Methodological quality and reporting standards in systematic reviews with meta-analysis of physical activity studies: a report from the Strengthening the Evidence in Exercise Sciences Initiative (SEES Initiative)

Short title: Methodological quality and reporting standards: The SEES Initiative

Nórton Luís Oliveira ^{1,2}, Cíntia Ehlers Botton ³, Angélica Trevisan De Nardi ^{2,4}, Daniel Umpierre ^{1,2,4,5}

Affiliations:

¹National Institute of Science and Technology for Health Technology Assessment (IATS/HCPA), Hospital de Clínicas de Porto Alegre, Clinical Research Center, Porto Alegre, RS, Brazil

²Exercise Pathophysiology Research Laboratory, Hospital de Clínicas de Porto Alegre, Clinical Research Center, Porto Alegre, RS, Brazil

³Universidade Federal de Pelotas, Pelotas, RS, Brazil

⁴Graduate Program in Cardiology and Cardiovascular Sciences, Universidade Federal do Rio Grande do Sul, Porto Alegre, RS, Brazil

⁵Department of Public Health, Universidade Federal do Rio Grande do Sul, Porto Alegre, RS, Brazil

Corresponding author:

Nórton Luís Oliveira

Exercise Pathophysiology Research Laboratory, Hospital de Clínicas de Porto Alegre, Clinical Research Center - Rua Ramiro Barcelos, 2350 - CEP: 90035-903 - Porto Alegre, RS, Brazil (nloedfisica@gmail.com).

This paper will be submitted to a peer-reviewed journal

Abstract

Background: Suboptimal practices have been observed in methodological conduct, transparency, and report quality of systematic reviews with meta-analysis (SRMAs) from different areas of the health sciences. To our knowledge, there is no such data on SRMAs from exercise sciences.

Objective: To assess the methodological and reporting standards in SRMAs of physical activity studies. This report presents the main results of the SEES initiative in 2019.

Design: The SEES Initiative uses a prospective systematic review methodology to implement post-publication surveillance of research practices in exercise sciences.

Data sources: During the year 2019, pre-specified searches were conducted monthly (PubMed/MEDLINE) in journals from the exercise sciences (n = 9) and medicine (n = 5). The assessments were independently conducted by two authors, based on 36 items/practices derived from established statements/tools (PRISMA, AMSTAR 2, ROBIS).

Eligibility criteria: SRMAs should summarize studies that had, at least, one arm consisting of physical activity interventions/exposures and one health or behavioral outcome. Our methodological protocol is publicly available (<https://osf.io/2cu8g/>).

Results: Out 1028 studies assessed for eligibility, 103 SRMAs were included. The minimum adherence was 13/36 items, whereas only one adhered to all items. Some highly contemplated items included: identification of title as SRMA (97.1%) and descriptions of the main outcome in the abstract (95.1%) and risk of bias (RoB) assessment (95.1%). Some poorly contemplated items included: publicly available protocol (4.9%), discussion of the results in light of RoB in studies included (32.0%) and data sharing statements (35.9%).

Conclusion: There is a suboptimal adherence to recommended practices on methodological quality and reporting standards in the SRMAs of physical activity intervention/exposure evaluated from the selected journals in 2019, which likely reduce the reproducibility and usefulness of these studies.

Keywords: systematic reviews with meta-analysis, methodological quality, reporting standards, exercise sciences

1 Introduction

2 Over the last decades, there has been a consistent increase in the publication of
3 systematic reviews with or without meta-analysis (SRMAs) in the health sciences [1–3]. This
4 type of literature review is accepted as a valuable source of evidence [4–6], being highly cited
5 [7]. Additionally, SRMAs are used to inform clinical decision-making, development of clinical
6 practice guidelines and health policies [2,8]. In the health sciences, numerous SRMAs have
7 been poorly conducted with low methodological quality and poor reporting [2,9–11]. In the
8 exercise sciences, SRMAs and clinical trials are essential to inform strategies of prevention
9 and treatment in several clinical conditions [12,13], however, methodological and
10 reproducibility issues have been scarcely addressed [14–16]. Such evidence of low-quality
11 practices affects reproducibility, since it makes findings less verifiable, credible, and
12 informative [17], with likely waste of financial or non-financial resources [18].

13 Several strategies have been developed to improve the quality of evidence syntheses,
14 such as open handbooks for standardized methods [19], registry platforms (e.g., International
15 prospective register of systematic review - PROSPERO), data sharing policies, reporting
16 guidelines and methodological quality assessment tools [20]. In this context, the “Preferred
17 Reporting Items for Systematic Reviews and Meta-Analyses Statement” (PRISMA Statement)
18 [21], a reporting guideline, and the “A Measurement Tool to Assess systematic Reviews”
19 (AMSTAR 2) [22], a methodological quality assessment tool, are carefully conceived
20 documents to improve the quality of SRMAs. However, reproducibility requires
21 methodological rigor, adequate reporting and transparency in the design, conduction,
22 analysis and critical appraisal of scientific evidence [18,23], denoting that adherence to
23 existent resources is of foremost importance. Since adherence is still quite low [1,2],
24 advocacy, education/training, implementation and surveillance are issues to be addressed in
25 a more intense and in-depth way [1,2,24,25].

26 Thus, we have developed and implemented a discipline-based initiative with a
27 surveillance approach to evaluate the practices from the publications in exercise sciences.
28 Specifically, our purpose was to prospective and systematically assess methodological and
29 reporting standards of systematic reviews with meta-analyses of physical activity from
30 selected journals in the field of exercise sciences and general medicine, as well as to
31 disseminate our findings publicly and provide direct feedback to authors and journal editors.
32 Therefore, the present report brings the results of our assessments in 2019.

Methods

Design and organization

The Strengthening the Evidence in Exercise Sciences Initiative (SEES Initiative) is a collaborative, nonprofit, living project for assessment of published research (i.e., randomized controlled trials and SRMAs) and dissemination of recommended practices of transparency, reproducibility and integrity. This project was launched in January 2019 and the results to be presented herein refer to the SRMAs assessments from January to December 2019. Our protocol is available at the Open Science Framework repository (OSF) [26].

The SEES Initiative is conducted by trained collaborators organized in three committees. The pre-assessment committee is responsible for literature searches and retrieval. The assessment committee carries out the assessment of eligibility criteria and conducts data extraction on randomized controlled trials and SRMAs. The post-assessment committee conducts the data management (i.e., analyses, storage and sharing) and dissemination of evidence appraisal. To ensure internal consistency, planned tasks of each committee follow standardized operational procedures (<https://osf.io/efgvv/>).

Our methodological design relates to a meta-research prospective approach, mostly regarding post-publication analyses and, therefore, submission to ethical committees is not applicable for this study design.

Eligibility criteria

We included SRMAs which synthesized primary studies having at least one intervention arm of physical activity or exercise programs with a research question related to health outcomes or health behaviors at a minimal extent. In addition, SRMAs of observational studies with well defined physical activity exposures were eligible for inclusion, as well as reviews with exploratory approaches (e.g., meta-regression) using meta-analytic techniques.

Literature search

We conducted searches in PubMed/MEDLINE between the 3rd to 7th day of each month of 2019. To reduce the burden with eligibility analyses, after conducting pilot tests and examining specific literature, search strategies (Supplementary material 1) were built using highly-sensitivity filters for SRMAs [27,28]. Moreover, we applied date filters restricting the searches for the two past months. For example, February was included in the search

1 conducted both in March as well as in April. As expected, this strategy resulted in duplicated
2 references because each month was looked up twice. Since MEDLINE/PubMed does have
3 some timing variability in indexing of references, we implemented such overlapping
4 procedure because non-overlapped searches yielded loss of some references in our pilot
5 tests.

6 The searches were conducted and retrieved on a monthly basis in 9 exercise science
7 journals (American Journal of Sports Medicine, British Journal of Sports Medicine, European
8 Journal of Preventive Cardiology, International Journal of Behavioral Nutrition and Physical
9 Activity, Journal of Physiotherapy, Journal of Science and Medicine in Sport, Medicine and
10 Science in Sports and Exercise, Scandinavian Journal of Medicine & Science in Sports and
11 Sports Medicine) and on a quarterly basis from 5 general medicine journals (Annals of Internal
12 Medicine, British Medical Journal, Journal of the American Medical Association, Lancet and
13 New England Journal of Medicine. The choice of these exercise science journals was guided
14 by the audience reach, for which we considered whether a given journal was linked to a
15 professional or scientific society, journal impact factor, and by the expected volume of SRMAs
16 publications. Regarding the general medicine journals our choice was based on the public
17 reach and media coverage that articles addressing physical activity exposures/interventions
18 usually achieve whenever published by these journals. The protocol document presents
19 further reasoning regarding the journals' choice [29].

21 **Screening**

22 Two trained collaborators (N.L.O and C.E.B) composed the assessment committee for
23 SRMAs and screened independently the title/abstracts for eligibility. The same collaborators
24 read the full-texts of potential eligible SRMAs for a final selection. Any disagreements were
25 solved by a third collaborator (D.U.). Full data set with publications assessed for eligibility is
26 publicly available on OSF (<https://osf.io/6jzb8/>).

28 **Data extraction (SRMAs assessment)**

29 We set up a standardized form to assess the methodological and reporting quality of
30 the SRMAs, which was completed independently by two collaborators (N.L.O and C.E.B).
31 Disagreements were resolved with participation of a third collaborator (D.U.). The assessment
32 form was constructed based on documents widely available and referred to in the literature.

The questions were therefore operationalized based on recommendation items from the PRISMA Statement [21], the AMSTAR 2 [22], and the “Risk of Bias in Systematic Reviews” (ROBIS) [30]. Supplementary material 2 presents the items included in the assessment form and respective guideline/tools of reference.

In this report, methodological quality refers to how well SRMA was planned and conducted, taking into account the processes of analysis, synthesis and critical appraisal of the selected evidence [22]. In addition, the quality of reporting refers to the details and completeness of the information throughout the research report [21].

Items and domains

The SRMAs (unit of analysis) were analysed by (1) cumulative number of items (from 0 to 36) in accordance to recommended practices, and (2) seven domains that aggregate these items based on contributions in methodological or reporting aspects. The seven domains were: transparency (n=4 items), completeness (n=11), methodological rigor (n=7), participants (n=2), interventions/exposures (n=2), outcome (n=5), and critical appraisal (n=5). Although the domains are intrinsically related, items were distributed in an exclusive approach to avoid overlapped counts (which is only present for one item associated with two domains in the SRMA assessment) (Supplementary material 2). We underscore that the “domains” are labelled as “components” both on the website and in the SEES initiative protocol [29].

Dissemination of individual results

After SRMAs assessments were completed, we disseminated reports by three means: (1) summarized reports at our website (www.sees-initiative.org); (2) deposit of full reports were filed in OSF (<https://osf.io/ntw7d/> folder: assessment reports/SRMA); (3) full reports sent by email to the corresponding authors. We encouraged authors to send us requests for clarification or correction.

Analyses

A general summary is presented in median and interquartile range (25th-75th percentiles). Data for each item evaluated are presented as absolute frequencies (n) and

percentages of SRMA publications that received each possible answer (for example, yes or no). Regarding the domains, the data are summarized as frequencies and percentages of SRMAs that scored as “Yes” or a “positive” evaluation in a specific number of items in each individual domain. In most items, the option “Yes” was indicative of a recommended practice (item) that was followed in the SRMA. Items that did not have a binary (yes/no) response option included those related to the time range of literature search, duplicated process for selection, extraction, and assessment of risk of biases (yes vs. partially yes), number of languages considered for study eligibility, description of risk of bias within studies, description for protocol deviations during the course of the review, discussion limitations at the study/outcome and/or at the review level. Another type of exception regards the occurrence of a potential spin bias, in which “No” is the positive (recommended) evaluation. We describe these exceptions in detail and their possible operationalization in Supplementary material 3.

Although we did compute a cumulative frequency of adherence to recommended practices per study, we chose not to disseminate study results primarily based on scores, grades, or single index, because the use of such metrics can hide critical flaws in important items for reproducibility, credibility and trustworthiness of the SRMA findings [22]. In this sense, we attempt to communicate such frequencies by also showing all the items included in the assessment.

The data was extracted and summarized in a Google Sheet generated by filling out our Google Form.

Results

A total of 1028 records were retrieved from the electronic literature searches performed monthly in 2019. Nine hundred and eighty five were screened by title and abstract, and 882 were excluded. Thus, after confirmation by full-text screening, 103 SRMAs were eligible for assessment (Figure 1). Only one SRMA included was published in general medicine journals. Most publications were from Sports Medicine (n=36) and British Journal of Sports Medicine (n=34) (Supplementary material 4).

Figure 1. Flow-diagram of study selection process

Overall, one SRMA addressed appropriately the 36 core-items (Figure 2). The median of items that received “Yes” or other “positive” evaluation was 28 items (24-29 items). The item regarding the identification as a systematic review, meta-analysis, or both in the study title (completeness domain) was the one with the highest adherence (100/103, 97.1%). On the other hand, the item regarding the existence of a publicly available methodological protocol (transparency domain) was the one with the lowest adherence (5/103, 4.9%) (Table 1).

(Studies in decreasing order of adherence to items)

Adherence ■ Yes ■ No

Figure 2. Overview of the adherence to recommended practices: items (n=36) per SRMA (n=103).

Transparency

For the four items of this domain, only 2.9% (3/103) of SRMAs received "yes" on all (full-yes) but 11.7% (12/103) received "no" on all (full-no). The item with the most positive evaluation was related to the availability of at least one full-search query (83/103, 80.6% of SRMAs) (Table 1). Just over a third of the reports made a statement regarding the data sharing (37/103, 35.9%) and just over half (57/103, 55.3%) were registered in a public database.

Completeness

Thirty SRMAs (29.1%) complied to all the 11 items of this domain. In addition, 7 items presented a high percentage of "yes" (93.2% to 97.1%) among the SRMAs. However, 31.1% (32/103) of SRMAs not reported detailed explanation of eligibility criteria for PICOS elements (Population, Intervention, Comparator, Outcome, Setting) and the abstract of 25.2% SRMAs (26/103) did not inform some of the key eligibility criteria for study selection (Table 1).

Participants and interventions/exposures

These two domains were composed of two items, one of which (i.e., full description PICOS elements in the results section) is repeated in both domains. Eighty eight (85.4%) SRMAs scored "yes" in this item. Also, 69 (67.0%) SRMAs received full-yes in the participants domain and 82 (79.6%) SRMAs in the interventions/exposures domain. However, 21.4% (22/103) of SRMAs did not provide a description regarding the population (participants) or main condition(s) addressed in the abstract section (Table 1).

Outcome

Thirty five SRMAs (34.0%) complied to all the 5 items of this domain. The item with the most positive evaluation was related to the result description for the main outcome in the abstract (98/103, 95.1%). However, almost half of SRMAs (47/103, 45.6%) failed to report a full description of individual results for studies composing the meta-analysis (i.e., effects size, imprecision measure and percentage weight) in the results section. Moreover, 31.1% (32/103) of SRMAs did not describe meta-analytic summary estimates as recommended (i.e.,

binary outcomes as frequencies with and without the event (or as proportions such as 12/45); continuous outcomes as the mean, standard deviation, and sample size for each group) (Table 1).

Methodological rigor

A full-yes was reached in 9.7% (10/103) of SRMAs in this domain of 7 items. The item with the most positive evaluation was related to the description of the risk of biases assessment in the methods section (98/103, 95.1%), but 35.0% (36/103) of SRMA did not perform this assessment in duplicate. In addition, most of SRMAs considered only one language to study eligibility (62/103, 60.2%) and almost half did not extract data in duplicate (47/103, 45.6%) (Table 1).

Critical appraisal

Fourteen SRMAs (13.6%) presented a full-yes in this domain. A high percentage of SRMAs described the risk of bias within included studies in the results section (89/103, 86.4%). However, only 32.0% (33/103) of SRMAs discussed the results in light of the risk of biases in individual studies. A description for non-planned modifications during the course of the review was not carried out in 17.5% (18/103) after comparing the published report to registration. Also, this comparison was not possible in 40.8% (42/103) of SRMAs due to the lack of registration (Table 1).

Table 1. Frequency distribution of recommended items by domains for the total number of Systematic Reviews with Meta-Analysis (SRMAs)

	SRMAs (n=103)
Domain: Transparency	
<i>Registration</i>	
No	46 (44.7%)
Yes	57 (55.3%)
<i>Protocol</i>	
No	98 (95.1%)
Yes	5 (4.9%)
<i>Available searches</i>	
No	20 (19.4%)
Yes	83 (80.6%)
<i>Data Statement</i>	

No	66 (64.1%)
Yes	37 (35.9%)

Domain: Completeness

Title as SRMA

No	3 (2.9%)
Yes	100 (97.1%)

Data sources (ab)

No	33 (32.0%)
Yes	70 (68.0%)

Key eligibility criteria (ab)

No	26 (25.2%)
Yes	77 (74.8%)

Number of included studies (ab)

No	5 (4.9%)
Yes	98 (95.1%)

Research question

No	29 (28.2%)
Yes	74 (71.8%)

PICOS explanation

No	32 (31.1%)
Yes	71 (68.9%)

Number of references

No	6 (5.8%)
Yes	97 (94.2%)

Description of sample sizes

No	5 (4.9%)
Yes	98 (95.1%)

Duration of included studies

Does not apply	16 (15.5%)
No	7 (6.8%)
Yes	80 (77.7%)

Sources of funding

No	7 (6.8%)
Yes	96 (93.2%)

Potential conflicts of interest

No	7 (6.8%)
Yes	96 (93.2%)

Domain: Participants

Description of participants (ab)

No	22 (21.4%)
Yes	81 (78.6%)

Detailed studies' characteristics

No	15 (14.6%)
Yes	88 (85.4%)

Domain: Intervention/exposure

Description of interventions/exposures (ab)

No	8 (7.8%)
----	----------

Yes	95 (92.2%)
<i>Detailed studies' characteristics</i>	
No	15 (14.6%)
Yes	88 (85.4%)
Domain: Outcome	
<i>Main outcome of interest (ab)</i>	
No	5 (4.9%)
Yes	98 (95.1%)
<i>Statistical methods</i>	
No	18 (17.5%)
Yes	85 (82.5%)
<i>Statistical heterogeneity</i>	
No	12 (11.7%)
Yes	91 (88.3%)
<i>Meta-analytic summary estimates</i>	
No	32 (31.1%)
Yes	71 (68.9%)
<i>Statistics per study</i>	
No	47 (45.6%)
Yes	56 (54.4%)
Domain: Methodological rigor	
<i>Searches in grey literature</i>	
No	49 (47.6%)
Yes	54 (52.4%)
<i>Searches from inception or with justification</i>	
No	8 (7.8%)
Yes	95 (92.2%)
<i>Number of languages</i>	
1	62 (60.2%)
2	11 (10.7%)
3	7 (6.8%)
4	1 (1.0%)
No restriction	19 (18.4%)
No statement	3 (2.9%)
<i>Study selection in duplicate</i>	
No	26 (25.2%)
Partial Yes (e.g., a sample of 50% of studies were checked by two independent researchers)	3 (2.9%)
Yes	74 (71.8%)
<i>Data extraction in duplicate</i>	
No	47 (45.6%)
Partial Yes (e.g., a sample of 50% of studies were checked by two independent researchers)	1 (1.0%)
Yes	55 (53.4%)
<i>Description of RoB assessment</i>	
No	5 (4.9%)
Yes	98 (95.1%)

RoB assessment in duplicate

No	36 (35.0%)
Yes	67 (65.0%)

Domain: Critical appraisal**RoB results within studies**

No	14 (13.6%)
Partial Yes (there are individual results without specification of specific criteria/domains)	17 (16.5%)
Yes	72 (69.9%)

Description of protocol deviations

No	18 (17.5%)
Unclear	42 (40.8%)
Yes	16 (15.5%)
Does not apply	27 (26.2%)

Presence of spin bias

No	82 (79.6%)
Yes	21 (20.4%)

Discussion addressing RoB

No	70 (68.0%)
Yes	33 (32.0%)

Limitations thoroughly addressed

No	6 (5.8%)
Yes, BOTH for study and review levels	74 (71.9%)
Yes, ONLY for the review level (limitation within or across studies not mentioned)	2 (1.9%)
Yes, ONLY for the study and/or outcome level (review processes not mentioned)	21 (20.4%)

ab: abstract; PICOS: acronym for Population, Intervention, Comparator/Control, Outcome, Setting; RoB : risk of bias

Answers to SEES emails

Nine of the 103 corresponding authors questioned the assessment of some items. The median of claims was 4 items, ranging from 1 to 10.

Discussion

This study presents the main report of methodological and reporting surveillance on SRMAs of physical activity intervention/exposure assessed by the SEES Initiative in the year of 2019. Across the seven domains operationalized to aggregate the items/practices recommended by the PRISMA Statement, ROBIS, and AMSTAR 2 tools, some of the most adopted practices were in the completeness domain: identification as a systematic review, meta-analysis, or both in the study title (97.1%) and more 2 items with 95.1%; outcome

domain: description of the main outcome of interest in the abstract (95.1%) and methodological rigor domain: description of RoB assessment (95.1%). In opposite, the existence of a publicly available methodological protocol (4.9%) (transparency domain), discussion of the results in light of RoB in individual studies (32.0%) (critical appraisal domain) and statement on data sharing (35.9%) (transparency domain) were among the practices least adhered to, therefore indicating possible emphases that can be addressed through educational resources and editorial policies.

Practices most and least present in assessed articles

In the transparency domain, modest existence of reviews registered in public databases and only a minimal fraction (5/103) presenting a detailed methodological protocol in addition to the registration may reduce the potential for methodological reproducibility [31]. Although most SRMAs retrieve disseminated and aggregated data, data sharing statements were still less than modest. There are similar results indicating low adherence to practices of registration [1,24] or data sharing [32] in other disciplines, denoting that transparent practices need more implementation in varied research fields.

We observed that most studies adhered to items under the completeness, participants and interventions/exposures domains. Taken together, these three domains included 14 items, which allow stakeholders to understand the scope of the study, the research question, its applicability and, finally, the replication of the study if desired [21]. However, our results indicate that the information on the eligibility criteria for the selection of individual studies were not adequately reported in a relevant number of SRMAs, either in abstracts (25.2%) or in methods (31.1%). Furthermore, 22.4% of SRMAs did not provide a clear description of participants or main conditions under study in the abstract section. We underscore the importance of such information since it is key for stakeholders to decide whether to access (possibly by payment) the article.

In domains of methodological rigor, outcome reporting, and critical appraisal, we combined items (practices) that could more directly affect the interpretation of evidence generated by SRMAs. Surprisingly, several reviews did lack duplicate processes in data extraction and risk of bias assessments. Although similar results were observed in meta-research studies in health sciences [1,24,33], it is widely recommended that these procedures be performed by two independent researchers or primarily conducted by one researcher and

1 fully verified by another one [22,30]. Additionally, we found that nearly two thirds of SRMAs
2 (60.3%) considered only the English language to study eligibility. Such restriction is commonly
3 applied in literature searches [34], but can introduce a systematic error, a language bias,
4 possibly modifying meta-analytic estimates [34,35]. In the outcome domain, items such as
5 individual results for each analysed study and meta-analytic summary estimates still need
6 substantial reporting improvement to facilitate the understanding and replicability of meta-
7 analyses. These items were considered as properly reported when all recommended
8 elements were addressed, therefore, our assessment may have been very strict. However, to
9 note, previous studies have also reported similar findings in the biomedical sciences [32,36].

10 In the critical appraisal domain, more than 80% of our sample presented the risk of
11 bias assessment, however, only a third (33/103) discussed their review findings in the context
12 of potential biases identified through the assessment. Part of our results are in line with data
13 by Buttner et al [37], who focused only on the risk of bias assessment, and also described a
14 high percentage of these assessments in SR and SRMAs of exercise interventions/exposures
15 published in a single journal. Moreover, studies in biomedical literature have also reported a
16 reduced use of this type of assessment in the discussion of SRMAs [9,33], therefore, in
17 agreement with our findings. In contrast, Buttner et al [37] found more favourable results,
18 indicating that risk of bias assessments were incorporated in the interpretation/discussion
19 sections of 86% of their sample. Although both studies, Buttner et al [37] and ours, have
20 distinct samples and criteria to assess this practice, we reason that, especially when risk of
21 bias is high in individual studies, such results should be thoroughly addressed in discussion
22 sections. Finally, we assessed whether SRMAs reported methodological changes during the
23 course of the review, which were not planned as described in the registry or methodological
24 protocol. These changes in design, conduct or analysis can introduce bias in the review
25 findings and therefore should be reported to the readers' knowledge [22,30]. Unfortunately,
26 our results show that 17.5% (18/103) of SRMAs made unplanned changes and did not describe
27 them in the final publication. Additionally, in 40.8% (42/103) of SRMAs it was not possible to
28 carry out this verification due to the lack of a public registration record, which again reinforces
29 the need for transparent practices at the pre-study stage.

31 **Strengths and Limitations**

Several strategies exist to improve the methodological and reporting quality of SRMAs in the biomedical literature [18,20,24,33,38–40]. Our study highlights that solely relying on the availability of these documents/tools could be insufficient for adoption. In exercise sciences there is a need for most journals to provide endorsement of the guidelines, which may trigger quality improvements [38,39]. The use of guidelines in the peer review process has also been proposed [24,33]. We point out that the main strength of this study is the “translational approach”, since we used established guidelines in a lively project, giving feedback to editors and authors (despite the low number of replies). In methodological grounds, we did not homogenize our assessments through a total score, allowing readers a more detailed view of adherence to items and domains. Indeed, generating such scores is not recommended, since it assumes that all items have the same weight [22]. Finally, we included some ROBIS items in our assessment form, therefore addressing more directly the methodological quality assessment of the SRMAs.

Some limitations should be considered. First, our sample of SRMAs represents only the year 2019 and only a few pre-selected scientific journals. Therefore, the results observed in our assessments are not promptly generalizable to other SRMA publications in exercise sciences. However, it seems unlikely that quality of reporting of previous SRMAs is superior to the current ones, given that our results are in line with several studies of this nature in other health disciplines. Second, the selected journals had features such as connection to a professional or scientific societies, considerable public reach and impact factor. Since these features increase visibility, we believe quality of studies is a common priority among the selected journals. Third, a possible limitation is the fact that our assessment form has not been formally validated. Although the form was elaborated on the basis of a reporting guideline for systematic reviews (i.e., PRISMA) and well planned methodological quality assessment tools (i.e., AMSTAR and ROBIS), testing the properties of our form would be important.

Conclusion

There was a heterogeneous pattern for adherence to reporting standards proposed by the PRISMA Statement in SRMAs in exercise sciences, with structure (completeness) at a better status than practices related to transparency and full information for adequate critical appraisal. This suboptimal (sometimes low) adherence to recommended practices on quality

1 of reporting and methodological rigor limit the reproducibility and credibility of the SRMAs
2 findings, and consequently, the evidence uptake for clinical decision-making. Therefore, the
3 implementation of strategies such as SEES, which proposes to monitor published research
4 continuously and discuss its findings openly and cordially with the authors and editors can be
5 a way to improve the synthesis of evidence in the field of exercise sciences. However, it seems
6 necessary to discuss aims, outcomes, and diversify strategies to increase engagement and
7 responsibility of all stakeholders.

8
9 **Data sharing statement:** All data generated by the SEES Initiative are publicly available via
10 Open Science Framework (<https://osf.io/ntw7d/>) under a CC-BY Attribution 4.0 International
11 license. Data sheets: Eligibility (<https://osf.io/6jzb8/>); SRMA assessments
12 (<https://osf.io/29f4a/>).

13
14 **Funding:** No funding has been specifically awarded to the SEES Initiative. The project is
15 housed at the Hospital de Clínicas de Porto Alegre (Porto Alegre, Brazil). NLO receives post-
16 doctoral fellowship funding support from the National Institute of Science and Technology for
17 Health Technology Assessment (IATS, Brazil). ADN receives doctoral funding support from the
18 Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES, Brazil – Finance Code
19 001). DU receives research productivity grant and support from the Conselho Nacional de
20 Desenvolvimento Científico e Tecnológico (CNPq foundation).

21
22 **Competing interests:** All authors have no competing interests to disclose.

Supplementary material 1

1. Search strategies for journals of exercise sciences:

MEDLINE/PubMed query for systematic review:

"Br J Sports Med"[Journal] OR "Am J Sports Med"[Journal] OR "Med Sci Sports Exerc"[Journal] OR "Eur J Prev Cardiol"[Journal] OR "Sports Med"[Journal] OR "Int J Behav Nutr Phys Act"[Journal] OR "J Physiother"[Journal] OR "J Sci Med Sport"[Journal] OR "Scand J Med Sci Sports"[Journal]

AND

(meta[tiab] OR synthesis[tiab] OR literature[tiab] OR published[tiab] OR meta-analysis[tiab] OR extraction[tiab] OR trials[tiab] OR search[tiab] OR MEDLINE[tiab] OR selection[tiab] OR sources[tiab] OR review[tiab] OR review[pt] OR articles[tiab] OR reviewed[tiab] OR english[tiab] OR language[tiab])

2. Search strategies for journals in general medicine:

MEDLINE/PubMed query for systematic review:

"Ann Intern Med"[Journal] OR "BMJ"[Journal] OR "JAMA"[Journal] OR "Lancet"[Journal] NOT ("Lancet Respir Med"[Journal] OR "Lancet Public Health"[Journal] OR "Lancet Psychiatry"[Journal] OR "Lancet Planet Health"[Journal] OR "Lancet Oncol"[Journal] OR "Lancet Neurol"[Journal] OR "Lancet Infect Dis"[Journal] OR "Lancet Haematol"[Journal] OR "Lancet HIV"[Journal] OR "Lancet Glob Health"[Journal] OR "Lancet Gastroenterol Hepatol"[Journal] OR "Lancet Diabetes Endocrinol"[Journal] OR "Lancet Child Adolesc Health"[Journal]) NOT ("Lancet Respir Med"[Journal] OR "Lancet Public Health"[Journal] OR "Lancet Psychiatry"[Journal] OR "Lancet Planet Health"[Journal] OR "Lancet Oncol"[Journal] OR "Lancet Neurol"[Journal] OR "Lancet Infect Dis"[Journal] OR "Lancet Haematol"[Journal] OR "Lancet HIV"[Journal] OR "Lancet Glob Health"[Journal] OR "Lancet Gastroenterol Hepatol"[Journal] OR "Lancet Diabetes Endocrinol"[Journal] OR "Lancet Child Adolesc Health"[Journal]) OR "N Engl J Med"[Journal]

AND

((meta[tiab] OR synthesis[tiab] OR literature[tiab] OR published[tiab] OR meta-analysis[tiab] OR extraction[tiab] OR trials[tiab] OR search[tiab] OR MEDLINE[tiab] OR selection[tiab] OR sources[tiab] OR review[tiab] OR review[pt] OR articles[tiab] OR reviewed[tiab] OR english[tiab] OR language[tiab]))

AND (exercise[tiab] OR "physical activity"[tiab] OR training[tiab] OR rehabilitation[tiab]))

Separated by journals:

"Ann Intern Med"[Journal] AND ((meta[tiab] OR synthesis[tiab] OR literature[tiab] OR published[tiab] OR meta-analysis[tiab] OR extraction[tiab] OR trials[tiab] OR search[tiab] OR MEDLINE[tiab] OR selection[tiab] OR sources[tiab] OR review[tiab] OR review[pt] OR articles[tiab] OR reviewed[tiab] OR english[tiab] OR language[tiab]) **AND (exercise[tiab] OR "physical activity"[tiab] OR training[tiab] OR rehabilitation[tiab]))**

1
2 "BMJ"[Journal] AND ((meta[tiab] OR synthesis[tiab] OR literature[tiab] OR published[tiab] OR meta-
3 analysis[tiab] OR extraction[tiab] OR trials[tiab] OR search[tiab] OR MEDLINE[tiab] OR selection[tiab]
4 OR sources[tiab] OR review[tiab] OR review[pt] OR articles[tiab] OR reviewed[tiab] OR english[tiab]
5 OR language[tiab]) AND (exercise[tiab] OR "physical activity"[tiab] OR training[tiab] OR
6 rehabilitation[tiab]))

7
8
9 "JAMA"[Journal] AND ((meta[tiab] OR synthesis[tiab] OR literature[tiab] OR published[tiab] OR meta-
10 analysis[tiab] OR extraction[tiab] OR trials[tiab] OR search[tiab] OR MEDLINE[tiab] OR selection[tiab]
11 OR sources[tiab] OR review[tiab] OR review[pt] OR articles[tiab] OR reviewed[tiab] OR english[tiab]
12 OR language[tiab]) AND (exercise[tiab] OR "physical activity"[tiab] OR training[tiab] OR
13 rehabilitation[tiab]))

14
15 "Lancet"[Journal] NOT (("Lancet Respir Med"[Journal] OR "Lancet Public Health"[Journal] OR "Lancet
16 Psychiatry"[Journal] OR "Lancet Planet Health"[Journal] OR "Lancet Oncol"[Journal] OR "Lancet
17 Neurol"[Journal] OR "Lancet Infect Dis"[Journal] OR "Lancet Haematol"[Journal] OR "Lancet
18 HIV"[Journal] OR "Lancet Glob Health"[Journal] OR "Lancet Gastroenterol Hepatol"[Journal] OR
19 "Lancet Diabetes Endocrinol"[Journal] OR "Lancet Child Adolesc Health"[Journal]) NOT ("Lancet Respir
20 Med"[Journal] OR "Lancet Public Health"[Journal] OR "Lancet Psychiatry"[Journal] OR "Lancet Planet
21 Health"[Journal] OR "Lancet Oncol"[Journal] OR "Lancet Neurol"[Journal] OR "Lancet Infect
22 Dis"[Journal] OR "Lancet Haematol"[Journal] OR "Lancet HIV"[Journal] OR "Lancet Glob
23 Health"[Journal] OR "Lancet Gastroenterol Hepatol"[Journal] OR "Lancet Diabetes
24 Endocrinol"[Journal] OR "Lancet Child Adolesc Health"[Journal]) AND (meta[tiab] OR synthesis[tiab]
25 OR literature[tiab] OR published[tiab] OR meta-analysis[tiab] OR extraction[tiab] OR trials[tiab] OR
26 search[tiab] OR MEDLINE[tiab] OR selection[tiab] OR sources[tiab] OR review[tiab] OR review[pt] OR
27 articles[tiab] OR reviewed[tiab] OR english[tiab] OR language[tiab]) AND (exercise[tiab] OR "physical
28 activity"[tiab] OR training[tiab] OR rehabilitation[tiab]))

29
30 "N Engl J Med"[Journal] AND ((meta[tiab] OR synthesis[tiab] OR literature[tiab] OR published[tiab]
31 OR meta-analysis[tiab] OR extraction[tiab] OR trials[tiab] OR search[tiab] OR MEDLINE[tiab] OR
32 selection[tiab] OR sources[tiab] OR review[tiab] OR review[pt] OR articles[tiab] OR reviewed[tiab] OR
33 english[tiab] OR language[tiab]) AND (exercise[tiab] OR "physical activity"[tiab] OR training[tiab] OR
34 rehabilitation[tiab]))

Supplementary material 2

Full assessment of SRMA with individual items, associated component labels (i.e., domains) and, whether applicable, the guideline or tool of reference. The items of interest to be assessed in SRMAs were selected based on key information that is (i) suitable to be presented in systematic reviews with exercise intervention or exposures related to physical activity; (ii) relevant to allow methodological reproducibility; and (iii) important to describe how evidence is weighted in by authors in order to provide well-reasoned knowledge translation.

Component	Item	Reference (guideline, checklist item; unless otherwise stated)
Transparency	Registration: Is the review registered in a public database?	PRISMA, 2 and 5 AMSTAR 2, 2
Transparency	Protocol: Is there referral of a publicly available methodological protocol? (Note: if so, you must also use the protocol to consult information about the review)	PRISMA, 5 AMSTAR 2, 2
Transparency	Methods: Is there at least one search query fully available? (Note: a full search query should allow complete replication)	PRISMA, 8
Transparency	Is there a statement regarding the data availability (data sharing plan)?	Not applicable
Completeness	Title: Is the study identified as a systematic review, meta-analysis, or both?	PRISMA, 1
Completeness	Abstract: Does the abstract list the data sources used in the review? (Note: if more than five databases were used, simplified or partial referral should be considered as "Yes")	PRISMA, 2
Completeness	Abstract: Does the abstract inform key eligibility criteria for study selection?	PRISMA, 2
Completeness	Abstract: Is there a description of the number of included studies?	Not applicable
Completeness	Introduction: Is there a description for the research question (with PICOS elements) or precisely stated objectives (with PICOS) ?	PRISMA, 4 AMSTAR 2, 1
Completeness	Methods: Is there a detailed explanation of eligibility criteria for PICOS elements? (Note: detailed explanation should allow complete replication)	PRISMA, 6 ROBIS, 1.3
Completeness	Results: Is there a full description regarding the numbers of references (retrieval, eligibility, synthesis)?	PRISMA, 17

Completeness	Results: Is there a description about the sample sizes of individual studies?	PRISMA, 18
Completeness	Results: Is there a description of study duration (follow-up lengths)?	PRISMA, 18 ROBIS, 3.2
Completeness	Is there a statement regarding the sources of funding? (Note: funding for the review itself)	PRISMA, 27 CONSORT, 25
Completeness	Did the review authors declare whether they had any conflicts of interest (COI)?	AMSTAR 2, 16
Participants	Abstract: Is there a description regarding the population (participants) or main condition(s) addressed in the review?	PRISMA, 2
Intervention / Exposure	Abstract: Is there a description regarding the interventions/exposures (or, broadly, independent variables) addressed in the review?	PRISMA, 2
Participants; Intervention / Exposure	Results: Is there a full description of characteristics? (Note: the available PICOS elements should be considered)	PRISMA, 18 AMSTAR 2, 8 ROBIS, 3.2
Outcome	Abstract: Is there a result description for the main outcome of interest?	PRISMA, 2
Outcome	Methods: Is there a description of the statistical combination (meta-analysis) regarding the effect measure (e.g., relative risk or mean difference), statistical method (e.g., inverse variance), and effects approach (fixed or random)?	PRISMA, 14 ROBIS, 3.3
Outcome	Methods: Is there a description regarding the assessment of statistical heterogeneity?	PRISMA, 14
Outcome	Results: Is there a minimally recommended description of meta-analytic summary estimates? (Note: binary outcomes as frequencies with and without the event (or as proportions such as 12/45); continuous outcomes as the mean, standard deviation, and sample size for each group)	PRISMA, 20, 21
Outcome	Results: Is there a full description of individual results for studies composing the meta-analysis? (Note: effects size, imprecision measure and percentage weight should be considered)	PRISMA, 20
Methodological rigor	Methods: Did the search strategy include non-published evidence? ("grey literature")	PRISMA, 7 AMSTAR 2, 4 ROBIS 2.1
Methodological rigor	Methods: If the search is restricted for evidence generated after 1980, is there an indirect or direct justification related to the time range?	PRISMA, 8 ROBIS, 2.4

Methodological rigor	Methods: How many languages were considered for study eligibility?	PRISMA, 6 ROBIS, 2.4
Methodological rigor	Methods: Was the study selection carried out in duplicate?	PRISMA, 10 AMSTAR 2, 5 ROBIS, 2.5
Methodological rigor	Methods: Was the data extraction carried out in duplicate?	PRISMA, 11 AMSTAR 2, 6
Methodological rigor	Methods: Is there a description of the assessment of risk of biases?	PRISMA, 12 ROBIS, 3.4
Methodological rigor	Methods: Was the assessment of risk of biases carried out in duplicate?	ROBIS, 3.5
Critical appraisal	Results: Is there a description of risk of bias within studies? (Note: your assessment should be based on the characteristic of the RoB tool)	PRISMA, 22
Critical appraisal	Results: Is there a description for non-planned modifications to the synthesis during the course of the review? (e.g.: change in eligibility criteria or RoB tools; please, what was changed and its justification [why] should be considered).	AMSTAR 2, 2 ROBIS, 4.2
Critical appraisal	Discussion: Is there a potential spin bias based on a specific reporting strategy to highlight that the experimental treatment (or condition of interest) leads to the hypothesized result?	Based on Boutron et al [41] ROBIS, C (Phase 3)
Critical appraisal	Discussion: Are the results discussed in light of the risk of biases in individual studies?	PRISMA, 25 AMSTAR 2, 13 ROBIS, 4.6 and A (Phase 3)
Critical appraisal	Discussion: Are limitations discussed at the study/outcome and/or at the review level?	PRISMA, 25 AMSTAR 2, 12 ROBIS, A (Phase 3)
Components are coded by colors, except the item related to full description of participants and intervention (white background), which accounts twice for both Participants and Interventions/Exposures.		

- 1 COI: conflict of interest; PICOS: acronym for Population, Intervention, Comparator/Control, Outcome,
- 2 Setting; RoB: risk of bias.
- 3

Component label	Number of items
Transparency	4

Completeness	11
Participants	2
Intervention / Exposure	2
Methodological rigor	7
Outcome	5
Critical appraisal	5
Note: As a same item (relating to both Intervention/Exposure and Participants components) accounts twice, the final sum totalizes 36 items.	

1

2

Supplementary material 3

Exceptions and operationalization of the answers of some items

Items where we considered more than one answer, the answer “No” or another answer as being "positive" (i.e., recommended practice):

1. Methods: If the search is restricted for evidence generated after 1980, is there an indirect or direct justification related to the time range? (Methodological rigor)

The possibilities considered “positive” were: “Yes” or “*The searches were carried out from the database inception (earliest date)*”. Therefore, these two possibilities were combined as “Yes”, when the result of this question was presented in binary form (i.e., yes or no).

2. Methods: Was the study selection carried out in duplicate? (Methodological rigor)

Methods: Was the data extraction carried out in duplicate? (Methodological rigor)

Methods: Was the assessment of risk of biases carried out in duplicate? (Methodological rigor)

The possibilities considered “positive” were: “Yes” or “*Partial Yes (e.g., a sample of 50% of studies were checked by two independent researchers)*”. Therefore, these two possibilities were combined as “Yes”, when the result of this question was presented in binary form (i.e., yes or no).

3. Results: Is there a description of risk of bias within studies? (Note: your assessment should be based on the characteristic of the RoB tool) (Critical appraisal)

The possibilities considered “positive” were: “Yes (*FULL description*)” or “*Partial Yes (there are individual results without specification of specific criteria/domains)*”.

Therefore, these two possibilities were combined as “Yes”, when the result of this question was presented in binary form (i.e., yes or no).

1 **4. Results: Is there a description for non-planned modifications to the synthesis during**
2 **the course of the review? (e.g.: change in eligibility criteria or RoB tools; please, what**
3 **was changed and its justification [why] should be considered) (Critical appraisal)**

4 The possibilities considered “positive” were: “Yes” or “Does not apply”. Here “Does
5 not apply” means that after comparing the article to the register or protocol it was not
6 identified any non-planned modification to be described in the main text. Therefore,
7 these two possibilities were combined as “Yes”, when the result of this question was
8 presented in binary form (i.e., yes or no).

10 **5. Results: Is there a description of study duration (follow-up lengths)? Completeness**

11 The possibilities considered “positive” were: “Yes” or “Does not apply”. Here “Does
12 not apply” was the possibility of choice when the SRMA assessed was from
13 observational studies. Therefore, these two possibilities were combined as “Yes”,
14 when the result of this question was presented in binary form (i.e., yes or no).

16 **6. Methods: How many languages were considered for study eligibility?**
17 **(Methodological rigor)**

18 The possibilities considered “positive” were: “No restriction” or “more than 1
19 language”. Therefore, these two possibilities were combined as “Yes”, when the result
20 of this question was presented in binary form (i.e., yes or no).

22 **7. Discussion: Is there a potential spin bias based on a specific reporting strategy to**
23 **highlight that the experimental treatment (or condition of interest) leads to the**
24 **hypothesized result? (Critical appraisal)**

25 The possibility considered “positive” was the answer “No”.

27 **8. Discussion: Are limitations discussed at the study/outcome and/or at the review**
28 **level? (Critical appraisal)**

29 There were three possibilities of “Yes”: “Yes, BOTH for study and review levels”; “Yes,
30 ONLY for the review level (limitation within or across studies not mentioned)”; “Yes,
31 ONLY for the review level (limitation within or across studies not mentioned)”. The only
32 one considered as “positive” was: “Yes, BOTH for study and review levels”.

Item where we considered more than one answer as “No”:

9. **Results: Is there a description for non-planned modifications to the synthesis during the course of the review? (e.g.: change in eligibility criteria or RoB tools; please, what was changed and its justification [why] should be considered) (Critical appraisal)**

The answer “Unclear” was combined with “No” when the result of this question was presented in binary form (i.e., yes or no). Here, “Unclear” was the possibility of choice when the SRMA assessed did not have a record or protocol to compare with the study report and verify the existence of non-planned modifications.

1 **Supplementary material 4**

2

3 **Table. Number of SRMAs assessed by journal**

Journal	SRMAs (n)
American Journal of Sports Medicine	4
British Journal of Sports Medicine	34
European Journal of Preventive Cardiology	6
International Journal of Behavioral Nutrition and Physical Activity	7
Journal of Physiotherapy	3
Journal of Science and Medicine in Sport	6
Medicine and Science in Sports and Exercise	1
Scandinavian Journal of Medicine & Science in Sports	5
Sports Medicine	36
The British Medical Journal	1
Total	103

4 Journals without SRMAs assessed: Annals of Internal Medicine, Journal of the American

5 Medical Association, Lancet, New England Journal of Medicine

References

1. Pussegoda K, Turner L, Garritty C, Mayhew A, Skidmore B, Stevens A, et al. Systematic review adherence to methodological or reporting quality. *Syst Rev*. 2017;6: 131.
2. Page MJ, Shamseer L, Altman DG, Tetzlaff J, Sampson M, Tricco AC, et al. Epidemiology and Reporting Characteristics of Systematic Reviews of Biomedical Research: A Cross-Sectional Study. *PLoS Med*. 2016;13: e1002028.
3. Bastian H, Glasziou P, Chalmers I. Seventy-five trials and eleven systematic reviews a day: how will we ever keep up? *PLoS Med*. 2010;7: e1000326.
4. Howick J, Chalmers I, Glasziou P, Greenhalgh T, Heneghan C, Liberati A, Moschetti I, Phillips B, and Thornton H. Explanation of the 2011 Oxford Centre for Evidence-Based Medicine (OCEBM) Levels of Evidence (Background Document). In: Oxford Centre for Evidence-Based Medicine [Internet]. [cited 12 Jul 2020]. Available: <http://www.cebm.net/index.aspx?o=5653>
5. Harbour R, Miller J. A new system for grading recommendations in evidence based guidelines. *BMJ*. 2001. pp. 334–336. doi:10.1136/bmj.323.7308.334
6. Krist AH, Wolff TA, Jonas DE, Harris RP, LeFevre ML, Kemper AR, et al. Update on the Methods of the U.S. Preventive Services Task Force: Methods for Understanding Certainty and Net Benefit When Making Recommendations. *Am J Prev Med*. 2018;54: S11–S18.
7. Patsopoulos NA, Analatos AA, Ioannidis JPA. Relative citation impact of various study designs in the health sciences. *JAMA*. 2005;293: 2362–2366.
8. Moher D, Tetzlaff J, Tricco AC, Sampson M, Altman DG. Epidemiology and Reporting Characteristics of Systematic Reviews. *PLoS Medicine*. 2007. p. e78. doi:10.1371/journal.pmed.0040078
9. Xu C, Liu Y, Jia P-L, Li L, Liu T-Z, Cheng L-L, et al. The methodological quality of dose-response meta-analyses needed substantial improvement: a cross-sectional survey and proposed recommendations. *J Clin Epidemiol*. 2019;107: 1–11.
10. Brugha TS, Matthews R, Morgan Z, Hill T, Alonso J, Jones DR. Methodology and reporting of systematic reviews and meta-analyses of observational studies in psychiatric epidemiology: Systematic review. *British Journal of Psychiatry*. 2012. pp. 446–453. doi:10.1192/bjp.bp.111.098103
11. Adie S, Ma D, Harris IA, Naylor JM, Craig JC. Quality of Conduct and Reporting of Meta-analyses of Surgical Interventions. *Annals of Surgery*. 2015. pp. 685–694. doi:10.1097/sla.0000000000000836
12. Pedersen BK, Saltin B. Exercise as medicine - evidence for prescribing exercise as therapy

- in 26 different chronic diseases. *Scandinavian Journal of Medicine & Science in Sports*. 2015. pp. 1–72. doi:10.1111/sms.12581
13. Bull FC, Al-Ansari SS, Biddle S, Borodulin K, Buman MP, Cardon G, et al. World Health Organization 2020 guidelines on physical activity and sedentary behaviour. *Br J Sports Med*. 2020;54: 1451–1462.
 14. Adherence to Consolidated Standards of Reporting Trials (CONSORT) Guideline Items in Randomized Trials of Physical Activity Published in 5 Sports Medicine Journals | Peer Review Congress. [cited 12 Jul 2020]. Available: <https://peerreviewcongress.org/prc17-0401>
 15. Silverman SR, Schertz LA, Yuen HK, Lowman JD, Bickel CS. Systematic review of the methodological quality and outcome measures utilized in exercise interventions for adults with spinal cord injury. *Spinal Cord*. 2012;50: 718–727.
 16. Abell B, Glasziou P, Hoffmann T. Reporting and replicating trials of exercise-based cardiac rehabilitation: do we know what the researchers actually did? *Circ Cardiovasc Qual Outcomes*. 2015;8: 187–194.
 17. Goodman SN, Fanelli D, Ioannidis JPA. What does research reproducibility mean? *Science Translational Medicine*. 2016. pp. 341ps12–341ps12. doi:10.1126/scitranslmed.aaf5027
 18. Glasziou P, Altman DG, Bossuyt P, Boutron I, Clarke M, Julious S, et al. Reducing waste from incomplete or unusable reports of biomedical research. *Lancet*. 2014;383: 267–276.
 19. Website. [cited 23 Dec 2020]. Available: Website. Higgins JPT, Thomas J, Chandler J, Cumpston M, Li T, Page MJ, Welch VA (editors). *Cochrane Handbook for Systematic Reviews of Interventions* version 6.1 (updated September 2020). Cochrane, 2020. Available from www.training.cochrane.org/handbook. (accessed 20 Dec 2020)
 20. Munafò MR, Nosek BA, Bishop DVM, Button KS, Chambers CD, Percie du Sert N, et al. A manifesto for reproducible science. *Nat Hum Behav*. 2017;1: e124.
 21. Liberati A. The PRISMA Statement for Reporting Systematic Reviews and Meta-Analyses of Studies That Evaluate Health Care Interventions: Explanation and Elaboration. *Annals of Internal Medicine*. 2009. p. W. doi:10.7326/0003-4819-151-4-200908180-00136
 22. Shea BJ, Reeves BC, Wells G, Thuku M, Hamel C, Moran J, et al. AMSTAR 2: a critical appraisal tool for systematic reviews that include randomised or non-randomised studies of healthcare interventions, or both. *BMJ*. 2017. p. j4008. doi:10.1136/bmj.j4008
 23. Ioannidis JPA, Greenland S, Hlatky MA, Khoury MJ, Macleod MR, Moher D, et al. Increasing value and reducing waste in research design, conduct, and analysis. *Lancet*. 2014;383: 166–175.
 24. Sharma S, Oremus M. PRISMA and AMSTAR show systematic reviews on health literacy and cancer screening are of good quality. *J Clin Epidemiol*. 2018;99: 123–131.

- 1 25. Blanco D, Altman D, Moher D, Boutron I, Kirkham JJ, Cobo E. Scoping review on
2 interventions to improve adherence to reporting guidelines in health research. *BMJ*
3 *Open*. 2019;9: e026589.
- 4 26. Umpierre D, De Nardi AT, Botton CE, Helal L, Santos LP, Oliveira NL. SEES-Initiative
5 Protocol. In: *Open Science Framework* [Internet]. 17 Apr 2019 [cited 8 Jan 2021].
6 doi:10.17605/OSF.IO/NTW7D
- 7 27. Boynton J, Glanville J, McDaid D, Lefebvre C. Identifying systematic reviews in MEDLINE:
8 developing an objective approach to search strategy design. *Journal of Information*
9 *Science*. 1998. pp. 137–154. doi:10.1177/016555159802400301
- 10 28. Lee E, Dobbins M, Decorby K, McRae L, Tirilis D, Husson H. An optimal search filter for
11 retrieving systematic reviews and meta-analyses. *BMC Med Res Methodol*. 2012;12: 51.
- 12 29. Umpierre D, De Nardi AT, Botton CE, Helal L, Santos LP, Oliveira NL. SEES-
13 Initiative_Protocol_2019-04-15. 2019. doi:10.17605/OSF.IO/NTW7D
- 14 30. Whiting P, Savović J, Higgins JPT, Caldwell DM, Reeves BC, Shea B, et al. ROBIS: A new
15 tool to assess risk of bias in systematic reviews was developed. *J Clin Epidemiol*. 2016;69:
16 225–234.
- 17 31. Stewart L, Moher D, Shekelle P. Why prospective registration of systematic reviews
18 makes sense. *Systematic Reviews*. 2012. doi:10.1186/2046-4053-1-7
- 19 32. Page MJ, Altman DG, Shamseer L, McKenzie JE, Ahmadzai N, Wolfe D, et al. Reproducible
20 research practices are underused in systematic reviews of biomedical interventions. *J*
21 *Clin Epidemiol*. 2018;94: 8–18.
- 22 33. Nissen T, Wayant C, Wahlstrom A, Sinnett P, Fugate C, Herrington J, et al. Methodological
23 quality, completeness of reporting and use of systematic reviews as evidence in clinical
24 practice guidelines for paediatric overweight and obesity. *Clin Obes*. 2017;7: 34–45.
- 25 34. Grégoire G, Derderian F, Le Lorier J. Selecting the language of the publications included
26 in a meta-analysis: Is there a tower of babel bias? *Journal of Clinical Epidemiology*. 1995.
27 pp. 159–163. doi:10.1016/0895-4356(94)00098-b
- 28 35. Pham B, Klassen TP, Lawson ML, Moher D. Language of publication restrictions in
29 systematic reviews gave different results depending on whether the intervention was
30 conventional or complementary. *Journal of Clinical Epidemiology*. 2005. pp. 769–776.e2.
31 doi:10.1016/j.jclinepi.2004.08.021
- 32 36. Wasiak J, Tyack Z, Ware R, Goodwin N, Faggion CM Jr. Poor methodological quality and
33 reporting standards of systematic reviews in burn care management. *Int Wound J*.
34 2017;14: 754–763.
- 35 37. Büttner F, Winters M, Delahunt E, Elbers R, Lura CB, Khan KM, et al. Identifying the
36 “incredible”! Part 2: Spot the difference - a rigorous risk of bias assessment can alter the
37 main findings of a systematic review. *British Journal of Sports Medicine*. 2020. pp. 801–

808. doi:10.1136/bjsports-2019-101675

38. Panic N, Leoncini E, de Belvis G, Ricciardi W, Boccia S. Evaluation of the endorsement of the preferred reporting items for systematic reviews and meta-analysis (PRISMA) statement on the quality of published systematic review and meta-analyses. *PLoS One*. 2013;8: e83138.
39. Cullis PS, Gudlaugsdottir K, Andrews J. A systematic review of the quality of conduct and reporting of systematic reviews and meta-analyses in paediatric surgery. *PLoS One*. 2017;12: e0175213.
40. Blanco D, Schroter S, Aldcroft A, Moher D, Boutron I, Kirkham JJ, et al. Effect of an editorial intervention to improve the completeness of reporting of randomised trials: a randomised controlled trial. *BMJ Open*. 2020;10: e036799.
41. Boutron I, Dutton S, Ravaud P, Altman DG. Reporting and interpretation of randomized controlled trials with statistically nonsignificant results for primary outcomes. *JAMA*. 2010;303: 2058–2064.