Evaluation of Spin in the Abstracts of Systematic Reviews and Meta-Analyses Focused on the Treatment of Acne Vulgaris: Cross-Sectional Analysis

Ryan Ottwell, BS; Taylor C Rogers, BS; J Michael Anderson, BS; Austin Johnson, BS; Matt Vassar, PhD
Oklahoma State University Center for Health Sciences, Tulsa, OK, United States

Corresponding Author:
Ryan Ottwell, BS
Oklahoma State University Center for Health Sciences
1111 W 17th St
Tulsa, OK, 74107
United States
Phone: 1 9189918718
Email: ryan.ottwell@okstate.edu

Abstract

**Background:** Spin is the misrepresentation of study findings, which may positively or negatively influence the reader's interpretation of the results. Little is known regarding the prevalence of spin in abstracts of systematic reviews, specifically systematic reviews pertaining to the management and treatment of acne vulgaris.

**Objective:** The primary objective of this study was to characterize and determine the frequency of the most severe forms of spin in systematic review abstracts and to evaluate whether various study characteristics were associated with spin.

**Methods:** Using a cross-sectional study design, we searched PubMed and EMBASE for systematic reviews focusing on the management and treatment of acne vulgaris. Our search returned 316 studies, of which 36 were included in our final sample. To be included, each systematic review must have addressed either pharmacologic or nonpharmacologic treatment of acne vulgaris. These studies were screened, and data were extracted in duplicate by two blinded investigators. We analyzed systematic review abstracts for the nine most severe types of spin.

**Results:** Spin was present in 31% (11/36) of abstracts. A total of 12 examples of spin were identified in the 11 abstracts containing spin, with one abstract containing two instances of spin. The most common type of spin, *selective reporting of or overemphasis on efficacy outcomes or analysis favoring the beneficial effect of the experimental intervention*, was identified five times (5/12, 42%). A total of 44% (16/36) of studies did not report a risk of bias assessment. Of the 11 abstracts containing spin, six abstracts (55%) had not reported a risk of bias assessment or performed a risk of bias assessment but did not discuss it. Spin in abstracts was not significantly associated with a specific intervention type, funding source, or journal impact factor.

**Conclusions:** Spin is present in the abstracts of systematic reviews and meta-analyses covering the treatment of acne vulgaris. This paper raises awareness of spin in abstracts and emphasizes the importance of its recognition, which may lead to fewer incidences of spin in future studies.

**KEYWORDS**

acne vulgaris; systematic review; abstracts; dermatology

Introduction

**Background**

Acne vulgaris is one of the most common medical diagnoses made by dermatologists each year, with over 50 million Americans affected by this condition [1]. In 2013 alone, more than 5 million people with acne vulgaris sought medical treatment, resulting in over US $1.2 billion spent annually on the disease [1]. This condition can cause physical and emotional discomfort for those affected and can lead to long-term scarring [2]. Owing to the prevalence of acne vulgaris, it is important for the providers to be well informed about the current treatment options that are supported with the highest level of evidence.

* JMIR Dermatol 2020;3(1):e16978 doi: 10.2196/16978

http://derma.jmir.org/2020/1/e16978/
Systematic reviews offer valuable insights to clinicians about current treatments [3]. Clinicians will often look specifically at the abstract of systematic reviews—which may be all that is accessible to them—as a succinct summary of results to guide treatment decisions. Barry et al [4] found that abstracts influenced the treatment decisions of nearly three-quarters of family physicians. Similarly, Marcelo et al [5] reported that nearly two-thirds of the residents used the abstracts of papers alone to guide them in their clinical decision making. Johnson et al [6] reported that the majority of nurse practitioner students found abstracts useful in guiding their clinical decision making. Haynes et al [7], investigating the usefulness of Medical Literature Analysis and Retrieval System Online (MEDLINE) in the clinical setting, showed that almost half of the searches performed influenced clinical decision making. They also found that, more often than not, decisions were made from limited information such as the abstract, article title, subject heading, or a combination of these. Therefore, it is imperative that the information provided in the abstracts of publications be representative of the study findings and devoid of spin. Boutron et al [8] defined spin as “a specific way of reporting, intentional or not, to highlight that the beneficial effect of the experimental treatment in terms of efficacy or safety is greater than that shown by the results.” Spin, specifically in abstracts, can influence the interpretation of data, which can lead to misinformation and, ultimately, misguided decisions about treatment.

**Objectives**

Spin has been demonstrated to be prevalent in the abstracts of randomized controlled trials [9-14]. Similar lines of inquiry have noted poor reporting quality in systematic reviews [15-18]. On the basis of these two complementary lines of work, we hypothesized that spin would also be present in the abstracts of systematic reviews on the topic of treatment of acne vulgaris. This study aimed to characterize and determine the frequency of spin in the abstracts [8]. The secondary objective was to evaluate whether various study characteristics corresponded with the presence of spin in systematic review abstracts focusing on acne vulgaris.

**Methods**

**Publication Search**

This study was performed in accordance with a previously written protocol available publicly on Open Science Framework (OSF) [19]. A search of PubMed and EMBASE, which incorporates MEDLINE, was performed on September 2, 2019. The search string for PubMed was as follows: “acne vulgaris”[MeSH Terms] OR acne vulgaris[Text Word] AND (systematic[sb] OR Meta-Analysis[ptyp]). The search string for EMBASE was as follows: ‘acne vulgaris’/exp AND (‘meta analysis’/de OR ‘systematic review’/de). These searches were modeled from the search strategies provided by a number of Cochrane systematic reviews on acne vulgaris [20,21]. Search results were then added to Rayyan [22], a screening platform.

**Training**

Before screening, training was conducted, which included face-to-face training sessions for the authors RO and CR. Author MV led the training sessions as his studies regarding the presence of spin in randomized controlled trials have previously been published. During these training sessions, the authors (RO and CR) received instruction and education using example abstracts and full studies from different areas of medicine. During this training, the most severe types of spin described by Yavchitz et al [8] were discussed, and agreement was reached on the definition of each item. A Google form containing the nine most severe types of spin was developed and pilot tested for ease of use and to ensure that all the necessary data elements were included. The Google form was pilot tested by the authors (RO and CR) on numerous papers known to contain spin to ensure the form contained all the items needed for data extraction and that the Google form worked correctly. Studies were then screened independently by title and abstract by RO and CR. To be included in this study, a published article had to be a systematic review or meta-analysis designed to address either the pharmacological or nonpharmacological treatment of acne vulgaris.

**Data Extraction**

Data extraction was performed in duplicate fashion using the Google form, and the investigators were blinded during the extraction process. After data extraction was completed, the investigators were unblinded and met to discuss any inconsistency until agreement was reached. On items for which agreement could not be reached, a third investigator was available for adjudication. The process of dual extraction of data was used to substantiate the work of each individual investigator. The items extracted on the Google form included the study title, journal title, number of studies included, types of subjects included, types of studies included, types of interventions, when the review was performed, databases searched, funding source, and if the review/primary outcome was significant. Furthermore, we evaluated whether the article contained any of the nine most severe examples of spin in abstracts [8].

To characterize and determine the frequency of spin in the abstracts of systematic reviews on the topic of the treatment of acne vulgaris, we employed the classification system previously developed by Yavchitz et al [8], specifically evaluating the top nine most severe examples of spin found in the abstracts of systematic reviews. These nine types of spin are outlined in Table 1.
Table 1. Frequency of each type of spin (N=12).

<table>
<thead>
<tr>
<th>Type of spin</th>
<th>Value, n (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Conclusion contains recommendations for clinical practice not supported by the findings.</td>
<td>0 (0)</td>
</tr>
<tr>
<td>Title claims or suggests a beneficial effect of the experimental intervention not supported by the findings.</td>
<td>0 (0)</td>
</tr>
<tr>
<td>Selective reporting of or overemphasis on efficacy outcomes or analysis favoring the beneficial effect of the experimental intervention.</td>
<td>5 (42)</td>
</tr>
<tr>
<td>Conclusion claims safety based on nonstatistically significant results with a wide confidence interval.</td>
<td>0 (0)</td>
</tr>
<tr>
<td>Conclusion claims the beneficial effect of the experimental treatment despite a high risk of bias in primary studies.</td>
<td>0 (0)</td>
</tr>
<tr>
<td>Selective reporting of or overemphasis on harm outcomes or analysis favoring the safety of the experimental intervention.</td>
<td>1 (83)</td>
</tr>
<tr>
<td>Conclusion extrapolates the review’s findings to a different intervention (ie, claiming efficacy of a specific intervention although the review covers a class of several interventions).</td>
<td>4 (33)</td>
</tr>
<tr>
<td>Conclusion extrapolates the review’s findings from a surrogate marker or a specific outcome to the global improvement of the disease.</td>
<td>2 (17)</td>
</tr>
<tr>
<td>Conclusion claims the beneficial effect of the experimental treatment despite reporting bias.</td>
<td>0 (0)</td>
</tr>
</tbody>
</table>

aMore than one type of spin may have been present in the same systematic review.

Data Synthesis

To evaluate the most common forms of spin within systematic review abstracts, we calculated frequencies and percentages. To evaluate particular study characteristics associated with spin in systematic review abstracts, we planned in our protocol, a priori, to conduct a logistic regression in which the presence or absence of spin would be coded as a dichotomous criterion variable (0=presence of spin and 1=no spin), and intervention type, journal impact factor, funding source, whether there was medical writer assistance, and whether the journal requires adherence to Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) or PRISMA for Abstracts (PRISMA-A) [23,24] would be included in the predictor set. However, the final sample size of 36 systematic reviews suggested that examining the associations between categorical variables using the Chi-square test and point biserial correlations between categorical and continuous variables would be more appropriate, and it was thus decided upon before data analysis was conducted. Only two systematic reviews were written by medical writers; thus, we did not include this variable. Furthermore, none of the systematic reviews reported the use of PRISMA-A, so this variable was also eliminated from the analysis. Interrater reliability was estimated using Gwet’s AC1. All statistical analyses were performed using Stata 15.1. [25].

Results

Simple Characteristics

Our search string retrieved 316 studies, of which 58 studies were deleted as duplicates and 204 studies were excluded by the title and abstract. The studies excluded from this study are mentioned in Figure 1. We further excluded 20 studies after screening the full text. This screening process left 36 studies that met the inclusion criteria (Figure 1). Among the 36 systematic reviews included, pharmacological intervention for acne vulgaris treatment was the most common intervention type (23/36, 64%), and 31% (11/36) of studies focused on nonpharmacological interventions. The most common source of funding was through public mechanisms (12/36, 33%); however, equally as many studies made no mention of funding (12/36, 33%). The median impact factor was 3.092 for the journals with the systematic reviews included in this study’s sample. Detailed characteristics of the included studies can be found in Table 2.
Figure 1. Flow diagram of study selection.

316 articles screened

56 articles retained for data extraction

36 articles from which data were extracted

**Exclusions**

(n=260, with rationale)
- 59 not related to treatment
- 58 duplicates
- 55 not related to acne vulgaris
- 31 cannot find full text
- 19 not specific to acne vulgaris
- 13 related to acne scars
- 11 related to PCOS
- 10 review articles
- 2 contained no abstract
- 2 related to treatment monitoring

**Exclusions**

(n=20, with rationale)
- 9 not specific to acne vulgaris
- 4 not related to treatment
- 3 review articles
- 2 related to acne scars
- 2 cannot find full text
Table 2. Characteristics of included studies (N=36).

<table>
<thead>
<tr>
<th>Characteristics</th>
<th>Total (N=36), n (%)</th>
<th>With spin (n=11), n (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Source of funding</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Public</td>
<td>12 (33)</td>
<td>3 (27)</td>
</tr>
<tr>
<td>Industry</td>
<td>6 (17)</td>
<td>2 (18)</td>
</tr>
<tr>
<td>Not funded</td>
<td>6 (17)</td>
<td>3 (27)</td>
</tr>
<tr>
<td>Funding not mentioned</td>
<td>12 (33)</td>
<td>3 (27)</td>
</tr>
<tr>
<td><strong>Intervention type</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pharmacologic</td>
<td>23 (64)</td>
<td>8 (73)</td>
</tr>
<tr>
<td>Nonpharmacologic</td>
<td>11 (31)</td>
<td>3 (27)</td>
</tr>
<tr>
<td>Combined</td>
<td>2 (6)</td>
<td>0 (0)</td>
</tr>
<tr>
<td><strong>Journal requirement for Preferred Reporting Items for Systematic Reviews and Meta-Analyses</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>21 (58)</td>
<td>7 (64)</td>
</tr>
<tr>
<td>No</td>
<td>15 (42)</td>
<td>4 (36)</td>
</tr>
<tr>
<td><strong>Journal requirement for Preferred Reporting Items for Systematic Reviews and Meta-Analyses for Abstracts</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>0 (0)</td>
<td>0 (0)</td>
</tr>
<tr>
<td>No</td>
<td>0 (0)</td>
<td>0 (0)</td>
</tr>
<tr>
<td><strong>Use of a medical writer</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>2 (6)</td>
<td>1 (9)</td>
</tr>
<tr>
<td>No</td>
<td>34 (94)</td>
<td>0 (0)</td>
</tr>
</tbody>
</table>

**Primary Outcome**

For the primary outcome, interrater reliability was assessed using Gwet’s AC1, which was 0.68 (95% CI 0.44-0.92). Each discrepancy was subsequently resolved by group discussion, achieving an agreement of 100%. Spin was determined to be present in 31% (11/36) of abstracts in this study’s sample. A total of 12 separate instances of spin were identified in the 11 abstracts containing spin, with one of the abstracts containing two examples of spin. The most common type of spin identified was type 3, *selective reporting of or overemphasis on efficacy outcomes or analysis favoring the beneficial effect of the experimental intervention* (5/12, 42%); followed by type 5, *conclusion claims the beneficial effect of the experimental treatment despite a high risk of bias in primary studies* (4/12, 33%); type 6, *conclusion extrapolates the review’s findings to a different intervention* (ie, claiming efficacy of one specific intervention although the review covers a class of several interventions; 2/12, 17%); and type 7, *selective reporting of or overemphasis on harm outcomes or analysis favoring the safety of the experimental intervention* (1/12, 8%). Table 1 summarizes this study’s findings for the top nine most severe types of spin in the abstracts of systematic reviews.

Although we found no instances of spin type number 4 regarding safety (Table 1), safety was not mentioned in 61% (22/36) of abstracts. For the purpose of this study, we also considered the discussion of adverse events or treatment side effects as the author inferring safety. A total of 16 studies did not provide a risk of bias assessment (16/36, 44%), and four studies assessed the risk of bias but did not discuss it (4/36, 11%). Of the 11 abstracts that did contain a form of spin, six (55%) did not report a risk of bias assessment or performed a risk of bias assessment but did not discuss it. If an article mentioned a risk of bias assessment but did not provide explanation of risk assessment or data concerning the risk of bias, then we considered this as assessed but not discussed. Spin in the abstracts was not significantly associated with a specific intervention type, funding source, or journal impact factor.

**Discussion**

**Principal Findings**

Spin is prevalent in the abstracts of systematic reviews and meta-analyses focused on the treatment of acne vulgaris. We identified spin in nearly one-third of the included studies in this study’s sample. The most common type of spin found was type 3, *selective reporting of or overemphasis on efficacy outcomes or analysis favoring the beneficial effect of the experimental intervention*. For example, Koo et al [26] performed a meta-analysis comparing the efficacy of oral antibiotics with that of oral contraceptives for the treatment of acne vulgaris at 3 and 6 months. The oral contraceptive pills were shown to be less effective than oral antibiotics at 3 months but were equivalent to oral antibiotics at 6 months. However, the author suggests in the abstract that oral contraceptive pills may be a better alternative to oral antibiotics while providing no data to support that oral contraceptive pills would be more effective.

We were limited in our assessment of spin type 5, *conclusion claims the beneficial effect of the experimental treatment despite a high risk of bias in primary studies*, as bias was not assessed or was assessed and not discussed, in many of the included studies. This issue is of concern as we noticed six of our 11
studies, which contained spin, either did not assess for the risk of bias or assessed for bias but did not discuss it or further account for it. Providing a risk of bias assessment in systematic reviews is important as it guides the reader in determining the extent to which the included study results should be trusted [27]. Another reason it is essential to perform a risk of bias assessment is that a study may have an important risk of bias, even though the study was performed to the highest possible standards. A few of the reviews included stated that their studies had methodological limitations but did not provide a proper risk of bias assessment. Without a risk of bias assessment, the reader does not have a promising tool to assess the validity of the study. Therefore, it is imperative that systematic reviews perform and discuss a risk of bias assessment of their primary studies.

**Recommendations**

To our knowledge, we are the first to investigate the presence of spin in the abstracts of systematic reviews and meta-analyses focused on acne vulgaris treatment options. In the field of dermatology, Motosko et al [28] found that 100% (20/20) of the randomized controlled trials on topical treatments for photoaged skin contained some form of spin, most broadly classified as either inappropriate statistical analyses or inappropriate interpretation of results. Other studies in various fields have previously demonstrated the presence of spin in the abstracts of randomized controlled trials. For example, Boutron et al [29], in the field of oncology, showed that spin in the abstract of a randomized controlled trial can influence clinicians’ understanding of the study’s results, to the degree that they are more likely to rate a treatment as beneficial, even though the primary outcome is statistically nonsignificant. This finding could have significant clinical implications across all fields of medicine and research in general. A total of four studies conducted by the members of our research team identified spin within abstracts of randomized controlled trials. Austin et al [9] identified spin in the abstracts of nearly half of randomized obesity trials. Cooper et al [10] demonstrated the presence of spin in 70% of the trials published in highly ranked otorhinolaryngology journals. Checketts et al [14] identified spin in 58% of the lower extremity joint trials. Kinder et al [11] reported spin being present in 23% abstracts of the randomized control trials in the field of anesthesiology. Outside of our team’s research, others have found similar results. Khan et al [30] reported spin in the abstracts of 53% of the cardiovascular randomized clinical trials. The prevalence of spin in these previously mentioned studies and others led us to ask the question whether spin was also prevalent in the abstracts of systematic reviews. Considering previous studies have shown that abstracts may influence clinical decision making [4-7], the presence of spin in the abstracts of systematic reviews may lead to the dissemination of misinformation and, ultimately, misguided decisions in clinical practice. The steps to minimize spin rest on the shoulders of all research stakeholders, including peer reviewers. Previous studies have demonstrated that reviewers often fail to recognize the misrepresentation of results in study abstracts. For example, Lazarus et al [31] demonstrated that peer reviewers failed to identify spin in abstract conclusions in 76% of the reports reviewed, and 15% of the reviewers actually requested the authors to add some type of spin. Therefore, the task of reducing spin includes the ability to recognize it. Increasing the knowledge of spin could be implemented through education and training so that reviewers, editors, and authors are better equipped to identify and eliminate spin. Guidelines have been developed and published, which will assist reviewers in the detection and interpretation of misleading claims [32,33]. Making use of these guidelines a common practice may help in identifying and eliminating spin.

The PRISMA Statement [34] is an evidence-based minimum set of items for reporting in systematic reviews and meta-analyses, which is widely used by authors, peer reviewers, and editors for reporting and critical appraisal of published systematic reviews [24]. PRISMA-A is an extension to the PRISMA Statement, which provides guidelines for reporting in the abstracts of systematic reviews. Beyond authors and reviewers, we advocate that clinicians, who ultimately make treatment decisions, should be trained to evaluate for the presence of spin in abstracts. Methodologists have performed interesting work to aid in establishing the confidence of the results and conclusions of systematic reviews; however, more work is needed regarding systematic review abstracts. A study performed by Gómez-García et al [35] found that, on average, only 57% of the PRISMA-A items were included in the abstracts of systematic reviews and meta-analyses covering psoriasis treatments. They reported that studies with low-risk bias and high methodological quality had significantly more PRISMA-A items reported compared with studies with high-risk bias and low methodological quality. On the basis of these findings, examining a study’s adherence to PRISMA-A may be a useful screening tool for journal editors, reviewers, and clinicians to assess a study’s methodological quality and risk of bias. Currently, the PRISMA-A guideline does not specifically address spin; therefore, we recommend mandatory adherence to PRISMA and a revision to the PRISMA extension, PRISMA-A, which focuses on the minimization of spin.

**Strengths and Limitations**

This study has several strengths and limitations. Regarding strengths, investigators underwent extensive training to ensure the understanding of the concept of spin in publications. A group consensus was reached on the definition of each of the nine types of spin in an effort to ensure standardized responses. We fostered an atmosphere of reproducibility by posting the full study protocol on OSF before extraction. Each data extractor was blinded during the screening and data extraction process. Data were then extracted in duplicate fashion in accordance with the recommendations of the Cochrane Handbook for Systematic Reviews of Interventions [27]. Despite these efforts, this study is not without limitations. The identification of spin is inherently subjective in nature. To reduce some of this subjectivity, data extractors underwent rigorous training before data extraction and then met to discuss and resolve any discrepancies between extraction responses. A third-party arbitrator was consulted where agreement could not be reached. Another limitation of this study was the low availability of studies that met our inclusion criteria, which limited our sample size. The results of this study are applicable to the field of dermatology with regard to the treatment of acne vulgaris alone. Further studies are needed to investigate the presence of spin.
in publications regarding other areas of dermatology. Finally, the results of this study should not be generalized across other timelines or publication types other than systematic reviews regarding acne vulgaris treatment options.

Conclusions

In conclusion, a sizable amount of spin was found in the abstracts of systematic reviews focused on the treatment of acne vulgaris. Moving forward, we recommend that peer reviewers and editors should be educated about the identification of spin. We also recommend the development of strict reporting guidelines for abstracts. Further research is needed to evaluate the types of spin that most significantly affect clinical decision making, which was outside of the purview of this study. We recommend future studies to assess the frequency of spin in other specialties.

Acknowledgments

All authors of this study had access to all data. This study was not funded by any institution or sponsor.

Authors' Contributions

RO and TR contributed equally to this paper. RO contributed to data extraction, manuscript writing, manuscript editing, and project design. TR contributed to data extraction, manuscript writing, manuscript editing, and project design. MA contributed to manuscript editing, data collection, analysis, and project design. AJ contributed to data analysis, manuscript editing, and approval. CC contributed to project design, training, and project oversight. MV contributed to project design, manuscript editing, project oversight, and training.

Conflicts of Interest

None declared.

References

1. American Academy of Dermatology. Skin Conditions by the Numbers URL: https://www.aad.org/media/stats/conditions/skin-conditions-by-the-numbers [accessed 2019-09-09]

http://derma.jmir.org/2020/1/e16978/


22. Rayyan QCRI. URL: https://rayyan.qcri.org/reviews/81224 [accessed 2019-09-10]


