Caries Research

Review Article

Caries Res
DOI: 10.1159/000531569

Received: June 28, 2022
Accepted: June 10, 2023
Published online: June ▪▪▪, 2023

Assessment of *Spin* in the Abstracts of Randomized Controlled Trials in Dental Caries with Statistically Nonsignificant Results for Primary Outcomes: A Methodological Study

Naichuan Su[a] Michiel W. van der Linden[a] Clovis M. Faggion Jr.[b] Geert J.M.G. van der Heijden[a]

[a]Department of Oral Public Health, Academic Centre for Dentistry Amsterdam (ACTA), University of Amsterdam and Vrije Universiteit Amsterdam, Amsterdam, The Netherlands; [b]Department of Periodontology and Operative Dentistry, Faculty of Dentistry, University Hospital Münster, Münster, Germany

Keywords
Spin · Dental caries · Randomized controlled trials · Misinterpretation · Misrepresentation

Abstract
The study aimed to assess the prevalence of *spin* in the titles and abstracts of randomized controlled trials (RCTs) in dental caries with statistically nonsignificant primary outcomes and to assess the risk indicators which may be associated with *spin*. Any original publication reporting a two-arm RCT in dental caries with clearly identified statistically nonsignificant primary outcomes published from January 1, 2015, until October 28, 2022, were included. PubMed was searched electronically to identify the eligible publications. The prevalence of *spin* in titles and abstracts were assessed and categorized into *spin* patterns based on a predetermined classification scheme. The association between *spin* and the potential risk indicators at study, author, journal, institutional, and national levels was assessed. A total of 234 eligible RCT publications were included. The prevalence of *spin* in the titles and abstracts was 3% (95% confidence interval [CI]: 2–6%) and 79% (95% CI: 74–84%), respectively. The most common *spin* patterns in the results and conclusion sections, respectively, were results focusing on statistically significant within-group comparisons (23%), and conclusions focusing only on statistically significant results without acknowledgment of statistically nonsignificant results for the primary outcomes (26%). The *spin* was significantly associated with number of study centers (single-center vs. multicenter) (OR = 2.131; 95% CI: 1.092–4.158; $p = 0.03$), trial designs (nonparallel designs vs. parallel designs) (OR = 0.395; 95% CI: 0.193–0.810; $p = 0.01$), and overall H index of institutions for last authors (OR = 0.998; 95% CI: 0.996–0.999; $p < 0.01$), while it was not significantly associated with the other indicators. In the RCT publications with statistically nonsignificant results for primary outcomes in dental caries, the prevalence of *spin* may be low in the titles but high in the abstracts. Single-center studies with parallel designs and a lower overall H index of institutions for last authors may be more likely to have *spin* in the abstracts.

karger@karger.com
www.karger.com/cre

Correspondence to:
Naichuan Su, n.su@acta.nl

Introduction

Spin originally was defined as polishing the truth and using dishonest, deceptive, and manipulative strategies to provide a certain interpretation of information in order to sway opinion in a desired way [Safire, 1996]. It has been first introduced in biomedicine since 2007 [Hróbjartsson and Gøtzsche, 2007] and systematically assessed and classified in biomedicine since 2010 [Boutron et al., 2010]. In medicine, *spin* is defined as reporting practices that distort the interpretation of research and create misleading conclusions intentionally or unintentionally by suggesting more favorable results [Chiu et al., 2017]. In randomized controlled trials (RCTs), *spin* is specifically defined as the use of specific reporting strategies to make the treatment, intentionally or unintentionally, more beneficial to readers with no consideration of the statistically nonsignificant difference for the primary outcome, or to distract the readers from statistically nonsignificant results [Boutron et al., 2010].

Spin is common in both abstracts and main texts of biomedical publications [Chiu et al., 2017]. The overall prevalence of *spin* in the abstracts of biomedical publications is around 57%, and the prevalence in the abstracts of RCTs with nonsignificant primary outcomes was even reported to be 61% based on a meta-research study [Chiu et al., 2017]. *Spin* has a negative impact on advancing healthcare practice and population health, adversely influences health policy planning, adds to research waste, decreases the reproducibility of research, hampers the progress of science, and reduces return-of-investment from research [Ochodo et al., 2013; Caulfield and Ogbogu, 2015; Chiu et al., 2017].

RCTs play an important role in providing evidence for clinical decision-making and are commonly used for shaping clinical practice guidelines [Cooper et al., 2019]. This is because in general RCTs provide the strongest evidence with regard to the effectiveness of treatments [Hackshaw et al., 2006]. Abstracts have been shown to directly influence readers' opinions on treatments or interventions [Khan et al., 2019]. Many clinicians have no access to full text and use only the abstract of a trial to inform their clinical decisions [Hopewell et al., 2008]. Therefore, an accurate presentation of the results in the abstracts of RCTs is essential for the interpretation and dissemination of the results and their application in clinical practice.

Several methodological studies have assessed *spin* in the titles and abstracts of RCTs in medicine, including dentistry [Boutron et al., 2010; Lockyer et al., 2013; Arunachalam et al., 2017; Khan et al., 2019; Eleftheriadi et al., 2020; Wu et al., 2020; Guo et al., 2021; Ito et al., 2021; Rassy et al., 2021]. The prevalence of *spin* in those reviews was high, ranging from 34% to 70%. In addition, several studies assessed risk indicators for the presence of *spin* [Arunachalam et al., 2017; Khan et al., 2019; Roszhart et al., 2020; Wu et al., 2020; Guo et al., 2021; Ito et al., 2021]. However, most of those studies focused on a limited number of risk indicators, notably funding status, number of study centers, international collaboration, citations of the studies, and impact factors of the journals. To our best knowledge, however, no studies have comprehensively assessed the risk indicators of *spin* at all possible levels including study, author, journal, institutional, and national levels.

Dental caries is the most common dental disease with a global prevalence of 35% in 2010 [Vos et al., 2012]. A large proportion of all dental research concerns dental caries. However, to date, no attention has been given to the assessment of *spin* in the abstracts of RCTs in dental caries. Therefore, the aim of the present study was to assess the prevalence of *spin* in the titles and abstracts of RCTs in dental caries with statistically nonsignificant primary outcomes and to investigate the risk indicators at study, author, journal, institutional, and national levels which may be associated with *spin*.

Methods

Search Strategy

Two separate search strategies were used to search for the relevant publications. In the first search strategy, we included Medical Subject Headings (Mesh) terms relevant to "dental caries" and "randomized controlled trials" and their synonyms as the free words (search strategy 1) to search for articles in PubMed without language restriction, on March 16, 2020. An update of the search was performed on October 28, 2022. In the second search strategy, we used the Cochrane Highly Sensitivity Search Strategy for identifying randomized trials in Medline: sensitivity- and precision-maximizing version [Lefebvre et al., 2011] (search strategy 2) to search for the relevant studies via PubMed on October 28, 2022. Both search strategies were limited to publications dated between 2015 and 2022. The search strategies are presented in online supplementary Table 1 (for all online suppl. material, see https://doi.org/10.1159/000531569).

Eligibility Criteria

Publications which met the following criteria were included:

1. They were any type of RCTs including parallel, crossover, or spit-mouth designs, with two arms only (one intervention group and one control group);
2. At least one of the participant, intervention, control, and outcomes components was relevant to dental caries. For example, if the target population of a study is patients with any type of caries, this study is considered to be eligible. If one of the

outcomes of a study is relevant to caries, such as DMFT and caries incidence, this study is also considered to be eligible;
3. The primary outcomes can be identified from the publications. Whenever the primary outcomes were not explicitly reported, the outcomes used for the sample size estimation were considered the primary outcomes. If outcomes were not stated in the sample size estimation, the primary outcomes were taken from the primary study objectives [Boutron et al., 2010];
4. The primary outcomes were statistically nonsignificant. If there was more than one primary outcome in a publication, at least one primary outcome was statistically nonsignificant;
5. The date of the publications published in print was later than January 1, 2015;
6. They were written in English.

Publications which met the following criteria were excluded:
1. Reporting equivalence or non-inferiority trials;
2. Duplicate publications. That is, if the exact same RCT study was published in more than one publication, only the earliest published study was included. However, if an RCT study was published in different publications but the reported outcomes or follow-up times in those publications were different, none of the publications was excluded.

Three reviewers independently assessed the titles, abstracts, and full texts of the identified publications (namely, N.S. and M.d.B. for publications published from January 1, 2015, to December 31, 2016, and N.S. and B.I. for publications published from January 1, 2017 to March 16, 2020). Since the present review only focused on the *spin* of the abstracts, full texts were not assessed if clear decisions for inclusion and exclusion could be made based on the titles and abstracts. Full texts were only obtained for further assessment if a clear decision from the titles and abstracts alone could not be made. Any disagreement between the reviewers was resolved by consensus discussion.

Data Extraction

Spin Patterns

The *spin* patterns were defined as the specific strategies of *spin* that authors used in the titles, and results and conclusion section of abstracts [Boutron et al., 2010]. The common *spin* patterns in titles and abstracts of RCTs are presented in Table 1. The *spin* patterns were identified and extracted from all the included publications. The classification scheme of *spin* patterns for RCTs used in the present study was developed by Boutron et al. [2010].

Extent (i.e., the Amount) of *Spin*

The classification method for the extent of *spin* was developed by Boutron et al. [2010]. The extent of *spin* was defined as the number of sections with *spin* in the abstracts [Boutron et al., 2010]. It focused on the amount of *spin* in the abstracts and was determined using four hierarchical categories: (1) no *spin* in either the results or conclusion section; (2) *spin* in the results section only; (3) *spin* in the conclusion section only; and (4) *spin* in both the results and conclusion sections [Boutron et al., 2010]. If more sections of abstracts had *spin* and if *spin* was present in the conclusion section, the extent of *spin* was considered to be higher.

Level (i.e., the Impact) of *Spin* in Conclusions

The classification method for the level of *spin* in conclusions was developed by Boutron et al. [2010]. The level of *spin* was defined as the extent to which the conclusion sections of the abstracts are misleading. It focused on the impact of *spin* on the conclusions of abstracts and was determined by the following indicators: (1) uncertainty in the framing of the conclusions; (2) recommendation for further trials; (3) acknowledgment of the statistically nonsignificant results for the primary outcomes; and (4) recommendation to use the treatment in clinical practice [Boutron et al., 2010]. The level of *spin* was classified into none, low, moderate, and high levels based on the four indicators. The classification scheme for the level of *spin* is presented in online supplementary Table 2. A high level of *spin* indicated that the conclusion of the abstract may be strongly misleading. A high level of *spin* was defined as no uncertainty in the framing of conclusions, no recommendations for further trials, and no acknowledgment of the statistically nonsignificant results for primary outcomes, or as the presence of recommendations to use the treatment in clinical practice. A moderate level of *spin* indicated that the conclusion of the abstract may be questionable and moderately misleading. A moderate level of *spin* was defined as uncertainty in the framing of conclusions or recommendations for further trials but no acknowledgment of the statistically nonsignificant results for the primary outcomes. A low level of *spin* indicated that the conclusion of the abstract may be slightly misleading. A low level of *spin* was defined as uncertainty in the framing of conclusions and recommendations for further trials or acknowledgment of the statistically nonsignificant results for the primary outcome [Boutron et al., 2010].

Risk Indicators Associated with *Spin*

The potential risk indicators that may be associated with *spin* were extracted from the included publications. The factors included five levels, i.e., study, author, journal, institutional, and national levels. The potential risk indicators were decided a priori by the team based on previous literature and our own assumptions.

The risk indicators at *study level* included the study funding, number of study centers where the studies were performed, structured abstracts, the reporting quality of the abstracts, involvement of a statistician or methodologist, and the trial design. A structured abstract was defined as an abstract with distinct and labeled sections (e.g., introduction, methods, results, and conclusions) for rapid comprehension [Sollaci and Pereira, 2004]. The reporting quality of the abstracts was assessed with the Consolidated Standards of Reporting Trials (CONSORT) checklist for the abstracts of RCTs with 17 items [Hopewell et al., 2008]. If an abstract reported a higher number of items on the CONSORT checklist, the reporting quality of the abstract was considered to be higher. The involvement of a statistician or methodologist was decided based on the affiliation information within the studies [Makou et al., 2021] and the acknowledgment section of the studies. An affiliation with a department/unit of biostatistics, epidemiology, mathematics, and research methodology was considered indicative of such involvement. The trial design of RCTs was dichotomized into parallel trials and nonparallel trials. Nonparallel trials included cross-over trials, split-mouth trials, and cluster trials.

The risk indicators at *author level* included the number of authors, conflict of interest among authors, international collaborations, H index of the first author, and H index of the last author. International collaborations were defined that the authors of a study came from at least two different countries. If an author had

Table 1. Prevalence of *spin* and *spin* patterns in the titles and abstracts of the included publications (N = 234)

Spin patterns	Number of publications with the presence of the *spin* pattern	Prevalence (95% CI)
Titles	**8**	**3% (2%, 6%)**
Title claims treatment effectiveness when none exists	1	0% (0%, 2%)
Title claims treatment safety	0	0% (0%, 1%)
Title focuses only on the statistically significant outcomes*	7	3% (1%, 6%)
Abstracts	**186**	**79% (74%, 84%)**
Results section of the abstracts	**104**	**44% (38%, 51%)**
Results focus on statistically significant primary outcomes and ignore the statistically nonsignificant primary outcomes*	25	11% (7%, 15%)
Results focus on statistically significant secondary outcomes	23	10% (6%, 14%)
Results focus on statistically significant within-group comparisons	53	23% (18%, 28%)
Results focus on statistically significant subgroup analysis	3	1% (0%, 3%)
Results focus on statistically significant modified population of analyses (e.g., statistically significant per-protocol analysis instead of nonstatistically intention-to-treat analysis)	1	0% (0%, 2%)
Trend toward significance or equivalent verbiage in results	18	8% (5%, 12%)
Other *spin* patterns	1	0% (0%, 2%)
Conclusion section of the abstracts	**169**	**72% (66%, 78%)**
Conclusions claim equivalence for statistically nonsignificant results	43	18% (14%, 24%)
Conclusions claim efficacy with no consideration of the statistically nonsignificant results for primary outcomes	46	20% (15%, 25%)
Conclusions focus only on statistically significant results (e.g., other significant primary outcomes, secondary outcomes, subgroup analyses, within-group comparisons, or modified population of analyses) without acknowledgment of statistically nonsignificant results for the primary outcomes	60	26% (20%, 32%)
Conclusions acknowledge statistically nonsignificant results for the primary outcomes but with emphasis on the beneficial treatment effect	15	6% (4%, 10%)
Conclusions acknowledge statistically nonsignificant results for the primary outcomes but emphasis in on other statistically significant results (e.g., other significant primary outcomes, secondary outcomes, subgroup analyses, within-group comparisons, or modified population of analyses)	14	6% (3%, 10%)
Trend toward significance or equivalent verbiage in conclusions	10	4% (2%, 7%)
Conclusions recommend to use the treatment	20	9% (5%, 13%)
Other *spin* patterns	15	6% (4%, 10%)

multiple affiliations from different countries, the first affiliation was considered the affiliation of the author when determining the international collaborations. H index is a quantitative metric for assessing the cumulative impact of a researcher's academic output and performance [Hirsch, 2005]. The H index is defined as the number of publications of a researcher which have been cited at least that same number of times [Hirsch, 2005]. A higher H index indicated that the researcher is more productive and influential. The H index can also be used to measure the research performance of a journal, an institution, or a country. The H index of an author was estimated based on the Author Search^{BETA} of the Web of Science Core Collection (WOSCC) database up to June 2021. Similar names were differentiated on the basis of the affiliation histories and research subjects.

The risk indicators at *journal level* included the 4-year impact factor, the H index of the journals, and the word limit of the abstracts of the journals. Both the 4-year impact factor in the publication year of the studies in 2020 and the H index of the journals were collected from the SCImago Journal & Country Rank website in June 2021, where the impact factor and H index were originally based on the information contained in the SCOPUS database. The word limit of the abstracts was collected from the instructions to authors of the journals.

The risk indicators at *institutional level* included the overall H indexes of the institutions of both first and last authors and the H indexes of the institutions in dental research of both first and last authors. If an author had multiple affiliations, the first affiliation was collected. The H index of the institutions was collected based on the SCOPUS affiliation identifier up to July 2021.

The risk indicators at *national level* included the overall H indexes of the countries of both first and last authors, H indexes of the countries in dental research of both first and last authors, and

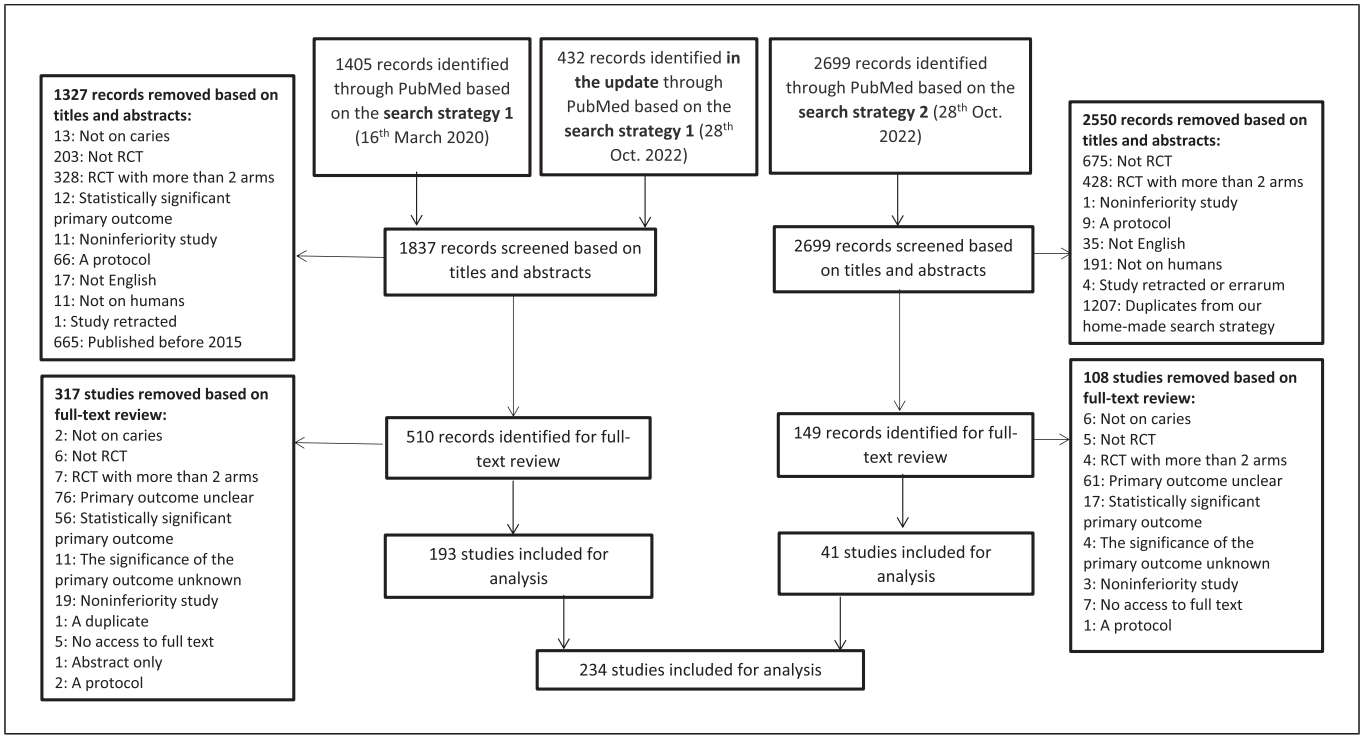

Fig. 1. Flow diagram of study inclusion (search strategy 1 used Medical Subject Headings [Mesh] terms relevant to "dental caries" and "randomized controlled trials" and their synonyms as the free words; search strategy 2 is Cochrane Highly Sensitivity Search Strategy for identifying randomized trials in Medline: sensitivity- and precision-maximizing version [2008 version]).

the economical level of the countries. The H indexes of the countries in 2020 were collected based on the SCImago Journal & Country website. The economical level of the countries was determined based on the criteria of the World Bank for the 2021 fiscal year which relied on the gross national income per capita of a country [The World Bank, 2021].

Two reviewers (N.S. and M.W.L.) independently assessed *spin* of the abstracts of all the included publications. The initial inter-rater agreement on the identification of each *spin* pattern was assessed with a kappa test. Three reviewers (N.S., M.B., and M.d.B.) extracted the factors associated with *spin* for the studies identified in the original search. Any disagreement between the reviewers was resolved by consensus discussion.

Statistical Analysis

The overall prevalence of *spin* in titles and abstracts, the prevalence of each *spin* pattern, the extent of *spin*, and the level of *spin* in conclusions of the abstracts were all presented with percentages and 95% confidence intervals (CIs). To assess potential risk indicators associated with *spin*, the extent of *spin* was regarded as the dependent variable and was dichotomized into high extent (*spin* in both results and conclusion sections) and low extent (no *spin* in either results or conclusion section, *spin* in results section only, or *spin* in conclusion section only). Univariate binary logistic regression analyses were then used to assess the association between the extent of *spin* and each potential risk indicator separately. Multicollinearity of the risk indicators which were significant based on the univariate analyses was tested using the variance inflation factor (VIF). When a VIF value of a risk indicator was higher than 10 [O'brien, 2007], collinearity was considered present and the risk indicator was excluded from the subsequent analysis. Subsequently, the multivariate binary logistic regression analysis with backward selection ($p > 0.05$ for removal) was performed to assess the association between the extent of *spin* and those factors with p values of <0.05 in the univariate analyses. SPSS 27.0 (IBM, New York, NY, USA) was used to perform both the descriptive and statistical analyses in the study.

Results

Based on search strategy 1, we identified a total of 1,405 records originally on 16th March 2021 and another 433 records on 28th October 2022 in the update. During the screening of the titles and abstracts, 510 of them satisfied the eligibility criteria. Eventually, 317 publications did not satisfy the eligibility criteria after carefully reading the full-text publications, and the other 193 publications were included. In search

Table 2. Extent of *spin* of the abstracts and level of *spin* in conclusions of the abstracts of the included publications (N = 234)

Extent of *spin*	Prevalence (95% CI)	No. of publications
No *spin* in either results or conclusion section	21% (16%, 26%)	48
Spin in results section only	7% (4%, 11%)	17
Spin in conclusion section only	35% (29%, 41%)	82
Spin in both results and conclusion sections	37% (31%, 44%)	87
Level of *spin*		
None	7% (4%, 11%)	16
Low	39% (33%, 46%)	92
Moderate	10% (6%, 14%)	23
High	44% (38%, 50%)	103

strategy 2 based on the Cochrane Highly Sensitivity Search Strategy, 2,699 records were identified and eventually, 41 publications were included. Therefore, a total of 234 publications were included in the present study (Fig. 1).

Most of the included studies (N = 200) were published in dental journals and were parallel RCTs (N = 175). The studies were performed in various countries such as Brazil (N = 37), India (N = 34), and the USA (N = 16). Online supplementary Table 3 presents the characteristics of the included publications.

Prevalence of Spin and Spin Patterns

The prevalence and 95% CI of each *spin* pattern in the 234 included publications are shown in Table 1. The inter-rater agreement on the identification of each *spin* pattern ranged from 76% to 100%, and the kappa values of each *spin* pattern ranged from 0.08 to 0.59 (online suppl. Table 4). The prevalence of *spin* in the titles was 3% (95% CI: 2–6%). The prevalence of *spin* in the abstracts was 79% (95% CI: 74–84%). The prevalence of *spin* in the results and conclusion sections of the abstracts was 44% (95% CI: 38–51%) and 72% (66–78%), respectively. The *spin* pattern with the highest prevalence in the results section was that results focused on statistically significant within-group comparisons (23%; 95% CI: 18–28%). In the conclusion section, the most prevalent *spin* pattern was that conclusions focused only on statistically significant results without acknowledgment of statistically nonsignificant results for the primary outcomes (26%; 95% CI: 20–32%).

Extent of Spin and Level of Spin in the Conclusions

The extent of *spin* in the abstracts is presented in Table 2. Out of the 234 abstracts, 186 (79%; 95% CI: 74–84%) had *spin* in at least one section of the abstracts. The level of *spin* in the conclusion sections of the 234 abstracts is also presented in Table 2. The majority of the publications had a high level of *spin* in the conclusion sections of the abstracts (44%; 95% CI: 38–50%).

Risk Indicators Associated with Spin

The characteristics of the potential risk indicators and their distribution over the extent of *spin* of the included abstracts are presented in online supplementary Table 5. In the univariate analysis, the extent of *spin* was significantly associated with the following risk indicators ($p < 0.05$): funding, number of study centers, reporting quality of abstracts, trial design, international collaborations, H index of the last authors, 4-year impact factor, H index of the journals, overall H index of institutions for the first authors and the last authors, H index of institutions in dental research for first authors and last authors, H index of countries in dental research for first authors, economical level of countries for the first authors and the last authors (Table 3). The VIF values of all the risk indicators except for H index of institutions in dental research for first authors were <10. Therefore, H index of institutions in dental research for first authors was removed from the subsequent multivariate regression analysis.

In the subsequent multivariate regression analysis with backward selection, the number of study centers (OR: 2.131; 95% CI: 1.092–4.158; $p = 0.03$), trial design (OR: 0.395; 95% CI: 0.193–0.810; $p = 0.01$), and the overall H index of institutions for the last authors (OR: 0.998; 95% CI: 0.996–0.999; $p < 0.01$) remained as independent risk indicators for the extent of *spin* (Table 3). The studies performed in single centers had 2.131 times higher odds to have high extent of *spin* in the abstracts than those in multiple centers. The studies with nonparallel designs (e.g., cross-over, split-

Table 3. Univariate and multivariate binary logistic regression analyses for the association between the risk indicators and the extent of *spin* (*N* = 234)

Risk indicators	Univariate analysis		Multivariate analysis (full model)		Multivariate analysis with backward selection	
	OR (95% CI)	*p* value	OR (95% CI)	*p* value	OR (95% CI)	*p* value
Study level						
Funding						
Funded	Ref.		Ref.			
Not funded/not reported	2.485 (1.443, 4.279)	<0.01	1.410 (0.734, 2.709)	0.30		
Number of study centers						
Multiple	Ref.		Ref.		Ref.	
Single	2.731 (1.445, 5.160)	<0.01	1.596 (0.773, 3.293)	0.206	2.131 (1.092, 4.158)	0.03
Structured abstracts						
No	Ref.					
Yes	1.974 (0.913, 4.270)	0.08				
Reporting quality of abstracts	0.850 (0.746, 0.969)	0.02	0.929 (0.806, 1.072)	0.32		
Involvement of a statistician or methodologist						
No	Ref.					
Yes	0.905 (0.495, 1.657)	0.75				
Trial design						
Parallel trials	Ref.		Ref.		Ref.	
Nonparallel trials	0.435 (0.222, 0.850)	0.02	0.432 (0.207, 0.900)	0.03	0.395 (0.193, 0.810)	0.01
Author level						
Number of authors	0.921 (0.834, 1.017)	0.11				
Conflict of interest						
Yes	Ref.					
No/not reported	0.964 (0.274, 3.392)	0.95				
International collaborations						
No	Ref.		Ref.			
Yes	0.348 (0.182, 0.667)	<0.01	0.573 (0.262, 1.252)	0.16		
H index of first author	0.966 (0.932, 1.002)	0.06				
H index of last author	0.965 (0.945, 0.986)	<0.01	0.990 (0.962, 1.020)	0.52		
Journal level						
4-year impact factor	0.775 (0.635, 0.945)	0.01	1.068 (0.795, 1.436)	0.66		
H index of the journals	0.990 (0.983, 0.997)	<0.01	0.994 (0.985, 1.003)	0.20		
Word limit of the abstracts						
≤250 words	Ref.					
>250 words	0.957 (0.556, 1.646)	0.87				
Institutional level						
Overall H index of institutions for first authors	0.998 (0.997, 1.000)	0.01	1.001 (0.998, 1.003)	0.56		
Overall H index of institutions for last authors	0.998 (0.996, 0.999)	<0.01	0.999 (0.996, 1.002)	0.45	0.998 (0.996, 0.999)	<0.01
H index of institutions in dental research for first authors	0.991 (0.984, 0.998)	0.01				
H index of institutions in dental research for last authors	0.988 (0.981, 0.995)	<0.01	0.998 (0.985, 1.010)	0.74		
National level						
Overall H index of countries for first authors	1.000 (0.999, 1.000)	0.15				
Overall H index of countries for last authors	1.000 (0.999, 1.000)	0.24				
H index of countries in dental research for first authors	0.995 (0.991, 1.000)	0.05	0.999 (0.992, 1.006)	0.83		
H index of countries in dental research for last authors	0.996 (0.992, 1.001)	0.12				
Economical level of countries for first author						
Middle/low income	Ref.		Ref.			
High income	0.531 (0.308, 0.917)	0.02	0.702 (0.124, 3.967)	0.69		
Economical level of countries for last author						
Middle/low income	Ref.		Ref.			
High income	0.512 (0.298, 0.879)	0.02	1.576 (0.270, 9.188)	0.61		

The dependent variable, extent of *spin*, was dichotomized into high extent (*spin* in both results and conclusion sections) and low extent (no *spin* in either results or conclusion section, *spin* in results section only, or *spin* in conclusion section only). Low extent of *spin* was the reference category.

mouth, or cluster design) had 0.395 times less odds to have high extent of *spin* in the abstracts than those with parallel designs. With the overall H index of institutions for the last authors increasing by one unit, the studies had 0.998 times less odds to have high extent of *spin*.

Discussion

Summary of the Main Findings

The present study shows that the prevalence of *spin* in the abstracts of RCTs in dental caries reporting statistically nonsignificant results for primary outcomes was high, with a prevalence of 79%. With a prevalence of 3%, *spin* in the titles was rare. A lower extent of *spin* was found for multicenter studies, studies with nonparallel designs, and studies in which the institutions for last authors had a higher H index. To our best knowledge, this is the first study focusing on *spin* in RCTs in dental caries.

Interpretation of the Main Findings

The important reasons for the high prevalence of *spin* in the abstracts of RCTs in dental caries may be authors' unawareness of *spin* in science, ignorance of scientific standards, young researchers' imitation of previous practice, intention to increase the acceptance rate of their papers in a higher journal, and intention to influence readers with more promising and attractive results and increase the citations of the papers [Fletcher and Black, 2007; Yavchitz et al., 2016]. Besides, the current rewarding system, which focuses more on the quantity of the publications rather than the quality of the publications may also cause *spin* [Weed, 1998]. The authors' financial, personal, and intellectual conflicts of interest may also contribute to *spin* in the publications [Fletcher and Black, 2007]. The high prevalence of *spin* in the publications may also be caused by the peer-reviewing period of the scientific journals [Lazarus et al., 2016]. Lazarus et al. [2016] assessed the impact of peer reviewers on *spin* in reports of nonrandomized studies on a therapeutic intervention based on 123 primary reports. It was reported that the authors added *spin* in the manuscripts requested by peer reviewers in 9% of the publications, and peer reviewers failed to identify *spin* in the abstract conclusions in 76% of the publications [Lazarus et al., 2016]. Several previously published methodological studies on *spin* in the abstracts of RCTs with statistically nonsignificant primary outcomes also reported a high prevalence of *spin* in various fields of medicine [Boutron et al., 2010; Arunachalam et al., 2017; Eleftheriadi et al., 2020; Wu et al., 2020; Guo et al., 2021; Rassy et al., 2021]. Among those previous studies, several studies assessed *spin* in dentistry. Wu et al. [2020] reported that the prevalence of *spin* in the abstracts was 69.9% in the 196 RCTs in periodontology and oral implantology. Guo et al. [2021] reported a prevalence of 62.2% in the abstracts of the 111 RCTs in orthodontics. Eleftheriadi et al. [2020] reported a prevalence of 61.7% in the abstracts in the 47 RCTs in dentistry. This indicated that the prevalence of *spin* in the abstracts of RCTs in medicine, including dental fields, was high, which is consistent with the findings of the present study.

In the present study, number of study centers was a significant independent risk indicator for the extent of *spin*. Wu et al. [2020] found that the studies performed in multiple centers had 0.28 times less odds to have *spin* than the studies performed in single centers using the multivariate analysis based on 196 RCTs in periodontology and implantology, which is consistent with the finding of the present study. However, Wu et al. [2020] reported that the number of study centers was not a significant risk indicator for the severity of *spin* (≤two *spin* patterns vs. >two *spin* patterns) when only the abstracts with *spin* were included. Besides, the trial design and the overall H index of institutions for last authors were significant risk indicators for extent of *spin*. To our best knowledge, this is the first study to find the potential link between trial design, research performance/influence of the institutions, and extent of *spin*. However, the possible mechanisms that can explain the association of number of study centers, trial designs, and the overall H index of institutions for last authors with the extent of *spin* are unclear so far. In the previous studies on *spin*, several other risk indicators at author and study levels were found to be significantly associated with *spin* in the multivariate analyses, such as international collaboration [Guo et al., 2021] and funding [Ito et al., 2021]. However, those two risk indicators were not significant independent risk indicators in the present study. This may be because more risk indicators at different levels were added in the present study and their effects on extent of *spin* were stronger and took over the role of those risk indicators which were significant in the previous studies.

Clinical Relevance of the Study

The present study can help increase the awareness of stakeholders in medicine and healthcare including clinicians, researchers, and policymakers about *spin* in scientific publications, in particular, in the publications performed in single centers, using parallel design and in which the institution of last authors has insufficient academic impact. The study can also help them identify different *spin* patterns that may occur in the abstracts when they read or write scientific publications and help them reduce *spin* when they write scientific papers. This can ensure accurate interpretation and dissemination of oral health research for clinicians and researchers, and make clinicians' decision-making and the shaping of clinical practice guidelines more reliable and unbiased. Eventually, it is beneficial for oral health outcomes of patients. Besides,

to date, the present study was the first study to comprehensively assess the risk indicators of *spin* at different levels, notably, study, author, journal, institutional, and national levels. In this case, the significant independent risk indicators for extent of *spin* can be identified more reliably.

Limitation of the Study

In interpreting the findings of the present study, some limitations should be taken into consideration. First, the subjectivity of the assessment of *spin* may be a limitation. The definitions of *spin* are various [Chiu et al., 2017], and there are no standardized guidelines and elaborations for the identification of *spin* in scientific publications. Therefore, the identification and interpretation of *spin* patterns in the RCTs may be arbitrary and vary among different teams. This may be also the main reason why the inter-rater agreement (i.e., the kappa values) was relatively low for the initial assessment of *spin* patterns in several previous methodological studies on *spin* [Nascimento et al., 2019] and the present study. However, in the present study, the consensus between the two reviewers on the *spin* patterns of each included study was reached based on the discussion after the initial assessment and only the agreed *spin* patterns were included in the final analysis. Therefore, we believe that the results of the present study were reliable. Second, we only assessed *spin* in the RCTs with statistically nonsignificant primary outcomes with two arms because those types of RCTs were most commonly used for the assessment of *spin* [Arunachalam et al., 2017; Wu et al., 2020; Guo et al., 2021; Ito et al., 2021]. However, this may hinder the generalizability of the findings to other types of RCTs, for instance, the RCTs with more than two arms, with no specific primary outcomes, or with statistically significant primary outcomes. Third, we used H index to quantify the research performance and influence of the authors because of the simplicity of calculation using the WOSCC database. However, some disadvantages of using H index were raised by previous researchers. For example, the H index, as a single-number criterion, may be too simple to assess a researcher reliably and comprehensively because the performance of a researcher is multidimensional [Bornmann and Marx, 2011]. The H index may disadvantage newcomers since their publications and citation rates are relatively low [Bornmann and Marx, 2011]. The H index also lacks sensitivity to performance changes of a researcher [Bornmann and Marx, 2011]. Therefore, H index may not truly reflect the research performance and influence of a researcher even if it has been widely used in research so far. Fourth, a smaller number of events relative to the high number of independent variables (i.e., risk indicators) are a common limitation for multivariate regression analysis. In the present study, the number of events (i.e., studies with a high extent of *spin*) was 87, while the number of the risk indicators included in the multivariate analysis was 14. Therefore, the present study did not meet the criterion of events per variable ≥10 [Peduzzi et al., 1996; Ogundimu et al., 2016]. This may be a limitation because this may yield biased estimates of regression coefficients of the risk indicators in the present study.

Recommendations

In future research, it is suggested to further explore the mechanisms for the association of number of study centers and H index of last authors with the extent of *spin*. Besides, to minimize the *spin* in scientific publications, it is recommended to develop a new reporting guideline for *spin* or expand the current reporting guidelines by adding specific items on avoiding *spin*.

Conclusion

The prevalence of *spin* may be high in the abstracts of the RCTs in dental caries with statistically nonsignificant results for primary outcomes. Number of study centers, trial designs, and the overall H index of institutions for the last authors may be the significant independent risk indicators for the extent of *spin*. Single-center studies, studies with parallel designs, and studies with a lower overall H index of institutions for the last authors may be more likely to have a higher extent of *spin* in the abstracts than multicenter studies, studies with nonparallel designs, and studies with higher overall H index of institutions for the last authors.

Acknowledgments

The authors gratefully acknowledge the three masters students from department of Oral Public Health, ACTA, Burcu Imac (B.I.), Marcel de Bruin (M.d.B.), and Merel Beks (M.B.), for their contributions on the screening of the eligible publications and the extraction of the risk indicators.

Conflict of Interest Statement

The authors have no conflicts of interest to declare.

Funding Sources

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Author Contributions

Naichuan Su contributed to the conception/design of the work, data acquisition, data analysis, data interpretation, and drafting the manuscript; Michiel W. van der Linden contributed to data acquisition, data interpretation, and revising the manuscript critically; Clovis M. Faggion Jr. contributed to conception of the work and revising the manuscript critically; and Geert J.M.G. van der Heijden contributed to conception/design of the work, data interpretation, and revising the manuscript critically.

References

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 Jun;265(6):1141–5.

Bornmann L, Marx W. The h index as a research performance indicator. Eur Sci Ed. 2011 Aug; 37(3):77–80.

Boutron I, Dutton S, Ravaud P, Altman DG. Reporting and interpretation of randomized controlled trials with statistically nonsignificant results for primary outcomes. JAMA. 2010 May;303(20):2058–64.

Caulfield T, Ogbogu U. The commercialization of university-based research: balancing risks and benefits. BMC Med Ethics. 2015 Oct;16(1):70.

Chiu K, Grundy Q, Bero L. 'Spin' in published biomedical literature: a methodological systematic review. PLoS Biol. 2017 Sep;15(9): e2002173.

Cooper CM, Gray HM, Ross AE, Hamilton TA, Bea Downs J, Wayant C, et al. Evaluation of spin in the abstracts of otolaryngology randomized controlled trials. Laryngoscope. 2018;129(9):2036–40.

Eleftheriadi I, Ioannou T, Pandis N. Extent and prevalence of spin in randomized controlled trials in dentistry. J Dent. 2020 Sep;100:103433.

Fletcher RH, Black B. "Spin" in scientific writing: scientific mischief and legal jeopardy. Med Law. 2007 Sep;26(3):511–25.

Guo F, Fang X, Li C, Qin D, Hua F, He H. The presence and characteristics of 'spin' among randomized controlled trial abstracts in orthodontics. Eur J Orthod. 2021 Oct;43(5):576–82.

Hackshaw AK, Paul EA, Davenport ES. Assessing the effectiveness of treatments. In: Hackshaw AK, Paul EA, Davenport ES, editors. Evidence-Based Dentistry, An Introduction. Oxford: Blackwell Munksgaard; 2006. p. 68–9.

Hirsch JE. An index to quantify an individual's scientific research output. Proc Natl Acad Sci U S A. 2005 Nov;102(46):16569–72.

Hopewell S, Clarke M, Moher D, Wager E, Middleton P, Altman DG, et al. CONSORT for reporting randomized controlled trials in journal and conference abstracts: explanation and elaboration. PLoS Med. 2008 Jan;5(1):e20.

Hróbjartsson A, Gøtzsche PC. Powerful spin in the conclusion of Wampold et al.'s re-analysis of placebo versus no-treatment trials despite similar results as in original review. J Clin Psychol. 2007 Apr;63(4):373–7.

Ito C, Hashimoto A, Uemura K, Oba K. Misleading reporting (spin) in noninferiority randomized clinical trials in oncology with statistically not significant results: a systematic review. JAMA Netw Open. 2021 Dec;4(12):e2135765.

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 May;2(5):e192622.

Lazarus C, Haneef R, Ravaud P, Hopewell S, Altman DG, Boutron I. Peer reviewers identified spin in manuscripts of nonrandomized studies assessing therapeutic interventions, but their impact on spin in abstract conclusions was limited. J Clin Epidemiol. 2016 Sep;77:44–51.

Lefebvre C, Manheimer E, Glanville J. Chapter 6: searching for studies. In: Higgins JPT, Green S, editors. Cochrane Handbook for Systematic Reviews of Interventions, version 5.1.0. Cochrane. 2011 (updated 2011 March) Available from: https://handbook-5-1.cochrane.org/chapter_6/6_4_11_search_filters.htm.

Lockyer S, Hodgson R, Dumville JC, Cullum N. "Spin" in wound care research: the reporting and interpretation of randomized controlled trials with statistically non-significant primary outcome results or unspecified primary outcomes. Trials. 2013 Nov;14:371.

Makou O, Eliades T, Koletsi D. Reporting, interpretation, and extrapolation issues (SPIN) in abstracts of orthodontic meta-analyses published from 2000 to 2020. Eur J Orthod. 2021 Mar;43(5):567–75.

Nascimento DP, Costa LOP, 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 Jun;100(10):1976–85.e18.

O'brien RM. A caution regarding rules of thumb for variance inflation factors. Qual Quant. 2007 Mar;41(5):673–90.

Ochodo EA, de Haan MC, Reitsma JB, Hooft L, Bossuyt PM, Leeflang MMG. Overinterpretation and misreporting of diagnostic accuracy studies: evidence of "spin". Radiology. 2013 May;267(2):581–8.

Ogundimu EO, Altman DG, Collins GS. Adequate sample size for developing prediction models is not simply related to events per variable. J Clin Epidemiol. 2016 Aug;76:175–82.

Peduzzi P, Concato J, Kemper E, Holford TR, Feinstein AR. A simulation study of the number of events per variable in logistic regression analysis. J Clin Epidemiol. 1996 Dec;49(12):1373–9.

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 Nov;139: 87–95.

Roszhart JI, Kumar SS, Allareddy V, Childs CA, Elangovan S. Spin in abstracts of randomized controlled trials in dentistry: a cross-sectional analysis. J Am Dent Assoc. 2020;151(1): 26–32.e3.

Safire W. The spinner spun. The New York Times [Internet]. 1996 Dec 22 [cited 2022 Apr 2022]. Available from: https://www.nytimes.com/1996/12/22/magazine/the-spinner-spun.html.

Sollaci LB, Pereira MG. The introduction, methods, results, and discussion (IMRAD) structure: a fifty-year survey. J Med Libr Assoc. 2004 Jul;92(3):364–7.

The World Bank [Internet]. World Bank Country and Lending Groups. 2021 [cited 2021 Jul 20]. Available from: https://datahelpdesk.worldbank.org/knowledgebase/articles/906519-world-bank-country-and-lending-groups.

Vos T, Flaxman AD, Naghavi M, Lozano R, Michaud C, Ezzati M, et al. Years lived with disability (YLDs) for 1160 sequelae of 289 diseases and injuries 1990–2010: a systematic analysis for the Global Burden of Disease Study 2010. Lancet. 2012 Dec; 380(9859):2163–96.

Weed DL. Preventing scientific misconduct. Am J Public Health. 1998 Jan;88(1):125–9.

Wu X, Yan Q, Fang X, Hua F, Shi B, Tu YK. Spin in the abstracts of randomized controlled trials in periodontology and oral implantology: a cross-sectional analysis. J Clin Periodontol. 2020 Jul;47(9): 1079–86.

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 Jul;75:56–65.