The Journal of Arthroplasty 40 (2025) 3335–3341

Contents lists available at ScienceDirect

The Journal of Arthroplasty

journal homepage: www.arthroplastyjournal.org

Systematic Review and Meta-Analysis

Quality Assessment of Computer-Assisted Navigation Knee Arthroplasty Studies

Antron Spooner, MD [a], Austin G. Helton [a,*], Vincent Lee [b], David Redden, PhD [b], Brent A. Ponce, MD [c], Justin P. Fernicola, MD [c], Randall J. Ruark, MD [c]

[a] Medical College of Georgia, Augusta, Georgia
[b] Edward Via College of Osteopathic Medicine - Auburn Campus, Auburn, Alabama
[c] Jack Hughston Memorial Hospital, Phenix City, Alabama

ARTICLE INFO

Article history:
Received 27 January 2025
Received in revised form
24 May 2025
Accepted 27 May 2025
Available online 6 June 2025

Keywords:
total knee arthroplasty
computer navigation
randomized control trial
transformed detsky
modified cochrane risk-of-bias

ABSTRACT

Background: Computer-navigated total knee arthroplasty (TKA) seeks to improve component alignment and reduce postoperative complications versus conventional TKA. However, the quality of studies evaluating outcomes varies. This review systematically assesses the methodological quality of randomized controlled trials (RCTs) comparing computer-navigated to conventional TKA using established quality assessment tools.

Methods: A systematic review was conducted by searching PubMed, Ovid (Embase), and MEDLINE in April 2024 using the terms "Computer Navigation and Knee Arthroplasty" or "Computer Assisted and Knee Arthroplasty." A total of 119 RCTs met the inclusion criteria. Trial quality was evaluated using the transformed Detsky index, which assesses reporting completeness, and the modified Cochrane risk-of-bias tool, which evaluates methodological rigor and biases.

Results: Of the 119 RCTs, 6.7% were from North America and 42% were from Asia. Single-center trials comprised 78%. The mean transformed Detsky score was 92%, with 98% of studies classified as high quality, reflecting strong adherence to reporting guidelines. Scores improved over time ($P = 0.003$) and were higher in multicenter studies ($P = 0.034$), with no differences by trial location ($P = 0.24$) or primary outcome ($P = 0.19$). However, the mean modified Cochrane risk-of-bias score was 6.2, with 27% rated as high quality, indicating deficiencies in allocation concealment, blinding, and bias mitigation. Modified Cochrane risk-of-bias scores improved over time ($P = 0.027$) and varied by location ($P < 0.001$), with the Australian trials scoring highest and European trials scoring lowest. There were no differences found between single- and multicenter trials ($P = 0.067$) or among primary outcomes ($P = 0.36$).

Conclusions: While RCT-reported standards have improved, methodological weaknesses persist. Future trials should emphasize robust randomization, blinding, and adequate sample sizes to strengthen validity. Multicenter collaborations, standardized reporting, and increased transparency in industry-funded research are essential for improving quality. Addressing these limitations will ensure reliable evidence for TKA decision-making.

Total knee arthroplasty (TKA) is the standard treatment for severe knee osteoarthritis unresponsive to conservative management, providing pain relief and functional improvement. Conventional TKA relies on manual implant alignment using visual inspection and poses a risk of malalignment [1,2]. Implant malalignment is typically defined as a deviation of more than ± 3° from the neutral mechanical axis in the coronal plane, which is linked to higher failure rates, worse functional scores, and increased risk of implant loosening and anterior knee pain [1,2].

Computer navigation emerged as a tool to enhance intraoperative accuracy, reducing alignment outliers and improving coronal plane alignment [3]. Numerous randomized controlled trials (RCTs) have compared its clinical outcomes to conventional TKA. However, despite being considered the gold standard for

One or more of the authors of this paper have disclosed potential or pertinent conflicts of interest, which may include receipt of payment, either direct or indirect, institutional support, or association with an entity in the biomedical field which may be perceived to have potential conflict of interest with this work. For full disclosure statements refer to https://doi.org/10.1016/j.arth.2025.05.108.

* Address correspondence to: Austin G. Helton, Medical College of Georgia, 6262 Veterans Pkwy, Columbus, GA 31909.

https://doi.org/10.1016/j.arth.2025.05.108
0883-5403/

clinical research, RCTs published in orthopaedics have historically demonstrated considerable variability in methodological quality [4–6]. Common issues like small sample sizes yielding low statistical power, inconsistent outcome measures, and heterogeneity in study design compromise the reliability of their findings. Consequently, RCTs evaluating computer navigation in TKA may be subject to the same limitations, making it difficult to draw definitive conclusions for clinical decision-making.

To improve methodological quality, many orthopaedic journals have adopted initiatives such as the Consolidated Standards of Reporting Trials (CONSORT) as publication requirements. These guidelines have led to notable improvements in trial transparency, methodology, and bias reduction [7,8,9–12]. However, compliance remains imperfect [7,8,13,14], with many RCTs still lacking detailed descriptions of randomization procedures, transparency about exclusions and dropouts, and sufficient power to detect meaningful clinical differences [7,8,9–12].

There were no studies that have systematically evaluated the methodological quality of RCTs comparing computer-navigated and conventional TKA. To address this gap, our study uses two established scoring systems—the Detsky index and the modified Cochrane risk-of-bias (mROB) assessment tool—to objectively assess the quality of these RCTs. The Detsky index was selected for its ability to evaluate reporting transparency, while the mROB provides a detailed assessment of bias and methodological rigor.

The primary aims of this study were: (1) to assess the methodological quality of RCTs comparing computer-navigated and conventional TKA using the Detsky index and mROB assessment tools; and (2)to identify factors influencing trial quality (i.e., study design, reporting standards, and regional trends over time).

We hypothesized that RCT quality has improved over time due to increased adherence to reporting standards. However, major variability will persist, with many studies still exhibiting limitations in reporting transparency, randomization procedures, and statistical power. Despite the adoption of CONSORT, inconsistencies will continue to impact study quality.

Methods

Search Strategy

A systematic review was performed using the 2020 Preferred Reporting Items for Systematic Review and Meta-Analysis guidelines [6]. A literature search was conducted on PubMed, Ovid (Embase), and MEDLINE for the terms (Computer Navigation and Knee Arthroplasty OR Computer Assisted and Knee Arthroplasty) published from inception until April 4, 2024. Results were filtered for RCTs, yielding 397 articles. Articles were screened via a 4-phase system (identification, screening, eligibility, and inclusion). This systematic review was registered on International Prospective Register of Systematic Reviews (CRD420251005325).

Eligibility Criteria

All search results were combined through EndNote (Clarivate, Philadelphia, Pennsylvania). Duplicated titles were internally removed, resulting in 222 articles. There were two reviewers (A.S. and A.G.H.) who independently screened the remaining articles by titles and abstracts for relevance. The title review resulted in 127 articles, which were further reduced to 123 after the abstract

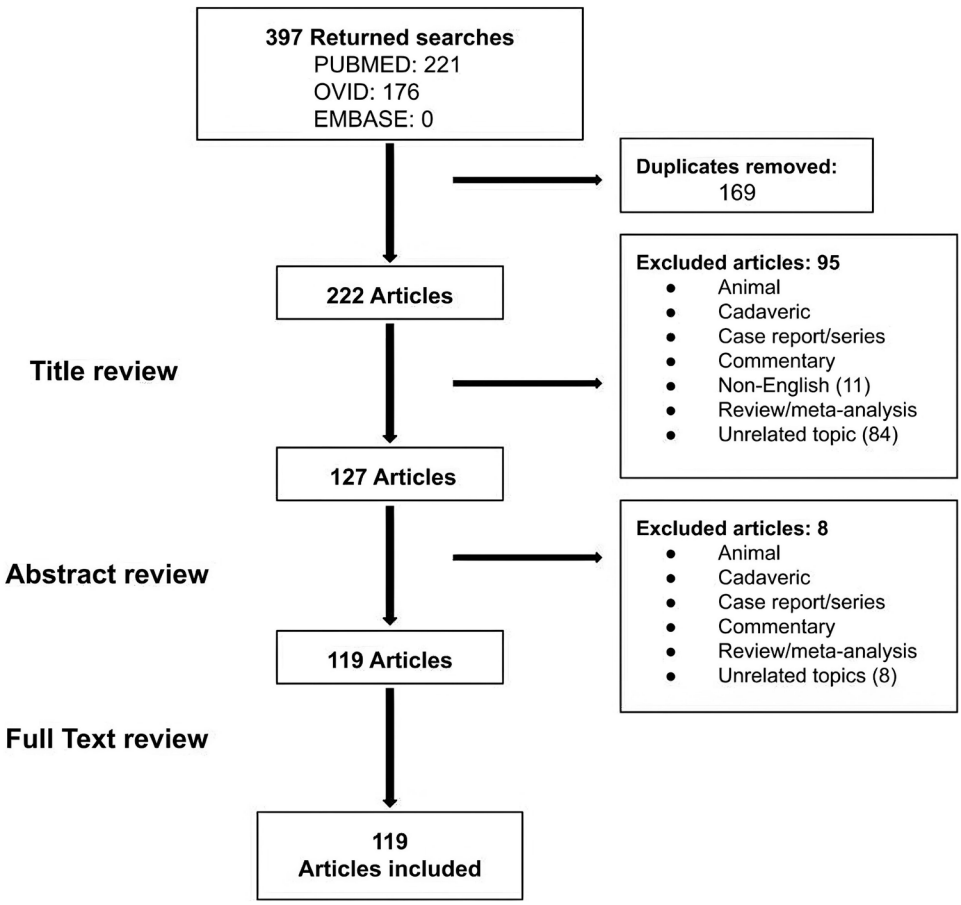

Figure 1. Consolidated Standards of Reporting Trials flow diagram.

review. There were four articles inaccessible, leaving 119 articles for final analysis (Figure 1). A third author (V.L.) was used in disagreements, and discrepancies were resolved by consensus. The inclusion criteria were RCTs mentioning computer-assisted TKA and navigation-assisted TKA (robotic, handheld, electromechanical, and pinless) to ensure a focused comparison of computer navigation versus conventional TKA. The exclusion criteria included non-English text, cohort studies, case–control studies, case series, case reports, meta-analyses, and reviews to limit bias and maintain methodological rigor. Additionally, unrelated studies mentioning hip or shoulder arthroplasty, tendon repairs, patient education, patient-specific instrumentation, unicompartmental, and minimally invasive studies without navigation were excluded.

Extracted Data

The following variables were extracted from each included RCT: year of publication, trial location, number of centers, primary outcome, reported conflict of interest, the inclusion of a CONSORT diagram, type of funding, the longest follow-up time point, and the impact factor of the journal in which the study was published.

Quality Assessment

The reviewers (A.S., A.G.H.) worked independently to conduct quality assessments for each study, with a third author (V.L.) resolving disagreements remaining after discussion. Trials were reviewed using the Detsky quality-of-reporting index and the mROB tools, as they are commonly used measurement instruments of peer-reviewed publications. Both scores were considered the two primary outcome measures. The Detsky score evaluates the quality of reporting based on 14 questions covering five categories, each worth four points for a total possible score of 20 or 21 [15,16]. The score was then converted into a percentage (mean-transformed Detsky score) due to negative trials being scored out of 21 to reduce the ROB in score compilation. Studies scoring > 75% on the transformed score were considered high quality based on prior literature [15].

The mROB assessment evaluates the methodological quality of the study based on the following 10 categories: (1) randomization, (2) allocation concealment, (3) orthopaedic surgeon or treatment provider blinding, (4) assessor blinding, (5) patient blinding, (6) patient follow-up, (7) selective outcome reporting, (8) objectivity of outcomes, (9) adequate sample size, and (10) surgeon experience with robotic TKAs, defined as successfully performing greater than 10 cases based on the current literature [17–19]. The maximum score on this scale is 10 points, indicating a low ROB. Trials scoring eight out of 10 points on the mROB assessment were considered high quality based on prior literature [7].

Data Analyses

The outcome variables (transformed Detsky score and mROB score) were summarized using measures of central location (sample means, sample medians) and dispersion or variability (sample SDs, sample interquartile ranges). The distributional shape of each outcome was assessed by histograms and normal probability plots (Figures 2 and 3). These statistical graphs indicated that any assumption of the outcomes being distributed normally was unreasonable; the graphs indicated skewness to the left of the median. Based upon these graphs, along with the fact that the outcomes are ordinal in nature, we chose to use the nonparametric Kruskal-Wallis procedure to test for differences in outcomes across several factors (funding source, year of study, etc.). For the two non-categorical factors (sample size and journal impact factor), we used Spearman correlation to assess whether they correlated with the outcomes. All analyses and graphs were constructed using SAS 9.4 (SAS Institute Inc, Cary, North Carolina). For interrater agreement, the exact agreements and confidence intervals for each criterion within both Detsky and mROB can be found in Supplementary Tables 1 and 2.

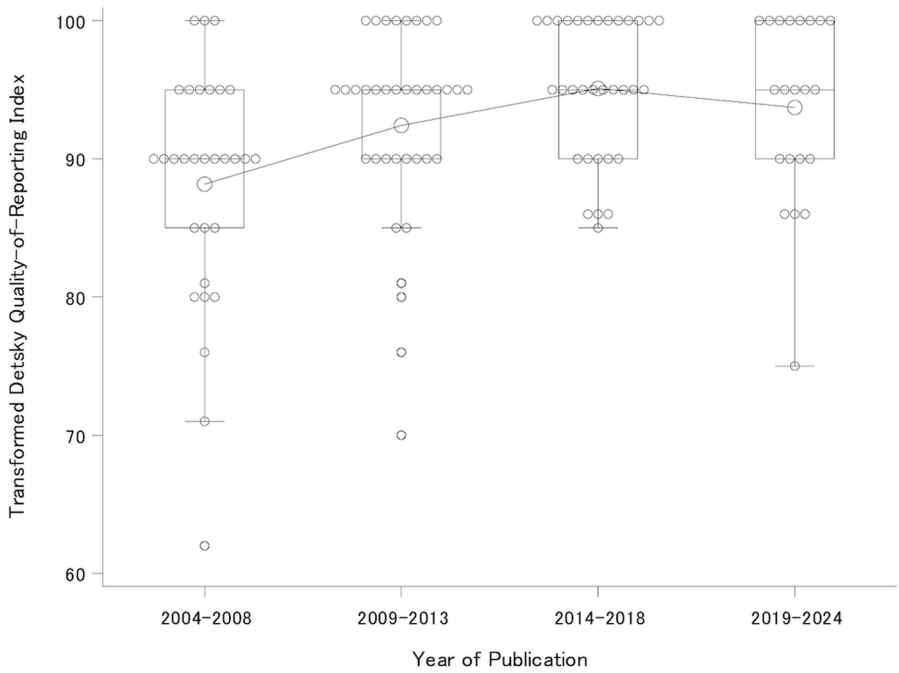

Figure 2. Mean transformed Detsky scores from 2004 to 2024.

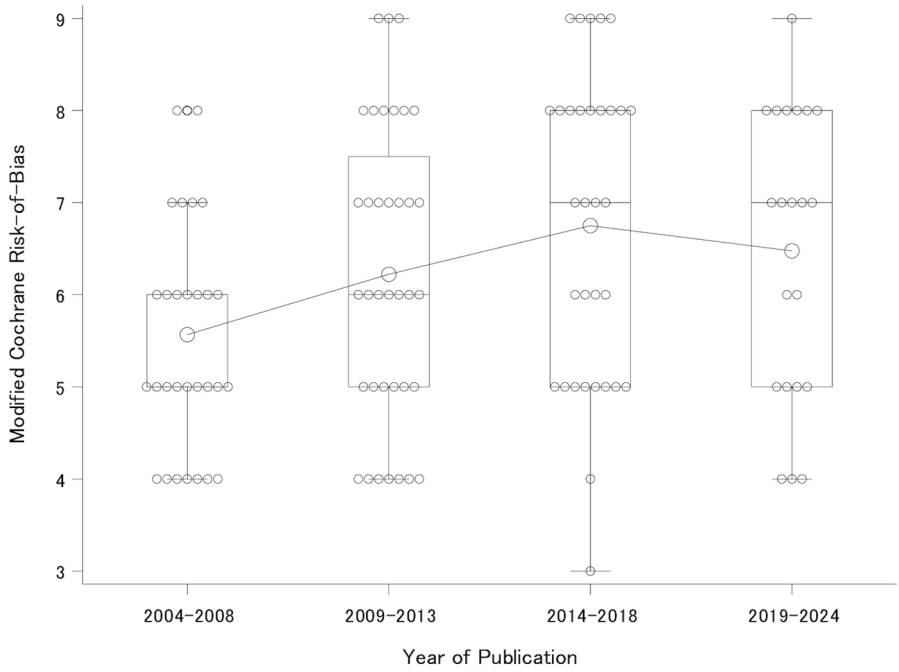

Figure 3. Mean modified Cochrane risk-of-bias scores from 2004 to 2024.

Institutional Review Board Approval

The study was exempt from institutional review board approval given the systematic review nature.

Results

Study Characteristics

Of the 397 articles identified from an initial search, 119 RCTs comparing computer-navigated TKA to conventional TKA were included in our analysis (Figure 1). The breakdown of the extracted data, excluding quality assessment findings, is available in Supplementary Table 3.

Assessment of the Detsky Index Quality Score

The mean transformed Detsky score was 92% ± 7.4, with 98% (117 of 119) rated as high quality. The study characteristics and their relation to the overall transformed Detsky score are shown in Table 1. There was a statistically significant difference between the 4-year periods ($P = 0.003$), with 2014 to 2018 having the highest and 2004 to 2008 having the lowest (95.1 ± 5.1 versus 88.2 ± 8.6). Studies involving multiple centers also showed higher scores (96.5 ± 3.4) than single-center studies (91.8 ± 7.4) ($P = 0.034$). Furthermore, the inclusion of a CONSORT diagram was associated with notably higher scores (96.0 ± 4.4) compared to studies that did not report one (89.9 ± 7.9) ($P < 0.001$). In contrast, there were no significant differences in transformed Detsky scores based on geographical location, primary outcome type, reported conflicts of interest, funding sources, follow-up duration, or impact factor.

Assessment of the Modified ROB Score

The mean modified ROB score was 6.2, with 27% (33 of 119) being considered high quality. The study characteristics and their relation to the overall transformed Detsky score are shown in Table 2. Statistically significant differences in mROB scores were found across year bins and geographical locations. Studies from the 2014 to 2018 period had the highest average mROB scores (6.8 ± 1.7), significantly different from other periods ($P = 0.027$). Geographically, the Australian trials had the highest average mROB scores (7.3 ± 1.4), while the European trials had the lowest (5.6 ± 1.4), with statistical significance in location ($P < 0.001$). The number of centers showed a trend toward higher scores for multiple-center studies (7.2 ± 1.5) compared to single-center studies (6.3 ± 1.5), but this was not statistically significant ($P = 0.067$). The reporting of a CONSORT diagram was associated with higher mROB scores (7.2 ± 1.5 versus 5.7 ± 1.4, $P < 0.001$). No significant differences were found related to primary outcome type, funding sources, or impact factor. There was a correlation with sample size (Spearman = 0.23, $P = 0.013$), suggesting that smaller, underpowered studies may have greater methodological weakness.

Discussion

Trends in Methodological Quality of RCTs

Over the past two decades, the quality of RCTs comparing computer-assisted navigation to conventional TKA has significantly improved [7,10–12]. The high mean-transformed Detsky score (92% ± 7.4) reflects advancements in trial reporting, with notable progress from 2014 to 2018. This mirrors the general improvement in RCT methodological quality recently seen in orthopaedics and the medical literature [7,10–12,20].

This improvement is driven by implementing standardized reporting guidelines like CONSORT, which enhance trial transparency and completeness by requiring explicit documentation of sequence generation, allocation concealment, and participant flow [7–11,13,14,21]. This reduces selection and attrition bias and improves outcome reliability [7–11,13,14,21]. Regulatory mandates, like trial registration, have further strengthened study quality [22]. Registration rates increased from approximately 16% in 2008 to

Table 1
Transformed Detsky Scores.

Variable	n	Mean ± SD	Median ± IQR	*P*-Value
Year				
1 (2004 to 2008)	30	88.2 ± 8.6	90.0 ± 10.0	0.003
2 (2009 to 2013)	36	92.4 ± 7.1	95.0 ± 5.0	
3 (2014 to 2018)	32	95.1 ± 5.1	95.0 ± 10.0	
4 (2019 to 2024)	21	93.7 ± 6.7	95.0 ± 10.0	
Location				
Asia	50	93.4 ± 7.0	95.0 ± 10.0	0.24
Australia	14	93.3 ± 5.3	92.5 ± 10.0	
Europe	46	90.3 ± 8.5	90.0 ± 9.0	
North American	8	94.4 ± 3.2	95.0 ± 2.5	
Number of centers				
Multiple	10	96.5 ± 3.4	95.0 ± 5.0	0.034
One	93	91.8 ± 7.4	95.0 ± 5.0	
Primary outcome				
Biomechanical	58	92.8 ± 7.6	95.0 ± 10.0	0.19
Intraoperative	8	88.8 ± 6.9	90.0 ± 10.0	
Late postoperative	34	92.8 ± 7.2	95.0 ± 10.0	
Multiple	18	90.9 ± 7.0	90.0 ± 5.0	
Reported COI				
N	99	92.4 ± 7.8	95.0 ± 10.0	0.33
Y	20	91.9 ± 5.0	92.5 ± 5.0	
CONSORT diagram				
N	73	89.9 ± 7.9	90.0 ± 9.0	<0.001
Y	46	96.0 ± 4.4	95.0 ± 5.0	
Funding				
Grant	21	94.1 ± 8.2	95.0 ± 5.0	0.25
Industry	26	92.2 ± 4.9	90.0 ± 5.0	
None	61	92.5 ± 6.5	95.0 ± 10.0	
Unknown	11	88.1 ± 12.9	90.0 ± 24.0	
Follow-up				
<4 weeks	33	90.5 ± 6.5	90.0 ± 9.0	0.26
1 to 12 months	25	93.1 ± 9.3	95.0 ± 10.0	
12 to 24 months	13	92.2 ± 8.7	95.0 ± 10.0	
24 to 36 months	17	93.0 ± 8.1	95.0 ± 10.0	
36 to 60 months	7	93.8 ± 7.5	95.0 ± 14.0	
5 years or more	21	93.2 ± 4.4	95.0 ± 5.0	
Sample size	Spearman correlation = 0.16			0.077
Journal impact factor	Spearman correlation = 0.042			0.65

IQR, interquartile range; COI, conflict of interest; CONSORT, Consolidated Standards of Reporting Trials.

Table 2
Modified Cochrane Risk-of-Bias Scores.

Variable	n	Mean ± SD	Median ± IQR	*P*-Value
Year				
1 (2004 to 2008)	30	5.6 ± 1.3	5.0 ± 1.0	0.027
2 (2009 to 2013)	36	6.2 ± 1.6	6.0 ± 2.5	
3 (2014 to 2018)	32	6.8 ± 1.7	7.0 ± 3.0	
4 (2019 to 2024)	21	6.5 ± 1.6	7.0 ± 3.0	
Location				
Asia	50	6.6 ± 1.6	7.0 ± 3.0	<0.001
Australia	14	7.3 ± 1.4	8.0 ± 1.0	
Europe	46	5.6 ± 1.4	5.0 ± 1.0	
North American	8	6.1 ± 1.4	7.0 ± 2.0	
Number of centers				
Multiple	10	7.2 ± 1.5	7.5 ± 2.0	0.067
One	93	6.3 ± 1.5	6.0 ± 2.0	
Primary outcome				
Biomechanical	58	6.3 ± 1.5	6.0 ± 3.0	0.36
Intraoperative	8	6.3 ± 1.2	6.0 ± 1.0	
Late postoperative	34	6.4 ± 1.8	6.5 ± 3.0	
Multiple	18	5.6 ± 1.7	5.0 ± 3.0	
Reported COI				
No	99	6.2 ± 1.6	6.0 ± 3.0	0.68
Yes	20	6.4 ± 1.7	7.0 ± 3.0	
CONSORT diagram				
No	73	5.7 ± 1.4	5.0 ± 2.0	<0.001
Yes	46	7.2 ± 1.5	8.0 ± 2.0	
Funding				
Grant	21	6.8 ± 1.6	7.0 ± 3.0	0.41
Industry	26	6.2 ± 1.5	6.0 ± 2.0	
None	61	6.2 ± 1.6	6.0 ± 3.0	
Unknown	11	5.9 ± 1.6	6.0 ± 4.0	
Follow-up				
<4 weeks	33	6.0 ± 1.3	6.0 ± 3.0	0.075
1 to 12 months	25	6.6 ± 1.8	7.0 ± 3.0	
12 to 24 months	13	6.6 ± 1.6	6.0 ± 2.0	
24 to 36 months	17	6.5 ± 1.8	7.0 ± 3.0	
36 to 60 months	7	7.1 ± 1.5	8.0 ± 2.0	
5 years or more	21	5.6 ± 1.5	5.0 ± 3.0	
Sample size	Spearman correlation = 0.23			0.013
Journal impact factor	Spearman correlation = 0.13			0.16

IQR, interquartile range; COI, conflict of interest; CONSORT, Consolidated Standards of Reporting Trials.

approximately 90% by 2019 in compliance with International Committee of Medical Journal Editors requirements [22]. Stricter protocols by regulatory bodies and funding agencies and greater collaboration with biostatisticians have also improved trial design [7,10–12,20].

Addressing Internal Validity and Bias Reduction

Our findings highlight a discrepancy between the Detsky index and mROB assessment tools. While Detsky primarily evaluates reporting completeness, the mROB places greater emphasis on internal validity domains such as allocation concealment, blinding, and attrition bias [11,12,15,16]. This explains why, despite 92% of RCTs in our study being classified as high quality based on the transformed Detsky index, only 27% met the threshold for high quality on the mROB scale.

Lower-scoring studies on the Detsky and transformed Detsky indices were more likely to lack clear sample size justifications, post hoc power calculations, well-defined inclusion or exclusion criteria, and a CONSORT diagram. Similarly, studies with lower mROB scores frequently failed to provide clear indications of patient blinding, surgeon experience, and outcome assessment blinding. Additional shortcomings included insufficient or unclear follow-up duration, a lack of transparency in sample size calculations, and the absence of a CONSORT diagram.

Inadequate sample size planning and high attrition rates likely contributed to lower mROB scores. Although 96.7% of RCTs (115 of 119) conducted an *a priori* power analysis, only 54.8% (63 of 115) enrolled an appropriate sample size, and just 87.3% (55 of 63) maintained an adequate number of patients at the final follow-up. Similarly, in a recent review, only 83% of studies met their recruitment targets, and just 57.8% retained a fully powered sample at the final follow-up [11]. Under-powered trials with significant attrition show low statistical reliability and validity of findings [23–27].

Researchers should collaborate with biostatisticians to design adequately powered studies by using realistic effect size assumptions in sample size calculations and accounting for potential follow-up losses through overrecruitment [25–27]. Pooling patients from multiple centers improves statistical power and outcome reliability by increasing sample size. Multicenter trials in our study demonstrated higher methodological quality than single-center trials, with higher Detsky (96.5 versus 91.8, $P = 0.034$) and slightly higher mROB scores (7.2 versus 6.3).

Integrating CONSORT guidelines into journal requirements can improve trial design and transparency. Enforcing stricter peer-review policies will help ensure reporting completeness and methodological rigor. In our study, the inclusion of a CONSORT diagram was associated with higher mROB scores (7.2 versus 5.7, $P < 0.001$), indicating improved methodological quality.

Geographic Variability in Study Quality

Geographic disparities in RCT quality were observed, particularly in mROB scores. The Australian RCTs achieved the highest mean mROB scores (7.3 ± 1.4), while the European RCTs had the lowest (5.6 ± 1.4). These discrepancies may be attributed to differences in regulatory oversight, funding availability, and adherence to reporting standards [5,16,28]. Furthermore, 42% of RCTs were conducted in Asia, whereas only 6.7% originated from North America. This reflects a global shift in surgical trial production over the last two decades, with a dramatic rise in output from Asia [28]. Additionally, the limited number of US-based RCTs in our study may be attributed to the recent Food and Drug Administration approval of robotic TKA systems, beginning with ROBODOC in 2008, followed by Mako (2016), Navio (2017), OMNIBotic (2017), and ROSA (2019) [29,30]. Differences in health care infrastructure and training also contribute to regional variations in methodological quality [28]. Addressing these disparities requires international collaboration. Research partnerships between established and emerging centers could help standardize reporting guidelines in underrepresented regions and promote global research standards [28].

Industry Sponsorship and Research Transparency

Industry funding is essential for orthopaedic research, supporting large well-powered studies. However, it also introduces potential biases that may affect study design, reporting, and interpretation [11,31–36]. In this study, only 16% of RCTs disclosed conflicts of interest, raising concerns about underreporting. While no significant correlation was found between industry funding and study quality, the literature suggests industry-sponsored trials are more likely to report favorable outcomes, contributing to selective outcome reporting bias [12,16,31–36]. Ezzet [33] found 93% of commercially funded hip research studies reported positive outcomes, compared to only 37% of independently funded studies. Furthermore, studies with negative or null findings are less likely to be published, exacerbating publication bias [31–36].

To improve transparency, industry-sponsored trials should implement independent data monitoring committees, enforce prespecified protocols, and mandate comprehensive reporting of all outcomes [31–36]. Journals and funding agencies should strengthen conflict-of-interest disclosure policies and require trial preregistration to improve accountability [31–36]. Using independent statistical analysis teams for industry-funded studies can also reduce bias [31–36].

Potential Limitations

This study has notable potential limitations. The Detsky and mROB assessment tools depend on reported trial information; incomplete reporting may lead to inaccurately low scores despite robust methodology. Additionally, surgical trials often have an inherent bias, as neither the surgeons nor patients can always be blinded [7,10,11]. This study focused solely on RCTs, excluding insights from high-quality cohort studies. Similar to Smith et al. and Imam et al., our review found that many trials had sample sizes below 100 patients, raising concerns about their statistical power [7,11]. Small trials are more vulnerable to type I and II errors, making their results less generalizable. Bhandari et al. [10] emphasized that underpowered trials often yield inconclusive results. This study focused exclusively on computer-navigated versus conventional TKA, which may limit the broader applicability of the findings. This narrow scope restricts external validity, making it difficult to generalize the results to other surgical procedures or methodologies. A post hoc power analysis was not conducted, as we believe statistical power is best discussed in terms of effect sizes detectable with 80% power and 119 observations. Consequently, the study had sufficient power to detect medium-to-large effect sizes but was underpowered for small effect sizes [37].

Conclusions

Enhancing the methodological rigor of future RCTs in orthopaedic surgery requires a multifaceted approach. Stronger randomization, centralized allocation concealment, and rigorous blinding should be prioritized to enhance trial validity. Addressing statistical power limitations through multicenter collaborations, adaptive trial methodologies, and refined sample size calculations will improve the generalizability of study findings. Additionally, implementing standardized reporting guidelines, increasing transparency in industry-sponsored research, and fostering international partnerships will further elevate trial quality. By integrating these measures, future orthopaedic RCTs can generate high-quality evidence to guide surgical decision-making and improve patient outcomes.

CRediT authorship contribution statement

Antron Spooner: Writing – review & editing, Writing – original draft, Methodology, Investigation, Conceptualization. **Austin G. Helton:** Writing – review & editing, Writing – original draft, Visualization, Investigation, Data curation. **Vincent Lee:** Writing – review & editing, Writing – original draft, Validation. **David Redden:** Formal analysis, Data curation. **Brent A. Ponce:** Writing – review & editing, Supervision, Methodology, Conceptualization. **Justin P. Fernicola:** Writing – review & editing.

References

[1] Jeffery RS, Morris RW, Denham RA. Correlation between alignment and aseptic loosening in total knee arthroplasty. Clin Orthop Relat Res 1991;273:178–83.
[2] Barrack RL, Schrader T, Bertot AJ, Wolfe MW, Myers L. Component rotation and anterior knee pain after total knee arthroplasty. Clin Orthop Relat Res 2001;392:46–55. https://doi.org/10.1097/00003086-200111000-00006.
[3] Zhao L, Xu F, Lao S, Zhao J, Wei Q. Comparison of the clinical effects of computer-assisted and traditional techniques in bilateral total knee arthroplasty: a meta-analysis of randomized controlled trials. PLoS One 2020;15:e0239341. https://doi.org/10.1371/journal.pone.0239341.
[4] Choong PF, Dowsey MM, Stoney JD. Does accurate anatomical alignment result in better function and quality of life? Comparing conventional and computer-assisted total knee arthroplasty. J Arthroplasty 2009;24:560–9. https://doi.org/10.1016/j.arth.2008.02.018.
[5] Reito A, Raittio L, Helminen O. Revisiting the sample size and statistical power of randomized controlled trials in orthopaedics after 2 decades. JBJS Rev 2020;8:e0079. https://doi.org/10.2106/JBJS.RVW.19.00079.
[6] Marcano-Fernández F, Camara-Cabrera J, Madden K, Johal H, Nadeem IM, Kapoor R, et al. A systematic review of outcome measures in orthopaedic trauma trials: what are we measuring? Indian J Orthop 2022;56:1316–26. https://doi.org/10.1007/s43465-022-00667-8.
[7] Smith CS, Mollon B, Vannabouathong C, Fu JM, Sales B, Bhandari M, et al. An assessment of randomized controlled trial quality in the journal of bone & joint surgery: update from 2001 to 2013. J Bone Joint Surg Am 2020;102:e116. https://doi.org/10.2106/JBJS.18.00653.
[8] Moher D, Hopewell S, Schulz KF, Montori V, Gøtzsche PC, Devereaux PJ, et al. CONSORT 2010 explanation and elaboration: updated guidelines for reporting parallel group randomised trials. BMJ 2010;340:c869. https://doi.org/10.1136/bmj.c869.
[9] Checketts JX, Sims MT, Detweiler B, Middlemist K, Jones J, Vassar M. An evaluation of reporting guidelines and clinical trial registry requirements among orthopaedic surgery journals. J Bone Joint Surg Am 2018;100:e15. https://doi.org/10.2106/JBJS.17.00529.
[10] Bhandari M, Richards RR, Sprague S, Schemitsch EH. The quality of reporting of randomized trials in the Journal of Bone and Joint Surgery from 1988 through 2000. J Bone Joint Surg Am 2002;84:388–96. https://doi.org/10.2106/00004623-200203000-00009.
[11] Imam N, Sudah SY, Shaikh SZ, Bonney AA, Nicholson AD, Namdari S, et al. The rising quality of randomized controlled trials in *the Journal of bone & joint*

surgery: an updated analysis from 2014 to 2022. JB JS Open Access 2024;9: e23.00079. https://doi.org/10.2106/JBJS.OA.23.00079.
[12] Dodwell E, Dua S, Dulai SK, Astone K, Mulpuri K. The quality of randomized controlled trials in pediatric orthopaedics: are we improving? J Pediatr Orthop 2015;35:536–45. https://doi.org/10.1097/BPO.0000000000000324.
[13] Adie S, Harris IA, Naylor JM, Mittal R. CONSORT compliance in surgical randomized trials: are we there yet? A systematic review. Ann Surg 2013;258: 872–8. https://doi.org/10.1097/SLA.0b013e31829664b9.
[14] Dunford B, Bacani R, Anderson JM, Checketts J, Scott J, Torgerson T, et al. Reporting quality and adherence to the CONSORT checklist of randomized controlled trials underpinning the American academy of orthopaedic surgeons upper extremity clinical practice guidelines. J Am Osteopathic Acad Orthop 2022;6.
[15] Detsky AS, Naylor CD, O'Rourke K, McGeer AJ, L'Abbé KA. Incorporating variations in the quality of individual randomized trials into meta-analysis. J Clin Epidemiol 1992;45:255–65. https://doi.org/10.1016/0895-4356(92)90085-2.
[16] Higgins JP, Altman DG, Gøtzsche PC, Jüni P, Moher D, Oxman AD, et al. The Cochrane Collaboration's tool for assessing risk of bias in randomised trials. BMJ 2011;343:d5928. https://doi.org/10.1136/bmj.d5928.
[17] Dragosloveanu S, Petre MA, Capitanu BS, Dragosloveanu CDM, Cergan R, Scheau C. Initial learning curve for robot-assisted total knee arthroplasty in a dedicated orthopedics center. J Clin Med 2023;12:6950. https://doi.org/10.3390/jcm12216950.
[18] Bolam SM, Tay ML, Zaidi F, Sidaginamale RP, Hanlon M, Munro JT, et al. Introduction of ROSA robotic-arm system for total knee arthroplasty is associated with a minimal learning curve for operative time. J Exp Orthop 2022;9:86. https://doi.org/10.1186/s40634-022-00524-5.
[19] Bell C, Grau L, Orozco F, Ponzio D, Post Z, Czymek M, et al. The successful implementation of the Navio robotic technology required 29 cases. J Robot Surg 2022;16:495–9. https://doi.org/10.1007/s11701-021-01254-z.
[20] Kilicoglu H, Jiang L, Hoang L, Mayo-Wilson E, Vinkers CH, Otte WM. Methodology reporting improved over time in 176,469 randomized controlled trials. J Clin Epidemiol 2023;162:19–28. https://doi.org/10.1016/j.jclinepi.2023.08.004.
[21] Schulz KF, Altman DG, Moher D, CONSORT Group. CONSORT 2010 statement: updated guidelines for reporting parallel group randomized trials. Obstet Gynecol 2010;115:1063–70. https://doi.org/10.1097/AOG.0b013e3181d9d421.
[22] Robinson NB, Fremes S, Hameed I, Rahouma M, Weidenmann V, Demetres M, et al. Characteristics of randomized clinical trials in surgery from 2008 to 2020: a systematic review. JAMA Netw Open 2021;4:e2114494. https://doi.org/10.1001/jamanetworkopen.2021.14494.
[23] Sterne JAC, Savović J, Page MJ, Elbers RG, Blencowe NS, Boutron I, et al. RoB 2: a revised tool for assessing risk of bias in randomised trials. BMJ 2019;366: l4898. https://doi.org/10.1136/bmj.l4898.
[24] Page MJ, Higgins JP, Clayton G, Sterne JA, Hróbjartsson A, Savović J. Empirical evidence of study design biases in randomized trials: systematic review of meta-epidemiological studies. PLoS One 2016;11:e0159267. https://doi.org/10.1371/journal.pone.0159267.
[25] Schulz KF, Grimes DA. Allocation concealment in randomised trials: defending against deciphering. Lancet 2002;359:614–8. https://doi.org/10.1016/S0140-6736(02)07750-4.
[26] Biau DJ, Jolles BM, Porcher R. P value and the theory of hypothesis testing: an explanation for new researchers. Clin Orthop Relat Res 2010;468:885–92. https://doi.org/10.1007/s11999-009-1164-4.
[27] Pildal J, Hróbjartsson A, Jørgensen KJ, Hilden J, Altman DG, Gøtzsche PC. Impact of allocation concealment on conclusions drawn from meta-analyses of randomized trials [published correction appears in Int J Epidemiol. 2008 Apr;37 (2):422]. Int J Epidemiol 2007;36:847–57. https://doi.org/10.1093/ije/dym087.
[28] Pronk AJM, Roelofs A, Flum DR, Bonjer HJ, Abu Hilal M, Dijkgraaf MGW, et al. Two decades of surgical randomized controlled trials: worldwide trends in volume and methodological quality. Br J Surg 2023;110:1300–8. https://doi.org/10.1093/bjs/znad160.
[29] Urish KL, Conditt M, Roche M, Rubash HE. Robotic total knee arthroplasty: surgical assistant for a customized normal kinematic knee. Orthopedics 2016;39:e822–7. https://doi.org/10.3928/01477447-20160623-13.
[30] Siddiqi A, Horan T, Molloy RM, Bloomfield MR, Patel PD, Piuzzi NS. A clinical review of robotic navigation in total knee arthroplasty: historical systems to modern design. EFORT Open Rev 2021;6:252–69. https://doi.org/10.1302/2058-5241.6.200071.
[31] Khan SN, Mermer MJ, Myers E, Sandhu HS. The roles of funding source, clinical trial outcome, and quality of reporting in orthopedic surgery literature. Am J Orthop (Belle Mead NJ) 2008;37:E205–12.
[32] Cho MK, Shohara R, Schissel A, Rennie D. Policies on faculty conflicts of interest at US universities. JAMA 2000;284:2203–8. https://doi.org/10.1001/jama.284.17.2203.
[33] Ezzet KA. The prevalence of corporate funding in adult lower extremity research and its correlation with reported results. J Arthroplasty 2003;18(7 Suppl 1):138–45. https://doi.org/10.1016/s0883-5403(03)00289-4.
[34] Lundh A, Lexchin J, Mintzes B, Schroll JB, Bero L. Industry sponsorship and research outcome. Cochrane Database Syst Rev 2017;2:MR000033. https://doi.org/10.1002/14651858.MR000033.pub3.
[35] Murad MH, Chu H, Lin L, Wang Z. The effect of publication bias magnitude and direction on the certainty in evidence. BMJ Evid Based Med 2018;23: 84–6. https://doi.org/10.1136/bmjebm-2018-110891.
[36] Ioannidis JPA. The importance of potential study biases in randomized controlled trials. JAMA 2018;319:627–8.
[37] Turner RM, Bird SM, Higgins JP. The impact of study size on meta-analyses: examination of underpowered studies in Cochrane reviews. PLoS One 2013;8:e59202. https://doi.org/10.1371/journal.pone.0059202.

Appendix

Supplementary Table 1
Agreement for Detsky Score.

Criteria	N	Number With Agreement	Percent Agree	Exact 95% Confidence Interval
Random assignment	119	119	100	97.0, 100.0
Randomization described	119	116	97.5	92.8, 99.5
Concealment explained	119	109	91.6	85.1, 95.0
Outcomes clearly explained	119	119	100	97.0, 100.0
Objectives outcomes	119	119	100	97.0, 100.0
Blinding documented	119	109	91.6	85.1, 95.9
Clear inclusion or exclusion criteria	119	113	95.0	89.4, 98.1
Reasons for exclusion	119	105	88.2	81.1, 93.4
Treatment explained	119	119	100	97.0, 100.0
Control explained	119	119	100	97.0, 100.0
Statistical analysis included	119	119	100	97.0, 100.0
Confidence intervals included	119	115	96.6	91.6, 99.1
Sample size justified	119	118	99.2	95.4, 100.0
CONSORT diagram	119	119	100	97.0, 100.0

CONSORT, Consolidated Standards of Reporting Trials.

Supplementary Table 2
Agreement for Modified Cochrane Risk-of-Bias Scores.

Criteria	N	Number With Agreement	Percent Agree	Exact 95% Confidence Interval
Random assignment	119	119	100	97.0, 100.0
Concealment	119	117	98.3	94.1, 100.0
Surgeon experienced	119	118	99.2	95.4, 100.0
Assessor blinded	119	117	98.3	94.1, 100.0
Patient blinded	119	108	90.8	84.1, 95.3
Adequate follow-up	119	110	92.4	88.1, 96.5
Free of selective reporting	119	119	100	97.0, 100.0
Outcomes were objective	119	119	100	97.0, 100.0
Adequate sample size	119	117	98.3	94.1, 100.0
Surgeon experienced with method	119	118	99.2	95.4, 100.0

Supplementary Table 3
Characteristics of the 119 Trials.

Variable	Frequency	%
Year		
2004 to 2008	30	25.2
2009 to 2013	36	30.3
2014 to 2018	32	26.9
2019 to 2024	21	17.7
Location		
Undetermined	1	0.8
Asia	50	42.0
Australia	14	11.8
Europe	46	38.7
North America	8	6.7
Number of centers		
Not stated	16	13.5
Multiple	10	8.4
One	93	78.2
Primary outcome		
Biomechanical	58	48.7
Early postoperative outcomes	1	0.8
Intraoperative outcomes	8	6.7
Late postoperative outcomes	34	28.6
Multiple	18	15.1
Reported COI		
N	99	83.2
Y	20	16.8
CONSORT diagram		
N	73	61.3
Y	46	38.7
Funding		
Grant	21	17.7
Industry	26	21.9
None	61	51.3
Unknown	11	9.2
Follow-up time		
Not specified	3	2.5
<4 weeks	33	27.7
1 to 12 months	25	21.0
12 to 24 months	13	10.9
24 to 36 months	17	14.3
36 to 60 months	7	5.9
5 years or more	21	17.7

COI, conflict of interest; CONSORT, Consolidated Standards of Reporting Trials.