Is water exchange superior to water immersion in detecting adenomas during colonoscopies? Results from a Bayesian network meta-analysis

Xin Shi1,*, Dan Tian2,*, Xiaofei Ye3,*, Qiong Wu1, Yanglin Pan1, Zhiping Yang1 and Daiming Fan1

1State Key Laboratory of Cancer Biology, National Clinical Research Center for Digestive Diseases and Xijing Hospital of Digestive Diseases, Fourth Military Medical University, Xi’an, China
2Office of Educational Administration, Fourth Military Medical University, Xi’an, China
3Department of Health Statistics, Second Military Medical University, Shanghai, China

*These authors contributed equally to this work

Correspondence to: Zhiping Yang, email: zpyang@fmmu.edu.cn
Yanglin Pan, email: yanglinpan@hotmail.com

Keywords: colonoscopy; water exchange; water immersion; colorectal adenoma; network meta-analysis

Received: February 14, 2018   Accepted: May 08, 2018   Published: July 17, 2018

Copyright: Shi et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License 3.0 (CC BY 3.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

ABSTRACT

Aim: Water-assisted colonoscopy (water exchange [WE] and water immersion [WI]) has been shown to improve the adenoma detection rate. However, few studies have compared these two methods head-to-head. Thus, we conducted a network meta-analysis to integrate both direct and indirect evidence comparing the effectiveness of these two procedures.

Method: We searched PubMed, Web of Science, and the Cochrane Central Register of Controlled Trials for original papers and abstracts published up to March 2018. Randomized controlled trials (RCTs) reporting data in accordance with the eligibility criteria were included in this study. We performed a Bayesian random effects network meta-analysis with mixed comparisons.

Results: Twenty-nine studies (n = 11464 patients) including 6 direct and 23 indirect comparisons were included in this network meta-analysis. There was a statistically significant difference in the efficacy of adenoma detection when WE was compared with WI (risk ratio [RR]: 1.2, 95% credible interval [CrI]: 1.1–1.3), air insufflation (AI; RR: 1.3, 95% CrI: 1.1–1.4), and carbon dioxide (CO2) insufflation (RR: 1.2, 95% CrI: 1.1–1.5). The different methods were ranked in order from the most to least effective in adenoma detection as follows: WE, WI, AI, and CO2. Moreover, although there were no significant differences in pain scores, willingness to repeat, caecal intubation rate, or total procedure time between WI and WE colonoscopy, WE required a longer caecal intubation time than WI.

Conclusion: This network meta-analysis supposes that WE may be superior to WI in detecting adenomas during colonoscopies without affecting other technical features or patient acceptance.

INTRODUCTION

A colonoscopy is considered the gold standard for diagnosing colonic diseases and an important therapeutic modality [1]. It can reduce the risk of death from colorectal cancer by detecting tumours at an earlier, more treatable stage and removing precancerous adenomas [2, 3]. The adenoma detection rate (ADR), defined as the
proportion of patients with at least one adenoma detected during a screening colonoscopy in average-risk patients >50 years, has been shown to be inversely associated with the development of interval colorectal cancer [4, 5]. In our previous study, the recurrence rates of adenomas among post-polypectomy patients were high [6]. Thus, it is essential to improve the ADR to benefit patients undergoing a colonoscopy [7, 8].

Traditionally, a diagnostic colonoscopy begins with air insufflation (AI) to distend the colonic lumen to permit visualization and passage. Gas insufflation with either air or carbon dioxide (CO2) can be painful and poorly tolerated by patients [9, 10]. Due to these undesirable outcomes, investigators adopted the use of water-aided methods in lieu of gas insufflation to improve comfort during the insertion procedure. Water-aided methods comprise two major categories, namely, water immersion (WI) and water exchange (WE). WI is characterized by suction removal of the infused water during the withdrawal phase of a colonoscopy, and WE is characterized by suction removal of the infused water predominantly during the insertion phase [11, 12].

Recent studies revealed that, compared with AI, WE significantly improved the ADR [13, 14]. However, previous systematic reviews and meta-analyses showed controversial results regarding the differences between the ADRs of water infusion (water immersion or water exchange methods) and standard air insufflation for colonoscopies [15–20] (Supplementary Table 1). Hafner S [15] found that adenoma detection rate was slightly improved with water infusion (risk ratio 1.16, 95% CI 1.04–1.30, P = 0.007). But others showed no significant statistical differences. To our knowledge, there are few studies with head-to-head comparisons of the ADRs associated with WE and WI. Thus, we conducted a network meta-analysis to determine the comparative effectiveness of different water-aided methods during colonoscopy.

**METHODS**

The methodology and reporting of this study complies with the PRISMA statement for reporting systematic reviews that incorporated network meta-analyses of health care interventions [21].

**Eligibility criteria**

For this network meta-analysis, we only considered randomized controlled trials (RCTs) irrespective of publication status or date of publication. We excluded studies of other design because of the risk of bias associated with such studies. We included RCTs with participants who underwent a colonoscopy with or without sedation. We included RCTs that compared WI or WE with AI or CO2, and we excluded other assisted colonoscopy methods such as cap-assisted WI, water-assisted plus indigo carmine, and water-assisted plus CO2.

Studies reporting both included and excluded populations were considered only if subject data specifically meeting the inclusion criteria for this network meta-analysis could be extracted (e.g., data for subjects who received WE as a single method in a trial comparing WE alone vs. WE with additional CO2). Abstracts reporting unpublished studies were also considered if sufficient data were reported.

**Information sources and search**

A bibliographic search was performed using PubMed, Web of Science, and the Cochrane Central Register of Controlled Trials to identify original studies analysing the efficacy of water-assisted colonoscopy. The search strategy was as follows: colonoscopy, water and randomized or randomized trial or clinical trial. Moreover, reference lists from the retrieved articles, systematic reviews and meta-analyses were evaluated to identify additional relevant studies. To obtain the maximum number of papers, no publication date restrictions were imposed in any database. The last search was completed in March 2018. Only English language publications were included in this study.

**Study selection**

Two investigators (X.S. and D.T.) independently identified trials for inclusion by screening titles and abstracts yielded by the search. To potentially include additional studies, we searched for full-length articles of all references from the studies that the investigators had identified in the initial search. We selected trials for inclusion based on the review of these full-length articles. Queries concerning inclusion were resolved by discussion and consensus between the two investigators.

**Data extraction**

Two investigators (X.S. and D.T.) independently extracted the following data from each study and summarized the data in tables: first author’s name, publication year, country, number of centres, registration numbers, interventions, indications, sedation, sample size, percentage of males, age of each intervention, number of participants per group, number of participants per group with at least one detected adenoma, pain score, willingness to repeat colonoscopy, number of participants per group who completed caecal intubation, caecal intubation time, and total procedure time. Differences in opinion between the investigators were resolved through discussion. Pain scores were derived from 0 to 10 visual analogue or numeric rating scales with 0 = none and 10 = most severe pain; if an included study reported 0–5 or other visual analogue scales, then the scores were converted. For continuous data, if an included RCT with a sample size larger than 100 reported only the median and interquartile
range (IQR), then the median was regarded as the mean, and the standard deviation was calculated as IQR/1.35 [22].

Risk of bias assessment

We assessed the risk of bias as described in the Cochrane Handbook for Systematic Reviews of Interventions [23]. Bias risk assessment was performed with reference to the following domains: 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 detectable biases. Each risk of bias component was rated as low risk, unclear risk, or high risk of bias. A study with sufficient data to fulfill the criteria for the quality item was marked “low risk of bias”. A study in which the reported data did not fulfill the criteria for the quality item was marked “high risk of bias”. “Unclear risk of bias” was marked for studies that did not report the necessary data to assess the quality item.

Outcome measures

The primary outcome measure was the ADR, which was defined as the proportion of patients in whom at least one adenoma was detected during a colonoscopy. Secondary outcome measures were the mean and maximum pain scores (defined as the mean and maximum pain scores of patients during the colonoscopy), willingness to repeat rate (defined as the proportion of patients who were willing to repeat the procedure after the colonoscopy), caecal intubation rate, caecal intubation time, and total procedure time.

Statistical analysis

Given the limited number of studies that directly compared WE and WI, we used a network meta-analytical approach to simultaneously combine direct comparisons of these interventions within studies and, if available, indirect comparisons between studies. Whenever possible, we used results from intention-to-treat analyses. Network meta-analysis was conducted by using a Bayesian Markov-chain Monte Carlo method, and data were fitted in R statistical software version 3.2.2 (R Foundation for Statistical Computing, Vienna, Austria). Analytical results were presented as risk ratios (RRs) and mean differences (MDs) with 95% credible intervals (CrIs). A network diagram was used for each outcome to present the comparisons between different water-assisted colonoscopy methods and common AI or CO2. In these diagrams, nodes represented different methods of colonoscopies, lines indicated direct comparisons and line thickness was proportional to the number of available studies. We ranked methods of colonoscopy based on analysis of ranking probabilities derived from Monte Carlo simulations [24]. The probabilities of different rankings of the same intervention were summed to 100%. Inconsistency between direct and indirect evidence was assessed by a “node-splitting” approach [25].

We did not perform any subgroup analysis or sensitivity analysis due to the lack of data. We conducted Egger’s regression asymmetry test to examine potential publication bias in terms of ADR by using STATA software version 12.0 (Stata Corporation, College Station, Texas, USA).

RESULTS

Study selection

Figure 1 shows the detailed steps of the study selection process. The literature search yielded 479 potentially relevant studies. Of these, 32 potentially eligible studies were retrieved from the electronic databases, and 2 additional relevant studies were identified through a manual search of the reference section of identified articles and former meta-analyses, yielding a total of 34 studies. After excluding 5 studies based on the predefined inclusion criteria, 29 studies were included in the network meta-analysis.

Study characteristics and network geometry

Study characteristics and outcome measures of the 29 included studies are summarized in Table 1 and Supplementary Table 2. Six trials [26–31] including 2088 patients (18.2%) were analysed in per-protocol, others were analysed in intention-to-treat. The studies were published between 2009 [32] and 2017 [13, 14, 33]. Study population sizes ranged from 23 [34] to 3303 [13], and 11464 patients were included in total, which equated to 1998 patients in the WI group, 3778 patients in the WE group, 4431 patients in the AI group, and 696 patients in the CO2 group. The proportion of male patients ranged from 30.4% [34] to 100% [31], and the mean patient age varied from 51.1 [35] to 69 [34] years in the WI group, 36 [36] to 65.6 [27] years in the WE group, 40.2 [36] to 67 [34] years in the AI group, and 54.3 [26] to 63.1 [27] years in the CO2 group.

In 18 studies evaluating the ADR [13, 14, 27, 29–31, 34, 37–47], these rate ranged from 25% [46] to 54.5% [34] in the WI group, 18.3% [13] to 67.5% [27] in the WE group, 13.3% [13] to 46% [41] in the AI group, and 26.5% [30] to 57.8% [27] in the CO2 group. Of 27 studies evaluating pain scores [13, 26–40, 42–52], 13 studies evaluated the mean pain score [27, 29–32, 34–36, 40, 44, 46, 50, 52], which ranged from 1.3 [32] to 4.1 [31, 35, 50] in the WI group, 1.3 [29] to 4.3 [27] in the WE group, 2.3 [29, 30] to 6.4 [35] in the AI group, and 2 [30] to 4.8 [27] in the CO2 group, whereas 14 studies evaluated the
maximum pain score [13, 26, 28, 33, 37–39, 42, 43, 45, 47–49, 51], which ranged from 2.3 [43] to 4 [49] in the WI group, 1.1 [28] to 3.3 [33] in the WE group, 2.9 [28] to 5.7 [26] in the AI group, and 2.9 [26] to 4.7 [33] in the CO₂ group. In 21 studies evaluating the willingness to repeat rate [13, 14, 26, 28, 32, 34, 35, 38–40, 42–52], this rate ranged from 72.4% [42] to 100% [34, 38, 52] in the WI group, 76% [39] to 100% [38, 52] in the WE group, 48% [39] to 98.9% [38] in the AI group, and 88.7% [40] to 93.8% [26] in the CO₂ group. In 27 studies evaluating caecal intubation time [13, 14, 26–33, 35–49, 51, 52], this timing ranged from 5.6 minutes [44] to 34 minutes [47] in the WI group, 6.9 minutes [41] to 17.5 minutes [52] in the WE group, 4.6 minutes [44] to 37 minutes [47] in the AI group, and 5 minutes [40] to 9.8 minutes [33] in the CO₂ group. In 22 studies evaluating total procedure time [14, 26, 30–37, 39–49, 51], the total timing ranged from 14 minutes [34] to 56 minutes [47] in the WI group, 15.7 minutes [36] to 29 minutes [39] in the WE group, 13.1 minutes [44] to 56 minutes [47] in the AI group, and 13 minutes [40] to 20.5 minutes [49] in the CO₂ group.

Figure 2 shows the network of comparisons for each outcome of interest. In relation to ADR, 3 trials compared WI and WE, 7 trials compared WE and AI, 1 trial compared WE and CO₂, 13 trials compared WI and AI, 2 trials compared WI and CO₂, and 2 trials compared AI and CO₂.

Quality of trials

The quality of the studies included in the network meta-analysis, as assessed using a risk of bias assessment tool, is shown in Figure 3 and Supplementary Figure 1. In relation to random sequence generation, most of the trials were rated as having a “low risk of bias” (19 of 29 trials), and regarding allocation concealment, 50% of trials were rated as having a “low risk of bias” (15 of 29 trials) because they used sealed opaque envelopes. Trials rated as having an “unclear risk of bias” had issues related to blinding of participants and personnel and of outcome assessment (27 and 26, respectively). In terms of attrition bias, reporting bias, and other biases, all trials were rated as having a “low risk of bias” except one [29], which was rated as having a “high risk of bias” in relation to reporting bias.

Meta-analysis

Figure 4 shows the results of the relative effect between any pair of interventions and each outcome of interest. There was a statistically significant difference in the ADR when WE was compared with WI (RR: 1.2, 95% CrI: 1.1–1.3), AI (RR: 1.3, 95% CrI: 1.1–1.4), and CO₂ insufflation (RR: 1.2, 95% CrI: 1.1–1.5). In terms of mean and maximum pain scores, willingness to repeat, caecal

Figure 1: Study flowchart.
intubation rate, and total procedure time, there were no significant differences between WI and WE colonoscopy. However, WE required a longer caecal intubation time than WI (MD: 3.3, 95% CrI: 1.5–5.1). Ranking probabilities for all methods are presented in Figure 5. The efficacies of the different methods in adenoma detection were ranked in order from the most to least effective as follows: WE (100%), WI (82%), AI (67%), and CO₂ (73%).

Node-splitting analysis comparing results between direct and indirect estimates did not produce significant evidence of inconsistency, except when the caecal intubation time of WE was compared with that of AI or CO₂ (both $P = 0.04$, see Supplementary Table 3). We then compared WE with AI or CO₂ using an inconsistency model (MD: 2.66, 95% CrI: 1.05–4.19 and MD: 1.73, 95% CrI: 2.38–4.96, respectively).

Egger’s test of ADR for publication bias showed that the risk of having missed or overlooked trials was insignificant: WI vs. WE, $P = 0.264$; WI vs. AI, $P = 0.331$; WE vs. AI, $P = 0.726$. With regard to the remaining comparisons, we were unable to evaluate publication bias due to the small number of included studies.

Subgroup or sensitivity analysis was not conducted because of the limited data available.

Table 1: Study characteristics

<table>
<thead>
<tr>
<th>Author</th>
<th>Year</th>
<th>Country</th>
<th>Centres</th>
<th>Registration No.</th>
<th>Interventions</th>
<th>Indications</th>
<th>Sedation</th>
<th>Sample size</th>
<th>Male (%)</th>
<th>Age (y, mean ± SD)</th>
<th>WI</th>
<th>WE</th>
<th>AI</th>
<th>CO₂</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cadoni S [14]</td>
<td>2017</td>
<td>Italy, Czech</td>
<td>3</td>
<td>NCT02041507</td>
<td>WI, WE, AI</td>
<td>Scr</td>
<td>On-demand</td>
<td>1224</td>
<td>54.9</td>
<td>61 ± 6.3</td>
<td>61.4 ± 6.2</td>
<td>60.9 ± 6.2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Jia H [13]</td>
<td>2017</td>
<td>China</td>
<td>6</td>
<td>NR</td>
<td>WE, AI</td>
<td>Scr, Sur, and Dx</td>
<td>Sedated</td>
<td>3303</td>
<td>51</td>
<td>-</td>
<td>50.3 ± 13.7</td>
<td>50.3 ± 13.9</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cadoni S [33]</td>
<td>2017</td>
<td>Italy, Czech</td>
<td>2</td>
<td>NCT02409979</td>
<td>WE, CO₂</td>
<td>NR</td>
<td>Unsedated</td>
<td>403</td>
<td>61.3</td>
<td>-</td>
<td>65.6 ± 12</td>
<td>63.1 ± 12.6</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hsieh YH [37]</td>
<td>2017</td>
<td>China</td>
<td>2</td>
<td>NCT01894191</td>
<td>WI, WE, AI</td>
<td>Scr, Sur, and Dx</td>
<td>Sedated</td>
<td>651</td>
<td>51.6</td>
<td>55.9 ± 10.2</td>
<td>55.7 ± 10.6</td>
<td>54.8 ± 10.5</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Azari M [27]</td>
<td>2016</td>
<td>Japan</td>
<td>2</td>
<td>NCT000009706</td>
<td>WE, CO₂</td>
<td>NR</td>
<td>Unsedated</td>
<td>287</td>
<td>48.3</td>
<td>54.3 ± 12.4</td>
<td>-</td>
<td>55 ± 10.7</td>
<td>54.3 ± 11.7</td>
<td></td>
</tr>
<tr>
<td>Xu X [26]</td>
<td>2016</td>
<td>China</td>
<td>1</td>
<td>NR</td>
<td>WI, AI, CO₂</td>
<td>Dx</td>
<td>Unsedated</td>
<td>92</td>
<td>53.3</td>
<td>-</td>
<td>36 ± 13</td>
<td>40.2 ± 14.6</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fai P [36]</td>
<td>2015</td>
<td>Czech</td>
<td>1</td>
<td>NCT01933867</td>
<td>WE, AI</td>
<td>Dx</td>
<td>On-demand</td>
<td>576</td>
<td>58.7</td>
<td>60 ± 10.8</td>
<td>59 ± 11.3</td>
<td>59 ± 11.3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cadoni S [48]</td>
<td>2015</td>
<td>Italy</td>
<td>2</td>
<td>NCT01781650</td>
<td>WI, WE, AI</td>
<td>Scr, and Dx</td>
<td>On-demand</td>
<td>624</td>
<td>58.7</td>
<td>59 ± 12.3</td>
<td>58 ± 15.5</td>
<td>58 ± 12.7</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Wang X [28]</td>
<td>2015</td>
<td>China</td>
<td>1</td>
<td>NCT01735266</td>
<td>WE, AI</td>
<td>94-95% Dx</td>
<td>Unsedated</td>
<td>300</td>
<td>47.3</td>
<td>-</td>
<td>46.6 ± 11.8</td>
<td>48.6 ± 13.2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Miroslav V [50]</td>
<td>2014</td>
<td>Yugoslavia</td>
<td>1</td>
<td>NR</td>
<td>WI, AI</td>
<td>Scr, Sur, and Dx</td>
<td>Sedated</td>
<td>122</td>
<td>47.5</td>
<td>54.4 ± 14.8</td>
<td>-</td>
<td>58 ± 13</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hsieh YH [38]</td>
<td>2014</td>
<td>China</td>
<td>1</td>
<td>NCT01535326</td>
<td>WI, WE, AI</td>
<td>Scr, Sur, and Dx</td>
<td>Minimal</td>
<td>270</td>
<td>62.2</td>
<td>54.3 ± 11.4</td>
<td>56.9 ± 10.3</td>
<td>56.5 ± 13.7</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cadoni S [29]</td>
<td>2014</td>
<td>Italy</td>
<td>2</td>
<td>NCT0146319</td>
<td>WE, AI</td>
<td>Scr, Sur, and Dx</td>
<td>On-demand</td>
<td>672</td>
<td>60.3</td>
<td>-</td>
<td>58 ± 12.4</td>
<td>60 ± 12.3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Leung J [39]</td>
<td>2013</td>
<td>USA</td>
<td>1</td>
<td>NCT01383252</td>
<td>WE, AI</td>
<td>NR</td>
<td>Unsedated</td>
<td>100</td>
<td>97</td>
<td>-</td>
<td>61 ± 7</td>
<td>60 ± 6.7</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Amato A [40]</td>
<td>2013</td>
<td>Italy</td>
<td>1</td>
<td>NCT01259583</td>
<td>WI, AI, CO₂</td>
<td>Scr, Sur, and Dx</td>
<td>Sedated</td>
<td>341</td>
<td>64.5</td>
<td>60 ± 11.5</td>
<td>-</td>
<td>60 ± 13.4</td>
<td>61.5 ± 14</td>
<td></td>
</tr>
<tr>
<td>Luo H [51]</td>
<td>2013</td>
<td>China</td>
<td>1</td>
<td>NCT01485133</td>
<td>WE, AI</td>
<td>Scr, Sur, and Dx</td>
<td>Sedated</td>
<td>110</td>
<td>30.9</td>
<td>-</td>
<td>55.8 ± 11</td>
<td>54.6 ± 12</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bayopyarmacna P [35]</td>
<td>2013</td>
<td>Indonesia</td>
<td>1</td>
<td>NCT01341847</td>
<td>WI, AI</td>
<td>Dx</td>
<td>Unsedated</td>
<td>110</td>
<td>65.5</td>
<td>51 ± 14.7</td>
<td>-</td>
<td>50.4 ± 15.9</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hsieh YH [52]</td>
<td>2013</td>
<td>China</td>
<td>1</td>
<td>NCT0090555</td>
<td>WI, WE, AI</td>
<td>Scr, Sur, and Dx</td>
<td>Minimal</td>
<td>200</td>
<td>62</td>
<td>53 ± 12</td>
<td>57 ± 11</td>
<td>57 ± 13</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Portocarrerro DJ [34]</td>
<td>2012</td>
<td>USA</td>
<td>2</td>
<td>NR</td>
<td>WI, AI</td>
<td>Scr</td>
<td>Unsedated</td>
<td>23</td>
<td>30.4</td>
<td>69 ± 10</td>
<td>-</td>
<td>67 ± 15</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fai P [30]</td>
<td>2012</td>
<td>Czech</td>
<td>1</td>
<td>NCT01440543</td>
<td>WI, AI, CO₂</td>
<td>Scr, Sur, and Dx</td>
<td>Minimal</td>
<td>420</td>
<td>51.7</td>
<td>60.1 ± 13.9</td>
<td>-</td>
<td>58 ± 13.8</td>
<td>59.4 ± 14.5</td>
<td></td>
</tr>
<tr>
<td>Hsieh YH [44]</td>
<td>2011</td>
<td>China</td>
<td>1</td>
<td>NR</td>
<td>WI, AI</td>
<td>Dx</td>
<td>Minimal</td>
<td>153</td>
<td>56.9</td>
<td>52.4 ± 13.5</td>
<td>-</td>
<td>56.3 ± 13.2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ramirez FC [41]</td>
<td>2011</td>
<td>USA</td>
<td>1</td>
<td>NR</td>
<td>WE, AI</td>
<td>Scr</td>
<td>Minimal</td>
<td>368</td>
<td>96.5</td>
<td>-</td>
<td>60 ± 0.5</td>
<td>59.3 ± 0.5</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pohl J [42]</td>
<td>2011</td>
<td>Germany</td>
<td>1</td>
<td>DRKS00000431</td>
<td>WI, AI</td>
<td>Scr, Sur, and Dx</td>
<td>On-demand</td>
<td>116</td>
<td>73.3</td>
<td>62.7 ± 9.7</td>
<td>-</td>
<td>61.7 ± 11.5</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Leung J [45]</td>
<td>2011</td>
<td>USA</td>
<td>1</td>
<td>NCT00920751</td>
<td>WI, AI</td>
<td>Scr, Sur</td>
<td>On-demand</td>
<td>100</td>
<td>99</td>
<td>60.7 ± 8.1</td>
<td>-</td>
<td>58.3 ± 7</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Leung FW [47]</td>
<td>2010</td>
<td>USA</td>
<td>1</td>
<td>NCT00747084</td>
<td>WI, AI</td>
<td>Scr, Sur, and Dx</td>
<td>Sedated</td>
<td>82</td>
<td>NA</td>
<td>66 ± 8.6</td>
<td>-</td>
<td>66.8 ± 8.4</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Leung CW [31]</td>
<td>2010</td>
<td>USA</td>
<td>1</td>
<td>NCT0067177</td>
<td>WI, AI</td>
<td>Scr, Sur, and Dx</td>
<td>Minimal</td>
<td>229</td>
<td>100</td>
<td>62.5 ± 8.9</td>
<td>-</td>
<td>62.5 ± 8.9</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Radaelli F [46]</td>
<td>2010</td>
<td>Italy</td>
<td>1</td>
<td>NCT0090555</td>
<td>WI, AI</td>
<td>Scr, Sur, and Dx</td>
<td>On-demand</td>
<td>230</td>
<td>58.3</td>
<td>58.4 ± 11.5</td>
<td>-</td>
<td>58.8 ± 13.3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ransibrahmanakul K [45]</td>
<td>2010</td>
<td>USA</td>
<td>1</td>
<td>NCT00641282</td>
<td>WI, AI</td>
<td>Scr, Sur</td>
<td>Minimal</td>
<td>62</td>
<td>98.4</td>
<td>61 ± 7.9</td>
<td>-</td>
<td>61 ± 7.8</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Leung JW [32]</td>
<td>2009</td>
<td>USA</td>
<td>1</td>
<td>NCT00785889</td>
<td>WI, AI</td>
<td>Scr, Sur</td>
<td>Minimal</td>
<td>56</td>
<td>91.1</td>
<td>60 ± 6.6</td>
<td>-</td>
<td>59 ± 8.6</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

USA, United States of America; NR, not reported; WI, water immersion; WE, water exchange; AI, air insufflation; CO₂, carbon dioxide; Scr, screening; Sur, surveillance; Dx, diagnosis.
DISCUSSION

Colonoscopies are widely used for prevention and early detection of colorectal cancer. The detection and removal of adenomas, which are major precursor lesions for colorectal cancer, is considered a crucial aspect of cancer prevention [53]. An increase in the ADR has been shown to reduce risks of interval colorectal cancer and colorectal cancer mortality [54], and the ADR has been codified as a key quality indicator for colonoscopy procedures [55, 56]. It has been estimated that every 1% increase in the ADR predicts a 3% decrease in the risk of colorectal cancer.

Figure 2: Network of comparisons included in the analyses. Solid lines represent direct comparisons within the RCTs. Numbers denote trials that compared corresponding interventions. (A) ADR; (B) mean pain scores; (C) maximum pain scores; (D) willingness to repeat; (E) caecal intubation rate; (F) caecal intubation time; and (G) total procedure time.

Figure 3: Risk of bias graph.
of interval cancer and a 5% decrease in the risk of a fatal interval cancer [57]. Recently published quality guidelines by the European Society of Gastrointestinal Endoscopy recommend that endoscopists achieve an ADR of ≥25% [58], and the American Society of Gastrointestinal Endoscopy recommended an ADR ≥30% in male patients and ≥20% in female patients [55].

The use of water in colonoscopies was first introduced in 1984 as a method for difficult sigmoid circumstances due to diverticulosis [59]. WE was first introduced in 2007 as a modification of WI [60]. The initial purpose of WI was to allow colonoscopists to correctly identify the colonic lumen in order to pass the colonoscope through sigmoid colons that were deformed by diverticulosis and, as a secondary endpoint, to provide a more comfortable procedure for the patients [15]. Thus far, comparative studies have not been able to distinguish the differences in the ADRs between these
two colonoscopy approaches. Our network meta-analysis showed that, compared with WI colonoscopy, WE colonoscopy statistically significantly improved the ADR (RR: 1.2, 95% CrI: 1.1–1.3) but did not significantly affect the patients’ pain scores, willingness to repeat rate, caecal intubation rate, or total procedure time. However, WE significantly prolonged the caecal intubation time (MD: 3.3, 95% CrI: 1.5–5.1).

To date, there have been two head-to-head comparisons of WE, WI and AI using the ADR as the primary outcome. Cadoni et al [14] conducted an RCT with 408 patients per group, and ADR in the WE group was numerically higher than in the WI group (49.3% vs. 43.4%, \( P = 0.28 \)). However, this difference was not significant. Another study enrolled 651 participants (217 per group) and found that the ADRs in the WE and WI groups were 49.8% and 40.6%, respectively [37]. Although the ADR of the WE group was numerically higher, this difference was also not significant (\( P = 0.064 \)). Therefore, distinguishing the difference between the ADRs...
of the WE and WI groups may require a larger sample size. In addition, according to a previous study, a sample of 2447 patients per group would be required to show a significant difference in the ADR between WE and WI [14]. Despite the indirect nature of network meta-analysis, we found that WE method of colonoscopy improved the ADR when compared with WI.

The mechanism of the potential superiority of WE over WI is not fully understood. There were some plausible mechanisms of how WE could significantly increase adenoma detection. As shown in Supplementary Table 2 and previous reports [13, 14, 37, 38, 41], WE significantly increased the quality of colon cleanliness to excellent levels both in the entire colon and the ascending colon, and this quality ultimately led to an increase in the ADR. The WE method involves continuous water infusion and suction of residual faeces and air to clear the view, which enables better visualization of adenomas. The turbulence created at the tip of the colonoscope effectively dislodges residual faeces and debris adhered to the mucosa to facilitate removal by aspiration [14]. Inadequate or fair bowel preparation has been proven to be associated with an increased miss rate of adenoma detection [61, 62]. Thus, bowel cleanliness might have contributed to the increased ADR associated with the WE method. The reduced use of suction during the withdrawal step of the WE procedure has been shown to result in few collapses of the lumen or contractions of the colon; moreover, colonoscopists can concentrate on inspecting the colon instead of being distracted by the need to suction residual faeces and water [63]. The longer insertion time required by WE may also play a role in detecting more adenomas during insertion [11]. In addition, Hsieh et al [38] found that WE (24.4%) achieved a higher ADR during insertion than WI (15.6%) and AI (14.4%).
Previous studies reported that age ≥50 years, male gender, body mass index ≥25 kg/m², Boston Bowel Preparation Scale (BBPS) score ≥6, indications of screening and surveillance, WE colonoscopy, and inspection time >8 minutes were significant predictors of increased adenoma detection [13, 37]. We noticed that the included studies showed a wide variation in these variables. Indications such as screening, surveillance, or diagnosis varied among studies, and the recruited patients with a history of colorectal adenoma remained at high risk for colorectal adenoma recurrence [64]. There were three trials [14, 34, 41] only included patients underwent screening colonoscopy, and five articles [26, 33, 36, 44] only included patients underwent diagnosis colonoscopy. Others included patients with indications for screening, surveillance, or diagnosis. This may contribute to potential selection bias. There was a wide variation in gender distribution, which was important because the ADR of men was significantly higher than that of women. Most of patients (≥90%) included in six trials [31–32, 39, 41, 43, 45] were male which could also contribute potential selection bias in our study. However, the data were sparse, and thus, we could not conduct further analysis to control for these confounding factors.

Recent studies revealed that the pain scores associated with WE were considerably lower than those associated with WI [37, 38, 48, 49, 52, 65]. Our study showed that although the differences of the pain scores between WI and WE were not significant, the pain score of WE method was numerically lower. As shown in Figure 5, WE was considered as the least painful method for the ranking probabilities of mean and maximum pain scores (57% and 94% respectively). We needed to be cautious in interpreting the outcome of these pain scores because the pain scores most likely differed between participants who underwent the procedure with or without sedation. As the primary endpoint of the current study was ADR, we included patients regardless of whether they had been sedated. Thus, including patients with full and deep sedation in our study may have lessened the impact of the pain-alleviating effect of WE compared to WI.

In our analysis, we noticed that the pain score from CO₂ insufflation was second only to that of AI. Previous studies reported that compared with AI or WE colonoscopy, CO₂ insufflation could significantly reduce post-colonoscopy discomfort and pain but not reduce insertion pain during the colonoscopy [49, 66–70]. CO₂ is absorbed from the bowel more rapidly than air, which results in a reduction in post-colonoscopy gas volume, and is transported by the blood to the lungs, where it is exhaled [71]. Thus, the purpose of using CO₂ in the procedure is to prevent or minimize post-colonoscopy colonic distension. A recent study suggested that the combination of WE during insertion and CO₂ during withdrawal appears to be the optimal choice for decreasing pain during the examination and bloating and other outcomes after the procedure [33].

Our study included the following potential limitations: (1) a limited number of studies directly comparing WE and WI hindered us from conducting subgroup and sensitivity analysis; (2) the inherent indirectness of comparisons in a network meta-analytical approach; (3) in most of the studies, the endoscopists were not blinded to the patient outcomes. Nevertheless, our study is provocative enough to prompt an appropriate response of prospective randomized studies, and consideration given to evaluate the WE method by colonoscopists with low ADR is recommended [72]. Until additional data are available, we believe that WE may be considered superior to WI during colonoscopy.

### Abbreviations

WI: water immersion; WE: water exchange; AI: air insufflation; CO₂: carbon dioxide; ADR: adenoma detection rate; RCT: randomized controlled trial; RR: risk ratio; MD: mean difference; CrI: credible interval.

### Author contributions

Conception and design: Yanglin Pan and Zhiping Yang; analysis and interpretation of the data: Xin Shi, Dan Tian, and Xiaofei Ye; drafting of the article: Xin Shi, Qiong Wu, and Zhiping Yang; critical revision of the article for important intellectual content: Yanglin Pan; and final approval of the article: Daiming Fan.

### CONFLICTS OF INTEREST

All authors declare no conflicts of interest. The authors alone are responsible for the content and writing of the paper.

### FUNDING

This study was not supported by any pharmaceutical corporation or company; the publication fee was paid by the authors’ institution.

### REFERENCES


