Full Length Research Paper

A systematic review of quality of randomized controlled trials of glycyrrhizin acid to treat hepatitis in China

Ming-Hua Zheng¹*, Shi-Cheng Yu², Ke-Qing Shi¹, Xiao-Zhi Jin¹ and Yong-Ping Chen¹

¹Department of Infection and Liver Diseases, First Affiliated Hospital of Wenzhou Medical College, No. 2 fuxue lane, Wenzhou 325000, Zhejiang, China.
²Department of Cardiology, First Affiliated Hospital of Wenzhou Medical College, Wenzhou 325000, China.

Accepted 10 September, 2009

To assess the quality of Randomized Controlled Trials (RCTs) for variety of preparations made from Glycyrrhizin Acid (GA), an ingredient found in Chinese herb Glycyrrhiza that is widely used in the treatment of hepatitis in clinical practice in China. From electronic searches of PubMed, CNKI, TCMDS, CBMdisc and Chongqing VIP database, we identified 93 reported RCTs between 1991 and 2007 that GA was the subject of the trial. The quality of each trial was assessed using the number of Consolidated Standards for Reporting of Trials (CONSORT) checklist items included, the frequency of allocation concealment and a 5-point quality assessment instrument (Jadad scale). 72.22% (26/36) of the CONSORT checklist items was included in the RCTs. Only 3 (3.22%) of 93 RCTs reported allocation concealment by sealed envelopes. 59 (63.44%) trials described baseline demographic and clinical characteristics of each group, but only 7 (7.52%) reported statistical results of baseline data. Although 77 (82.79%) trials described the number of participants in each group included in each analysis, none of the trials adopted intention-to-treat analysis. Information regarding adverse events was reported in 36 (38.70%) of the RCTs. The quantity of RCTs increased from 1 in 1993 to 15 in 2007, but the average points assessed by CONSORT in each year didn’t elevate. At the same time, the overall quality of the trials was low as assessed by the Jadad scale. Although the quantity of RCTs related to GA mainly against hepatitis increased in recent years, the methodological quality of these reports was low. Therefore, the efficacy and safety of GA being used in clinical practice need to be further confirmed by carefully designed clinical research.

Key words: Traditional Chinese medicine, glycyrrhizin acid, randomized controlled trial, evaluation of quality.

INTRODUCTION

Glycyrrhizin Acid (GA), isolated from the Chinese medical herb Glycyrrhiza, was shown to exhibit multiple activities of anti-inflammatory, anti-virus (Guoyuan and Yajuan, 1991), protecting hepatic cell membrane and improving liver function in vitro and in vivo (Su et al., 1982; Zhao et al., 1983; Wang et al., 1990; Wu et al., 1991). Glycyrrhizin has been reported to suppress hepatic inflammation with an effect to improve the elevated ALT levels and histologic findings of the liver. Diammonium Glycyrrhizinate (DG) injection (Chia-tai Tianqing Pharmaceutical Co., Ltd) with the main ingredient of GA has been widely used in China since 1995 after the approval of State Administration of Pharmacy and Ministry of Health People’s Republic of China. Thereafter, a variety of medications made from GA have been manufactured by other Chinese pharmaceutical companies and marketed for treatment of liver diseases, such as Compound Glycyrrhizin Injection, Compound Glycyrrhizin Tablets, Diammonium Glycyrrhizinate Capsules, Ammonium Glycyrrhizinate Injection.

Since these medications were introduced in the middle of 1990s, a large number of related clinical trials were conducted and the reports of trials including some Randomized Controlled Trials (RCTs) were published in Chinese academic journals, such as Chinese journal of infectious diseases (Zhang et al., 1996). Although preliminary research indicates that ingredient of GA could improve the clinical symptoms and outcomes of patients with hepatitis and liver function abnormalities, the effectiveness

*Corresponding author. E-mail: ypchen77@hotmail.com. Tel: (86) 577-88078232. Fax: (86) 577-88078262.
of GA in the clinical setting could not be confirmed due to the lack of the convincing evidence. So it is premature to make a conclusion that GA can play a role in protecting the hepatic cell membrane, anti-inflammatory, anti-virus, improving liver function, or regulating immunity.

It is well known that the most rigorous method for evaluation of any therapeutic intervention is RCT. Meanwhile, the quality of RCT is crucial for providing reliable groundwork in medical decision making based on the principle of Evidence-Based Medicine (EBM). Moher et al. (2002) assessed the quality of 251 RCTs in pedia-tric complementary and alternative medicine in 2002 and found there was an increase over time in number of the items of CONSORT checklist included in the reports. Meanwhile, 83.1% of RCTs reported unclear allocation concealment. As for RCTs of Traditional Chinese Medicine (TCM), Chang et al. (2006) reviewed 365 articles published in 4 Chinese journals of TCM from 2000 to 2005 submitted by four colleges of TCM separately. They conclude that although methodological quality has been improved over the years, many problems remain unsolved. For example, most of the trials did not describe the randomization, only few trials used the allocation concealment and blinded method, no description of comparability of baseline data, lack of inclusion and exclusion criteria, or short of the estimation of sample size.

We have assessed the quality of RCTs of Glycyrrhizin Acid effect to against hepatitis by comparing the reported RCTs with the gold standard (CONSORT guideline and Jadad scale) in order to evaluate the RCTs of GA. The assessment enables readers to judge to what extent that the reported RCTs are internally valid and free of bias. At the same time, the quality of clinical trials of TCM will be improved over the years, many problems remain unsolved. For example, most of the trials did not describe the randomization, only few trials used the allocation concealment and blinded method, no description of comparability of baseline data, lack of inclusion and exclusion criteria, or short of the estimation of sample size.

RESULTS

Database searching identified 1060 reports related to GA. A total of 338 clinical trials were screened, from which 93 RCTs were identified for further study (Figure 1) (Guo et al., 1994; He, 1994; Zhang et al., 1994; He et al., 1995; Zhang and Zhang, 1995; Zhang and Zhu, 1995; Huang et al., 1996; Li et al., 1997; Li and Ha, 1998; Wu, 1998; Yuan, 1998; Zhang and Sun, 1998; Cheng and Yang, 1999; Cheng et al., 1999; Han et al., 1999; Zhu, 1999; Jin and Gu, 2000; Liang et al., 2001; Ling et al., 2001; Ma and Zhou, 2001; Wen, 2001; Guo et al., 2002; Li, 2002; Li, 2003; Liu et al., 2003; Lu, 2003; Luo, 2003; Qiang and Jing, 2003; Wang et al., 2003; Zeng and Xie, 2003; Han and Wang, 2004; Huang et al., 2004; Li et al., 2004; Nong et al., 2004; Wan, 2004; Yan, 2004; Cheng et al., 2005; Dai et al., 2005; Hu and Li, 2005; Huang and Liu, 2005; Ling, 2005; Lu and Wan, 2005; Peng, 2005; Qiu et al., 2005; Sheng, 2005; Wang et al., 2005; Yang et al., 2005; Yuan et al., 2005; Zeng, 2005; Zhang and Tang, 2005; Cheng, 2006; Cheng et al., 2006; Guo, 2006; Huang, 2006; Ling and Jiang, 2006; Qiu et al., 2006; Sun et al., 2006; Sun and Dun, 2006; Wen and Guo, 2006; Yu et al., 2006; Zhang and Li, 2006; Zhang and Zhang, 2006; Zhong, 2006; Gao, 2007; Gong, 2007; Huang et al., 2007; Li, 2007; Ling, 2007; Luo et al., 2007; Tan et al., 2007; Xin et al., 2007; Yang, 2007; Zhou et al., 2007; Chen et al., 2008; Li et al., 2008; Wang et al., 2008). Descriptive information was collected and then assessment of the 93 RCTs followed. All of RCTs were published between 1993 to 2007. The sum of RCTs in each year ranged from 1 to 18. During the years from 1993 to 2003, the sum of the reports was kept steady at 3.7 per year, but it rapidly increased to 14 per year between 2004 and 2007. 72.22% (26/36) of the CONSORT checklist items was included in the reports of GA RCTs (Table 1). The items which should be included but not described in the reports were the sample size, the explanation of any interim analyses and stopping rules, the randomization (allocation concealment and implementation), the blinding

MATERIALS AND METHODS

Five electronic databases, including PubMed, CNKI (China National Knowledge Infrastructure), TCMDS (Traditional Chinese Medicine Database System), CBMdisc (China Biomedicine Database disc) and Chongqing VIP database (a full text issues database of China), were thoroughly searched by using the keywords Glycyrrhizin Acid (MeSH Subjects) or Diammonium Glycyrrhizinate, Compound Glycyrrhizin Injection for human clinical trials. The time frame of the reported RCTs was limited between January, 1991 and December, 2007. Conference proceedings in China were also manually searched. The search results were incorporated to a reference database and further screened. In addition, a non-electronic search was conducted for those studies cited in relevant references, but not included in the five databases. All of the RCTs listed in this study clearly stated that the use of GA given alone or concurrently with other medication as treatment groups (Figure 1).

A comprehensive quality assessment of each report was completed using three steps. First, the revised CONSORT statement checklist was modified so that multiple items were listed separately, which resulted in 36 items (Altman et al., 2001; Moher et al., 2001). Each item was assigned a ‘yes’ or ‘no’ response depending on whether or not the author had reported it. Second, the reporting of allocation concealment was assessed as ‘adequate’, ‘inadequate’, or ‘unclear’. Third, the Jadad scale, which contains two questions each on randomization and blinding and one question on reporting of dropouts and withdrawals, was used to assess quality. Each question contains a ‘yes’ or ‘no’ response option. In total, five points can be awarded with higher points indicating superior quality. After that, the sum of RCTs published in each year was tabulated and the average points assessed by CONSORT statement checklist were calculated (Table 1).

Two reviewers completed the assessments independently. We conducted formal training prior to evaluating the RCTs for each of the three steps. Discrepancies were resolved by consensus between the two raters. We compared the number of the CONSORT checklist criteria included in each report. We also assessed the percentage of studies that reported unclear allocation concealment and the specific item and overall quality score derived from the Jadad scale.
63.44% (59/93) of reports described baseline demographic and clinical characteristics of each group, the rest simply stated in the article that there was no significant difference between two groups; whereas only 2 (2.15%) studies reported statistical analysis of baseline data.

ding, the description of protocol deviations from study as planned, the ancillary analyses (subgroup analyses, adjusted analysis), the interpretation sources of potential bias or imprecision and the dangers associated with multiplicity of analyses and outcomes.
Table 1. CONSORT checklist criteria included in 93 reports of RCTs about Glycyrrhizin acid.

<table>
<thead>
<tr>
<th>Total number of items</th>
<th>Sub-items reported</th>
<th>Number of reported RCTs [n(%)]</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Abstract</strong></td>
<td>1. How participants were allocated to interventions</td>
<td>43(46.23%)</td>
</tr>
<tr>
<td><strong>Introduction</strong></td>
<td>2. Scientific background and explanation of rationale Methods</td>
<td>57(61.29%)</td>
</tr>
<tr>
<td></td>
<td>3.1. Eligibility criteria for participants</td>
<td>91(97.84%)</td>
</tr>
<tr>
<td></td>
<td>3.2. The settings and locations where the data were collected</td>
<td>80(86.02%)</td>
</tr>
<tr>
<td></td>
<td>4.1. Precise details of the interventions intended for each group</td>
<td>91(97.84%)</td>
</tr>
<tr>
<td></td>
<td>4.2. How and when they were actually administered</td>
<td>93(100%)</td>
</tr>
<tr>
<td></td>
<td>5. Specific objectives and hypotheses</td>
<td>93(100%)</td>
</tr>
<tr>
<td></td>
<td>6 Clearly defined primary and secondary outcome measures and, when applicable, any methods used to enhance the quality of measurements</td>
<td>93(100%)</td>
</tr>
<tr>
<td></td>
<td>7.1. How sample size was determined</td>
<td>0(0%)</td>
</tr>
<tr>
<td></td>
<td>7.2. Explanation of any interim analyses and stopping rules</td>
<td>0(0%)</td>
</tr>
<tr>
<td></td>
<td>8. Method used to generate the random allocation sequence, including details of any restriction</td>
<td>2(2.15%)</td>
</tr>
<tr>
<td></td>
<td>9. Method used to implement the random allocation sequence (e.g., numbered containers or central telephone), clarifying whether the sequence was concealed until interventions were assigned</td>
<td>2(2.15%)</td>
</tr>
<tr>
<td></td>
<td>10.1. Who generated the allocation sequence</td>
<td>0(0%)</td>
</tr>
<tr>
<td></td>
<td>10.2. Who enrolled participants and who assigned</td>
<td>0(0%)</td>
</tr>
<tr>
<td></td>
<td>11. Whether or not participants, those administering the interventions and those assessing the outcomes, were blinded to group assignment. If done, how the success of blinding was evaluated</td>
<td>0(0%)</td>
</tr>
<tr>
<td></td>
<td>2.1. Statistical methods used to compare groups for primary outcomes)</td>
<td>53(56.98%)</td>
</tr>
<tr>
<td></td>
<td>12.2. Methods for additional analyses, such as subgroup analyses and adjusted analyses</td>
<td>8(8.60%)</td>
</tr>
<tr>
<td><strong>Results</strong></td>
<td>13.1. Flow of participants through each stage</td>
<td>93(100%)</td>
</tr>
<tr>
<td></td>
<td>13.2. Receiving intended treatment, completing the study protocol</td>
<td>87(93.54%)</td>
</tr>
<tr>
<td></td>
<td>13.3. Analyzed for the primary outcome</td>
<td>91(97.84%)</td>
</tr>
<tr>
<td></td>
<td>13.4. Describe protocol deviations from study as planned, together with reasons</td>
<td>0(0%)</td>
</tr>
<tr>
<td></td>
<td>14.1. Dates defining the periods of recruitment</td>
<td>66(70.96%)</td>
</tr>
<tr>
<td></td>
<td>14.2. Dates of follow-up</td>
<td>70(75.26%)</td>
</tr>
<tr>
<td></td>
<td>15. Baseline demographic and clinical characteristics of each group</td>
<td>9(63.44%)</td>
</tr>
<tr>
<td></td>
<td>16.1. Number of participants (denominator) in each group included in each analysis</td>
<td>89(95.69%)</td>
</tr>
<tr>
<td></td>
<td>16.2. Whether the analysis was by &quot;intention-to-treat&quot;. State the results in absolute numbers when feasible</td>
<td>77(82.79%)</td>
</tr>
<tr>
<td></td>
<td>17.1. For each primary and secondary outcome, a summary of results for each group</td>
<td>88(94.62%)</td>
</tr>
<tr>
<td></td>
<td>17.2. The estimated effect size</td>
<td>93(100%)</td>
</tr>
<tr>
<td></td>
<td>18.1. Subgroup analyses</td>
<td>0(0%)</td>
</tr>
<tr>
<td></td>
<td>18.2 Adjusted analysis</td>
<td>0(0%)</td>
</tr>
<tr>
<td></td>
<td>19. All important adverse events or side effects in each intervention group</td>
<td>36(38.70%)</td>
</tr>
<tr>
<td><strong>Discussion</strong></td>
<td>20.1. Interpretation of the results, taking into account study hypotheses</td>
<td>80(86.02%)</td>
</tr>
<tr>
<td></td>
<td>20.2. Sources of potential bias or imprecision</td>
<td>0(0%)</td>
</tr>
<tr>
<td></td>
<td>20.3. The dangers associated with multiplicity of analyses and outcomes Universality</td>
<td>0(0%)</td>
</tr>
<tr>
<td></td>
<td>21. Generalization (external validity) of the trial findings summarization</td>
<td>6(70.96%)</td>
</tr>
<tr>
<td></td>
<td>22. General interpretation of the results in the context of current evidence</td>
<td>24(25.80%)</td>
</tr>
</tbody>
</table>

Note: Boldface indicating the ratio has not reached 50%.
Table 2. Quality of reports of 93 RCTs about Glycyrrhizin acid using the Jadad assessment scale and the adequacy of allocation concealment.

<table>
<thead>
<tr>
<th>Total number of items</th>
<th>Score number of reported RCTs [n (%)]</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>2</td>
</tr>
<tr>
<td>Randomization</td>
<td>5 (5.37%)</td>
</tr>
<tr>
<td>Double-blinding</td>
<td></td>
</tr>
<tr>
<td>Withdrawals/dropouts</td>
<td>17 (18.27%)</td>
</tr>
<tr>
<td>Allocation concealment</td>
<td></td>
</tr>
</tbody>
</table>

46.23% (43/93) of reports stated in the abstract that the participants were allocated randomly for the trials, the rest mentioned in other sections of the article. But only 2 (2.15%) described the method used to generate the random allocation sequence and the method used to implement the random allocation sequence. Neither of the reports detailed the allocation sequence, the assignment of enrolled participants, nor whether the participants who were dispensing the medication and assessing the outcomes were blinded to group assignments. Information regarding adverse events was reported in 38.70% (36/93) of the RCTs. The reports achieved their maximum possible total scores as assessed using the Jadad scale (Table 2). 5 (5.37%) of 93 RCTs achieved a score of 2.80 (86.02%) were awarded a score of 1 (only referring to “randomization”) and 8 (8.60%) were given a score of 0 because of using quasi-randomization. Only 17 (18.27%) reported withdrawals or dropouts. None of reports described double blinding and allocation concealment adequately. The sum of RCTs in each year from 1993 to 2007 was presented graphically in Figure 2. It shows that the number of the RCTs conducted before 2003 was kept low, it increased significantly from 10 to 18 between 2002 and 2005. The average point of 93 RCTs assessed by CONSORT is 18.17 with total maximum of 36. The average point in each year fluctuated between 9 and 19.06 (Figure 3). No noticeable improvement of the quality of GA RCTs was shown in recent years.
DISCUSSION

Generally speaking, there is a couple of stages in adopting an herbal medication in the practice of TCM. First, research activities focused on the screening, extracting and testing relevant active substances; then followed by a series of RCTs to demonstrate its clinical efficacy and possible toxicity in human. During the years, there is an increasing amount of systemic reviews or meta-analyses aimed at such RCTs were published (Li et al., 2005). Many have indicated that the quality of such clinical trials related to TCM was low (Yu et al., 1994; Mao et al., 2007).

With the development of EBM in China, researchers of TCM began to realize the importance of confirming the efficacy and safety of TCM through a series of standardized and verifiable clinical trials. In 2000, the WHO Traditional Medicine Program convened experts who represented various disciplines to adapt the guidelines for the clinical evaluation of traditional medicine. After the State Food and Drug Administration (SFDA) of China developed and revised Good Clinical Practice (GCP) regulations in 1998, the clinical trial is required for manufacturing, marketing and adopting of TCM. As an agent of protecting liver cell membrane, the efficacy and safety of GA are being evaluated by a large number of clinical trials. However, the efficacy of GA has not been adequately and systematically evaluated due to the paucity of information on the quality of such clinical trials. In this study, two-thirds (26 out of 36) of the CONSORT checklist items was included in the RCTs of GA. Unfortunately, the results from Jadad scale and allocation concealment suggest that the validity of GA RCTs is questionable. All of them were scored poorly when examined with Jadad Quality Scale (score < 3). This is particularly true when our attention focuses on the aspects of how randomization was reported. Only two reports documented how the random numbers were generated and none had clear allocation concealment. It is worthy to mention that 18.27% of the reports were published in the key Chinese academic journal “Chinese Journal of Integrated Traditional and Western Medicine” and “Chinese Journal of Evidence-Based Medicine”.

In a report on the assessment of 7422 RCTs published in the Chinese Journal of Evidence-Based Medicine in 2007 (Mao et al., 2007), the average number of CONSORT items included was 39.4%. The mean Jadad score was 1.03 ± 0.61 in all the trials with 1 RCT with 5
analyses that quoted such RCTs published in recent years are not reliable. We reviewed one meta-analysis on the efficacy of Diammonium Glycyrrhizinate Injection in the treatment of chronic hepatitis published in the Chinese Journal of Infectious Diseases (Chinese core journal) (Qin et al., 2005), and found that 10 of 24 RCTs cited in the meta-analysis were actually not randomized. There was apparent difference in the sample size calculation (11%), randomization sequence (79%), allocation concealment (03%), implementation of the random allocation sequence (00%), analysis of intention to treat (00%), were incompletely reported. The CONSORT result reported in this study is higher than the results observed elsewhere, whereas the score of Jadad is lower. There is evidence to suggest that journals using the CONSORT statement, compared to those not doing so, would have higher quality reports of RCTs (Jadad et al., 1996).

Evidently, examining the quality of published reports is “after the fact” when the trials were already completed. Therefore, the way to improve the quality of RCTs, especially in the field of TCM, is not only to endorse the CONSORT statement, but also to enforce the audit of clinical trials.

The low quality of RCTs reflected ill-designed and poorly executed trials. Consequently, it also indicated that the outcomes of many systemic reviews or meta-analyses that quoted such RCTs published in recent years are not reliable. We reviewed one meta-analysis on the efficacy of Diammonium Glycyrrhizinate Injection in the treatment of chronic hepatitis published in the Chinese Journal of Infectious Diseases (Chinese core journal) (Qin et al., 2005), and found that 10 of 24 RCTs cited in the meta-analysis were actually not randomized. There was apparent difference in the sample size between intervention and control group. In some “RCTs”, the cases recruited in the control group are only two third of the intervention group. Therefore, the conclusions of such meta-analysis and alike are needed to be interpreted cautiously.

We noticed that there were few published reports of GA RCTs prior to 2002 that was contrasted by a rapid increase in the number of the reports between 2002 and 2005. This might also reflect that the research activities on GA were less organized at the national level, which resulted in substandard trials and wasted tremendous resources.

Meanwhile, with the popularization of CONSORT checklist in China, many researchers finished their reports by referring to the items provided by CONSORT checklist. Therefore, “randomization” appeared in more and more reports, but unfortunately methods on how to conduct “randomization” were not described clearly in most of RCTs. Successful randomization should describe the method used to generate the random allocation sequence, including details of any restriction and the method used to implement the random allocation sequence (e.g. numbered containers or central telephone), should clarify whether the sequence was concealed un-

ACKNOWLEDGEMENTS

The authors thank Yong Zhang PhD, China Pharmaceutical University, for his kindly help in Figures.

REFERENCES

Cheng Y, Ye M, Lei, Yuan H (1999). Observation of Effect of Diam-
Sheng X (2005). The clinical effects of compound glycyrrhizic acid on 86