REASONS TO BELIEVE—BIOSTATISTICS & METHODOLOGY FOR THE NEUROSURGEON

Reporting Quality of Systematic Review Abstracts Published in Leading Neurosurgical Journals: A Research on Research Study

**BACKGROUND:** Systematic review (SR) abstracts are frequently relied upon to guide clinical decision-making. However, there is mounting evidence that the quality of abstract reporting in the medical literature is suboptimal.

**OBJECTIVE:** To appraise SR abstract reporting quality in neurosurgical journals and identify factors associated with improved reporting.

**METHODS:** This study systematically surveyed SR abstracts published in 8 leading neurosurgical journals between 8 April 2007 and 21 August 2017. Abstracts were identified through a search of the MEDLINE database and their reporting quality was determined in duplicate using a tool derived from the Preferred Reporting Items for Systematic Reviews and Meta-analyses for Abstracts (PRISMA-A) statement. All SR abstracts that provided comparison between treatment strategies were eligible for inclusion. Descriptive statistics were utilized to identify factors associated with improved reporting.

**RESULTS:** A total of 257 abstracts were included in the analysis, with a mean of 22.8 (±25.3) included studies. The overall quality of reporting in included abstracts was suboptimal, with a mean score of 53.05% (±11.18). Reporting scores were higher among abstracts published after the release of the PRISMA-A guidelines (M = 56.52; 21.74-73.91) compared with those published beforehand (M = 47.83; 8.70-69.57; U = 4346.00, z = −4.61, P < .001). Similarly, both word count (r = 0.338, P < .001) and journal impact factor (r = 0.199, P = .001) were associated with an improved reporting score.

**CONCLUSION:** This study demonstrates that the overall reporting quality of abstracts in leading neurosurgical journals requires improvement. Strengths include the large number abstracts assessed, and its weaknesses include the fact that only neurosurgery-specific journals were surveyed. We recommend that attention be turned toward strengthening abstract submission and peer-review processes.

**KEYWORDS:** Neurosurgery, PRISMA, Reporting quality, Research on research, Systematic review

---

**ABBR EVIATIONS:** MA, meta-analysis; PRISMA-A, Preferred Reporting Items for Systematic Reviews and Meta-analyses for Abstracts; SR, systematic review

Supplemental digital content is available for this article at www.neurosurgery-online.com.
reader. In addition to frequently providing the only accessible component of an SR upon which clinical decisions are based,10–12 abstracts can improve the precision of electronic literature searches,13–15 permit a brief assessment of study validity and applicability,16,17 and facilitate peer review.14

Despite their importance, SR abstracts have historically been inaccurately18 and incompletely reported9,19–22 and are frequently misinterpreted by clinicians.23 It was in this context that the Preferred Reporting Items for Systematic Reviews and Meta-analyses for Abstracts (PRISMA-A) tool24 was developed as an extension to the original PRISMA statement.25 The PRISMA-A tool includes 12 items that should be reported in SRs and is aimed at improving the quality of abstract reporting.25 This checklist has been formally endorsed by The Lancet, The British Medical Journal, and PLoS Medicine, among many other journals. Within the field of neurosurgery, it has been endorsed by the Neurosurgery journal and is recommended by The Journal of Neurosurgery Publishing Group. Previous studies have adapted the guideline to appraise the quality of abstracts in conference proceedings26 and the dental,27 general medical,28–30 oncological,31 and mental health32 literature. However, to date, there has not been any published assessment of the reporting quality of SR abstracts in specialty neurosurgical journals. Therefore, we sought to evaluate the completeness of SR abstract reporting in 8 leading neurosurgery journals and to determine factors associated with superior reporting.

METHODS

Design and Objectives

This research-on-research study systematically surveyed SR abstracts published in 8 leading neurological journals between 8 April 2007 and 21 August 2017 according to a prespecified search protocol.24 The quality of reporting in included studies was determined using a tool derived from the PRISMA-A statement. We aimed to; (1) assess the quality of SR reporting in leading neurological journals, and (2) identify factors associated with improved reporting, including the effect of PRISMA-A guidelines on reporting quality.

Eligibility Criteria

Given that the PRISMA-A statement was primarily designed to guide reporting of interventional SRs,24 we restricted inclusion to SRs that provided comparison between treatment and control, or between two or more treatment strategies, and included more than one study. An SR was defined as any review that attempts to identify, appraise, and synthesize all of the primary research using an explicit and prespecified methodology that is aimed at minimizing bias. This definition is based upon that proposed by the Cochrane Collaboration.33

All SRs published in Neurosurgery, World Neurosurgery, Journal of Neurosurgery, Journal of Neurosurgery: Pediatrics, Journal of Neurosurgery: Spine, Neurosurgical Focus, Stereotactic & Functional Neurosurgery, or British Journal of Neurosurgery were considered eligible for inclusion. These journals were selected on the basis of having an exclusively neurological focus, and an impact factor of greater than one in the Journal Citation Reports database, which reflects the approach taken in a number of similar studies.20,34–37 Inclusion was restricted to exclusively neurological journals given the lack of a unifying definition of what constitutes a “neurosurgical study” and the attendant risk of selection bias associated with attempting to identify relevant studies from general surgical or medical journals.

We excluded all narrative reviews, abstracts of conference proceedings, primary research studies, or SRs that did not directly compare treatment strategies. Clinical guidelines were excluded unless they contained an SR component that was reported in the abstract. SRs containing a majority of animal or laboratory research were also ineligible for inclusion in this review.

Information Sources

The MEDLINE database was searched from 8 April 2007 to 8 April 2017 without restriction of language. The following search strategy was employed: (((((((((“J Neurosurg”[Journal]) OR “World Neurosurg”[Journal]) OR “J Neurosurg Pediatr”[Journal]) OR “J Neurosurg Spine”[Journal]) OR “J Neurosurg Focus”[Journal]) OR “Stereotact Funct Neurosurg”[Journal]) OR “Neurosurgery”[Journal])) AND (((((((((“Review”[Title/Abstract]) OR “meta-analysis”[Title/Abstract]) OR “systematic review”[Title/Abstract]) OR “systematic review”[Title/Abstract]) OR “meta-analysis”[Title/Abstract]) OR “meta-analysis”[Title/Abstract]) OR “meta-analysis”[Title/Abstract]) OR “meta-analysis”[Title/Abstract]) OR “meta-analysis”[Title/Abstract]) OR “meta-analysis”[Title/Abstract]))) AND (((((((“meta-analysis”[Title/Abstract]) OR “systematic review”[Title/Abstract]) OR “systematic review”[Title/Abstract]) OR “meta-analysis”[Title/Abstract]) OR “meta-analysis”[Title/Abstract]) OR “meta-analysis”[Title/Abstract]) OR “meta-analysis”[Title/Abstract]) OR “meta-analysis”[Title/Abstract]))). A single author (T.J.O.) updated this search on 21 August 2017 using the same terms.

Study Selection and Data Extraction

Two authors (T.J.O. and T.L.B.) screened all articles independently and in duplicate for inclusion in this study. Disagreement was resolved through discussion, with contested citations referred to an independent author.

Data from included studies were extracted onto a predefined data collection form. Extracted data included journal and year of publication, number of authors, journal impact factor, abstract word count, area of neurosurgery, whether or not MA was included, total number of included studies and patients, design of included studies, and whether or not the research included formal quality assessment and documented adherence to a reporting guideline. The journal impact factor closest in time to the date of publication of the abstract was also extracted.

Assessment of Reporting Quality

An assessment of SR reporting quality was performed independently and in duplicate by 2 authors (T.J.O. and T.L.B.) using a predefined coding manual adapted from the PRISMA-A checklist. All of the 12 domains outlined in the PRISMA-A statement25 and previous similar evaluations by Rice et al32 and Bigna et al30 were included. Given that structured abstracts have been associated with improved reporting35 and are strongly recommended in the PRISMA25 and PRISMA-A24 checklists, this was included as an additional quality item. Each domain received a score of 0 or 1 (not-reported/reported) for items with 2 outcomes, and 0, 1, or 2 (not-reported/incompletely reported/completely reported) for domains with 3 possibilities. A composite score (out of 23) was then determined for each article by summing individual scores for each of the 13 domains. This was then expressed as a percentage, with each domain being weighted equally when determining the overall percentage score. Discrepant judgments
were resolved by discussion, with referral of contested articles to a third author when necessary.

**Statistical Analysis**

Descriptive categorical data were summarized using absolute values and proportions, with numerical data being expressed according to their distribution. Dichotomous variables were assessed for independence using either a Pearson Chi-Square test or Fisher’s exact test depending on the expected cell frequencies.

Comparisons of central tendency between sets of continuous data were based on distribution. Normally distributed data were compared between groups with an independent 2-sample t-test, whereas a Mann–Whitney test was used to compare data from non-normal distributions. Mean values are presented with the standard deviation, while median values are accompanied by the 25th and 75th quartiles.

Where data were normally distributed, without outlier values, the presence and direction of a linear relationship between continuous variables was determined using the Pearson product–moment correlation. The strength of the association was described according to Cohen, with coefficient values of 0.1 to 0.3 considered small in strength, 0.3 to 0.5 described as moderate and >0.5 as strong.

Given that the “structured” quality item is not part of the PRISMA extension for abstracts, an additional analysis was conducted without this domain to assess the extent to which abstracts have complied with the PRISMA-A guidelines. Similarly, although the inclusion of registration and funding information has been deemed to be an essential component of SR abstracts, some may argue that their frequent presence in the main text of an SR obviates this requirement. Post-hoc sensitivity analyses were therefore performed to determine the quality of reporting once the “structure,” “registration,” and “funding” items had been excluded.

Cohen’s k was used to determine inter-rater agreement for domains with 2 outcomes, whereas weighted k was used when there were 3 potential scores. As described by Altman, coefficient values of <0.2 were considered poor, with 0.21 to 0.4 described as fair, 0.41 to 0.6 described as moderate, 0.61 to 0.8 described as good, and 0.81 to 1.00 as very good.

Statistical analyses were conducted with Statistical Package for the Social Sciences (SPSS) Version 24 (IBM Corporation, Armonk, NY, USA). Graphs were produced using GraphPad Prism 7.0 (GraphPad Software Inc., CA, USA) and Microsoft Excel 2012 (Microsoft Corporation, WA, USA). A P-value of <0.05 was considered to be statistically significant.

**RESULTS**

**Abstract Selection**

The initial database search identified 636 citations, from which 9 duplicate articles were removed. Of the remaining 627 studies, a total of 295 were excluded after independent scrutiny of their titles and abstracts. The most common reasons for exclusion were the absence of a comparative approach and not comprising an SR. A further 75 articles were excluded after a full-text evaluation of the remaining 332 studies. Therefore, a total of 257 abstracts met criteria for inclusion in this review (Figure 1). A full list of these studies is available online in Supplemental Digital Content 2.

**Characteristics of Included Abstracts**

A plurality (29.8%) of the studies were published in *World Neurosurgery* and originated in North America (57.8%). Spine (33.7%) and neurovascular (19.0%) were the most frequently represented neurosurgical subspecialties, and most articles (81.8%) were published between 2012 and 2017.

The mean number of studies included in each SR was 22.8 (±25.3), describing a median of 1115.0 patients (493.5-2266.0). The median number of contributing authors was 6 (4.0-7.0), and the mean length of each abstract was 286.43 (±70.8) words. The mean impact factor for each of the included journals between 2007 and 2017 was 2.6 (±0.9), and over half (62.3%) of the analyzed papers included an MA.

Among papers that documented utilization of a reporting guideline (42.8%), PRISMA (31.5%) was the most commonly cited. The majority (64.2%) of SRs included formal quality assessment using a standardized tool, of which the Cochrane (10.9%), Grading of Recommendations Assessment, Development, and Evaluation Working Group (11.7%), and Newcastle Ottawa Scale (11.3%) tools were the most frequently reported. The characteristics of included abstracts are detailed in Table 1.

**Quality of Abstract Reporting**

In assessing consistency between scorers, Cohen’s k revealed “very good” agreement for domains with both 2 (k = 0.958, P < .001) and 3 potential outcomes (k_w = 0.812, P < .001).

A majority of the abstracts were structured (89.5%) and appropriately identified their content in the title (84.8%). Similarly, 69.3% of papers adequately explained the research question with respect to the participants, intervention(s), comparator(s), and outcome(s) under investigation. However, the overall quality of reporting in included abstracts was suboptimal, with a mean percentage score of 53.05 (±11.18).

In total, 138 (53.7%) papers adequately described their inclusion criteria and less than one-third (31.9%) reported the databases and dates of the literature search. Approximately one-tenth (10.9%) of papers documented a reproducible risk of bias tool.

Included studies and baseline participant characteristics were sufficiently reported in 156 (60.7%) papers. Approximately two-thirds (66.9%) of analyzed studies adequately synthesized their main results using appropriate summary measures. Although 180 abstracts (70.0%) described the direction or size of the effect, only 33 studies (12.8%) reported both. One study (0.4%) summarized important strengths and limitations, 74 studies (28.4%) described one of those 2 items, and the remainder (71.2%) reported neither. Approximately one-half (49.4%) of studies concluded with an interpretation of the salient results and the implications of these findings.

One study (0.4%) declared its funding source(s). Five studies (1.9%) documented a registration number and registry name.
The quality of reporting in included abstracts is summarized in Figure 2 and Table 2.

Post-hoc sensitivity analysis revealed a mean reporting quality score of 51.39% (±11.19) once the “structured” quality item had been excluded from the dataset. There was an improvement in mean percentage reporting score with the further exclusion of “registration” and “funding” domains (59.29% ±12.78).

Factors Associated with Improved Reporting

We observed an improvement in reporting quality score values among abstracts published after the release of the PRISMA-A guidelines (M = 56.52%; 47.83-60.87) compared with those published beforehand (M = 47.83%; 39.13-56.52) (U = 4346.00, z = -6.84, P < .001; Figure 3). There was no difference in percentage reporting score between papers with 1 to 6 authors (M = 56.52%; 47.83-60.87) and those with 7 to 17 authors (M = 52.17%; 43.47-60.87; U = 6975.00, z = -0.68, P = .498). There was also no observed difference in reporting score between abstracts published in journals that endorse the PRISMA statement (M = 52.17%; 43.48-60.87) compared with those published in journals that do not (M = 52.17%; 47.82-56.52; U = 3211.00, z = -1.00, P = .319), or in those published in journals that refer to the PRISMA guidelines in their “Instructions to Authors” (M = 56.52%; 47.83-60.87) and those that do not (M = 52.17%; 43.48-60.87; U = 6411.00, z = -2.70, P = .721).

A Pearson’s product–moment correlation revealed a moderate, positive correlation between word count and reporting score (r = 0.338, P < .001). Journal impact factor was also positively correlated with reporting score (r = 0.199, P = .001). The
TABLE 1. Characteristics of Included Abstracts

<table>
<thead>
<tr>
<th>Journal</th>
<th>Number of included studies [M (R)]</th>
<th>Number of patients [M (R)]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Neurorsurgery [n (%)]</td>
<td>32 (12.5)</td>
<td>25 (70.8)</td>
</tr>
<tr>
<td>World Neurosurg [n (%)]</td>
<td>77 (30.0)</td>
<td>22.8 (±0.9)</td>
</tr>
<tr>
<td>J Neurosurg</td>
<td>33 (12.8)</td>
<td>25.3 (±10.1)</td>
</tr>
<tr>
<td>J Neurosurg—Spine [n (%)]</td>
<td>37 (14.4)</td>
<td>25.3 (±10.1)</td>
</tr>
<tr>
<td>J Neurosurg—Pediatr [n (%)]</td>
<td>18 (7.0)</td>
<td>25.3 (±10.1)</td>
</tr>
<tr>
<td>Neurosurg Focus [n (%)]</td>
<td>27 (10.5)</td>
<td>25.3 (±10.1)</td>
</tr>
<tr>
<td>Neurosurg Rev [n (%)]</td>
<td>17 (6.6)</td>
<td>25.3 (±10.1)</td>
</tr>
<tr>
<td>Stereot Func Neurosurg [n (%)]</td>
<td>2 (0.8)</td>
<td>25.3 (±10.1)</td>
</tr>
<tr>
<td>Br J Neurosurg [n (%)]</td>
<td>14 (5.4)</td>
<td>25.3 (±10.1)</td>
</tr>
</tbody>
</table>

Number of authors [M (R)] 6.0 (1-16)
Abstract word count [x, (s)] 286.43 (±70.8)


Quality of Abstract Reporting

Two primary research studies, and one editorial have previously evaluated the reporting quality of the main text of neurosurgical literature reviews. The presently described evaluation is focused on the reporting quality of a large number of SR abstracts published in leading neurosurgical journals. It utilized a predefined quality assessment scale that was adapted from the PRISMA-A checklist and compared favorably to existing studies with documented inter-rater reliability.

We found that the quality of SR abstract reporting in the neurosurgical literature is suboptimal, with a mean percentage quality score of 53.05 (±11.18) and only 3 of the 13 items receiving a judgment of “completely reported” in more than 50% of articles. These results are broadly consistent with previous studies that have documented poor abstract reporting across various medical specialties.

Informative titles and structured abstracts are known to improve indexing, electronic searching, and empower stakeholders to rapidly identify relevant articles. In this sample of neurosurgical studies, we observed that 15.2% of review titles did not identify themselves as containing an SR and/or MA, and 10.5% were unstructured. Similarly, 30.8% of papers did not adequately describe the objectives of the review, and almost half (46.5%) of all papers did not record the eligibility criteria for article inclusion. These omissions preclude an informed interpretation of the external validity of the study’s results.

The PRISMA-A guidelines recommend that abstracts document the databases that are accessed and any date ranges that are incorporated into the search strategy. This facilitates an assessment of how comprehensive and recent a search is, and permits subsequent replication. We found that this information was reported in only 31.9% of neurosurgical abstracts, which compares unfavorably with the oncology, mental health, and dental literature.
Approximately one-tenth (10.9%) of papers documented the utilization of a specific risk of bias assessment tool, which is significantly less than that reported in a survey of the general medical literature (55.39%), but consistent with a similar mental health study (14.29%).

Reporting the number of included studies, participants, and baseline demographic and clinical characteristics is needed to assess the validity and applicability of an SR’s results, and are a central component in the PRISMA-A guidelines. In the current series, 60.9% of abstracts (n = 156) reported both the number and characteristics of included studies, with a further 29.6% (n = 76) reporting one of these 2 domains. These findings are concordant with assessments of the general medical literature. First, by Bigna et al., who found that 61.27% reported both domains, and, second, by Tsou et al. who observed that 90% of papers reported at least one of the two.

Two-thirds (66.7%) of surveyed studies adequately synthesized their results, which is slightly less than what has been documented in the mental health (76.19%) and general medical (78.43%) literature, but compares favorably to 93 oral implantology SRs surveyed by Kiriakou and colleagues (51.1%). The description of effect was inadequately reported, with 12.8% of papers expressing the size and direction of the effect. This level of compliance is markedly inferior to that reported by Beller et al. in a study of 182 general medical abstracts, which found that 58% documented the direction of effect in words and numbers and approximately three-quarters of MAs included either a measure of numerical effect or uncertainty. Although it appears to be commonly neglected, the PRISMA-A recommendation for describing the direction and size of the effect has been made in the context of research demonstrating that only 61% of clinicians can correctly identify the direction of effect from SR abstracts.

In addition to succinctly summarizing the important results from the SR, abstracts should also provide a brief description of the strengths and limitations of the appraised literature to enable busy decision-makers to evaluate the validity and applicability of the findings. We found that only one study (0.4%) documented both strengths and limitations, while slightly more than one-quarter (28.4%) reported one of these two domains. This is significantly less than the 35.5% of oral implantology SRs that reported limitations, and the 30.9% and 43% of general medical, and 19% of mental illness studies that reported both domains.

Documentation of research sponsorship and study registration increases SR transparency. It is well known that sponsored primary studies and SRs are more likely to report favorable outcomes for the sponsor. This led the PRISMA-A working group to recommend documentation of sponsorship in the abstract. We
found that only one study (0.4%) complied with this suggestion. Similarly, despite growing acceptance of study registries, only 1.9% reported registration in such a database in the abstract.

**Factors Associated with Improved Reporting**

Although there has been a significant increase in the proportion of surgical journals endorsing the PRISMA guidelines, it is clear from this and other studies that isolated endorsement of the statement, or passive mention in the author instructions, is insufficient to improve the quality of reporting. There is some evidence to suggest that active enforcement of the checklists is effective, and we would advocate for incorporation of reporting guidelines into submission and peer-review processes.

The association between word count and reporting quality is controversial. While some have previously ascribed poor reporting to stringent word limits, others have demonstrated that CONSORT checklist items can been sufficiently adhered to in abstracts of 250 to 300 words. Indeed, Bigna et al reported that abstracts of <300 words had improved reporting quality when compared with those containing more than 300 words.

---

**TABLE 2. Scoring Tool Adapted from PRISMA-A Checklist and Summary of Abstract Reporting Quality**

<table>
<thead>
<tr>
<th>Item</th>
<th>Description</th>
<th>n (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Title</strong></td>
<td>0 = does not include reference to SR or MA</td>
<td>39 (15.2)</td>
</tr>
<tr>
<td></td>
<td>1 = includes SR, MA, or both</td>
<td>218 (84.8)</td>
</tr>
<tr>
<td><strong>Structured</strong></td>
<td>0 = No</td>
<td>27 (10.5)</td>
</tr>
<tr>
<td></td>
<td>1 = Yes</td>
<td>230 (89.5)</td>
</tr>
<tr>
<td><strong>Objectives</strong></td>
<td>0 = No mention of objectives</td>
<td>13 (5.1)</td>
</tr>
<tr>
<td></td>
<td>1 = 1-2 items from population, intervention, comparator, and outcomes</td>
<td>66 (25.7)</td>
</tr>
<tr>
<td></td>
<td>2 = 3-4 from population, intervention, comparator, and outcomes</td>
<td>178 (69.3)</td>
</tr>
<tr>
<td><strong>Eligibility criteria</strong></td>
<td>0 = No mention of eligibility criteria</td>
<td>41 (16.0)</td>
</tr>
<tr>
<td></td>
<td>1 = 1-2 items from population, intervention, comparator, and outcomes, or alternative characteristics such as language, year of publication, types of studies.</td>
<td>78 (30.4)</td>
</tr>
<tr>
<td></td>
<td>2 = 3-4 items from population, intervention, comparator, and outcomes</td>
<td>138 (53.7)</td>
</tr>
<tr>
<td><strong>Information sources</strong></td>
<td>0 = No mention of information sources</td>
<td>52 (20.2)</td>
</tr>
<tr>
<td></td>
<td>1 = Reports databases or date range or last search date</td>
<td>123 (47.9)</td>
</tr>
<tr>
<td></td>
<td>2 = Reports databases and dates (range or last search date)</td>
<td>82 (31.9)</td>
</tr>
<tr>
<td><strong>Risk of bias assessment</strong></td>
<td>0 = No mention of risk of bias assessment</td>
<td>183 (71.2)</td>
</tr>
<tr>
<td></td>
<td>1 = Mentions that risk of bias assessment was undertaken, but not the name of the tool or method of assessment</td>
<td>46 (17.9)</td>
</tr>
<tr>
<td></td>
<td>2 = Mentions that risk of bias assessment was undertaken, and includes the name of the tool, or the methods behind the assessment, or both</td>
<td>28 (10.9)</td>
</tr>
<tr>
<td><strong>Included studies</strong></td>
<td>0 = No mention of number of included studies</td>
<td>25 (9.7)</td>
</tr>
<tr>
<td></td>
<td>1 = Includes number of included studies or participant characteristics, but not both</td>
<td>76 (29.6)</td>
</tr>
<tr>
<td></td>
<td>2 = Includes number of included studies and participant characteristics</td>
<td>156 (60.7)</td>
</tr>
<tr>
<td><strong>Synthesis of results</strong></td>
<td>If MA:</td>
<td></td>
</tr>
<tr>
<td></td>
<td>0 = No report of results</td>
<td>9 (3.5)</td>
</tr>
<tr>
<td></td>
<td>1 = Reports summary measures or confidence intervals (not both)</td>
<td>76 (29.6)</td>
</tr>
<tr>
<td></td>
<td>2 = Reports summary measures and confidence intervals</td>
<td>172 (66.9)</td>
</tr>
<tr>
<td><strong>Description of effect</strong></td>
<td>0 = No report of effect size (description or direction)</td>
<td>44 (17.1)</td>
</tr>
<tr>
<td></td>
<td>1 = Reports effect size in words (direction) or numbers (size)</td>
<td>180 (70.0)</td>
</tr>
<tr>
<td></td>
<td>2 = Reports effect size in words (direction) and numbers (size)</td>
<td>33 (12.8)</td>
</tr>
<tr>
<td><strong>Strengths and limitations of evidence</strong></td>
<td>0 = No mention of strengths or weaknesses</td>
<td>183 (71.2)</td>
</tr>
<tr>
<td></td>
<td>1 = Identifies strengths or weaknesses</td>
<td>73 (28.4)</td>
</tr>
<tr>
<td></td>
<td>2 = Identifies strengths and weaknesses</td>
<td>1 (0.4)</td>
</tr>
<tr>
<td><strong>Interpretations</strong></td>
<td>0 = Does not provide a conclusion that involves interpretation or implication</td>
<td>1 (0.4)</td>
</tr>
<tr>
<td></td>
<td>1 = Provides conclusion that involves interpretation or implication</td>
<td>129 (50.2)</td>
</tr>
<tr>
<td></td>
<td>2 = Provides conclusion that involves interpretation and implication</td>
<td>127 (49.4)</td>
</tr>
<tr>
<td><strong>Funding</strong></td>
<td>0 = Funding not referred to</td>
<td>256 (99.6)</td>
</tr>
<tr>
<td></td>
<td>1 = Reports sources of funding</td>
<td>1 (0.4)</td>
</tr>
<tr>
<td><strong>Registration number</strong></td>
<td>0 = Registration not referred to</td>
<td>252 (98.1)</td>
</tr>
<tr>
<td></td>
<td>1 = Reports registry name (eg, PROSPERO) or registration number</td>
<td>0 (0.0)</td>
</tr>
<tr>
<td></td>
<td>2 = Reports registry name (eg, PROSPERO) and registration number</td>
<td>5 (1.9)</td>
</tr>
<tr>
<td><strong>Overall score [x ± (s)]</strong></td>
<td>53.05 ± 11.18</td>
<td></td>
</tr>
</tbody>
</table>

In the present study, we found that increased abstract word count was associated with superior reporting quality. As such, although the balance between brevity and the inclusion of sufficient detail remains unclear, the evidence seems to support an abstract word limit of not less than 300 words and leniency in its implementation.

Although existing studies have not observed a positive correlation between impact factor and improved reporting, it is consistent with existing research demonstrating an association between impact factor and journal quality. Similarly, the inverse association between the number of studies included in an SR and reporting quality was unexpected and has not been reproduced in previous studies. Although the magnitude of this relationship was small, it requires clarification in future research.

Strengths and Limitations

The presently described study had a number of strengths. It comprised of a large cohort of surgical abstracts and utilized a sensitive search, which spanned a broad range of dates and thereby facilitated an investigation of temporal changes in abstract reporting quality. The study was methodologically robust, with study selection and quality assessment conducted independently and in duplicate by two authors. Cohen’s $\kappa$ revealed an unusually high level of agreement between assessors when compared with similar existing studies.

The limitations of this study should also be acknowledged. First, although our target population comprises all comparative neurosurgical SRs, inclusion was limited to a nonrandom sample of articles published in leading exclusively neurosurgical journals. Although these journals were identified using a large database, it is possible that some eligible journals were not identified in our search and, thus, were inadvertently omitted from this analysis. Second, some may argue that the weighting of each PRISMA-A domain should reflect the perceived importance of the recommendation. However, Beller and colleagues do not place discrepant amounts of emphasis on different domains, and our approach is consistent with all existing studies in this area.

CONCLUSION

This study has indicated that the quality of abstract reporting in leading neurosurgical journals requires improvement. Given that abstracts frequently comprise the only component of an article that is accessed by clinicians and policy makers, SR authors and journal editors must ensure that they are reported accurately and comprehensively, and in accordance with recommended guidelines. More than simply endorsing reporting guidelines, or referring to them in author instructions, we would advocate that this be achieved through their incorporation into article submission and peer-review processes. The results of this study also serve as a caution to readers to thoroughly evaluate the full text of SRs prior to clinical implementation of findings reported in an abstract.

Disclosure

The authors have no personal, financial, or institutional interest in any of the drugs, materials, or devices described in this article.

REFERENCES


Some historians consider the Romantic period among or perhaps the most important moment in the evolution of thought in modern history. The Romantics comprised painters, novelists, poets, musicians, and philosophers. They challenged the very idea of “rules” and championed “truth” through beauty, emotion, reverence, innocence, and the natural world. One of the most common themes in Romantic art is that of the imposing beauty and mystery of the wilderness. In reaction to the ideas of the Enlightenment, with its emphasis on rational thought, science, and atheism, the Romantics sought and evangelized for the same feelings of awe and reverence once limited to a religious experience to be found in nature. This painting, *Wanderer Above the Sea of Fog*, by Caspar David Friedrich, depicts this search and experience. Public Domain. Details throughout issue text largely from https://www.youtube.com/watch?v=V6fTJtwLGQ.