Identification of the optimal cognitive drugs among Alzheimer’s disease: a Bayesian meta-analytic review

Jinghong Liang
Jiayu Li
Ruixia Jia
Yingquan Wang
Rongkun Wu
Lei Hang
Yong Xu

Department of Child Health, Jiangsu Key Laboratory of Preventive and Translational Medicine for Geriatric Diseases, School of Public Health, Soochow University, Suzhou, China

Purpose: The increasing prevalence of Alzheimer’s disease (AD) demands more effective drugs, which are still unclear. The aim of this study is to compare the effectiveness of six drugs, such as donepezil, rivastigmine, galantamine, memantine, huperzine-A, and tacrine, in senior AD patients and identify the most effective one to improve patients’ cognitive function.

Methods: A system of search strategies was used to identify relevant studies including randomized controlled trials and clinical controlled trials evaluating the efficacy of six drugs in patients with AD. We updated relevant studies that were published before March 2018 as full-text articles. Using Bayesian network meta-analysis (NMA), we ranked cognitive ability objectively based on Mini–Mental State Examination (MMSE). Pairwise and NMAs were sequentially performed for the efficacy of drugs compared to each drug or control group through the trials included.

Results: Among the 35 trials included, no obvious heterogeneity ($I^2=0.0\%$, $P=0.583$) was revealed according to the pooled data for cognition in NMA and the mean difference (MD) of memantine (MD=1.7, 95% CI: 0.73, 2.8) showed that the memantine was significantly efficacious in the treatment group in terms of MMSE. Followed by galantamine, huperzine-A, rivastigmine, tacrine, and donepezil.

Conclusion: As the first NMA comparing the major drugs in market for AD, our study suggests that memantine might have a more significant benefit on cognition than other five drugs available.

Keywords: Alzheimer’s disease, cognitive drugs, Bayesian network meta-analysis

Introduction

Alzheimer’s disease (AD) is a neurologically degenerative disease that would obtain progressive development but is concealed in the early stage. The disease is clinically characterized by memory impairment, aphasia, disability, visual impairment, executive dysfunction and personality, and behavioral changes. Patients suffering from AD have a poor self-reliance ability, and their disease imposes a heavy burden on their families, caregivers, health care system, and even society.

The world’s aging population currently comprises nearly 900 million people, most of whom are living in relatively poor countries. In the mean time, more than 47 million people around the world are suffering from AD and, by 2050, the number is estimated to reach 131.5 million. The estimated worldwide cost of dementia totals US $818 billion. AD organizations should focus on how we can slow down the progressive cognitive dysfunctions, maintain functional status, improve the quality of life,
minimize adverse events (AEs), modulate caregiver stress, and relieve the economic burden on the family.\(^6\)

At present, it is generally acknowledged in the world that the pathological cause of AD is the absence of neurotransmitter acetylcholine (Ach) in the brain of AD patients, which may result in a decreased cognitive function and memory loss.\(^7\)\(^-\)\(^9\) The current pharmacological therapy of the disease is mainly achieved by increasing the Ach level in patients through inhibiting cholinesterase (CHE). The acetylcholinesterase inhibitor (AChEI) takes extensive effect in the treatment of AD patients through inhibiting acetylcholinesterase’s activity, delaying the rate of hydrolysis of Ach, and increasing the level of Ach’s synaptic gap.\(^9\)-\(^11\) As an early cholinesterase inhibitor (ChEI) to treat patients with mild-to-moderate AD, tacrine has basically been replaced for now by three first-tier drugs,\(^12\) ChEIs donepezil, rivastigmine, galantamine, and memantine, an uncompetitive antagonist of N-methyl-D-aspartate (NMDA) glutamate receptors. As a potent, reversible, and selective inhibitor of Ach esterase (AChE) derived from the Chinese herb *Huperzia serrata*, huperzine-A was identified by scientists in China in the 1980s and its improvement in cognitive ability on AD has been approved by some studies.\(^13\)-\(^15\) The abovementioned drugs were supported by comprehensive clinical data that demonstrate their safety and symptomatic efficacy in treating cognitive dysfunction or other major domains of AD.\(^16\)-\(^24\)

Network meta-analysis (NMA) can be helpful to summarize the performances and ranking of two or more drugs. Given the doubts on the methodological issues of sample size, relevant outcomes, and heterogeneity sources, NMA, however, may obtain more accurate and reliable results than traditional meta-analysis.\(^25\) It also enhances the relative effectiveness of inference with each drug through direct and indirect data.\(^26\)-\(^27\) In this study, we used this novel differential meta-analysis method to estimate the comparative efficacy of cognitive drugs by comparisons or control group (CG) for AD. Our aim is to provide relatively effective, safe, and comparative evidence to identify the optimal drug for AD patients.

**Methods**

**Search strategy**

A systematic search for highly relevant publications was executed by using strategy electronic databases (the Cochrane Central Register of Controlled Trials, EMBASE, PubMed, as well as four Chinese medical databases, such as China National Knowledge Infrastructure database, Chinese Biomedical Literature database, Wanfang database, and Web of Science). Relevant studies before March 1, 2018, were exhaustively searched. The search strategy was based on medical subject heading (Mesh) terms or Emtree terms combined with Boolean logical operators. The search strategy was built by crossing key search terms using the following search phrases regardless of publication time: “Alzheimer’s disease”, “cholinesterase inhibitors”, “donepezil”, “rivastigmine”, “galantamine”, “memantine”, “huperzine-A”, “tacrine”, and “randomized controlled trial”.

The EndNote X7 literature management software (Thompson ISI Research Soft, Philadelphia, PA, USA) was used to screen and manage search records. No language restrictions were implemented, and non-English studies were translated to examine their full text. The bibliographies of all selected articles were additionally screened, and manual search was carried out to ensure a complete identification of all eligible studies such as studies in reports and reference lists of identified studies from previous systematic reviews. The proceedings from major international conferences in AD and medicine were also searched. A citation analysis was performed to identify newer studies that had cited former ones. Preparations were made in NMA according to the PRISMA guidelines.\(^28\)

**Selection criteria**

Three authors, respectively, reviewed full manuscripts of eligible studies and extracted correlated information including study characteristics and measured outcomes.

Two authors independently screened the title and abstract according to the predefined criteria for inclusion initially. Identical studies were excluded in the literature management software at the same time. Additionally, one author retrieved and perused the full texts of preliminary relevant studies identified in the preceding step for compliance with eligibility criteria and data extraction. We excluded studies published only as abstract without additional available data.

We used population, interventions, comparators, study designs (PICOS) criteria. The PICOS criteria are as follows.

**Population**

Senior patients diagnosed with AD by using any scientific criteria were included. (Individual research had different measurement methods, but it does not affect the normal assessment of AD.)

**Intervention**

Studies with AD patients treated by any new or former CHE inhibitors, such as donepezil, galantamine, rivastigmine,
huperzine-A, tacrine, and another drug, memantine, were included.

Comparison
Cognitive drugs, CG alone, or in any combination (no treatment, placebo, best supportive care) were compared.

Outcome
The primary outcome was evaluated by the validated assessment of Mini–Mental State Examination (MMSE), as the efficacy of cognitive drugs. This endpoint was used by an adequate amount of included trials; thus, our NMA could be conducted.

Study design
We restricted to randomized controlled trials (RCTs) as the golden standard for examining cognitive drugs. Other types of trials were not allowed in our NMA.

Risk of bias (ROB) and quality appraisal
We appraised the ROB by the Cochrane risk of bias tool. Two authors independently appraised the ROB using the items reported in the Cochrane risk of bias tool, which includes seven items regarding ROB performed in Review Manager (Version 5.3). The ROB include the following seven items: random sequence generation (selection bias), allocation concealment (selection bias), blinding of participants and personnel (performance bias), blinding of outcome assessment (detection bias), incomplete outcome data (attrition bias), selective reporting (reporting bias), and other bias. Based on the above domains, the included RCTs were classified as being at low risk, high risk, or unclear risk.

As an essential scatter plot to estimate the effect of cognitive drugs, the comparison-adjusted funnel plot was drawn to estimate the magnitude of funnel plot asymmetry for MMSE. The abscissa axis and vertical axis of the funnel plot, respectively, represented the standard mean difference (SMD) and the standard error of SMD. Publication bias identification depends on whether the scatter of the study is symmetrical in the inverted funnel. This plot stands up to the examination of heterogeneity and different types of bias, such as selective reporting, publication, and funding biases.

Assessment of inconsistency
We used “node-splitting” approach to examine the inconsistency in the network by comparing the differences between direct and indirect evidence and the side-splitting approach to detect comparisons where direct estimations disagree with indirect evidence from entire network. P-rank scores were generated to determine the probability of inconsistency.

Data abstraction and outcome measure
We conducted a rigorous process to extract relevant data. We first analyzed the summary data and demographic characteristics of each study. Based on the predefined strategy and criteria, two authors reviewed the included studies to collect descriptive data from eligible studies with information as year of publication, study origin (country), people’s characteristics, diagnosis, specific drug therapy and its CG, and duration of therapy. The primary outcome was MMSE, which evaluated the cognitive domain. Scores and cognitive ability were proportional. The mean and SD of the change from baseline were extracted. To ensure data precision, check was necessary before the further analysis could be realized for other authors.

Statistical analyses
We first performed a conventional pairwise meta-analysis by synthesizing all direct evidence. Publication bias and small-study effects were examined visually using comparison-adjusted funnel plots under the random effect model. The bias of magnitude heterogeneity variance parameter was quantified using the I² statistics, which values over 50% indicating substantial heterogeneity. We produced network plot of each drug as a visual representation that offers the evidence base and a concise description of their characteristics.

Statistically, we evaluated a common between-study variance parameter because all the treatments were pharmacological. The arm-specific mean differences (MDs) from baseline were calculated by the random effect models with inverse variance method that was used for continuous outcomes. The MDs were concluded by calculating the difference value between experiment group and CG. The following 95% CI was calculated as a measure of an estimate uncertainty. The mean and SD of the MMSE at the last observation of the studies were extracted, and the standardized mean change (Hedges’ adjusted g) was computed from baseline to follow-up as the gist of difference between the treatment groups.

Transitivity assumption was evaluated by exploring potential treatment effects’ modifiers through the available direct comparisons.

A Bayesian hierarchical random effects meta-analysis was performed for each treatment comparison, as we anticipate clinical and methodological between-study heterogeneity. Bayesian approaches can compare multiple pharmacology
at one time by enabling simultaneous comparisons between multiple drugs by the way of forming a connected network while integrating direct and indirect estimates and using the methodology of multivariate meta-analysis. Our NMA was performed for cognitive outcome measure with Bayesian framework using the Markov chain Monte Carlo method with a random effect mode. We initially set up three Markov chains to simulate an accurate estimate for statistic model and the number of iterations was 25,000, of which the first 5,000 were annealed to eliminate the influence of the initial value.

The probability of which was the most efficacious drug derived from the proportion of the best ranking in all simulation operations. Probability values were summarized and reported as surface under the cumulative ranking (SUCRA), which was generated by summarizing the cumulative ranking probability based on all of the abovementioned simulation operations. The SUCRA line was used to provide hierarchy probabilities in which the outcome 1 was possibly the best while 0 represented the worst. The above analyses were performed using the Gemtc package in R (X64 Version 3.32) and STATA Version 14.0 (StataCorp LP, College Station, TX, USA).

Results
Baseline characteristics of included studies and ROB quality assessment
Figure 1 summarizes the selection process. A total of 35 studies were identified based on our initial search item. Eight studies were retrieved through manual search from varying online databases between 2000 and 2018. One hundred eight articles were removed for similarity. 2,734 articles were assessed for relevance by screening the title and abstract. At last, 35 articles were included in our Bayesian NMA through the above rigorous eligible criteria. All participating authors agreed on the methodology for selection and assessment.

Table 1 presents the baseline data of demographic characteristics from 35 trials included. A total of 9,820 AD patients underwent Population, Intervention, Outcomes strategies. The ChEI group holds an overwhelming majority of sample size including 2,585 of donepezil, 2,397 of rivastigmine, and 2,418 of galantamine. A relatively small sample size for memantine, huperzine-A, and tacrine was 1,557, 258, and 605, respectively. The trials were published between 1992 and 2015, and majority of them were from the USA (n=4,684, 47.6%). The mean age of all samples ranged from 63.8 to 84.9 years. Fifty-nine percentage (n=5,816) of the participants were female, and the average scores of MMSE for all samples ranged from 7.25 to 23.1 at baseline. There are totally 27 two-arm trials and two three-arm trials.

We assessed the qualities of the included studies based on the Cochrane Collaboration tool for assessing ROB. Only one RCT was classified to be high risk in “allocation concealment” aspects. And three trials were judged as moderate/unclear ROB studies. A total of 16 RCTs reported adequate random sequence generation, 27 RCTs reported blind method, and seven studies described their approaches of allocation concealment. 45.7% of the included studies indicated a low ROB for random sequence generation, 77.1% of the included studies indicated a low ROB for blinding, 54.3% of the included studies indicated a low ROB for incomplete data outcome, and 0.03% of the included studies indicated a low ROB for selective reporting and other bias. Figure S1 shows the degree of ROB for all studies included, and Figure S2 is formed by summarizing the total ROB of Figure S1 based on the seven items.

NMA results
Primary outcome
The network plot of cognitive outcome is shown in Figure 2. Node sizes and edge widths in network plot indicated the number of drugs compared and the available direct comparisons between pairs of drugs, respectively. Each cognitive drug had at least one controlled trial. Donepezil had five closed loops between galantamine, rivastigmine, memantine, huperzine-A, and placebo while tacrine lacked a closed loop with others, thus only representing the existence of indirect evidence.

Relevant data about cognition were recorded from 35 studies included, of which 13 studies were for donepezil, nine studies were for rivastigmine, five studies were for galantamine, eight studies were for memantine, six studies were for huperzine-A, and four studies were for tacrine. The absence of obvious heterogeneity ($I^2=0.00\%$, $P=0.563$) was shown by preliminary meta-analysis (Table 2). The funnel plot showed a symmetric distribution (Figure S3), and the effect size of all the included studies was spread in the inverted funnel indicating no publication bias. At last, 35 eligible studies were finally included and reflected in the network relationship plot (Figure 2). Our analysis revealed that memantine, galantamine, huperzine-A, and rivastigmine had significantly greater improvement than CG (Table 2). The SUCRA plot (Figure 3) presented the probability of rank for each cognitive drug, which indicated that memantine was the highest in probability among all the six cognitive drugs (SUCRA=70.5%), followed by galantamine (SUCRA=65.4%), huperzine-A.
Optimal cognitive drugs among Alzheimer’s disease

(SUCRA=53.3%), rivastigmine (SUCRA=52.1%), and tacrine (SUCRA=44.9%). In contrast, donepezil seemed to have the lowest probability (SUCRA=43.4%). No significant inconsistency was detected among various treatments whose $P$-value was lower than 0.05 (Figure S4A and B).

**Discussion**
Our NMA was conducted on the RCTs of the cognitive drugs for AD patients in which indirect evidence was used to appraise and compare the relative effectiveness of cognitive drugs across trials simultaneously. Our study attempted...
<table>
<thead>
<tr>
<th>Intervention (mean dose/day)</th>
<th>Study</th>
<th>Age (years), mean±SD</th>
<th>Gender (female, %)</th>
<th>Baseline MMSE, mean±SD</th>
<th>Trial duration (weeks)</th>
<th>Country</th>
</tr>
</thead>
<tbody>
<tr>
<td>D (10 mg/day)</td>
<td>SL Rogers, 1998</td>
<td>74.60±0.60</td>
<td>60/97</td>
<td>19.17±0.37</td>
<td>24</td>
<td>USA</td>
</tr>
<tr>
<td></td>
<td>Peter Johansen, 2006</td>
<td>74.10±7.60</td>
<td>40/59</td>
<td>18.80±4.80</td>
<td>24</td>
<td>USA</td>
</tr>
<tr>
<td></td>
<td>SE Black, 2010</td>
<td>78.00±8.00</td>
<td>48/128</td>
<td>7.50±0.25</td>
<td>20</td>
<td>Canada</td>
</tr>
<tr>
<td></td>
<td>Fred Andersen, 2012</td>
<td>80.80±6.80</td>
<td>28/62</td>
<td>23.20±4.20</td>
<td>52</td>
<td>Norway</td>
</tr>
<tr>
<td></td>
<td>Hong-Mei, Tu, 2015</td>
<td>75.50±3.50</td>
<td>36/24</td>
<td>18.20±4.30</td>
<td>12</td>
<td>China</td>
</tr>
<tr>
<td></td>
<td>Qun-Sheng, Qin, 2015</td>
<td>72.40±1.10</td>
<td>21/29</td>
<td>17.60±1.80</td>
<td>48</td>
<td>China</td>
</tr>
<tr>
<td>R (2.4 mg/day)</td>
<td>Yahya Karamana, 2004</td>
<td>74.11±0.87</td>
<td>11/13</td>
<td>11.40±0.20</td>
<td>52</td>
<td>Turkey</td>
</tr>
<tr>
<td></td>
<td>Yu Nakamura, 2011</td>
<td>75.10±6.90</td>
<td>92/195</td>
<td>16.40±3.09</td>
<td>24</td>
<td>Japan</td>
</tr>
<tr>
<td></td>
<td>Bengt Winblad, 2007</td>
<td>74.20±7.70</td>
<td>103/200</td>
<td>16.60±2.90</td>
<td>24</td>
<td>USA</td>
</tr>
<tr>
<td></td>
<td>V Kumar, R, 1999</td>
<td>74.30±0.37</td>
<td>38/50</td>
<td>20.20±0.22</td>
<td>26</td>
<td>USA</td>
</tr>
<tr>
<td>G (18 mg/day)</td>
<td>Klaus Hager, 2014</td>
<td>73.00±8.90</td>
<td>353/671</td>
<td>19.00±4.12</td>
<td>108</td>
<td>China</td>
</tr>
<tr>
<td></td>
<td>Rong-Ping, Chen, 2014</td>
<td>64.50±2.40</td>
<td>38/22</td>
<td>18.20±3.10</td>
<td>12</td>
<td>China</td>
</tr>
<tr>
<td></td>
<td>Jean-Marc Orgogozo, 2014</td>
<td>76.60±6.50</td>
<td>72/75</td>
<td>16.90±2.60</td>
<td>28</td>
<td>Belgium, Switzerland</td>
</tr>
<tr>
<td>(20 mg/day)</td>
<td>Maurice W, 2014</td>
<td>78.80±7.20</td>
<td>149/6</td>
<td>20.80±3.80</td>
<td>24</td>
<td>USA</td>
</tr>
<tr>
<td></td>
<td>Chris Fox, 2012</td>
<td>84.90±6.70</td>
<td>20/52</td>
<td>7.30±6.20</td>
<td>12</td>
<td>Canada</td>
</tr>
<tr>
<td></td>
<td>Anton P, 2008</td>
<td>74.90±7.64</td>
<td>100/117</td>
<td>16.70±3.68</td>
<td>24</td>
<td>USA</td>
</tr>
<tr>
<td></td>
<td>David Wilkinson, 2011</td>
<td>74.00±9.00</td>
<td>50/83</td>
<td>16.70±2.40</td>
<td>52</td>
<td>UK</td>
</tr>
<tr>
<td></td>
<td>Tao Wang, 2013</td>
<td>65.70±12.50</td>
<td>4/7</td>
<td>14.40±4.60</td>
<td>24</td>
<td>China</td>
</tr>
<tr>
<td></td>
<td>Huperzine-A (0.2 mg/day)</td>
<td>76.80±3.50</td>
<td>22/13</td>
<td>18.80±2.30</td>
<td>16</td>
<td>China</td>
</tr>
<tr>
<td></td>
<td>Zhi-Qian Xu, 2011</td>
<td>72.30±6.90</td>
<td>26/13</td>
<td>16.52±2.17</td>
<td>12</td>
<td>China</td>
</tr>
<tr>
<td></td>
<td>Liu Fugen, 1995</td>
<td>77.80±7.60</td>
<td>2/3</td>
<td>15.70±6.10</td>
<td>40</td>
<td>China</td>
</tr>
<tr>
<td></td>
<td>Zhou, Zhi-Qiang, 2004</td>
<td>70.20±12.30</td>
<td>5/3</td>
<td>14.80±6.20</td>
<td>24</td>
<td>China</td>
</tr>
<tr>
<td></td>
<td>Feng-Ling, Wang, 2011</td>
<td>70.90±5.20</td>
<td>20/10</td>
<td>14.64±4.23</td>
<td>12</td>
<td>China</td>
</tr>
<tr>
<td>Tacrine (80 mg/day)</td>
<td>Philip C, 1994</td>
<td>76.00±141</td>
<td>54/46</td>
<td>16.80±0.56</td>
<td>12</td>
<td>UK</td>
</tr>
<tr>
<td></td>
<td>H Alliana S, 1999</td>
<td>76.00±8.49</td>
<td>50/62</td>
<td>17.10±4.20</td>
<td>12</td>
<td>France</td>
</tr>
<tr>
<td></td>
<td>Wen-Jang, Wong, 1998</td>
<td>73.60±53-94</td>
<td>36/34</td>
<td>16.20±4.70</td>
<td>18</td>
<td>Taiwan, China</td>
</tr>
<tr>
<td></td>
<td>Egger SA, 1992</td>
<td>66.40±40-86</td>
<td>50/39</td>
<td>17.50±0.61</td>
<td>18</td>
<td>Italy</td>
</tr>
<tr>
<td>R vs D vs G</td>
<td>E Aguglia, 2004</td>
<td>R: 78.02±6.40</td>
<td>R: 33/36</td>
<td>R: 20.37±3.56</td>
<td>18</td>
<td>Italy</td>
</tr>
<tr>
<td></td>
<td>D: 77.00±6.40</td>
<td>D: 36/64</td>
<td>D: 20.49±3.82</td>
<td>36</td>
<td>Italy</td>
<td></td>
</tr>
<tr>
<td></td>
<td>G: 77.00±5.90</td>
<td>G: 31/69</td>
<td>G: 21.47±3.80</td>
<td>36</td>
<td>Italy</td>
<td></td>
</tr>
<tr>
<td>R vs D</td>
<td>Rafael Biesa, 2006</td>
<td>68.70±5.30</td>
<td>15/18</td>
<td>15.30±3.30</td>
<td>15.50±3.20</td>
<td>NR</td>
</tr>
<tr>
<td>R vs D vs G</td>
<td>Aurelia Santoro, 2014</td>
<td>75.90±7.30</td>
<td>18/25</td>
<td>15.30±3.30</td>
<td>15.50±3.20</td>
<td>NR</td>
</tr>
<tr>
<td>R vs D</td>
<td>G: 76.10±6.40</td>
<td>48/57</td>
<td>15.46±4.03</td>
<td>36</td>
<td>Italy</td>
<td></td>
</tr>
<tr>
<td>R vs D</td>
<td>R: 76.10±6.80</td>
<td>R: 91/201</td>
<td>R: 19.63±4.31</td>
<td>36</td>
<td>Italy</td>
<td></td>
</tr>
<tr>
<td>R vs D</td>
<td>Roger Bullock, 2005</td>
<td>75.90±6.60</td>
<td>154/341</td>
<td>15.15±3.00</td>
<td>15.13±2.90</td>
<td>104</td>
</tr>
<tr>
<td>G vs D</td>
<td>Gordon Wilcock, 2003</td>
<td>74.10±53-88</td>
<td>157/342</td>
<td>15.15±3.00</td>
<td>15.13±2.90</td>
<td>104</td>
</tr>
<tr>
<td>Huperzine-A vs D</td>
<td>Wei-Wu, 2013</td>
<td>51–86</td>
<td>15/10</td>
<td>19.70±3.40</td>
<td>19.50±3.60</td>
<td>24</td>
</tr>
<tr>
<td>R vs memantine</td>
<td>Hsi-Hua, Huang, 2012</td>
<td>72.20±3.30</td>
<td>13/18</td>
<td>12.52±4.34</td>
<td>12.73±4.57</td>
<td>24</td>
</tr>
<tr>
<td>Memantine vs D</td>
<td>Wen-Guang, Liu, 2011</td>
<td>77.20±4.70</td>
<td>19/27</td>
<td>15.30±2.50</td>
<td>15.40±2.60</td>
<td>24</td>
</tr>
</tbody>
</table>

**Abbreviations:** Con, control; D, donepezil; Exp, experiment; G, galantamine; MMSE, Mini–Mental State Examination; NR, not report; R, rivastigmine.
to summarize available data from published randomized placebo-controlled studies to suggest that memantine, galantamine, huperzine-A, and rivastigmine have a significant benefit in improving cognitive ability in patients with AD. The most probable best drug for cognitive decline is memantine (SUCA$=70.5\%$). Our NMA interpreted with caution to identify the optimal drug to treat AD patients.

Our results showed the consistency with previous research$^{43}$ that the efficacy of galantamine is greater than that of rivastigmine followed by donepezil, but it should be noted that some studies indicated that these three drugs have no significant difference in improving the cognitive function of AD patients in a short duration.$^{44}$

The trials in previous meta-analyses$^{45-51}$ (PubMed search March 1, 2017) only investigated the efficacy of ChEIs, huperzine-A, and tacrine, respectively, and lacked a synthesized analysis among them. By contrast, our NMA assessed major cognitive drugs by applying six pairwise MDs. The integration of indirect comparisons in our NMA led to greater statistical precision in scientific comparisons of cognitive drugs against a CG. Previous studies have consistently demonstrated that almost all of these six drugs have beneficial effects on senior AD patients,$^{50,52-56}$ in particular memantine.$^{21,22}$

AD is a progressive neurodegenerative disorder, and it is still incurable. The effect of ChEIs may be considerable, but any drug that could possibly slow down the progression of AD patients is worth introducing. Major studies demonstrate that ChEIs can improve the cognition function of AD patients,$^{45,46}$ while some other studies show the contrary conclusion.$^{57-59}$ Besides, the dose of ChEIs is proportional to the cognitive ability of AD patients.$^{50,61}$ It should be noted that the efficacy of a series of nondrug interventions to treat AD patients and their effects on cognition have been proved by relevant researches, such as estrogen replacement therapy,$^{62,63}$ psychotherapy,$^{64-66}$ computerized cognitive training.$^{67,68}$

By contrast, pharmacological intervention, one of the biological therapies, is currently the most stable and effective treatment for AD patients. They will bring some adverse effects while objectively improving the cognitive ability of AD patients.

Tacrine is the first drug approved by the Food and Drug Administration for the treatment of AD to be applied in early stage. It is rarely used in clinical practice nowadays because of its dose-limiting hepatotoxicity.$^{69-71}$ Donepezil, rivastigmine, and galantamine were the second-generation ChEIs, and rivastigmine is the most effective for patients with terminal severe AD.$^{72,73}$ Donepezil is the best choice for most mild-to-moderate AD patients due to its excellence in effectiveness.$^{74}$ Previous studies consistently demonstrated

---

**Table 2** The forest plot cognitive outcome based on MMSE

<table>
<thead>
<tr>
<th>Compare with control group</th>
<th>Number of patients</th>
<th>Heterogeneity ($I^2$) (%)</th>
<th>Mean difference (95% CI)</th>
<th>$P$-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Overall</td>
<td>9,820</td>
<td>0.00</td>
<td>0.88 (0.36, 1.78)</td>
<td>0.027</td>
</tr>
<tr>
<td>Donepezil</td>
<td>2,585</td>
<td>6.80</td>
<td>0.20 (−0.75, 1.10)</td>
<td>0.524</td>
</tr>
<tr>
<td>Rivastigmine</td>
<td>2,397</td>
<td>0.00</td>
<td>0.73 (0.08, 1.40)</td>
<td>0.041</td>
</tr>
<tr>
<td>Galantamine</td>
<td>2,418</td>
<td>0.00</td>
<td>1.30 (0.15, 2.50)</td>
<td>0.03</td>
</tr>
<tr>
<td>Memantine</td>
<td>1,557</td>
<td>0.00</td>
<td>1.70 (0.73, 2.80)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Huperzine-A</td>
<td>258</td>
<td>0.00</td>
<td>0.90 (0.12, 1.70)</td>
<td>0.037</td>
</tr>
<tr>
<td>Tacrine</td>
<td>605</td>
<td>12.40</td>
<td>0.44 (−0.30, 1.20)</td>
<td>0.216</td>
</tr>
</tbody>
</table>

**Abbreviation:** MMSE, Mini–Mental State Examination.
that the levels of Ach in the brain increased immediately after administering oral donepezil. What is different from previous studies is that we concluded that not all ChEIs have a significant increase in the cognitive abilities of AD patients as donepezil (MD = 0.2, 95% CI: -0.75, 1.1) and tacrine (MD = 0.44, 95% CI: -0.3, 1.2).

As the new generation of CHIEs, the effect of memantine is designed as a NMDA receptor antagonist to have a greater extent than any previous ChEIs. Despite this, it seems illegitimate that memantine is not significantly more effective than placebo according to certain studies. That is probably because there are different studies presenting a difference in methodological design. Some differences were mentioned in a long-term treatment for the memantine and donepezil’s outcomes, which are probably because of the difference between studies. A few studies reveal no differences between the effect of AChEs on cognition even if there is a decline in the cognitive function in MMSE. Moreover, the majority of studies suggest that the treatment from cognitive drugs is a long-term instead of temporary process.

It is worth mentioning that a peculiar drug – huperzine-A – the most successful drug for treating AD developed in China, is a reversible and highly effective selective ChEI. Its co-efficiency on safety and endurability is relatively high unlike other drugs analyzed in our study. Even if there were AEs, most adverse effects were mildly and generally diminished by the time as long as the treatment was not paused, but the previous reviews were inconclusive because those drugs had not been compared directly. There are also some studies analyzing the combination of each ChEIs showing that the efficacy of combinational therapy is superior to single drug such as donepezil in combination with memantine, the outcome of which exhibited a preferable improvement than donepezil alone. However, we did not include this type of studies because these data from such studies directly influence the analysis of the evidence for the effects of single drug.

**Strength**

Instead of merely grouping various drugs into pharmacological therapy, the greatest strength of the NMA is that it assessed each drug individually and compared major cognitive drugs from all generations simultaneously. Our NMA applied a trial sequential analysis to detect the internal robustness and reliability of evidence from each study that
can produce relatively scientific interpretation in terms of causal relationships. We performed an extensive search strategy across several databases and sources to obtain an adequate number of eligible studies regardless of the languages they were written. We also extensively searched the bibliographies of published studies. When it comes to methodology, our NMA demonstrates a series of preponderance of Bayesian NMA for comparing various cognitive drugs and for evaluating the relative effectiveness of multifarious drugs. In this context, the results of our NMA are likely to be more useful for decision makers, service commissioners, and caregivers when they are making choices among different alternatives than results from multiple separate traditional meta-analyses, because several relevant outcomes have been assessed simultaneously.90

Limitation
The limitation of our study also needs to be acknowledged. First, the quality of the included studies was uneven because we have to make sure that the number of studies is enough to conduct our NMA. Second, the direct evidence of rivastigmine in our included studies was relatively small; thus, pooled analysis yielded heterogeneity, although it has a certain number of indirect evidence when compared with other cognitive drugs. Third, we did not analyze the dose of each drug neither the stage-dependent effects (mild, moderate, and severe), which represent subject qualifications or constraints that are typical for these studies. This was because only three included studies have described of them. Fourth, our NMA used the same outcome as the basis, and the necessary data we must extract for further assessment such as the mean, SD, and sample size values at baseline. Last but not least, quite a number of studies do not have the abovementioned data, which makes available studies even less. And finally, our analyses were only based on a limited cognitive assessment instrument (MMSE). The abovementioned reasons can also explain why huperzine-A and tacrine hold a relevant limited sample size compared with other drugs. At last, we have not extracted the number of patients who have been observed in trials of numerous AEs or other reasons because not all included studies have described it except some high-quality studies.

Conclusion
Our NMA suggested that memantine was the optimum cognitive drug for patients with AD, whereas donepezil is the worst one. In the future, there is a need to include more studies of high methodological quality related to comprehensive cognitive drugs to help establish a more extensive literature foundation.

Acknowledgments
We would like to gratefully acknowledge the help by the following authors: Jia-Yan Qiu, Deng-Juan Qian, and Wen-Hao Du. This research did not receive any funding from agencies in the public, commercial, or not-for-profit sectors.

Author contributions
JHL conducted the database search, screened and extracted data for the meta-analysis, prepared extracted data for the procedures, and had primary responsibility in writing this article. RXJ and YQW performed statistical analysis and interpretation of data. HBZ, RKW, and LH contributed to the discussion and editing. JYL and YX critically revised the draft manuscript. All authors contributed toward data analysis, drafting and critically revising the paper and agree to be accountable for all aspects of the work.

Disclosure
The authors report no conflicts of interest in this work.

References


Supplementary materials

Figure S1 Risk of bias assessment.

Figure S2 Risk of bias assessment (summary graph).

Figure S3 Funnel plot.

Abbreviations: SE, standard error; SMD, standard mean difference.
**Figure S4 (A,B)** Inconsistency plot.

**Abbreviations:** CG, control group; Don, donepezil; Gal, galantamine; H-A, huperzine-A; Mem, memantine; Riv, rivastigmine; Tac, tacrine.