SYSTEMATIC REVIEW

Adherence to the PRISMA statement and its association with risk of bias in systematic reviews published in rehabilitation journals: A meta-research study

Tiziano Innocenti\textsuperscript{a,b,*}, Daniel Feller\textsuperscript{c}, Silvia Giagio\textsuperscript{d,e}, Stefano Salvioli\textsuperscript{b,f}, Silvia Minnucci\textsuperscript{g}, Fabrizio Brindisino\textsuperscript{g,h}, Carola Cosentino\textsuperscript{f}, Leonardo Piano\textsuperscript{i}, Alessandro Chiarotto\textsuperscript{a,j}, Raymond Ostelo\textsuperscript{a,k}

\textsuperscript{a} Department of Health Sciences, Faculty of Science, Vrije Universiteit Amsterdam, Amsterdam Movement Sciences, the Netherlands
\textsuperscript{b} GIMBE Foundation, Bologna, Italy
\textsuperscript{c} Provincial Agency for Health of the Autonomous Province of Trento, Trento, Italy
\textsuperscript{d} Division of Occupational Medicine, IRCCS Azienda Ospedaliero-Universitaria di Bologna, Bologna, Italy
\textsuperscript{e} Department of Biomedical and Neuromotor Sciences (DIBINEM), Alma Mater Studiorum, University of Bologna, Bologna, Italy
\textsuperscript{f} Department of Neuroscience, Rehabilitation, Ophthalmology, Genetics, Maternal and Child Health, University of Genoa, Genoa, Italy
\textsuperscript{g} Department of Clinical Science and Translational Medicine, University of Rome Tor Vergata, Rome, Italy
\textsuperscript{h} Department of Medicine and Health Science “Vincenzo Tiberio”, University of Molise, Campobasso, Italy
\textsuperscript{i} Unit of Rehabilitation and Functional Recovery, Fondazione dei Santi Lorenzo e Teobaldo, Rodengo, Italy
\textsuperscript{j} Department of General Practice, Erasmus MC, University Medical Center, Rotterdam, the Netherlands
\textsuperscript{k} Department of Epidemiology and Data Science, Amsterdam UMC, Location VUMC, Amsterdam Movement Sciences research institute, the Netherlands

Received 17 November 2021; received in revised form 2 September 2022; accepted 3 October 2022
Available online 14 October 2022

KEYWORDS
Meta-research; Randomised controlled trial; Rehabilitation; Reporting guidelines; Research quality; Risk of bias

Abstract

Background: Systematic reviews (SRs) and meta-analyses are essential resources for the clinicians. They allow to evaluate the strengths and the weaknesses of the evidence to support clinical decision-making if they are adequately reported. Little is known in the rehabilitation field about the completeness of reporting of SRs and its relationship with the risk of bias (ROB).

Objectives: Primary: 1) To evaluate the completeness of reporting of systematic reviews (SRs) published in rehabilitation journals by evaluating their adherence to the PRISMA 2009 checklist, 2) To investigate the relationship between ROB and completeness of reporting. Secondary: To study the association between completeness of reporting and journals and study characteristics.
Introduction

Systematic reviews (SRs) and meta-analyses provide the highest quality of scientific evidence and are essential resources shaping the clinical decision-making process. They represent an efficient way for clinicians to keep up to date with the current evidence and provide a starting point for developing clinical guidelines. Therefore, adequate reporting of SR methods and results is essential for evaluating the strengths and weaknesses of the evidence provided. Moreover, complete and transparently reported research aids reproducibility and critical appraisal. Indeed, critical appraisal of study methods enables assessment of the extent to which results and author interpretations overestimate or underestimate study effects and highly depend on what is reported in research articles. For example, if a study has better reporting of all necessary information, this could positively impact the transparency of the risk of bias (ROB) evaluation.

Problems in the reporting of a scientific study can affect research in different ways. For example, it is known that study methods are frequently not described in adequate detail and that results are presented ambiguously, incompletely, or selectively. The consequence is that many reports cannot be used for replication purposes, are potentially a waste of resource, or could even be harmful. Reporting guidelines (RGs) have been developed to help authors optimally report key aspects of their research in their manuscripts.

The lack of standardisation and poor quality of reporting among SRs and meta-analyses led to the development of the Quality of Reporting of Meta-analyses (QUOROM) statement in 1999. In 2009, an updated version of the QUOROM, the Preferred Reporting Items for Systematic Reviews and Meta-analyses (PRISMA) statement was launched: it is a consensus-based minimum set of items to be reported in SRs with or without meta-analyses. It consists of 27 items that should be reported in the title, abstract, methods, results, and discussion of a SR and meta-analysis, including source of funding. The PRISMA statement was updated in 2020.

The ROB in systematic reviews (ROBIS) tool, specifically designed to assess ROB in SRs was published in 2016. Following a domain-based approach, ROBIS covers the evaluation of the internal validity of the review process and the relevance of the review question for its users. It comprises four specific domains (eligibility criteria, identification and selection of studies, data collection and study appraisal, synthesis and findings) and results in ROB judgment (both on each domain and overall) which can be “high,” “low,” or “unclear.”

Previous studies have assessed the adherence to the PRISMA 2009 checklist in published SRs in the medical field; one study in the field of orthopaedic surgery analysed the quality of reporting of relevant studies which were published in the top five highest impact factor (IF) orthopaedic journals and found that only 68% of items of the PRISMA 2009 statement were reported. Little is known about reporting in the rehabilitation field: a previous study confirmed that most of the SRs authors (~ 65%) publishing in high impact rehabilitation journals did not mention the use of PRISMA, and approximately 40% of those who declared using PRISMA did not do so appropriately.

To our knowledge, completeness of reporting and the relationship between ROB and completeness of reporting for SRs published in rehabilitation journals have not been systematically evaluated.

The primary objectives of this meta-research study were: 1) to evaluate the completeness of reporting in SRs published in rehabilitation journals through the evaluation of the adherence to the PRISMA checklist; 2) to investigate the relationship between the completeness of reporting and the ROB assessed with the ROBIS tool. The secondary objective was to study the association between the completeness of the reporting and the characteristics of the SRs and journals, such as impact factor, publication medium, publication year, and protocol registration. We believe that a better understanding of these topics is important to help researchers and readers in producing and assessing the evidence in a more accurate manner.
Methods

A cross-sectional analysis was conducted for a random sample of 200 SRs published between 2011 and 2020 in all 68 rehabilitation journals indexed under the “rehabilitation” category in InCites Journal Citation Report. The 68 rehabilitation journals indexed in InCites and their characteristics are reported in the online supplementary material (see Table S.1). The project protocol was publicly posted before data extraction (prospectively registered) on the medRxiv preprint server. Given that a specific reporting checklist for meta-research studies is currently under development, we followed the PRISMA 2020 checklist in the reporting of this manuscript.

Study selection criteria

Potentially eligible SRs (with or without meta-analysis) should have been published between 2011 and 2020 as full-text scientific articles in any of the 68 rehabilitation journals indexed under the “rehabilitation” category in InCites Journal Citation Report. A standard or consensus definition of a SR does not exist. Therefore we considered a review to be systematic if it displayed the following characteristics:

- A clearly defined research question
- Declaration of the sources that were searched
- Inclusion and exclusion criteria and selection methods
- Critical appraisal and reporting of the quality/ROB of the included studies
- Information about data analysis and synthesis

Narrative reviews, mixed-methods reviews, qualitative evidence synthesis, umbrella reviews, scoping reviews, editorials, letters, and news reports were excluded.

Study selection process

Journal tags for the journals were identified in Medline, and a detailed search strategy was created to find all SRs published from 2011 to 2020 in this database (see online supplementary material Table S.2). The search was performed only in Medline, as all rehabilitation journals are indexed in this database. Titles, abstracts, and full texts of all articles retrieved were screened for eligibility in a double-blinded process by two independent reviewers who selected potential articles based on study selection criteria. A third reviewer resolved any disagreement. After this process, 200 reports (our study sample) were randomly selected by an independent author using a free web service (Research Randomizer - www.randomizer.org) to create the random sequence list. Randomisation was stratified by publication date and journal ranking (quartile range; Q1, Q2, Q3, Q4 — based on InCites) to include an equal number of studies from 2011 to 2015 (n=25 for each quartile range) and from 2016 to 2020 (n=25 for each quartile range).

Data extraction

Full texts were stored in EndNote X7 (Thomas Reuters, Philadelphia, Pennsylvania, USA). A data extraction form was used, and the following characteristics were extracted: first author; publication year; journal characteristics (open access vs hybrid); country; rehabilitation field; protocol registration (yes/no); completeness of reporting (see below); ROB with the ROBIS tool (see below). Two reviewers performed data extraction and the ROB assessment for each SR, each one independently; when necessary, any disagreement was resolved by a third reviewer. For the publication option (open access or hybrid) we checked if journals had changed their policy over time (by checking the journal website or asking the editor directly).

Assessing the completeness of reporting

Completeness of reporting was calculated as adherence to the 27-item PRISMA 2009 checklist (from here onwards indicated as PRISMA checklist). Discrepancies between authors on the application of checklist items were resolved through consensus by referring to the published explanation of the PRISMA checklist. Inter-rater agreement was evaluated for 30 SRs (15% of the entire sample) which were scored independently by two authors with post-graduate training in clinical epidemiology and critical appraisal. This method was chosen to ensure reproducibility and to align the appraisers. However, PRISMA checklist is not an assessment checklist, so it does not need to be performed independently by two authors.

Following the explanation and elaboration statements of the PRISMA checklist, each item was marked with “1” if it was well described, with “0” if incomplete or missing, and with “NA” if not applicable. An item was defined as “NA” when the authors did not describe it, and if it was clear (by reading the protocol and/or the full text) that such information was justifiably missing (e.g. if authors did not perform any meta-analyses, item 23 was not applicable). Adherence to each item and overall adherence to the PRISMA checklist for each study were calculated (from 0% to 100% with 0 representing zero adherence and 100% full adherence), weighting against the number of applicable items. As the aim of this study was to investigate completeness of reporting, the authors of the studies included were not contacted for information omitted from manuscripts.

Assessing the risk of bias

The ROBIS tool was used to assess the ROB in the studies included. For each SR included, a ROB judgment was assigned for each of the four domains (1. the specification of study; 2. methods used to identify and/or select studies eligibility criteria; 3. methods used to collect data and appraise studies; 4. synthesis and findings). Possible assessment was “high,” “low,” or “some concerns.” An overall ROB score for each study was assigned as suggested by the instructions of the tool. The results were graphically summarised through the ROB graph obtained with the ROBVIS Tool.

Data analysis

The primary analysis addressed:

- The completeness of reporting in each study (see above) calculated as the overall adherence to the PRISMA checklist. It was estimated (in percentage) as the total number
of items described and reported out of the total number of applicable items.

- The completeness of reporting for each item and each section of the PRISMA checklist in all studies, calculated as (for each item; in percentage) the number of times that one item is described and reported out of the total number of studies in which the item could potentially appear in.

- The relationship between the completeness of reporting in each study (see above) and studies with low or high ROB. We hypothesised that studies with domains with a high ROB display poorer reporting than those with a low ROB. The relationship was investigated through descriptive statistics and linear regression analyses (both univariable and multivariable) between the overall adherence to the PRISMA checklist as the dependent variable and the ROB (for each ROBIS domain and overall ROB) as independent variables.

Secondary analysis:

- Linear regression analyses (both univariable and multivariable) were performed to assess the association between the completeness of reporting (calculated as the overall adherence to the PRISMA checklist — see above) as the dependent variable and the following characteristics as independent variables: publication year; journal ranking (quartile range: Q1, Q2, Q3, Q4); publication options (open access vs hybrid); study protocol prospective registration (no vs yes).

For each linear regression model, the assumptions of linearity, homoscedasticity, independence, and normality were checked. For all regression analyses, these assumptions were met.

Results

The study selection process is summarised in Fig. 1. Online supplementary material includes characteristics of included systematic reviews (Table S.3) and their complete references (Table S.4).

Completeness of reporting (adherence to the PRISMA checklist)

The mean overall adherence among our sample was 61.4%, ranging from 100% to 16% (See online supplementary material Figure S.1). Overall, the mean adherence varied across items (Table 1). The most frequently reported item was #1 in 99.5% of cases; the least reported one was the ROB across studies (#22) in 12% of cases. Some items were not applicable in most SRs (Table 1), which may have influenced the calculation of adherence in these cases (e.g. item #16 concerning the description of additional analyses, where adherence was 35% and NA=120). In general, the highest adherence rates were in the title item/section (99.5%) and discussion section (83%), the lowest in the funding section (in which there is only item #27) (39%) (Fig. 2).

Relationship between risk of bias and completeness of reporting

Overall, studies with high ROB showed poorer reporting than those with low ROB (Table 2). The results of ROBIS analysis are reported in supplementary material Table S.5. The descriptive analysis showed that reporting of the methods section presented the largest differences between studies at high and low ROB in all ROBIS domains (Table 2).

The univariable linear regression analysis showed a negative association between high ROB and completeness of the reporting (adherence to the PRISMA checklist) in all domains (see supplementary material Table S.6). Still, when we entered all independent variables in the same model, the multivariable linear regression analyses (Table 3 — primary analysis) reported an association between completeness of reporting and ROB in domains 1, 2, and overall ROB.

Having a high ROB in domain 1 (B= -5.7%; 95%CI -10.1, -1.3) and domain 2 (B= -5.4%; 95%CI -9.7, -1.1) are significant predictors of lower overall adherence to the PRISMA checklist. If we look at the overall ROB, having a high overall ROB is a significant predictor of lower adherence (B= -7.1%; 95%CI -12.1, -2.0).
Table 1  Mean adherence across each item of the PRISMA 2009 Checklist in systematic reviews published in rehabilitation journals (n = 200).

<table>
<thead>
<tr>
<th>PRISMA 2009 Items</th>
<th>Studies where item was not applicable (n)</th>
<th>Mean adherence (%) calculated in studies where item was applicable</th>
</tr>
</thead>
<tbody>
<tr>
<td>Title</td>
<td>0</td>
<td>99.5</td>
</tr>
<tr>
<td>Abstract Structured summary</td>
<td>0</td>
<td>64.0</td>
</tr>
<tr>
<td>Introduction Rationale</td>
<td>0</td>
<td>97.5</td>
</tr>
<tr>
<td>Objectives</td>
<td>0</td>
<td>52.5</td>
</tr>
<tr>
<td>Methods Protocol and registration</td>
<td>0</td>
<td>12.5</td>
</tr>
<tr>
<td>Eligibility criteria</td>
<td>0</td>
<td>71.5</td>
</tr>
<tr>
<td>Information sources</td>
<td>0</td>
<td>61.0</td>
</tr>
<tr>
<td>Search</td>
<td>0</td>
<td>49.0</td>
</tr>
<tr>
<td>Study selection</td>
<td>0</td>
<td>75.0</td>
</tr>
<tr>
<td>Data collection process</td>
<td>0</td>
<td>44.5</td>
</tr>
<tr>
<td>Data items</td>
<td>0</td>
<td>59.0</td>
</tr>
<tr>
<td>Risk of bias in individual studies</td>
<td>0</td>
<td>77.0</td>
</tr>
<tr>
<td>Summary measures</td>
<td>21</td>
<td>52.5</td>
</tr>
<tr>
<td>Synthesis of results</td>
<td>23</td>
<td>48.0</td>
</tr>
<tr>
<td>Risk of bias across studies</td>
<td>0</td>
<td>17.5</td>
</tr>
<tr>
<td>Additional analyses</td>
<td>120</td>
<td>35.0</td>
</tr>
<tr>
<td>Results Study selection</td>
<td>0</td>
<td>83.5</td>
</tr>
<tr>
<td>Study characteristics</td>
<td>0</td>
<td>94.5</td>
</tr>
<tr>
<td>Risk of bias within studies</td>
<td>0</td>
<td>85.0</td>
</tr>
<tr>
<td>Results of individual studies</td>
<td>0</td>
<td>57.0</td>
</tr>
</tbody>
</table>
Univariable linear regression models showed a significant relationship between journal IF, publication year, study protocol registration, and completeness of the reporting (see supplementary material Table S.7). When we entered all independent variables in the multivariable linear regression model, only journal IF and protocol registration were significant predictors (see Table 3 — secondary analysis). Studies published in fourth quartile journals displayed a lower overall adherence ($B=-7.2\%$; $95\%CI=1.3, 13.2$) than those published in first quartile journals; the overall adherence increased ($B=11.9\%$; $95\%CI=5.9, 18.0$) if the SR protocol was registered.

### Discussion

This study evaluated the completeness of reporting in a sample of SRs published in rehabilitation journals, and the existing correlations with ROB and other important characteristics were investigated. Our results showed that the completeness of the reporting was suboptimal: in fact, the overall adherence to the PRISMA checklist was 61.4%, with a range from 100% to 16% (see supplementary material Figure S.1). These results are in line with others in different biomedical research fields such as vascular surgery, emergency medicine, and nursing that showed a 59% to 70% range of adherence to the PRISMA 2009.

To our knowledge, this is the first study to compare the completeness of reporting to the ROB in the rehabilitation...
field. Completeness of the reporting showed a relationship with the ROB: the higher the ROB, the poorer the completeness of reporting. However, the potential causality in this association remains to be proven. We investigated this relationship through multivariable regression analysis (Table 2) with all ROB domains and overall ROB score in the same model. We adopted this solution to clearly show how reporting can be influenced by both individual domains (especially domains 1 and 2) and the overall score without multicollinearity. Tunis et al. evaluated the same association in SRs published in major radiology journals. Despite differences in methods concerning our study (smaller sample size and the use of the AMSTAR tool), they found a strong association between completeness of reporting and higher study quality.

Our secondary analysis showed that completeness of reporting is also associated with the IF and registration of the study protocol. Although IF metrics do not indicate the quality of the studies published in a journal, our results suggest that reporting is more complete in journals with a high IF. Completeness of reporting is strongly associated with the registration of the study protocol: there was an 11.95% increase in overall adherence to the PRISMA checklist if the protocol was registered.

Publication of the study protocol is one of the most critical aspects that emerge from our study. Only 15.5% (31/200 – see supplementary material Table S.3) of the studies included had published the protocol. Item 5 of the PRISMA statement was the least reported item. This remarkable observation is also confirmed by other studies. In gastroenterology and hepatology journals it was reported in only 4.4% of studies, while it was 27% in orthodontic journals. 15% of the SRs in radiology journals mentioned research protocols. This happens despite methodological standards explicitly requiring the study protocol registration and publication in SRs.

There is much room for improvement for all actors in the evidence production process (i.e. editors, peer-reviewers and researchers). Journal editors should personally check the peer-reviews and encourage peer-reviewers to pay more attention to the reporting issues. Peer-reviewers should check the accuracy of the reporting in the studies, while researchers need to improve compliance with the reporting standard and increase the readability and transparency of their research reports. Moreover, the researchers play an important role in this topic. They should not simply report their manuscript following the reporting guidelines (e.g. the PRISMA statement for systematic review authors) but they should provide appendices and raw data in sufficient detail to allow peer-reviewers and readers to assess their manuscripts and to make their results applicable.

Despite some strategies adopted by journals to improve RGs use (such as requiring authors to follow the RG in the “Instructions to authors”), these do not appear to be consistently effective. Journal editors may need to adopt electronic systems (e.g. through artificial intelligence) to check reporting accuracy in a scientific paper. These systems may run checks during the submission process, detecting and prompting for corrections before submission and explicitly requiring that authors follow RGs indications.
To overcome this issue, we have tried to de-
and we believe... papers.

being consistent with previous meta-research studies,
and not for evaluating quality. We are aware that despite
We used the PRISMA checklist to assess the completeness of
this may represent a limitation.

Study limitations

We used the PRISMA checklist to assess the completeness of reporting, despite the checklist being a guide for writing and not for evaluating quality. We are aware that despite being consistent with previous meta-research studies, this may represent a limitation. To overcome this issue, we have tried to define how we have reached the scores clearly, and we ran a preliminary pilot test to increase inter-assessor agreement, especially on items requiring interpretation (sometimes indicated by phrases like “if relevant” or “if applicable”). Moreover, between 2011 and 2020 many other SRs have been published and our sample (200 hits) may not be representative of all studies published in such time frame. However, our choice is based on the sample size used in other similar meta-research studies and we believe it is representative of the characteristic of the SRs published in the rehabilitation journals.

Conclusions

The completeness of reporting in SRs published in rehabilitation journals is suboptimal. High ROB is associated with poorer completeness of reporting, while the registration of study protocols and journal ranking are also related to more complete reporting. Journal editors, peer-reviewers, and researchers play an essential role in increasing the readability and transparency of the research reports; as such, they should make every effort to improve the quality of research by adopting strategies (e.g. electronic systems and artificial intelligence) to check reporting accuracy in scientific papers.

Conflicts of interest

The authors have no financial or personal relationships that could cause a conflict of interest regarding this article.

Acknowledgements

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

Supplementary materials

Supplementary material associated with this article can be found in the online version at doi:10.1016/j.bjpt.2022.100450.

References

4. Moher D, Liberati A, Tetzlaff J, Altman DG. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA