BJR

**SPINE**

# The safety and accuracy of robot-assisted pedicle screw internal fixation for spine disease

## A META-ANALYSIS

**W. Li,**
**G. Li,**
**W. Chen,**
**L. Cong**

*From China Medical University First Hospital, Shenyang, China*

## Aims

The aim of this study was to systematically compare the safety and accuracy of robot-assisted (RA) technique with conventional freehand with/without fluoroscopy-assisted (CT) pedicle screw insertion for spine disease.

## Methods

A systematic search was performed on PubMed, EMBASE, the Cochrane Library, MEDLINE, China National Knowledge Infrastructure (CNKI), and WANFANG for randomized controlled trials (RCTs) that investigated the safety and accuracy of RA compared with conventional freehand with/without fluoroscopy-assisted pedicle screw insertion for spine disease from 2012 to 2019. This meta-analysis used Mantel-Haenszel or inverse variance method with mixed-effects model for heterogeneity, calculating the odds ratio (OR), mean difference (MD), standardized mean difference (SMD), and 95% confidence intervals (CIs). The results of heterogeneity, subgroup analysis, and risk of bias were analyzed.

## Results

Ten RCTs with 713 patients and 3,331 pedicle screws were included. Compared with CT, the accuracy rate of RA was superior in Grade A with statistical significance and Grade A + B without statistical significance. Compared with CT, the operating time of RA was longer. The difference between RA and CT was statistically significant in radiation dose. Proximal facet joint violation occurred less in RA than in CT. The postoperative Oswestry Disability Index (ODI) of RA was smaller than that of CT, and there were some interesting outcomes in our subgroup analysis.

## Conclusion

RA technique could be viewed as an accurate and safe pedicle screw implantation method compared to CT. A robotic system equipped with optical intraoperative navigation is superior to CT in accuracy. RA pedicle screw insertion can improve accuracy and maintain stability for some challenging areas.

Cite this article: *Bone Joint Res* 2020;9(10):653–666.

**Keywords:** Meta-analysis, Randomized controlled trial, Robot-assisted, Freehand, Pedicle screw insertion.

Correspondence should be sent to Lin Cong; email: chinaconglin@outlook.com; chinaconglin@163.com

doi: 10.1302/2046-3758.910.BJR-2020-0064.R2

*Bone Joint Res* 2020;9(10):653–666.

## Article focus

- Robot-assisted (RA) surgical technique is applied increasingly in the field of spine surgery, but the evidence for its effectiveness is not solid enough.
- The aim of this study was to systematically compare the safety and accuracy of RA technique with conventional freehand with/without fluoroscopy-assisted (CT) pedicle screw insertion for spine disease.

## Key messages

- The RA technique could be viewed as an accurate and safe pedicle screw implantation method compared to CT.

- A robotic system equipped with optical intraoperative navigation is superior to CT in accuracy.
- RA pedicle screw insertion can improve accuracy and maintain stability for some challenging areas.

## Strengths and limitations

- Ten randomized controlled trials (RCTs) with similar inclusion and exclusion criteria and scrupulous patient selection were collected on the same topic of research without language restrictions.
- Subgroup analyses comparing different types of robots and different operative levels were performed.
- All of the included RCTs failed to report the comparison of costs and reimbursement.

## Introduction

Pedicle screw insertion is one of the most commonly used methods in spinal surgery. Due to the complex anatomical structure of the spine, the instability of human hands, and the long operating time,[1] the difficulty and risk of conventional spine surgery remain high.

In recent decades, robotic surgery has been a flourishing technique in the field of surgery. Many kinds of robots have been developed, such as Mazor robotics (SpineAssist robot and Renaissance), Robotic Surgical Assistant (ROSA), and 'TiRobot' orthopaedic robot (Tinavi Robot). During the development of spinal surgery robots, many studies have been published evaluating the accuracy and safety of spine robotic screw insertion surgery. Some studies supported robots as more accurate and safer technologies. Khan et al[2] and Stull et al[3] concluded that robotic spine surgery can be associated with a safer and more accurate surgical procedure. Staartjes et al[4] concluded that postoperative revisions for screw malposition occurred less in robotic spine surgery, but there was still other research suggesting that more studies were needed to determine which was better: robot-assisted (RA) or conventional freehand with/without fluoroscopy-assisted technique (CT). Ghasem et al[5] and Joseph et al[6] suggested that radiation exposure, length of stay, operating time, and real-time cost-efficacy of robotic spine surgery remained unclear. Other studies suggested that application range, application depth, and real-time cost-efficacy of robotic surgery technique should be further studied and analyzed.[7-9] Additionally, after scrupulous deliberation we found that some published meta-analyses have their own shortcomings. The latest meta-analysis comparing RA technique and freehand pedicle screw implantation purely based on randomized controlled trials (RCTs) was performed by Li et al.[10] We hold the opinion that the results of radiation time and dose in their outcomes were developed from inappropriate data and comparative method.

Considering this, the debate between RA and CT techniques still continues. We performed this meta-analysis to systematically compare safety and accuracy of RA with CT pedicle screw insertion for spine disease. In this meta-analysis, we compared the data more rigorously based on more RCTs and applied subgroup analyses comparing different types of robots and different operative levels.

## Methods

**Search methods and eligibility criteria.** RCTs without language restrictions were identified. A systematic search was performed on PubMed, EMBASE, the Cochrane Library, MEDLINE, China National Knowledge Infrastructure (CNKI), and WANFANG for RCTs, which investigated RA with CT pedicle screw insertion for spine disease from 2012 to 2019. The following keywords were applied to the search: "robot-assisted"; "conventional freehand"; "pedicle screw insertion"; "spine"; and "randomized controlled trial".

Published trials were systematically inspected according to the following criteria: recruiting adult patients who underwent RA with CT pedicle screw insertion for spine disease; reporting at least one outcome of interest; and a postoperative follow-up period of at least one year. Trials were excluded if interventions were different from the previous description, or if original data were lost after confirmation with the corresponding author.

**Data extraction and statistical analyses.** Two researchers (WL and GL) independently extracted the data, including the information of trials, inclusion criteria, participant characteristics, outcomes of interest, and duration of follow-up period. The outcomes in this meta-analysis include: the accuracy of screw insertion evaluated by Gertzbein–Robbins Grade A criteria; the accuracy of screw insertion evaluated by Gertzbein–Robbins Grade A + B criteria; operating time; radiation dose; radiation time; proximal facet joint violation; intraoperative blood loss; postoperative visual analogue scale (VAS)–back and postoperative VAS–leg; postoperative stay; and postoperative Oswestry Disability Index (ODI). The outcomes in subgroup analysis include: the screw insertion accuracy of the SpineAssist-assisted (SRA), Renaissance-assisted (RRA), and Tinavi robot-assisted (TRA) groups evaluated by Gertzbein–Robbins Grade A criteria; the screw insertion accuracy of the SRA, RRA, and TRA groups evaluated by Gertzbein–Robbins Grade A + B criteria; operating time of the TRA and RRA groups; the screw insertion accuracy of operative thoracolumbar level and operative lumbar level evaluated by Gertzbein–Robbins Grade A criteria; and the screw insertion accuracy of operative thoracolumbar level and operative lumbar level evaluated by Gertzbein–Robbins Grade A + B criteria.

The continuous outcomes are presented as mean difference (MD) and 95% confidence interval (CI), and odds ratio (OR) and 95% CIs are presented for dichotomous outcomes. Standardized mean difference (SMD) is presented for the same type of continuous outcomes with different units. The chi-squared test and $I^2$ during each analysis were utilized and evaluated for heterogeneity.

If the p-value was 0.05 or less, statistical heterogeneity existed. In this situation, a random-effects model was utilized. We used RevMan software (v. 5.3; Copenhagen: The Nordic Cochrane Centre, The Cochrane Collaboration, London, UK) to perform all analyses. Statistical significance was considered when $p < 0.05$.

**Search methods and data collection.** The data were abstracted from included RCTs, and included the following classifications: attributions of participants, the number of participants, and the loss to follow-up; study attributions; interventions; outcomes and subgroup analysis outcomes; ORs, MD, and SMD; and 95% CI.

**Assessment of risk of bias.** The risk of bias in the included RCTs was evaluated by The Cochrane Collaboration's tool. The classifications of bias were based on 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. The results were compared afterwards.

## Results

**RCT selection and characteristics.** A total of 161 RCTs with potential eligibility were retrieved. Of these, 18 RCTs were removed for duplication, and 123 RCTs were excluded on the titles/abstracts. Of the remaining 20 RCTs , ten were excluded after full-text screening due to ineligible comparison, data duplication, and the absence of relevant reported outcomes. At the final stage, ten RCTs that met the eligibility criteria were included in this meta-analysis (Figure 1).

The characteristics of the included trials are shown in Table I. In all included RCTs, the sample size of patients who needed pedicle screw insertion were mustered ranging from 40 to 234 patients. Explicit inclusion/exclusion criteria were exhibited in all of the included RCTs. Four studies investigated the safety and effectiveness of TRA minimally invasive pedicle screw insertion compared with fluoroscopy-assisted pedicle screw insertion.[11-14] In addition, the Mazor robot-assisted (MRA) minimally invasive pedicle screw insertion was employed compared with conventional open freehand pedicle screw insertion in three studies,[15-17] and fluoroscopy-assisted pedicle screw insertion in three other studies.[18-20] The operative data, clinical outcomes, and postoperative recovery were anatomized in at least one year of follow-up.

**Risk of bias.** The risk of bias in the included trials was evaluated by The Cochrane Collaboration's tool. All the trials affirmed randomization, but three trials did not describe the approach of random sequence generations.[11,19,20] Six studies[12–14,16–18] reported allocation concealment. All of these studies included trials that failed to report the blinding of personnel and participants, which were evaluated to high risk of bias. Seven trials[11–14,16,17,19] reported the blinding of outcome assessment. Incomplete outcome data (attrition bias) and selective reporting (reporting bias) were evaluated to high risk in the trial by Roser et al.[20] The risk of bias detail for included trials is presented in Figure 2.

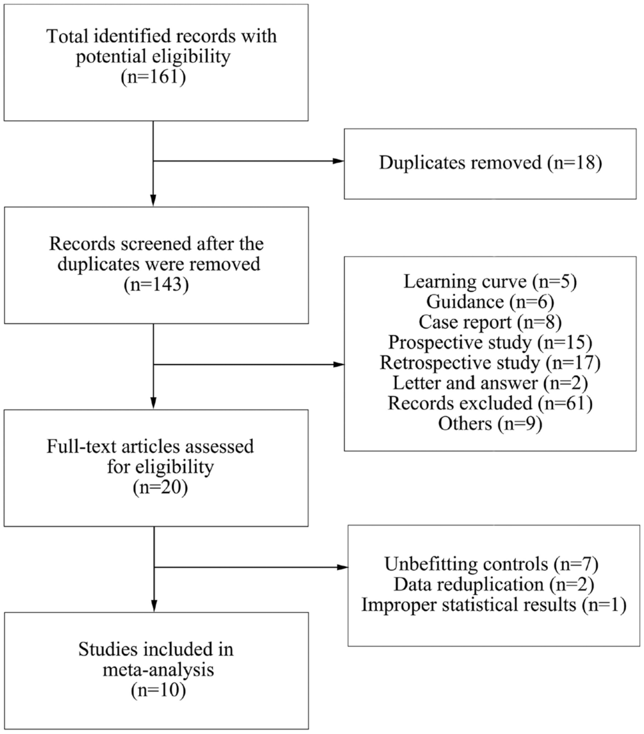


Fig. 1

Flowchart showing the selection process of randomized controlled trials (RCTs) for meta-analysis.

**Meta-analysis results.** Compared with CT, the accuracy rate of RA was superior according to the Gertzbein–Robbins Grade A (RA group: 90.73%; CT group: 84.37%; OR: 2.45; 95% CI: 1.29 to 4.65; p = 0.006, Z test; $I^2$ = 77%). The accuracy rate of RA was better than CT without statistical significance (RA group: 95.50%; CT group: 92.69%; OR: 2.61; 95% CI: 0.95 to 7.23; p = 0.060, Z test; $I^2$ = 64%) (Figure 3). The difference in radiation dose between RA and CT was statistically significant (SMD: -1.30; 95% CI: -2.01 to -0.60; $p < 0.001$, Z test; $I^2$ = 97%). Radiation time was not statistically significant between RA and CT (MD: -2.80; 95% CI: -27.81 to 22.21; p = 0.830, Z test; $I^2$ = 98%). The operating time of RA was longer than that of CT (MD: 9.11; 95% CI: 3.69 to 14.53; p = 0.001, Z test; $I^2$ = 39%). Proximal facet joint violation of RA occurred less than for CT (OR: 0.05; 95% CI: 0.01 to 0.28; $p < 0.001$, Z test; $I^2$ = 0%). No significant differences were found between RA and CT in intraoperative blood loss (MD: -51.47; 95% CI: -112.51 to 9.57; p = 0.100, Z test; $I^2$ = 90%), postoperative VAS-back (MD: -0.15; 95% CI: -0.34 to 0.04; p = 0.120, Z test; $I^2$ = 36%), postoperative VAS-leg (MD: -0.06, 95% CI: -0.20 to 0.09; p = 0.430, Z test; $I^2$ = 0%), and postoperative stay (MD: -0.36; 95% CI: -1.03 to 0.31; p = 0.300, Z test; $I^2$ = 62%). The postoperative ODI of RA was smaller than that of CT

**Table I.** Characteristics of the included trials.

| Trial | Intervention (Exp/Clt) | Sample size (Exp/Clt), n | Mean age, yrs (SD) | | | Sex(female/male), n | Operative level | Outcome |
|---|---|---|---|---|---|---|---|---|
| | | | Overall | Exp | Clt | | | |
| Feng 2019[11] | Tinavi robot-assisted pedicle screw insertion/fluoroscopy-assisted pedicle screw insertion. | 80 (40/40) | 67.7 (6.9) | 67.6 (6.5) | 67.9 (7.3) | 55/25 | L2 to L5 | ASP, operative time, radiation dose, blood loss, postoperative stay |
| Fu 2017[15] | Renaissance-assisted pedicle screw insertion/conventional freehand pedicle screw insertion. | 50 (24/26) | 68.0 (4.5) | 67.0 (4.0) | 69.0 (5.0) | 27/23 | L3 to S1 | Operative time, VAS, ODI, blood loss |
| Han 2019[12] | Tinavi robot-assisted pedicle screw insertion/fluoroscopy-assisted pedicle screw insertion. | 234 (115/119) | 55.4 (12.4) | 54.6 (11.3) | 56.1 (13.4) | 121/133 | Thoracolumbar spine | ASP, operative time, radiation dose and time, blood loss, postoperative stay, proximal facet joint violation |
| Hyun 2017[18] | Renaissance-assisted pedicle screw insertion/fluoroscopy-assisted pedicle screw insertion. | 60 (30/30) | 66.7 (8.5) | 66.5 (8.1) | 66.8 (8.9) | 43/17 | Lumbar spine | ASP, operative time, radiation dose, VAS, ODI, postoperative stay, proximal facet joint violation |
| Kim 2015[17] | Renaissance-assisted pedicle screw insertion/conventional freehand pedicle screw insertion. | 40 (20/20) | 64.6 (10.3) | 64.4 (11.9) | 64.7 (8.6) | 21/19 | L2 to S1 | ASP, operative time |
| Roser 2013[20] | SpineAssist-assisted pedicle screw insertion/fluoroscopy-assisted pedicle screw insertion. | 28 (18/10) | N/A | N/A | N/A | N/A | T1 to S1 | ASP, radiation dose, radiation time |
| Ringel 2012[19] | SpineAssist-assisted pedicle screw insertion/fluoroscopy-assisted pedicle screw insertion. | 60 (30/30) | 67.5 (N/A) | 68.0 (N/A) | 67.0 (N/A) | 34/26 | L2 to S1 | ASP, radiation time, postoperative stay |
| Kim 2018[16] | Renaissance-assisted pedicle screw insertion/conventional freehand pedicle screw insertion. | 78 (37/41) | 65.7 (9.5) | 65.4 (10.4) | 66.0 (8.6) | 37/41 | L2 to S1 | ASP, operative time, VAS, ODI, proximal facet joint violation |
| Tian 2016[13] | Tinavi robot-assisted pedicle screw insertion/fluoroscopy-assisted pedicle screw insertion. | 40 (23/17) | 54.0 (11.9) | 55.5 (10.1) | 52.0 (14.1) | 23/17 | Lumbar spine | ASP, operative time |
| Xu 2018[14] | Tinavi robot-assisted pedicle screw insertion/fluoroscopy-assisted pedicle screw insertion. | 43 (24/19) | 46.4 (10.5) | 45.2 (9.2) | 50.6 (10.7) | 18/25 | Thoracolumbar spine | ASP, operative time, VAS, ODI, blood loss |

ASP, accuracy of screw placement; Ctl, control group; Exp, experimental group; N/A, not available; ODI, Oswestry Disability Index; VAS, visual analogue scale.

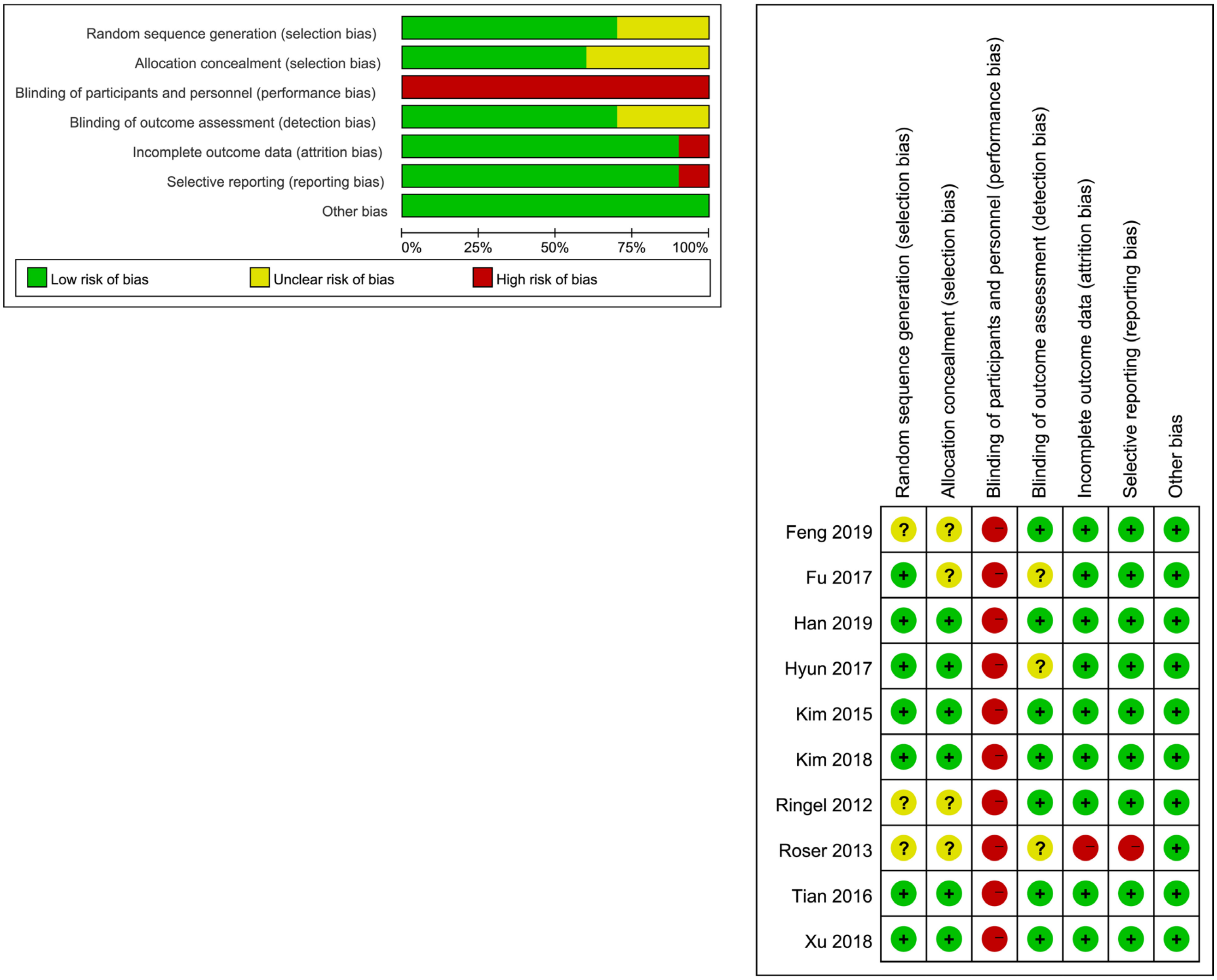
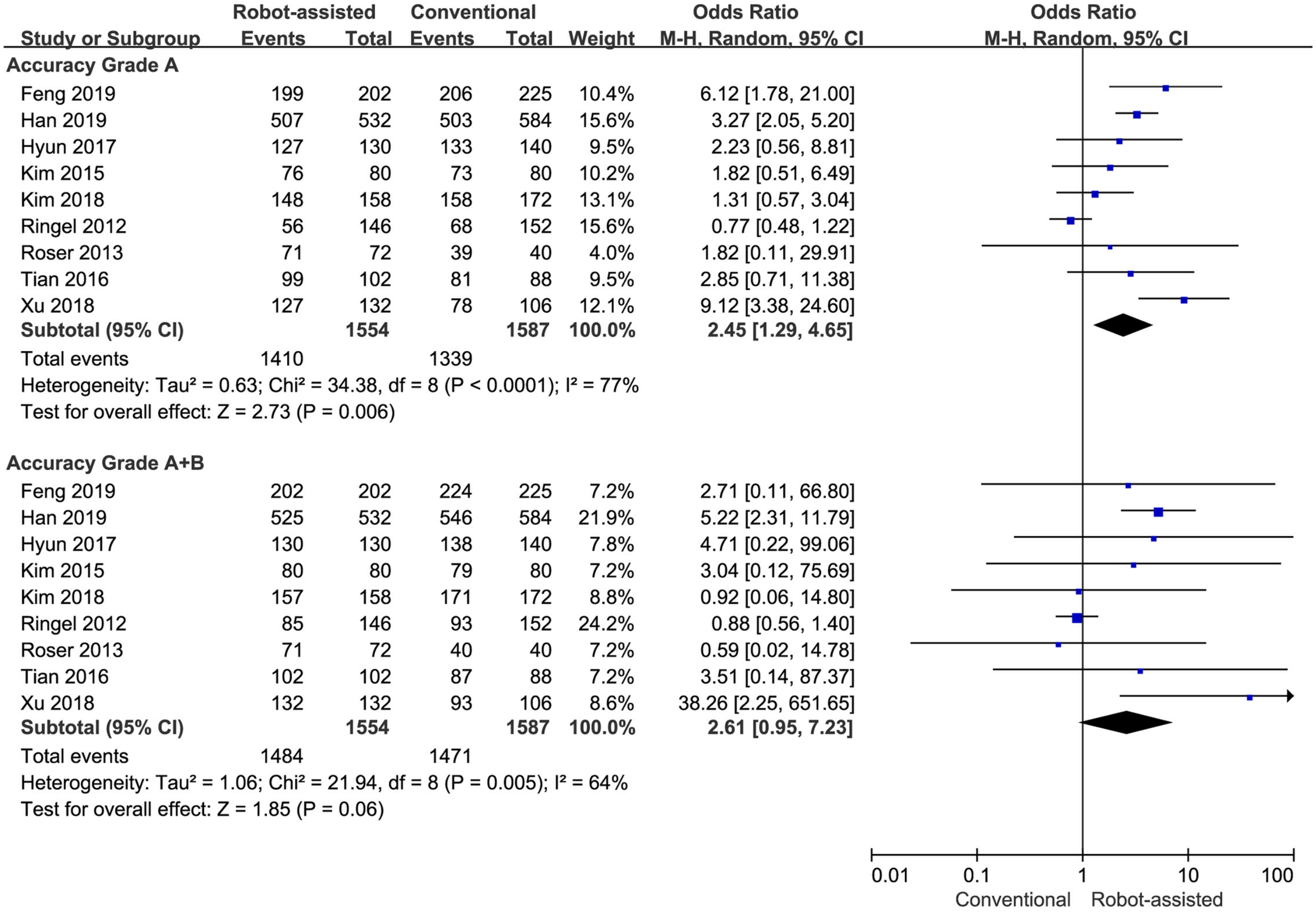


Fig. 2

Risk of bias summary showing details of the risk of bias in each of the included trials. Red colour indicates high risk, yellow indicates unclear risk, and green colour indicates low risk.

(MD: -2.22; 95% CI: -3.83 to -0.61; $p = 0.007$, Z test; $I^2 = 3\%$) (Figure 4).

**Subgroup analysis results.** In this meta-analysis, subgroup analyses based on different types and generations of robots (SRA vs freehand, RRA vs freehand, and TRA vs freehand) were performed. In subgroup analyses, the accuracy rate of SRA was better than CT without statistical significance in the Grade A (SRA group: 58.26%; CT group: 55.73%; OR: 0.79; 95% CI: 0.50 to 1.24; $p = 0.300$, Z test; $I^2 = 0\%$) and Grade A + B criteria (SRA group: 71.56%; CT group: 69.27%; OR: 0.88; 95% CI: 0.55 to 1.38; $p = 0.570$, Z test; $I^2 = 0\%$). The accuracy rate of RRA was better than CT without statistical significance in the Grade A (RRA group: 95.38%; CT group: 92.86%; OR: 1.59; 95% CI: 0.86 to 2.97; $p = 0.140$, Z test; $I^2 = 0\%$) and Grade A + B criteria (RRA group: 99.73%; CT group: 98.98%; OR: 2.38; 95% CI: 0.46 to 12.30; $p = 0.300$, Z test; $I^2 = 0\%$). The accuracy rate of the TRA group was better than CT with statistical significance in Grade A (TRA group: 96.28%; CT group: 86.54%; OR: 4.11; 95% CI: 2.81 to 6.00; $p < 0.001$, Z test; $I^2 = 27\%$) and Grade A + B criteria (TRA group: 99.28%; CT group: 94.72%; OR: 6.53; 95% CI: 3.17 to 13.45; $p < 0.001$, Z test; $I^2 = 0\%$) (Figure 5). The operating time of the TRA group was longer than CT (MD: 8.83; 95% CI: 0.45 to 17.20; $p = 0.040$, Z test; $I^2 = 35\%$), while the difference between the operating times of the RRA and CT groups was not statistically significant (MD: 12.29; 95% CI: -5.04 to 29.62; $p = 0.160$, Z test; $I^2 = 57\%$) (Figure 6). Subgroup analysis based on operative level (operative thoracolumbar level and operative lumbar level) was also performed. Compared with CT, the accuracy rate of RA was superior according to the Grade A (RA group: 95.79%; CT group: 84.93%; OR: 4.45; 95% CI: 2.00 to 9.93; $p < 0.001$, Z test; $I^2 = 45\%$) and Grade A + B criteria (RA group: 98.91%; CT group: 93.01%; OR: 5.41; 95% CI: 0.97 to 30.05; $p = 0.050$, Z test; $I^2 = 47\%$)



| Study or Subgroup | Robot-assisted Events | Robot-assisted Total | Conventional Events | Conventional Total | Weight | Odds Ratio M-H, Random, 95% CI |
|---|---|---|---|---|---|---|
| **Accuracy Grade A** | | | | | | |
| Feng 2019 | 199 | 202 | 206 | 225 | 10.4% | 6.12 [1.78, 21.00] |
| Han 2019 | 507 | 532 | 503 | 584 | 15.6% | 3.27 [2.05, 5.20] |
| Hyun 2017 | 127 | 130 | 133 | 140 | 9.5% | 2.23 [0.56, 8.81] |
| Kim 2015 | 76 | 80 | 73 | 80 | 10.2% | 1.82 [0.51, 6.49] |
| Kim 2018 | 148 | 158 | 158 | 172 | 13.1% | 1.31 [0.57, 3.04] |
| Ringel 2012 | 56 | 146 | 68 | 152 | 15.6% | 0.77 [0.48, 1.22] |
| Roser 2013 | 71 | 72 | 39 | 40 | 4.0% | 1.82 [0.11, 29.91] |
| Tian 2016 | 99 | 102 | 81 | 88 | 9.5% | 2.85 [0.71, 11.38] |
| Xu 2018 | 127 | 132 | 78 | 106 | 12.1% | 9.12 [3.38, 24.60] |
| **Subtotal (95% CI)** | | **1554** | | **1587** | **100.0%** | **2.45 [1.29, 4.65]** |
| Total events | 1410 | | 1339 | | | |
| Heterogeneity: Tau² = 0.63; Chi² = 34.38, df = 8 (P < 0.0001); I² = 77% | | | | | | |
| Test for overall effect: Z = 2.73 (P = 0.006) | | | | | | |
| **Accuracy Grade A+B** | | | | | | |
| Feng 2019 | 202 | 202 | 224 | 225 | 7.2% | 2.71 [0.11, 66.80] |
| Han 2019 | 525 | 532 | 546 | 584 | 21.9% | 5.22 [2.31, 11.79] |
| Hyun 2017 | 130 | 130 | 138 | 140 | 7.8% | 4.71 [0.22, 99.06] |
| Kim 2015 | 80 | 80 | 79 | 80 | 7.2% | 3.04 [0.12, 75.69] |
| Kim 2018 | 157 | 158 | 171 | 172 | 8.8% | 0.92 [0.06, 14.80] |
| Ringel 2012 | 85 | 146 | 93 | 152 | 24.2% | 0.88 [0.56, 1.40] |
| Roser 2013 | 71 | 72 | 40 | 40 | 7.2% | 0.59 [0.02, 14.78] |
| Tian 2016 | 102 | 102 | 87 | 88 | 7.2% | 3.51 [0.14, 87.37] |
| Xu 2018 | 132 | 132 | 93 | 106 | 8.6% | 38.26 [2.25, 651.65] |
| **Subtotal (95% CI)** | | **1554** | | **1587** | **100.0%** | **2.61 [0.95, 7.23]** |
| Total events | 1484 | | 1471 | | | |
| Heterogeneity: Tau² = 1.06; Chi² = 21.94, df = 8 (P = 0.005); I² = 64% | | | | | | |
| Test for overall effect: Z = 1.85 (P = 0.06) | | | | | | |

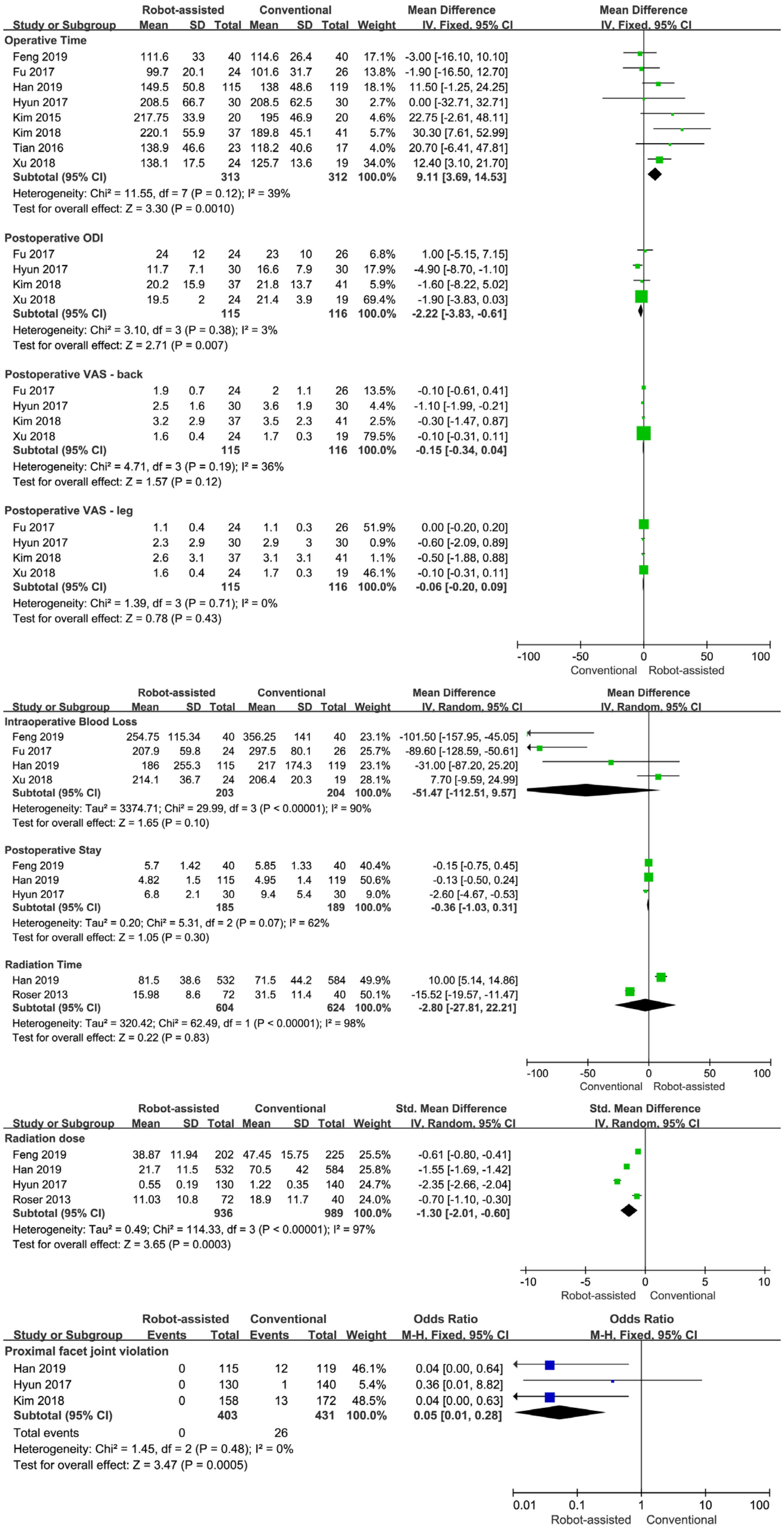
Fig. 3

Pooled results of the robot-assisted group and the conventional freehand with/without fluoroscopy-assisted group. The results are shown as follows: Grade A accuracy and Grade A + B accuracy. CI, confidence interval; M-H, Mantel–Haenszel.

when the operative level was the thoracolumbar level. However, the difference in accuracy rates of the Grade A (RA group: 86.19%; CT group: 83.90%; OR: 1.77; 95% CI: 0.91 to 3.45; p = 0.090, Z test; $I^2 = 62\%$) and Grade A + B criteria (RA group: 92.42%; CT group: 92.42%; OR: 0.98; 95% CI: 0.63 to 1.52; p = 0.940, Z test; $I^2 = 0\%$) was not statistically significant between RA and CT when the operative level was the lumbar level. (Figure 7)

## Discussion

This meta-analysis included ten RCTs that compared RA with CT pedicle screw insertion for spine disease, including 713 patients (361 in RA group and 352 in CT group) and 3,331 pedicle screws (1,656 in RA group and 1,675 in CT group). It revealed that RA technique had a better accuracy rate than CT group utilizing Gertzbein–Robbins Grade A criteria with statistical significance and Gertzbein–Robbins Grade A + B criteria. We also found some interesting outcomes in the subgroup analyses; there were no significant differences between the MRA group and CT group in the Grade A or Grade A + B criteria accuracy rate, meanwhile the TRA group was still superior to the CT group in accuracy rate. The reasons for those outcomes may be the differences of preoperative preparation and image-forming method between two kinds of robots, position changing of the patient, and real-time navigation monitoring. The preoperative preparation and image-forming method of Mazor robot was a CT scan of the object region, which was uploaded to the proprietary software before surgery, and the optimal screw dimensions and positioning relative to the patient's anatomy were established by the preoperative CT scan in 3D views. However, the position of the patient was supine during the preoperative CT scan and the patient was prone during the operation, which may cause the planned trajectory - according to the preoperative CT - to not reach 100% fit to the actuality. If there is an error between the 3D reconstruction image and the actual anatomical position due to the change of body position, the screw placement may fail. As for the Tinavi Robot system equipped with optical intraoperative navigation, surgeons could drill through the guiding tube into the pedicle either percutaneously or via an open approach under real-time navigation monitoring. In addition, as the Tinavi Robot system has the function of intraoperative fine-tuning, if the screw position verified by fluoroscopy



| | Robot-assisted | | | Conventional | | | | Mean Difference |
|---|---|---|---|---|---|---|---|---|
| **Study or Subgroup** | **Mean** | **SD** | **Total** | **Mean** | **SD** | **Total** | **Weight** | **IV, Fixed, 95% CI** |
| **Operative Time** | | | | | | | | |
| Feng 2019 | 111.6 | 33 | 40 | 114.6 | 26.4 | 40 | 17.1% | -3.00 [-16.10, 10.10] |
| Fu 2017 | 99.7 | 20.1 | 24 | 101.6 | 31.7 | 26 | 13.8% | -1.90 [-16.50, 12.70] |
| Han 2019 | 149.5 | 50.8 | 115 | 138 | 48.6 | 119 | 18.1% | 11.50 [-1.25, 24.25] |
| Hyun 2017 | 208.5 | 66.7 | 30 | 208.5 | 62.5 | 30 | 2.7% | 0.00 [-32.71, 32.71] |
| Kim 2015 | 217.75 | 33.9 | 20 | 195 | 46.9 | 20 | 4.6% | 22.75 [-2.61, 48.11] |
| Kim 2018 | 220.1 | 55.9 | 37 | 189.8 | 45.1 | 41 | 5.7% | 30.30 [7.61, 52.99] |
| Tian 2016 | 138.9 | 46.6 | 23 | 118.2 | 40.6 | 17 | 4.0% | 20.70 [-6.41, 47.81] |
| Xu 2018 | 138.1 | 17.5 | 24 | 125.7 | 13.6 | 19 | 34.0% | 12.40 [3.10, 21.70] |
| **Subtotal (95% CI)** | | | **313** | | | **312** | **100.0%** | **9.11 [3.69, 14.53]** |
| Heterogeneity: Chi$^2$ = 11.55, df = 7 (P = 0.12); I$^2$ = 39% | | | | | | | | |
| Test for overall effect: Z = 3.30 (P = 0.0010) | | | | | | | | |
| **Postoperative ODI** | | | | | | | | |
| Fu 2017 | 24 | 12 | 24 | 23 | 10 | 26 | 6.8% | 1.00 [-5.15, 7.15] |
| Hyun 2017 | 11.7 | 7.1 | 30 | 16.6 | 7.9 | 30 | 17.9% | -4.90 [-8.70, -1.10] |
| Kim 2018 | 20.2 | 15.9 | 37 | 21.8 | 13.7 | 41 | 5.9% | -1.60 [-8.22, 5.02] |
| Xu 2018 | 19.5 | 2 | 24 | 21.4 | 3.9 | 19 | 69.4% | -1.90 [-3.83, 0.03] |
| **Subtotal (95% CI)** | | | **115** | | | **116** | **100.0%** | **-2.22 [-3.83, -0.61]** |
| Heterogeneity: Chi$^2$ = 3.10, df = 3 (P = 0.38); I$^2$ = 3% | | | | | | | | |
| Test for overall effect: Z = 2.71 (P = 0.007) | | | | | | | | |
| **Postoperative VAS - back** | | | | | | | | |
| Fu 2017 | 1.9 | 0.7 | 24 | 2 | 1.1 | 26 | 13.5% | -0.10 [-0.61, 0.41] |
| Hyun 2017 | 2.5 | 1.6 | 30 | 3.6 | 1.9 | 30 | 4.4% | -1.10 [-1.99, -0.21] |
| Kim 2018 | 3.2 | 2.9 | 37 | 3.5 | 2.3 | 41 | 2.5% | -0.30 [-1.47, 0.87] |
| Xu 2018 | 1.6 | 0.4 | 24 | 1.7 | 0.3 | 19 | 79.5% | -0.10 [-0.31, 0.11] |
| **Subtotal (95% CI)** | | | **115** | | | **116** | **100.0%** | **-0.15 [-0.34, 0.04]** |
| Heterogeneity: Chi$^2$ = 4.71, df = 3 (P = 0.19); I$^2$ = 36% | | | | | | | | |
| Test for overall effect: Z = 1.57 (P = 0.12) | | | | | | | | |
| **Postoperative VAS - leg** | | | | | | | | |
| Fu 2017 | 1.1 | 0.4 | 24 | 1.1 | 0.3 | 26 | 51.9% | 0.00 [-0.20, 0.20] |
| Hyun 2017 | 2.3 | 2.9 | 30 | 2.9 | 3 | 30 | 0.9% | -0.60 [-2.09, 0.89] |
| Kim 2018 | 2.6 | 3.1 | 37 | 3.1 | 3.1 | 41 | 1.1% | -0.50 [-1.88, 0.88] |
| Xu 2018 | 1.6 | 0.4 | 24 | 1.7 | 0.3 | 19 | 46.1% | -0.10 [-0.31, 0.11] |
| **Subtotal (95% CI)** | | | **115** | | | **116** | **100.0%** | **-0.06 [-0.20, 0.09]** |
| Heterogeneity: Chi$^2$ = 1.39, df = 3 (P = 0.71); I$^2$ = 0% | | | | | | | | |
| Test for overall effect: Z = 0.78 (P = 0.43) | | | | | | | | |

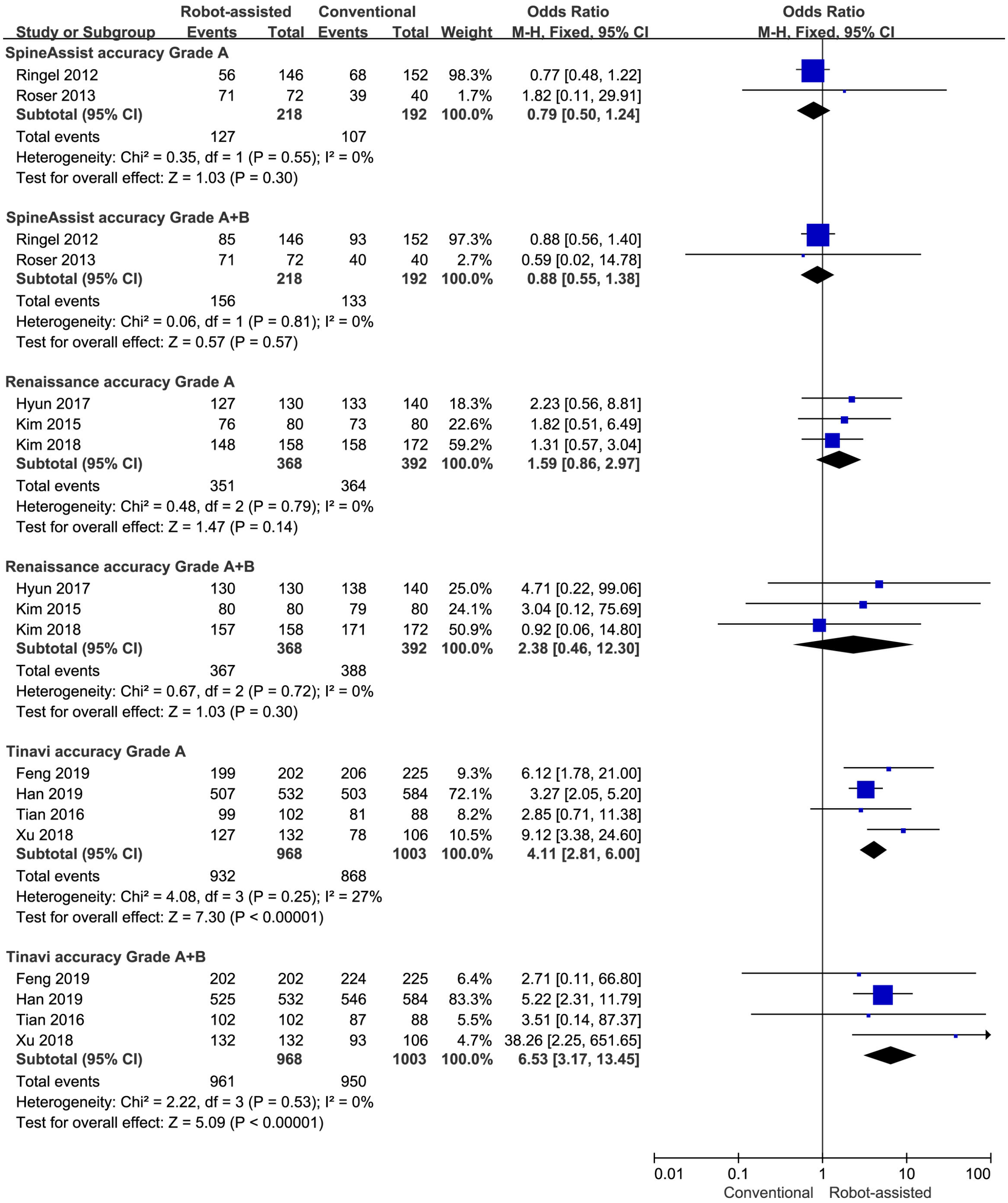
| | Robot-assisted | | | Conventional | | | | Mean Difference |
|---|---|---|---|---|---|---|---|---|
| **Study or Subgroup** | **Mean** | **SD** | **Total** | **Mean** | **SD** | **Total** | **Weight** | **IV, Random, 95% CI** |
| **Intraoperative Blood Loss** | | | | | | | | |
| Feng 2019 | 254.75 | 115.34 | 40 | 356.25 | 141 | 40 | 23.1% | -101.50 [-157.95, -45.05] |
| Fu 2017 | 207.9 | 59.8 | 24 | 297.5 | 80.1 | 26 | 25.7% | -89.60 [-128.59, -50.61] |
| Han 2019 | 186 | 255.3 | 115 | 217 | 174.3 | 119 | 23.1% | -31.00 [-87.20, 25.20] |
| Xu 2018 | 214.1 | 36.7 | 24 | 206.4 | 20.3 | 19 | 28.1% | 7.70 [-9.59, 24.99] |
| **Subtotal (95% CI)** | | | **203** | | | **204** | **100.0%** | **-51.47 [-112.51, 9.57]** |
| Heterogeneity: Tau$^2$ = 3374.71; Chi$^2$ = 29.99, df = 3 (P < 0.00001); I$^2$ = 90% | | | | | | | | |
| Test for overall effect: Z = 1.65 (P = 0.10) | | | | | | | | |
| **Postoperative Stay** | | | | | | | | |
| Feng 2019 | 5.7 | 1.42 | 40 | 5.85 | 1.33 | 40 | 40.4% | -0.15 [-0.75, 0.45] |
| Han 2019 | 4.82 | 1.5 | 115 | 4.95 | 1.4 | 119 | 50.6% | -0.13 [-0.50, 0.24] |
| Hyun 2017 | 6.8 | 2.1 | 30 | 9.4 | 5.4 | 30 | 9.0% | -2.60 [-4.67, -0.53] |
| **Subtotal (95% CI)** | | | **185** | | | **189** | **100.0%** | **-0.36 [-1.03, 0.31]** |
| Heterogeneity: Tau$^2$ = 0.20; Chi$^2$ = 5.31, df = 2 (P = 0.07); I$^2$ = 62% | | | | | | | | |
| Test for overall effect: Z = 1.05 (P = 0.30) | | | | | | | | |
| **Radiation Time** | | | | | | | | |
| Han 2019 | 81.5 | 38.6 | 532 | 71.5 | 44.2 | 584 | 49.9% | 10.00 [5.14, 14.86] |
| Roser 2013 | 15.98 | 8.6 | 72 | 31.5 | 11.4 | 40 | 50.1% | -15.52 [-19.57, -11.47] |
| **Subtotal (95% CI)** | | | **604** | | | **624** | **100.0%** | **-2.80 [-27.81, 22.21]** |
| Heterogeneity: Tau$^2$ = 320.42; Chi$^2$ = 62.49, df = 1 (P < 0.00001); I$^2$ = 98% | | | | | | | | |
| Test for overall effect: Z = 0.22 (P = 0.83) | | | | | | | | |

| | Robot-assisted | | | Conventional | | | | Std. Mean Difference |
|---|---|---|---|---|---|---|---|---|
| **Study or Subgroup** | **Mean** | **SD** | **Total** | **Mean** | **SD** | **Total** | **Weight** | **IV, Random, 95% CI** |
| **Radiation dose** | | | | | | | | |
| Feng 2019 | 38.87 | 11.94 | 202 | 47.45 | 15.75 | 225 | 25.5% | -0.61 [-0.80, -0.41] |
| Han 2019 | 21.7 | 11.5 | 532 | 70.5 | 42 | 584 | 25.8% | -1.55 [-1.69, -1.42] |
| Hyun 2017 | 0.55 | 0.19 | 130 | 1.22 | 0.35 | 140 | 24.7% | -2.35 [-2.66, -2.04] |
| Roser 2013 | 11.03 | 10.8 | 72 | 18.9 | 11.7 | 40 | 24.0% | -0.70 [-1.10, -0.30] |
| **Subtotal (95% CI)** | | | **936** | | | **989** | **100.0%** | **-1.30 [-2.01, -0.60]** |
| Heterogeneity: Tau$^2$ = 0.49; Chi$^2$ = 114.33, df = 3 (P < 0.00001); I$^2$ = 97% | | | | | | | | |
| Test for overall effect: Z = 3.65 (P = 0.0003) | | | | | | | | |

| | Robot-assisted | | Conventional | | | Odds Ratio |
|---|---|---|---|---|---|---|
| **Study or Subgroup** | **Events** | **Total** | **Events** | **Total** | **Weight** | **M-H, Fixed, 95% CI** |
| **Proximal facet joint violation** | | | | | | |
| Han 2019 | 0 | 115 | 12 | 119 | 46.1% | 0.04 [0.00, 0.64] |
| Hyun 2017 | 0 | 130 | 1 | 140 | 5.4% | 0.36 [0.01, 8.82] |
| Kim 2018 | 0 | 158 | 13 | 172 | 48.5% | 0.04 [0.00, 0.63] |
| **Subtotal (95% CI)** | | **403** | | **431** | **100.0%** | **0.05 [0.01, 0.28]** |
| Total events | 0 | | 26 | | | |
| Heterogeneity: Chi$^2$ = 1.45, df = 2 (P = 0.48); I$^2$ = 0% | | | | | | |
| Test for overall effect: Z = 3.47 (P = 0.0005) | | | | | | |

Fig. 4

Pooled results of the robot-assisted group and the conventional freehand with/without fluoroscopy-assisted group. The results are shown as follows: radiation dose, radiation time, operating time, proximal facet joint violation, intraoperative blood loss, postoperative visual analogue scale (VAS)-back, postoperative VAS-leg, postoperative stay, and postoperative Oswestry Disability Index (ODI). CI, confidence interval; IV, inverse variance; M-H, Mantel–Haenszel.

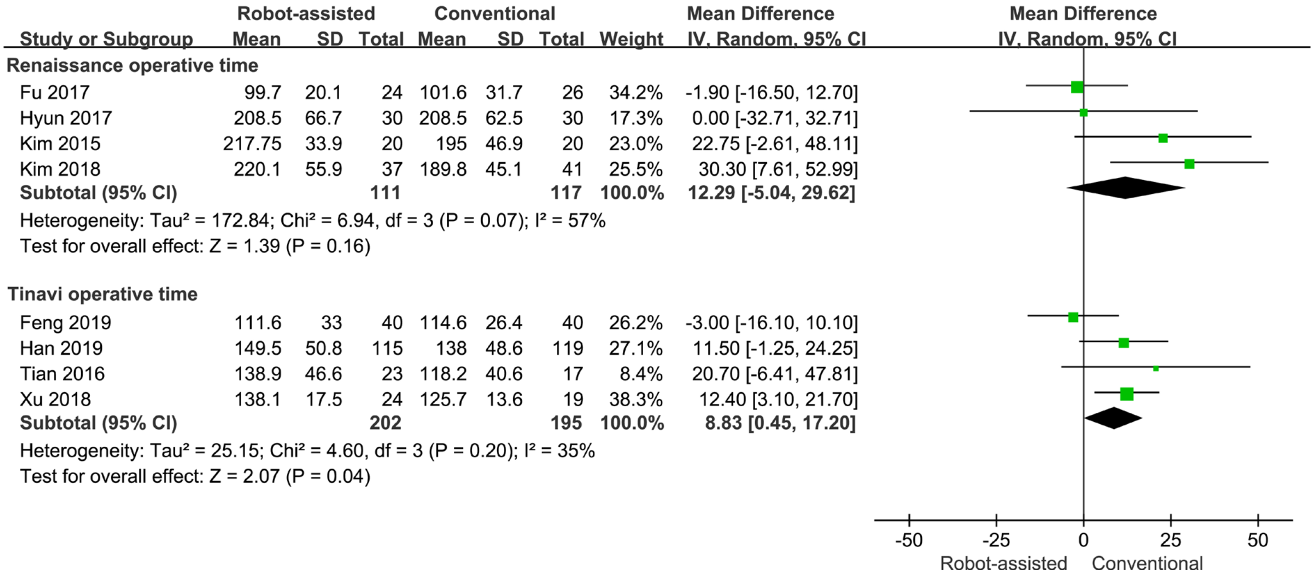


| Study or Subgroup | Robot-assisted Events | Robot-assisted Total | Conventional Events | Conventional Total | Weight | Odds Ratio M-H, Fixed, 95% CI |
|---|---|---|---|---|---|---|
| **SpineAssist accuracy Grade A** | | | | | | |
| Ringel 2012 | 56 | 146 | 68 | 152 | 98.3% | 0.77 [0.48, 1.22] |
| Roser 2013 | 71 | 72 | 39 | 40 | 1.7% | 1.82 [0.11, 29.91] |
| **Subtotal (95% CI)** | | **218** | | **192** | **100.0%** | **0.79 [0.50, 1.24]** |
| Total events | 127 | | 107 | | | |
| Heterogeneity: Chi² = 0.35, df = 1 (P = 0.55); I² = 0% | | | | | | |
| Test for overall effect: Z = 1.03 (P = 0.30) | | | | | | |
| **SpineAssist accuracy Grade A+B** | | | | | | |
| Ringel 2012 | 85 | 146 | 93 | 152 | 97.3% | 0.88 [0.56, 1.40] |
| Roser 2013 | 71 | 72 | 40 | 40 | 2.7% | 0.59 [0.02, 14.78] |
| **Subtotal (95% CI)** | | **218** | | **192** | **100.0%** | **0.88 [0.55, 1.38]** |
| Total events | 156 | | 133 | | | |
| Heterogeneity: Chi² = 0.06, df = 1 (P = 0.81); I² = 0% | | | | | | |
| Test for overall effect: Z = 0.57 (P = 0.57) | | | | | | |
| **Renaissance accuracy Grade A** | | | | | | |
| Hyun 2017 | 127 | 130 | 133 | 140 | 18.3% | 2.23 [0.56, 8.81] |
| Kim 2015 | 76 | 80 | 73 | 80 | 22.6% | 1.82 [0.51, 6.49] |
| Kim 2018 | 148 | 158 | 158 | 172 | 59.2% | 1.31 [0.57, 3.04] |
| **Subtotal (95% CI)** | | **368** | | **392** | **100.0%** | **1.59 [0.86, 2.97]** |
| Total events | 351 | | 364 | | | |
| Heterogeneity: Chi² = 0.48, df = 2 (P = 0.79); I² = 0% | | | | | | |
| Test for overall effect: Z = 1.47 (P = 0.14) | | | | | | |
| **Renaissance accuracy Grade A+B** | | | | | | |
| Hyun 2017 | 130 | 130 | 138 | 140 | 25.0% | 4.71 [0.22, 99.06] |
| Kim 2015 | 80 | 80 | 79 | 80 | 24.1% | 3.04 [0.12, 75.69] |
| Kim 2018 | 157 | 158 | 171 | 172 | 50.9% | 0.92 [0.06, 14.80] |
| **Subtotal (95% CI)** | | **368** | | **392** | **100.0%** | **2.38 [0.46, 12.30]** |
| Total events | 367 | | 388 | | | |
| Heterogeneity: Chi² = 0.67, df = 2 (P = 0.72); I² = 0% | | | | | | |
| Test for overall effect: Z = 1.03 (P = 0.30) | | | | | | |
| **Tinavi accuracy Grade A** | | | | | | |
| Feng 2019 | 199 | 202 | 206 | 225 | 9.3% | 6.12 [1.78, 21.00] |
| Han 2019 | 507 | 532 | 503 | 584 | 72.1% | 3.27 [2.05, 5.20] |
| Tian 2016 | 99 | 102 | 81 | 88 | 8.2% | 2.85 [0.71, 11.38] |
| Xu 2018 | 127 | 132 | 78 | 106 | 10.5% | 9.12 [3.38, 24.60] |
| **Subtotal (95% CI)** | | **968** | | **1003** | **100.0%** | **4.11 [2.81, 6.00]** |
| Total events | 932 | | 868 | | | |
| Heterogeneity: Chi² = 4.08, df = 3 (P = 0.25); I² = 27% | | | | | | |
| Test for overall effect: Z = 7.30 (P < 0.00001) | | | | | | |
| **Tinavi accuracy Grade A+B** | | | | | | |
| Feng 2019 | 202 | 202 | 224 | 225 | 6.4% | 2.71 [0.11, 66.80] |
| Han 2019 | 525 | 532 | 546 | 584 | 83.3% | 5.22 [2.31, 11.79] |
| Tian 2016 | 102 | 102 | 87 | 88 | 5.5% | 3.51 [0.14, 87.37] |
| Xu 2018 | 132 | 132 | 93 | 106 | 4.7% | 38.26 [2.25, 651.65] |
| **Subtotal (95% CI)** | | **968** | | **1003** | **100.0%** | **6.53 [3.17, 13.45]** |
| Total events | 961 | | 950 | | | |
| Heterogeneity: Chi² = 2.22, df = 3 (P = 0.53); I² = 0% | | | | | | |
| Test for overall effect: Z = 5.09 (P < 0.00001) | | | | | | |

Fig. 5

Pooled results of the subgroup analysis (SpineAssist-assisted group, Renaissance-assisted group, Tinavi robot-assisted group, and conventional freehand with/without fluoroscopy-assisted group). The results are shown as follows: Grade A accuracy and Grade A + B accuracy. CI, confidence interval; M-H, Mantel–Haenszel.

is not satisfactory, the path should be positioned again after fine tuning of the robot, or positioning image acquisition should be restarted.[21] Numerically speaking, the accuracy of the RRA group was much better than that of the SRA group, but there were no statistical differences of accuracy between the experimental group and the control group of different generations. With the accumulation and development of surgical experience and skills, CT surgery technology and surgical equipment are also constantly upgraded to overcome the difficult learning



| Study or Subgroup | Robot-assisted Mean | Robot-assisted SD | Robot-assisted Total | Conventional Mean | Conventional SD | Conventional Total | Weight | Mean Difference IV, Random, 95% CI |
|---|---|---|---|---|---|---|---|---|
| **Renaissance operative time** | | | | | | | | |
| Fu 2017 | 99.7 | 20.1 | 24 | 101.6 | 31.7 | 26 | 34.2% | -1.90 [-16.50, 12.70] |
| Hyun 2017 | 208.5 | 66.7 | 30 | 208.5 | 62.5 | 30 | 17.3% | 0.00 [-32.71, 32.71] |
| Kim 2015 | 217.75 | 33.9 | 20 | 195 | 46.9 | 20 | 23.0% | 22.75 [-2.61, 48.11] |
| Kim 2018 | 220.1 | 55.9 | 37 | 189.8 | 45.1 | 41 | 25.5% | 30.30 [7.61, 52.99] |
| **Subtotal (95% CI)** | | | **111** | | | **117** | **100.0%** | **12.29 [-5.04, 29.62]** |
| Heterogeneity: Tau² = 172.84; Chi² = 6.94, df = 3 (P = 0.07); I² = 57% | | | | | | | | |
| Test for overall effect: Z = 1.39 (P = 0.16) | | | | | | | | |
| **Tinavi operative time** | | | | | | | | |
| Feng 2019 | 111.6 | 33 | 40 | 114.6 | 26.4 | 40 | 26.2% | -3.00 [-16.10, 10.10] |
| Han 2019 | 149.5 | 50.8 | 115 | 138 | 48.6 | 119 | 27.1% | 11.50 [-1.25, 24.25] |
| Tian 2016 | 138.9 | 46.6 | 23 | 118.2 | 40.6 | 17 | 8.4% | 20.70 [-6.41, 47.81] |
| Xu 2018 | 138.1 | 17.5 | 24 | 125.7 | 13.6 | 19 | 38.3% | 12.40 [3.10, 21.70] |
| **Subtotal (95% CI)** | | | **202** | | | **195** | **100.0%** | **8.83 [0.45, 17.20]** |
| Heterogeneity: Tau² = 25.15; Chi² = 4.60, df = 3 (P = 0.20); I² = 35% | | | | | | | | |
| Test for overall effect: Z = 2.07 (P = 0.04) | | | | | | | | |

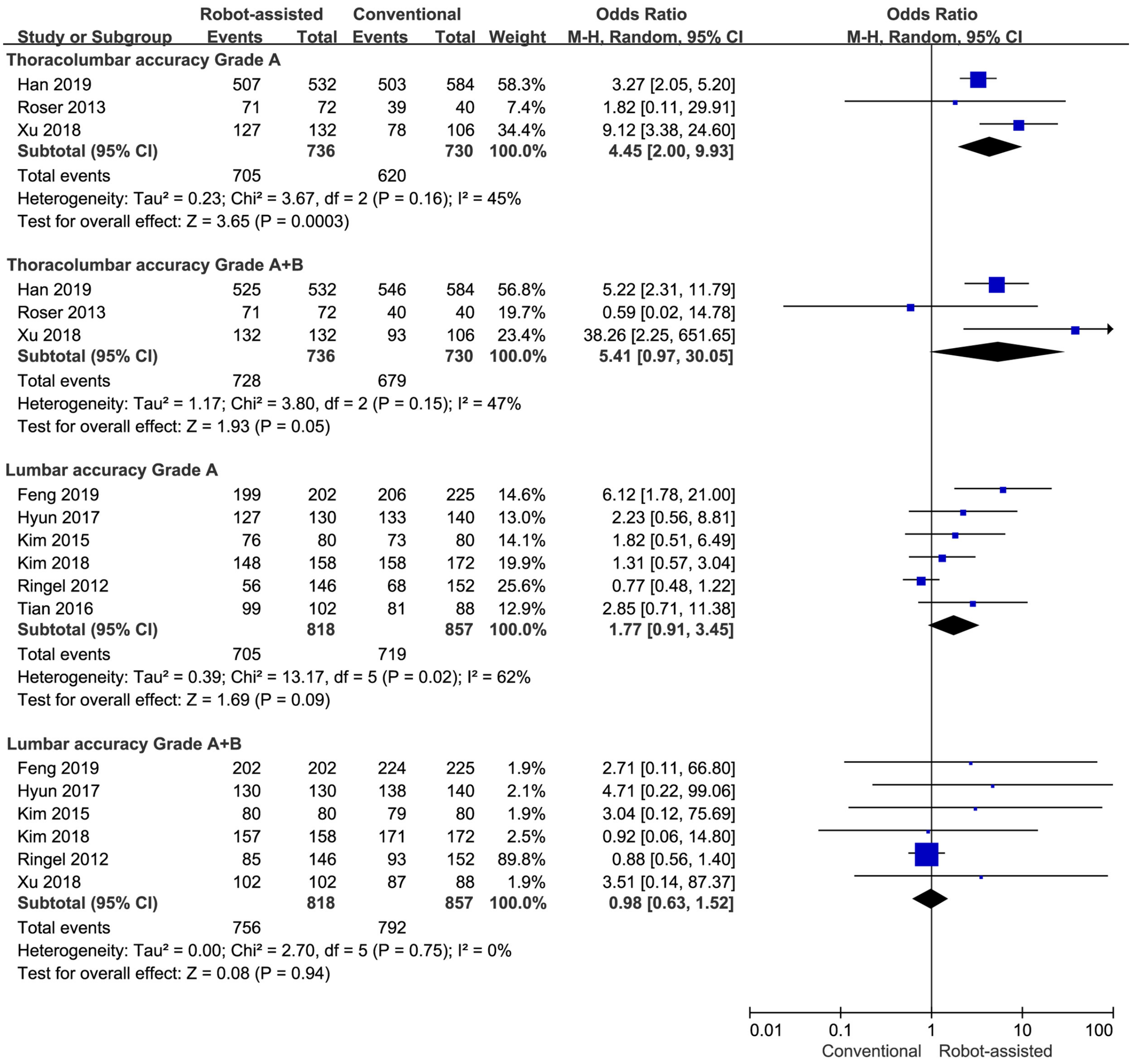
Fig. 6

Pooled results of the subgroup analysis (operating time of renaissance-assisted group, Tinavi robot-assisted group, and conventional freehand with/without fluoroscopy-assisted group). CI, confidence interval; IV, inverse variance.

curve.[22-25] We suspected that this may be the reason why the surgical accuracy of the non-robot group was also improved.

However, the above-mentioned outcomes in this paper do not mean that Tirobot is more advanced than Mazor robot in accuracy of pedicle screw insertion in a real sense. We suspected that the superiority of accuracy in the TRA group could be misleading, since the accuracies of the control group in Renaissance studies are generally better than those in Tinavi studies according to the Grade A (control group’s accuracy in Renaissance studies: 92.9%; control group’s accuracy in Tinavi studies: 86.5%) and Grade A + B criteria (control group’s accuracy in Renaissance studies: 99.0%; control group’s accuracy in Tinavi studies: 94.7%). In order to find out which kind of robot is more accurate, it is necessary to conduct a direct comparison of different types of robots with high-quality and large sample sizes. At present, there is no such RCT for horizontal comparison of multiple kinds of robots. So, there are no data to support which kind of robot is more accurate. With the further upgrading of robot technology, the new generation of robots are equipped with intraoperative navigation technology, such as MazorX, ROSA, and Tinavi. Many studies have mentioned that accuracy has been further improved due to the emergence of optical intraoperative navigation robotic technology.[26-29] We performed a brief overview of the current spinal surgery robots shown in Table II, and some research suggests that the MazorX will have a stronger registration software and be equipped with intraoperative navigation technology at the same time, which makes us look forward to a brighter performance from the new generation of Mazor. However there are still some vision-blind areas of optical navigation. In some areas, the mechanical arm cannot be detected by infrared ray since the distance from tracker frame to NDI Polari is too far to maintain image quality and too close to maintain operator's range, so the appropriate distance between those two still needs to be adjusted according to the performer's ability. Additionally, sometimes a proximal facet joint with steep angles may cause the slipping of screw insertion, leading to proximal facet joint violation. At present, we suspect that accuracy of screw insertion is relatively better in robot systems equipped with infrared intraoperative navigation; nevertheless, more independent high-quality RCTs with comparison between different types of robots are still needed as the basis for this further research.

There was no statistically significant difference in accuracy between the RA group and CT group when the pedicle screws were inserted on lumbar spine alone. One reason is that the diameter of the pedicles in the lumbar spine is greater and risk of vital structure damage is reduced.[30] In addition, the surgeon becomes more experienced and the imaging-assisted technique is upgraded continuously, which may be the other two reasons for the substantial improvement in accuracy of CT. But compared with CT, the accuracy rate of RA was better according to the Grade A and Grade A + B criteria when the operative level was the thoracolumbar level. This is because the thoracic pedicles are relatively thin, which means the technical requirements of freehand pedicle screw insertion are relatively high.[31] Misplacement rates of freehand with/without fluoroscopy-assisted pedicle screw still remain high in the thoracic spine.[32] The learning curve for the thoracic pedicle screw insertion is also steeper, and it takes a long time for the surgeon to learn and master. The surgeon is required to maintain a consistently high level



| | Robot-assisted | | Conventional | | | Odds Ratio |
|---|---|---|---|---|---|---|
| Study or Subgroup | Events | Total | Events | Total | Weight | M-H, Random, 95% CI |
| **Thoracolumbar accuracy Grade A** | | | | | | |
| Han 2019 | 507 | 532 | 503 | 584 | 58.3% | 3.27 [2.05, 5.20] |
| Roser 2013 | 71 | 72 | 39 | 40 | 7.4% | 1.82 [0.11, 29.91] |
| Xu 2018 | 127 | 132 | 78 | 106 | 34.4% | 9.12 [3.38, 24.60] |
| **Subtotal (95% CI)** | | **736** | | **730** | **100.0%** | **4.45 [2.00, 9.93]** |
| Total events | 705 | | 620 | | | |
| Heterogeneity: Tau² = 0.23; Chi² = 3.67, df = 2 (P = 0.16); I² = 45% | | | | | | |
| Test for overall effect: Z = 3.65 (P = 0.0003) | | | | | | |
| **Thoracolumbar accuracy Grade A+B** | | | | | | |
| Han 2019 | 525 | 532 | 546 | 584 | 56.8% | 5.22 [2.31, 11.79] |
| Roser 2013 | 71 | 72 | 40 | 40 | 19.7% | 0.59 [0.02, 14.78] |
| Xu 2018 | 132 | 132 | 93 | 106 | 23.4% | 38.26 [2.25, 651.65] |
| **Subtotal (95% CI)** | | **736** | | **730** | **100.0%** | **5.41 [0.97, 30.05]** |
| Total events | 728 | | 679 | | | |
| Heterogeneity: Tau² = 1.17; Chi² = 3.80, df = 2 (P = 0.15); I² = 47% | | | | | | |
| Test for overall effect: Z = 1.93 (P = 0.05) | | | | | | |
| **Lumbar accuracy Grade A** | | | | | | |
| Feng 2019 | 199 | 202 | 206 | 225 | 14.6% | 6.12 [1.78, 21.00] |
| Hyun 2017 | 127 | 130 | 133 | 140 | 13.0% | 2.23 [0.56, 8.81] |
| Kim 2015 | 76 | 80 | 73 | 80 | 14.1% | 1.82 [0.51, 6.49] |
| Kim 2018 | 148 | 158 | 158 | 172 | 19.9% | 1.31 [0.57, 3.04] |
| Ringel 2012 | 56 | 146 | 68 | 152 | 25.6% | 0.77 [0.48, 1.22] |
| Tian 2016 | 99 | 102 | 81 | 88 | 12.9% | 2.85 [0.71, 11.38] |
| **Subtotal (95% CI)** | | **818** | | **857** | **100.0%** | **1.77 [0.91, 3.45]** |
| Total events | 705 | | 719 | | | |
| Heterogeneity: Tau² = 0.39; Chi² = 13.17, df = 5 (P = 0.02); I² = 62% | | | | | | |
| Test for overall effect: Z = 1.69 (P = 0.09) | | | | | | |
| **Lumbar accuracy Grade A+B** | | | | | | |
| Feng 2019 | 202 | 202 | 224 | 225 | 1.9% | 2.71 [0.11, 66.80] |
| Hyun 2017 | 130 | 130 | 138 | 140 | 2.1% | 4.71 [0.22, 99.06] |
| Kim 2015 | 80 | 80 | 79 | 80 | 1.9% | 3.04 [0.12, 75.69] |
| Kim 2018 | 157 | 158 | 171 | 172 | 2.5% | 0.92 [0.06, 14.80] |
| Ringel 2012 | 85 | 146 | 93 | 152 | 89.8% | 0.88 [0.56, 1.40] |
| Xu 2018 | 102 | 102 | 87 | 88 | 1.9% | 3.51 [0.14, 87.37] |
| **Subtotal (95% CI)** | | **818** | | **857** | **100.0%** | **0.98 [0.63, 1.52]** |
| Total events | 756 | | 792 | | | |
| Heterogeneity: Tau² = 0.00; Chi² = 2.70, df = 5 (P = 0.75); I² = 0% | | | | | | |
| Test for overall effect: Z = 0.08 (P = 0.94) | | | | | | |

Fig. 7

Pooled results of the subgroup analysis (operative level thoracolumbar group, operative level lumbar group, and conventional freehand with/without fluoroscopy-assisted group). The results are shown as follows: Grade A accuracy and Grade A + B accuracy. CI, confidence interval; M-H, Mantel–Haenszel.

**Table II.** Overview of current commonly used spinal robotics.

| Name | Production place | Certification | Application |
|---|---|---|---|
| Mazor robotics (SpineAssist/Renaissance) | Israel | CE (CONFORMITE EUROPEENNE);<br>FDA (Food and Drug Administration) | Assisting pedicle screw insertion |
| Robotic Surgical Assistant (ROSA) | France | CE (CONFORMITE EUROPEENNE);<br>FDA (Food and Drug Administration) | Assisting pedicle screw insertion |
| TiRobot orthopaedic robot (Tinavi Robot) | China | CFDA<br>(China Food and Drug Administration) | Assisting surgeries of limbs, pelvis, and all spinal segments |

of performance and mental concentration during long-time and challenging operations. In our opinion, using robots is more advantageous for this difficult pedicle screw insertion surgery. Using robots for pedicle screw insertion can not only improve accuracy, but may also maintain stability. We also found that the operating time

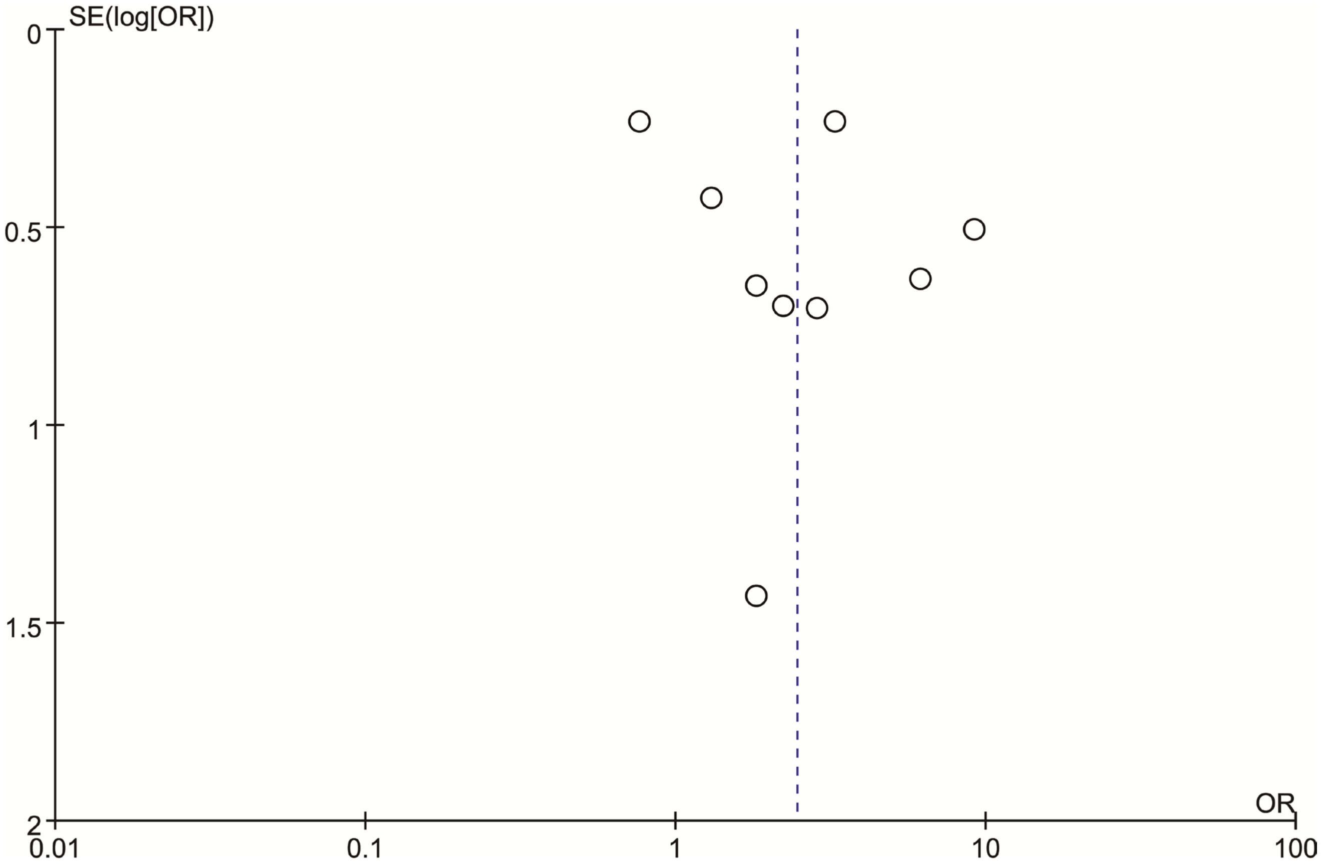


Fig. 8

Funnel plot of robot-assisted versus conventional freehand with/without fluoroscopy-assisted technique. OR, odds ratio.

of the TRA group was longer than that of the CT group. While the difference between the MRA group and CT was not statistically significant for operating time, we suspect the reason that the operating time of the TRA group was longer than CT is that more intraoperative planning is needed when employing the TiRobot system. The problem existed when we applicated TiRobot system in practice. When more than three segments of the lumbar vertebra or more than five segments of the thoracic vertebra were involved in the operation, multiple instances of intraoperative planning are required. Intraoperative planning increases the duration of surgery, and an increased duration of surgery means an increased risk. Strictly speaking, the current spine robotic devices on the market which are composed of optical navigation and robotic arms are not really robots. In the future, a lot of work should be done to further improve the robotic devices; our research is also focused on how to combine preoperative planning and intraoperative planning to reduce intraoperative planning time and autonomous learning of robots. In the future, surgical robots will be combined with artificial intelligence. These will then be able to not only achieve minimally invasive surgery (MIS) and navigation guidance, but also independently analyze the patients' characteristics and carry out preoperative planning and intraoperative operations autonomously.

In our study, the difference between the RA group and CT group had no statistical significance in both postoperative VAS–back and postoperative VAS–leg, but postoperative ODI of the RA group was smaller. Degree of pain and results of treatment were scored by VAS, which is by far the most frequently used assessment instrument. A preponderance of evidence[33-37] demonstrated that the ODI had been exploited to evaluate pain-related disability in people with various kinds of lower back pain. A few different versions have been exploited since 1980, when the first version was published. Version 2.0 is now recommended for general use. The ODI covers one item on pain and nine items on activities of daily living (personal care, lifting, walking, sitting, standing, sleeping, sex life, social life, and travelling).[38-40] In our opinion, RA technique supported execution of MIS thus reducing intraoperative injury and improving postoperative recovery of the patient. We considered this was primarily why the postoperative ODI of RA was statistically smaller compared with CT.

The first systematic review and meta-analysis to investigate the difference of accuracy between RA and conventional freehand pedicle screw insertion was published in 2016 by Liu et al,[41] whose included studies comprised both RCTs and cohort studies, leading to less powerful results. Gao et al[42] first published a meta-analysis and systematic review to compare RA and freehand with/without fluoroscopy-assisted pedicle screw implantation purely based on several RCTs conducted in 2018. Due to the lack of sufficient pertinent RCTs, the statistically pooled results of RA and conventional freehand pedicle screw insertion in the treatment of spinal diseases were not comprehensive in these two reviews. Before our study, the study by Gao et al[42] reported that compared with conventional freehand with/without fluoroscopy-assisted pedicle screw insertion, a smaller intraoperative radiation dose was significantly associated with RA surgery. The study by Li et al[10] reported that RA pedicle screw insertion was associated with significantly less intraoperative radiation exposure time and a lower intraoperative radiation dose. We think that these two results were based on an incorrect comparison method. In the study by Gao et al,[42] they used the RCTs conducted by Hyun et al[18] and Roser et al[20] to compare radiation dose. However, the original data of the former and latter studies were C-arm-registered fluoroscopy exposure per screw and total radiation exposure for robotic-assisted surgeries, respectively. The same applies to the study by Li et al,[10] in which the authors used radiation time of C-arm fluoroscopy with each screw[18] and that of C-arm fluoroscopy with full operation[20] to compare for radiation dose and radiation time. We attempted to contact the authors, but unfortunately the original data were not available. We arrived at the conclusion that the credibility of the meta-analysis by Gao et al[42] on radiation dose, and the meta-analysis by Li et al[10] on radiation dose and radiation time, were questionable.

Radiation exposure was most likely dependent on the type of robot and variability of surgeon.[6,43,44] The radiation dose may be dramatically increased through optical intraoperation navigation technique compared with conventional open freehand pedicle screw insertion.[45-48] Although the surgical team have utilized many approaches of radiation protection to decrease the high exposure risk (e.g. whole-body apron, thyroid shield, safe distance, or shield room), the patient was still very exposed.[49-51] In this meta-analysis, the difference in radiation dose between RA and CT was statistically significant ($p < 0.001$, Z test; $I^2 = 97\%$). Due to differences in evaluation method, environment, measuring equipment, and robot types in each included RCT comparing the radiation dose between the two intervention methods, qualitative conclusions have been drawn showing that the radiation dose in robotic surgery was lower than that in conventional freehand with/without fluoroscopy-assisted pedicle screw insertion. However the range of each result varied greatly in numerical value, and some results even differed by orders of magnitude. We think the reason for this outcome may be the application of optical navigation. Because of Mazor's working theory, the combination of preoperative CT and intraoperative imaging can greatly reduce the radiation exposure of the medical team and patients, while the optical navigation robot system conducts intraoperative guidance with the aid of optical navigation, which still needs further research on the radiation exposure of the medical team and patients. Furthermore, Kaminski et al[48] demonstrated that radiation dose imparted to patients not only depended on operating time, surgical technique, and acquisition protocol but also on the patient's body mass index (BMI). After the different units in radiation dose were removed by employing SMD, the difference in radiation dose between RA and CT was statistically significant ($p < 0.001$, Z test; $I^2 = 97\%$). Nevertheless, due to its significant heterogeneity, this conclusion still lacked clinical significance. The difference in radiation time between the RA and CT groups also showed no clinical significance ($p = 0.830$, Z test; $I^2 = 98\%$). More independent high-quality RCTs using sufficiently large sample sizes with radiation time and dose are needed.

The aim of this study was to systematically compare safety and accuracy of RA with CT pedicle screw insertion for spine disease. Many published studies have performed the comparison of RA with CT pedicle screw insertion for spine disease. However, most of these were based on non-RCTs. Our statistical efficiency is superior to these studies. The studies of Liu et al,[41] Gao et al,[42] and Li et al[10] are meta-analyses on this subject; the former was the first published study based on RCTs and cohort studies, and the latter two were based on pure RCTs. As the same type of meta-analysis, our study contained more RCTs with different types and generations of robotics, and performed many refined subgroup analyses based on types of robots, generations of Mazor robot, and operative levels. Furthermore, we corrected the methods they used to extract data and perform comparisons. After the data from ten high-quality included trials were pooled, the results suggested that the RA group had a superior accuracy rate utilizing both Gertzbein–Robbins Grade A criteria and Gertzbein–Robbins Grade A + B criteria, smaller postoperative ODI, and less proximal facet joint violation. In the subgroup analysis, the results suggested that robot equipped with optical intraoperative navigation was more accurate in pedicle screw insertion. The advantage of this meta-analysis was that the RCTs were collected on the same topic of research. In each included RCT, the patients were selected discreetly and the inclusion and exclusion criteria of patients were similar in each trial included in this meta-analysis, which ensured our statistical efficacy. Scrupulous patient selection is essential to acquiring a good clinical result. Pooling RCTs that have high methodological qualities can imply a lower risk of bias. By considering all included trials without language restrictions, this meta-analysis avoided outcomes distorted by language bias. But there were still several limitations in this meta-analysis. We could not collect all the negative conclusions that existed in any unpublished studies, which may lead to an overestimation of the authentic effectiveness of interventions. But the publication bias could be seen as not statistically significant through funnel plot generated via RevMan 5.3 (Copenhagen: The Nordic Cochrane Centre, The Cochrane Collaboration) (Figure 8). Besides, all of the included trials failed to report the comparison of costs and reimbursement. As is well known, medical problems are not only a scientific issue but also constitute a social problem. Malham and Wells-Quinn[52] suggested that the cost of a robotic arm equipped with intraoperative navigation requires additional data to justify its expense. Although robotic surgery technology will be at its peak one day, its expenses can only be accepted by a limited number people. That situation will also be the opposite of the original intention of this technology, which is to serve the general public. So far, the question of robot operation costs and reimbursement is still pending in China.

All in all, so far robotics offers a tool to improve the skills of surgeons, who must have the ability and experience to perform CT pedicle screw insertion. However, we do not recommend the use of robots to complete MIS for surgeons who have no experience in CT pedicle screw insertion. The experience of conventional freehand pedicle screw insertion will improve the surgeons' understanding of the core of this surgical procedure and anatomical structure from useless rote memorization to substantial cognition. The CT pedicle screw insertion technique can improve the surgical skills and proficiency of the surgeon, making a more perfect combination of the surgeon and the robot in robotic surgery, which allows surgeons to better understand the advantages of robotic surgery. In addition, robotic surgery technology is still in its infancy at the present stage. Our study concluded that the accuracy of robotic screw insertion was improved, but there is still a long way to go in achieving authentically robot-intelligent MIS. While the design and application of robots have further improved, changing surgical methods is also one of the methods by which surgeons may improve the level of MIS. Further improvement of robots, continuous emergence of large data sample size RCT comparing different kinds of robots, and further comparison of costs and reimbursement are the problems that need further study. Robotic surgery might change the models and future trends of surgery: spinal surgery robots and even surgical robots could be combined with artificial intelligence in the future. Through deep learning, the current RA pedicle screw insertion could be transformed into robotic automatic pedicle screw insertion. At that point, a robot will be able to perform the surgery under the supervision of a qualified operator. Alternatively, by making full use of the advantages of big data, it is possible to achieve remotely operated robotic surgery in the future, which will further promote the rational use of medical resources and the reform of surgical techniques.

In conclusion, the outcomes of this meta-analysis suggest that the RA technique could be viewed as accurate and safe pedicle screw implantation compared to CT. The results of the subgroup analyses showed that a robot system equipped with optical intraoperative navigation was more accurate in pedicle screw insertion than CT. Using robots for pedicle screw insertion can improve accuracy and maintain stability for some challenging areas such as the thoracic vertebra. At present, there are no published RCTs reporting the cost-effectiveness analyses. More independent high-quality RCTs using sufficiently large sample sizes with radiation time, radiation dose, and cost-effectiveness analyses are needed. More comparisons between different types of robots should also be conducted.

## References

1. **Ondeck NT**, **Bohl DD**, **McLynn RP**, **et al**. Longer Operative Time Is Associated With Increased Adverse Events After Anterior Cervical Diskectomy and Fusion: 15-Minute Intervals Matter. *Orthopedics*. 2018;41(4):e483–e488.
2. **Khan A**, **Meyers JE**, **Siasios I**, **Pollina J**. Next-Generation Robotic Spine Surgery: First Report on Feasibility, Safety, and Learning Curve. *Oper Neurosurg (Hagerstown)*. 2019;17(1):61–69.
3. **Stull JD**, **Mangan JJ**, **Vaccaro AR**, **Schroeder GD**. Robotic guidance in minimally invasive spine surgery: a review of recent literature and commentary on a developing technology. *Curr Rev Musculoskelet Med*. 2019;12(2):245–251.
4. **Staartjes VE**, **Klukowska AM**, **Schröder ML**. Pedicle Screw Revision in Robot-Guided, Navigated, and Freehand Thoracolumbar Instrumentation: A Systematic Review and Meta-Analysis. *World Neurosurg*. 2018;116:433–443.e8.
5. **Ghasem A**, **Sharma A**, **Greif DN**, **Alam M**, **Maaieh MA**. The arrival of robotics in spine surgery: a review of the literature. *Spine*. 2018;43(23):1670–1677.
6. **Joseph JR**, **Smith BW**, **Liu X**, **Park P**. Current applications of robotics in spine surgery: a systematic review of the literature. *Neurosurg Focus*. 2017;42(5):E2.
7. **Fiani B**, **Quadri SA**, **Farooqui M**, **et al**. Impact of robot-assisted spine surgery on health care quality and neurosurgical economics: a systemic review. *Neurosurg Rev*. 2020;43(1):17–25.
8. **Huang J**, **Li Y**, **Huang L**. Spine surgical robotics: review of the current application and disadvantages for future perspectives. *J Robot Surg*. 2020;14(1):11–16.
9. **Shweikeh F**, **Amadio JP**, **Arnell M**, **et al**. Robotics and the spine: a review of current and ongoing applications. *Neurosurg Focus*. 2014;36(3):E10.
10. **Li HM**, **Zhang R-J**, **Shen C-L**. Accuracy of pedicle screw placement and clinical outcomes of robot-assisted technique versus conventional Freehand technique in spine surgery from nine randomized controlled trials: a meta-analysis. *Spine*. 2020;45(2):E111–E119.
11. **Feng S**, **Tian W**, **Sun Y**, **et al**. Effect of robot-assisted surgery on lumbar pedicle screw internal fixation in patients with osteoporosis. *World Neurosurg*. 2019;125(6):e1057–e1062.
12. **Han X**, **Tian W**, **Liu Y**, **et al**. Safety and accuracy of robot-assisted versus fluoroscopy-assisted pedicle screw insertion in thoracolumbar spinal surgery: a prospective randomized controlled trial. *J Neurosurg Spine*. 2019:615–622.
13. **Tian W**, **Fan MX**, **Han XG**, **Zhao JW**, **Liu YJ**. Pedicle screw insertion in spine: a randomized comparison study of robot-assisted surgery and fluoroscopy-guided techniques. *Orthop Surg*. 2016;11(2):4–10.
14. **Xu P**, **Ge P**, **Zhang RJ**, **et al**. Effect of robot assisted pedicle screw fixation in the treatment of thoracolumbar fracture. *Orthop Surg*. 2018;39(6):687–690.
15. **Fu S**, **Shao SZ**, **Wang LQ**, **Wang YN**, **Hou HT**. Pedicle screw insertion in spine: a randomized comparison study of robot-assisted surgery and fluoroscopy-guided techniques. *Orthop Surg*. 2017;3(2):70–76.
16. **Kim H-J**, **Kang K-T**, **Chun H-J**, **et al**. Comparative study of 1-year clinical and radiological outcomes using robot-assisted pedicle screw fixation and freehand technique in posterior lumbar interbody fusion: a prospective, randomized controlled trial. *Int J Med Robot*. 2018;14(4):e1917.
17. **Kim H-J**, **Lee SH**, **Chang B-S**, **et al**. Monitoring the quality of robot-assisted pedicle screw fixation in the lumbar spine by using a cumulative summation test. *Spine*. 2015;40(2):87–94.
18. **Hyun S-J**, **Kim K-J**, **Jahng T-A**, **Kim H-J**. Minimally invasive robotic versus open Fluoroscopic-guided spinal instrumented fusions: a randomized controlled trial. *Spine*. 2017;42(6):353–358.

19. **Ringel F**, **Stüer C**, **Reinke A**, **et al**. Accuracy of robot-assisted placement of lumbar and sacral pedicle screws: a prospective randomized comparison to conventional freehand screw implantation. *Spine*. 2012;37(8):E496–501.
20. **Roser F**, **Tatagiba M**, **Maier G**. Spinal robotics: current applications and future perspectives. *Neurosurgery*. 2013;72 Suppl 1:12–18.
21. **Tian W**, **Liu Y-J**, **Liu B**, **et al**. Guideline for thoracolumbar pedicle screw placement assisted by orthopaedic surgical robot. *Orthop Surg*. 2019;11(2):153–159.
22. **Sharif S**, **Afsar A**. Learning curve and minimally invasive spine surgery. *World Neurosurg*. 2018;119:472–478.
23. **Qian L**, **Jiang C**, **Sun P**, **et al**. A comparison of the biomechanical stability of pedicle-lengthening screws and traditional pedicle screws: an in vitro instant and fatigue-resistant pull-out test. *Bone Joint J*. 2018;4:516–521.
24. **Schatlo B**, **Martinez R**, **Alaid A**, **et al**. Unskilled unawareness and the learning curve in robotic spine surgery. *Acta Neurochir*. 2015;157(10):1819–1823.
25. **Hu X**, **Lieberman IH**. What is the learning curve for robotic-assisted pedicle screw placement in spine surgery? *Clin Orthop Relat Res*. 2014;472(6):1839–1844.
26. **Khan A**, **Meyers JE**, **Yavorek S**, **et al**. Comparing next-generation robotic technology with 3-dimensional computed tomography navigation technology for the insertion of posterior pedicle screws. *World Neurosurg*. 2019;123:e474–e481.
27. **Chenin L**, **Capel C**, **Fichten A**, **Peltier J**, **Lefranc M**. Evaluation of screw placement accuracy in circumferential lumbar arthrodesis using robotic assistance and intraoperative Flat-Panel computed tomography. *World Neurosurg*. 2017;105:86–94.
28. **Lefranc M**, **Peltier J**. Accuracy of thoracolumbar transpedicular and vertebral body percutaneous screw placement: coupling the Rosa® Spine robot with intraoperative flat-panel CT guidance--a cadaver study. *J Robot Surg*. 2015;9(4):331–338.
29. **Lonjon N**, **Chan-Seng E**, **Costalat V**, **et al**. Robot-assisted spine surgery: feasibility study through a prospective case-matched analysis. *Eur Spine J*. 2016;25(3):947–955.
30. **Youkilis AS**, **Quint DJ**, **McGillicuddy JE**, **Papadopoulos SM**. Stereotactic navigation for placement of pedicle screws in the thoracic spine. *Neurosurgery*. 2001;48(4):771–778.
31. **Huang H**, **Nightingale RW**, **Dang ABC**. Biomechanics of coupled motion in the cervical spine during simulated whiplash in patients with pre-existing cervical or lumbar spinal fusion: a finite element study. *Bone Joint Res*. 2018;7(1):28–35.
32. **Perna F**, **Borghi R**, **Pilla F**, **et al**. Pedicle screw insertion techniques: an update and review of the literature. *Musculoskelet Surg*. 2016;100(3):165–169.
33. **Collins SL**, **Moore RA**, **McQuay HJ**. The visual analogue pain intensity scale: what is moderate pain in millimetres? *Pain*. 1997;72(1-2):95–97.
34. **de Williams AC**, **Davies HTO**, **Chadury Y**. Simple pain rating scales hide complex idiosyncratic meanings. *Pain*. 2000;85(3):457–463.
35. **Hirschfeld G**, **Zernikow B**. Cut points for mild, moderate, and severe pain on the vas for children and adolescents: what can be learned from 10 million ANOVAs? *Pain*. 2013;154(12):2626–2632.
36. **Reed MD**, **Van Nostran W**. Assessing pain intensity with the visual analog scale: a plea for uniformity. *J Clin Pharmacol*. 2014;54(3):241–244.
37. **Schechter NL**. From the Ouchless place to comfort central: the evolution of a concept. *Pediatrics*. 2008;122(Suppl 3):S154–S160.
38. **Fairbank JC**, **Couper J**, **Davies JB**, **O'Brien JP**. The Oswestry low back pain disability questionnaire. *Physiotherapy*. 1980;66(8):271–273.
39. **Fairbank JC**, **Pynsent PB**. The Oswestry Disability Index. *Spine*. 2000;25(22):2940–2952.
40. **Smeets R**, **Köke A**, **Lin C-W**, **Ferreira M**, **Demoulin C**. Measures of function in low back pain/disorders: low back pain rating scale (LBPRS), Oswestry disability index (ODI), progressive Isoinertial lifting evaluation (pile), Quebec back pain disability scale (QBPDS), and Roland-Morris disability questionnaire. *Arthritis Care Res*. 2011;63(S11 ):S158–S173.
41. **Liu H**, **Chen W**, **Wang Z**, **et al**. Comparison of the accuracy between robot-assisted and conventional freehand pedicle screw placement: a systematic review and meta-analysis. *Int J Comput Assist Radiol Surg*. 2016;11(12):2273–2281.
42. **Gao S**, **Lv Z**, **Fang H**. Robot-assisted and conventional freehand pedicle screw placement: a systematic review and meta-analysis of randomized controlled trials. *Eur Spine J*. 2018;27(4):921–930.
43. **Kantelhardt SR**, **Martinez R**, **Baerwinkel S**, **et al**. Perioperative course and accuracy of screw positioning in conventional, open robotic-guided and percutaneous robotic-guided, pedicle screw placement. *Eur Spine J*. 2011;20(6):860–868.
44. **Onen MR**, **Simsek M**, **Naderi S**. Robotic spine surgery: a preliminary report. *Turk Neurosurg*. 2014;24(4):512–518.
45. **Allam Y**, **Silbermann J**, **Riese F**, **Greiner-Perth R**. Computer tomography assessment of pedicle screw placement in thoracic spine: comparison between free hand and a generic 3D-based navigation techniques. *Eur Spine J*. 2013;22(3):648–653.
46. **Oertel MF**, **Hobart J**, **Stein M**, **Schreiber V**, **Scharbrodt W**. Clinical and methodological precision of spinal navigation assisted by 3D intraoperative O-arm radiographic imaging. *J Neurosurg Spine*. 2011;14(4):532–536.
47. **Pennington Z**, **Cottrill E**, **Westbroek EM**, **et al**. Evaluation of surgeon and patient radiation exposure by imaging technology in patients undergoing thoracolumbar fusion: systematic review of the literature. *Spine J*. 2019;19(8):1397–1411.
48. **Kaminski L**, **Cordemans V**, **Cartiaux O**, **Van Cauter M**. Radiation exposure to the patients in thoracic and lumbar spine fusion using a new intraoperative cone-beam computed tomography imaging technique: a preliminary study. *Eur Spine J*. 2017;26(11):2811–2817.
49. **Balter S**. Methods for measuring fluoroscopic skin dose. *Pediatr Radiol*. 2006;36 Suppl 2(Suppl 2):136–140.
50. **Harstall R**, **Heini PF**, **Mini RL**, **Orler R**. Radiation exposure to the surgeon during fluoroscopically assisted percutaneous vertebroplasty: a prospective study. *Spine*. 2005;30(16):1893–1898.
51. **Mulconrey DS**. Fluoroscopic radiation exposure in spinal surgery: in vivo evaluation for operating room personnel. *Clin Spine Surg*. 2016;29(7):E331–335.
52. **Malham GM**, **Wells-Quinn T**. What should my hospital buy next?-Guidelines for the acquisition and application of imaging, navigation, and robotics for spine surgery. *J Spine Surg*. 2019;5(1):155–165.

**Author information:**

- W. Li, MD, Orthopaedic Surgeon
- L. Cong, MD, PhD, Orthopaedic Surgeon
  Department of Orthopedic Surgery, The First Hospital of China Medical University, Shenyang, China.
- G. Li, MD, Gynaecologist, Department of Obstetrics and Gynecology, Shengjing hospital of China Medical University, Shenyang, China.
- W. Chen, MD, Attending Physician, Disease Control and Prevention Center, China Railway Shenyang Bureau Group Corporation, Shengyang, China.

**Author contributions:**

- W. Li: Conceptualized and designed the study, Acquired and interpreted the data, Drafted the manuscript.
- G. Li: Acquired the data.
- W. Chen: Acquired and interpreted the data.
- L. Cong: Conceptualized, designed, and supervised the study, Interpreted the data, Critically reviewed the manuscript.

**Funding statement:**

- This study was supported by the National Natural Science Foundation of China (No. 81871803). The funder had no role in the study design, data collection and analysis, decision to publish, or preparation of the manuscript.No benefits in any form have been received or will be received from a commercial party related directly or indirectly to the subject of this article.

**ICMJE COI statement**

- None declared.

**Acknowledgements**

- We would like to thank Xiaomei Wu for her contribution to this meta-analysis in the field of statistics.

**Ethical review statement**

- None declared.