JB&JS OPEN ACCESS

# The Rising Quality of Randomized Controlled Trials in *The Journal of Bone & Joint Surgery*

## An Updated Analysis from 2014 to 2022

Nareena Imam, BA, Suleiman Y. Sudah, MD, Siraj Z. Shaikh, BS, Ashley A. Bonney, BS, Allen D. Nicholson, MD, Surena Namdari, MD, and Mariano E. Menendez, MD

*Investigation performed at Monmouth Medical Center, Long Branch, New Jersey*

**Background:** Previous reports found that 40% of randomized controlled trials (RCTs) published in *The Journal of Bone & Joint Surgery (JBJS)* from 1988 to 2000 and 47% of those published from 2001 to 2013 were of high quality. The purpose of this study was to assess the quality of RCTs published from 2014 to 2022 in *JBJS* and to compare these findings with those of prior analyses in order to identify trends over time and areas for continued improvement.

**Methods:** PubMed was searched for the term "randomized controlled trial" to identify studies published in *JBJS* from 2014 to 2022. Each included RCT was evaluated with use of the Detsky score and a risk-of-bias assessment modified from the Cochrane tool. These evaluations were then compared with previous evaluations of RCTs from the 1988 to 2000 and 2001 to 2013 periods with use of independent-sample t tests. A transformed Detsky score of >75% and a modified risk-of-bias score of ≥8 were defined as being indicative of high quality.

**Results:** A total of 218 RCTs were published in *JBJS* from 2014 to 2022. An a priori sample size was calculated in 183 studies (83.9%). A total of 152 (83.1%) of the 183 studies enrolled the calculated number of patients, of which 126 (82.9%) maintained an adequate number at the time of final follow-up. Most RCTs were conducted at a single center (146 of 218; 67%), evaluated a surgical intervention (162 of 218; 74%), and reported positive results (142 of 218; 65%). The mean transformed Detsky score was 85% ± 10% (95% confidence interval, 83.7% to 86.3%), with 82% of trials (179 of 218) scored as high quality. The mean transformed Detsky score from 2014 to 2022 was higher than that from 1988 to 2000 and that from 2001 to 2013 (85% versus 76% and 68%, respectively; $p < 0.001$). The mean modified risk-of-bias score was 7 ± 1, with 42% of trials (92 of 218) scored as high quality. RCTs published from 2014 to 2022 had a higher mean modified risk-of-bias score than those published from 2001 to 2013 (7 ± 1 versus 6 ± 1; $p < 0.001$). Compared with the 2001 to 2013 and 2014 to 2022 periods, the 1988 to 2000 period had a greater proportion of trials that reported positive results (51% and 65% versus 82%, respectively; $p < 0.001$) and that included data from multiple centers (31% and 33% versus 67%; $p < 0.001$).

**Conclusions:** The quality of RCTs published in *JBJS* from 2014 to 2022 has improved from that reported previously, as demonstrated by the increases in the modified risk-of-bias score and transformed Detsky score from prior periods. This may be the result of journal policies such as the requirements of CONSORT adherence and prospective trial registration. Investigators should focus on improving the clarity of reporting, limiting attrition bias, and making efforts to blind support staff in order to increase the quality of future RCTs.

**Clinical Relevance:** Improving the quality of RCTs is crucial given their potential to influence current clinical practice.

Evidence-based medicine requires high-quality clinical research. Because randomized controlled trials (RCTs) represent the study design yielding the highest level of evidence[1], they are commonly utilized as the basis for recommendations in the American Academy of Orthopaedic Surgeons (AAOS) clinical practice guidelines and to guide clinical decision-making. Accordingly, the number of RCTs published in orthopaedic journals has been on the rise[2-4]. However, inherent biases and methodological errors can limit the accuracy of their findings[2,5], and unclear reporting

**Disclosure:** The **Disclosure of Potential Conflicts of Interest** forms are provided with the online version of the article (http://links.lww.com/JBJSOA/A599).

of even high-quality RCTs may undermine their perceived quality.

Instruments such as the Cochrane risk-of-bias tool have been developed for assessors to evaluate the quality of published RCTs[6]. Previous studies that utilized these tools have demonstrated variable quality among RCTs in multiple surgical specialties[7-9], including orthopaedics[2,10-12]. In an effort to improve the quality of the reporting of RCTs, the Consolidated Standards of Reporting Trials (CONSORT) statement has become a requirement for publication in many orthopaedic journals, including *The Journal of Bone & Joint Surgery (JBJS)*[13]. The implementation of these policies has been associated with improvements in the reporting quality of RCTs[14,15].

*JBJS* has one of the highest impact factors among orthopaedic journals and has considerable influence on clinical practice[16]. Despite this, Bhandari et al. found that only 40% of RCTs published in *JBJS* from 1988 to 2000 were of high quality[10]. In fact, >50% of the trials were limited by a lack of concealed randomization, a lack of blinding of the outcome assessors, or failure to report the reasons for excluding patients. In an updated analysis, Smith et al. reported an encouraging increase in the quality of trials published from 2001 to 2013; however, an increase in the number of smaller, single-center studies was also observed[2]. The purpose of the present study was to assess the quality of RCTs published in *JBJS* from 2014 to 2022 and to compare these findings with those of prior reports in order to identify trends over time and areas for continued improvement.

## Materials and Methods

### *Systematic Search*

The electronic PubMed database was searched on November 20, 2022, with use of the search term "randomized controlled trial" to identify RCTs published in *JBJS* from 2014 to 2022. Searches were performed with no restriction on language or publication format by 2 independent reviewers. The results were uploaded to a Microsoft Excel file (version 16.43; Microsoft).

### *Inclusion and Exclusion Criteria*

A study was considered eligible if the study design was a prospective RCT and the study had been published in *JBJS* between 2014 and 2022. Studies that were not available as full text; that had a case-control, cohort, case series, or systematic review study design; or that were nonrandomized prospective trials were excluded. The abstracts of potentially eligible studies were independently screened by 2 authors, who then performed a full-text review of the remaining studies to determine final inclusion. Consensus was reached between the reviewers through discussion.

### *Data Extraction*

Data were extracted from each of the included studies by 1 investigator, confirmed by a second investigator, and compiled in a Microsoft Excel document. Data included the background of the first author (i.e., surgeon or nonsurgeon), the year of publication, sample size, the number of centers involved (multiple centers or a single center), the category of intervention as defined previously by Bhandari et al.[10] (fracture treatment, treatment of degenerative disease of the spine and joints, diagnostic test, thrombosis prevention, pain management, or other), body region (upper extremity, long bones of the lower extremity, spine, hip and knee, or foot and ankle), whether or not the study received funding, and the direction of the results (positive if significant or negative if not significant).

### *Evaluation Using the Detsky Score and the Modified Cochrane Risk-of-Bias Tool*

Two authors independently reviewed the full text of each included RCT to evaluate the study with use of the Detsky index quality score and a risk-of-bias assessment modified from the Cochrane tool, as was done previously by Smith et al.[2].

The Detsky score evaluates the quality of trial reporting and consists of 14 questions covering 5 categories[17]. To standardize the final Detsky score, a transformed score based on a maximum score of 100 was calculated, as was done previously by Bhandari et al.[10]. A transformed Detsky score of >75% was indicative of high quality. The modified Cochrane risk-of-bias assessment tool comprises 10 categories: randomization, allocation concealment, surgeon or treatment provider blinding, assessor blinding, patient blinding, patient follow-up, selective outcome reporting, objectivity of outcomes, adequate sample size, and surgeon experience with the treatment[2,6]. Each category was assigned 1 point, for a maximum score of 10 points, which represents a low risk of bias. A score of ≥8 on this scale was indicative of high quality.

Although the 2 raters did not receive specific training for the use of the Detsky score or the modified risk-of-bias tool, they carefully reviewed the instructions for these guidelines.

### *Statistical Analysis*

Interrater agreement for the modified risk-of-bias assessment and the Detsky score was calculated using an intraclass correlation coefficient (ICC) with a 95% confidence interval (CI). An ICC of 0 to 0.2 was defined as poor; 0.21 to 0.40, fair; 0.41 to 0.60, moderate; 0.61 to 0.80, substantial; and ≥0.81, nearly perfect agreement[10].

Categorical variables are presented as proportions and continuous variables, as means and standard deviations (SDs), except as noted. Normality was assessed with use of a Shapiro-Wilk test. Univariate analysis was performed with use of a chi-square test or Fisher test for categorical variables and with use of an independent-sample Student t test or a Mann-Whitney U test for continuous variables. Regression analyses were performed to evaluate whether there was a significant change in the transformed Detsky scores or modified risk-of-bias scores over time among studies published from 2014 to 2022. The change in the number of publications per year was assessed with use of a linear regression analysis.

To determine if the quality of RCTs published in *JBJS* had changed from that previously reported for the 2001 to

**TABLE I Characteristics of the Included Trials**

| | No. (%) of Trials |
|---|---|
| Publication year | |
| 2014 | 38 (17.4%) |
| 2015 | 26 (11.9%) |
| 2016 | 14 (6.4%) |
| 2017 | 25 (11.4%) |
| 2018 | 23 (10.6%) |
| 2019 | 31 (14.2%) |
| 2020 | 28 (12.8%) |
| 2021 | 19 (8.7%) |
| 2022 | 14 (6.4%) |
| No. of centers | |
| One | 146 (67.0%) |
| Multiple | 72 (33.0%) |
| Reported positive results | |
| Yes | 142 (65.1%) |
| No | 76 (34.9%) |
| Evaluated a surgical intervention | |
| Yes | 162 (74.3%) |
| No | 56 (25.7%) |
| Reported financial support | |
| Yes | 103 (47.2%) |
| No | 115 (52.8%) |
| Surgeon as first author | |
| Yes | 195 (89.4%) |
| No | 23 (10.6%) |

2013[2] and 1988 to 2000[10] time periods, mean transformed Detsky scores and modified risk-of-bias scores were compared with use of independent-sample t tests. Because the time period evaluated in this study (9 years) was shorter than that in the previous analyses (13 years), only time-normalized variables were compared. Significance was set at $p < 0.05$.

*Source of Funding*

No external funding was received for this investigation.

## Results

### *Study Identification*

A total of 218 RCTs were published in *JBJS* from 2014 to 2022 and included in the analysis. *JBJS* published the highest number of RCTs in 2014 (38 RCTs; 17% of the total). The median sample size of the included studies was 104 patients (interquartile range [IQR], 61 to 182). The mean sample size (and SD) of the included studies was 157 ± 184 patients (range, 14 to 1,244 patients).

An a priori sample size was calculated with use of a power analysis in 183 studies (84%); 152 (83%) of these studies enrolled the calculated number of patients, of which 83% (126 of 152) maintained an adequate number at the time of final follow-up. A total of 146 (67%) of the 218 RCTs were conducted at a single center. Most RCTs evaluated a surgical intervention (162 of 218; 74%) and reported positive results (142 of 218; 65%). Over half of the RCTs (115 of 218; 53%) did not report financial support. The first author was a surgeon in 89% of the trials (195 of 218). The characteristics of the included studies are presented in Table I.

### *Detsky Index Quality Score*

The frequencies of nontransformed and transformed Detsky scores are shown in Figures 1-A and 1-B. The mean transformed Detsky score was 85% ± 10% (95% CI, 83.7% to 86.3%; Fig. 2).

One-half percent of trials (1 of 218) had a transformed Detsky score of ≤50%, 17% (38 of 218) had a score of ≤75% but

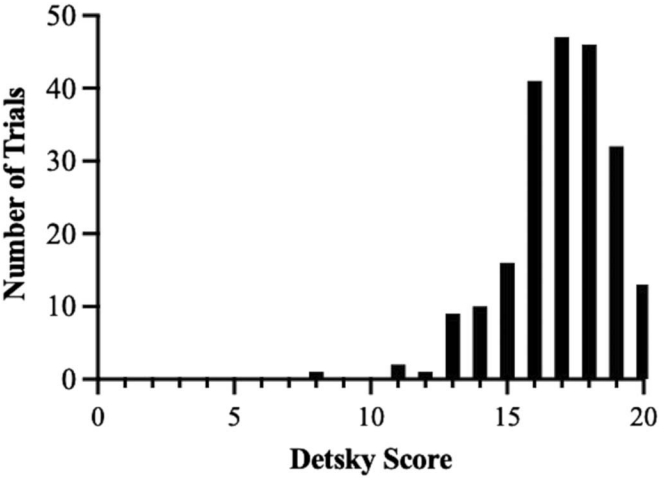


Fig. 1-A

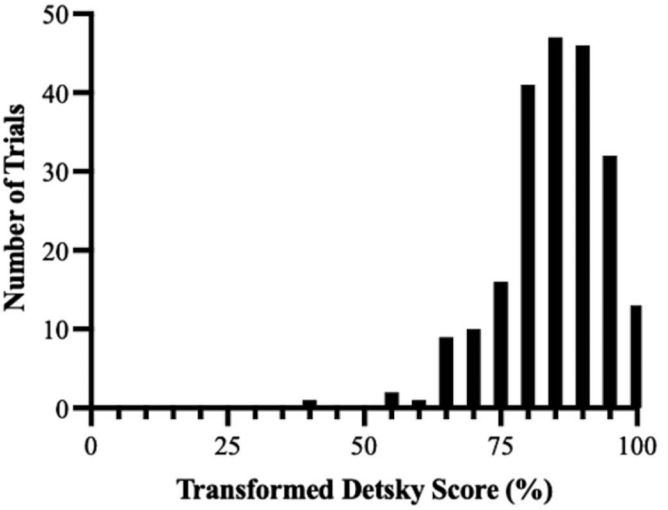


Fig. 1-B

**Figs. 1-A and 1-B** The frequency of Detsky scores before (**Fig. 1-A**) and after (**Fig. 1-B**) transformation among RCTs published in *JBJS* from 2014 to 2022.

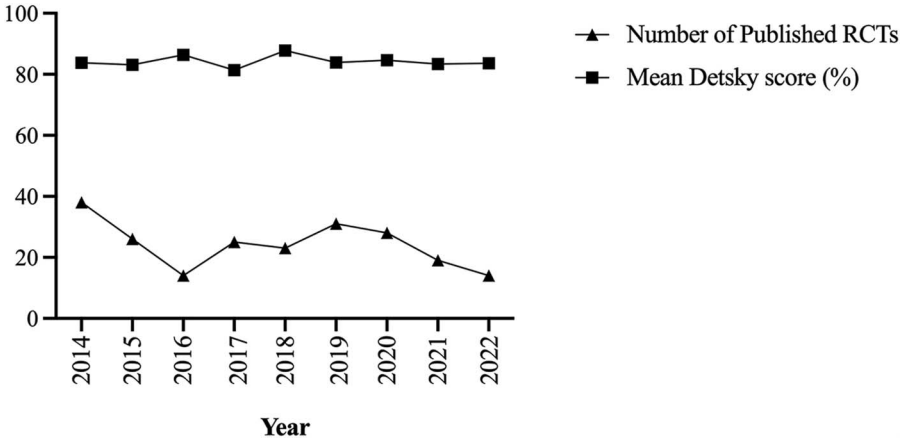


Fig. 2

The number of RCTs published in *JBJS* and the mean transformed Detsky score by year.

>50%, and 82% (179 of 218) had a score of >75%. The ICC for the Detsky score was 0.66 (95% CI, 0.46 to 0.78). A regression analysis of the transformed Detsky scores demonstrated that scores did not increase over time from 2014 to 2022.

There was no significant difference in transformed Detsky scores between trials that did and did not have a surgeon as the first author (mean score, 85% ± 10% versus 85% ± 10%) or between trials that evaluated a surgical intervention and those that did not (85% ± 9% versus 85% ± 11%). Similarly, there was no significant difference in transformed Detsky scores between trials that cited financial support and those that did not (85% ± 9% versus 84% ± 10%). The mean transformed Detsky score of trials reporting positive results did not significantly differ from that of trials reporting negative results (84% ± 10% versus 86% ± 10%). No significant difference in the mean transformed Detsky score was found between multicenter and single-center trials (84% ± 9% versus 85% ± 10%).

Transformed Detsky scores did not significantly differ by intervention category or body region. Mean transformed Detsky scores by intervention category and body region are presented in Tables II and III, respectively.

The number of trials fully meeting the criteria of each component of the Detsky score are shown in Table IV. Randomization was concealed in 56% of trials (123 of 218) and outcome assessors were blinded in 55% (120 of 218) of trials.

A comparison of the characteristics of RCTs published from 2014 to 2022 with the characteristics of RCTs published in the 1988 to 2000 and 2001 to 2013 time periods is presented in Table V. The mean transformed Detsky score from 2014 to 2022 was significantly higher than that from 1988 to 2000 and that from 2001 to 2013 (85% versus 76% and 68%, respectively; $p < 0.001$). The proportion of RCTs with a Detsky score of >75% from 2014 to 2022 was significantly higher than that from 1988 to 2000 and that from 2001 to 2013 (82% versus 40% and 47%; $p < 0.001$).

Compared with the 2001 to 2013 and 2014 to 2022 periods, the 1988 to 2000 period had a greater proportion of trials that reported positive results (51% and 65% versus 82%, respectively; $p < 0.001$) and a greater proportion of trials that included data from multiple centers (31% and 33% versus

**TABLE II Mean Transformed Detsky Score by the Category of Intervention***

| Intervention | No. of Trials | Mean Transformed Detsky Score *(%)* |
|---|---|---|
| Treatment of degenerative disease | 53 (24.3%) | 84.1 ± 10.2 |
| Fracture treatment | 47 (21.6%) | 83.6 ± 9.0 |
| Pain management | 29 (13.3%) | 85.9 ± 8.7 |
| Thrombosis prevention | 5 (2.3%) | 87.0 ± 7.6 |
| Other | 84 (38.5%) | 85.0 ± 10.1 |

*The values are given as the count, with the percentage in parentheses, or as the mean ± standard deviation. No studies met the inclusion criteria for diagnostic test and therefore the category was excluded.

**TABLE III Mean Transformed Detsky Score by Body Region***

| Body Region | No. of Trials | Mean Transformed Detsky Score *(%)* |
|---|---|---|
| Hip and knee | 105 (48.2%) | 85.8 ± 9.5 |
| Upper extremity | 58 (26.6%) | 84.2 ± 7.6 |
| Spine | 19 (8.7%) | 82.6 ± 10.6 |
| Foot and ankle | 17 (7.8%) | 86.8 ± 8.1 |
| Lower extremity | 11 (5.0%) | 81.8 ± 13.8 |
| Uncategorized | 8 (3.7%) | 76.9 ± 15.8 |

*The values are given as the count, with the percentage in parentheses, or as the mean ± standard deviation.

**TABLE IV Number of Trials Fully Meeting Detsky Score Components**

| Component of Scale | No. of Trials Meeting All Criteria (4 of 4) | No. of Trials Meeting <4 Criteria |
|---|---|---|
| Randomization | 123 | 95 |
| Outcome | 120 | 98 |
| Eligibility | 189 | 29 |
| Intervention | 141 | 77 |
| Statistics | 129 | 89 |

67%; $p < 0.001$). The proportion of studies that reported receiving funding did not significantly differ among the 1988 to 2000, 2001 to 2013, and 2014 to 2022 periods (57% versus 53% versus 47%, respectively).

### *Modified Risk-of-Bias Score*

The frequency of modified risk-of-bias scores is presented in Figure 3. The mean modified risk-of-bias score was 7 ± 1 (Fig. 4). The overall ICC for the modified risk-of-bias score was 0.69 (95% CI, 0.48 to 0.77). A total of 92 (42%) of the 218 studies had a modified risk-of-bias score of ≥8. There was no significant difference in the mean modified risk-of-bias score over time from 2014 to 2022.

The mean modified risk-of-bias score did not significantly differ between surgical and nonsurgical trials (7 ± 1 versus 7 ± 2), between trials that reported receiving funding and those that did not (7 ± 1 versus 7 ± 1), or between trials that reported positive results and those that reported negative results (7 ± 1 versus 7 ± 1). Multicenter trials did not have a higher mean modified risk-of-bias score than single-center trials (7 ± 1 versus 7 ± 1).

There was no significant difference in the mean modified risk-of-bias score by the category of intervention. Trials that did not have a body region specified had significantly lower modified risk-of-bias scores than those that did ($p = 0.016$). Mean modified risk-of-bias scores by intervention category and body region are provided in Tables VI and VII, respectively.

The mean modified risk-of-bias score was not available for the 1988 to 2000 period. Compared with 2001 to 2013, the 2014 to 2022 period had a significantly higher mean modified risk-of-bias score (6 ± 1 versus 7 ± 1; $p < 0.001$).

## Discussion

This analysis of RCTs published in *JBJS* from 2014 to 2022 demonstrated a continued increase in quality when compared with the characteristics previously reported for the 1988 to 2000 and 2001 to 2013 periods. We found a small resurgence in the proportion of multicenter trials and a large increase in the reporting of positive results from the 2001 to 2013 period to the 2014 to 2022 period. Although the average sample size of present-day RCTs had decreased from that in the 2001 to 2013 period (albeit not significantly), there was an increased likelihood of trials having conducted an a priori power analysis and having maintained an adequate sample size at the time of final follow-up.

The Detsky index quality score consists of 14 questions covering 5 categories and evaluates the quality of trial reporting[17]. Bhandari et al. standardized the final Detsky score by calculating a transformed score that was based on a maximum of 100; a transformed score of 75% was established

**TABLE V Comparison of RCTs from 2014 to 2022 with RCTs from 1998 to 2000 and from 2001 to 2013***

| | 1988-2000 | 2001-2013 | 2014-2022 | P Value |
|---|---|---|---|---|
| No. of published trials | 72 | 285 | 218 | |
| Publications per year | 5.5 | 21.9 | 24.2 | 0.261 |
| Mean transformed Detsky score† | 68.1% ± 1.6% | 76.2% ± 0.7% | 84.6% ± 9.6% | **<0.001** |
| Mean modified risk-of-bias score† | NR | 6.40 ± 0.71 | 7.3 ± 1.3 | **<0.001** |
| Trials reporting positive results‡ | 81.9% | 50.5% | 65.1% | **<0.001** |
| Multicenter trials‡ | 67% | 31% | 33.0% | **<0.001** |
| Reported receiving funding‡ | 56.9% | 52.6% | 47.2% | 0.263 |
| Sample size§ | NR | 165.6 (17.9) | 156.8 (12.5) | 0.705 |
| Trials reporting an a priori power analysis‡ | NR | 67.7% | 83.9% | **<0.001** |
| Study type *(no. [%] of trials)* | | | | **<0.001** |
| Surgical | 30 (41.7%) | 183 (64.2%) | 162 (74.3%) | |
| Nonsurgical | 16 (22.2%) | 45 (15.8%) | 38 (17.4%) | |
| Drug | 26 (36.1%) | 57 (20.0%) | 18 (8.3%) | |

*NR = not reported. Bold indicates significance. †Values are given as the mean ± standard deviation. ‡Values are given as the proportion of the total number of published studies for that period. §Values are given as the mean, with the standard error in parentheses.

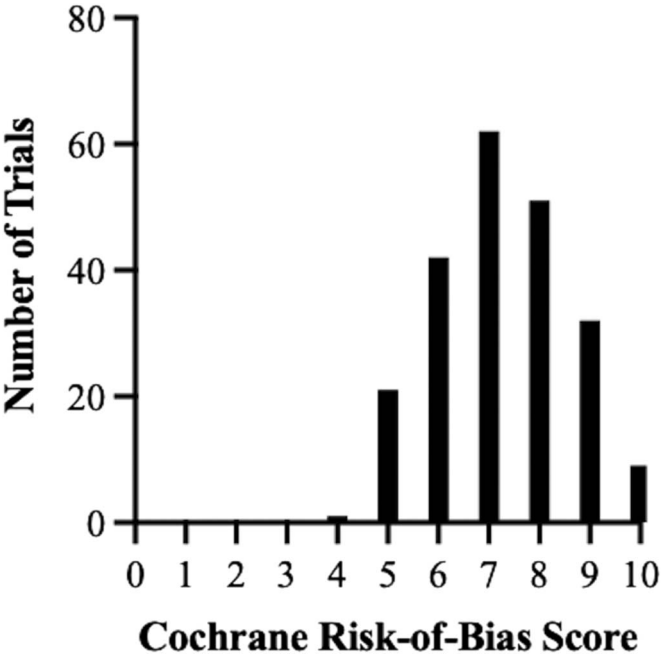


Fig. 3

The frequency of modified risk-of-bias scores among RCTs published in *JBJS* from 2014 to 2022.

as the benchmark above which a study may be considered high quality[10]. In the present study, we found that 82.1% of RCTs published in *JBJS* from 2014 to 2022 had a transformed Detsky score of >75% and were therefore of high quality. This proportion was significantly higher than that from 1998 to 2000 (40.3% of *JBJS* RCTs) and that from 2001 to 2013 (46.7%)[2,10]. Additionally, the mean transformed Detsky scores for RCTs from 2014 to 2022 across all categories of intervention and body region were similarly high. Likewise, although Smith et al. determined that a nonsurgical trial design and financial support were associated with higher transformed Detsky scores from 2001 to 2013[2], these effects no longer persisted from 2014 to 2022. Altogether, these findings indicate that the review process at *JBJS* has been effective at preventing selection bias on the basis of these factors.

On the basis of the modified Cochrane risk-of-bias assessment, we found that trials published from 2014 to 2022 were less likely to be biased than those from 2001 to 2013 (mean score, $7 \pm 1$ versus $6 \pm 1$; $p < 0.001$) but that less than half of the trials from 2014 to 2022 had the sufficiently low levels of bias needed to be considered high quality. Despite improvements in the overall quality of the published studies, these findings highlight that investigators need to continue to strive for transparency and reduction of bias in their RCTs. To aid in this process, in 2016, *JBJS* began requiring the prospective registration of clinical trials in a public registry[13]. Prospective registration involves the public release of methodology and outcomes prior to the enrollment of patients and is thought to reduce selective outcome reporting (SOR), which occurs when original outcome variables are selectively chosen for publication on the basis of their significance[18,19]. SOR is a likely consequence of the higher probability of publication for studies reporting positive results[19]. Interestingly, the proportion of trials with positive results published in *JBJS* increased by nearly 15 percentage points from the 2001 to 2013 period to the 2014 to 2022 period. Despite this trend, from 2014 to 2022, the modified Cochrane risk-of-bias score—which incorporates an assessment of SOR—did not significantly differ between trials reporting positive results and those reporting negative results. Taken together, these findings indicate that the policies to reduce SOR may have been successful. However, given that there are 9 other components to the risk-of-bias score, it is unclear which methodological factors among trials with positive results and those with negative results contributed to their observed similarity in quality. In addition, the rising proportion of trials with positive results may reflect the increased involvement of the industry, as commercially funded studies are more likely to report favorable findings[20,21]. Although the

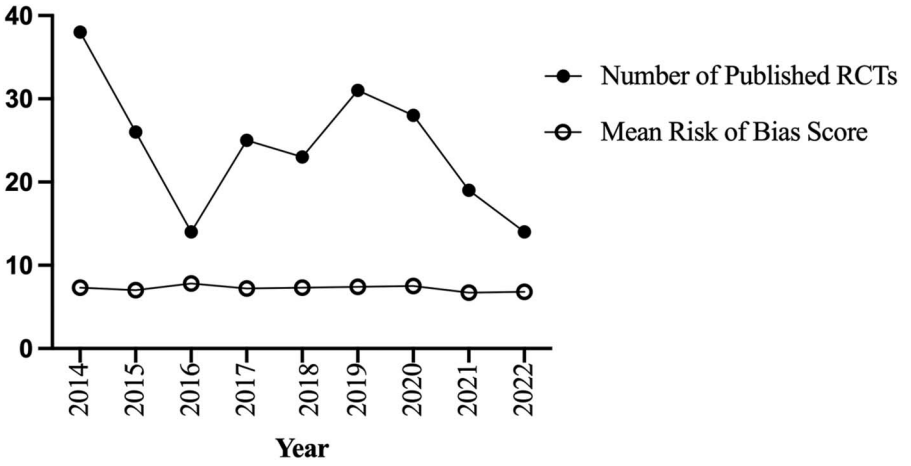


Fig. 4

The number of RCTs published in *JBJS* and the mean modified risk-of-bias score by year.

**TABLE VI Mean Modified Risk-of-Bias Score by the Category of Intervention***

| Intervention | No. of Trials | Mean Modified Risk-of-Bias Score |
|---|---|---|
| Treatment of degenerative disease | 53 (24.3%) | 6.9 ± 2.0 |
| Fracture treatment | 47 (21.6%) | 6.2 ± 1.8 |
| Pain management | 29 (13.3%) | 7.3 ± 1.7 |
| Thrombosis prevention | 5 (2.3%) | 7.8 ± 2.0 |
| Other | 84 (38.5%) | 7.0 ± 1.9 |

*The values are given as the count, with the percentage in parentheses, or as the mean ± standard deviation.

proportion of studies citing financial support did not significantly differ between 1998 to 2000, 2001 to 2013, and 2014 to 2022, it is possible that the proportion of funded trials receiving payments from the industry has increased. The Detsky score and the modified Cochrane risk-of-bias assessment tool do not directly evaluate the effect of financial conflicts of interest on the risk of bias, and thus conclusions cannot be drawn regarding the influence of such factors.

A number of other policies implemented by *JBJS* editors may have also influenced the rising quality of published RCTs, the earliest of which was the requirement for RCTs to follow CONSORT guidelines[13,22]. The CONSORT statement is a 25-item checklist published in 1996 to improve the quality of reporting in RCTs[22,23]. Although a formal analysis of reporting quality in *JBJS* following the adoption of CONSORT guidelines has not yet been performed, evaluations of RCTs in other journals have shown improvement in RCT reporting quality following the inclusion of CONSORT guidelines in journal policy[24].

Despite significantly reduced bias among trials published from 2014 to 2022 compared with that among trials from 2001 to 2013, only 42.2% of RCTs from 2014 to 2022 were regarded as high quality on the basis of the modified Cochrane risk-of-bias tool. This value is appreciably lower than that calculated with use of the Detsky score, with which 82.1% of studies were found to be of high quality. Although the 2 scales overlap in many crucial domains, the tools assign different weights to each item. The modified risk-of-bias tool also uniquely considers the impacts of attrition bias and surgeon experience with the intervention. In the present analysis, 83.9% of RCTs (183 of 218) conducted an a priori power analysis, but only 83.1% of the studies that conducted such an analysis (152 of 183) enrolled the appropriate sample size and only 82.9% of these studies (126 of 152) maintained an adequate number at the time of final follow-up. This exemplifies a considerable degree of attrition bias, as only 57.8% of RCTs (126 of 218) were adequately powered at the time of final follow-up. Notably, the Detsky score evaluates patient eligibility, including whether patients were excluded and why as well as whether a sample size calculation was performed a priori, but it does not directly assess attrition bias. Most RCTs published from 2014 to 2022 (86.7%; 189 of 218) achieved a perfect score in the "eligibility" component of the Detsky score but performed worse in the remaining 4 categories (randomization, outcome, intervention, and statistics). Future RCTs are recommended to perform a priori sample-size calculations and to over-enroll in order to ensure that there is an adequate number of patients at the time of follow-up, but this may be particularly challenging for trials with extended follow-up durations or funding limitations. Smith et al. advocated for more trials to include data from multiple centers in order to address concerns of underpowered analyses[2], but the proportion of trials published from 2014 to 2022 that were multicenter studies (33%) remained similar to the proportion reported for the 2001 to 2013 period (31%), which represents a 51% reduction from the 1998 to 2000 period (67%). Although the nature of surgery may make it difficult to blind patients and outcome assessors, attempts to blind other support staff, such as data analysts, should be made and described in future trials.

### *Limitations*

This study had limitations. First, although 2 reviewers assessed the quality of each trial with substantial agreement, subjectivity may be inherent to the use of these tools. Despite the substantial interobserver agreement portrayed by the ICC scores, the raters were not professionally trained in the use of these tools. Additionally, only trials published in *JBJS* were included in the analysis, which limits the generalizability of these findings. However, RCTs in *JBJS* are likely of higher quality than those published in other orthopaedic journals as indicated by *JBJS's* relatively higher impact factor and greater influence on clinical practice[2,10]. Furthermore, the quality assigned by the assessment depends on the reporting quality of each trial. Trials with poor reporting of sound methodology may have inappropriately been assigned lower quality scores[2], but this scoring may have been limited by adherence to the required CONSORT guidelines.

**TABLE VII Mean Modified Risk-of-Bias Score by Body Region***

| Body Region | No. of Trials | Mean Modified Risk-of-Bias Score |
|---|---|---|
| Hip and knee | 105 (48.2%) | 7.3 ± 1.9 |
| Upper extremity | 58 (26.6%) | 6.5 ± 1.6 |
| Spine | 19 (8.7%) | 6.6 ± 2.2 |
| Foot and ankle | 17 (7.8%) | 7.1 ± 1.9 |
| Lower extremity | 11 (5.0%) | 6.3 ± 2.2 |
| Uncategorized | 8 (3.7%) | 5.5 ± 1.1 |

*The values are given as the count, with the percentage in parentheses, or as the mean ± standard deviation.

*Conclusions*

The quality of RCTs published in *JBJS* from 2014 to 2022 has improved from that reported previously, likely as a result of journal policies such as the requirements of CONSORT adherence and prospective trial registration. Investigators should focus on improving the clarity of reporting, limiting attrition bias, and making efforts to blind support staff in order to increase the quality of future RCTs. ■

Nareena Imam, BA[1]
Suleiman Y. Sudah, MD[2]
Siraj Z. Shaikh, BS[3]
Ashley A. Bonney, BS[4]
Allen D. Nicholson, MD[2]
Surena Namdari, MD[5]
Mariano E. Menendez, MD[6]

[1]Department of Orthopedic Surgery, Rutgers Robert Wood Johnson Medical School, New Brunswick, New Jersey

[2]Department of Orthopedic Surgery, Monmouth Medical Center, Long Branch, New Jersey

[3]Department of Orthopedic Surgery, Rutgers New Jersey Medical School, Newark, New Jersey

[4]Department of Orthopedic Surgery, Hackensack Meridian School of Medicine, Nutley, New Jersey

[5]Department of Orthopaedic Surgery, Rothman Orthopaedic Institute, Sidney Kimmel Medical College, Thomas Jefferson University, Philadelphia, Pennsylvania

[6]Oregon Shoulder Institute, Medford, Oregon

Email for corresponding author: ni60@rwjms.rutgers.edu

## References

**1.** The Journal of Bone & Joint Surgery. JBJS, Inc. Journals Level of Evidence. Accessed 2023 Feb 22. https://journals.lww.com/jbjsjournal/pages/journals-level-of-evidence.aspx.

**2.** Smith CS, Mollon B, Vannabouathong C, Fu JM, Sales B, Bhandari M, Whelan DB. 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 Oct 21;102(20):e116.

**3.** Bederman SS, Wright JG. Randomized trials in surgery: how far have we come? J Bone Joint Surg Am. 2012 Jul 18;94(Suppl 1):2-6.

**4.** Cunningham BP, Harmsen S, Kweon C, Patterson J, Waldrop R, McLaren A, McLemore R. Have levels of evidence improved the quality of orthopaedic research? Clin Orthop Relat Res. 2013 Nov;471(11):3679-86.

**5.** McLeod RS, Wright JG, Solomon MJ, Hu X, Walters BC, Lossing AI. Randomized controlled trials in surgery: Issues and problems. Surgery. 1996 May;119(5):483-6.

**6.** Higgins JP, Altman DG, Gøtzsche PC, Jüni P, Moher D, Oxman AD, Savovic J, Schulz KF, Weeks L, Sterne JA; Cochrane Bias Methods Group; Cochrane Statistical Methods Group. The Cochrane Collaboration's tool for assessing risk of bias in randomised trials. BMJ. 2011 Oct 18;343:d5928.

**7.** Alam M, Rauf M, Ali S, Nodzenski M, Minkis K. A systematic review of reporting in randomized controlled trials in Dermatologic Surgery: Jadad scores, power analysis, and sample size determination. Dermatol Surg. 2014 Dec;40(12):1299-305.

**8.** Kiehna EN, Starke RM, Pouratian N, Dumont AS. Standards for reporting randomized controlled trials in neurosurgery. J Neurosurg. 2011 Feb;114(2):280-5.

**9.** Welk B, Afshar K, MacNeily AE. Randomized controlled trials in pediatric urology: room for improvement. J Urol. 2006 Jul;176(1):306-9, discussion 309-10.

**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(3):388-96.

**11.** McCormick F, Cvetanovich GL, Kim JM, Harris JD, Gupta AK, Abrams GD, Romeo AA, Provencher MT. An assessment of the quality of rotator cuff randomized controlled trials: utilizing the Jadad score and CONSORT criteria. J Shoulder Elbow Surg. 2013 Sep;22(9):1180-5.

**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 Jul-Aug;35(5):536-45.

**13.** Journal of Bone & Joint Surgery. Instructions for Authors. 2023 Sep 27. Accessed 2023 Feb 24. https://journals.lww.com/jbjsjournal/Pages/Instructions-for-Authors.aspx.

**14.** Moher D, Jones A, Lepage L; CONSORT Group (Consolidated Standards for Reporting of Trials). Use of the CONSORT statement and quality of reports of randomized trials: a comparative before-and-after evaluation. JAMA. 2001 Apr 18;285(15):1992-5.

**15.** Agha RA, Fowler AJ, Limb C, Whitehurst K, Coe R, Sagoo H, Jafree DJ, Chandrakumar C, Gundogan B. Impact of the mandatory implementation of reporting guidelines on reporting quality in a surgical journal: A before and after study. Int J Surg. 2016 Jun;30:169-72.

**16.** Clarivate. JCR 2023 Statistics. Accessed 15 Dec 2023. https://clarivate.com/products/scientific-and-academic-research/research-analytics-evaluation-and-management-solutions/journal-citation-reports/infographic/.

**17.** 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 Mar;45(3):255-65.

**18.** Hahn S, Williamson PR, Hutton JL, Garner P, Flynn EV. Assessing the potential for bias in meta-analysis due to selective reporting of subgroup analyses within studies. Stat Med. 2000 Dec 30;19(24):3325-36.

**19.** Hopewell S, Loudon K, Clarke MJ, Oxman AD, Dickersin K. Publication bias in clinical trials due to statistical significance or direction of trial results. Cochrane Database Syst Rev. 2009 Jan 21;2009(1):MR000006.

**20.** 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 Dec;37(12):E205-12, discussion E212.

**21.** Amiri AR, Kanesalingam K, Cro S, Casey AT. Does source of funding and conflict of interest influence the outcome and quality of spinal research? Spine J. 2014 Feb 1;14(2):308-14.

**22.** Moher D, Hopewell S, Schulz KF, Montori V, Gøtzsche PC, Devereaux PJ, Elbourne D, Egger M, Altman DG. CONSORT 2010 explanation and elaboration: updated guidelines for reporting parallel group randomised trials. BMJ. 2010 Mar 23;340:c869.

**23.** CONSORT. Welcome to the CONSORT Website. Accessed 2022 Dec 1. http://www.consort-statement.org.

**24.** Hopewell S, Ravaud P, Baron G, Boutron I. Effect of editors' implementation of CONSORT guidelines on the reporting of abstracts in high impact medical journals: interrupted time series analysis. BMJ. 2012 Jun 22;344:e4178.