Clinical efficacy of robotic spine surgery: an updated systematic review of 20 randomized controlled trials

Wen-xi Sun1, Wei-qiang Huang1, Hua-yang Li1, Hong-shen Wang1, Sheng-li Guo2, Jie Dong3, Bo-lai Chen1* and Yong-peng Lin1*

1State Key Laboratory of Traditional Chinese Medicine Syndrome, Guangdong Provincial Hospital of Chinese Medicine, Guangzhou University of Chinese Medicine, Guangzhou, China
2State Key Laboratory of Traditional Chinese Medicine Syndrome, Guangdong Provincial Hospital of Chinese Medicine, Guangzhou University of Chinese Medicine, Guangzhou, China
3Huizhou Hospital of Guangzhou University of Chinese Medicine, Huizhou, China

*© B Chen and Y Lin contributed equally to this work

Correspondence should be addressed to B Chen or Y Lin
Email: drlinyp@gzucm.edu.cn or chenbolai@gzucm.edu.cn

Keywords
- robotic spine surgery
- pedicle screw placement
- clinical efficacy
- randomized controlled trial
- systematic review

Introduction

Miscellaneous spinal disorders, including degenerative diseases, deformities, and trauma, are among the major causes of disability worldwide (1, 2). The persistently high morbidity rates make it become an important contributor to the global healthcare burden (3). Surgical intervention is the main strategy when patients with severe neurological impairments and unresponsive to conservative therapeutic measures. Since Hibbs and Albee first reported spinal fusion surgery in 1911, this operative procedure has gradually become the most widely used for treating spinal pathologies (4). Pedicle screw placement played an important role in the reconstruction of spinal stability, promoting fusion and early rehabilitation, and has become an essential procedure in spine surgery (5). Rajaee et al. (6) revealed that with the widespread application of pedicle screws, the number of spinal fusions has also increased dramatically by 2.4 times annually from 1998 to 2008.

Conventional freehand screw placement relies on anatomical landmarks, intraoperative fluoroscopy, and clinical experience of the surgeons. For inexperienced doctors, the risk of screw placement failure is higher, which can lead to neurovascular injuries and other complications (7, 8). The pedicle screw misplacement rates of conventional techniques are 30% and 55% in the lumbar and thoracic spines, respectively (9, 10, 11). This contrasts the reported high success rate of robot-assisted pedicle screw placement from 91.5% to 94.4% (12). In nearly two decades, robotic-assisted spine surgery (RSS) has represented a revolution in new surgical...
technologies. Despite being an emerging technology, it has been widely adopted for various surgical procedures (13, 14). The precise positioning performance of robots has greatly promoted the development of minimally invasive surgery (15, 16). Recently, some new spinal robotic systems are introduced into clinical use (17, 18).

Robot-assisted surgical technology is generally thought to be more accurate and stable, and several recent data and meta-analysis confirmed the benefits of robotic surgery in accurate pedicle screw placement (19, 20). However, previous meta-analyses have focused on the accuracy of screw placement, with insufficient attention to clinical outcomes and complications, and have not analyzed the differences in outcomes between different robot types and surgical segments. Therefore, we comprehensively included high-quality randomized controlled trials (RCTs) for a comprehensive in-depth systematic review and meta-analysis. This study aimed to further clarify whether RSS could enhance efficacy and decrease complications by improving the accuracy of the pedicle screw. In addition, this study focused on whether different types of robots, operating sites, and regions/countries affect clinical outcomes. It contributes to the further understanding of the advantages and deficiencies of RSS and provides suggestions on the future directions of this field.

Methods

The work has been described following AMSTAR-2 (21) and PRISMA 2020 (22). The study is listed in the PROSPERO database (registration number: CRD42022375991).

Search strategy

Robot-assisted pedicle screw placement was the subject of an electronic literature search from October 1990 to October 2022 in eight databases, including PubMed, The Cochrane Library, Embase, Web of Science Core Collection, China National Knowledge Infrastructure, China Biology Medicine, Wanfang Digital Periodicals (WAN FANG), and VIP databases. Orthopedic robotic surgery, pedicle screws, randomized controlled trials, and other terms were included in the search, which were only restricted to English and Chinese. A similar approach was used for additional electronic databases. References of eligible studies were also searched. The complete search strategy for PubMed is provided in Supplementary File 1 (see section on supplementary materials given at the end of this article). The whole texts of any eligible studies, as well as their titles and abstracts, were independently evaluated by two researchers (W-X S and W-Q H). Disagreements were discussed with a third researcher (Y-P L).

Inclusion and exclusion criteria

The PICOS principle was used to set the following inclusion criteria: (i) participants: the participants who required pedicle screw insertion including degenerative spinal disease or spinal fractures, and whose age was restricted to 18–60; (ii) Interventions: the robot-assisted technique for treating diseases was used in the experimental group (RA group); (iii) comparisons: the conventional freehand technique was used in the control group (FH group); (iv) results: at least two of the following were used as outcome indicators: the number of screws in grade A and grade A+B positions, visual analog scale (VAS) of low back pain and Oswestry disability index (ODI) scores, radiation dose and exposure time, intraoperative blood loss, operating time, and hospital stay. Based on the Gertzbein–Robbins classification (9), postoperative follow-up CT was used to assess the accuracy of pedicle screw placement; and (v) RCTs were acceptable as study designs.

The exclusion criteria trials were as follows: participants with a history of prior spinal surgery or combined spinal tumors, infections, rheumatic immune diseases, etc. Other exclusion factors include studies by the same authors in different languages, studies conducted on the same subjects during the same period, and studies reported by the same author.

Data extraction

Two researchers (W-XS and H-YL) independently gathered baseline data and outcome markers. Disagreements were settled by conversing with a third researcher (Y-P L) or seeking outside counsel. Baseline information includes the first author, publication year, study country, research design, sex, robot type, screw section, etc. Clinical indicators (VAS scores of low back pain, ODI scores, grade A and A+B screw positions) were the main outcome markers of interest. Safety factors, such as radiation dose, radiation exposure time, intraoperative blood loss, operation time, and hospital stay, served as secondary outcome indicators.

Statistical analysis

Risk ratios (RRs) and 95% CIs were calculated for dichotomous variables. The mean difference (MD) was calculated for continuous variables, and the standard mean difference (SMD) and its 95% CI were used as statistics when the calculation method and units varied. A \( \chi^2 \) test with \( \alpha = 0.05 \) was used to examine heterogeneity between trials. If \( I^2 \leq 50 \) and \( P > 0.05 \), heterogeneity between trials was deemed unimportant, and a fixed effects model was used for the meta-analysis. A meta-analysis employing a random-effects model was performed if \( I^2 > 50 \) and \( P \leq 0.05 \), which indicated considerable heterogeneity, with \( P > 0.05 \) suggesting statistically significant differences in the outcome markers.
Egger’s regression was used to examine funnel plots to determine heterogeneity and bias, adding Begg’s test where there >10 studies were included. Sensitivity analysis (leave-one-out method) was used to evaluate the influence of each study to confirm the consistency of the meta-analysis’s conclusions. We will do subgroup analysis for the region, robot type, and screw section for the key outcome indicators. At \( P < 0.05 \), the differences between the analysis subgroups are deemed significant. R 4.2.1 was used to conduct meta-analyses (www.r-project.org).

The interquartile range and the median of some of the data in the available literature indicate that the data were not normally distributed. Box–Cox and Quantile estimates, which more precisely estimated means and s.d., were suggested by McGrath et al. (23) to modify the data. Using an online scientific calculator (https://smcgrath.shinyapps.io/estmeansd/), these data were converted into means and s.d. The S.E., which was calculated for the meta-analysis, could be applied to both the RA and freehand groups, using the P-value and sample size to calculate the S.D. to complete the data (24). This was done because some data lacked S.D., so the S.E. was assumed to be equal for each group.

Results

Study search

In total, 197 studies were retrieved, 68 of which were repeatedly published. Ultimately, 20 RCTs were selected, 9 of which were written in English (20, 25, 26, 27, 28, 29, 30, 31, 32) and 11 were in Chinese (33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43). Quality assessments of the studies are detailed in Figs. 1 and 2. Figure 3 shows the flow diagram for the study selection process. Three studies (44, 45, 46) were eliminated because they contained trials from the same period or the literature was written by the same authors, and one article (47) with an unclear indication that was a prospective research was also eliminated. In total, 1582 patients were registered, with 787 in the RA group and 795 in the FH group. The publication years were mostly in the range of 2012–2022.

![Figure 1](https://smcgrath.shinyapps.io/estmeansd/) Risk of bias summary 1.

Clinical efficacy evaluation

VAS scores of low back pain

Nine studies (20, 25, 28, 29, 35, 36, 37, 39, 41) discussed the VAS scores of low back pain, which included 490 patients, i.e., 241 in the RA group and 249 in the FH group.

![Figure 2](https://smcgrath.shinyapps.io/estmeansd/) Risk of bias summary 2.
records identified through database searching (n=137)
Additional records identified through other sources (n=0)
Records after duplicates eliminated (n=120)
Records screened (n=29)
Records excluded (n=36)
Full-text articles assessed for eligibility (n=91)
Full-text articles excluded (n=71)
*reviews/Meta-Analysis (n=32)
Different study design (n=15)
No data extractable (n=3)
Irrelevant (n=14)
Animal experiment (n=11)
Including similar subjects (n=1)
Studies included in quantitative synthesis (n=20)

Figure 3
Flow diagram of study selection.

The two groups had high levels of heterogeneity (P < 0.05, I² = 80%), and a random-effects model was used. The results of the study (Fig. 4) indicated a statistical difference between the RA group and FH group (MD = −0.44, 95% CI (−0.89, 0.02), P < 0.01). No differences were found between the subgroups based on country, robot type, and screw segment (P > 0.05) (Supplementary Fig. 1).

ODI scores
In this study, 430 patients (RA group, n = 213; FH group, 217) (20, 25, 26, 29, 35, 37, 39, 41). The two groups did not differ from one another (P > 0.05, I² = 38%), and a fixed effects model was used. According to the analysis (Fig. 5), no difference was found between the two groups (MD = −0.79, 95% CI (−1.62, 0.04), P > 0.05). The country, robot type, and screw segment did not show differences between subgroups (Supplementary Fig. 2). These results suggest that ODI scores did not correlate with country, robot type, or screw segment.

Pedicle screw placement accuracy

Grade A screw position
In total, 19 studies (20, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42) had grade A screw position. The total includes 6452 screws, of which 3200 were created with robot assistance and 3252 with manual methods. By using a random-effects model for meta-analysis, differences between the two groups that were statistically significant (P < 0.05, I² = 81%) were examined. In this study, 1.13 times as many screws in grade A position in the RA group as there were in the FH group (Fig. 6) (RR = 1.13, 95% CI (1.07, 1.19), P < 0.01). Depending on the country and screw segment, differences were found between the subgroups (P < 0.01) (Supplementary Figs. 3 and 4). Based on the robot type, no differences were found between the groupings (P > 0.05) (Supplementary Fig. 5). These results indicate that the number of grade A screw positions was related to the country and screw segment and not to the type of robot.

Grade A+B screw position
A total of 6260 screws, of which 3108 were created with robot assistance and 3152 with manual methods, were described in 18 studies (20, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42). According to the analysis (Fig. 7), the RA group had 1.04 times more screws

Table 1 Demographic characteristics of the studies included in the meta-analysis.

<table>
<thead>
<tr>
<th>Study</th>
<th>Country</th>
<th>Study type</th>
<th>Sex</th>
<th>Cases</th>
<th>Average age (years)</th>
<th>BMI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cui et al. (25)</td>
<td>China</td>
<td>RCT</td>
<td>M 10</td>
<td>RA 23</td>
<td>51.3 ± 9.8</td>
<td>RA 1.03</td>
</tr>
<tr>
<td>Fan et al. (26)</td>
<td>China</td>
<td>RCT</td>
<td>F 38</td>
<td>FH 25</td>
<td>54.1 ± 10.2</td>
<td>FH 1.07</td>
</tr>
<tr>
<td>Feng et al. (20)</td>
<td>China</td>
<td>RCT</td>
<td>RA 31</td>
<td>FH 49</td>
<td>63.45 ± 4.56</td>
<td>RA 0.05</td>
</tr>
<tr>
<td>Han et al. (27)</td>
<td>Korea</td>
<td>Korea</td>
<td>RA 121</td>
<td>RA 121</td>
<td>54.6 ± 1.13</td>
<td>FH 1.00</td>
</tr>
<tr>
<td>Hyun et al. (28)</td>
<td>Korea</td>
<td>Korea</td>
<td>RA 17</td>
<td>RA 43</td>
<td>66.5 ± 8.1</td>
<td>RA 0.00</td>
</tr>
<tr>
<td>Kim et al. (29)</td>
<td>Korea</td>
<td>Korea</td>
<td>RA 41</td>
<td>RA 17</td>
<td>65.4 ± 10.4</td>
<td>FH 1.00</td>
</tr>
<tr>
<td>Li et al. (30)</td>
<td>Korea</td>
<td>RCT</td>
<td>RA 7</td>
<td>RA 10</td>
<td>47.4 ± 12.9</td>
<td>FH 1.00</td>
</tr>
<tr>
<td>Ringel et al. (31)</td>
<td>Germany</td>
<td>RCT</td>
<td>RA 26</td>
<td>RA 34</td>
<td>68 ± 6</td>
<td>RA 0.05</td>
</tr>
<tr>
<td>Roser et al. (32)</td>
<td>Germany</td>
<td>RCT</td>
<td>RA N/A</td>
<td>RA N/A</td>
<td>68 ± 6</td>
<td>RA 0.05</td>
</tr>
<tr>
<td>Wang et al. (33)</td>
<td>China</td>
<td>RCT</td>
<td>RA 63</td>
<td>RA 19</td>
<td>42.7 ± 8.0</td>
<td>FH 1.00</td>
</tr>
<tr>
<td>Hou et al. (34)</td>
<td>China</td>
<td>RCT</td>
<td>RA 49</td>
<td>RA 13</td>
<td>43.1 ± 8.91</td>
<td>RA 0.00</td>
</tr>
<tr>
<td>Wang &amp; Kedong (35)</td>
<td>China</td>
<td>RCT</td>
<td>RA 27</td>
<td>RA 32</td>
<td>55.7 ± 2.6</td>
<td>RA 0.00</td>
</tr>
<tr>
<td>Huang et al. (36)</td>
<td>China</td>
<td>RCT</td>
<td>RA 33</td>
<td>RA 27</td>
<td>54.32 ± 6.54</td>
<td>RA 0.05</td>
</tr>
<tr>
<td>Xu et al. (37)</td>
<td>China</td>
<td>RCT</td>
<td>RA 27</td>
<td>RA 25</td>
<td>45.2 ± 9.2</td>
<td>RA 0.00</td>
</tr>
<tr>
<td>Wei et al. (38)</td>
<td>China</td>
<td>RCT</td>
<td>RA 17</td>
<td>RA 28</td>
<td>55.5 ± 10.1</td>
<td>RA 0.00</td>
</tr>
<tr>
<td>Zhai et al. (39)</td>
<td>China</td>
<td>RCT</td>
<td>RA 17</td>
<td>RA 20</td>
<td>16.7 ± 7.0</td>
<td>RA 0.00</td>
</tr>
<tr>
<td>Zhang et al. (40)</td>
<td>China</td>
<td>RCT</td>
<td>RA 47</td>
<td>RA 43</td>
<td>54.82 ± 3.52</td>
<td>RA 0.00</td>
</tr>
<tr>
<td>Yang et al. (40)</td>
<td>China</td>
<td>RCT</td>
<td>RA 47</td>
<td>RA 39</td>
<td>45.27 ± 7.71</td>
<td>RA 0.00</td>
</tr>
<tr>
<td>Cao et al. (41)</td>
<td>China</td>
<td>RCT</td>
<td>RA 19</td>
<td>RA 11</td>
<td>16.08 ± 2.31</td>
<td>RA 0.00</td>
</tr>
<tr>
<td>Li et al. (42)</td>
<td>China</td>
<td>RCT</td>
<td>RA 19</td>
<td>RA 11</td>
<td>16.08 ± 2.31</td>
<td>RA 0.00</td>
</tr>
</tbody>
</table>

FH, freehand; RA, robot-assisted; RCT, randomized controlled trial.

www.efortopenreviews.org
in grade A+B position than the FH group (RR = 1.04, 95% CI (1.02, 1.06), \( P < 0.01 \)). Differences were found between subgroups based on country, robot type, and screw segment (\( P < 0.01 \)) (Supplementary Figs. 6, 7, and 8). These results suggest that the number of grade A+B screw position correlated with the country, robot type, and screw segment.

### Safety indicators

#### Radiation dose

Five studies (27, 32, 35, 40, 43) reported the radiation dose and analyzed a total of 500 patients, with 250 in each group. As shown in Fig. 8, the radiation dose was lower in the RA group than in the FH group (SMD = −1.78, 95% CI (−3.00, −0.95), \( P < 0.01 \)).

#### Radiation exposure time

Five studies (27, 30, 31, 32, 36) reported the radiation exposure time and included a total of 399 patients, with 196 in the RA group and 203 in the FH group. The analysis showed (Fig. 9) that the duration of radiation exposure time in the RA group did not differ from that in the FH group (SMD = −0.08, 95% CI (−0.93, 0.77), \( P < 0.01 \)).

### Intraoperative blood loss

Twelve articles (20, 25, 26, 27, 30, 35, 36, 37, 39, 40, 41, 42) reported intraoperative blood loss and included a total of 840 patients, with 414 in the RA group and 426 in the FH group. Heterogeneity between the two groups was high (\( P < 0.05, I^2 = 98\% \)), and meta-analysis was performed using a random-effects model. The analysis showed (Fig. 10) that the incidence of intraoperative blood loss was less in the RA group than in the FA group (MD = −60.55, 95% CI (−112.10, −9.01), \( P < 0.01 \)).

### Operation time

Seventeen articles (20, 25, 26, 27, 28, 29, 30, 31, 34, 35, 36, 37, 38, 39, 41, 42, 43) reported on the operation time and included a total of 1142 patients, with 565 in the RA group and 577 in the FH group. Heterogeneity between the two groups was high (\( P < 0.05, I^2 = 97\% \)), and meta-analysis was performed using a random-effects model.

### Table 3 Complications reported in the studies.

<table>
<thead>
<tr>
<th>Complications</th>
<th>RA (n = 787)</th>
<th>FH (n = 795)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Screw-related</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Proximal facet violations</td>
<td>5 (0.64%)</td>
<td>33 (4.15%)</td>
</tr>
<tr>
<td>Screw repair during operation</td>
<td>16 (2.03%)</td>
<td>49 (6.16%)</td>
</tr>
<tr>
<td>Postoperative screw loosening</td>
<td>0</td>
<td>1 (0.13%)</td>
</tr>
<tr>
<td>Injury of vertebral artery</td>
<td>0</td>
<td>1 (0.13%)</td>
</tr>
<tr>
<td>Nerve root injury</td>
<td>2 (0.25%)</td>
<td>10 (1.26%)</td>
</tr>
<tr>
<td>Screw-unrelated</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Infection</td>
<td>2 (0.25%)</td>
<td>5 (0.63%)</td>
</tr>
<tr>
<td>Numbrness in the back and lower limbs</td>
<td>0 (0.13%)</td>
<td>0</td>
</tr>
<tr>
<td>Delayed wound healing</td>
<td>0</td>
<td>1 (0.13%)</td>
</tr>
<tr>
<td>Transfusion events</td>
<td>10 (1.27%)</td>
<td>14 (1.76%)</td>
</tr>
<tr>
<td>Cerebrospinal fluid leak</td>
<td>1 (0.13%)</td>
<td>3 (0.38%)</td>
</tr>
<tr>
<td>Decreased muscle strength</td>
<td>1 (0.13%)</td>
<td>1 (0.13%)</td>
</tr>
<tr>
<td>Dural injury</td>
<td>0</td>
<td>1 (0.13%)</td>
</tr>
<tr>
<td>Total</td>
<td>38 (4.83%)</td>
<td>119 (14.97%)</td>
</tr>
</tbody>
</table>

FH, freehand; RA, robot assisted.
The analysis showed (Fig. 11) a statistical difference between the RA and FH groups; however, it could not be concluded that the operating time was shorter in the RA group than in the FH group (MD = −2.15, 95% CI (−19.50, 15.20), P < 0.01).

**Hospital stay**

Nine studies (25, 26, 27, 28, 30, 31, 36, 39, 42) reported the hospital stay and included a total of 1142 patients, 565 in the RA group and 577 in the FH group. Heterogeneity between the two groups was high (P < 0.05, $I^2 = 86\%$) and meta-analysis was performed using a random-effects model. The analysis showed (Fig. 12) that the hospital stay was shorter in the RA group than in the FH group (SMD = −0.75, 95% CI (−1.40, −0.11), P < 0.01).

**Sensitivity analyses and publication bias**

VAS scores of low back pain, ODI scores, radiation dose, radiation exposure time, intraoperative blood loss, and hospital stay funnel plots suggested overall basic symmetry and further Egger’s regression (P > 0.05), indicating that the findings were robust (Supplementary Fig. 9). Funnel plots of screw position grade A suggested overall asymmetry, further Egger’s regression (P < 0.05), and Begg’s test (P > 0.05), indicating that the findings were not robust. Grade A+B funnel plot suggested overall asymmetry, followed by Egger’s regression (P < 0.05), and Begg’s test (P < 0.05) indicated that the study results were not robust. Operation time funnel plot suggested overall asymmetry, further Egger’s regression (P > 0.05), and Begg’s test (P < 0.05), indicating that the study results may have publication bias (Supplementary Fig. 10).

Heterogeneity of VAS scores of low back pain, grade A screw position, grade A+B screw position, radiation dose, radiation exposure time, operation time, and hospital stay were high. After the removal of Hyun, the results of the sensitivity analysis showed that the heterogeneity of ODI scores decreased to 0 (28), suggesting that this study may have had a greater effect on the ODI scores, and the findings did not change after removal. Further sensitivity analysis of intraoperative blood loss showed a decrease in heterogeneity after Cui removal (25), suggesting that this study may have had a greater effect on the intraoperative blood loss results; however, the conclusions were not altered after removal.
Discussion

In this meta-analysis, the study participants were from different areas across three countries. The main diagnoses were lumbar spondylolisthesis, thoracolumbar fracture, scoliosis, cervical spondylosis, and spinal stenosis, which cover almost all common spinal diseases aside from tumors. The surgical segments covered the whole spine. Compared with the published meta-analysis, this study included the largest sample size of RCTs. Articles were carefully re-selected after the initial screening to minimize selection bias, reasons that studies had been...
culled including similar studies from the same hospital and literatures with overlapping study periods. Moreover, the recent findings of relevant studies on the two latest robots were included to further enhance the reliability of the meta-analysis.

Our results contribute further evidence supporting that the accuracy of robot-assisted pedicle screw placement (grades A and A+B) was superior to the freehand group, which reflects the main advantages of RSS that the robot can work consistently and accurately and eliminate medical errors (e.g. screw misplacement) due to surgeons’ inexperience and mental fatigue. A large, multicenter retrospective analysis showed that RSS was more accurate than freehand screw placement (48). This conclusion has also been confirmed by further prospective RCTs (19, 25, 26). Some studies point out the advantages of robotic-assisted screw placement accuracy in difficult surgeries such as scoliosis (39, 41). Although RSS has higher accuracy of pedicle screw placement, we did not find this apple on translating into clinical benefits.

Clinical efficacy is one of the most important dimensions in evaluating an innovative technique and should be considered a primary outcome indicator. Our findings suggest no significant differences in VAS and ODI scores between the RSS group and the conventional surgery group. This illustrates that greater precision in screw placement is not equivalent to better patient outcomes. In clinical practice, some patients with poor screw position grades (grade C or even grade D) do not have corresponding clinical symptoms during follow-up, which is consistent with the findings of previous studies (29, 49). Remarkably, the accuracy of the freehand screw placement was greatly dependent upon the surgeon’s experience and awareness of intraoperative fluoroscopy outcome. Experienced surgeons can insert screws more efficiently and safely (50). In addition, adequate nervous decompression is the key to ensure therapeutic potency. However, the robot cannot provide substantial assistance. This would explain why more accurate screw placement did not lead to better results.

The RSS had a lower radiation dose; however, no difference was found in the radiation exposure time between the groups. Adequate preoperative planning in the RSS group can reduce the number of intraoperative radiographs, and the surgeon can insert screws step-by-step under robotic guidance without the need for additional fluoroscopy (51). A prospective RCT found that the radiation dose in the freehand group was almost four times higher than that of the RSS group (28). Some types of robots require a three-dimensional CT in the prone position during surgery, which prolongs intraoperative radiation exposure (52). However, overall, the radiation dose for robot-assisted screw placement is reduced.

The RSS group had less intraoperative blood loss and shorter hospital stays, although the operation time did not differ between the groups. Most of the RSS is minimally invasive surgery; thus, it can reduce the incision and avoid prolonged muscle distraction, which contributes...
to reduced postoperative pain and shorter hospital stays (25). The large variation in the operation time between different spinal procedures and the long preparation and commissioning time required for the robot upfront prolong the duration of robotic surgery. In the future, the operation time can still be further reduced by optimizing the preparation process and skilled use of the robot.

Perioperative adverse events of RSS are another important concern that we should pay more attention; however, not all RCTs had reported complications. Some of them were unable to confirm whether all complications were disclosed. As a result, a complication summary was only conducted using the limited data available, not a meta-analysis. Collectively, the common complications of RSS including proximal facet violations, pedicle screw misplacement or loosening, vertebral artery and nerve root injury, infection, delayed wound healing, transfusion, durotomy, and cerebrospinal fluid leak (20, 25, 26, 27, 28, 29, 30, 31, 32), and the overall complication rate of RSS was 4.83%, with a screw-related complication rate of 2.92%, which was significantly lower than that in the traditional surgery group.

In addition, RSS has some flaws and questions. (i) At this stage, most of the clinical applications of RSS are focused on pedicle screw placement. Only a small number of cases of robotic-assisted vertebroplasty and spinal endoscopic surgery have been reported. Moreover, few or no studies have focused on other critical steps of spine surgery, such as robot-assisted spinal decompression, discectomy, interbody fusion, and osteotomy. (ii) Although RSS performs excellently in screw insertion, some problems with the practical application of orthopedic robots must be considered, which consisted mostly of unstable operation, time-consuming image

![Figure 10](https://example.com/fig10.png)

**Figure 10**
Intraoperative blood loss forest plot.

![Figure 11](https://example.com/fig11.png)

**Figure 11**
Operation time forest plot.
alignment, and lateral slippage of the positioning screw (53). Robotic malfunction or operational error may cause serious complications and increase the operative duration. Page et al. found that the operative time was significantly longer in the RSS group than in the conventional surgery group (54). Such conditions could be associated with different design and manufacturing philosophy of different robot types, which require further exploration (55). (iii) The high cost of purchase and maintenance of robots is the major obstacle that limits its generalized application. Peter et al. found that the cost of a robot in lumbar fusion surgery was significantly higher than that of conventional open and minimally invasive surgery (28). Another study showed that robot-assisted minimally invasive transfemoral lumbar interbody fusion was cost-effective in 63% of simulations when patients were willing to pay $50 000/quality-adjusted life year (QALY) (56). (iv) In addition, the learning curves in RSS should not go unnoticed. RSS takes approximately 25 practice operations for a surgeon to become proficient (53). The learning curve for RSS has many variables, and similar nuances and inconsistencies in different types of robots are a relatively important factor among them (57).

This study is limited by the use of freehand screw placement, which comprised two different treatment paradigms, i.e., conventional percutaneous fluoroscopy-assisted screw placement and open surgical freehand screw placement, which has some consequences on the clinical efficacy evaluation. Their case numbers were not large enough to permit a subgroup analysis, and this may also contribute to some of the shortcomings in this study.

Conclusion
The accuracy of screw placement is the major advantage of RSS. The potential to decrease blood loss and hospital stays was noted. Moreover, the complication rates of RSS appear to be lower than conventional techniques. However, these strengths did not translate to better clinical efficacy. Therefore, how to broaden the application of RSS reasonably needs further research for greater clinical effectiveness, which included but is not limited to robot-assisted spinal decompression and tumors resection. More attention is needed to the efficacy, complications, and cost–effectiveness ratio. Further integration is still needed to enable automatic registration and automatic planning of surgical paths and improve the automation and intelligence of orthopedic surgical robots. Large, multicenter, high-quality, RCTs with uniform metrics are still needed to provide more robust evidence.

Supplementary materials
This is linked to the online version of the paper at https://doi.org/10.1530/EOR-23-0125.

ICMJE Conflict of interest Statement
The authors declare that there is no conflict of interest that could be perceived as prejudicing the impartiality of the research reported.

Funding Statement
This work was partially supported by a grant from National Natural Science Foundation of China (No. 82004385 to YPL), Excellent Talents Program of Guangdong Provincial Hospital of Chinese Medicine (No. 212022ZY17 for LYP), The State Key Laboratories Development Program of Guangzhou (No.202102010012 to CBL), The Innovation Team Project of Guangdong Provincial Department of Education (No. 2021KXTD020 to CBL), and Administration of Traditional Chinese Medicine of Guangdong Province (o.20190407090652 to DJ).

Author contribution statement
W-x Sun: investigation, data curation, formal analysis, writing the original draft; W-q Huang: investigation, data curation, formal analysis; H-y Li: investigation, data curation; H-s Wang: methodology, reviewing and editing; S-l Guo: analysis and interpretation of data and supervision; J Dong: methodology; B-l Chen: conceptualization, methodology, writing, reviewing and editing, and supervision; Y-p Lin: conceptualization, methodology, writing, reviewing, and editing, supervision.
Acknowledgements
We appreciate the team at the Guangzhou Key Laboratory of Integrated Traditional Chinese and Western Medicine for Prevention and Treatment of Lumbar Degenerative Diseases.

References


