

Title: Quality of systematic reviews on physiotherapy interventions for musculoskeletal disorders is critically low: a meta-epidemiological study

Authors:

1. Nicola Ferri, PT, MSc, Department of Biomedical and Neuromotor Sciences (DIBINEM), Alma Mater Studiorum University of Bologna, Bologna, Italy; Division of Occupational Medicine, IRCCS University Hospital of Bologna S Orsola-Malpighi Polyclinic, Bologna, Italy, nicola.ferri11@unibo.it
2. Elisa Ravizzotti, PT, MSc, Department of Neuroscience, Rehabilitation, Ophthalmology, Genetics and Maternal Child Health (DINO GMI), University of Genoa, Genoa, Italy, elisa.ravizzotti@gmail.com
3. Alessandro Bracci, PT, MSc, Department of Quality of Life Sciences, University of Bologna, Rimini, Italy, alessandro.bracci7@unibo.it
4. Giulia Carreras, PhD, Oncologic Network, Prevention and Research Institute (ISPRO), Florence, Italy, giulia.carreras@gmail.com
5. Paolo Pillastrini, PT, MSc, Department of Biomedical and Neuromotor Sciences (DIBINEM), Alma Mater Studiorum University of Bologna, Bologna, Italy; Division of Occupational Medicine, IRCCS University Hospital of Bologna S Orsola-Malpighi Polyclinic, Bologna, Italy, paolo.pillastrini@unibo.it
6. Mauro Di Bari, MD, PhD, Department of Experimental and Clinical Medicine, University of Florence, Florence, Italy; Unit of Geriatrics, Department of Medicine and Geriatrics, Azienda Ospedaliero-Universitaria Careggi, Florence, Italy, mauro.dibari@unifi.it

Corresponding author:

Name	Nicola Ferri
Department	Department of Biomedical and Neuromotor Sciences (DIBINEM)
Institution	University of Bologna
Country	Italy
Email	nicola.ferri11@unibo.it

ABSTRACT

Question: How good is the quality of systematic reviews on the effectiveness of physiotherapy for musculoskeletal conditions? Are there any factors associated with quality?

Design: This is a meta-epidemiological study on systematic reviews with meta-analysis (SR-MA) of randomised controlled trials (RCT).

Methods: MEDLINE, Cochrane Database of Systematic Reviews (CDSR), CINAHL, and PEDro were searched for SR-MA of RCT on physiotherapy in musculoskeletal disorders in the last ten years. Two independent researchers screened and extracted the records and analysed the full-texts. The quality of SR-MA was quantified with AMSTAR-2 tool on a sample of 100 studies, randomly selected from the records retrieved. Disagreements were solved by consensus.

Results: The number of eligible publications increased over the past ten years. However, the methodological quality was critically low in as many as 90% of the studies retrieved and did not increase with time. The last author's H-index was the only quality predictor among the variables analysed.

Conclusion: The methodological quality of the SR-MA of RCT is unacceptably low. Given the frequent application of physiotherapy in musculoskeletal disorders, there is an urgent need to improve secondary research by adopting more rigorous methods.

Keywords

Musculoskeletal Diseases; Physical Therapy; Systematic Review; Evidence-Based Medicine

Registration: Open Science Framework (<https://osf.io/bc8zw/>).

INTRODUCTION

Systematic reviews with meta-analysis (SR-MA) are essential to Evidence-Based Medicine, as they are considered the best synthesis of intervention studies – in particular, of randomised clinical trials (RCT) – and a fundamental source of evidence for guideline developers and policymakers.¹ The number of this type of studies has been dramatically increasing in recent years, up to an estimate of 80 new SR-MA published daily.²

The area of physiotherapy has clinical, social, and economic relevance within the scientific literature. It has been estimated that one-third of all the people in the world will need physiotherapy at some point in their lives.³ Out of the variety of conditions that can benefit from physiotherapy, the most frequent indication to physiotherapy is represented by musculoskeletal disorders, which affect 1.71 billion people (95% CI 1.68 to 1.80) across the world, with low back pain being the leading condition.³

Despite an increasing information overload, little is known about the overall quality of SR-MA and the impact of their results and conclusions on clinical practice. This can undermine the progress and credibility of research and be an obstacle to reducing the gap between researchers and clinicians. Recent methodological studies, analysing SR-MA quality in specific medical areas, found that quality was low or critically low in several, diverse fields, such as treatments for Alzheimer's disease (82.4%),⁴ acupuncture (99.1%),⁵ Chinese herbal medicine (99.3%),⁶ and surgical adverse events (100%).⁷

The "Publish or Perish" paradigm is one of the elements behind quantity over quality.⁸ For this reason, there is still debate about the strengths and limitations of current bibliometrics in representing the quality and impact of scientific research in academics.⁹⁻¹² We have recently conducted a survey confirming that more than 50% of decision-makers have difficulty choosing the best evidence on a given topic: the methodological quality, the reputation of the authors and the journal, and the type of primary studies included were among the features thought to be important by respondents.¹³

The research questions were:

1. How good is the quality of SR-MA of RCT on the effectiveness of physiotherapy for musculoskeletal conditions published in the last ten years?
2. Is the quality associated with any publication factor?

METHOD

Design

This is an observational study in the physiotherapy evidence synthesis field. Since there are no specific guidelines for meta-epidemiological studies, and the Methodological STudy reporting Checklist (MISTIC) is still under development,¹⁴ we followed the Cochrane Handbook for guidance in the selection and extraction processes,¹⁵ and the PRISMA 2020 statement for reporting, as applicable.¹⁶ This meta-epidemiological study was prospectively registered on the Open Science Framework (<https://osf.io/bc8zw/>).

Population

Our study focuses on musculoskeletal conditions defined and categorised by the International Classification of Diseases 11th Revision (ICD-11).¹⁷

Intervention

We included all types of treatment that physiotherapists in any part of the world can administer, considering the heterogeneity of norms and cultures: thus, acupuncture, dry needling, yoga, and percutaneous TENS were included, among other approaches. All studies on prevention in healthy subjects were excluded.

Control

We included all types of comparisons, because our objective was to study the methodological quality, not to perform a synthesis of treatment effectiveness.

Outcome measures

For the same reason, we had no interest in considering only specific outcomes. The only restriction applied was to studies broadly addressing the clinical effectiveness of physiotherapy: thus, economic studies were excluded.

Study design

We included only SR-MA of RCT because they are considered as the best summary of the available scientific evidence, a claim which justifies their high impact on clinical practice in questions concerning treatment effectiveness. SR-MA incorporating also non-randomised intervention studies were included only if a meta-analysis of RCT was reported separately. We excluded network meta-analyses, scoping reviews, rapid reviews, and other forms of evidence synthesis different from

Cochrane's definition of SR.¹⁵

Searching

We searched MEDLINE (via PubMed), Cochrane Database of Systematic Reviews (via Cochrane Library), CINAHL (via EBSCOHost), and PEDro, from December 2012 to December 2022; the search strategy included both structured terms and free texts linked by logical operators (see Appendix 1 on the eAddenda). Then, a semi-automated deduplication via EndNote^a was performed, followed by a manual check. One record was randomly selected if the same paper was published in multiple journals. The strategy for full-text retrieval included contacting authors of the manuscript when necessary.

Screening

Two independent researchers selected records by Title and Abstracts, based on the inclusion and exclusion criteria as described above. The same process was repeated on the full-texts retrieved. Any disagreement was solved by consensus.

Sampling

We randomised all the full-texts using an Excel^b function: a final sample of 100 SR-MA was deemed as adequate for the purposes of the study.

Extraction

A researcher extracted data of interest (Table 1) from the SR-MA included, using a predefined Excel^b form. A second reviewer then checked all the extractions; conflicts were solved by consensus.

Table 1. Extraction data

<i>General variables</i>	<i>Authors variables</i>	<i>Journal variables</i>	<i>Study variables</i>
First author	Country	Journal IF ^c	Type of SR-MA ^d
Year of publication ^c	Number of authors ^c	Journal Q (Scimago) ^c	Publication policy ^d
Study title	First author h-index ^c	Journal Q (JCR) ^c	Total studies included ^c
Journal name	Last author h-index ^c		Study design included ^d
			Results ^d
			Conclusions ^d

IF: Impact Factor, Q: Quartile, JCR: Journal Citation Report, SR-MA: Systematic Review with Meta-Analysis.
c: continuous data, d: dichotomous data.

Type of SR-MA identifies Cochrane vs. non-Cochrane review. Journals' impact factor (IF) was assessed at the time of the publication via the Clarivate Journal Citation Report and the same was done for journals' Quartile via Scimago and Clarivate. When a journal was classified as belonging to different discipline areas, the quartile of the discipline closer to physiotherapy was considered. First and last authors' h-indexes were assessed via Scopus. The journal's access policy was determined by consulting the Directory of Open Access Journals, Clarivate Journal Citation Report, Scopus, the journals' site, and the possibility of a full-text free download. The number of the studies included refers to the papers included in the SR, not only in meta-analyses. In extracting results and conclusions, only the primary outcome was considered; if more than one primary outcome was present or no primary outcome was mentioned, we extracted the first that was reported. We also considered the results favorable if a statistical significance was obtained and the conclusions favorable if a recommendation toward the intervention was given. Country refers to the corresponding author's primary affiliation; we categorised this data into the continent for statistical analysis.

AMSTAR-2 assessment

A Measurement Tool to Assess Systematic Reviews - version 2 (AMSTAR-2) is a critical appraisal tool for systematic reviews, consisting of 16 items on methodological domains with a yes/no and partial yes rating.¹⁸ Items number 2, 4, 7, 9, 11, 13, and 15 are considered as critical domains and, therefore, they are weighted more in the final quality assessment. The psychometric characteristics of AMSTAR-2 have been studied and validated, comparing it with pre-existing tools.¹⁹

Three researchers performed deep training following the AMSTAR-2 publication, guidance, and online educational support. A piloting test was completed before the assessments, and there was a discussion with AMSTAR-2's authors on some judgment calls to enhance the most accurate interpretation and consistency. Then, two independent researchers assessed each SR included, and any conflict was solved by consensus. A summary score was not calculated, whereas, as recommended by Shea et al.,¹⁸ four levels of quality (i.e., critically low, low, moderate, and high) were identified.

Data analysis

Data were analysed with descriptive statistics: we used mean and standard deviation for continuous data (or median and interquartile range for non-normal distribution), counts and percentages for categorical data. All the variables were grouped according to the AMSTAR-2 score, and statistical

tests were performed to analyse any significant differences.

With the aim of preserving the ordinal distribution of the four AMSTAR-2 levels, an ordinal logistic regression analysis models was built to identify predictors of SR-MA quality among the variables extracted. Statistical analyses were performed with Stata 17 software^c.

RESULTS

The search strategy retrieved 2,067 records, of which 151 were duplicates, and 2 were retracted papers. Thus, 1,914 records were screened by title and abstract, leading to 475 full texts assessed for eligibility. Finally, 395 SR-MA were included (Figure 1), of which a random sample of 100 was obtained. We reported details of the studies included (see Appendix 2 on the eAddenda), and the reasons for full-text exclusions (see Appendix 3 on the eAddenda).

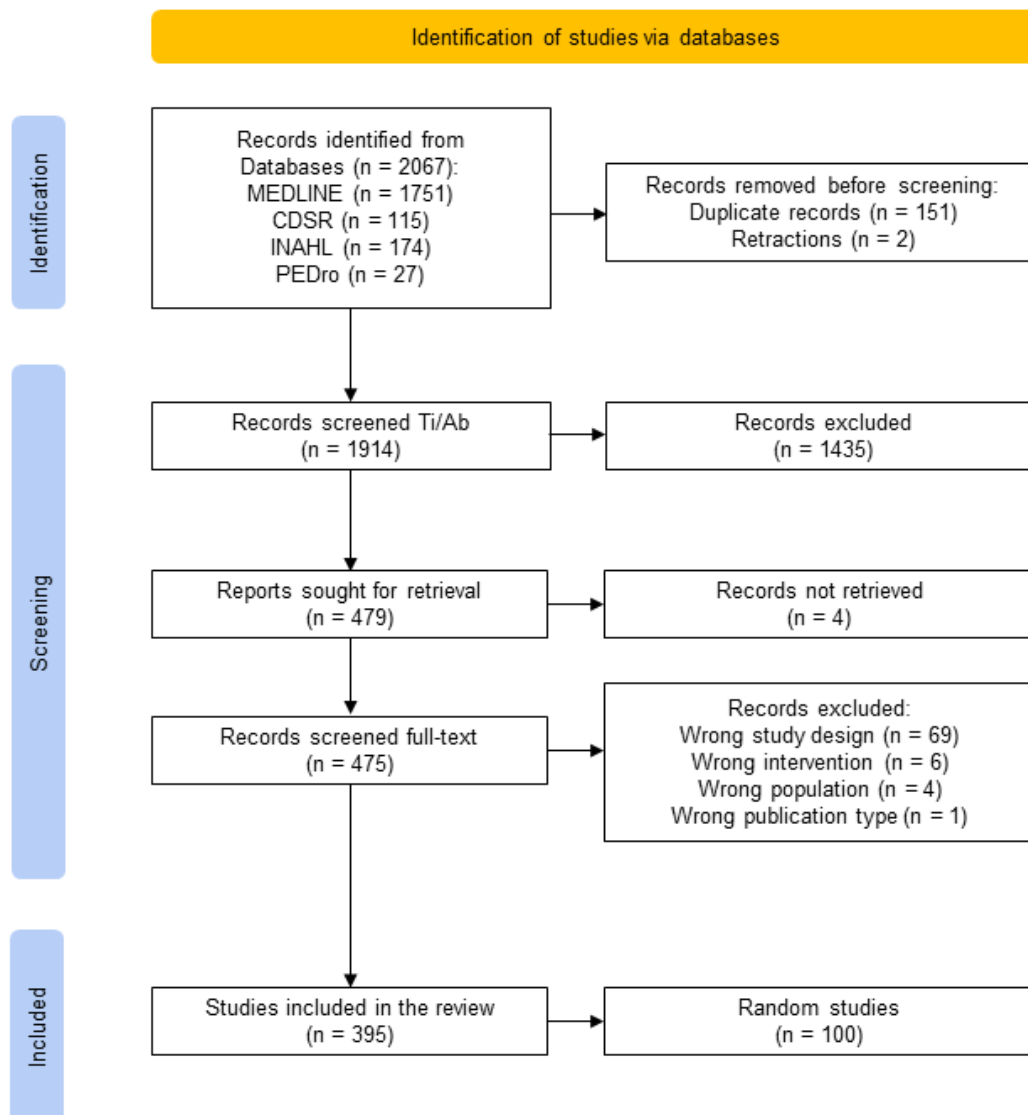


Figure 1. PRISMA flow diagram.

Overall, as many as 90% of the studies selected were judged as of critically low quality. Moreover, although the number of publications increased markedly over the past ten years, the methodological quality of the studies remained always markedly unsatisfactory (Table 2).

Table 2. Descriptive statistics

		AMSTAR-2					
		Critically low	Low	Moderate	High	p-value	
Total		90	4	2	4		
<i>Journal variables</i>							
Year of publication, n (%)							
2013	4	3 (75)	0 (0)	1 (25)	0 (0)	0.200*	
2014	4	3 (75)	0 (0)	0 (0)	1 (25)		
2015	5	5 (100)	0 (0)	0 (0)	0 (0)		
2016	10	9 (90)	0 (0)	1 (10)	0 (0)		
2017	12	11 (91.7)	0 (0)	0 (0)	1 (8.3)		
2018	12	10 (83.3)	1 (8.3)	0 (0)	1 (8.3)		
2019	13	11 (84.6)	1 (7.7)	0 (0)	1 (7.7)		
2020	12	10 (83.3)	2 (16.7)	0 (0)	0 (0)		
2021	13	13 (100)	0 (0)	0 (0)	0 (0)		
2022	15	15 (100)	0 (0)	0 (0)	0 (0)		
Journal IF, median (IQR)	2.99 (2.36-4.77)	2.90 (2.18-4.06)	4.14 (3.09-17.55)	6.10 (5.94-6.26)	7.25 (6.39-7.82)	0.004**	
Quartile (Scimago)							
1	68	58 (85.3)	4 (5.9)	2 (2.9)	4 (5.9)	0.828*	
2	21	21 (100)	0 (0)	0 (0)	0 (0)		
3	9	9 (100)	0 (0)	0 (0)	0 (0)		
4	1	1 (100)	0 (0)	0 (0)	0 (0)		
Quartile (JCR)							
1	50	41 (82)	3 (6)	2 (4)	4 (8)	0.668*	
2	26	25 (96.1)	1 (3.9)	0 (0)	0 (0)		
3	11	11 (100)	0 (0)	0 (0)	0 (0)		
4	5	5 (100)	0 (0)	0 (0)	0 (0)		
<i>Authors variables</i>							
Country, n (%)							
Africa	1	1 (100)	0 (0)	0 (0)	0 (0)	0.935*	
Asia	27	26 (96.3)	1 (3.7)	0 (0)	0 (0)		
Central America	2	2 (100)	0 (0)	0 (0)	0 (0)		
Europe	34	29 (85.3)	2 (5.9)	2 (5.9)	1 (2.9)		
North America	6	5 (83.3)	0 (0)	0 (0)	1 (16.7)		
Oceania	14	13 (92.9)	0 (0)	0 (0)	1 (7.1)		
South America	14	12 (85.7)	1 (7.1)	0 (0)	1 (7.1)		
Number of authors, mean (SD)	5.56 (1.80)	5.42 (1.82)	6.50 (1.00)	7.5 (0.71)	6.75 (0.96)	0.326***	
H-index first author, median (IQR)	4.50 (2.00-8.00)	4.00 (2.00-8.00)	16.50 (9.50-24.00)	18.50 (2.00-35.00)	6.50 (5.00-8.00)	0.045**	

H-index last author, median (IQR)		16.50 (8.00-30.50)	14.50 (7.00-28.00)	49.00 (26.00-84.50)	33.50 (29.00-38.00)	31.50 (22.50-46.00)	0.014**
<i>Study variables</i>							
Systematic review type							
	non-Cochrane	90	86 (95.6)	4 (4.4)	0 (0)	0 (0)	<0.001*
	Cochrane	10	4 (40)	0 (0)	2 (20)	4 (40)	
Publication Type							
	non-Open Access	62	55 (88.7)	3 (4.8)	1 (1.6)	3 (4.8)	0.867*
	Open Access	38	35 (92.1)	1 (2.6)	1 (2.6)	1 (2.6)	
Number of studies, median (IQR)		11.50 (7.00-19.50)	12.00 (7.00-20.00)	15.00 (7.5-33.00)	8.00 (3.00-13.00)	8.00 (6.50-19.00)	0.677**
Type of studies							
	also non-RCTs	15	12 (80)	2 (13.3)	1 (6.7)	0 (0)	0.083*
	only RCT	85	78 (91.8)	2 (2.4)	1 (1.2)	4 (4.7)	
Results favorable to experimental treatment							
	no	40	36 (90)	2 (5)	0 (0)	2 (5)	0.644*
	yes	60	54 (90)	2 (3.3)	2 (3.3)	2 (3.3)	
Conclusions favorable to experimental treatment							
	no	33	28 (84.8)	2 (6.1)	0 (0)	3 (9.1)	0.184*
	yes	67	62 (92.5)	2 (3)	2 (3)	1 (1.5)	

* chi-squared test; ** Kruskal-Wallis test; *** one-way ANOVA test.

IQR: interquartile range, SD: standard deviation, JCR: Journal Citation Report.

As shown in Table 2, every continent was represented in our sample, yet with different distributions. After exclusion of 8 journals that had no IF at the publication date, the IF ranged from 0.445 to 30.313 and the differences between the four AMSTAR-2 levels were statistically significant. The H-index of the first and last authors ranged from 0 to 39 and 1 to 100, respectively; these variables had a significantly different distribution according to the AMSTAR-2 quality scores. The only four studies rated as high quality were all Cochrane SR-MA.

Analysing the single items (Figure 2), 93% of the studies did not explain the reason for the eligibility criteria of study designs (item 3), 78% did not report the list of the studies excluded (item 7), and 90% did not check funding sources of the primary studies included (item 10).

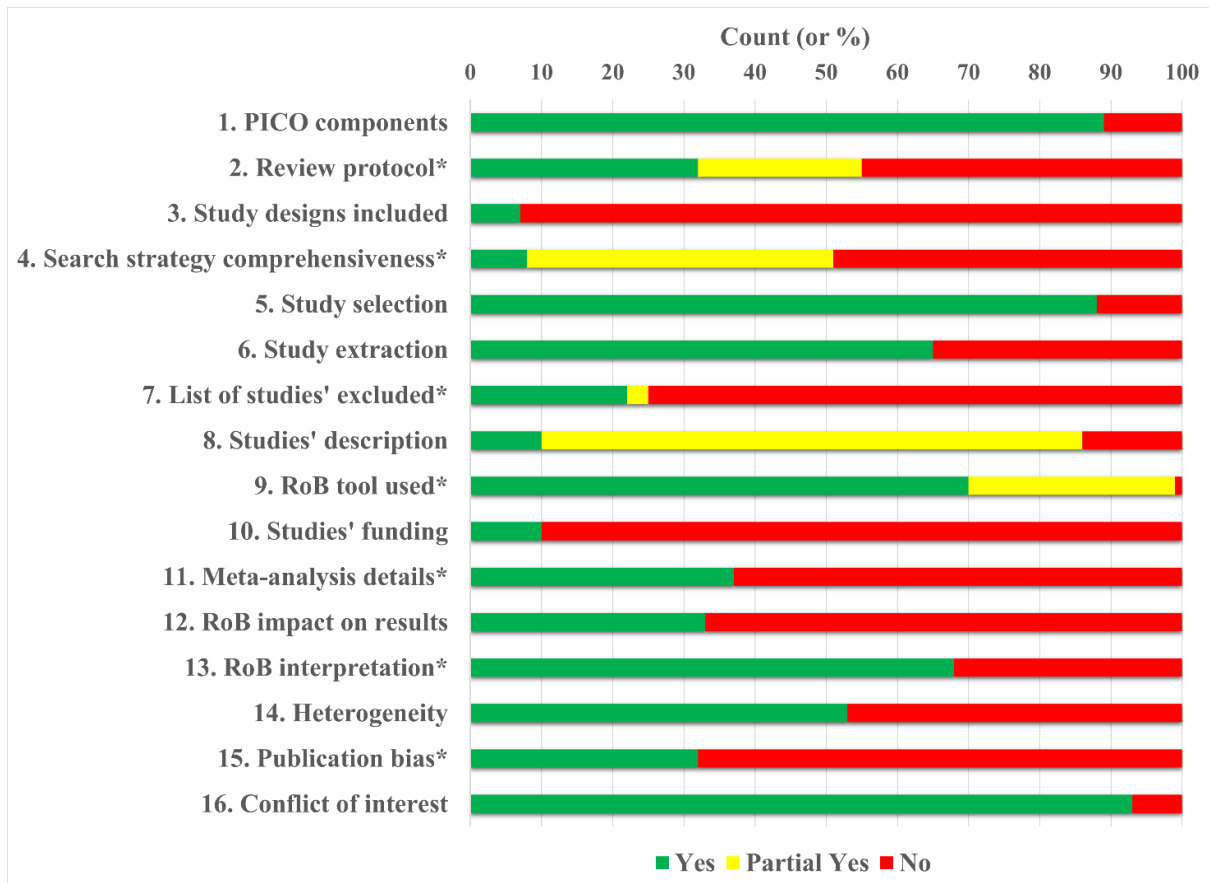


Figure 2. AMSTAR-2 items assessment. Asterisks identify the seven critical domains.

Ordered logistic regression

An ordered logistic regression model for analysing AMSTAR-2 scores was built (Table 3). Covariates that satisfied the proportional odds assumption were identified, and the best fitting model in terms of Akaike Information Criterion was selected. Results are displayed as proportional odds ratios.

Table 3. Multivariable ordinal logistic regression predicting the methodological quality of 100 systematic reviews with meta-analysis of randomised clinical trials on musculoskeletal physiotherapy.

Variable	OR (95% CI)
Journal IF	1.00 (0.99-1.01)
H-index last author	1.04 (1.01-1.06)
Number of authors	1.53 (0.99-2.39)
Intercept 1	5.8
Intercept 2	6.44
Intercept 3	6.88

OR: odds ratio, CI: confidence interval.

The journal's IF was not a predictor of quality. Conversely, the odds of a high AMSTAR-2 level versus the combined critically low, low, and moderate AMSTAR-2 level were 1.04 times greater for each unit increase in the H-index of the last author, assuming that all other variables in the model were held constant. Because of the proportional odds assumption, the same increase of 1.04 times would represent the risk of being in the combined AMSTAR-2 high and moderate categories versus the critically low and low ones. Increasing the number of authors by one resulted in 53% higher odds for high AMSTAR-2 than the other categories, with a borderline significant association.

DISCUSSION

In this representative sample of SR-MA of RCT on musculoskeletal physiotherapy published in the last ten years, the methodological quality, assessed with a standardised and recognised tool, was critically low in as many as 90% of the studies retrieved and did not show any time-dependent increase. Of the candidate predictors, only the last author's H-index predicted the quality level of the SR-MA.

The practice of healthcare professions implies an obligation of means toward the patient, which, in turn, translates into acting with prudence and diligence. To achieve this, both for clinical and legal purposes healthcare providers should mandatorily apply the best scientific evidences available. Yet, what if the evidence is of low quality?

This meta-research strengthens the findings of Riley et al.²⁰, who reported low quality on a sample of 24 SRs on a similar topic. We suggest being cautious about the evidence synthesis quality, which could be a critical issue in implementing Evidence-Based Medicine in musculoskeletal physiotherapy, both in choosing the best evidence and for the risk of not properly considering the quality of conduction during knowledge translation.

Sometimes, assessing methodological quality is challenging, because it also depends on the quality of reporting; our difficulty is supported by a recent study that found a significant association between the quality of reporting and the risk of bias of SR.²¹ In our study, these challenges were overcome thanks to the extensive training and piloting preliminary conducted and the consensus-based decision-making process. Retrospectively, our choices were consistent with the latest published recommendations.²² We also found a tendency for a floor effect of the AMSTAR-2 tool, as described in a previous research:²³ thus, one might wonder whether indeed the SR-MA were of low quality or the instrument was too rigid and demanding. However, a high standard is definitively

desirable and, after all, we did find some high-quality reviews, a finding that clearly indicate that such levels of methodological quality can be achieved. Thus, we argue that methodological improvement is a priority.

The H-index of the last author was the only variable with a predictive value for quality. This finding is consistent with the last author being usually a senior researcher, responsible for the choice and training of collaborators and for the overall scientific architecture of the study. In a way, it also represents a small comfort on the goodness of the index, knowing all its recent limitations and criticalities within academia.⁹⁻¹²

Main strengths of this study are the comprehensive and rigorous selection of the studies to analyse, the application of a recognised, standardised tool for quality assessment, and the use of a statistical approach that accurately reflects the ordinal nature of the outcome variable. At the same time, some limitations must be acknowledged. First, in the expectation of a possible time-dependent trend in methodological quality, we included a broad and representative number of studies published in the last ten years: however, the AMSTAR-2 tool was published only in 2017, more recently than the oldest studies we selected. Nevertheless, the tool was designed to assess quality, not as guidance to conduction of a SR; moreover, a previous version of AMSTAR and the Cochrane Handbook were both already available before 2017.^{24,25} Year of publication trended to be negatively correlated with quality, although this variable could not be included in the multivariable ordinal model due to a violation of the assumption of proportionality. Second, to optimise resources, we did not analyse the entire sample of 395 SR-MA, but only a share of them. However, because the sample was drawn at random and included more than a quarter of the studies retrieved, we are confident of its representativeness and, consequently, of the validity of our results, which, indeed, are well consistent with previous studies. Last, the open-access variable was attributed not to the journal, but rather to each specific article, considering the free access to it in the year of publication. We realise that the dichotomisation open-access yes/no does not fully capture the heterogeneity of publication policies and may result in less sensitive data for this variable.

Evidence-based medicine involves considering patient expectations and seeking the best scientific evidence to support clinical reasoning and decision-making. This research alerts us to the very low quality of SR-MA on treatment effectiveness for such a largely prevalent condition as musculoskeletal disorders, which represent the first indication to physiotherapy in the world. As a result, our findings may directly impact on important issues, such as the delivery of optimal care to

patients and the liability of healthcare providers. A better adherence to the most rigorous scientific methods is imperative to improve the quality of scientific research and, thus, its impact on clinical practice.

What was already known on this topic: Systematic reviews have a fundamental role in summarising evidence on a given topic, thus providing guidance for best practices and guideline development. However, the methodological quality of these SR-MA on musculoskeletal conditions is uncertain.

What this study adds: Over the past ten years, nine out of ten systematic reviews are of critically low quality; this directly impacts the Evidence-Based Practice of physiotherapists and brings the urgency of improving the quality of secondary literature. The h-index of the last author is the only predictor of quality in our sample of 100 SRs.

FOOTNOTES

Funding: This research received no specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Competing interests: Authors have no conflict of interest to declare.

Acknowledgments: We would like to thank Dr Beverly Shea for her feedback on some AMSTAR-2 details.

CRedit author statement

Conceptualisation: all authors; methodology: NF, MDB, GC; investigation: NF, ER, AB; data curation: NF, ER, AB, GC; formal analysis: GC, MDB; supervision: PP, MDB; writing – original draft: NF, MDB; writing – review and editing: all authors.

^a EndNote Web, Clarivate, Philadelphia, USA

^b Microsoft Office Professional Plus 2021: Microsoft Excel. Microsoft.

^c StataCorp. 2021. Stata Statistical Software: Release 17. College Station, TX, USA: StataCorp LLC

REFERENCES

1. Clarke J. What is a systematic review? *Evid Based Nurs*. 2011;14(3):64-64.
doi:10.1136/ebn.2011.0049
2. Hoffmann F, Allers K, Rombey T, et al. Nearly 80 systematic reviews were published each day: Observational study on trends in epidemiology and reporting over the years 2000-2019. *J Clin Epidemiol*. 2021;138:1-11. doi:10.1016/j.jclinepi.2021.05.022
3. Cieza A, Causey K, Kamenov K, Hanson SW, Chatterji S, Vos T. Global estimates of the need for rehabilitation based on the Global Burden of Disease study 2019: a systematic analysis for the Global Burden of Disease Study 2019. *Lancet Lond Engl*. 2021;396(10267):2006-2017.
doi:10.1016/S0140-6736(20)32340-0
4. Zhong CCW, Zhao J, Wong CHL, et al. Methodological quality of systematic reviews on treatments for Alzheimer's disease: a cross-sectional study. *Alzheimers Res Ther*. 2022;14(1):159.
doi:10.1186/s13195-022-01100-w
5. Ho L, Ke FYT, Wong CHL, et al. Low methodological quality of systematic reviews on acupuncture: a cross-sectional study. *BMC Med Res Methodol*. 2021;21(1):237.
doi:10.1186/s12874-021-01437-0
6. Cheung AKL, Wong CHL, Ho L, Wu IXY, Ke FYT, Chung VCH. Methodological quality of systematic reviews on Chinese herbal medicine: a methodological survey. *BMC Complement Med Ther*. 2022;22(1):48. doi:10.1186/s12906-022-03529-w
7. Zhou X, Li L, Lin L, Ju K, Kwong JSW, Xu C. Methodological quality for systematic reviews of adverse events with surgical interventions: a cross-sectional survey. *BMC Med Res Methodol*. 2021;21(1):223. doi:10.1186/s12874-021-01423-6
8. Siegel MG, Brand JC, Rossi MJ, Lubowitz JH. "Publish or Perish" Promotes Medical Literature Quantity Over Quality. *Arthrosc J Arthrosc Relat Surg Off Publ Arthrosc Assoc N Am Int Arthrosc Assoc*. 2018;34(11):2941-2942. doi:10.1016/j.arthro.2018.08.029
9. Curry S. Let's move beyond the rhetoric: it's time to change how we judge research. *Nature*. 2018;554(7691):147-147. doi:10.1038/d41586-018-01642-w
10. Hunter PR. Bibliometrics, research quality, and neglected tropical diseases. *The Lancet*. 2009;373(9664):630-631. doi:10.1016/S0140-6736(09)60397-4
11. Singh HP. Alternative research bibliometrics: It's about quality and not quantity. *Shoulder Elb*. 2022;14(2):121-122. doi:10.1177/17585732211058453
12. Weightman AL, Butler CC. Using bibliometrics to define the quality of primary care research. *BMJ*. 2011;342:d1083. doi:10.1136/bmj.d1083
13. Decision makers find it difficult to compare and select similar systematic reviews based on

- quality, methods and results: a cross-sectional survey. doi:10.21203/rs.3.rs-2416773/v1
14. Lawson DO, Puljak L, Pieper D, et al. Reporting of methodological studies in health research: a protocol for the development of the Methodological Study reportIng Checklist (MISTIC). *BMJ Open*. 2020;10(12):e040478. doi:10.1136/bmjopen-2020-040478
 15. Cochrane Handbook for Systematic Reviews of Interventions. Accessed August 10, 2023. <https://training.cochrane.org/handbook/current>
 16. Page MJ, McKenzie JE, Bossuyt PM, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *BMJ*. 2021;372:n71. doi:10.1136/bmj.n71
 17. ICD-11. Accessed July 4, 2023. <https://icd.who.int/en>
 18. Shea BJ, Reeves BC, Wells G, 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;358:j4008. doi:10.1136/bmj.j4008
 19. Lorenz RC, Matthias K, Pieper D, et al. A psychometric study found AMSTAR 2 to be a valid and moderately reliable appraisal tool. *J Clin Epidemiol*. 2019;114:133-140. doi:10.1016/j.jclinepi.2019.05.028
 20. Riley S, Swanson BT, Sawyer SF, Brismée JM, Staysniak G. Should low-quality evidence dominate high-level evidence? A systematic review and meta-analysis of systematic reviews of musculoskeletal physical therapy interventions. *J Man Manip Ther*. 29(4):203-215. doi:10.1080/10669817.2020.1839728
 21. Innocenti T, Feller D, Giagio S, et al. Adherence to the PRISMA statement and its association with risk of bias in systematic reviews published in rehabilitation journals: A meta-research study. *Braz J Phys Ther*. 2022;26(5):100450. doi:10.1016/j.bjpt.2022.100450
 22. De Santis KK, Pieper D, Lorenz RC, Wegewitz U, Siemens W, Matthias K. User experience of applying AMSTAR 2 to appraise systematic reviews of healthcare interventions: a commentary. *BMC Med Res Methodol*. 2023;23:63. doi:10.1186/s12874-023-01879-8
 23. Lorenz RC, Matthias K, Pieper D, et al. AMSTAR 2 overall confidence rating: lacking discriminating capacity or requirement of high methodological quality? *J Clin Epidemiol*. 2020;119:142-144. doi:10.1016/j.jclinepi.2019.10.006
 24. Shea BJ, Grimshaw JM, Wells GA, et al. Development of AMSTAR: a measurement tool to assess the methodological quality of systematic reviews. *BMC Med Res Methodol*. 2007;7(1):10. doi:10.1186/1471-2288-7-10
 25. Higgins JPT, Green S. Cochrane Handbook for Systematic Reviews of Interventions. Version 5.1.0. The Cochrane Collaboration; 2011. www.handbook.cochrane.org