Review Article

Low prevalence of spin in conclusions of interventional pediatric orthopedic studies

John Amen, PhD., FRCS.[1] ⓘ, Sherif Ahmed Kamel, MD. (Ortho), FRCSEd. (Tr&Ortho).[1] ⓘ, Tamer A. El-Sobky, PhD.[1] ⓘ

[1]Department of Orthopedic Surgery, Faculty of Medicine, Ain Shams University, Cairo, Egypt.

*Corresponding author:

Tamer A. El-Sobky,
Department of Orthopedic
Surgery, Faculty of Medicine,
Ain Shams University, Cairo,
Egypt.

tamer.ahmed@med.asu.edu.eg

ABSTRACT

Interpretation bias as a factor of research reporting quality has not been thoroughly investigated in the conclusions of pediatric orthopedic publications. Our objective was to investigate the prevalence, subtypes, and severity of research reporting bias or spin of the conclusions in full-texts and abstracts of published studies investigating the effects of treatment/intervention in the pediatric orthopedic literature. We systematically searched ten high-ranking orthopedic journals on MEDLINE/PubMed. Inclusion criteria were pediatric orthopedic studies investigating the effects of treatment/intervention. We used descriptive statistics to report the prevalence, subtype, and severity of reporting bias in the studies' conclusions according to validated classification criteria. We checked the results to ensure that data were neither misreported nor misinterpreted/extrapolated in the conclusions of the full-texts and their abstracts. Out of 93 included studies in the final analysis, 17 (18%) had at least one count of bias. Nine (10%) studies had bias in both full-text and the corresponding abstract conclusions. In four (4%) studies, bias was restricted to conclusions of abstracts only, and in four (4%) studies was restricted to a bias criterion assigned to the classification criteria of conclusions of full-text only and not the abstract. We analyzed 2511 spin/bias items across 93 studies and reported 30 (1%) counts of bias in conclusions of full-text and/or abstracts. The intervention was surgical in (71%) of studies. Interventional pediatric orthopedic studies published in high-ranking journals showed a low prevalence of reporting bias, namely, misleading reporting, misinterpretation, and inadequate extrapolation of conclusions. A comparative analysis with lower-ranking journals as a control group may reveal if our favorable results are an attribute of journal rank/quality. In general, editorial policies should emphasize skilled interpretation and extrapolation of research results.

Keywords: Bias, Guideline adherence, Orthopedic procedures, Pediatric orthopedics, Research design, Research reporting quality

INTRODUCTION

The medical literature is the primary source of information for academics, clinical practice physicians, health policymakers, and the public. Several factors can erode trust and reliance on the published literature. Enhancing academic research – design, performance, and writing skills, especially for early-career researchers – was cited as a means of improving research quality and boosting trust in the medical literature.[1,2] Control of research reporting quality by eliminating sources of bias and enhancing impartiality in research writing is important for the trustworthiness of the findings and conclusions of published medical literature.[3] Research bias includes reporting bias, interpretation bias, and publication bias, among other sources of

bias related to different stages of the research process, such as study design bias, recruitment bias etcetera.[3] There is a general consensus that validated assessment tools of bias/spin in published interventional research fall under three main categories, namely, (a) misleading reporting, (b) misleading interpretation, and (c) inappropriate extrapolation.[4,5]

While research reporting bias refers to selective and inaccurate reporting of study results/correlations, interpretation bias refers to misinterpretation or non-contextual extrapolation of the conclusions of a published article.[4,5] Multiple factors can drive reporting and interpretation bias, such as authors' own professional preferences or aspirations, conflicts of interest, and academic and societal pressures etcetera. This may occur consciously or subconsciously.[6-8] Such biases are occasionally attributable to substandard research writing skills in drawing proper inferences and clinical contextual interpretations of the resultant statistical values.[2,9,10] Authors may unjustifiably use causal language to interpret *assumingly* positive statistical correlations between intervention and effect where they actually are associative or uncertainly defined at best. [2,9,10] Research interpretation bias, also known as *spin,* has been documented across various medical[5,7] and surgical disciplines.[5,11] Research interpretation bias has also been documented in abstracts reporting on the treatment of various orthopedic disorders such as trauma,[12] arthroscopy and sports medicine,[13] and Achilles tendon injuries,[14] among others. Furthermore, a systematic review (SR) of surgical randomized controlled trials (RCTs) found that spin or interpretation bias was more likely to be reported in the abstracts than in their corresponding full-text articles.[11] Alarmingly, interpretation bias or spin of media coverage of medicine was found to be attributable to the interpretation bias initially found in the conclusions of the published articles.[15] Interpretation bias of research writers/producers can take various forms. For instance, overstated conclusions that exaggerate the favorable effects of an intervention/treatment without being supported by the results, undue implications that embrace a larger or different study population/setting, and unjustified clinical practice recommendations can mislead readers, distort the scientific literature, and waste research resources.[5,9-11,13] For example, causal language 53%,[5] claim of equivalence/non-inferiority versus control for a negative endpoint 15%,[12] and selective reporting of the efficacy of outcomes 6.8%[13] were the most prevalent types of spin identified in conclusions/results of included interventional studies in medicine and orthopedic surgery. Contrastingly, interpretation bias of research readers/consumers can misinterpret conclusions made in the article that is being cited and produce the same detrimental consequences on the scientific literature in general[16] and orthopedic literature in specific.[17] Moreover, authors of citing articles show a specific predilection for referencing articles with positive statistical significance.[18] This represents an additional source of bias in the published

biomedical literature.[18] The previous drawbacks prompted various academic interventions aimed at controlling bias in the published medical literature.[19,20] However, these interventions, namely, prospective trial registration and editorial control, enjoyed little success.[19,20]

Interpretation bias as a factor of research reporting quality has not been thoroughly investigated in the conclusions of pediatric orthopedic publications. The primary objective of this study was to report the prevalence and subtypes of interpretation bias/spin of the conclusions in abstracts and full-texts of published studies investigating the effects of treatment/intervention in pediatric orthopedic literature. The secondary objective was to report the severity of bias/spin of conclusions in published abstracts.

MATERIALS AND METHODS

We searched ten established and high-impact general orthopedic journals, three of which were Pediatric Orthopedic Society journals. According to Scopus's updated CiteScore methodology, all except two included journals were Quartile 1 (Q1). The two exceptions were Q2 journals, which were the official publications of leading, highly sought-after continental and international pediatric orthopedic societies. In addition, all included journals were listed in the release of Journal Citation Reports. Demographics of included journals are shown in Table 1. Inclusion criteria were primary interventional – randomized or non-randomized – pediatric orthopedic studies investigating the effects of a treatment or an intervention (surgical or non-surgical). That is, matched comparative studies, classic interventional, and controlled clinical trials. We excluded classic – placebo/no treatment – RCTs. However, interventional studies that did not fit the traditional RCTs' strict definition or principles were still considered based on our inclusion criteria (see working definitions).

In addition, we excluded case reports, editorials, narrative reviews, diagnostic, prognostic, outcome instrument validation studies, SRs/meta-analyses, those whose study samples consisted exclusively of adults, and studies that were exclusively focused on cost analysis without reporting outcomes and irretrievable studies. We searched the MEDLINE/PubMed using the PubMed Advanced Search Builder. We conducted the initial search on February 19, 2023, and set the publication date to 3 years (2020–2022). We included a sample of 93 pediatric orthopedic studies in the final analysis. The flow diagram and detailed results of the literature search and identification process are shown in Figure 1 and Additional File 1, respectively.

Screening and data analysis

We screened the full-texts (and their abstracts) of 93 included studies and focused on the results and conclusion sections. The results were carefully weighed against the conclusions

Table 1: Demographics of included journals https://www.scopus.com/sources.uri

Journal abbreviation	Coverage	Quartile	Cite Score 2021	H-index	SCImago Journal Rank 2021
1. J Child Orthop	2007-2021	Q2	2.7	40	0.59
2. J Bone Joint Surg Am	1946-2021	Q1	8.0	274	1.91
3. J Am Acad Orthop Surg	1998-2021	Q1	4.9	118	1.27
4. J Pediatr Orthop B	1989, 1992-2021	Q2	2.7	54	0.45
5. J Pediatr Orthop	1981-2021	Q1	3.6	99	0.98
6. Arch Orthop Trauma Surg	1989-2021	Q1	4.5	82	0.89
7. Clin Orthop Relat Res	1963-2021	Q1	4.6	217	1.17
8. Knee Surg Sports Traumatol Arthrosc	1993-2021	Q1	7.3	134	1.48
9. Int Orthop	1977-2021	Q1	5.8	96	0.96
10. Bone Joint J formerly: Journal of Bone and Joint Surgery - Series B	2013-2021	Q1	7.5	189	2.17
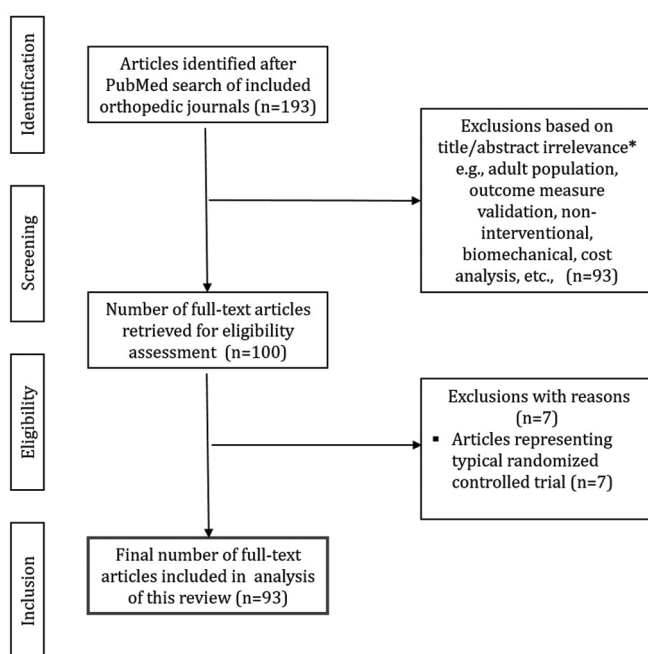
Figure 1: Flow diagram of the literature search and identification process. *Details can be found in Additional file 1.

reported in the full-text articles and their abstracts concerning misleading reporting, misinterpretation, and inadequate extrapolation of data. We used descriptive statistics to report the prevalence, subtype, and severity of the above-noted research reporting bias or spin found in the conclusions according to classification criteria proposed by Lazarus *et al.*, 2015,[5] [Additional File 1]. We compared the reported interpretation bias of full-text conclusions to those found in their abstracts. We established correlations between the recorded prevalence and subtype. The severity of biases in conclusions of abstracts was classified categorically according to Lazarus *et al.*, 2015,[5] as follows; (a) low level of bias/spin was defined as spin reported with uncertainty in the framing and recommendations for further trials, (b) moderate level

as spin reported with some uncertainty in the framing or recommendations for further trials, and (c) high level as spin reported without any uncertainty or recommendations for further trials. T.A.E. did data identification and screening with assistance from the coauthors. The study sample was assigned to authors in equal numbers. The authors independently extracted and analyzed the data. The senior author, T.A.E, also performed data extraction, analysis, and revision/verification for the whole study sample. The results were compared and discrepancies were resolved by agreement.

Spin classification system and definitions

The spin classification for non-randomized studies assessing therapeutic interventions with detailed explanatory notes of each type of bias/spin is shown in Additional File 1.[5]

(a) Spin was defined as the presentation of research data in a manner that exaggerates the beneficial outcomes or underestimates the harmful effects of a treatment/intervention(s), either intentional or unintentional. This occurs by means of under-reporting, inappropriate interpretation, or misleading extrapolation of study data.[4,5]

(b) Non-RCTs: Any clinical trial investigating the effect of treatment/intervention (surgical or medical, and so forth) on specified health-related outcome measures that do not fit the classic criteria of RCTs as randomization employing mathematical techniques such as the use of a random numbers table to assign patients to treatment/intervention or control groups. Typically, RCTs are prospectively designed and aim at establishing a causal interpretation for an intervention, whereas the control group receives a placebo or supportive treatment. The assessor of the outcomes should typically be blinded to treatment allocation (impractical in surgical interventions, and so forth), among other essential criteria. Randomized studies that focused on investigating the comparative effectiveness and/or

safety of two or more techniques/interventions were considered matched comparative studies.[21,22]

(c) Matched comparative studies: As above, but with two treatment/intervention groups typically lacking a control group. The patients are typically allocated to either of the intervention groups through randomization in studies conducted prospectively and through a bias-free method for allocating patients in studies conducted retrospectively. Matched comparative studies aim to study comparative effectiveness, outcome responsiveness, and/or safety for at least two treatment or intervention modalities.

(d) Classic interventional clinical trial: Any clinical trial investigating the effect of a treatment/intervention (surgical or medical, etcetera) on specified health-related outcome measures but with a single arm, that is, neither a control group nor plausibly randomization. Trials can be prospective or retrospective.

(e) Controlled clinical trials: Any clinical trial investigating the effect of a treatment/intervention (surgical or medical and so forth) on specified health-related outcome measures with a *normal* control group.

(f) Causal language: We defined causal language as the use of wording or phrases that unjustifiably assert or strongly imply a causal relationship between an intervention and an effect or outcome.[5,23,24]

RESULTS

The final list of the 93 included studies in this critical analysis study is shown in Additional File 1. Demographics of included articles are shown in Table 2. The demographics of the included pediatric orthopedic subspecialties were as follows: Congenital and developmental disorders ($n = 25$) (27%), trauma ($n = 23$) (25%), spine ($n = 18$) (19%), hip, for example, slipped capital femoral epiphysis (SCFE), Perthes ($n = 9$) (10%), arthroscopy and sports injuries ($n = 8$) (9%), neuromuscular ($n = 5$) (5%), bone tumors ($n = 3$) (3%), and other ($n = 2$) (2%).

Our classification identified the presence of at least one count of bias in 17 (18%) studies, 14 (15%) of which had only one count of bias, while 3 (3%) studies had two counts of bias. Nine (10%) studies had bias in both full-text and the corresponding abstract conclusions. In 4 (4%) studies, bias was restricted to conclusions of abstracts only, and in 4 (4%) studies was restricted to a bias criterion assigned to the classification criteria of conclusions of full-text only and not the abstract raw data file. Twenty-seven spin/bias items or outcomes were reported per full-text and corresponding abstract for each of the 93 included studies. A total of 2511 spin/bias items were analyzed across 93 included studies. We reported 30 (1%) counts of bias in conclusions of full-texts and/or abstracts. The severity of bias/spin in conclusions was mild level ($n = 21$; 70%), moderate level ($n = 5$; 17%),

Table 2: Demographics of 93 included articles

Category	N (%)
Study design	
Prospective	37 (40)
Retrospective	53 (57)
Prospective and retrospective	1 (1)
Other e.g., cross sectional*	2 (2)
Study type	
Matched comparative study	79 (85)
Classic interventional clinical trial	11 (12)
Controlled clinical trial	1 (1)
Other	2 (2)
Type of control[a]	
Surgical	51/82 (62)
Conservative	24/82 (29.5)
Surgical and conservative	2/82 (2.5)
Medical	1/82 (1)
Educational, behavioral etc.	3/82 (4)
Normal	1/82 (1)
Type of intervention	
Surgical	66 (71)
Conservative	18 (19.5)
Surgical and conservative	4 (4.5)
Medical	1 (1)
Medical and surgical	1 (1)
Educational, behavioral etc.	3 (3)
Sample size, total patients[b]	1,966,546[c]
Follow-up period range (Y)[b]	0-45
Conflict of interest declared	92 (99)
Conflict of interest present[d]	23/91 (25)[e]
Multicenter study	20 (21.5)

N, number of studies; * both studies included an intervention. However, they did not typically fit in any of the standard types of interventional studies. Hence they were classified under "other"; [a]not applicable in 11 studies per study design i.e., single-arm intervention (no control); [b]Study's raw data with explanatory comments; [c]a large nationwide epidemiologic retrospective study included 1,957,131 patients in the control group; [d]of one author at least; [e]one study declared conflicts of interest but the COI forms were inaccessible. The distribution of included articles across journals were (*N*= 93); Journal of pediatric orthopaedics (*n*=40), The bone joint journal (*n*=12), Journal of pediatric orthopaedics Part B (*n*=9), The Journal of bone and joint surgery American volume (*n*=9), International orthopaedics (*n*=8), Clinical orthopaedics and related research (*n*=8), Knee surgery sports traumatology arthroscopy (*n*=6), The journal of the American academy of orthopaedic surgeons (*n*=1).

and high level ($n = 4$; 13%). Characteristics of reporting and interpretation bias in conclusions of included full-text articles and their abstracts are shown in Table 3. Selective citation of other studies under the category of misleading reporting in Table 3 was not reported, as it was beyond the scope of the present study. The study's raw data with explanatory comments are shown in raw data file. A graphical abstract of discrepancies between results and conclusions in interventional pediatric orthopedic research is shown in Figure 2.

Table 3: Reporting and interpretation bias of conclusions reported in 93 full-text articles and their abstracts (Lazarus *et al.*, 2015).

Type of spin or interpretation bias	Full-texts[a]	Abstracts[a]
Misleading reporting		
Not reporting adverse events or lack of focus on harm	0	0
Selective reporting	0	0
Misleading description of study design	0	1
Use of linguistic spin	1 (1)	1 (1)
No consideration of the limitations	5 (5)	Not applicable
Inadequate interpretation		
Claim an effect for non-statistically significant results	0	0
Claim an equivalence for non-statistically significant results despite a wide confidence interval	0	0
Ruling out safety for non-statistically significant results	0	0
Causal language or causal claim	1 (1)	1 (1)
Claim of any significant difference despite lack of statistical test	0	0
Focus on statistical significance instead of clinical relevance	3 (3)	4 (4)
Inadequate extrapolation		
Inadequate extrapolation to a larger population, intervention or outcome	2 (2)	2 (2)
Inadequate implication for clinical practice	4 (4)	5 (5)
Other	0	0
Total spin counts=30[b]	16	14

[a]All data are presented as a number of counts and percentages of each spin/bias item across 93 included studies, [b]Total spin counts (*n*=30) represented (1%) of the total number of spin/bias items (*n*=2511) analyzed across 93 included studies. Total spin counts were distributed as follows: Misleading reporting 8/837 (1%), Inadequate interpretation 9/1116 (1%), Inadequate extrapolation 13/558 (2%)
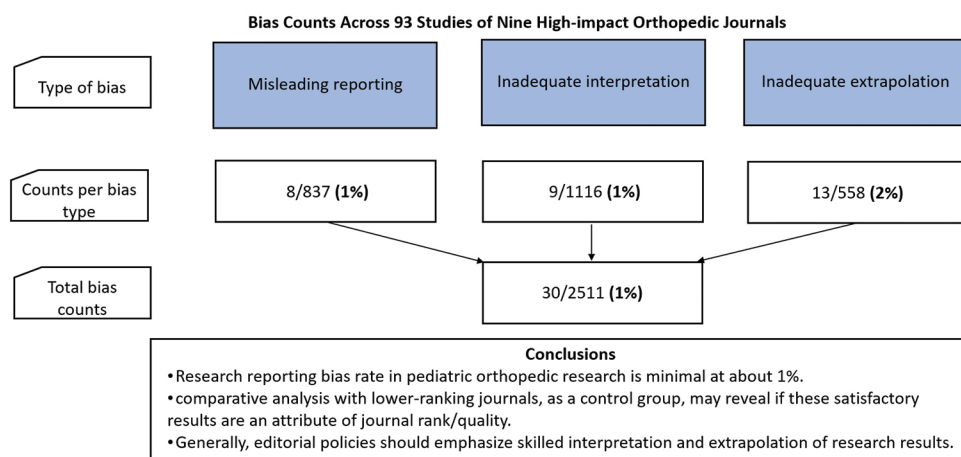

Figure 2: Graphical abstract. Discrepancies between results and conclusions in interventional pediatric orthopedic research.

DISCUSSION

Our study investigated the prevalence, subtypes, and severity of misleading reporting, inadequate interpretation and extrapolation of the conclusions in full-texts and corresponding abstracts of 93 pediatric orthopedic interventional studies published in high-impact factor orthopedic journals. Most studies (82%) were free of any form of bias/spin. Furthermore, most of the reported counts of bias (87%) were of low-to-moderate severity. These overall satisfactory outcomes may be ascribed to the high-impact society journals constituting our study sample. Contrastingly, higher degrees of research interpretation bias was recorded in SRs and RCTs, reporting on the treatment of orthopedic trauma,[12,25] arthroscopy and sports medicine,[13] and Achilles tendon injuries.[14] Although one of those studies targeted high-impact journals, interpretation bias was still unacceptably high.[12] Similarly, considerable research interpretation and extrapolation bias have been reported in high-impact journals of both medical[26] and surgical disciplines.[27]

This underscores the fact that research interpretation bias is prevalent across the medical literature, including high-impact journals.[26,27] Hypothetically, bias in research

reporting should be less prevalent in high-impact journals than in low-impact journals. Conclusions and recommendations from studies published in high-impact journals are usually the main guide for clinical practice decision-making, public health planning, and national and global health policies. However, the high journal impact factor is not necessarily associated with high reporting quality in published articles.[12,14,17,27-29]

The most common types of spin or interpretation bias reported in our study were; no consideration of the limitations, inadequate implication for clinical practice, and focus on statistical significance instead of clinical relevance. Misleading reporting, such as selective reporting and use of linguistic spin, and inadequate interpretation of statistically non-significant results, has been widely cited as an important cause of bias in conclusions of diverse medical and surgical studies as; orthopedic surgery,[12,25] bariatric surgery,[30] urology,[31] cardiology,[32] and oncology,[33] among others.[27,34] This contrasts with our study, where the degree of spin was generally low, with no reported spin in relation to the "claim of an effect or equivalence for non-statistically significant results." Statistical significance was more likely to be reported in abstracts of surgical randomized trials than their corresponding full-texts.[11] Besides, noticeable inconsistencies between the reporting quality of abstracts and corresponding full-text articles were found in pediatric orthopedic and physical rehabilitation publications.[35,36] Our study found negligible differences between spin counts reported in abstracts and full-texts.

Using causal language, which unjustifiably and clearly asserts or strongly implies a cause-effect relationship between an intervention and an outcome, can mislead readers and misguide clinical practice recommendations and health policies. However, occasionally, there is no clear line of demarcation between causal and non-causal language, namely, the presence of mere association or correlation between an intervention and an outcome.[23,37] In addition, it may be challenging to judge if the authors' intent in the conclusion was to convey causation or association.[23,24] The fact that a considerable percentage of biomedical authors are non-native English speakers adds to the challenge. In our study, the use rate of unjustifiable causal language was extremely low. Nonetheless, the unjustifiable use of causal language remains problematic in observational research across the biomedical literature,[23,24,37,38] including orthopedics.[39,40]

Pre-publication peer review is considered a means of achieving scientific quality control. It can potentially protect against misinterpreted, over-interpreted, and biased conclusions.[41] However, the peer review system itself is not without deficiencies. Difficulties in standardizing quality assessment measures for peer review reports,[42] insufficient peer review training programs,[43] shortage of qualified peer

reviewers,[44] and peer review bias[45,46] have all been cited as potential causes for deficiencies of the peer review system.

Suggested strategies to improve research reporting bias and misinterpretation/extrapolation of conclusions include (a) integrating peer review – critical appraisal skills – and research writing skills into under- and postgraduate medical curricula, (b) incentivizing and officially recognizing extra-curricular learning activities aimed at improving peer review and research writing skills of early career researchers, for example, courses, workshops, co-reviewing, and (c) raising awareness of research reporting bias among editors and peer reviewers, for example, journal instructions to authors and reviewers.

Strengths and limitations

The strengths of our study were the large and reasonably representative sample size, namely, 93 pediatric orthopedic studies and the use of established classification criteria for bias/spin. However, we acknowledge limitations. Classically, the SR process follows a dual or multiple reviewer/coder approach across all phases, for example, data identification, screening, and extraction/quality assessment. This is aimed at ensuring methodological rigor and validity of the findings. However, some studies have suggested that SRs could be conducted with a modified or parallelized approach, where coders work simultaneously – instead of sequentially – toward achieving study objectives, for example, a single coder may be assigned to a particular SR phase. This is provided the single coder was experienced in the clinical subject matter of the SR.[47] Waffenschmidt *et al.*, in 2019, found no or negligible impact of missing studies on the findings of meta-analyses when an experienced single coder conducted the screening. Contrastingly, missing studies would have changed the findings substantially if a less experienced single coder conducted the screening.[47] In our study, two authors are pediatric orthopedic surgeons, the study's subject matter. T.A.E. is a senior faculty at the university's orthopedic department and primarily focuses on pediatric orthopedic clinical practice/research. This senior author was principally in charge of spin/bias coding (data extraction, analysis) and revision of the whole study sample. In contrast, the mid-career authors were independently assigned equal numbers of the study sample. The total number of spin counts was insufficiently low to allow for meaningful subgroup analytic statistics with respect to research types, journal impact factors, and conflict of interest.[14,26] The severity of spin was categorical and somewhat subjective. We excluded classic placebo – no treatment – RCTs investigating the effects of treatment/intervention. However, the included study types – interventional single-armed and matched comparative studies – contribute significantly to the published pediatric orthopedic literature and remain a valuable source of clinical practice guidance.[48] This may be attributed to the impractical prerequisites and ethical restraints of conducting classic RCTs in pediatric orthopedic

surgical interventions.[49] Another limitation of our study was that we did not include non-experimental (observational) pediatric orthopedic research. This was beyond the scope of our study.

CONCLUSION

There was a low prevalence of reporting bias, namely, misleading reporting, misinterpretation, and inadequate extrapolation of conclusions, in 93 full-texts and their abstracts of interventional pediatric orthopedic studies published in high-ranking journals. A comparative study with lower-ranking journals, as a control group may reveal if our favorable results are an attribute of journal rank/quality. We recommend that authors, reviewers, and editors check the results section to ensure that – numerical – data are neither misreported nor misinterpreted and are adequately extrapolated in the conclusions, preferably in accordance with relevant reporting guidelines.

AVAILABILITY OF DATA AND MATERIALS

All data generated or analyzed during this study are included in this published article and its additional information files.

AUTHORS' CONTRIBUTIONS

All authors have contributed significantly. TAE conceptualized and designed the study. TAE, JA, and SAK contributed to data acquisition analysis and interpretation. TAE wrote the manuscript draft. All authors revised the manuscript critically for important intellectual content and approved the version to be published. All authors agree to be accountable for all aspects of the work to ensure that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved. All authors have critically reviewed and approved the final draft and are responsible for the manuscript's content and similarity index.

ETHICAL APPROVAL

The Institutional Review Board approval is not required.

DECLARATION OF PATIENT CONSENT

Patient's consent not required as there are no patients in this study.

USE OF ARTIFICIAL INTELLIGENCE (AI) ASSISTED TECHNOLOGY FOR MANUSCRIPT PREPARATION

The authors confirm that there was no use of artificial intelligence (AI)-assisted technology for assisting in the writing or editing of the manuscript and no images were manipulated using the AI.

CONFLICTS OF INTEREST

Tamer A. El-Sobky is a reviewer for the Journal of Children's Orthopedics, Journal of Pediatric Orthopedics Part B, and Bone and Joint Journal. John Amen and Sherif Ahmed Kamel have nothing to declare.

FINANCIAL SUPPORT AND SPONSORSHIP

This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

SUPPLEMENTARY INFORMATION

• Additional File 1: The literature search and identification process and the spin classification system, with examples used to assess included studies.

• Raw Data File: The study's raw data are presented with explanatory comments.

REFERENCES

1. Buffalari D, Fernandes JJ, Chase L, Lom B, McMurray MS, Morrison ME, *et al.* Integrating research into the undergraduate curriculum: 1. Early research experiences and training. J Undergrad Neurosci Educ 2020;19:A52-63.
2. El-Sobky TA. An author's guide to mastering academic writing skills: Discussion of a medical manuscript. J Musculoskelet Surg Res 2021;5:227-34.
3. Warden G. Definitions of bias in clinical research. Methods Mol Biol 2021;2249:35-52.
4. Yavchitz A, Ravaud P, Altman DG, Moher D, Hrobjartsson A, Lasserson T, *et al*. A new classification of spin in systematic reviews and meta-analyses was developed and ranked according to the severity. J Clin Epidemiol 2016;75:56-65.
5. Lazarus C, Haneef R, Ravaud P, Boutron I. Classification and prevalence of spin in abstracts of non-randomized studies evaluating an intervention. BMC Med Res Methodol 2015;15:85.
6. van der Steen JT, Ter Riet G, van den Bogert CA, Bouter LM. Causes of reporting bias: A theoretical framework. F1000Res 2019;8:280.
7. Riediger ND, Bombak AE, Mudryj A, Bensley J, Ankomah S. A systematic search and qualitative review of reporting bias of lifestyle interventions in randomized controlled trials of diabetes prevention and management. Nutr J 2018;17:83.
8. Lundh A, Lexchin J, Mintzes B, Schroll JB, Bero L. Industry sponsorship and research outcome. Cochrane Database Syst Rev 2017;2:MR000033.
9. Berner D, Amrhein V. Why and how we should join the shift from significance testing to estimation. J Evol Biol 2022;35:777-87.
10. Greenland S, Senn SJ, Rothman KJ, Carlin JB, Poole C, Goodman SN, *et al*. Statistical tests, P values, confidence intervals, and power: A guide to misinterpretations. Eur J Epidemiol 2016;31:337-50.
11. Assem Y, Adie S, Tang J, Harris IA. The over-representation of significant p values in abstracts compared to corresponding

full texts: A systematic review of surgical randomized trials. Contemp Clin Trials Commun 2017;7:194-9.

12. Shepard S, Checketts J, Eash C, Austin J, Arthur W, Wayant C, *et al.* Evaluation of spin in the abstracts of orthopedic trauma literature: A cross-sectional review. Injury 2021;52:1709-14.

13. Chellamuthu G, Muthu S, Damodaran UK, Rangabashyam R. "Only 50% of randomized trials have high level of confidence in arthroscopy and sports medicine"-a spin-based assessment. Knee Surg Sports Traumatol Arthrosc 2021;29:2789-98.

14. Carr M, Dye D, Arthur W, Ottwell R, Detweiler B, Stotler W, *et al.* Evaluation of spin in the abstracts of systematic reviews and meta-analyses covering treatments for Achilles tendon ruptures. Foot Ankle Orthop 2021;6.

15. Yavchitz A, Boutron I, Bafeta A, Marroun I, Charles P, Mantz J, *et al.* Misrepresentation of randomized controlled trials in press releases and news coverage: A cohort study. PLoS Med 2012;9:e1001308.

16. Pavlovic V, Weissgerber T, Stanisavljevic D, Pekmezovic T, Milicevic O, Lazovic JM, *et al.* How accurate are citations of frequently cited papers in biomedical literature? Clin Sci (Lond) 2021;135:671-81.

17. Gazendam A, Cohen D, Morgan S, Ekhtiari S, Ghert M. Quotation errors in high-impact-factor orthopaedic and sports medicine journals. JB JS Open Access 2021;6.

18. Duyx B, Urlings MJ, Swaen GM, Bouter LM, Zeegers MP. Scientific citations favor positive results: A systematic review and meta-analysis. J Clin Epidemiol 2017;88:92-101.

19. Ghannad M, Yang B, Leeflang M, Aldcroft A, Bossuyt PM, Schroter S, *et al.* A randomized trial of an editorial intervention to reduce spin in the abstract's conclusion of manuscripts showed no significant effect. J Clin Epidemiol 2021;130:69-77.

20. Won J, Kim S, Bae I, Lee H. Trial registration as a safeguard against outcome reporting bias and spin? A case study of randomized controlled trials of acupuncture. PLoS One 2019;14:e0223305.

21. Rouse DJ, Hirtz DG, Thom E, Varner MW, Spong CY, Mercer BM, *et al.* A randomized, controlled trial of magnesium sulfate for the prevention of cerebral palsy. N Engl J Med 2008;359:895-905.

22. Kendall JM. Designing a research project: Randomised controlled trials and their principles. Emerg Med J 2003;20:164-8.

23. Haber NA, Wieten SE, Rohrer JM, Arah OA, Tennant PW, Stuart EA, *et al.* Causal and associational language in observational health research: A systematic evaluation. Am J Epidemiol 2022;191:2084-97.

24. Han MA, Leung G, Storman D, Xiao Y, Srivastava A, Talukdar JR, *et al.* Causal language use in systematic reviews of observational studies is often inconsistent with intent: A systematic survey. J Clin Epidemiol 2022;148:65-73.

25. Jones C, Rulon Z, Arthur W, Ottwell R, Checketts J, Detweiler B, *et al.* Evaluation of spin in the abstracts of systematic reviews and meta-analyses related to the treatment of proximal humerus fractures. J Shoulder Elbow Surg 2021;30:2197-205.

26. Chow R, Huang E, Fu S, Kim E, Li S, Sodhi J, *et al.* Spin in randomized controlled trials in obstetrics and gynecology: A systematic review. Womens Health Rep (New Rochelle) 2022;3:795-802.

27. Arunachalam L, Hunter IA, Killeen S. Reporting of

randomized controlled trials with statistically nonsignificant primary outcomes published in high-impact surgical journals. Ann Surg 2017;265:1141-5.

28. Nascimento DP, Gonzalez GZ, Araujo AC, Costa LO. Journal impact factor is associated with PRISMA endorsement, but not with the methodological quality of low back pain systematic reviews: A methodological review. Eur Spine J 2020;29:462-79.

29. Riado Minguez D, Kowalski M, Vallve Odena M, Longin Pontzen D, Jelicic Kadic A, Jeric M, *et al.* Methodological and reporting quality of systematic reviews published in the highest ranking journals in the field of pain. Anesth Analg 2017;125:1348-54.

30. Rassy N, Rives-Lange C, Carette C, Barsamian C, Moszkowicz D, Thereaux J, *et al.* Spin occurs in bariatric surgery randomized controlled trials with a statistically nonsignificant primary outcome: A systematic review. J Clin Epidemiol 2021;139:87-95.

31. Wu J, Ho W, Klotz L, Yuan M, Lee JY, Krakowsky Y. Assessing "Spin" in urology randomized controlled trials with statistically nonsignificant primary outcomes. J Urol 2023;209:494-503.

32. Khan MS, Lateef N, Siddiqi TJ, Rehman KA, Alnaimat S, Khan SU, *et al.* Level and prevalence of spin in published cardiovascular randomized clinical trial reports with statistically nonsignificant primary outcomes: A systematic review. JAMA Netw Open 2019;2:e192622.

33. Wayant C, Margalski D, Vaughn K, Vassar M. Evaluation of spin in oncology clinical trials. Crit Rev Oncol Hematol 2019;144:102821.

34. Otte WM, Vinkers CH, Habets PC, van IJzendoorn DG, Tijdink JK. Analysis of 567,758 randomized controlled trials published over 30 years reveals trends in phrases used to discuss results that do not reach statistical significance. PLoS Biol 2022;20:e3001562.

35. Kamel SA, El-Sobky TA. Reporting quality of abstracts and inconsistencies with full text articles in pediatric orthopedic publications. Res Integr Peer Rev 2023;8:11.

36. Nascimento DP, Costa LO, Gonzalez GZ, Maher CG, Moseley AM. Abstracts of low back pain trials are poorly reported, contain spin of information and are inconsistent with the full text: An overview study. Arch Phys Med Rehabil 2019;1976-85.e18.

37. Olarte Parra C, Bertizzolo L, Schroter S, Dechartres A, Goetghebeur E. Consistency of causal claims in observational studies: A review of papers published in a general medical journal. BMJ Open 2021;11:e043339.

38. Chiu K, Grundy Q, Bero L. 'Spin' in published biomedical literature: A methodological systematic review. PLoS Biol 2017;15:e2002173.

39. Barakat N, Novicoff WM, Werner BC, Browne JA. High prevalence of causal language and inferences in observational hip and knee arthroplasty database studies: A review of papers published across four orthopaedic journals. J Arthroplasty 2023;38:945-9.

40. Gleadhill C, Lee H, Kamper SJ, Cashin A, Hansford H, Traeger AC, *et al.* Mixed messages: Most spinal pain and osteoarthritis observational research is unclear or misaligned. J Clin Epidemiol 2023;155:39-47.

41. Barroga E. Innovative strategies for peer review. J Korean Med

Sci 2020;35:e138.

42. Superchi C, González JA, Solà I, Cobo E, Hren D, Boutron I. Tools used to assess the quality of peer review reports: A methodological systematic review. BMC Med Res Methodol 2019;19:48.

43. Galipeau J, Moher D, Campbell C, Hendry P, Cameron DW, Palepu A, *et al*. A systematic review highlights a knowledge gap regarding the effectiveness of health-related training programs in journalology. J Clin Epidemiol 2015;68:257-65.

44. Fernandez-Llimos F. Pharmacy Practice 2018 peer reviewers. Peer review and publication delay. Pharm Pract (Granada) 2019;17:1502.

45. Pier EL, Brauer M, Filut A, Kaatz A, Raclaw J, Nathan MJ, *et al*. Low agreement among reviewers evaluating the same NIH grant applications. Proc Natl Acad Sci U S A 2018;115:2952-7.

46. Manchikanti L, Kaye AD, Boswell MV, Hirsch JA. Medical journal peer review: Process and bias. Pain Physician 2015;18:E1-14.

47. Waffenschmidt S, Knelangen M, Sieben W, Bühn S, Pieper D. Single screening versus conventional double screening for study selection in systematic reviews: A methodological systematic review. BMC Med Res Methodol 2019;19:132.

48. Luksameearunothai K, Chaudhry Y, Thamyongkit S, Jia X, Hasenboehler EA. Assessing the level of evidence in the orthopaedic literature, 2013-2018: A review of 3449 articles in leading orthopaedic journals. Patient Saf Surg 2020;14:23.

49. El-Sobky TA, El-Sayed M. Situational ethics of study designs for orthopaedic surgery interventions in children with cerebral palsy. Curr Orthop Pract 2019;30:178-80.