# Clinical and radiological outcomes of robotic-assisted versus conventional total knee arthroplasty: a systematic review and meta-analysis of randomized controlled trials

Pakpoom RUANGSOMBOON [1,2], Onlak RUANGSOMBOON [3,4], Chaturong PORNRATTANAMANEEWONG [2], Rapeepat NARKBUNNAM [2], and Keerati CHAREANCHOLVANICH [2]

[1] Institute of Health Policy, Management and Evaluation, University of Toronto, Ontario, Canada; [2] Department of Orthopaedic Surgery, Faculty of Medicine, Siriraj Hospital, Mahidol University, Thailand; [3] Department of Health Research Methods, Evidence, and Impact, McMaster University, Ontario, Canada; [4] Department of Emergency Medicine, Faculty of Medicine, Siriraj Hospital, Mahidol University, Thailand
Correspondence: Onlak Ruangsomboon, doctor.mo@yahoo.com
Submitted 2022-10-06. Accepted 2023-01-26.

**Background and purpose** — Robotic-assisted total knee arthroplasty (RATKA) is an alternative surgical treatment method to conventional total knee arthroplasty (COTKA) that may deliver better surgical accuracy. However, its impact on patient outcomes is uncertain. The aim of this systematic review of randomized controlled trials (RCTs) is to evaluate whether RATKA could improve functional and radiological outcomes compared with COTKA in adult patients with primary osteoarthritis of the knee.

**Methods** — We searched Ovid MEDLINE, EMBASE, Scopus, and the Cochrane Library to identify published RCTs comparing RATKA with COTKA. 2 reviewers independently screened eligible studies, reviewed the full texts, assessed risk of bias using the Risk of Bias 2.0 tool, and extracted data. Outcomes were patient-reported outcomes, range of motion, and mechanical alignment (MA) deviation and outliers, and complications.

**Results** — We included 12 RCTs involving 2,200 patients. RATKA probably results in little to no effect on patient-reported outcomes (mean difference (MD) in the Western Ontario and McMaster Universities Osteoarthritis Index (WOMAC) score of –0.35 (95% confidence interval [CI] –0.78 to 0.07) and range of motion (MD –0.73°; CI –7.5° to 6.0°) compared with COTKA. However, RATKA likely results in a lower degree of MA outliers (risk ratio 0.43; CI 0.27 to 0.67) and less deviation from neutral MA (MD –0.94°; CI –1.1° to –0.73°). There were no differences in revision rate or major adverse effects associated with RATKA.

**Conclusion** — Although RATKA likely results in higher radiologic accuracy than COTKA, this may not be clinically meaningful. Also, there is probably no clinically important difference in clinical outcomes between RATKA and COTKA, while it is as yet inconclusive regarding the revision and complication rates due to insufficient evidence.

Osteoarthritis (OA) of the knee is the most common cause of knee pain in the elderly, with an estimated global prevalence of 23% (**1**). This condition leads to significant disability, affecting patients' quality of life. Conventional total knee arthroplasty (COTKA) is usually offered to patients whose conservative treatment has failed. Although the results of COTKA are mostly favorable, surgical issues such as poor soft-tissue tension balancing and malalignment of the prosthesis may negatively impact the outcomes (**2**). Postoperative prosthesis alignment and patient-reported outcome measures (PROMs) are among important outcomes affected by such issues (**3**). Consequently, various techniques and surgical guiding tools have been utilized for these critical surgical steps. Robotic-assisted total knee arthroplasty (RATKA) incorporates navigation technology to visualize preoperative planning. It can help to more precisely locate the hip center, guide bone cutting, and evaluate soft tissue and overall tension (**4**). In a large number of observational studies and their systematic reviews, RATKA has been shown to improve both prosthesis placement accuracy and some PROMs (**5-8**). During the past 4 years, there have been several published randomized controlled trials (RCTs) comparing RATKA with COTKA, yet no systematic review and meta-analysis of RCTs existed. With higher quality of evidence, such reviews can provide higher confidence in the pooled estimates and possibly solidify the synthesized results from observational studies. Therefore, we conducted this systematic review of RCTs to evaluate whether RATKA could improve clinical and functional outcomes, as well as radiological outcomes, compared with COTKA in adult patients with primary knee OA.

## Methods

### Inclusion criteria

This review was conducted and reported following the Cochrane Handbook for Systematic Reviews of Interventions and the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guideline (9,10). We included RCTs comparing RATKA with COTKA in adult patients with primary knee OA. We excluded studies investigating patients with secondary OA and studies of other robotic-assisted surgical interventions that were not RATKA. Studies that did not directly compare RATKA with COTKA were also excluded. We included studies that evaluated outcomes of interest for this review. The main outcomes of this review were clinical and functional outcomes, which are PROMs (the Western Ontario and McMaster University Osteoarthritis Index [WOMAC] score [11], the Knee Society Score [KSS] [12], the Hospital for Special Surgery [HSS] score [13], and the Oxford Knee Score [OKS]) and range of motion (ROM), and radiological outcome assessed by the deviation from neutral mechanical axis (MA) and the proportion of outliers (a deviation of > ± 3°) in MA alignment, which was the same angle as mechanical hip–knee–ankle and mechanical femorotibial angle (14). We focused on MA alignment because it is an important radiographic parameter of knee arthroplasty, designed to promote symmetrical and balanced load distribution between the medial and lateral compartments to decrease component wear and potential loosening (15). Also included as the review outcome was major complication secondary to the interventions, including periprosthetic infection, revision rate, and mortality. Other clinically important outcomes were also assessed where applicable. They were pain, quality of life, intraoperative blood loss, operative time, and hospital length of stay. Studies that did not measure at least one of the main outcomes of interest were excluded.

### Database search methods

We performed a comprehensive search of Ovid MEDLINE (1946 to August 23, 2022), EMBASE (1974 to August 23, 2022), Scopus (1966 to August 23, 2022), and Cochrane Library (1908 to August 23, 2022). Search terms used for the 4 databases are presented as Supplementary material. We did not apply any restrictions as we preferred the search to be sensitive. However, only studies conducted in humans and in English were included. The search results are shown in Figure 1.

### Selection of studies and data extraction

2 reviewers (PR and OR) independently screened the titles and abstracts after removal of duplicates. Articles considered potentially eligible by either of the reviewers were included in the full-text review. The 2 reviewers also reviewed the full texts to confirm the studies' eligibility. Included study data was then also abstracted by 2 data abstractors independently.

They resolved through discussion any discordance in any of the steps. The data abstracted from each study included the first author's identification, year of publication, study setting, study population, details of intervention and its comparator, source of funding, and results for the outcomes of interest.

### Assessment of risk of bias

2 reviewers (OR and CP) independently assessed the 7-domain risks of bias for each study based on the Cochrane Collaboration's tool for assessing risk of bias in randomized trials (16). Moreover, because outcomes from the same study may have different risks of bias, each outcome of the included studies was assessed for risk of bias separately by using the Risk of Bias (RoB) 2.0 tools (10). Their discordance was resolved through consensus.

### Quality of the evidence

2 reviewers (OR and CP) independently judged the quality of evidence related to each main study outcome using the 5 GRADE domains (10). They resolved any disagreements through discussion and provided justifications for any decisions to downgrade the quality of evidence. Interpretation and implications of the pooled estimates and their 95% confidence intervals (CIs) were based on the minimal important differences (MIDs) derived from the literature and approved by experienced knee surgeons (RN and KC). The MIDs for the WOMAC score, KSS, and HSS score were 15 (17), 6 (18), and 13 points (19), respectively, while a 30° difference in ROM (20) and a 3° difference in MA deviation (21) were considered clinically important.

### Statistics

All analyses were performed using the RevMan 5.4.1 (Cochrane Collaboration, Oxford, UK) software. The random-effects models were obtained to conduct the meta-analyses to take into account between-study heterogeneity should this exist. We used the generic inverse variance method and analyzed continuous outcomes as mean difference (MD) and 95% CIs. For dichotomous outcomes, the results were analyzed and presented as risk ratio (RR) and 95% CI. We reported complications only descriptively and did not perform a meta-analysis due to variations in reporting and because no trials reported major adverse events leading to significant morbidity or mortality. For all outcomes, we selected the longest follow-up time-point in each study to include in this review.

We visually inspected the forest plots of the pooled analyses and explored the $I^2$ statistics and the chi-square test to assess signs of heterogeneity (10). For substantial and considerable heterogeneity ($I^2 > 50\%$), we explored sources of heterogeneity by conducting sensitivity and/or subgroup analyses as appropriate. Sensitivity analyses were performed by using the fixed-effects models and excluding studies with high risk of bias to prove the robustness of the results and to explore sources of heterogeneity where applicable. Also, another
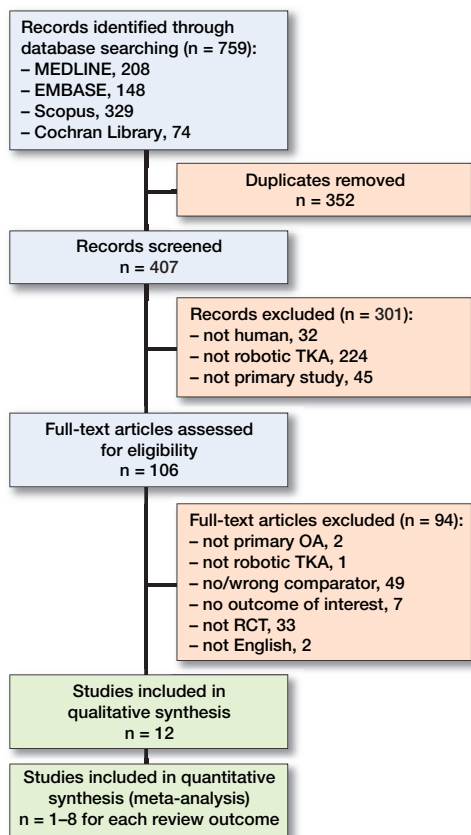
Figure 1. The PRISMA flow chart of study selection and inclusion. TKA = total knee arthroplasty, OA = osteoarthritis, RCT = randomized controlled trial.

sensitivity analysis was performed whereby included studies were only those involving commonly and widely used robotic systems. Moreover, 2 subgroup analyses were conducted, if applicable, based on (i) the mean time of follow-up (< 1 year versus ≥ 1 year) as clinical outcomes might change differently between the 2 study arms at longer follow-up time-points and (ii) the prosthetic type (cruciate-retaining [CR] versus posterior-stabilizing [PS]) as they may deliver different outcome measures.

### Ethics, registration, data sharing, funding, and disclosures
Ethical approval was not required for our study. The protocol was registered in the PROSPERO database (CRD42022358321). All methods were carried out in accordance with relevant guidelines and regulations. Requests for data not shown in the body of this manuscript can be made to the corresponding author. This is an unfunded study. The authors have no competing interests or other interests that might be perceived to influence the results reported in this paper. There is a potential conflict of interest that may influence the discussion because 2 of the authors who supervised the clinical context of the article, RN and KC, are orthopedic surgeons experienced and specialized in robotic knee surgery. All authors declare no financial support from the companies



A. Random sequence generation (selection bias)
B. Allocation concealment (selection bias)
C. Blinding of participants and personnel (performance bias)
D. Blinding of outcome assessment (detection bias)
E. Incomplete outcome data (attrition bias)
F. Selective reporting (reporting bias)
G. Other bias
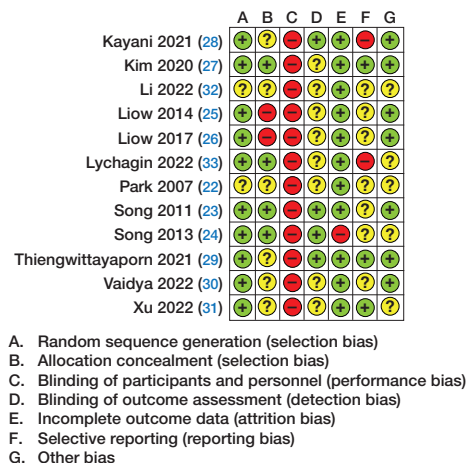
Figure 2. Study-level risk of bias summary.

that produce and/or distribute the drugs, devices, or materials described in this report. Completed disclosure forms for this article following the ICMJE template are available on the article home page, doi: 10.2340/17453674.2023.9411

## Results

### Results of the search and characteristics of included studies
Of the 759 citations retrieved from electronic databases, 407 articles remained after duplicates were removed. We assessed 106 full texts and included 12 eligible RCTs in the final analysis of this review (22-33). Details of the study screening and selection process, as well as reasons for exclusion at each stage, are presented in Figure 1. Characteristics of included studies are presented in Table 1.

All 12 included studies are parallel-group RCTs published during 2007–2020. 10 studies were conducted in Asia: South Korea (n = 4) (22-24,27), China (n = 2) (31,32), Singapore (n = 2) (25,26), Thailand (n = 1) (29), and India (n = 1) (30), and 2 were from Europe: Russia (n = 1) (33), and England (n = 1) (28). Their sample sizes ranged from 30 to 1,406, with participants' mean age ranging from 60 to 77 years. 1 study (26) was a secondary analysis of the previous study by the same author group (25) but at a longer follow-up time. Therefore, we obtained outcomes only from Liow et al., 2017 (26) for the meta-analyses to avoid sample duplications. ROBODOC (Curexo Technology, Fremont, CA, USA) was the most common robotic system employed (n = 6) (22-27), followed by NAVIO (Smith & Nephew, Memphis, TN, USA) (n = 2) (29,30) and Mako (Stryker, Portage, MI, USA) (n = 1) (28). The 3 other studies from China and Russia utilized locally invented robotic apparatus (31-33). More studies reported using PS (22,25–27,29-31) than CR prostheses (23,24,28). All studies employed neutral mechanical alignment concept, except Kayani et al. (28), who focused on functional align-

Table 1. Characteristics of included trials

| Author, year | Study setting | Population | Robot system/ Type [b] | Knee alignment concept | Surgical technique | Detail of prosthesis |
|---|---|---|---|---|---|---|
| Park 2007 (22) | South Korea, single center, 2 surgeons | N = 62 – RATKA: n = 32 mean age 62.7 – COTKA: n = 30 mean age 67.8 | ROBODOC with ORTHODOC (Curexo Technology, USA) / ARS | Mechanical alignment both groups | Conventional technique for soft tissue balancing | Zimmer LPS (PS) – No details of patella procedure |
| Song 2011 (23) | South Korea, Single center, 1 surgeon | N = 30 (60 knees) mean age 67 – RATKA: n = 30 – COTKA: n = 30 | ROBODOC with ORTHODOC (Curexo Technology, USA) / ARS | Mechanical alignment both groups | Tensor device (Stryker Howmedica Osteonics) for soft tissue balancing | Zimmer NexGen (CR) – No details of patella procedure |
| Song 2013 (24) | South Korea, Single center, 1 surgeon | N = 100 – RATKA: n = 50 mean age 66.1 – COTKA: n = 50 mean age 64.8 | ROBODOC with ORTHODOC (Curexo Technology, USA) / ARS | Mechanical alignment both groups | Measure resection technique | Zimmer NexGen (CR) – No details of patella procedure |
| Liow 2014 (25) | Singapore, single center, 1 surgeon | N = 60 – RATKA: n = 31 mean age 67.5 – COTKA: n = 29 mean age 68.3 | ROBODOC with ORTHODOC (Curexo Technology, USA) / ARS | Mechanical alignment both groups | Measure resection technique and soft tissue balancing | Zimmer NexGen LPS-Flex (PS) – Patelloplasty |
| Liow 2017 [a] (26) | Singapore, single center, 1 surgeon | N = 60 – RATKA: n = 31 mean age 67.5 – COTKA: n = 29 mean age 68.3 | ROBODOC with OR-THODOC (Curexo Technology, USA) / ARS | Mechanical alignment both groups | Measure resection technique and soft tissue balancing | Zimmer NexGen LPS-Flex (PS) – Patelloplasty |
| Kim 2020 (27) | South Korea, single center, 1 surgeon | N = 1,406 – RATKA: n = 700 (750 knees), mean age 60 – COTKA: n = 706 (766 knees) mean age 61 | ROBODOC with ORTHODOC (Curexo Technology, USA) / ARS | Mechanical alignment (neutral mechanical axis) both groups | Measure resection technique | Stryker Duracon (PS) – No resurfacing patella |
| Kayani 2021 (28) | England, single center, 2 surgeons | N = 30 – RATKA: n = 15 mean age 68.7 – COTKA: n = 15 mean age 67.9 | Mako (Mako Surgical Corp, Stryker, USA) / HRS | RATKA: functional COTKA – mechanical alignment | Soft tissue balancing | Triathlon cruciate-retaining knee (Stryker) (CR) – Patella resurfacing |
| Thieng-wittaya-porn 2021 (29) | Thailand, single center, 1 surgeon | N = 152 – RATKA: n = 75 mean age 69 – COTKA: n = 77 mean age 69.1 | NAVIO (Smith &Nephew, USA) / HRS | Mechanical alignment both groups | Soft tissue balancing using Robotic record | Legion PS total knee (Smith&Nephew) – Non-resurfacing patella |
| Vaidya 2022 (30) | India, single center | N = 60 – RATKA: n = 30 mean age 62.2 – COTKA: n = 30 mean age 59.9 | NAVIO (Smith &Nephew, USA) / HRS | Mechanical alignment both groups | Measure resection technique | Anthem PS prosthesis (Smith&Nephew) – No details of patella procedure |
| Xu 2022 (31) | China, single center, 2 surgeons | N = 72 – RATKA: n = 37 mean age 64.5 – COTKA: n = 35 mean age 63.4 | YUANHUA-TKA (YUANHUA-TKA, Shenzhen, China) / unknown type | Mechanical alignment both groups | Measure resection technique and soft tissue release if necessary | Unique knee, fixed platform and PCL sacrificing (PS) – No details of patella procedure |
| Li 2022 (32) | China, multicenter, multisurgeons | N = 150 – RATKA: n = 73 mean age 68 – COTKA: n=77 mean age 69 | HURWA system (HURWA, BEIJING HURWA-ROBOT Technology Co. LTd) / unknown type | Mechanical alignment both groups | No details of soft tissue release technique | Legion (Smith&Nephew) – Constraint unknown – No details of patella procedure |
| Lychagin 2022 (33) | Russia, single center | N = 68 mean age 67.3 – RATKA: n = 33 – COTKA: n = 35 | TSolution-One (THINK Surgical Inc.) / ARS | Mechanical alignment both groups | No details of soft tissue release technique | Not reported – No details of patella procedure |

Table 1 continued

| Author year | Follow-up time | Outcomes | Conflict of interest |
|---|---|---|---|
| Park 2007 | RATKA: mean 45 months COTKA: mean 49.3 months | – Knee society score (KSS), Knee functional score (KFS), range of motion (ROM),Tibiofemoral angle (TFA)<br>– Radiographic results; AP and lateral femoral flexion angle (FFA) and tibial angle (TA)<br>– Complication in RATKA | None declared |
| Song 2011 | Mean 16 months (minimum 12 months) | – Clinical assessment: ROM, Hospital for Special Surgery (HSS) score, Western Ontario and McMaster University score (WOMAC)<br>– Radiographic assessment: mechanical axis (MA), coronal and sagittal inclination on femoral and tibial side<br>– Complication, operative time | None declared |
| Song 2013 | Mean 65 months (minimum 41 months) | – Clinical assessment: ROM, HSS, WOMAC<br>– Radiographic assessment: postop. MA, coronal and sagittal inclination on femoral and tibial side<br>– Complication, operative time | Some authors reported receiving benefits from Curexo Tech. |
| Liow 2014 | 6 months | – Clinical assessment: KSS, KFS, Oxford knee score (OKS), SF-36, ROM at 6 months<br>– Radiographic assessment: postop. MA, AP femoral flexion and tibia angle, Lateral femoral flexion and tibia angle at 1 month<br>– Complication, operative time, length of stay (LOS) | None declared |
| Liow 2017 [a] | 24 months | – Clinical assessment: KSS, KFS, OKS, SF-36, ROM, satisfaction<br>– Complication | None declared |
| Kim 2020 | RATKA: mean 13 years COTKA: mean 14 years | – Clinical assessment: KSS, WOMAC, pain VAS, UCLA activity score, ROM<br>– Radiographic assessment: TFA, femoral and tibial component position and outliers (coronal and sagittal plane), aseptic loosening, rotational alignment<br>– Survivorship, complication, operative time | None declared |
| Kayani 2021 | 28 days | – Inflammatory markers<br>– Radiographic assessment: accuracy of achieving limb alignment and component positioning<br>– Complication, operative time | None declared |
| Thieng-wittaya-porn 2021 | 6 weeks | – Radiographic assessment: hip-knee-ankle (HKA) angle, TFA, coronal femoral and tibial component angle, sagittal femoral and tibial component angle at 6 weeks<br>– Operative time | Funding: institutional support |
| Vaidya 2022 | Postoperation | – Radiographic assessment: MA, coronal femoral and tibial alignment, sagittal femoral and tibial alignment, femoral component rotation | None declared |
| Xu 2022 | 90 days | – Clinical assessment: KSS, WOMAC, ROM<br>– Radiographic assessment: HKA angle, frontal femoral and tibial component angle, lateral femoral and tibial component angle<br>– Complication, operative time, LOS | None declared |
| Li 2022 | Mean 3 months | – Clinical assessment: ROM, HSS, KSS, WOMAC, SF-36<br>– Radiographic assessment: HKA angle | Funding: Beijing and China science institutions |
| Lychagin 2022 | 1 year | – Clinical assessment: ROM, pain VAS, OKS, WOMAC<br>– Radiographic assessment: mechanical axis | None declared |

ment. 3 studies only focused on radiographic results (**28-30**), while the rest reported both clinical and radiological outcomes. The mean follow-up time of the studies with clinical outcomes ranged from 3 months to 13 years. 2 studies received funding from governmental and non-profit organizations (**29,32**). Another study had a potential financial conflict of interest with the robot manufacturer (**24**).

## Risk of bias in included studies

Our judgement on the study-level risk of bias is presented in Figure 2 with justifications for our decision provided in Table 4 (see Supplementary material). All studies had a high risk of performance bias because the intervention cannot be blinded to the providers. Most studies failed to report their method of allocation concealment and whether assessors of some of the outcomes were blinded. Moreover, only 5 studies had their trials registered and can be assessed for reporting bias (**27-29,31,33**).

The overall risk of bias for each study outcome is presented in Table 2. We also presented the outcome-level risk of bias of each study with the forest plot for that outcome.

## Effects of interventions

The summary of findings for all outcomes, including the certainty of evidence and reasons for downgrading the evidence, are presented in the "summary of findings" table (see Appendix) and Table 5 (see Supplementary material), respectively.

Table 1. Footnotes

Abbreviations: RATKA = robotic-assisted total knee arthroplasty, COTKA = conventional total knee arthroplasty, HRS = haptic robotic systems, ARS = autonomous robotic systems, CR = cruciate-retaining, PS = posterior stabilizing
[a] secondary analysis of Liow, 2014
[b] The type of each robotic system was reported based on Lang J E, Mannava S, Floyd A J, Goddard M S, Smith B P, Mofidi A, et al. Robotic systems in orthopaedic surgery. J Bone Joint Surg Br 2011; 93(10): 1296-9.

Table 2. Risk of bias of included trials (outcome-level)

| Study | | Domain 1: Randomization process | Domain 2: Deviation from intervention | Domain 3: Missing outcome data | Domain 4: Measurement of outcome | Domain 5: Reported results | Overall risk of bias per outcome |
|---|---|---|---|---|---|---|---|
| Park, 2007 | KSS, ROM | Some [a] | Low | Low | Some [b] | Some [c] | Some concerns |
| | Radiographic results, complication | Some [a] | Low | Low | Low | Some [c] | Some concerns |
| Song, 2011 | HSS, WOMAC | Low | Low | Low | Some [b] | Some [c] | Some concerns |
| | Radiographic results, ROM, complication | Low | Low | Low | Low | Some [c] | Some concerns |
| Song, 2013 | HSS, WOMAC | Low | Low | High [d] | Some [b] | Some [c] | High |
| | ROM, complication | Low | Low | Low | Low | Some [c] | Some concerns |
| | Radiographic results | Low | Low | High [d] | Low | Some [c] | High |
| Liow, 2014 | KSS, ROM | High [e] | Low | Low | Some [b] | Some [c] | High |
| | Radiographic results, complication | High [e] | Low | Low | Low | Some [c] | High |
| Liow, 2017 | KSS, ROM | High [e] | Low | Low | Some [b] | Some [c] | High |
| | Complication | High [e] | Low | Low | Low | Some [c] | High |
| Kim, 2020 | KSS, WOMAC | Low | Low | Low | Some [b] | Low | Some concerns |
| | Radiographic results, ROM, complication | Low | Low | Low | Low | Low | Low |
| Kayani, 2021 | Radiologic results | Some [a] | Low | Low | Low | High [f] | High |
| Thiengwittaya-porn, 2021 | Radiologic results | Some [a] | Low | Low | Low | Low | Some concerns |
| Vaidya, 2022 | Radiologic results | Some [a] | Low | Low | Some [b] | Some [c] | Some concerns |
| Xu, 2022 | Radiographic results, complication | Some [a] | Low | Low | Low | Low | Some concerns |
| | KSS, WOMAC, ROM | Some [a] | Low | Low | Some | Low | Some concerns |
| Li, 2022 | HSS, KSS, WOMAC | Some [a] | Low | Low | Some [b] | Some [c] | Some concerns |
| | Radiographic results, ROM | Some [a] | Low | Low | Low | Some [c] | Some concerns |
| Lychagin, 2022 | WOMAC | Low | Low | Low | Low | High [g] | High |
| | Radiographic results, ROM | Low | Low | Low | Some [b] | High [g] | High |

Abbreviations: KSS = Knee Society Score, ROM = range of motion, HSS = Hospital for Special Surgery, WOMAC = Western Ontario and McMaster University Osteoarthritis Index.
[a] Some concern – allocation concealment method not provided.
[b] Some concern – patient-reported outcome or outcome assessed by healthcare assessors with no information if the intervention was blinded, or unblinded intervention but without objective evidence of influence on the outcomes.
[c] Some concern – no protocol available.
[d] High rate of missing outcome.
[e] Unconcealed allocation.
[f] Protocol did not include radiologic results as study outcomes and they may be subject to multiple analyses and selective reporting.
[g] Not all secondary outcomes in the protocol were reported, thereby having a risk of selective reporting

### PROMs: WOMAC, KSS, HSS, and OKS

6 studies reported the WOMAC score in 1,744 participants (23,24,27,31-33). As shown in Figure 3A, the WOMAC score was on average 0.35 points lower (CI –0.78 to 0.07, moderate certainty) in the RATKA than in the COTKA group ($I^2$ = 3%, p = 0.4). Sensitivity analyses found the point estimates to be in the same direction, with the 95% CIs all excluding the MID (Figures 6–8, see Appendix). 5 studies, including 1,685 patients, reported the KSS (22,26,27,31,32). Compared with COTKA, RATKA had 0.89 points higher KSS on average (CI –0.18 to 2.0, moderate certainty) as presented in Figure 3B ($I^2$ = 25%, p = 0.3). The sensitivity analysis with commonly used robotic systems found similar results (Figure 7, see Appendix); however, the analysis excluding a study with high risk of bias found both the point estimate and the 95%CI favoring COTKA (Figure 6, see Appendix), while the result from the analysis employing the fixed-effects model favored RATKA

(Figure 8, see Appendix). For HSS score, the combined results from 3 studies (23,24,32) of 256 participants with no heterogeneity ($I^2$ = 0%, p = 1) found an average of 0.67 HSS points higher in the RATKA group (CI –0.99 to 2.3, moderate certainty) (Figure 3C). The 3 sensitivity analyses gave similar results (Figures 6–8, see Appendix). There were no subgroup effects based on follow-up time or prosthesis type for any of these PROMs (Figures 9–10, see Appendix). 2 studies (25,33) with 128 participants reported different versions of the OKS, one with lower scores referring to better outcomes and the other one in the other direction; thus, they were not meta-analyzed. Regardless, they both found no between-group difference in the OKS.

### Range of motion

Results from 8 studies (22-24,26,27,31-33) with 2,013 patients were combined with considerable heterogeneity ($I^2$ = 99%, p

**A**

| WOMAC Study | RATKA mean (SD) | n | Conventional TKA mean (SD) | n | Weight % | Difference mean (95%CI) |
|---|---|---|---|---|---|---|
| Song 2011 (23) | 11 (4.5) | 30 | 13 (6.6) | 30 | 2.1 | −2.0 (−4.9 to 0.86) |
| Song 2013 (24) | 29 (4.4) | 29 | 30 (7.5) | 24 | 1.5 | −1.1 (−4.5 to 2.3) |
| Kim 2020 (27) | 18 (14) | 674 | 19 (15) | 674 | 7.0 | −1.0 (−2.6 to 0.55) |
| Lychagin 2022 (33) | 1.3 (0.5) | 33 | 1.6 (0.6) | 35 | 89.2 | −0.25 (−0.51 to 0.01) |
| Li 2022 (32) | 81 (46) | 69 | 72 (44) | 74 | 0.1 | 8.5 (−6.4 to 23) |
| Xu 2022 (31) | 13 (13) | 37 | 19 (31) | 35 | 0.1 | −6.9 (−18 to 4.3) |
| **Total** | | **872** | | **872** | | **−0.35 (−0.78 to 0.07)** |

Heterogeneity: Tau = 0.03; Chi² = 5.2, df = 5 (p = 0.4); I² = 3%
Test for overall effect Z = 1.7 (p = 0.1)

**B**

| KSS Study | Conventional TKA mean (SD) | n | RATKA mean (SD) | n | Weight % | Difference mean (95%CI) |
|---|---|---|---|---|---|---|
| Park 2007 (22) | 91 (4.9) | 30 | 92 (2.9) | 32 | 20.3 | −0.70 (−2.7 to 1.3) |
| Liow 2017 (25) | 88 (11) | 29 | 82 (15) | 31 | 2.6 | 6.1 (−0.41 to 13) |
| Kim 2020 (27) | 92 (6.0) | 674 | 93 (5.0) | 674 | 59.0 | −1.0 (−1.6 to −0.41) |
| Li 2022 (32) | 68 (10) | 74 | 70 (7.3) | 69 | 11.8 | −2.2 (−5.1 to 0.67) |
| Xu 2022 (31) | 66 (7.4) | 35 | 67 (10) | 37 | 6.3 | −1.0 (−5.1 to 3.1) |
| **Total** | | **842** | | **843** | | **−0.89 (−2.0 to 0.18)** |

Heterogeneity: Tau = 0.42; Chi² = 5.3, df = 4 (p = 0.3); I² = 25%
Test for overall effect Z = 1.6 (p = 0.1)

**C**

| HSS score Study | Conventional TKA mean (SD) | n | RATKA mean (SD) | n | Weight % | Difference mean (95%CI) |
|---|---|---|---|---|---|---|
| Song 2011 (23) | 95 (5.0) | 30 | 95 (4.0) | 30 | 52.6 | −0.50 (−2.8 to 1.8) |
| Song 2013 (24) | 95 (6.7) | 24 | 96 (4.0) | 29 | 29.7 | −1.0 (−4.1 to 2.1) |
| Li 2022 (32) | 79 (12) | 74 | 80 (12) | 69 | 17.8 | −0.65 (−4.6 to 3.3) |
| **Total** | | **128** | | **128** | | **−0.67 (−2.3 to 0.99)** |

Heterogeneity: Tau = 0.00; Chi² = 0.07, df = 2 (p = 0.97); I² = 0%
Test for overall effect Z = 0.8 (p = 0.4)

**D**

| ROM Study | Conventional TKA mean (SD) | n | RATKA mean (SD) | n | Weight % | Difference mean (95%CI) |
|---|---|---|---|---|---|---|
| Park 2007 (22) | 122 (17) | 30 | 118 (9.0) | 32 | 11.9 | 4.0 (−2.8 to 11) |
| Song 2011 (23) | 129 (13) | 30 | 129 (14) | 30 | 12.0 | 0.0 (−6.7 to 6.7) |
| Song 2013 (24) | 129 (12) | 50 | 128 (5.1) | 50 | 13.0 | 1.0 (−2.7 to 4.7) |
| Liow 2017 (26) | 126 (10) | 29 | 117 (16) | 31 | 12.0 | 8.2 (1.6 to 15) |
| Kim 2020 (27) | 128 (7.0) | 724 | 125 (6.0) | 724 | 13.5 | 3.0 (2.3 to 3.7) |
| Xu 2022 (31) | 120 (19) | 35 | 115 (11) | 37 | 11.8 | 5.0 (−2.1 to 12) |
| Li 2022 (32) | 112 (19) | 74 | 115 (18) | 69 | 12.2 | −2.9 (−9.0 to 3.2) |
| Lychagin 2022 (33) | 116 (1.3) | 35 | 127 (2.1) | 33 | 13.5 | −11 (−12 to −10) |
| **Total** | | **1,007** | | **1,006** | | **0.73 (−6.0 to 7.5)** |

Heterogeneity: Tau = 88; Chi² = 685, df = 7 (p < 0.001); I² = 99%
Test for overall effect Z = 0.21 (p = 0.8)

**E**

| Implant survival Study | RATKA mean (SD) | n | Conventional TKA mean (SD) | n | Weight % | Risk ratio (95%CI) |
|---|---|---|---|---|---|---|
| Kim 2020 (27) | 15 | 750 | 15 | 766 | 100.0 | 1.0 (0.5 to 2.1) |
| **Total** | **15** | **750** | **15** | **766** | | **1.0 (0.5 to 2.1)** |

Heterogeneity: Not applicable
Test for overall effect Z = 0.06 (p = 0.95)

Figure 3. Forest plots showing the Western Ontario and McMaster University Osteoarthritis Index scores (A), Knee Society Scores (B), Hospital for Special Surgery scores (C), range of motion (D), and implant survivorship (E) between robotic-assisted total knee arthroplasty (RATKA) and conventional total knee arthroplasty (conventional TKA) using random-effects model and inverse variance method. SD = standard deviation, CI = confidence interval. Note: outcome-level risk of bias legend (A) randomization process, (B) deviation from intended interventions, (C) missing outcome data, (D) measurement of outcome, and (E) selective outcome reporting.

< 0.001). ROM for RATKA was on average 0.73° lower than that of the COTKA group (CI −7.5° to 6.0°, low certainty) (Figure 3D). The sensitivity analysis with the fixed-effects model gave both the point estimate and the 95% CI favor-ing RATKA (Figure 8, see Appendix), while the results of the other 2 sensitivity analyses were in the opposite direction (Figures 6–7, see Appendix). The large degree of uncertainties disappeared in the 2 latter sensitivity analyses (I² decreased to

**A**

| MA deviation Study | RATKA mean (SD) | n | Conventional TKA mean (SD) | n | Weight % | Difference mean (95%CI) |
|---|---|---|---|---|---|---|
| Song 2011 (23) | 0.2 (1.6) | 30 | 1.2 (2.1) | 30 | 4.8 | −1.0 (−1.9 to −0.06) |
| Song 2013 (24) | 0.5 (1.4) | 50 | 1.2 (2.9) | 50 | 5.4 | −0.70 (−1.6 to 0.19) |
| Liow 2014 (25) | 1.3 (0.9) | 31 | 1.8 (1.2) | 29 | 14.8 | −0.50 (−1.0 to 0.04) |
| Kim 2020 (27) | 2.0 (2.0) | 724 | 3.0 (3.0) | 724 | 62.3 | −1.0 (−1.3 to −0.74) |
| Li 2022 (32) | 1.8 (1.6) | 69 | 3.0 (2.7) | 74 | 8.0 | −1.2 (−2.0 to −0.49) |
| Vaidya 2022 (30) | 1.8 (1.2) | 30 | 3.0 (2.4) | 30 | 4.7 | −1.2 (−2.2 to −0.24) |
| **Total** | | **934** | | **937** | | **−0.94 (−1.1 to −0.73)** |

Heterogeneity: Tau = 0.00; Chi² = 3.9, df = 5 (p = 0.6); I² = 0%
Test for overall effect Z = 8.9 (p < 0.001)

Difference mean (95%CI) — Favors RATKA / Favors conventional TKA

Risk of bias A B C D E

**B**

| MA outlier (> 3°) Study | RATKA events | n | Conventional TKA events | n | Weight % | Risk ratio (95%CI) |
|---|---|---|---|---|---|---|
| Song 2011 (23) | 0 | 30 | 7 | 30 | 2.4 | 0.07 (0.00 to 1.1) |
| Song 2013 (24) | 0 | 50 | 12 | 50 | 2.5 | 0.04 (0.00 to 0.66) |
| Liow 2014 (25) | 0 | 31 | 4 | 29 | 2.3 | 0.10 (0.01 to 1.9) |
| Kim 2020 (27) | 29 | 724 | 43 | 724 | 31.6 | 0.67 (0.43 to 1.1) |
| Thiengwittayaporn 2021 (29) | 4 | 75 | 12 | 77 | 12.6 | 0.34 (0.12 to 1.0) |
| Li 2022 (32) | 13 | 69 | 27 | 74 | 26.7 | 0.52 (0.29 to 0.92) |
| Vaidya 2022 (30) | 1 | 30 | 8 | 30 | 4.5 | 0.13 (0.02 to 0.94) |
| Xu 2022 (31) | 6 | 35 | 12 | 34 | 17.4 | 0.49 (0.21 to 1.2) |
| **Total** | | **1,044** | | **1,048** | | **0.43 (0.27 to 0.67)** |

Heterogeneity: Tau = 0.11; Chi² = 10, df = 7 (p = 0.2); I² = 31%
Test for overall effect Z = 3.7 (p < 0.001)

Risk ratio (95%CI) — Favors RATKA / Favors conventional TKA

Risk of bias A B C D E

Figure 4. Forest plots showing deviation from neutral mechanical axis (A) and the proportion of outliers (deviation > 3°) in the mechanical axis alignment (B) between robotic-assisted total knee arthroplasty (RATKA) and conventional total knee arthroplasty (conventional TKA) using random-effects model and inverse variance method. For abbreviations and Risk of Bias A–E, see Figure 3.

**A**

| Blood loss Study | RATKA mean (SD) | n | Conventional TKA mean (SD) | n | Weight % | Difference mean (95%CI) |
|---|---|---|---|---|---|---|
| Song 2011 (23) | 587 (385) | 30 | 816 (425) | 30 | 22.2 | −229 (−435 to −24) |
| Song 2013 (24) | 613 (318) | 50 | 933 (467) | 50 | 25.1 | −320 (−477 to −163) |
| Kim 2020 (27) | 261 (107) | 700 | 255 (110) | 706 | 30.7 | 6.0 (−5.4 to 17) |
| Xu 2022 (31) | 933 (454) | 37 | 863 (458) | 35 | 21.9 | 70 (−141 to 280) |
| **Total** | | **817** | | **821** | | **−114 (−298 to 69)** |

Heterogeneity: Tau = 28,471; Chi² = 22, df = 3 (p < 0.001); I² = 86%
Test for overall effect Z = 1.2 (p = 0.2)

Difference mean (95%CI) — Favors RATKA / Favors conventional TKA

Risk of bias A B C D E

**B**

| Operative time Study | RATKA mean (SD) | n | Conventional TKA mean (SD) | n | Weight % | Difference mean (95%CI) |
|---|---|---|---|---|---|---|
| Kayani 2021 (28) | 61 (3.1) | 15 | 62 (3.4) | 15 | 14.7 | −1.0 (−3.3 to 1.3) |
| Kim 2020 (27) | 97 (11) | 700 | 69 (6.3) | 706 | 14.8 | 28 (27 to 29) |
| Liow 2017 (36) | 91 (10) | 31 | 93 (14) | 29 | 14.3 | −2.0 (−8.2 to 4.2) |
| Song 2011 (23) | 95 (18) | 30 | 70 (15) | 30 | 13.9 | 25 (17 to 33) |
| Song 2013 (24) | 99 (11) | 50 | 74 (10) | 50 | 14.6 | 25 (21 to 29) |
| Thiengwittayaporn 2021 (29) | 70 (12) | 75 | 62 (10) | 77 | 14.6 | 8.2 (4.7 to 12) |
| Xu 2022 (31) | 154 (21) | 37 | 115 (30) | 35 | 13.0 | 39 (27 to 51) |
| **Total** | | **938** | | **942** | | **17 (4.7 to 30)** |

Heterogeneity: Tau = 273; Chi² = 663, df = 6 (p < 0.001); I² = 99%
Test for overall effect Z = 2.7 (p = 0.007)

Difference mean (95%CI) — Favors RATKA / Favors conventional TKA

Risk of bias A B C D E

**C**

| Hospital stay Study | RATKA mean (SD) | n | Conventional TKA mean (SD) | n | Weight % | Difference mean (95%CI) |
|---|---|---|---|---|---|---|
| Liow 2017 (36) | 5.2 (2.3) | 31 | 5.8 (3.8) | 29 | 37.1 | −0.60 (−2.2 to 1.0) |
| Xu 2022 (31) | 9.1 (2.0) | 37 | 8.4 (2.1) | 35 | 62.9 | 0.70 (−0.25 to 1.7) |
| **Total** | | **68** | | **64** | | **0.22 (−1.0 to 1.5)** |

Heterogeneity: Tau = 0.39; Chi² = 1.9, df = 1 (p = 0.2); I² = 47%
Test for overall effect Z = 0.35 (p = 0.7)

Difference mean (95%CI) — Favors RATKA / Favors conventional TKA

Risk of bias A B C D E

Figure 5. Forest plots showing intraoperative blood loss (A), operative time (B), and hospital length of stay (C) between robotic-assisted total knee arthroplasty (RATKA) and conventional total knee arthroplasty (conventional TKA) using random-effects model and inverse variance method. For abbreviations and Risk of Bias A–E, see Figure 3.

7–12%), which was most likely because 1 study (33) with an extremely discordant result was excluded from both sensitivity analyses. No subgroup differences were seen either by follow-up time or prosthetic type (Figures 9–10, see Appendix).

### Mechanical alignment deviation and outliers

6 RCTs involving 1,871 participants (23-25,27,30,32) found the deviation from neutral MA of 0.94° lower in RATKA (CI –1.1° to –0.73°, moderate certainty) with minimal heterogeneity ($I^2$ = 0%, p = 0.6) and all sensitivity analyses delivering concordant results (Figures 6–8, see Appendix). Moreover, pooled results from 2,092 participants from 8 studies (23,24,26,27,29-32) showed that the risk of MA outliers with RATKA was 0.43 times the risk with COTKA (CI 0.27 to 0.67, moderate certainty, $I^2$ = 31%, p = 0.2), which translates to a number needed to treat of 45 (CI 36 to 77). The sensitivity analyses found similar results with a similar degree of heterogeneity. There was a significant subgroup interaction (p = 0.03) based on the prosthetic used, with both subgroup estimates favoring RATKA but with a stronger effect in the CR subgroup (Figure 10, see Appendix).

### Complications

7 studies of 1,772 patients reported complications associated with RATKA, 2 of which (22,23) reported only adverse effects of RATKA. The details of the complications are reported descriptively in Table 3 (see Appendix). No studies reported the rate of periprosthetic infection, severe morbidity or mortality secondary to the interventions. Only 1 study (27) reported long-term implant survivorship at 15 years, defined as any revision or aseptic loosening. They found no significant between-group difference (98% survivorship in both groups; RR 1.02; CI 0.50 to 2.1, low certainty) (Figure 3E).

### Other outcomes

4 studies with 1,638 participants (23,24,27,31) found an overall 114 mL lower blood loss in favor of RATKA (CI –298 to 69 mL, $I^2$ = 86%) (Figure 5A). 7 studies with 1,880 participants (23-25,27-29,31) showed higher operative time for RATKA compared with COTKA (MD 17 minutes; CI 5 to 30, $I^2$ = 99%) (Figure 5B). Hospital length of stay was similar between the 2 interventions based on 2 studies with 102 participants (25,31) (MD 0.22 days; CI –1.0 to 1.5, $I^2$ = 47%) (Figure 5C). Sensitivity analyses found all results in the same direction as the primary ones (Figures 6–8, see Appendix).

Some of the proposed review outcomes were not meta-analyzed. Only 1 study (33) reported pain scale as a continuous outcome (lower pain scale in the RATKA group at 12 days); thus, a meta-analysis was not possible. 2 studies (26,32) reported quality of life assessed with SF-36; however, they were not combined because they reported summary statistics of different sub-categories of SF-36. Nonetheless, both studies delivered similar results with insignificant differences between the 2 groups.

## Discussion

From this systematic review and meta-analysis, RATKA probably results in little to no difference in the WOMAC and KSS scores compared with COTKA. The evidence also suggests that RATKA may not increase the HSS score or ROM. However, RATKA likely results in a significant reduction in the proportion of MA outliers compared with COTKA, although it probably does not reduce MA deviation to an important effect size. Complication rates were generally similar with no major complications reported and no difference in the rate of implant survivorship based on 1 long-term study. Overall, RATKA may provide more accuracy regarding radiological alignment; however, it did not demonstrate superior clinical or functional effects.

There were no statistically significant differences in WOMAC and KSS between the 2 interventions with moderate certainty and the 95% CIs excluding MIDs. Although sensitivity analyses found variability in the results, all estimates and their uncertainties help confirm that there probably were no clinically meaningful differences. For HSS, the quality of evidence was also at the moderate level with concordant results among included studies, although optimal information size was not met. The summarized results from all analyses on HSS were concordant with WOMAC and KSS, showing no evidence of important effects. Our pooled results of PROMs were similar to previous meta-analyses of observational studies with overall non-clinically important differences between RATKA and COTKA (5,7).

The primary analysis found an insignificant increase in ROM favoring COTKA. Sensitivity analyses found variable results, with some in favor of RATKA and others of COTKA. Nevertheless, despite conflicting results between various analyses, all 95% CIs were still largely lower than the MID, confirming that there were probably no important differences in ROM between the 2 interventions. Nonetheless, we should keep in mind when interpreting the results of these functional outcome parameters that there probably was a ceiling effect causing them to have low discriminatory power (34,35).

Our meta-analysis of RCTs found a largely lower proportion of MA outliers in RATKA compared with COTKA, which was concordant with previous meta-analyses of observational studies (5,7). However, we also found that the mean difference in MA deviation, despite the angle being significantly lower in RATKA, failed to meet the MID of 3°. This discordance between the 2 analyses could have been because some trials did not report the angle itself but only presented the proportion of outliers, thus affecting the amount of raw information for the continuous outcome analysis. Therefore, it is still unclear whether RATKA is superior to COTKA in delivering more accurate and clinically meaningful radiologic results. Regardless, it is important to note that the large effect size for MA

outliers translates to a risk difference of only 22 per 1,000 and a number needed to treat of 45, which may be considered non-clinically important. In fact, it is questionable whether the 3° difference or any difference in MA deviation is at all important, as many previous studies have questioned its association with long-term functional outcomes. Some studies and reviews have reported that MA deviation was not significantly associated with long-term functional outcomes and implant survivorship; however, the quality of evidence was low due to a limited number of studies, small sample sizes, and other study limitations (36,37). Consequently, further studies are still required to answer whether RATKA could deliver better radiological outcomes and if such radiographic accuracy leads to important clinical outcomes.

Nevertheless, we found a subgroup effect based on the type of implant with the CR prosthesis yielding a larger treatment effect than PS. However, this interaction could have also been due to scarce events in the studies that employed a CR prosthesis, with the 95% CI of the effect too wide to draw any meaningful conclusion. Therefore, we cannot be certain that there was a true difference in the outcome between the subgroups from this meta-analysis. With an ongoing debate on whether the type of prosthesis has any impact on patient outcomes, further studies and future reviews are required to deliver a more definitive answer.

No study reported mortality, and only one reported long-term implant survivorship, which was probably because this review was conducted at an early stage of the development of evidence, and trials with long follow-up time are most likely ongoing. Therefore, it is still inconclusive and not yet appropriate to conclude whether there were differences regarding major adverse effects. Moreover, we found that RATKA may increase the operative time and decrease intraoperative blood loss, but the evidence is very uncertain due to study bias, inconsistency, and imprecision. Also, the procedural time and amount of bleeding are operator-dependent and depend on the surgeons' learning curve. For hospital length of stay, the amount and quality of evidence was too low to determine any conclusion. Similarly, there is not enough evidence regarding other core outcome domains of patients with OA of the knee (38), namely quality of life, pain as a separate modality, and patients' global assessment of the target joint, to allow for the proper conduct of a meta-analysis.

With all things considered, there has not been enough high-quality evidence supporting the use of RATKA, despite its widespread use in many institutions. In fact, the expense of investing in this technology is extremely substantial, with an average added annualized cost ranging from 4,000 to 71,000 US dollars depending on surgical volume (39,40). Therefore, we should be concerned not only about the clinical effectiveness but also the cost-effectiveness of this intervention. Economic decision analyses based on appropriate perspectives are mandatory to help decide whether one should incorporate such a costly intervention into clinical practice.

This review has some limitations. Although we could obtain a satisfactory number of included trials, most of them were small studies with generally moderate to high risk of bias. They also employed robotic apparatus from different manufacturers that may aim for and provide different outcomes. Furthermore, these robotic systems were from different calendar years; thus, there may be differences among the interventions following technological evolution that might have been a source of heterogeneity among the studies included. Regardless, the random-effects models were obtained to conduct the meta-analyses to capture this potential between-study heterogeneity. Also, most studies were from Asian countries, potentially limiting the generalizability of this review's results to other settings. Additionally, the generalizability may also be limited to only the robotic procedure aimed at neutral MA, as all trials included in the analysis employed the concept of mechanical rather than functional alignment. Moreover, we could not identify many reports on many important clinical outcomes, such as long-term survivorship and quality of life. Some review outcomes and sensitivity analyses were analyzed from less than 5 studies. The small number of studies and combined participants could have led to imprecise and biased measures of consistency and estimates of treatment effects. Also, we could not properly create funnel plots to assess reporting bias due to a small number of studies for the review outcomes. The inclusion of future RCTs assessing all relevant mandatory outcome domains with low risk of bias and high quality in reporting that could yield optimal information size may help to potentially lead to more definitive conclusions.

## Conclusion

RATKA probably yields more precise prosthesis alignment compared with COTKA based on MA outliers. However, there is probably no clinically important difference in MA deviation, PROMs, and ROM, and there was insufficient evidence on implant survivorship and complication rate. Future well-designed RCTs focusing on longer follow-up periods that evaluate all important outcome domains are required to improve the quality and extent of the current evidence. Health technology assessments and cost-analysis studies and reviews should also be performed to guide decision-making in implementing this novel intervention.

## Supplementary data

Tables 4 and 5, Search strategies, and PRISMA checklist are available as Supplementary material on the article homepage, doi: 10.2340/17453674.2023.9411

PR and OR provided research questions, conducted data collection, and wrote and finalized the manuscript. CP provided methodological guidance to the review and graded the quality and certainty of the evidence. RN monitored the data-collection process. RN and KC, as experienced certified

1. **Cui A, Li H, Wang D, Zhong J, Chen Y, Lu H.** Global, regional prevalence, incidence and risk factors of knee osteoarthritis in population-based studies. EClinicalMedicine 2020; 29-30:100587. doi: 10.1016/j.eclinm.2020.100587

2. **Johnston H, Abdelgaied A, Pandit H, Fisher J, Jennings L M.** The effect of surgical alignment and soft tissue conditions on the kinematics and wear of a fixed bearing total knee replacement. J Mech Behav Biomed Mater 2019; 100: 103386. doi: 10.1016/j.jmbbm.2019.103386.

3. **Choong P F, Dowsey M M, Stoney J D. Does accurate anatomical alignment result in better function and quality of life?** Comparing conventional and computer-assisted total knee arthroplasty. J Arthroplasty 2009; 24(4): 560-9. doi: 10.1016/j.arth.2008.02.018.

4. **Khlopas A, Sodhi N, Sultan A A, Chughtai M, Molloy R M, Mont M A.** Robotic arm-assisted total knee arthroplasty. J Arthroplasty 2018; 33(7): 2002-6. doi: 10.1016/j.arth.2018.01.060.

5. **Agarwal N, To K, McDonnell S, Khan W.** Clinical and radiological outcomes in robotic-assisted total knee arthroplasty: a systematic review and meta-analysis. J Arthroplasty 2020; 35(11): 3393-409.e2. doi: 10.1016/j.arth.2020.03.005.

6. **Onggo J R, Onggo J D, De Steiger R, Hau R.** Robotic-assisted total knee arthroplasty is comparable to conventional total knee arthroplasty: a meta-analysis and systematic review. Arch Orthop Trauma Surg 2020; 140(10): 1533-49. doi: 10.1007/s00402-020-03512-5.

7. **Zhang J, Ndou W S, Ng N, Gaston P, Simpson P M, Macpherson G J, et al.** Robotic-arm assisted total knee arthroplasty is associated with improved accuracy and patient reported outcomes: a systematic review and meta-analysis. Knee Surg Sports Traumatol Arthrosc 2022; 30(8): 2677-95. doi: 10.1007/s00167-021-06464-4.

8. **Mullaji A B, Khalifa A A.** Is it prime time for robotic-assisted TKAs? A systematic review of current studies. J Orthop 2022; 34: 31-9. doi: 10.1016/j.jor.2022.07.016.

9. **Page M J, McKenzie J E, Bossuyt P M, Boutron I, Hoffmann T C, Mulrow C D, et al.** The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. BMJ 2021; 372: n71. doi: https://doi.org/10.1136/bmj.n71

10. **Higgins J P T, Sterne J A C, Savović J, Page M J, Hróbjartsson A, Boutron I, et al.** A revised tool for assessing risk of bias in randomized trials In: Chandler J, McKenzie J, Boutron I, Welch V, editors. Cochrane Methods. Cochrane Database of Systematic Reviews 2016, Issue 10 (Suppl. 1). dx.doi.org/10.1002/14651858.CD201601.

11. **Bellamy N, Buchanan W W, Goldsmith C H, Campbell J, Stitt L W.** Validation study of WOMAC: a health status instrument for measuring clinically important patient relevant outcomes to antirheumatic drug therapy in patients with osteoarthritis of the hip or knee. J Rheumatol 1988; 15(12): 1833-40.

12. **Scuderi G R, Bourne R B, Noble P C, Benjamin J B, Lonner J H, Scott W N.** The new Knee Society Knee Scoring System. Clin Orthop Relat Res 2012; 470(1): 3-19. doi: 10.1007/s11999-011-2135-0.

13. **Insall J N, Ranawat C S, Aglietti P, Shine J.** A comparison of four models of total knee-replacement prostheses. J Bone Joint Surg Am 1976; 58(6): 754-65.

14. **Sappey-Marinier E, Pauvert A, Batailler C, Swan J, Cheze L, Servien E, et al.** Kinematic versus mechanical alignment for primary total knee arthroplasty with minimum 2 years follow-up: a systematic review. SICOT J 2020; 6: 18. doi: 10.1051/sicotj/2020014.

15. **Rivière C, Iranpour F, Auvinet E, Howell S, Vendittoli P A, Cobb J, et al.** Alignment options for total knee arthroplasty: a systematic review. Orthop Traumatol Surg Res 2017; 103(7): 1047-56. doi: 10.1016/j.otsr.2017.07.010.

16. **Higgins J P T, Altman D G, Gøtzsche PC, Jüni P, Moher D, Oxman A D, et al.** The Cochrane Collaboration's tool for assessing risk of bias in randomised trials. BMJ 2011; 343: d5928. doi: https://doi.org/10.1136/bmj.d5928.

17. **Lee W C, Kwan Y H, Chong H C, Yeo S J.** The minimal clinically important difference for Knee Society Clinical Rating System after total knee arthroplasty for primary osteoarthritis. Knee Surg Sports Traumatol Arthrosc 2017; 25(11): 3354-9. doi: 10.1007/s00167-016-4208-9.

18. **Escobar A, Riddle D L.** Concordance between important change and acceptable symptom state following knee arthroplasty: the role of baseline scores. Osteoarthritis Cartilage 2014; 22(8): 1107-10. doi: 10.1016/j.joca.2014.06.006.

19. **Long W J, Bryce C D, Hollenbeak C S, Benner R W, Scott W N.** Total knee replacement in young, active patients: long-term follow-up and functional outcome: a concise follow-up of a previous report. J Bone Joint Surg Am 2014 ; 96(18): e159. doi: 10.2106/JBJS.M.01259.

20. **Escobar A, García Pérez L, Herrera-Espiñeira C, Aizpuru F, Sarasqueta C, Gonzalez Sáenz de Tejada M, et al.** Total knee replacement: minimal clinically important differences and responders. Osteoarthritis Cartilage 2013; 21(12): 2006-12. doi: 10.1016/j.joca.2013.09.009.

21. **Lee B S, Cho H I, Bin S I, Kim J M, Jo B K.** Femoral component varus malposition is associated with tibial aseptic loosening after TKA. Clin Orthop Relat Res 2018; 476(2): 400-7. doi: 10.1007/s11999.0000000000000012.

22. **Park S E, Lee C T.** Comparison of robotic-assisted and conventional manual implantation of a primary total knee arthroplasty. J Arthroplasty 2007; 22(7): 1054-9. doi: 10.1016/j.arth.2007.05.036.

23. **Song E K, Seon J K, Park S J, Jung W B, Park H W, Lee G W.** Simultaneous bilateral total knee arthroplasty with robotic and conventional techniques: a prospective, randomized study. Knee Surg Sports Traumatol Arthrosc 2011; 19(7): 1069-76. doi: 10.1007/s00167-011-1400-9.

24. **Song E K, Seon J K, Yim J H, Netravali N A, Bargar W L.** Robotic-assisted TKA reduces postoperative alignment outliers and improves gap balance compared to conventional TKA. Clin Orthop Relat Res 2013; 471(1): 118-26. doi: 10.1007/s11999-012-2407-3.

25. **Liow M H L, Xia Z, Wong M K, Tay K J, Yeo S J, Chin P L.** Robot-assisted total knee arthroplasty accurately restores the joint line and mechanical axis: a prospective randomised study. J Arthroplasty 2014; 29(12): 2373-7. doi: 10.1016/j.arth.2013.12.010.

26. **Liow M H L, Goh G S H, Wong M K, Chin P L, Tay D K J, Yeo S J.** Robotic-assisted total knee arthroplasty may lead to improvement in quality-of-life measures: a 2-year follow-up of a prospective randomized trial. Knee Surg Sports Traumatol Arthrosc 2017; 25(9): 2942-51. doi: 10.1007/s00167-016-4076-3.

27. **Kim Y H, Yoon S H, Park J W.** Does robotic-assisted TKA result in better outcome scores or long-term survivorship than conventional TKA? A randomized, controlled trial. Clin Orthop Relat Res 2020; 478(2): 266-75. doi: 10.1097/CORR.0000000000000916.

28. **Kayani B, Tahmassebi J, Ayuob A, Konan S, Oussedik S, Haddad F S.** A prospective randomized controlled trial comparing the systemic inflammatory response in conventional jig-based total knee arthroplasty versus robotic-arm assisted total knee arthroplasty. Bone Joint J 2021; 103-B(1): 113-22. doi: 10.1302/0301-620X.103B1.BJJ-2020-0602.R2.

29. **Thiengwittayaporn S, Uthaitas P, Senwiruch C, Hongku N, Tunyasuwanakul R.** Imageless robotic-assisted total knee arthroplasty accurately restores the radiological alignment with a short learning curve: a randomized controlled trial. Int Orthop 2021; 45(11): 2851-8. doi: 10.1007/s00264-021-05179-y.

30. **Vaidya N V, Deshpande A N, Panjwani T, Patil R, Jaysingani T, Patil P.** Robotic-assisted TKA leads to a better prosthesis alignment and a better joint line restoration as compared to conventional TKA: a prospective randomized controlled trial. Knee Surg Sports Traumatol Arthrosc 2022; 30(2): 621-6. doi: 10.1007/s00167-020-06353-2.

31. **Xu J, Li L, Fu J, Xu C, Ni M, Chai W, et al.** Early clinical and radiographic outcomes of robot-assisted versus conventional manual total knee arthroplasty: a randomized controlled study. Orthop Surg 2022; 14(9): 1972-80. doi: 10.1111/os.13323.

32. **Li Z, Chen X, Wang X, Zhang B, Wang W, Fan Y, et al.** HURWA robotic-assisted total knee arthroplasty improves component positioning and alignment: a prospective randomized and multicenter study. J Orthop Translat 2022; 33: 31-40. doi: 10.1016/j.jot.2021.12.004.

33. **Lychagin A, Elizarov M, Gritsyuk A, Rukin Y, Elizarov P, Rokity-anskaya A, et al.** Robot-assisted knee arthroplasty: randomized clinical trial. Open Access Maced J Med Sci 2022; 10(B): 559-64. https://doi.org/10.3889/oamjms.2022.8685.

34. **Eckhard L, Munir S, Wood D, Talbot S, Brighton R, Walter B, et al.** The ceiling effects of patient reported outcome measures for total knee arthroplasty. Orthop Traumatol Surg Res 2021; 107(3): 102758. doi: 10.1016/j.otsr.2020.102758.

35. **Na S E, Ha C W, Lee C H.** A new high-flexion knee scoring system to eliminate the ceiling effect. Clin Orthop Relat Res 2012; 470(2): 584-93. doi: 10.1007/s11999-011-2203-5.

36. **Hadi M, Barlow T, Ahmed I, Dunbar M, McCulloch P, Griffin D.** Does malalignment affect revision rate in total knee replacements: a systematic review of the literature. Springerplus 2015; 4: 835. doi: 10.1186/s40064-015-1604-4.

37. **Abdel M P, Ollivier M, Parratte S, Trousdale R T, Berry D J, Pagnano M W.** Effect of postoperative mechanical axis alignment on survival and functional outcomes of modern total knee arthroplasties with cement: a concise follow-up at 20 years. J Bone Joint Surg Am 2018; 100(6): 472-8. doi: 10.2106/JBJS.16.01587.

38. **Smith T O, Hawker G A, Hunter D J, March L M, Boers M, Shea B J, et al.** The OMERACT-OARSI core domain set for measurement in clinical trials of hip and/or knee osteoarthritis. J Rheumatol 2019; 46(8): 981-9. doi: 10.3899/jrheum.181194.

39. **Vermue H, Tack P, Gryson T, Victor J.** Can robot-assisted total knee arthroplasty be a cost-effective procedure? A Markov decision analysis. Knee 2021; 29: 345-52. doi: 10.1016/j.knee.2021.02.004.

40. **Rajan P V, Khlopas A, Klika A, Molloy R, Krebs V, Piuzzi N S.** The cost-effectiveness of robotic-assisted versus manual total knee arthroplasty: a Markov model-based evaluation. J Am Acad Orthop Surg 2022; 30(4): 168-76. doi: 10.5435/JAAOS-D-21-00309.

# Appendix

**Table 3. Complications reported in included studies TO APPENDIX**

| Study | Robotic-assisted total knee arthroplasty | Conventional total knee arthroplasty |
| --- | --- | --- |
| Park, 2007 | n = 6 (group sample = 32)<br>1 superficial infection, 1 patellar tendon rupture, 1 patella dislocation, 1 postoperative supracondylar fracture, 1 patellar fracture, and 1 peroneal injury | Not reported |
| Song, 2011 | n = 0 (group sample = 30)<br>"no major adverse events, such as deep infection or loosening requiring revision" | Not reported |
| Song, 2013 | n = 11 (group sample = 50)<br>*Local complication (n = 6):* 2 superficial infection, 1 seroma at pin site, 2 patellar tendon abrasion, 1 skin rash.<br>*Systemic complications (n = 5):* 1 lymphatic edema, 1 crepitus, 2 arrhythmia, and 1 pleural effusion. | n = 11 (group sample = 50)<br>*Local complication (n = 3):* 2 superficial infection, 1 incisional skin sloughing<br>*Systemic complications (n = 8):* 2 gastrointestinal issue,1 cerebrovascular accident, 3 pleural effusion, and 2 foot numbness |
| Liow, 2014 [a] | n = 2 (group sample = 31) likely duplicates of Liow, 2017<br>1 deep vein thrombosis, 1 superficial wound infection | n = 2 (group sample = 29) likely duplicates of Liow, 2017<br>1 deep vein thrombosis, 1 postoperative delirium |
| Liow, 2017 | n = 4 (group sample = 31)<br>1 deep vein thrombosis, 1 superficial wound infection, 1 deep vein thrombosis and septic arthritis, and 1 persistent lateral knee pain | n = 2 (group sample = 29)<br>1 deep vein thrombosis, 1 postoperative delirium |
| Kim, 2020 | n = 4 (group sample = 724)<br>4 superficial infection | n = 8 (group sample = 724)<br>4 superficial infection, 4 motion limitation |
| Kayani, 2021 | n = 0 (group sample = 15)<br>"There was no postoperative complication" | n = 0 (group sample = 15)<br>"There was no postoperative complication" |
| Xu, 2022 | n = 10 (group sample = 37)<br>only deep vein thrombosis reported | n = 11 (group sample = 35)<br>only deep vein thrombosis reported) |
| Total | 35 per 919 | 32 per 853 |

[a] Sample not included in the total row as they were duplicates of Liow, 2017

**SUMMARY OF FINDINGS: Robotic-assisted total knee arthroplasty compared with conventional total knee arthroplasty for adult patients with primary osteoarthritis of the knee**

| Outcomes | Risk with RATKA | Risk with Conventional TKA | Relative effect (95% CI) | Participants (studies) | Certainty of the evidence (GRADE) | Comments |
|---|---|---|---|---|---|---|
| WOMAC score (lower = better) time: 3 months to 13 years | Mean score range 1.3–81 points | 1.6–72 points | MD –0.35 (–0.78 to 0.07) | 1,744 (6 RCTs) | ⊕⊕⊕◯ MODERATE [a] | MID = 15 |
| KSS score (higher = better) time: 3 months to 13 years | Mean score range 67–93 points | 66–92 points | MD 0.89 (–0.18 to 2.0) | 1,685 (5 RCTs) | ⊕⊕⊕◯ MODERATE [b,c] | MID = 6 |
| HSS score (higher = better) time: 3 to 65 months | Mean score range 80–96 points | 80–95 points | MD 0.67 (–0.99 to 2.3) | 256 (3 RCTs) | ⊕⊕⊕◯ MODERATE [d,e] | MID= 13 |
| Range of motion (higher = better) time: 3 months to 13 years | Mean range 115°–129° | 112°–129° | MD –0.73° (–7.5° to 6.0°) | 2,013 (8 RCTs) | ⊕⊕◯◯ LOW [f,g] | MID = 30 |
| Deviation from neutral mechanical axis time: postoperative to 13 years | Mean deviation range 0.2°–2° | 1.2°–3° | MD –0.94° (–1.1° to –0.73°) | 1,871 (6 RCTs) | ⊕⊕⊕◯ MODERATE [h,i] | MID = 3 |
| Mechanical axis outliers time: postoperative to 13 years | 125 per 1,048 | 53 per 1,044 | RR 0.43 (0.27 to 0.67) | 2,092 (8 RCTs) | ⊕⊕⊕◯ MODERATE [j,k,l] | Risk difference 22 (13–28) fewer/1,000 Number needed to treat = 45 (36–77) |
| Adverse effects | 35 per 919 | 32 per 853 | n/a | 1,772 (7 RCTs) | n/a | 2 studies only reported complications in the RATKA arm. No significant morbidity or mortality reported |
| Implant survivorship | 15 per 750 | 15 per 766 | RR 1.02 (0.50 to 2.1) | 1,516 (1 RCT) | ⊕⊕◯◯ LOW [m] | Risk difference 0 (–10 to 21)/1,000 |

Abbreviations: TKA = total knee arthroplasty, RATKA = robotic-assisted total knee arthroplasty, CI = Confidence interval, WOMAC = Western Ontario and McMaster University Osteoarthritis Index, KSS = Knee Society Score, HSS = Hospital for Special Surgery, MD = Mean difference, RR = Risk ratio, MID = Minimal important difference.

**GRADE Working Group grades of evidence**

*High certainty:* We are very confident that the true effect lies close to that of the estimate of the effect.

*Moderate certainty:* We are moderately confident in the effect estimate: The true effect is likely to be close to the estimate of the effect, but there is a possibility that it is substantially different.

*Low certainty:* Our confidence in the effect estimate is limited: The true effect may be substantially different from the estimate of the effect.

*Very low certainty:* We have very little confidence in the effect estimate: The true effect is likely to be substantially different from the estimate of effect.

[a] 2 studies had high risk of bias, 1 of which had the highest contribution to the effect estimate.

[b] 1 study had high risk of bias (allocation concealment) that was important, but it contributed to only 0.7% weight of the effect estimate.

[c] most point estimates in the same direction, CIs overlap, $I^2$ = 25% with insignificant Chi-squared test but it was quite large due to the study with high risk of bias.

[d] 1 study with high contribution (29.7% weight) had high risk of bias (large missing outcome proportion) but its result was similar to the other studies.

[e] optimal information size using the rule of thumb ( n= 400) was not met but CI covers no effect and excludes important clinical benefit.

[f] 2 studies had high risk of bias, 1 of which (Lychagin) had a relatively high contribution (37.3%) to the effect estimate and changed the overall direction of effect.

[g] most point estimates in a different direction with non-overlapping confidence intervals, large $I^2$ of 99%. The heterogeneity could be explained by Lychagin's study (risk of bias and type of robot used).

[h] 1 study with high risk (allocation concealment) which was important and contributed to 14.8 % weight of the result, but excluding it did not change the pooled result.

[i] CI excludes no effect but did not include important difference.

[j] 1 study had high risk of bias (allocation concealment) that was important, but it contributed to only 0.09% weight of the effect estimate.

[k] point estimates in the same direction with mostly overlapping CI, although $I^2$ is relatively large (31%), which was due to studies with scarce events.

[l] optimal information size using the rule of thumb (events = 300) was not met but CI excludes no effect and includes a large benefit.

[m] optimal information size using the rule of thumb (events = 300) was not met and CI includes no effect and also includes a large benefit and harm.

**A**

| WOMAC Study | RATKA mean (SD) | n | Conventional TKA mean (SD) | n | Weight % | Difference mean (95%CI) | Difference mean (95%CI) | Risk of bias A B C D E |
|---|---|---|---|---|---|---|---|---|
| Song 2011 (23) | 11 (4.5) | 30 | 13 (6.6) | 30 | 22.2 | –2.0 (–4.9 to 0.86) | | |
| Kim 2020 (27) | 18 (14) | 674 | 19 (15) | 674 | 75.6 | –1.0 (–2.6 to 0.55) | | |
| Li 2022 (32) | 81 (46) | 69 | 72 (44) | 74 | 0.8 | 8.5 (–6.4 to 23) | | |
| Xu 2022 (31) | 13 (13) | 37 | 19 (31) | 35 | 1.4 | –6.9 (–18 to 4.3) | | |
| Total | | 810 | | 813 | | –1.2 (–2.6 to 0.12) | | |

Heterogeneity: Tau = 0.00; Chi² = 3.0, df = 3 (p = 0.4); I² = 0%
Test for overall effect Z = 1.8 (p = 0.07)

**B**

| KSS Study | Conventional TKA mean (SD) | n | RATKA mean (SD) | n | Weight % | Difference mean (95%CI) | Difference mean (95%CI) | Risk of bias A B C D E |
|---|---|---|---|---|---|---|---|---|
| Park 2007 (22) | 91 (4.9) | 30 | 92 (2.9) | 32 | 7.4 | –0.70 (–2.7 to 1.3) | | |
| Kim 2020 (27) | 92 (6.0) | 674 | 93 (5.0) | 674 | 87.1 | –1.0 (–1.6 to –0.41) | | |
| Li 2022 (32) | 68 (10) | 74 | 70 (7.3) | 69 | 3.7 | –2.2 (–5.1 to 0.67) | | |
| Xu 2022 (31) | 66 (7.4) | 35 | 67 (10) | 37 | 1.8 | –1.0 (–5.1 to 3.1) | | |
| Total | | 813 | | 812 | | –1.0 (–1.6 to –0.47) | | |

Heterogeneity: Tau = 0.00; Chi² = 0.75, df = 3 (p = 0.9); I² = 0%
Test for overall effect Z = 3.6 (p < 0.001)

**C**

| HSS score Study | Conventional TKA mean (SD) | n | RATKA mean (SD) | n | Weight % | Difference mean (95%CI) | Difference mean (95%CI) | Risk of bias A B C D E |
|---|---|---|---|---|---|---|---|---|
| Song 2011 (23) | 95 (5.0) | 30 | 95 (4.0) | 30 | 74.7 | –0.50 (–2.8 to 1.8) | | |
| Li 2022 (32) | 79 (12) | 74 | 80 (12) | 69 | 25.3 | –0.65 (–4.6 to 3.3) | | |
| Total | | 104 | | 99 | | –0.54 (–2.5 to 1.4) | | |

Heterogeneity: Tau = 0.00; Chi² = 0.00, df = 1 (p = 0.95); I² = 0%
Test for overall effect Z = 0.5 (p = 0.6)

**D**

| ROM Study | Conventional TKA mean (SD) | n | RATKA mean (SD) | n | Weight % | Difference mean (95%CI) | Difference mean (95%CI) | Risk of bias A B C D E |
|---|---|---|---|---|---|---|---|---|
| Park 2007 (22) | 122 (17) | 30 | 118 (9.0) | 32 | 4.0 | 4.0 (–2.8 to 11) | | |
| Song 2011 (23) | 129 (13) | 30 | 129 (14) | 30 | 4.1 | 0.0 (–6.7 to 6.7) | | |
| Song 2013 (24) | 129 (12) | 50 | 128 (5.1) | 50 | 12.1 | 1.0 (–2.7 to 4.7) | | |
| Kim 2020 (27) | 128 (7.0) | 724 | 125 (6.0) | 724 | 71.2 | 3.0 (2.3 to 3.7) | | |
| Xu 2022 (31) | 120 (19) | 35 | 115 (11) | 37 | 3.7 | 5.0 (–2.1 to 12) | | |
| Li 2022 (32) | 112 (19) | 74 | 115 (18) | 69 | 4.9 | –2.9 (–9.0 to 3.2) | | |
| Total | | 943 | | 942 | | 2.5 (1.1 to 3.9) | | |

Heterogeneity: Tau = 0.59; Chi² = 5.7, df = 5 (p = 0.3); I² = 12%
Test for overall effect Z = 3.5 (p < 0.001)

**E**

| MA deviation Study | RATKA mean (SD) | n | Conventional TKA mean (SD) | n | Weight % | Difference mean (95%CI) | Difference mean (95%CI) | Risk of bias A B C D E |
|---|---|---|---|---|---|---|---|---|
| Song 2011 (23) | 0.2 (1.6) | 30 | 1.2 (2.1) | 30 | 5.7 | –1.0 (–1.9 to –0.06) | | |
| Song 2013 (24) | 0.5 (1.4) | 50 | 1.2 (2.9) | 50 | 6.3 | –0.70 (–1.6 to 0.19) | | |
| Kim 2020 (27) | 2.0 (2.0) | 724 | 3.0 (3.0) | 724 | 73.1 | –1.0 (–1.3 to –0.74) | | |
| Li 2022 (32) | 1.8 (1.6) | 69 | 3.0 (2.7) | 74 | 9.4 | –1.2 (–2.0 to –0.49) | | |
| Vaidya 2022 (30) | 1.8 (1.2) | 30 | 3.0 (2.4) | 30 | 5.5 | –1.2 (–2.2 to –0.24) | | |
| Total | | 903 | | 908 | | –1.0 (–1.2 to –0.79) | | |

Heterogeneity: Tau = 0.00; Chi² = 0.94, df = 4 (p = 0.9); I² = 0%
Test for overall effect Z = 8.8 (p < 0.001)

**F**

| MA outlier (> 3°) Study | RATKA events | n | Conventional TKA events | n | Weight % | Risk ratio (95%CI) | Risk ratio (95%CI) | Risk of bias A B C D E |
|---|---|---|---|---|---|---|---|---|
| Song 2011 (23) | 0 | 30 | 7 | 30 | 2.4 | 0.07 (0.00 to 1.1) | | |
| Song 2013 (24) | 0 | 50 | 12 | 50 | 2.5 | 0.04 (0.00 to 0.66) | | |
| Kim 2020 (27) | 29 | 724 | 43 | 724 | 32.6 | 0.67 (0.43 to 1.1) | | |
| Thiengwittayaporn 2021 (29) | 4 | 75 | 12 | 77 | 12.8 | 0.34 (0.12 to 1.0) | | |
| Li 2022 (32) | 13 | 69 | 27 | 74 | 27.4 | 0.52 (0.29 to 0.92) | | |
| Vaidya 2022 (30) | 1 | 30 | 8 | 30 | 4.6 | 0.13 (0.02 to 0.94) | | |
| Xu 2022 (31) | 6 | 35 | 12 | 34 | 17.7 | 0.49 (0.21 to 1.2) | | |
| Total | | 1,013 | | 1,019 | | 0.44 (0.28 to 0.70) | | |

Heterogeneity: Tau = 0.11; Chi² = 9, df = 6 (p = 0.2); I² = 33%
Test for overall effect Z = 3.5 (p < 0.001)

**G**

| Blood loss Study | RATKA mean (SD) | n | Conventional TKA mean (SD) | n | Weight % | Difference mean (95%CI) | | Risk of bias A B C D E |
|---|---|---|---|---|---|---|---|---|
| Song 2011 (23) | 587 (385) | 30 | 816 (425) | 30 | 24.3 | −229 (−435 to −24) | | +++?? |
| Kim 2020 (27) | 261 (107) | 700 | 255 (110) | 706 | 52.0 | 6 (−5.4 to 17) | | +++?+ |
| Xu 2022 (31) | 933 (454) | 37 | 863 (458) | 35 | 23.7 | 70 (−141 to 280) | | ?+?+?+ |
| Total | | 767 | | 771 | | −36 (−175 to 102) | | |

Heterogeneity: Tau = 9,592; Chi² = 5.4, df = 2 (p = 0.07); I² = 63%
Test for overall effect Z = 0.51 (p = 0.6)

−500  −250  0  250  500
Favors RATKA    Favors conventional TKA

**H**

| Operative time Study | RATKA mean (SD) | n | Conventional TKA mean (SD) | n | Weight % | Difference mean (95%CI) | | Risk of bias A B C D E |
|---|---|---|---|---|---|---|---|---|
| Kim 2020 (27) | 97 (11) | 700 | 69 (6.3) | 706 | 27.1 | 28 (27 to 29) | | +++?+ |
| Song 2011 (23) | 95 (18) | 30 | 70 (15) | 30 | 24.3 | 25 (17 to 33) | | ?+?+?? |
| Thiengwittayaporn 2021 (29) | 70 (12) | 75 | 62 (10) | 77 | 26.6 | 8.2 (4.7 to 12) | | ?+++?+ |
| Xu 2022 (31) | 154 (21) | 37 | 115 (30) | 35 | 21.9 | 39 (27 to 51) | | ?+?+?+ |
| Total | | 842 | | 848 | | 24 (12 to 37) | | |

Heterogeneity: Tau = 157; Chi² = 117, df = 3 (p < 0.001); I² = 97%
Test for overall effect Z = 3.7 (p < 0.001)

−50  −25  0  25  50
Favors RATKA    Favors conventional TKA

Figure 6. Forest plots showing the Western Ontario and McMaster University Osteoarthritis Index scores (A), Knee Society Scores (B), Hospital for Special Surgery scores (C), range of motion (D), deviation from neutral mechanical axis (E), mechanical axis alignment outliers (F), intraoperative blood loss (G) and operative time (H) between robotic-assisted total knee arthroplasty (RATKA) and conventional total knee arthroplasty (conventional TKA) after excluding studies with high risk of bias using random-effects model and inverse variance method. For abbreviations and Risk of Bias A–E, see Figure 3.

**A**

| WOMAC Study | RATKA mean (SD) | n | Conventional TKA mean (SD) | n | Weight % | Difference mean (95%CI) | | Risk of bias A B C D E |
|---|---|---|---|---|---|---|---|---|
| Song 2011 (23) | 11 (4.5) | 30 | 13 (6.6) | 30 | 19.6 | −2.0 (−4.9 to 0.86) | | +++?? |
| Song 2013 (24) | 29 (4.4) | 29 | 30 (7.5) | 24 | 13.8 | −1.1 (−4.5 to 2.3) | | ++−?? |
| Kim 2020 (27) | 18 (14) | 674 | 19 (15) | 674 | 66.6 | −1.0 (−2.6 to 0.55) | | +++?+ |
| Total | | 733 | | 728 | | −1.2 (−2.5 to 0.05) | | |

Heterogeneity: Tau = 0.00; Chi² = 0.37, df = 2 (p = 0.8); I² = 0%
Test for overall effect Z = 1.9 (p = 0.06)

−20  −10  0  10  20
Favors RATKA    Favors conventional TKA

**B**

| KSS Study | Conventional TKA mean (SD) | n | RATKA mean (SD) | n | Weight % | Difference mean (95%CI) | | Risk of bias A B C D E |
|---|---|---|---|---|---|---|---|---|
| Park 2007 (22) | 91 (4.9) | 30 | 92 (2.9) | 32 | 34.9 | −0.70 (−2.7 to 1.3) | | ?+++?? |
| Liow 2017 (25) | 88 (11) | 29 | 82 (15) | 31 | 6.8 | 6.1 (−0.41 to 13) | | −+++?? |
| Kim 2020 (27) | 92 (6.0) | 674 | 93 (5.0) | 674 | 58.3 | −1.0 (−1.6 to −0.41) | | +++?+ |
| Total | | 733 | | 737 | | −0.41 (−2.2 to 1.4) | | |

Heterogeneity: Tau = 1.4; Chi² = 4.6, df = 2 (p = 0.1); I² = 56%
Test for overall effect Z = 0.45 (p = 0.7)

−20  −10  0  10  20
Favors RATKA    Favors conventional TKA

**C**

| HSS score Study | Conventional TKA mean (SD) | n | RATKA mean (SD) | n | Weight % | Difference mean (95%CI) | | Risk of bias A B C D E |
|---|---|---|---|---|---|---|---|---|
| Song 2011 (23) | 95 (5.0) | 30 | 95 (4.0) | 30 | 63.9 | −0.50 (−2.8 to 1.8) | | +++?? |
| Song 2013 (24) | 95 (6.7) | 24 | 96 (4.0) | 29 | 36.1 | −1.0 (−4.1 to 2.1) | | ++−?? |
| Total | | 54 | | 59 | | −0.68 (−2.5 to 1.2) | | |

Heterogeneity: Tau = 0.00; Chi² = 0.07, df = 1 (p = 0.8); I² = 0%
Test for overall effect Z = 0.73 (p = 0.5)

−20  −10  0  10  20
Favors RATKA    Favors conventional TKA

**D**

| ROM Study | Conventional TKA mean (SD) | n | RATKA mean (SD) | n | Weight % | Difference mean (95%CI) | | Risk of bias A B C D E |
|---|---|---|---|---|---|---|---|---|
| Park 2007 (22) | 122 (17) | 30 | 118 (9.0) | 32 | 2.7 | 4.0 (−2.8 to 11) | | ?+++?? |
| Song 2011 (23) | 129 (13) | 30 | 129 (14) | 30 | 2.8 | 0.0 (−6.7 to 6.7) | | ++++? |
| Song 2013 (24) | 129 (12) | 50 | 128 (5.1) | 50 | 8.7 | 1.0 (−2.7 to 4.7) | | ++++? |
| Liow 2017 (26) | 126 (10) | 29 | 117 (16) | 31 | 2.9 | 8.2 (1.6 to 15) | | −+++?? |
| Kim 2020 (27) | 128 (7.0) | 724 | 125 (6.0) | 724 | 82.9 | 3.0 (2.3 to 3.7) | | +++++ |
| Total | | 863 | | 867 | | 2.9 (1.8 to 4.1) | | |

Heterogeneity: Tau = 0.29; Chi² = 4.3, df = 4 (p = 0.4); I² = 7%
Test for overall effect Z = 5-01 (p < 0.001)

−20  −10  0  10  20
Favors RATKA    Favors conventional TKA

**E**

| MA deviation Study | RATKA mean (SD) | n | Conventional TKA mean (SD) | n | Weight % | Difference mean (95%CI) | Risk of bias A B C D E |
|---|---|---|---|---|---|---|---|
| Song 2011 (23) | 0.2 (1.6) | 30 | 1.2 (2.1) | 30 | 5.2 | −1.0 (−1.9 to −0.06) | + + + + ? |
| Song 2013 (24) | 0.5 (1.4) | 50 | 1.2 (2.9) | 50 | 5.9 | −0.70 (−1.6 to 0.19) | + + + + ? |
| Liow 2014 (25) | 1.3 (0.9) | 31 | 1.8 (1.2) | 29 | 16.1 | −0.50 (−1.0 to 0.04) | − + + + ? |
| Kim 2020 (27) | 2.0 (2.0) | 724 | 3.0 (3.0) | 724 | 67.8 | −1.0 (−1.3 to −0.74) | + + + + + |
| Vaidya 2022 (30) | 1.8 (1.2) | 30 | 3.0 (2.4) | 30 | 5.1 | −1.2 (−2.2 to −0.24) | ? + + ? ? |
| Total | | 865 | | 863 | | −0.91 (−1.1 to −0.70) | |

Heterogeneity: Tau = 0.00; Chi² = 3.3, df = 4 (p = 0.5); I² = 0%
Test for overall effect Z = 8.3 (p < 0.001)

Favors RATKA — Favors conventional TKA

**F**

| MA outlier (> 3°) Study | RATKA events | n | Conventional TKA events | n | Weight % | Risk ratio (95%CI) | Risk of bias A B C D E |
|---|---|---|---|---|---|---|---|
| Song 2011 (23) | 0 | 30 | 7 | 30 | 8.3 | 0.07 (0.00 to 1.1) | + + + + ? |
| Song 2013 (24) | 0 | 50 | 12 | 50 | 8.4 | 0.04 (0.00 to 0.66) | + + + + ? |
| Liow 2014 (25) | 0 | 31 | 4 | 29 | 8.0 | 0.10 (0.01 to 1.9) | − + + + ? |
| Kim 2020 (27) | 29 | 724 | 43 | 724 | 36.2 | 0.67 (0.43 to 1.1) | + + + + + |
| Thiengwittayaporn 2021 (29) | 4 | 75 | 12 | 77 | 25.5 | 0.34 (0.12 to 1.0) | ? + + + + |
| Vaidya 2022 (30) | 1 | 30 | 8 | 30 | 13.5 | 0.13 (0.02 to 0.94) | ? + + ? ? |
| Total | | 940 | | 940 | | 0.25 (0.10 to 0.63) | |

Heterogeneity: Tau = 0.55; Chi² = 10, df = 5 (p = 0.07); I² = 51%
Test for overall effect Z = 3.0 (p = 0.003)

Favors RATKA — Favors conventional TKA

**G**

| Blood loss Study | RATKA mean (SD) | n | Conventional TKA mean (SD) | n | Weight % | Difference mean (95%CI) | Risk of bias A B C D E |
|---|---|---|---|---|---|---|---|
| Song 2011 (23) | 587 (385) | 30 | 816 (425) | 30 | 29.6 | −229 (−435 to −24) | + + + ? ? |
| Song 2013 (24) | 613 (318) | 50 | 933 (467) | 50 | 32.6 | −320 (−477 to −163) | + + − ? ? |
| Kim 2020 (27) | 261 (107) | 700 | 255 (110) | 706 | 37.8 | 6.0 (−5.4 to 17) | + + + ? + |
| Total | | 780 | | 786 | | −170 (−410 to 70) | |

Heterogeneity: Tau = 39,641; Chi² = 22, df = 2 (p < 0.001); I² = 91%
Test for overall effect Z = 1.4 (p = 0.2)

Favors RATKA — Favors conventional TKA

**H**

| Operative time Study | RATKA mean (SD) | n | Conventional TKA mean (SD) | n | Weight % | Difference mean (95%CI) | Risk of bias A B C D E |
|---|---|---|---|---|---|---|---|
| Thiengwittayaporn 2021 (29) | 70 (12) | 75 | 62 (10) | 77 | 16.8 | 8.2 (4.7 to 12) | ? + + + + |
| Song 2013 (24) | 99 (11) | 50 | 74 (10) | 50 | 16.8 | 25 (21 to 29) | + + − ? ? |
| Song 2011 (23) | 95 (18) | 30 | 70 (15) | 30 | 16.0 | 25 (17 to 33) | + + + ? ? |
| Liow 2017 (36) | 91 (10) | 31 | 93 (14) | 29 | 16.4 | −2.0 (−8.2 to 4.2) | − + + ? ? |
| Kim 2020 (27) | 97 (11) | 700 | 69 (6.3) | 706 | 17.0 | 28 (27 to 29) | + + + ? + |
| Kayani 2021 (28) | 61 (3.1) | 15 | 62 (3.4) | 15 | 17.0 | −1.0 (−3.3 to 1.3) | ? + + + − |
| Total | | 901 | | 907 | | 14 (0.41 to 27) | |

Heterogeneity: Tau = 275; Chi² = 657, df = 5 (p < 0.001); I² = 99%
Test for overall effect Z = 2.0 (p = 0.04)

Favors RATKA — Favors conventional TKA

Figure 7. Forest plots showing the Western Ontario and McMaster University scores Osteoarthritis Index scores (A), Knee Society Scores (B), Hospital for Special Surgery scores (C), range of motion (D), deviation from neutral mechanical axis (E), mechanical axis alignment outliers (F), intraoperative blood loss (G) and operative time (H) between robotic-assisted total knee arthroplasty (RATKA) and conventional total knee arthroplasty (conventional TKA) including only studies that employed commonly-used robotic systems using random-effects model and inverse variance method. For abbreviations and Risk of Bias A–E, see Figure 3.

**A**

| WOMAC Study | RATKA mean (SD) | n | Conventional TKA mean (SD) | n | Weight % | Difference mean (95%CI) | Risk of bias A B C D E |
|---|---|---|---|---|---|---|---|
| Song 2011 (23) | 11 (4.5) | 30 | 13 (6.6) | 30 | 0.8 | −2.0 (−4.9 to 0.86) | + + + ? ? |
| Song 2013 (24) | 29 (4.4) | 29 | 30 (7.5) | 24 | 0.6 | −1.1 (−4.5 to 2.3) | + + − ? ? |
| Kim 2020 (27) | 18 (14) | 674 | 19 (15) | 674 | 2.8 | −1.0 (−2.6 to 0.55) | + + + ? + |
| Lychagin 2022 (33) | 1.3 (0.5) | 33 | 1.6 (0.6) | 35 | 95.8 | −0.25 (−0.51 to 0.01) | + + + + − |
| Li 2022 (32) | 81 (46) | 69 | 72 (44) | 74 | 0.0 | 8.5 (−6.4 to 23) | ? + + ? ? |
| Xu 2022 (31) | 13 (13) | 37 | 19 (31) | 35 | 0.1 | −6.9 (−18 to 4.3) | ? + + ? + |
| Total | | 872 | | 872 | | −0.29 (−0.55 to −0.03) | |

Heterogeneity: Chi² = 5.2, df = 5 (p = 0.4); I² = 3%
Test for overall effect Z = 2.2 (p = 0.03)

Favors RATKA — Favors conventional TKA

**B**

| KSS Study | Conventional TKA mean (SD) | n | RATKA mean (SD) | n | Weight % | Difference mean (95%CI) | Risk of bias A B C D E |
|---|---|---|---|---|---|---|---|
| Park 2007 (22) | 91 (4.9) | 30 | 92 (2.9) | 32 | 7.4 | –0.70 (–2.7 to 1.3) | |
| Liow 2017 (25) | 88 (11) | 29 | 82 (15) | 31 | 0.7 | 6.1 (–0.41 to 13) | |
| Kim 2020 (27) | 92 (6.0) | 674 | 93 (5.0) | 674 | 86.5 | –1.0 (–1.6 to –0.41) | |
| Li 2022 (32) | 68 (10) | 74 | 70 (7.3) | 69 | 3.7 | –2.2 (–5.1 to 0.67) | |
| Xu 2022 (31) | 66 (7.4) | 35 | 67 (10) | 37 | 1.8 | –1.0 (–5.1 to 3.1) | |
| Total | | 842 | | 843 | | –0.97 (–1.5 to –0.42) | |

Heterogeneity: Chi² = 5.3, df = 4 (p = 0.3); I² = 25%
Test for overall effect Z = 3.5 (p < 0.001)

–20 –10 0 10 20
Favors RATKA   Favors conventional TKA

**C**

| HSS score Study | Conventional TKA mean (SD) | n | RATKA mean (SD) | n | Weight % | Difference mean (95%CI) | Risk of bias A B C D E |
|---|---|---|---|---|---|---|---|
| Song 2011 (23) | 95 (5.0) | 30 | 95 (4.0) | 30 | 52.6 | –0.50 (–2.8 to 1.8) | |
| Song 2013 (24) | 95 (6.7) | 24 | 96 (4.0) | 29 | 29.7 | –1.0 (–4.1 to 2.1) | |
| Li 2022 (32) | 79 (12) | 74 | 80 (12) | 69 | 17.8 | –0.65 (–4.6 to 3.3) | |
| Total | | 128 | | 128 | | –0.67 (–2.3 to 0.99) | |

Heterogeneity: Chi² = 0.07, df = 2 (p = 0.97); I² = 0%
Test for overall effect Z = 0.8 (p = 0.4)

–20 –10 0 10 20
Favors RATKA   Favors conventional TKA

**D**

| ROM Study | Conventional TKA mean (SD) | n | RATKA mean (SD) | n | Weight % | Difference mean (95%CI) | Risk of bias A B C D E |
|---|---|---|---|---|---|---|---|
| Park 2007 (22) | 122 (17) | 30 | 118 (9.0) | 32 | 0.6 | 4.0 (–2.8 to 11) | |
| Song 2011 (23) | 129 (13) | 30 | 129 (14) | 30 | 0.6 | 0.0 (–6.7 to 6.7) | |
| Song 2013 (24) | 129 (12) | 50 | 128 (5.1) | 50 | 1.9 | 1.0 (–2.7 to 4.7) | |
| Liow 2017 (26) | 126 (10) | 29 | 117 (16) | 31 | 0.6 | 8.2 (1.6 to 15) | |
| Kim 2020 (27) | 128 (7.0) | 724 | 125 (6.0) | 724 | 57.8 | 3.0 (2.3 to 3.7) | |
| Xu 2022 (31) | 120 (19) | 35 | 115 (11) | 37 | 0.5 | 5.0 (–2.1 to 12) | |
| Li 2022 (32) | 112 (19) | 74 | 115 (18) | 69 | 0.7 | –2.9 (–9.0 to 3.2) | |
| Lychagin 2022 (33) | 116 (1.3) | 35 | 127 (2.1) | 33 | 37.3 | –11 (–12 to –10) | |
| Total | | 1,007 | | 1,006 | | –2.3 (–2.8 to –1.8) | |

Heterogeneity: Chi² = 685, df = 7 (p < 0.001); I² = 99%
Test for overall effect Z = 8.9 (p < 0.001)

–20 –10 0 10 20
Favors RATKA   Favors conventional TKA

**E**

| MA deviation Study | RATKA mean (SD) | n | Conventional TKA mean (SD) | n | Weight % | Difference mean (95%CI) | Risk of bias A B C D E |
|---|---|---|---|---|---|---|---|
| Song 2011 (23) | 0.2 (1.6) | 30 | 1.2 (2.1) | 30 | 4.8 | –1.0 (–1.9 to –0.06) | |
| Song 2013 (24) | 0.5 (1.4) | 50 | 1.2 (2.9) | 50 | 5.4 | –0.70 (–1.6 to 0.19) | |
| Liow 2014 (25) | 1.3 (0.9) | 31 | 1.8 (1.2) | 29 | 14.8 | –0.50 (–1.0 to 0.04) | |
| Kim 2020 (27) | 2.0 (2.0) | 724 | 3.0 (3.0) | 724 | 62.3 | –1.0 (–1.3 to –0.74) | |
| Li 2022 (32) | 1.8 (1.6) | 69 | 3.0 (2.7) | 74 | 8.0 | –1.2 (–2.0 to –0.49) | |
| Vaidya 2022 (30) | 1.8 (1.2) | 30 | 3.0 (2.4) | 30 | 4.7 | –1.2 (–2.2 to –0.24) | |
| Total | | 934 | | 937 | | –0.94 (–1.1 to –0.73) | |

Heterogeneity: Chi² = 3.9, df = 5 (p = 0.6); I² = 0%
Test for overall effect Z = 8.9 (p < 0.001)

–4 –2 0 2 4
Favors RATKA   Favors conventional TKA

**F**

| Blood loss Study | RATKA mean (SD) | n | Conventional TKA mean (SD) | n | Weight % | Difference mean (95%CI) | Risk of bias A B C D E |
|---|---|---|---|---|---|---|---|
| Song 2011 (23) | 587 (385) | 30 | 816 (425) | 30 | 0.3 | –229 (–435 to –24) | |
| Song 2013 (24) | 613 (318) | 50 | 933 (467) | 50 | 0.5 | –320 (–477 to –163) | |
| Kim 2020 (27) | 261 (107) | 700 | 255 (110) | 706 | 98.9 | 6.0 (–5.4 to 17) | |
| Xu 2022 (31) | 933 (454) | 37 | 863 (458) | 35 | 0.3 | 70 (–141 to 280) | |
| Total | | 817 | | 821 | | 3.8 (–7.5 to 15) | |

Heterogeneity: Chi² = 22, df = 3 (p < 0.001); I² = 86%
Test for overall effect Z = 0.66 (p = 0.5)

–500 –250 0 250 500
Favors RATKA   Favors conventional TKA

**G**

| Operative time Study | RATKA mean (SD) | n | Conventional TKA mean (SD) | n | Weight % | Difference mean (95%CI) | Risk of bias A B C D E |
|---|---|---|---|---|---|---|---|
| Xu 2022 (31) | 154 (21) | 37 | 115 (30) | 35 | 0.4 | 39 (27 to 51) | |
| Thiengwittayaporn 2021 (29) | 70 (12) | 75 | 62 (10) | 77 | 5.0 | 8.2 (4.7 to 12) | |
| Song 2013 (24) | 99 (11) | 50 | 74 (10) | 50 | 3.7 | 25 (21 to 29) | |
| Song 2011 (23) | 95 (18) | 30 | 70 (15) | 30 | 0.9 | 25 (17 to 33) | |
| Liow 2017 (36) | 91 (10) | 31 | 93 (14) | 29 | 1.6 | –2.0 (–8.2 to 4.2) | |
| Kim 2020 (27) | 97 (11) | 700 | 69 (6.3) | 706 | 76.7 | 28 (27 to 29) | |
| Kayani 2021 (28) | 61 (3.1) | 15 | 62 (3.4) | 15 | 11.6 | –1.0 (–3.3 to 1.3) | |
| Total | | 938 | | 942 | | 23 (22 to 24) | |

Heterogeneity: Chi² = 664, df = 6 (p < 0.001); I² = 99%
Test for overall effect Z = 57 (p < 0.001)

–50 –25 0 25 50
Favors RATKA   Favors conventional TKA

**H**

| Hospital stay Study | RATKA mean (SD) | n | Conventional TKA mean (SD) | n | Weight % | Difference mean (95%CI) | Risk of bias A B C D E |
|---|---|---|---|---|---|---|---|
| Liow 2017 (36) | 5.2 (2.3) | 31 | 5.8 (3.8) | 29 | 25.9 | −0.60 (−2.2 to 1.0) | |
| Xu 2022 (31) | 9.1 (2.0) | 37 | 8.4 (2.1) | 35 | 74.1 | 0.70 (−0.25 to 1.7) | |
| Total | | 68 | | 64 | | 0.36 (−0.45 to 1.2) | |

Heterogeneity: Chi² = 1.9, df = 1 (p = 0.2); I² = 47%
Test for overall effect Z = 0.87 (p = 0.4)



Figure 8. Forest plots showing the Western Ontario and McMaster University Osteoarthritis Index scores (A), Knee Society Scores (B), Hospital for Special Surgery scores (C), range of motion (D), deviation from neutral mechanical axis (E), intraoperative blood loss (F), operative time (G), and hospital length of stay (H) between robotic-assisted total knee arthroplasty (RATKA) and conventional total knee arthroplasty (conventional TKA) analyzed with the fixed-effects models using fixed-effects model and inverse variance method. For abbreviations and Risk of Bias A–E, see Figure 3.

**A**

| WOMAC Study | RATKA mean (SD) | n | Conventional TKA mean (SD) | n | Weight % | Difference mean (95%CI) | Risk of bias A B C D E |
|---|---|---|---|---|---|---|---|
| **FU < 1 year** | | | | | | | |
| Li 2022 (32) | 81 (46) | 69 | 72 (44) | 74 | 0.1 | 8.5 (−6.4 to 23) | |
| Xu 2022 (31) | 13 (13) | 37 | 19 (31) | 35 | 0.1 | −6.9 (−18 to 4.3) | |
| Subtotal | | 106 | | 109 | 0.2 | −0.0 (−15 to 15) | |

Heterogeneity: Tau = 73; Chi² = 2.6, df = 1 (p = 0.1); I² = 62%
Test for overall effect Z = 0.0 (p = 1)

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| **FU ≥ 1 year** | | | | | | | |
| Song 2011 (23) | 11 (4.5) | 30 | 13 (6.6) | 30 | 2.1 | −2.0 (−4.9 to 0.86) | |
| Song 2013 (24) | 29 (4.4) | 29 | 30 (7.5) | 24 | 1.5 | −1.1 (−4.5 to 2.3) | |
| Kim 2020 (27) | 18 (14) | 674 | 19 (15) | 674 | 7.0 | −1.0 (−2.6 to 0.55) | |
| Lychagin 2022 (33) | 1.3 (0.5) | 33 | 1.6 (0.6) | 35 | 89.2 | −0.25 (−0.51 to 0.01) | |
| Subtotal | | 766 | | 763 | 99.8 | −0.29 (−0.55 to −0.03) | |

Heterogeneity: Tau = 0.0; Chi² = 2.5, df = 3 (p = 0.5); I² = 0%
Test for overall effect Z = 2.2 (p = 0.03)

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| Total | | 872 | | 872 | | −0.35 (−0.78 to 0.07) | |

Heterogeneity: Tau = 0.03; Chi² = 5.2, df = 5 (p = 0.4); I² = 3%
Test for overall effect Z = 1.7 (p = 0.1)
Test for subgroup differences: Chi² = 0.0, df = 1 (p = 0.97); I² = 0%

**B**

| KSS Study | Conventional TKA mean (SD) | n | RATKA mean (SD) | n | Weight % | Difference mean (95%CI) | Risk of bias A B C D E |
|---|---|---|---|---|---|---|---|
| **FU < 1 year** | | | | | | | |
| Li 2022 (32) | 68 (10) | 74 | 70 (7.3) | 69 | 11.8 | −2.2 (−5.1 to 0.67) | |
| Xu 2022 (31) | 66 (7.4) | 35 | 67 (10) | 37 | 6.3 | −1.0 (−5.1 to 3.1) | |
| Subtotal | | 109 | | 106 | 18.1 | −1.8 (−4.2 to 0.54) | |

Heterogeneity: Tau = 0.0; Chi² = 0.22, df = 1 (p = 0.6); I² = 0%
Test for overall effect Z = 1.5 (p = 0.1)

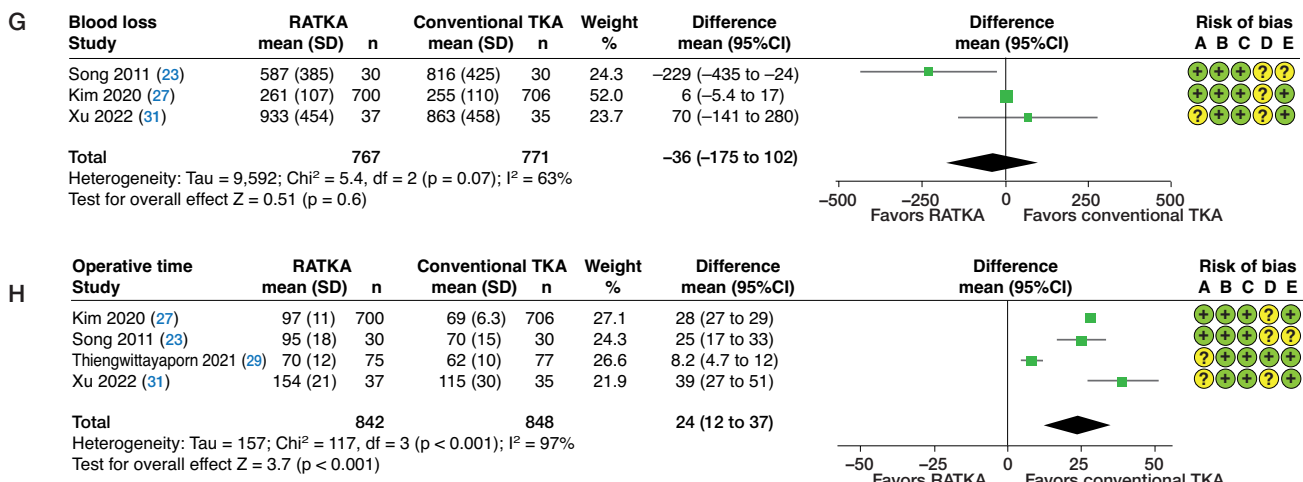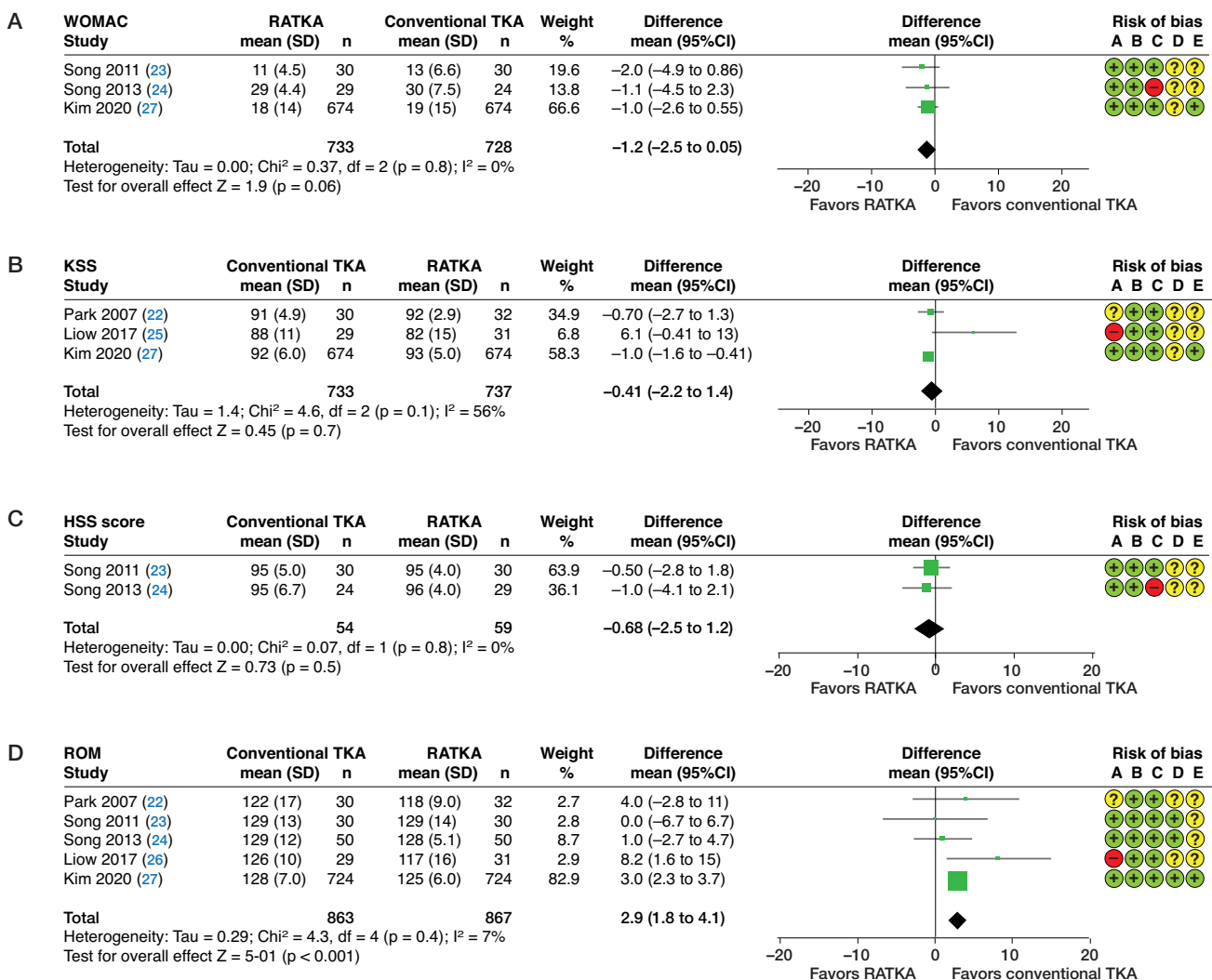| | | | | | | | |
|---|---|---|---|---|---|---|---|
| **FU ≥ 1 year** | | | | | | | |
| Park 2007 (22) | 91 (4.9) | 30 | 92 (2.9) | 32 | 20.3 | −0.70 (−2.7 to 1.3) | |
| Liow 2017 (25) | 88 (11) | 29 | 82 (15) | 31 | 2.6 | 6.1 (−0.41 to 13) | |
| Kim 2020 (27) | 92 (6.0) | 674 | 93 (5.0) | 674 | 59.0 | −1.0 (−1.6 to −0.41) | |
| Subtotal | | 733 | | 737 | 81.9 | −0.41 (−2.2 to 1.4) | |

Heterogeneity: Tau = 1.4; Chi² = 4.6, df = 2 (p = 0.1); I² = 56%
Test for overall effect Z = 0.45 (p = 0.7)

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| Total | | 842 | | 843 | | −0.89 (−2.0 to 0.18) | |

Heterogeneity: Tau = 0.42; Chi² = 5.3, df = 4 (p = 0.3); I² = 25%
Test for overall effect Z = 1.6 (p = 0.1)
Test for subgroup differences: Chi² = 0.85, df = 1 (p = 0.4); I² = 0%

**C**

| HSS score Study | Conventional TKA mean (SD) | n | RATKA mean (SD) | n | Weight % | Difference mean (95%CI) | Risk of bias A B C D E |
|---|---|---|---|---|---|---|---|
| **FU < 1 year** | | | | | | | |
| Li 2022 (32) | 79 (12) | 74 | 80 (12) | 69 | 17.8 | −0.65 (−4.6 to 3.3) | |
| Subtotal | | 74 | | 69 | 17.8 | −0.65 (−4.6 to 3.3) | |

Heterogeneity: Not applicable
Test for overall effect Z = 0.32 (p = 0.8)

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| **FU ≥ 1 year** | | | | | | | |
| Song 2011 (23) | 95 (5.0) | 30 | 95 (4.0) | 30 | 52.6 | −0.50 (−2.8 to 1.8) | |
| Song 2013 (24) | 95 (6.7) | 24 | 96 (4.0) | 29 | 29.7 | −1.0 (−4.1 to 2.1) | |
| Subtotal | | 54 | | 59 | 82.2 | −0.68 (−2.5 to 1.2) | |

Heterogeneity: Tau = 0.00; Chi² = 0.07, df = 1 (p = 0.8); I² = 0%
Test for overall effect Z = 0.73 (p = 0.5)

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| Total | | 128 | | 128 | | −0.67 (−2.3 to 0.99) | |

Heterogeneity: Tau = 0.00; Chi² = 0.07, df = 2 (p = 0.97); I² = 0%
Test for overall effect Z = 0.8 (p = 0.4)
Test dor subgroup differences: Chi² = 0.00, df = 1 (p = 0.99); I² = 0%

**D**

| ROM Study | Conventional TKA mean (SD) | n | RATKA mean (SD) | n | Weight % | Difference mean (95%CI) | | Risk of bias A B C D E |
|---|---|---|---|---|---|---|---|---|
| **FU < 1 year** | | | | | | | | |
| Park 2007 (22) | 122 (17) | 30 | 118 (9.0) | 32 | 11.9 | 4.0 (−2.8 to 11) | | |
| Song 2011 (23) | 129 (13) | 30 | 129 (14) | 30 | 12.0 | 0.0 (−6.7 to 6.7) | | |
| Song 2013 (24) | 129 (12) | 50 | 128 (5.1) | 50 | 13.0 | 1.0 (−2.7 to 4.7) | | |
| Li 2022 (32) | 112 (19) | 74 | 115 (18) | 69 | 12.2 | −2.9 (−9.0 to 3.2) | | |
| Xu 2022 (31) | 120 (19) | 35 | 115 (11) | 37 | 11.8 | 5.0 (−2.1 to 12) | | |
| Subtotal | | 219 | | 218 | 60.9 | 1.1 (−1.4 to 3.6) | | |

Heterogeneity: Tau = 0.0; Chi² = 3.6, df = 4 (p = 0.5); I² = 0%
Test for overall effect Z = 0.87 (p = 0.4)

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| **FU ≥ 1 year** | | | | | | | | |
| Liow 2017 (26) | 126 (10) | 29 | 117 (16) | 31 | 12.0 | 8.2 (1.6 to 15) | | |
| Kim 2020 (27) | 128 (7.0) | 724 | 125 (6.0) | 724 | 13.5 | 3.0 (2.3 to 3.7) | | |
| Lychagin 2022 (33) | 116 (1.3) | 35 | 127 (2.1) | 33 | 13.5 | −11 (−12 to −10) | | |
| Subtotal | | 788 | | 788 | 39.1 | −0.26 (−12 to 11) | | |

Heterogeneity: Tau = 99; Chi² = 674, df = 2 (p < 0.001); I² = 100%
Test for overall effect Z = 0.04 (p = 0.97)

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| Total | | 1,007 | | 1,006 | | 0.73 (−6.0 to 7.5) | | |

Heterogeneity: Tau = 88; Chi² = 685, df = 7 (p < 0.001); I² = 99%
Test for overall effect Z = 0.21 (p = 0.8)
Test for subgroup differences: Chi² = 0.05, df = 1 (p = 0.8); I² = 0%

−20  −10  0  10  20
Favors RATKA   Favors conventional TKA

Figure 9. Forest plots showing subgroup analyses based on mean follow-up time for the Western Ontario and McMaster University Osteoarthritis Index scores (A), Knee Society Scores (B), Hospital for Special Surgery scores (C), and range of motion (D) between robotic-assisted total knee arthroplasty (RATKA) and conventional total knee arthroplasty (conventional TKA) using random-effects model and inverse variance method. For abbreviations and Risk of Bias A–E, see Figure 3.

**A**

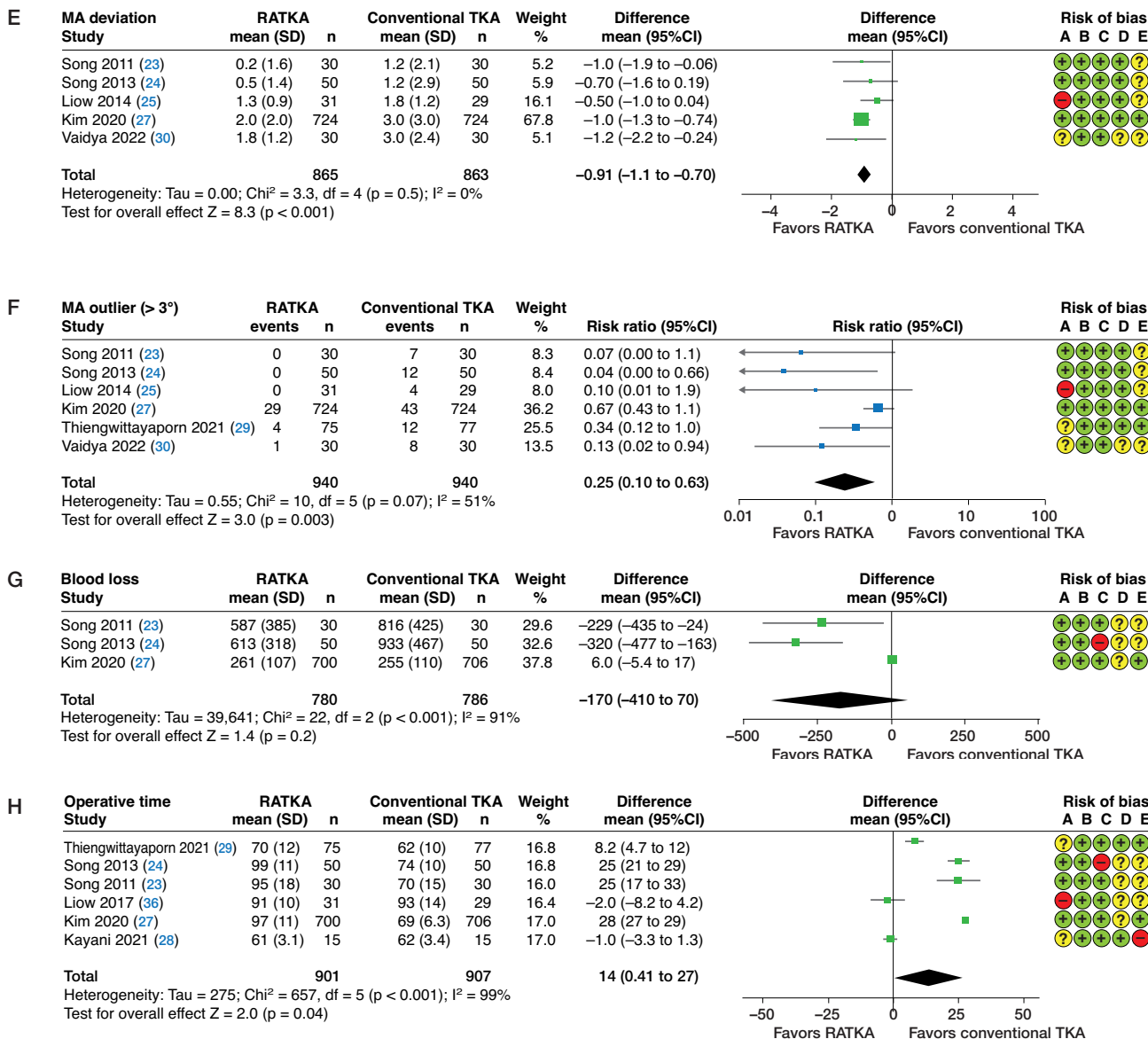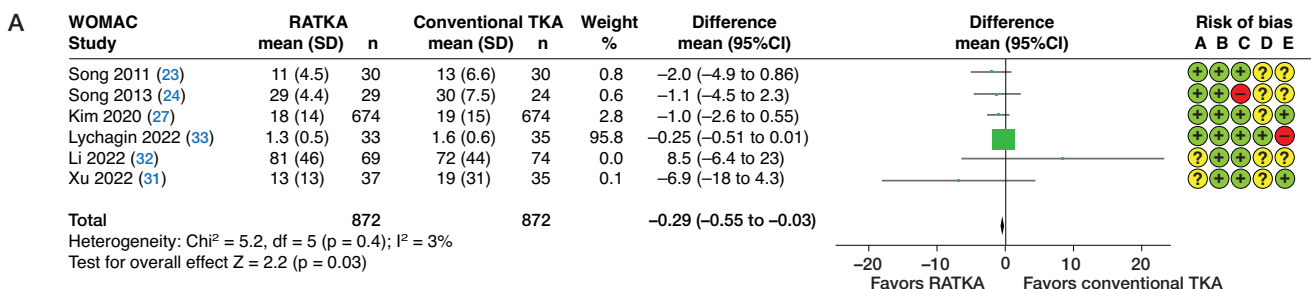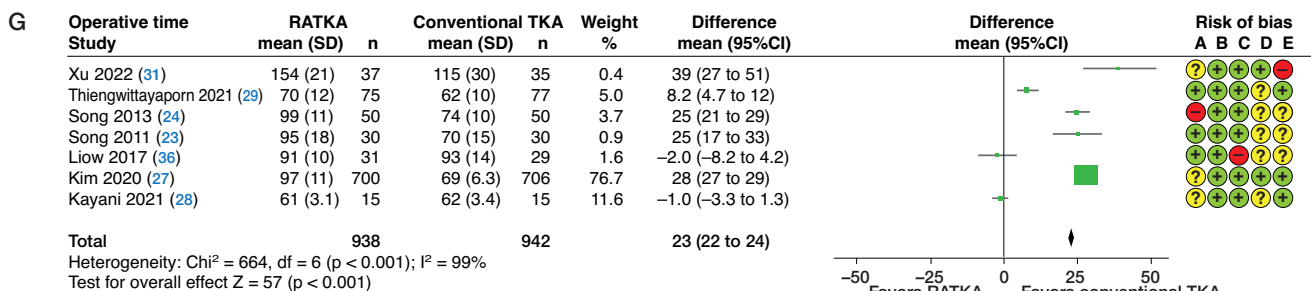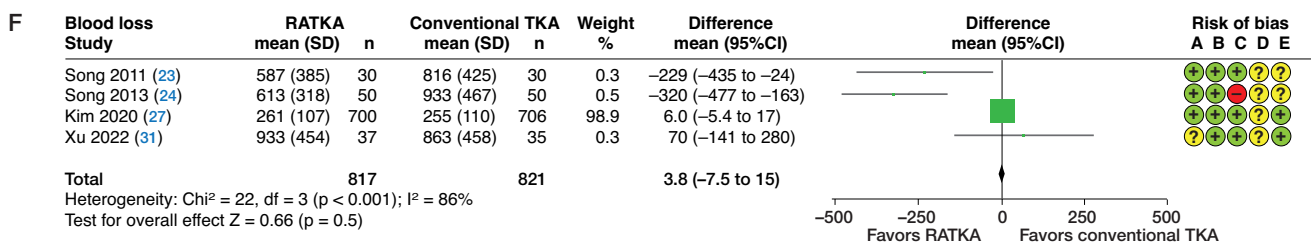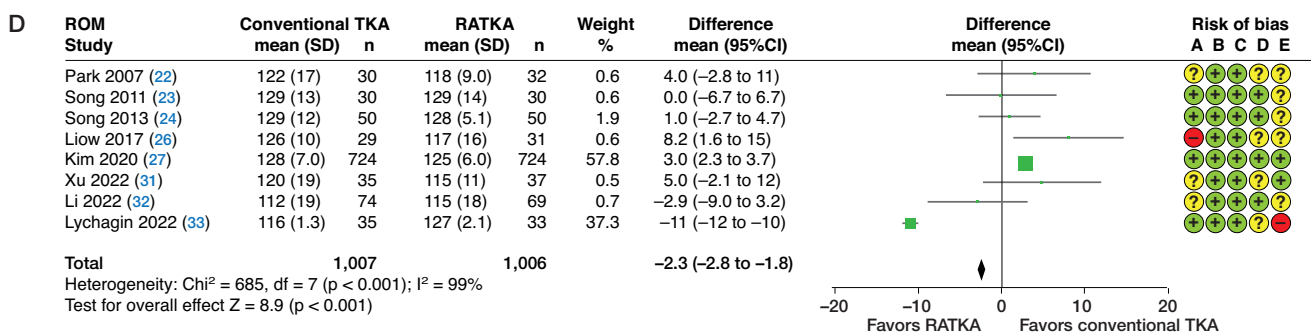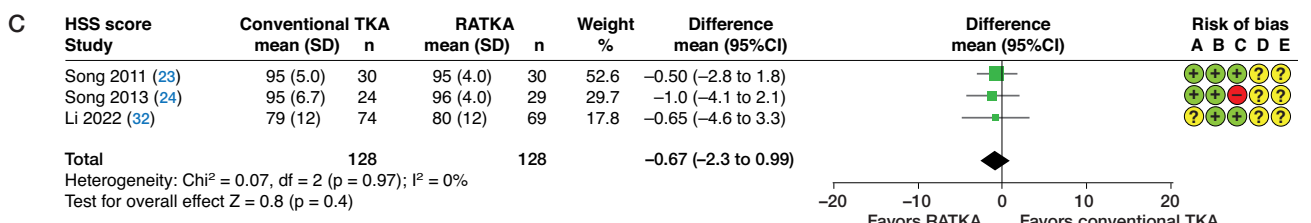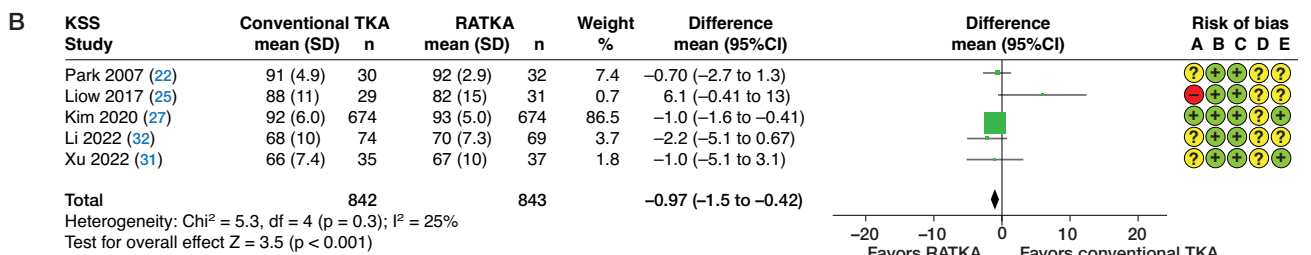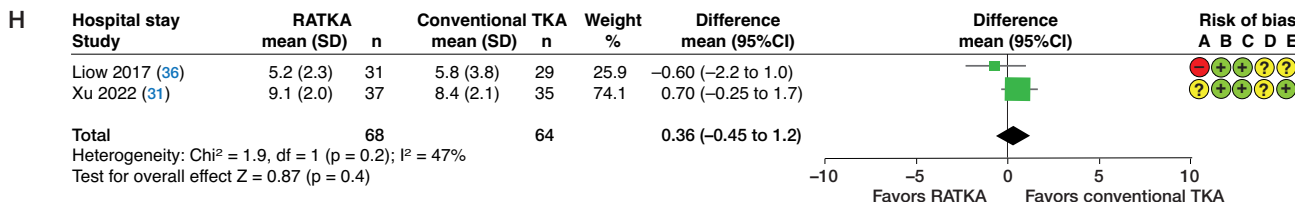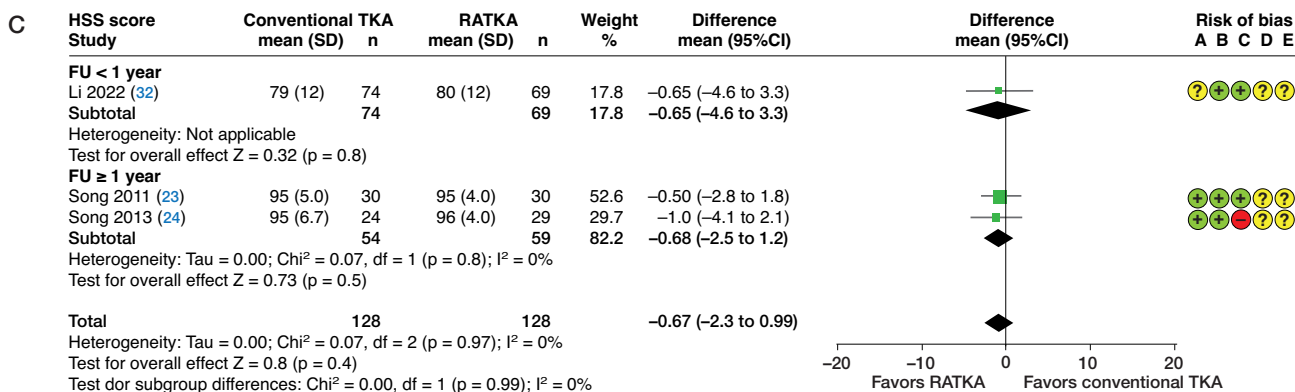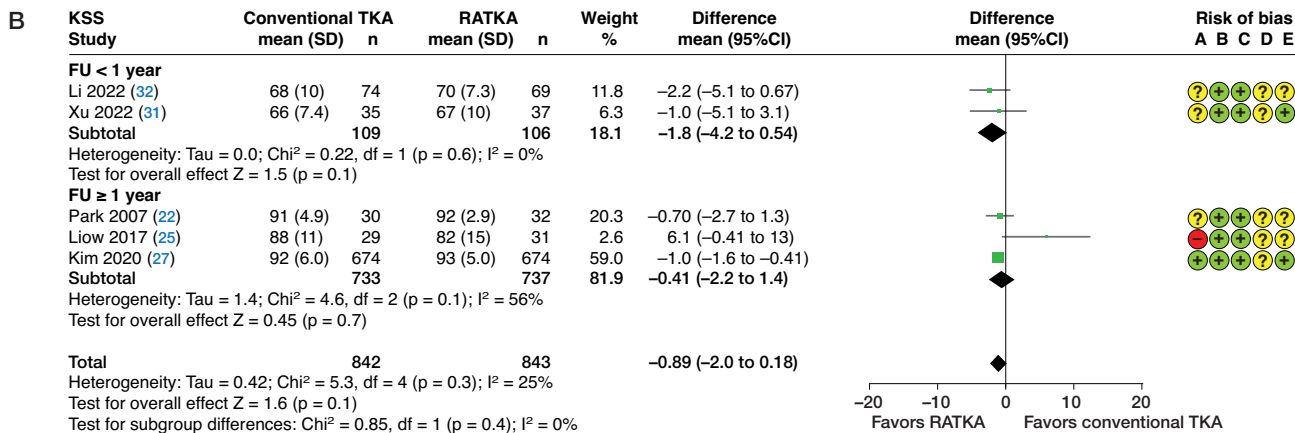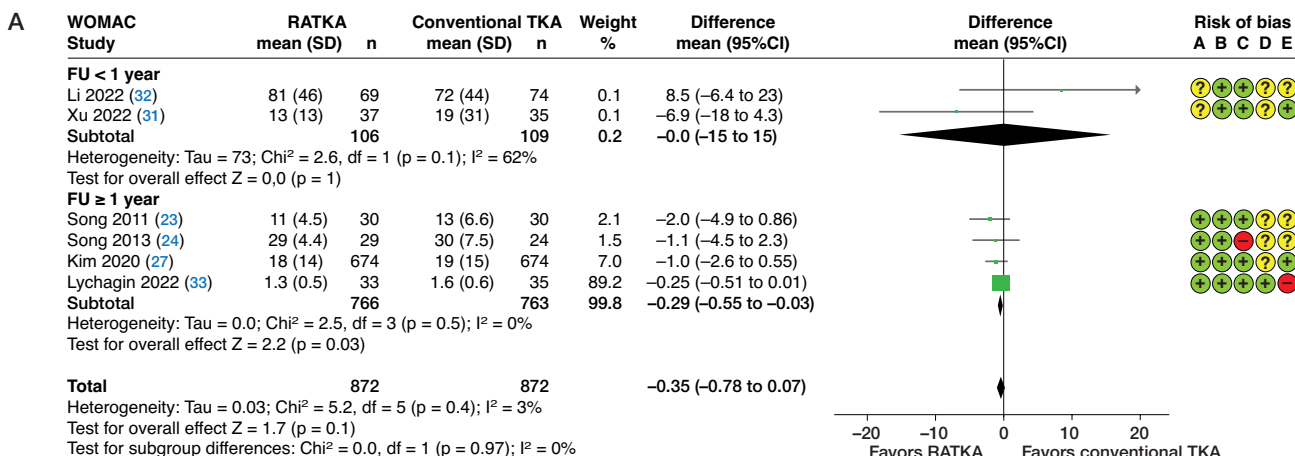| WOMAC Study | RATKA mean (SD) | n | Conventional TKA mean (SD) | n | Weight % | Difference mean (95%CI) | | Risk of bias A B C D E |
|---|---|---|---|---|---|---|---|---|
| **Crucite-retaining** | | | | | | | | |
| Song 2011 (23) | 11 (4.5) | 30 | 13 (6.6) | 30 | 19.3 | −2.0 (−4.9 to 0.86) | | |
| Song 2013 (24) | 29 (4.4) | 29 | 30 (7.5) | 24 | 13.6 | −1.1 (−4.5 to 2.3) | | |
| Subtotal | | 59 | | 54 | 33.0 | −1.6 (−3.8 to 0.56) | | |

Heterogeneity: Tau = 0.00; Chi² = 0.16, df = 1 (p = 0.7); I² = 0%
Test for overall effect Z = 1.5 (p = 0.1)

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| **Posterior–stabilizing** | | | | | | | | |
| Kim 2020 (27) | 18 (14) | 674 | 19 (15) | 674 | 65.8 | −1.0 (−2.6 to 0.55) | | |
| Xu 2022 (31) | 13 (13) | 37 | 19 (31) | 35 | 1.2 | −6.9 (−18 to 4.3) | | |
| Subtotal | | 711 | | 709 | 67.0 | −1.2 (−3.4 to 0.96) | | |

Heterogeneity: Tau = 0.65; Chi² = 1.0, df = 1 (p = 0.3); I² = 4%
Test for overall effect Z = 1.1 (p = 0.3)

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| Total | | 770 | | 763 | | −1.3 (−2.5 to −0.02) | | |

Heterogeneity: Tau = 0.00; Chi² = 1.3, df = 3 (p = 0.7); I² = 0%
Test for overall effect Z = 2.0 (p = 0.05)
Test for subgroup differences: Chi² = 0.07, df = 1 (p = 0.8); I² = 0%

−20  −10  0  10  20
Favors RATKA   Favors conventional TKA

**B**

| ROM Study | Conventional TKA mean (SD) | n | RATKA mean (SD) | n | Weight % | Difference mean (95%CI) | | Risk of bias A B C D E |
|---|---|---|---|---|---|---|---|---|
| **Crucite-retaining** | | | | | | | | |
| Song 2011 (23) | 129 (13) | 30 | 129 (14) | 30 | 0.9 | 0.0 (−6.7 to 6.7) | | |
| Song 2013 (24) | 129 (12) | 50 | 128 (5.1) | 50 | 3.0 | 1.0 (−2.7 to 4.7) | | |
| Subtotal | | 80 | | 80 | 4.0 | 0.77 (−2.5 to 4.0) | | |

Heterogeneity: Tau = 0.0; Chi² = 0.06, df = 1 (p = 0.8); I² = 0%
Test for overall effect Z = 0.46 (p = 0.6)

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| **Posterior–stabilizing** | | | | | | | | |
| Park 2007 (22) | 122 (17) | 30 | 118 (9.0) | 32 | 0.9 | 4.0 (−2.8 to 11) | | |
| Liow 2017 (26) | 126 (10) | 29 | 117 (16) | 31 | 1.0 | 8.2 (1.6 to 15) | | |
| Kim 2020 (27) | 128 (7.0) | 724 | 125 (6.0) | 724 | 93.3 | 3.0 (2.3 to 3.7) | | |
| Xu 2022 (31) | 120 (19) | 35 | 115 (11) | 37 | 0.8 | 5.0 (−2.1 to 12) | | |
| Subtotal | | 818 | | 824 | 96.0 | 3.1 (2.4 to 3.7) | | |

Heterogeneity: Tau = 0.0; Chi² = 2.7, df = 3 (p = 0.4); I² = 0%
Test for overall effect Z = 9.1 (p < 0.001)

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| Total | | 898 | | 904 | | 3.0 (2.3 to 3.6) | | |

Heterogeneity: Tau = 0.0; Chi² = 4.6, df = 5 (p = 0.5); I² = 0%
Test for overall effect Z = 9.0 (p < 0.001)
Test for subgroup differences: Chi² = 1.9, df = 1 (p = 0.2); I² = 46%

−20  −10  0  10  20
Favors RATKA   Favors conventional TKA

**C**

| MA deviation Study | RATKA mean (SD) | n | Conventional TKA mean (SD) | n | Weight % | Difference mean (95%CI) | Difference mean (95%CI) | Risk of bias A B C D E |
|---|---|---|---|---|---|---|---|---|
| **Crucite-retaining** | | | | | | | | |
| Song 2011 (23) | 0.2 (1.6) | 30 | 1.2 (2.1) | 30 | 5.2 | −1.0 (−1.9 to −0.06) | | + + + + ? |
| Song 2013 (24) | 0.5 (1.4) | 50 | 1.2 (2.9) | 50 | 5.9 | −0.70 (−1.6 to 0.19) | | + + + + ? |
| Subtotal | | 80 | | 80 | 11.1 | −0.84 (−1.5 to −0.19) | | |

Heterogeneity: Tau = 0.00; Chi² = 0.2, df = 1 (p = 0.7); I² = 0%
Test for overall effect Z = 2.5 (p = 0.01)

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| **Posterior–stabilizing** | | | | | | | | |
| Liow 2014 (25) | 1.3 (0.9) | 31 | 1.8 (1.2) | 29 | 16.1 | −0.50 (−1.0 to 0.04) | | − + + + ? |
| Kim 2020 (27) | 2.0 (2.0) | 724 | 3.0 (3.0) | 724 | 67.8 | −1.0 (−1.3 to −0.74) | | + + + + + |
| Vaidya 2022 (30) | 1.8 (1.2) | 30 | 3.0 (2.4) | 30 | 5.1 | −1.2 (−2.2 to −0.24) | | ? + + ? ? |
| Subtotal | | 785 | | 783 | 88.9 | −0.88 (−1.2 to −0.53) | | |

Heterogeneity: Tau = 0.04; Chi² = 3.0, df = 2 (p = 0.2); I² = 34%
Test for overall effect Z = 4.9 (p < 0.001)

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| Total | | 865 | | 863 | | −0.91 (−1.1 to −0.70) | | |

Heterogeneity: Tau = 0.00; Chi² = 3.3, df = 4 (p = 0.5); I² = 0%
Test for overall effect Z = 8.3 (p < 0.001)
Test for subgroup differences: Chi² = 0.01, df = 1 (p = 0.9); I² = 0%

−4 −2 0 2 4
Favors RATKA    Favors conventional TKA

**D**

| MA outlier (> 3°) Study | RATKA events | n | Conventional TKA events | n | Weight % | Risk ratio (95%CI) | Risk ratio (95%CI) | Risk of bias A B C D E |
|---|---|---|---|---|---|---|---|---|
| **Crucite-retaining** | | | | | | | | |
| Song 2011 (23) | 0 | 30 | 7 | 30 | 4.6 | 0.07 (0.00 to 1.1) | | + + + + ? |
| Song 2013 (24) | 0 | 50 | 12 | 50 | 4.7 | 0.04 (0.00 to 0.66) | | + + + + ? |
| Subotal | | 80 | | 80 | 9.3 | 0.05 (0.01 to 0.38) | | |

Heterogeneity: Tau = 0.00; Chi² = 0.06, df = 1 (p = 0.8); I² = 0%
Test for overall effect Z = 2.9 (p = 0.003)

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| **Posterior–stabilizing** | | | | | | | | |
| Kim 2020 (27) | 29 | 724 | 43 | 724 | 34.8 | 0.67 (0.43 to 1.1) | | + + + + + |
| Liow 2014 (25) | 0 | 31 | 4 | 29 | 4.5 | 0.10 (0.01 to 1.9) | | − + + + ? |
| Thiengwittayaporn 2021 (29) | 4 | 75 | 12 | 77 | 19.2 | 0.34 (0.12 to 1.0) | | ? + + + + |
| Vaidya 2022 (30) | 1 | 30 | 8 | 30 | 8.2 | 0.13 (0.02 to 0.94) | | ? + + ? ? |
| Xu 2022 (31) | 6 | 35 | 12 | 34 | 24.1 | 0.49 (0.21 to 1.2) | | ? + + + + |
| Subtotal | | 895 | | 894 | 90.7 | 0.48 (0.30 to 0.78) | | |

Heterogeneity: Tau = 0.06; Chi² = 4.9, df = 4 (p = 0.3); I² = 19%
Test for overall effect Z = 3.0 (p = 0.003)

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| Total | | 975 | | 974 | | 0.35 (0.18 to 0.66) | | |

Heterogeneity: Tau = 0.25; Chi² = 10, df = 6 (p = 0.1); I² = 41%
Test for overall effect Z = 3.2 (p = 0.001)
Test for subgroup differences: Chi² = 4.6, df = 1 (p = 0.03); I² = 78%

0.01 0.1 0 10 100
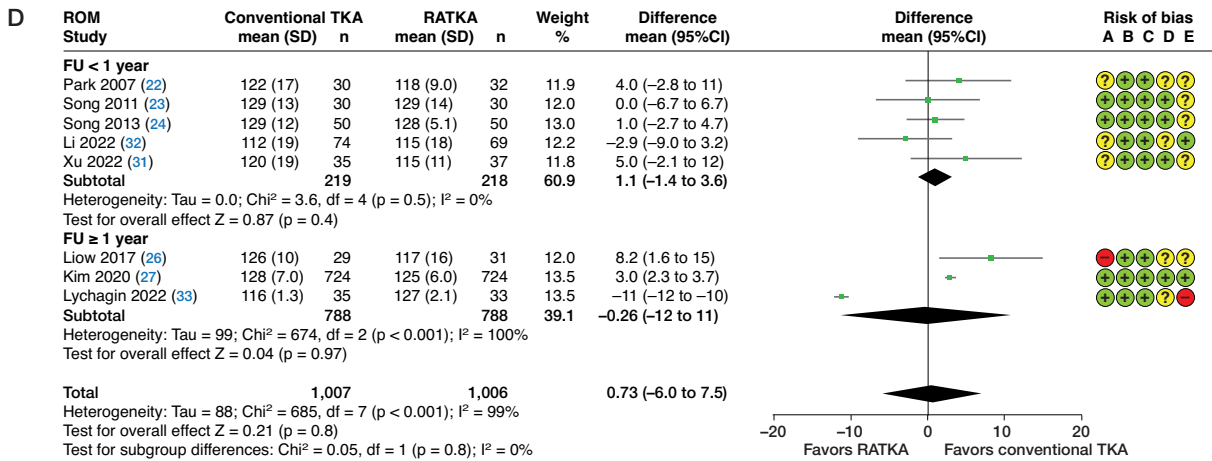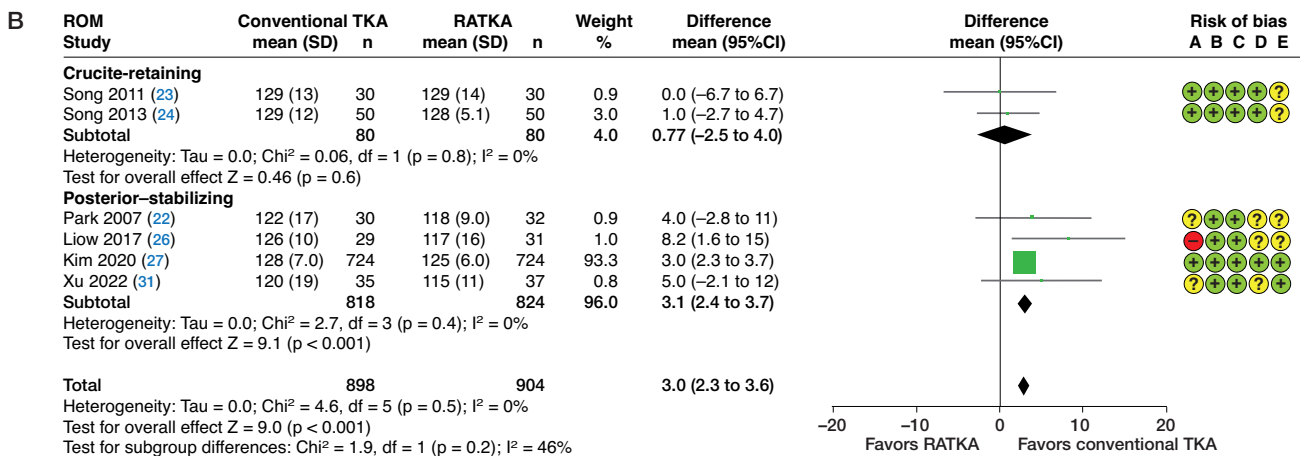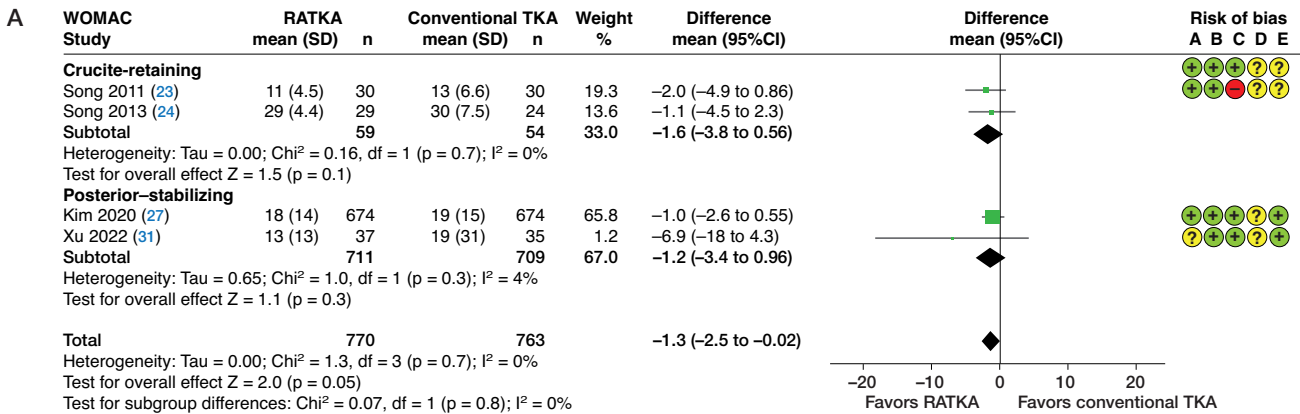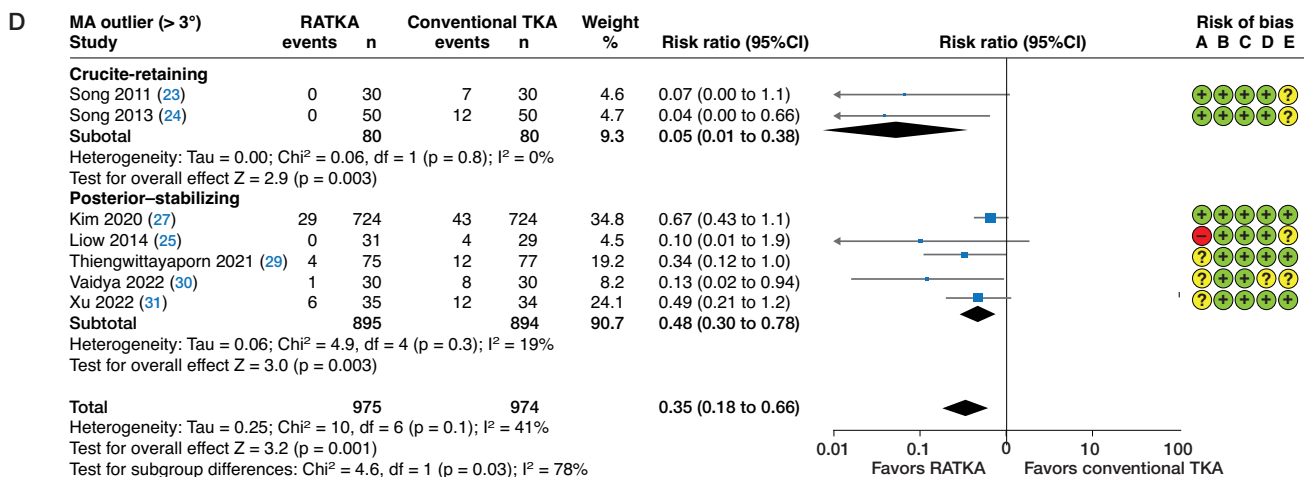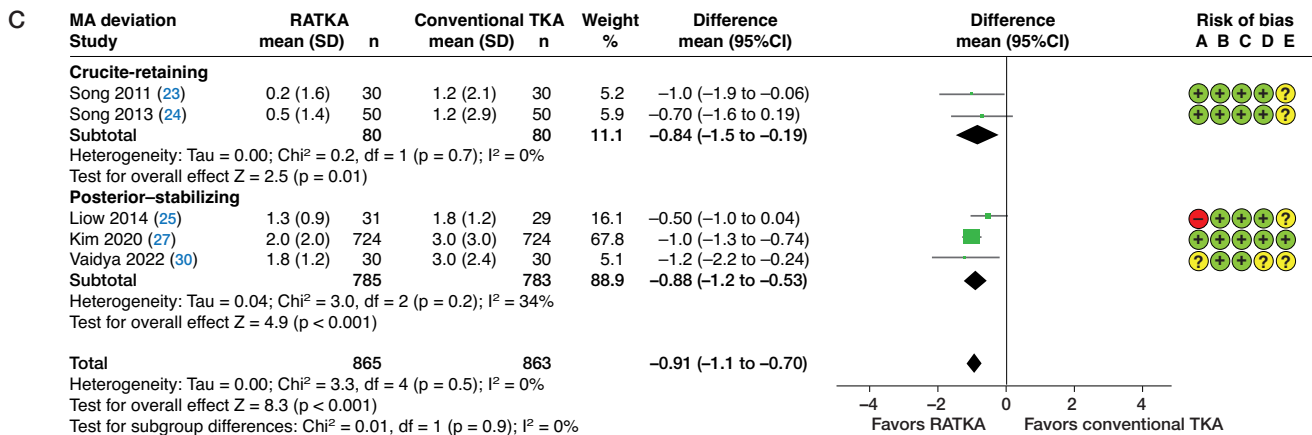Favors RATKA    Favors conventional TKA

Figure 10. Forest plots showing subgroup analyses based on type of prosthesis for the Western Ontario and McMaster University Osteoarthritis Index scores (A), range of motion (B), deviation from neutral mechanical axis (C), and mechanical axis alignment outliers (D) between robotic-assisted total knee arthroplasty (RATKA) and conventional total knee arthroplasty (conventional TKA) using random-effects model and inverse variance method. For abbreviations and Risk of Bias A–E, see Figure 3.