Analysis of the Evidence Underpinning the American Academy of Orthopedic Surgeons Knee Osteoarthritis Clinical Practice Guidelines

Philo Waters, BS,* Reece Anderson, MPH, J. Michael Anderson, DO, Jared Scott, DO, Byron Detweiler, DO, Sam Streck, BS, Micah Hartwell, PhD, Trevor Torgerson, DO, and Matt Vassar, PhD

Context: Clinical practice guidelines (CPGs) are vital to establishing a standardized and evidence-based approach in medicine. These guidelines rely on the use of methodologically sound clinical trials, and the subsequent reporting of their methodology.

Objective: To evaluate the completeness of randomized controlled trials (RCTs) underpinning CPGs published by the American Academy of Orthopedic Surgeons (AAOS) for management of osteoarthritis of the knee.

Data Sources: We searched the most recent AAOS CPGs for surgical and nonsurgical management of osteoarthritis of the knee for RCTs. To estimate the necessary sample size, we performed a power analysis using OpenEpi 3.0 (openepi.com).

Study Selection: Two authors independently screened the reference sections of the included CPGs. Included studies met the definition of an RCT, were retrievable in the English language, and were cited in at least one of the included CPGs.

Study Design: Meta-Analysis

Level of Evidence: Level 1a

Data Extraction: We performed double-blind screening and extraction of RCTs included in the AAOS CPGs. We evaluated each RCT for adherence to the Consolidated Standards of Reporting Trials (CONSORT) 2010 checklist. A multiple regression analysis was conducted to assess CONSORT adherence against characteristics of included studies (ie, type of intervention, funding source, etc).

Results: Our study included 179 RCTs. The overall adherence was 68.5% with significant differences between those published before and since the development of the 2010 CONSORT guidelines ($P = 0.02$). We found that RCTs receiving funding from industry/private sources as well as studies that included a conflict of interest statement showed more completeness than RCTs that reported receiving no funding ($P < 0.01$).

Conclusion: We found suboptimal CONSORT adherence for RCTs cited in AAOS CGPs for management of osteoarthritis of the knee. Therefore, the CPGs are likely supported by outdated evidence and lack of high-quality reporting. It is important that evidence used to guide clinical decision making be of the highest quality in order to optimize patient outcomes. In order for clinicians to confer the greatest benefits to their patients, CPGs should provide the totality of evidence and emphasize emerging high-quality RCTs to ensure up-to-date, evidence-based clinical decision-making.

Keywords: adherence; clinical practice guidelines; CONSORT; knee osteoarthritis; reporting

*Address correspondence to Philo Waters, BS, Oklahoma State University Center for Health Sciences, 1111 W 17th St., Tulsa, OK 74107 (email: pwaters@okstate.edu) (Twitter: @PWaters04).

All authors are listed in the Authors section at the end of this article.

The following authors declared potential conflicts of interest: M.H. has received grants from the National Institute of Justice. M.V. has received grants from the National Institute on Drug Abuse, National Institute on Alcohol Abuse and Alcoholism, US Office of Research Integrity, and the Oklahoma Center for Advancement of Science and Technology. All outside of the present work.

DOI: 10.1177/19417381221112674

Clinical practice guidelines (CPGs) are fundamental in establishing a unified, evidence-based approach to patient care across all medical specialties. CPGs provide a summation of the most up-to-date evidence on a given disease process and offer recommendations that are intended to reflect the most current understanding on a clinical problem at the time of its creation. However, these guidelines are reliant on the publication and subsequent location of high-quality, unbiased research. This limitation has resulted in the development of guidelines in multiple fields of medicine based on evidence with questionable methodological quality.[2,26,30,35] Because previous studies have shed light on the less than desirable quality of the evidence underpinning current CPG recommendations, it would be fair to question whether the CPGs in the field of orthopaedic surgery may be suffering from the same shortcomings.

CPG development has proved a difficult task in orthopaedic surgery, with many leaders in the field critiquing their creation and subsequent implementation process.[16,31] These concerns have proved to be beyond anecdotal. Another recent study found that only 18% of recommendations are defined as "strong" and supported by level 1 evidence.[29] However, other researchers have suggested that even the highest levels of evidence supporting orthopaedic CPGs may be compromised by publication bias, financial conflicts of interest, underpowered conclusions, and low reproducibility.[4,5,11,33] Orthopaedic CPGs have found difficulties not just in production but also in adherence to the guideline's recommendations. For example, great variability in adherence to recommendations has been well established in the management of knee osteoarthritis, with authors reporting adherence rates low as 21% in some situations.[3,18] More worrisome may be the disconnect between adherence from physicians and recommendation adoption by insurance providers, placing potential unjust cost on the patient and healthcare system.[27,40] Despite these limitations, the development of methodologically sound studies which serve as the evidentiary foundation on which CPG recommendations are established is critical to ensure clinicians, patients, and health policy makers are informed of the strongest supporting evidence when making critical clinical decisions.

As shown, the depth of investigation into improvement of orthopaedic CPGs is robust and demonstrates the community's concern for producing high-quality CPGs in the future. One way this improvement begins is with the creation and subsequent location of methodologically sound, high-quality randomized controlled trials (RCTs), which are considered as level I evidence in the field of orthopaedic surgery.[14] To date, no study has investigated the methodological quality of the current RCTs supporting the American Academy of Orthopaedic Surgery (AAOS) CPG recommendations. By furthering our knowledge of the quality of key studies underpinning these recommendations, we hope to identify gaps in RCT methodology and reporting, with the goal of bettering the strength of CPG recommendations in orthopaedic surgery. Therefore, the primary objective of the present study was to evaluate the variability in methodological quality of RCTs supporting CPG recommendations by specifically evaluating RCTs cited within the AAOS CPG for surgical and nonsurgical management of knee osteoarthritis.

METHODS

Data extraction was pilot tested in accordance with the prespecified methodology detailed in this protocol. This study was exempt from institutional review board oversight because it did not qualify as human subject research. To facilitate reproducibility and transparency of our results, we have made available all study materials via the Open Science Framework.[1]

Outcomes

Our primary objective was to evaluate the methodological quality and reporting of RCTs that support the recommendations from the AAOS surgical management of osteoarthritis of the knee and osteoarthritis of the knee CPGs.[15,42]

Identification of CPGs

We identified the above-mentioned guidelines using OrthoGuidelines.org, a website established by the AAOS to improve visibility and ease of access to all published recommendations. From this website, 1 investigator obtained the guidelines relating to surgical and nonsurgical management of osteoarthritis of the knee.

Identification of RCTs

After CPGs were obtained, 2 investigators independently screened the reference sections of the included CPGs to identify RCTs cited within the guidelines. To be considered for inclusion, selected studies were required to (1) meet the definition of an RCT, as defined by the International Committee of Medical Journal Editors (ICMJE)[12]; (2) be retrievable in the English language; and (3) be cited in at least one of the included CPGs. We used the kappa statistic to measure interrater reliability during the screening process. A kappa statistic ≥0.9 was required before proceeding with final data extraction as outlined below. If the kappa statistic was <0.9, investigators would reconvene for additional training and standardization of responses before proceeding.

CONSORT

The Consolidated Standards of Reporting Trials (CONSORT) statement, consisting of 25 items, provides guidance on proper methodological reporting of RCTs.[32] The rationale behind including this assessment tool is supported by the roughly 50% of core medical journals indexed through the Abridged Index Medicus on PubMed who currently endorse that authors adhere to the rigorous reporting requirements outlined within this checklist.[9] RCTs included in our study were scored based on adherence to each checklist item in a similar fashion as that used in previous investigations.[13,24,25] One point was awarded for full compliance with a given checklist item, 0.5 points for partial compliance, and 0 points for noncompliance. An overall final score was determined for each RCT based on the degree

by which authors adequately completed the requirements outlined in the CONSORT statement. An overall CONSORT score was calculated out of a total of 31 possible points.

Data Extraction and Scoring

Data extraction was performed by 2 independent authors in a blinded and duplicate fashion. Before commencement of data extraction, these authors completed training exercises which provided detail regarding the use of the CONSORT checklist, as well as instruction on the use of a pilot tested Google form used to catalogue authors' responses. To ensure consistency and accuracy of extraction between investigators, both authors extracted data from 5 RCTs using the CONSORT tool. After this exercise, the authors held a consensus meeting to resolve discrepancies in form responses. Similar to screening of CPG reference sections, the same interrater reliability kappa statistic ≥0.9 was used to ensure consistency between investigators performing data extraction. Following this training, authors continued to extract data from the remaining RCTs. In addition to extraction of CONSORT items, authors were prompted to extract the following study characteristics: year of publication, participant population, intervention, sample size, and mean follow-up. All data extraction was conducted in a duplicate, blinded manner. Disagreements between investigators were resolved by a third investigator, if necessary.

Statistical Analysis

To facilitate reproducibility and transparency of our results, we employed a 2-factor extraction by independent authors and repeated analyses by independent and blinded statisticians. Results were reported using descriptive statistics. A multiple regression model was constructed to evaluate the relationships between CONSORT completion and other extracted study characteristics accounted for variance in CONSORT scores. All analyses were computed using Stata 16.1 (Stata Corp, LLC, College Station, TX, USA).

Sample Size Determination

To estimate the necessary sample size, we performed a power analysis using OpenEpi 3.0 (openepi.com). We considered an RCT to have "adequately" complied with the CONSORT guidelines if ≥75% of CONSORT items were sufficiently met, as done in previously published investigations on CONSORT adherence.[25] Estimated parameters using a population size of 355 RCTs included a hypothesized percentage frequency of 37% for "adequate" adherence to CONSORT reporting (based on data obtained by Ngah et al[25]), a confidence limit of 5%, and a design factor of 1, which is used in random sampling. Based on these factors, we anticipated a sample size of 179 RCTs. Following screening of the AAOS CPG references for the surgical and nonsurgical management for osteoarthritis of the knee, we used the random number function of Excel to generate a random sample of 179 RCTs to be analyzed.

RESULTS

From the reference sections of the AAOS CPGs for the surgical and nonsurgical management of osteoarthritis of the knee, we identified 443 unique citations. Of these citations, 355 were found to be RCTs (Table 1). These RCTs were randomly assigned to yield the 179 required RCTs which were included in our final analysis (Figure 1).

Study Characteristics

The 179 RCTs were published between 1986 and 2015; 112 RCTs (of 179; 62.6%) included in our analysis were published before 2010, the year in which the CONSORT guidelines were implemented in biomedicine. More specifically, 76 (42.5%) of the studies were published more than 15 years ago (2006), with 17 (9.5%) being published in the 20th century. Funding statements were provided in 115 RCTs (64.2%). Of the 115 RCTs reporting funding support, 34 (29.6%) reported industry/private funding. Seventeen RCTs (of 179; 9.5%) reported receiving no external funding, and 64 RCTs (of 179; 35.8%) did not provide a funding statement. Conflict of interest statements were included in 92 RCTs (of 179; 51.4%). The most common intervention investigated was a drug/pharmaceutical (78/179; 43.6%). The journals most commonly represented included *The Journal of Bone and Joint Surgery* (20/179; 11.2%), *The Journal of Arthroplasty* (14/179; 7.8%), *Annals of the Rheumatic Diseases* (12/179; 6.7%), and *Osteoarthritis and Cartilage* (11/179; 6.1%).

CONSORT

Mean CONSORT scores were calculated as a percentage completed of a total 31 items. Before data reconciliation, the initial extraction of the CONSORT checklist by the independent raters had an agreement of 76.69% (kappa = 0.60, $P < 0.01$). The mean adherence to CONSORT guidelines was 68.5% (SD = 15.7) (Table 2). Seven items were reported in less than 50% of the RCTs. Items with the lowest percentage adherence included item 24 (in which the full trial protocol can be accessed, if applicable), item 10 (which generated the random allocation sequence, enrolled participants and assigned participants to interventions), and item 23 (registration number and name of trial registry; Table 3). Eight items were reported in greater than 90% of the RCTs. Items with the highest percentage adherence included item 22 (interpretation consistent with results, balancing benefits and harms, and considering other relevant evidence), item 2a (scientific background and explanation of rationale), and item 12a (statistical methods used to compare groups for primary and secondary outcomes; Table 3). Mean percentage adherence for RCTs cited within the surgical management of osteoarthritis of the knee and treatment of osteoarthritis of the knee CPGs were 73.7% and 62.9%, respectively.

Table 1. Characteristics of the included clinical practice guidelines

Clinical Practice Guideline	Year of Publication	Geographical Region	References per Guideline	RCTs per Guideline	RCTs as a Proportion of All Studies Cited by CPGs
Surgical management of osteoarthritis of the knee	2015	United States	220	162	73.64%
Treatment of osteoarthritis of the knee	2013	United States	223	193	86.55%
Date range	2013 to 2015	Totals	443	355	80.1%

CPGs, clinical practice guidelines; RCTs, randomized controlled trials.

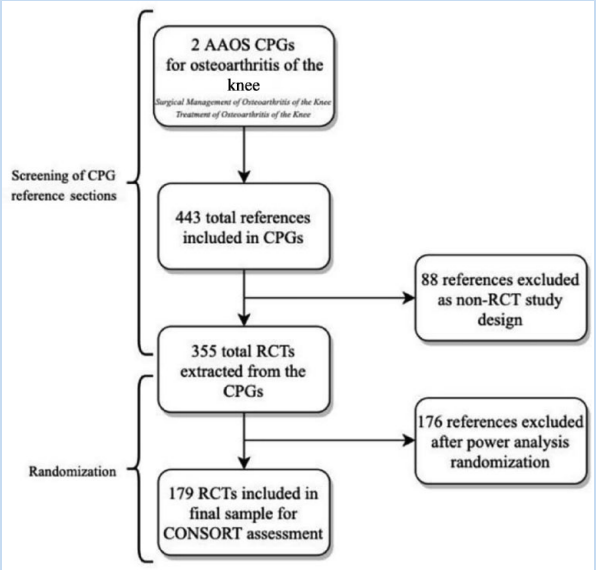

Figure 1. Flow diagram of search strategy for selected trials. AAOS, American Academy of Orthopedic Surgeons; CONSORT, Consolidated Standards of Reporting Trials; CPGs, clinical practice guidelines; RCT, randomized controlled trials.

Multiple Regression

We conducted a multiple regression analysis in which CONSORT percentages were regressed on the type of intervention, publication year, and the presence of funding disclosures and conflict of interest statements (Table 4). There was no difference in CONSORT adherence after stratifying RCTs by type of intervention (Table 4). Results from the Mann-Whitney *U* test revealed RCTs published after 2010 had a higher mean CONSORT score than studies published before 2010 (72.2% vs. 66.3%; $z = -2.317$). RCTs receiving industry funding had higher mean CONSORT adherence scores when compared to RCTs that reported no funding was received (SE = 0.32; $t = 2.75$; $P = 0.01$) (Table 4). Similarly, studies that reported funding support from multiple sources had an 11.7% higher CONSORT adherence score compared to studies that reported no funding was received (SE = 0.35; $t = 2.68$; $P = 0.01$). Trials that included a conflict of interest statement had better CONSORT adherence compared to studies that did not provide a conflict of interest statement (SE = 0.16; $t = 2.39$; $P = 0.01$).

DISCUSSION

We found that adherence to CONSORT reporting standards was suboptimal among RCTs cited as supporting evidence in the AAOS CPGs for the surgical and nonsurgical management of osteoarthritis of the knee. This finding has significant implications for orthopaedic surgeons and patients alike. Of utmost concern surrounds the use of RCTs, considered atop the hierarchy of evidence in orthopaedic surgery,[14] to help establish the most up-to-date, evidence-based CPG recommendations. Because the field of orthopaedic surgery places a heavy emphasis on evidence-based medical decision-making, it is essential that CPG authors are equipped with outcomes from the most reliable and adequately reported RCTs on which to base CPG recommendations. Therefore, we discuss our findings within the broader context of the literature and offer recommendations to better the quality of reporting of RCTs in orthopaedic surgery. Doing so would provide CPG developers with transparent and methodologically sound evidence on which CPG recommendations are based.

Our results demonstrated that RCTs cited as supporting evidence in the AAOS CPGs reported nearly 70% of CONSORT items. This finding is consistent with previous studies measuring CONSORT adherence in the biomedical literature. For example, a systematic review published by Montané et al,[23] which included trials investigating the efficacy of analgesics following traumatic and orthopaedic surgery, found less than one half of CONSORT items were adequately reported. Although these authors reported an overall low adherence to CONSORT items, the quality of reporting did improve over time. These results are similar to ours, as well as others in varying medical specialties,

Table 2. Completeness of reporting of RCTs

RCT Title	CONSORT Percentage Complete[a]	CPG
Evgeniadis et al, 2008	63.8	27355287
Bin et al, 2007	63.8	23996989
Zakeri et al, 2011	60.3	23996989
Newman et al, 2000	43.1	27355287
Mitchell et al, 1991	39.7	27355287
McIlwain et al, 1989	48.3	23996989
Brown et al, 1986	39.7	23996989
Gaffney et al, 1995	45.2	23996989
Ivey et al, 1994	44.8	27355287
Weidenhielm et al, 1993	50.0	27355287
Montgomery et al, 1996	31.0	27355287
Jones et al, 1996	55.2	23996989
Ettinger Jr et al, 1997	80.7	23996989
Sharrock et al, 1997	36.2	27355287
Schroeder-Boersch et al, 1998	39.7	27355287
Bucsi et al, 1998	53.5	23996989
Newman et al, 1998	60.3	27355287
Ritter et al, 1999	31.0	27355287
Kirkley et al, 1999	74.1	23996989
Wobig et al, 1999	61.7	23996989
Bensen et al, 1999	74.1	23996989
Deyle et al, 2000	77.6	23996989
Rindone et al, 2000	69.0	23996989
Niskanen et al, 2000	48.3	27355287
Das Jr et al, 2000	72.6	23996989
Gioe et al, 2000	53.3	27355287
Fransen et al, 2001	71.7	23996989
Reginster et al, 2001	85.5	23996989
JHyldahl et al, 2001	51.7	27355287
McKenna et al, 2001	67.2	23996989
McNamee et al, 2001	56.9	27355287
Ottillinger et al, 2001	66.7	23996989

(continued)

Table 2. (continued)

RCT Title	CONSORT Percentage Complete[a]	CPG
Chiu et al, 2001	48.3	27355287
Adalberth et al, 2001	56.9	27355287
Barrack et al, 2001	56.9	27355287
Maillefert et al, 2001	62.1	23996989
Bradley et al, 2002	84.5	23996989
Moseley et al, 2002	86.2	23996989
Khaw et al, 2002	69.0	27355287
Topp et al, 2002	53.5	23996989
Tanzer et al, 2002	51.7	27355287
Waters et al, 2003	72.4	27355287
Esler et al, 2003	51.7	27355287
Miller et al, 2003	69.0	23996989
Smith et al, 2003	77.6	23996989
Mayman et al, 2003	32.8	27355287
Gur et al, 2003	62.1	23996989
Pham et al, 2004	63.8	23996989
Caborn et al, 2004	69.0	23996989
Maruyama et al, 2004	37.9	27355287
Norgren et al, 2004	53.5	27355287
Miceli-Richard et al, 2004	64.5	23996989
Catani et al, 2004	34.5	27355287
Toda et al, 2004	69.0	23996989
Roth et al, 2004	86.2	23996989
Battisti et al, 2004	46.4	23996989
Vas et al, 2004	87.9	23996989
McAlindon et al, 2004	77.6	23996989
Burnett et al, 2004	53.5	27355287
Christensen et al, 2005	63.8	23996989
Borjesson et al, 2005	31.0	27355287
Decking et al, 2005	67.2	27355287
Mitchell et al, 2005	89.7	27355287

(continued)

Table 2. (continued)

RCT Title	CONSORT Percentage Complete[a]	CPG
Bennell et al, 2005	87.9	23996989
Lehmann et al, 2005	75.9	23996989
Schnitzer et al, 2005	79.0	23996989
Witt et al, 2005	82.8	23996989
Fleischmann et al, 2006	65.5	23996989
Kalairajah et al, 2005	63.8	27355287
Huang et al, 2005	48.3	23996989
Diracoglu et al, 2005	56.9	23996989
Denis et al, 2006	82.8	27355287
Brouwer et al, 2006	88.3	23996989
Petrella et al, 2006	77.4	23996989
Luyten et al, 2007	75.9	23996989
McKenna et al, 2001	70.7	23996989
Brouwer et al, 2006	76.7	23996989
Bingham et al, 2006	75.0	23996989
Perlman et al, 2006	87.9	23996989
Mazieres et al, 2007	79.0	23996989
Rother et al, 2007	75.9	23996989
Kim et al, 2007	41.4	27355287
Williamson et al, 2007	86.2	23996989
Puopolo et al, 2007	84.5	23996989
Toda et al, 2008	79.0	23996989
Hurley et al, 2007	88.7	23996989
Weiner et al, 2007	62.1	23996989
Beaupre et al, 2007	84.5	27355287
Mehta et al, 2007	77.6	23996989
Good et al, 2007	60.3	27355287
Arden et al, 2008	75.9	23996989
Kim et al, 2008	58.6	27355287
Fishman et al, 2007	72.4	23996989
Jan et al, 2008	74.2	23996989

(continued)

Table 2. (continued)

RCT Title	CONSORT Percentage Complete[a]	CPG
Lundsgaard et al, 2008	89.7	23996989
Raman et al, 2008	77.6	23996989
Beaulieu et al, 2008	75.9	23996989
Steffin et al, 2009	55.4	27355287
Lutzner et al, 2008	60.3	27355287
Williams et al, 2000	67.2	23996989
Kahan et al, 2009	93.6	23996989
Ravaud et al, 2009	96.8	23996989
Chevalier et al, 2009	83.3	23996989
Fu et al, 2009	74.1	27355287
Chevalier et al, 2010	85.5	23996989
Jan et al, 2009	77.4	23996989
Lin et al, 2009	87.1	23996989
Topp et al, 2009	53.2	27355287
Forestier et al, 2010	91.9	23996989
Barthel et al, 2009	77.6	23996989
Omonbude et al, 2010	65.0	27355287
Chao et al, 2010	58.1	23996989
Tao et al, 2009	60.3	23996989
Bennell et al, 2010	91.4	23996989
Pavelka et al, 2010	67.2	23996989
Kauppila et al, 2010	91.1	27355287
Trč et al, 2011	55.2	23996989
Lutzner et al, 2010	65.5	27355287
Jorgensen et al, 2010	74.2	23996989
Suarez-Almazor et al, 2010	82.8	23996989
Valtonen et al, 2010	79.3	27355287
Carli et al, 2010	85.0	27355287
Spreng et al, 2010	82.8	27355287
Gstoettner et al, 2011	55.4	27355287
Schnitzer et al, 2011	67.7	23996989

(continued)

Table 2. (continued)

RCT Title	CONSORT Percentage Complete[a]	CPG
Levy et al, 2010	70.7	23996989
Andersen et al, 2010	89.3	27355287
Tunay et al, 2010	53.5	23996989
Ishii et al, 2011	43.1	27355287
Fitzgerald et al, 2011	93.1	23996989
Sun et al, 2012	65.5	27355287
Bennell et al, 2011	88.7	23996989
Bliddal et al, 2011	79.3	23996989
Pavelka et al, 2011	93.6	23996989
Xie et al, 2012	62.9	27355287
Huang et al, 2011	75.9	23996989
Teixeira et al, 2011	74.1	23996989
Park et al, 2011	58.6	27355287
Schnitzer et al, 2012	75.9	23996989
Minns Lowe et al, 2012	93.1	23996989
Minns Lowe et al, 2012	93.1	27355287
Breeman et al, 2011	80.0	27355287
Sun et al, 2012	51.7	27355287
Roy et al, 2012	81.7	27355287
Atamaz et al, 2012	79.3	23996989
Meftah et al, 2012	36.2	27355287
McKay et al, 2012	74.1	27355287
Chia et al, 2013	63.8	27355287
Pietsch et al, 2013	60.3	27355287
Lizaur-Utrilla et al, 2014	56.9	27355287
Chen et al, 2012	72.4	27355287
Matassi et al, 2014	69.0	27355287
Brown et al, 2012	43.1	27355287
Chareancholvanich et al, 2013	74.2	27355287
Harsten et al, 2013	89.7	27355287
Yadeau et al, 2013	71.4	27355287

(continued)

Table 2. (continued)

RCT Title	CONSORT Percentage Complete[a]	CPG
Fernandez-Fairen et al, 2013	91.4	27355287
Roh et al, 2013	79.0	27355287
Hamilton et al, 2013	65.5	27355287
Pandit et al, 2013	80.0	27355287
Boonen et al, 2013	64.5	27355287
Thiengwittayaporn et al, 2013	56.9	27355287
Reinhardt et al, 2014	86.2	27355287
Ngasoongsong et al, 2013	86.2	27355287
Nakai et al, 2013	39.7	27355287
Isosifidis et al, 2014	53.5	27355287
Kim et al, 2014	60.3	27355287
Liu et al, 2014	82.8	27355287
Herrera et al, 2007	62.1	23996989
Giordano et al, 2009	75.9	23996989
Goregaonkar et al, 2009	81.0	23996989
Spangehl et al, 2015	90.0	27355287
Sarzaeem et al, 2014	62.1	27355287
Uesugi et al, 2014	74.1	27355287
Tanikawa et al, 2014	65.5	27355287
Ejaz et al, 2014	91.4	27355287
Pfitzner et al, 2014	72.4	27355287
Tsukada et al, 2014	80.4	27355287
Liu et al, 2014	72.4	27355287
Mean (SD)	68.5 (15.7)	

CONSORT, Consolidated Standards of Reporting Trials; CPG, clinical practice guideline; RCT, randomized controlled trial.
[a]CONSORT score based on number of items located in RCT out of total CONSORT items (31).

indicating that over recent years CONSORT adherence seems to be trending in the right direction. A possible explanation for this improvement may be that journals have begun endorsing the use of CONSORT for authors who submit for publication. Several studies have supported this idea.[8,22] For instance, a large systematic review which synthesized evidence from 16,604 RCTs concluded that endorsement of CONSORT by the journal was associated with more complete reporting of CONSORT items.[39] Another systematic review concluded that the inclusion of a statement recommending or requiring adherence to CONSORT within the journals' instructions for authors was associated with improved CONSORT reporting.[28] Despite these studies demonstrating the benefit of journal policy on checklist adherence, a 2018 study published in the *Journal of Bone and Joint Surgery* concluded that top orthopaedic surgery journals rarely recommended, and even less frequently required, adherence to reporting guidelines.[6] A more recent report in 2020 noted that less one half of orthopaedic surgery journals reference a reporting guideline in the instructions for authors.[10] Moreover, these same authors found that only 12% of

Table 3. Summary of CONSORT completeness scores for AAOS clinical practice guidelines

CONSORT Item	Per CPG 1 (Nonsurgical)		Per CPG 2 (Surgical)		Total	
	n	%	n	%	n	%
1a. Identification as a randomized trial in the title?	93	72.0 (45.1)	86	53.5 (50.2)	179	63.1 (48.4)
1b. Structured summary of trial design, methods, results, and conclusions?	93	88.2 (31.6)	86	49.4 (48.8)	179	69.6 (45.1)
2a. Scientific background and explanation of rationale?	93	100.0 (0.0)	86	98.8 (7.6)	179	99.4 (5.3)
2b. Specific objectives or hypotheses?	93	98.4 (11.5)	86	94.8 (21.7)	179	96.7 (17.3)
3a. Description of trial design including allocation ratio?	93	83.9 (32.3)	86	62.8 (43.5)	179	73.7 (39.4)
4a. Eligibility criteria for participants?	93	96.2 (18.4)	86	84.3 (34.5)	179	90.5 (27.9)
4b. Settings and locations where the data were collected?	93	50.5 (40.0)	86	41.9 (45.9)	179	46.4 (43.1)
5. Interventions for each group with sufficient details to allow replication	93	97.3 (15.4)	86	93.0 (25.6)	179	95.3 (21.0)
6a. Completely defined primary and secondary outcome measures	93	85.5 (22.8)	86	66.9 (26.1)	179	76.5 (26.1)
7a. How sample size was determined?	93	74.2 (44.0)	86	55.8 (50.0)	179	65.4 (47.7)
8a. Method used to generate the randomization	93	58.1 (49.1)	86	55.2 (49.7)	179	56.7 (49.3)
8b. Type of randomization; details of any restriction	93	49.5 (50.3)	86	27.3 (44.5)	179	38.8 (48.7)
9. Mechanism used to implement the random allocation sequence and steps taken to conceal sequence	93	38.7 (49.0)	86	49.4 (50.0)	179	43.9 (49.6)
10. Who generated the random allocation sequence, who enrolled participants, and who assigned participants to interventions?	93	20.4 (30.6)	86	12.2 (24.2)	179	16.5 (27.9)
11a. If done, who was blinded after assignment to interventions and how?	90	72.8 (42.5)	78	57.7 (48.4)	168	65.8 (45.8)
12a. Statistical methods used to compare groups for primary and secondary outcomes?	93	100.0 (0.0)	86	93.6 (24.0)	179	96.9 (16.9)
12b. Methods for additional analyses, such as subgroup analyses and adjusted analyses?	28	69.6 (45.8)	10	50.0 (52.7)	38	64.5 (47.8)
13a. Numbers of participants who were randomly assigned, received intended treatment and were analyzed for the primary outcome?	93	96.2 (15.2)	86	71.5 (42.4)	179	84.4 (33.6)
13b. For each group, losses and exclusions after randomization, together with reasons?	93	94.1 (21.9)	86	77.9 (41.7)	179	86.3 (33.9)
14a. Dates defining the periods of recruitment and follow-up?	93	40.9 (49.4)	86	59.3 (48.2)	179	49.7 (49.6)

(continued)

Table 3. (continued)

CONSORT Item	Per CPG 1 (Nonsurgical)		Per CPG 2 (Surgical)		Total	
	n	%	n	%	n	%
15. A table showing baseline demographic and clinical characteristics for each group?	93	92.5 (25.5)	86	79.7 (40.1)	179	86.3 (33.9)
16. For each group, number of participants (denominator) included in each analysis and whether the analysis was by original assigned groups?	93	92.0 (22.5)	86	66.9 (42.4)	179	79.9 (35.8)
17a. For each primary and secondary outcome, results for each group, and the estimated effect size and its precision (such as 95% CI)?	93	95.7 (15.9)	86	94.2 (16.1)	179	95.0 (16.0)
18. Results of any other analyses performed, including subgroup analyses and adjusted analyses, distinguishing prespecified from exploratory?	25	94.0 (22.0)	9	88.9 (33.3)	34	92.7 (25.0)
19. All important harms or unintended effects in each group?	93	74.7 (42.8)	86	54.1 (48.6)	179	64.8 (46.7)
20. Trial limitations, addressing sources of potential bias, imprecision and, if relevant, multiplicity of analyses?	93	61.8 (48.0)	86	60.5 (48.0)	179	61.2 (47.9)
21. Generalizability (external validity, applicability) of the trial findings?	93	88.2 (29.9)	86	80.8 (38.5)	179	84.6 (34.4)
22. Interpretation consistent with results, balancing benefits and harms, and considering other relevant evidence?	93	100.0 (0.0)	86	99.4 (5.4)	179	99.7 (3.7)
23. Registration number and name of trial registry?	93	20.4 (40.5)	86	19.8 (40.1)	179	20.1 (40.2)
24. Where the full trial protocol can be accessed, if available?	93	11.8 (26.0)	86	12.8 (27.9)	179	12.3 (26.8)
25. Sources of funding and other support and role of funders?	93	80.7 (39.7)	86	48.8 (50.3)	179	65.4 (47.7)
Total percentage complete, mean (SD)*		73.65 (12.49)		62.87 (17)		68.47 (15.74)

AAOS, American Academy of Orthopedic Surgeons; CONSORT, Consolidated Standards of Reporting Trials; CPG, clinical practice guideline; RCT, randomized controlled trial.
*Total percentage complete based on score/total for each RCT.

Table 4. Multiple-regression analysis of clinical trial publication characteristics

	No. (%) of Articles (n = 23)	Unadjusted Model Coefficient (SE)	*t* Value	*P* value	Consort Coefficients (SE)	Consort Standardized Coefficients	*t* Value	*P* value
Type of intervention								
Device	17 (9.5)	1 (Ref)	1 (Ref)	–	1 (Ref)	1 (Ref)	–	–
Drug	78 (43.6)	3.23 (4.0)	0.79	0.429	2.48 (3.59)	0.08	0.69	0.491
Surgical	34 (19.0)	–7.35 (4.53)	–1.62	0.107	–5.59 (4.01)	–0.14	–1.39	0.165
Combo/other	50 (27.9)	4.66 (4.28)	1.09	0.278	3.02 (3.75)	0.09	0.81	0.421
Publication year								
Before 2010	112 (62.6)	1 (Ref)	1 (Ref)	–	1 (Ref)	1 (Ref)	–	–
After 2010	67 (37.4)	5.9 (0.02)	2.46	0.015	5.61 (2.21)	0.17	2.54	0.012
Funding								
No funding received	13 (7.3)	1 (Ref)	1 (Ref)	–	1 (Ref)	1 (Ref)	–	–
No funding statement	64 (35.8)	0.67 (4.23)	0.16	0.875	–2.81 (4.11)	–0.09	–0.68	0.495
Industry/private	44 (24.6)	15.02 (4.39)	3.42	0.001	11.78 (4.28)	0.32	2.75	0.007
Other	58 (32.4)	16.34 (4.27)	3.83	0	11.66 (4.35)	0.35	2.68	0.008
COI statement								
No	87 (48.6)	1 (Ref)	1 (Ref)	–	1 (Ref)	1 (Ref)	–	–
Yes	92 (51.40)	6.81 (2.30)	2.95	0.004	5.06 (2.12)	0.16	2.39	0.018

COI, conflict of interest; CONSORT, Consolidated Standards of Reporting Trials.
CONSORT percentages were regressed on the type of intervention, publication year, and the presence of funding disclosures and conflict of interest statements, to evaluate whether the coded study characteristics accounted for variance in CONSORT scores.

orthopaedic surgery journals go as far as requiring authors to submit a reporting checklist at the time of submission. Because a substantial body of evidence supports the notion that journal policy may improve the quality of reporting of RCTs, we advise adopting a minimum standard for quality reporting, such as CONSORT 2010 guidelines, across journals within the orthopaedic literature.

Nearly two-thirds of the RCTs included in our analysis were published before 2010, with a significant portion being published over 20 years ago. For example, 2 RCTs (published in 1991[21] and 1997[36]) were cited as "high-quality evidence" to support the recommendation for the use of neuraxial anesthesia in patients undergoing total knee arthroplasty (TKA) procedures. Our results indicated that these trials failed adequately to report more than one half of CONSORT items. More recent RCTs (2017 and newer) provide updated patient outcome data following the use of neuraxial analgesia during TKA procedures.[19,37,38,41] Because the AAOS surgical management of osteoarthritis of the knee CPGs were published in 2013, the current recommendations do not account for evidence from these newer trials. While the strength or direction of CPG recommendations might not change in light of newer evidence, we contend that CPGs should be updated in a timely manner such that new information may be considered alongside established literature. In addition, when CPGs are not updated on a consistent basis, recommendations for or against new and emerging treatments are not available, which may hinder access to these treatments, because some insurance companies only cover guideline-recommended interventions.[40] For example, the current AAOS CPG for the surgical management of osteoarthritis of the knee makes no recommendations for the use of adductor canal and interspace between the popliteal artery and capsule of the posterior knee blocks. The popularity of these novel anesthesia options is rising due to their efficacy and muscle sparing properties.[7] Other authors support the idea that guidelines should be updated more frequently, and failure to do so may quickly result in recommendations that are out of date. For example, Martinez García et al reported that 1 out of every 5 CPG recommendations become out of date within as little as 3 years.[17] Scott et al further add to the argument that CPGs should be updated on a regular basis, but also contend that updates to CPGs should, at minimum, acknowledge any ongoing trials that address certain guideline recommendations, especially those in which the recommendation is inconclusive.[34] Our finding, that RCTs published after 2010 had significantly higher mean CONSORT scores compared to RCTs published before 2010, further supports these recommendations. Basing CPG recommendations on methodologically sound and sufficiently reported RCTs would provide an increased level of transparency and integrity to trial outcomes, thus improving the robustness of AAOS CPGs.

Strengths and Limitations

Our study is bolstered by strengths that allow for increased transparency, reproducibility, and internal validity. We accomplish these practices through publishing our protocol on Open Science Framework in order to allow for complete transparency in methodological practices. The double-blind, duplicate screening and extraction technique, considered the gold standard for extraction within meta research,[20] adds strength to our study through increasing the reliability of responses when assessing completeness of the CONSORT guideline checklist. Finally, the researchers involved in screening and extraction received substantial training on CONSORT criteria evaluation in order to strengthen the reliability of the study. Despite implementing the gold standard for data extraction, we acknowledge some limitations to our study. We analyzed a random sample of RCTs included within the AAOS CPGs, rather than the entire sample. However, we believe that our power analysis was sufficient to determine the true effect size. Additionally, this study analyzed only RCTs cited in the AAOS CPGs covering surgical and nonsurgical treatment of osteoarthritis of the knee. Therefore, our findings may not be applicable to the remaining AAOS CPGs or CPGs in other fields of medicine.

CONCLUSION

Our study found that RCTs cited in CPGs are lacking with regard to adherence to the CONSORT checklist. In particular, the RCTs published before 2010 are especially lacking as more journals are implementing and enforcing different reporting guidelines. Therefore, the AAOS CPGs are potentially outdated and lack high-quality reporting of evidence. For physicians to confer the greatest benefits to their patients, it is imperative that they are provided with thoroughly researched and up-to-date recommendations that account for newly published, high-quality studies. Our findings suggest that quality and completeness of reporting demonstrate a positive chronological association, further supporting the need for more frequent evaluation and amendment of CPGs, backed by studies demonstrating high quality of reporting.

AUTHORS

Philo Waters, BS (Office of Medical Student Research, Oklahoma State University Center for Health Sciences, Tulsa, Oklahoma); **Reece Anderson, MPH** (Office of Medical Student Research, Oklahoma State University Center for Health Sciences, Tulsa, Oklahoma); **J. Michael Anderson, DO** (Office of Medical Student Research, Oklahoma State University Center for Health Sciences, Tulsa, Oklahoma); **Jared Scott, DO** (Department of Orthopedic Surgery, Oklahoma State Medical Center, Tulsa, Oklahoma); **Byron Detweiler, DO** (Department of Orthopedic Surgery, Oklahoma State Medical Center, Tulsa, Oklahoma); **Sam Streck, BS** (Office of Medical Student Research, Oklahoma State University Center for Health Sciences, Tulsa, Oklahoma); **Micah Hartwell, PhD** (Office of Medical Student Research, Oklahoma State University Center for Health Sciences, Tulsa, Oklahoma); **Trevor Torgerson, DO** (Office of Medical Student Research, Oklahoma State University Center for

Health Sciences, Tulsa, Oklahoma); and **Matt Vassar, PhD** (Office of Medical Student Research, Oklahoma State University Center for Health Sciences, Tulsa, Oklahoma).

ORCID IDS

Philo Waters https://orcid.org/0000-0003-0918-1517
Reece Anderson https://orcid.org/0000-0001-6429-0033
J. Michael Anderson https://orcid.org/0000-0003-0811-243X
Micah Hartwell https://orcid.org/0000-0001-6810-6571
Trevor Torgerson https://orcid.org/0000-0001-7927-4060
Matt Vassar https://orcid.org/0000-0003-2859-6152

REFERENCES

1. Anderson M, Waters P. *AAOS Lower Extremity CPG - RCT Methodological Quality*. Published online June 3, 2021. https://osf.io/e8m2s/ (accessed June 20, 2021).
2. Aran G, Hicks C, Demand A, et al. Treating schizophrenia: the quality of evidence behind treatment recommendations and how it can improve. *BMJ Evid Based Med*. 2020;25:138-142.
3. Carlson VR, Ong AC, Orozco FR, et al. Compliance with the AAOS guidelines for treatment of osteoarthritis of the knee: a survey of the American Association of Hip and Knee Surgeons. *J Am Acad Orthop Surg*. 2018;26:103-107.
4. Checketts JX, Cook C, Vassar M. An evaluation of industry relationships among contributors to AAOS clinical practice guidelines and appropriate use criteria. *J Bone Joint Surg Am*. 2018;100:e10.
5. Checketts JX, Scott JT, Meyer C, et al. The robustness of trials that guide evidence-based orthopaedic surgery. *J Bone Joint Surg Am*. 2018;100:e85.
6. Checketts JX, Sims MT, Detweiler B, et al. An evaluation of reporting guidelines and clinical trial registry requirements among orthopaedic surgery journals. *J Bone Joint Surg Am*. 2018;100:e15.
7. Cullom C, Weed JT. Anesthetic and analgesic management for outpatient knee arthroplasty. *Curr Pain Headache Rep*. 2017;21:23.
8. Devereaux PJ, Manns BJ, Ghali WA, et al. The reporting of methodological factors in randomized controlled trials and the association with a journal policy to promote adherence to the Consolidated Standards of Reporting Trials (CONSORT) checklist. *Control Clin Trials*. 2002;23:380-388.
9. Endorsers. Consolidated Standards of Reporting Trial. http://www.consort-statement.org/about-consort/endorsers1 Accessed June 2, 2021.
10. Faroug R, Whittaker M. Reporting guidelines in orthopaedic journals: can ee do more to promote adoption? A systematic review. *J Surgery*. 2020;1(3):1013.
11. Fishbeck K, Checketts JX, Cooper CM, et al. Evaluation of the clarity and completeness of reporting in orthopedic clinical practice guidelines. *J Am Osteopath Assoc*. 2020;120:74-80.
12. International Committee of Medical Journal Editors. Recommendations: Clinical Trials. http://icmje.org/recommendations/browse/publishing-and-editorial-issues/clinical-trial-registration.html Accessed June 3, 2021.
13. Institute of Medicine, Board on Health Care Services, Committee on Standards for Systematic Reviews of Comparative Effectiveness Research. *Finding What Works in Health Care: Standards for Systematic Reviews*. Washington, DC: National Academies Press; 2011.
14. JBJS, Inc. *Journals Level of Evidence: JBJS*. https://journals.lww.com/jbjsjournal/pages/journals-level-of-evidence.aspx Accessed June 25, 2021.
15. Jevsevar DS. Treatment of osteoarthritis of the knee: evidence-based guideline, 2nd edition. *J Am Acad Orthop Surg*. 2013;21:571-576.
16. Lubowitz JH, McIntyre LF, Provencher MT, et al. AAOS rotator cuff clinical practice guideline misses the mark. *Arthroscopy*. 2012;28:589-592.
17. Martínez García L, Sanabria AJ, García Alvarez E, et al. The validity of recommendations from clinical guidelines: a survival analysis. *CMAJ*. 2014;186:1211-1219.
18. Meiyappan KP, Cote MP, Bozic KJ, et al. Adherence to the American Academy of Orthopaedic Surgeons clinical practice guidelines for nonoperative management of knee osteoarthritis. *J Arthroplasty*. 2020;35:347-352.
19. Memtsoudis SG, Poeran J, Zubizarreta N, et al. Do hospitals performing frequent neuraxial anesthesia for hip and knee replacements have better outcomes? *Anesthesiology*. 2018;129:428-439.
20. Michael Anderson J, Niemann A, Johnson AL, et al. Transparent, reproducible, and open science practices of published literature in dermatology journals: cross-sectional analysis. *JMIR Dermatol*. 2019;2:e16078.
21. Mitchell D, Friedman RJ, Baker JD III, et al. Prevention of thromboembolic disease following total knee arthroplasty. Epidural versus general anesthesia. *Clin Orthop Relat Res*. 1991;269:109-112.
22. 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;285:1992-1995.
23. Montané E, Vallano A, Vidal X, et al. Reporting randomised clinical trials of analgesics after traumatic or orthopaedic surgery is inadequate: a systematic review. *BMC Clin Pharmacol*. 2010;10:2.
24. Nagendran M, Harding D, Teo W, et al. Poor adherence of randomised trials in surgery to CONSORT guidelines for non-pharmacological treatments (NPT): a cross-sectional study. *BMJ Open*. 2013;3:e003898.
25. Ngah VD, Mazingisa AV, Zunza M, et al. A review of adherence and predictors of adherence to the CONSORT statement in the reporting of tuberculosis vaccine trials. *Vaccines*. 2020;8:770.
26. Nissen T, Wayant C, Wahlstrom A, et al. Methodological quality, completeness of reporting and use of systematic reviews as evidence in clinical practice guidelines for paediatric overweight and obesity. *Clin Obes*. 2017;7:34-45.
27. Ong KL, Niazi F, Lau E, et al. Knee OA cost comparison for hyaluronic acid and knee arthroplasty. *J Orthop Surg Res*. 2020;15:305.
28. Plint AC, Moher D, Morrison A, et al. Does the CONSORT checklist improve the quality of reports of randomised controlled trials? A systematic review. *Med J Aust*. 2006;185:263-267.
29. Reddy AK, Scott J, Checketts JX, et al. Levels of evidence backing the AAOS clinical practice guidelines. *J Orthopaed, Trauma Rehabil*. Epub ahead of print February 10, 2021:2210491721992533.
30. Ross A, Rankin J, Beaman J, et al. Methodological quality of systematic reviews referenced in clinical practice guidelines for the treatment of opioid use disorder. *PLoS One*. 2017;12:e0181927.
31. Sanders JO, Bozic KJ, Glassman SD, et al. Clinical practice guidelines: their use, misuse, and future directions. *J Am Acad Orthop Surg*. 2014;22:135-144.
32. Schulz KF, Altman DG, Moher D, CONSORT Group. CONSORT 2010 statement: updated guidelines for reporting parallel group randomised trials. *Int J Surg*. 2011;9:672-677.
33. Scott J, Checketts JX, Cooper CM, et al. An evaluation of publication bias in high-impact orthopaedic literature. *JB JS Open Access*. 2019;4:e0055.
34. Scott J, Checketts JX, Horn JG, et al. Knee osteoarthritis and current research for evidence - are we on the right way? *Int Orthop*. 2018;42:2105-2112.
35. Scott J, Howard B, Sinnett P, et al. Variable methodological quality and use found in systematic reviews referenced in STEMI clinical practice guidelines. *Am J Emerg Med*. 2017;35:1828-1835.
36. Sharrock NE, Go G, Williams-Russo P, et al. Comparison of extradural and general anaesthesia on the fibrinolytic response to total knee arthroplasty. *Br J Anaesth*. 1997;79:29-34.
37. Soffin EM, Memtsoudis SG. Anesthesia and analgesia for total knee arthroplasty. *Minerva Anestesiol*. 2018;84:1406-1412.
38. Turcotte JJ, Stone AH, Gilmor RJ, et al. The effect of neuraxial anesthesia on postoperative outcomes in total joint arthroplasty with rapid recovery protocols. *J Arthroplasty*. 2020;35:950-954.
39. Turner L, Shamseer L, Altman DG, et al. Does use of the CONSORT statement impact the completeness of reporting of randomised controlled trials published in medical journals? A Cochrane review. *Syst Rev*. 2012;1:60.
40. Vangsness CT Jr, Adamson TC III, Daley MJ. Consequences on private insurance coverage: the AAOS clinical practice guidelines and hyaluronic acid injections. *J Bone Joint Surg Am*. 2020;102:920-926.
41. Walker JB, Nguyen PL, Schmidt UH, et al. Postoperative outcomes associated with neuraxial vs general anesthesia following bilateral total knee arthroplasty. *J Arthroplasty*. 2017;32:3632-3636.
42. Weber KL, Jevsevar DS, McGrory BJ. AAOS clinical practice guideline: surgical management of osteoarthritis of the knee: evidence-based guideline. *J Am Acad Orthop Surg*. 2016;24:e94-e96.