>
<td>Dropouts and withdrawals</td>
<td>10</td>
<td>95 (167)</td>
</tr>
<tr>
<td>ITT analysis</td>
<td>6</td>
<td>32 (69)</td>
</tr>
<tr>
<td>Reporting adverse events</td>
<td>11</td>
<td>34 (160)</td>
</tr>
<tr>
<td>Contamination or cointervention</td>
<td>11</td>
<td>130 (130)</td>
</tr>
</tbody>
</table>