



ORIGINAL ARTICLE

Methodological quality of systematic reviews on treatments for Parkinson's disease: A cross-sectional study

Y. Zhang^{a,1}, Y. Lin^{b,1}, C.C.W. Zhong^{a,*}, F.F. Ho^c, I.X.Y. Wu^b, C. Mao^d, X. Yang^e, V.C.H. Chung^a

^a Jockey Club School of Public Health and Primary Care, The Chinese University of Hong Kong, Hong Kong, China

^b Xiangya School of Public Health, Central South University, Hunan, China

^c School of Chinese Medicine, The Chinese University of Hong Kong, Hong Kong, China

^d Department of Epidemiology, School of Public Health, Southern Medical University, Guangdong, China

^e State Key Laboratory of Dampness Syndrome of Chinese Medicine, The Second Affiliated Hospital of Guangzhou University of Chinese Medicine, Guangdong, China

Received 13 December 2022; accepted 6 March 2023

Available online 31 August 2024

KEYWORDS

Parkinson's disease;
Cross-sectional study;
Meta-analysis;
Bias;
Systematic reviews

Abstract

Background: Systematic reviews (SR) of high methodological quality can provide the best evidence for clinical practice. However, the methodological quality of SRs on Parkinson's disease treatments has not been evaluated comprehensively. The study aims to assess the methodological quality of a representative sample of SRs on Parkinson's disease treatments.

Methods: Four databases were searched to obtain potentially eligible SRs published between January 2016 and December 2021. A pre-designed questionnaire was used to extract the bibliographical characteristics of the included SRs. The AMSTAR-2 (Assessing the Methodological Quality of Systematic Reviews) tool was used to assess the methodological quality of SRs. Factors associated with methodological quality were assessed using multivariate regression analyses.

Results: A total of 119 eligible SRs were included and appraised. Only one SR (0.8%) was of high overall methodological quality. Four (3.4%) and 7 (5.9%) SRs were of moderate and low overall methodological quality, respectively. Among the appraised SRs, only 3 (2.5%) applied a comprehensive literature search strategy, 11 (9.2%) provided a list of excluded studies with justifications for exclusion, and 4 (3.4%) reported the sources of funding among the original studies included in the SR. Cochrane SRs and SRs published in journals with higher impact factors had relatively higher overall methodological quality.

* Corresponding author.

E-mail address: chenwenzhong@link.cuhk.edu.hk (C.C.W. Zhong).

¹ These authors contributed equally to this work.

PALABRAS CLAVE

Enfermedad de Parkinson;
Estudio transversal;
Metaanálisis;
Sesgo;
Revisiones sistemáticas

Conclusions: This study demonstrated that SRs on Parkinson's disease treatments are of low methodological quality. To enhance the quality and hence the trustworthiness of SRs, the protocols of future reviews should be designed and registered *a priori*, and researchers should conduct a comprehensive literature search, provide a list of excluded studies with justifications for exclusion, and report sources of funding for the included original studies.

© 2024 Sociedad Española de Neurología. Published by Elsevier España, S.L.U. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

Calidad metodológica de las revisiones sistemáticas sobre tratamientos para la enfermedad de Parkinson: un estudio transversal

Resumen

Antecedentes: Las revisiones sistemáticas (RS) de alta calidad metodológica pueden proporcionar la mejor evidencia para la práctica clínica. Sin embargo, la calidad metodológica de las RS sobre tratamientos de la enfermedad de Parkinson no ha sido evaluada de forma exhaustiva. El objetivo de este estudio es evaluar la calidad metodológica de una muestra representativa de RS sobre tratamientos para la enfermedad de Parkinson.

Métodos: Se realizaron búsquedas en 4 bases de datos para obtener RS potencialmente elegibles desde enero de 2016 hasta diciembre de 2021. Se utilizó un cuestionario prediseñado para extraer las características bibliográficas de las RS incluidas. Se utilizó la herramienta Assessing the Methodological Quality of Systematic Reviews (AMSTAR)-2 para evaluar la calidad metodológica de las RS. Los factores asociados con la calidad metodológica se evaluaron mediante análisis de regresión multivariable.

Resultados: Se incluyeron y evaluaron 119 RS elegibles. Solo una RS (0,8%) fue de alta calidad metodológica global. Cuatro (3,4%) y 7 (5,9%) RS fueron de calidad metodológica general moderada y baja, respectivamente. Entre las RS evaluadas, solo 3 (2,5%) aplicaron una estrategia de búsqueda bibliográfica exhaustiva, 11 (9,2%) proporcionaron una lista de estudios excluidos con justificaciones de la exclusión y 4 (3,4%) informaron de las fuentes de financiación de los estudios originales incluidos en las RS. Las RS Cochrane y las RS publicadas en revistas con factores de impacto más altos tuvieron una calidad metodológica global relativamente más alta. **Conclusiones:** Este estudio demostró que las RS sobre tratamientos de la enfermedad de Parkinson son de baja calidad metodológica. Para mejorar la calidad y, por tanto, la fiabilidad de las RS, las revisiones futuras deberían diseñar y registrar un protocolo *a priori*, realizar una búsqueda bibliográfica exhaustiva, proporcionar una lista de los estudios excluidos con las justificaciones de la exclusión e informar de las fuentes de financiación de los estudios originales incluidos.

© 2024 Sociedad Española de Neurología. Publicado por Elsevier España, S.L.U. Este es un artículo Open Access bajo la CC BY-NC-ND licencia (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

Introduction

Parkinson's disease is a progressive neurodegenerative condition characterized by bradykinesia, resting tremor (4–6 Hz), rigidity, and loss of postural reflexes. It is the second most common neurodegenerative disorder after Alzheimer disease.¹ The prevalence of Parkinson's disease increases with age. The global prevalence rate is estimated to be 0.3% in 2014, compared to 1% among those aged > 65 years and 3% among those older than 80 years.¹ Most patients initially present tremor, stiffness, and slowed movement in only one arm or one leg. Parkinson's disease caused 3.2 million disability-adjusted life-years lost and 211 296 deaths in 2016.² Most patients will lose their capacity to work as the disease advances, forcing them to deal with issues like lost pay and early retirement.³

Systematic reviews (SR) can serve as one of the most reliable sources of clinical evidence to inform practice.⁴ Since only SRs with adequate methodological quality can provide robust evidence for informing clinical practice, it is important for evidence-users to assess the methodological quality of SRs before adopting the findings in decision-making processes. Due to limitations in SR methodology, treatment effects may be over- or underestimated, resulting in misguided clinical decision-making.⁵ The Assessing the Methodological Quality of Systematic Reviews 2 (AMSTAR-2) tool is an up-to-date, validated instrument for assessing the methodological quality of SRs.⁶ While a wide range of treatments are currently prescribed in the management of Parkinson's disease, the methodological quality of SRs supporting these prescribing decisions has not been evaluated comprehensively using the current methodological benchmarks of AMSTAR-2.

This cross-sectional study aims to: (i) describe the bibliographical characteristics of SRs on treatments for Parkinson's disease, (ii) evaluate the methodological quality of SRs on treatments for Parkinson's disease using the AMSTAR-2 tool, and (iii) explore the association between bibliographical characteristics and methodological quality.

Methods

Eligibility criteria

In order to be considered eligible, SRs had to fulfill the following criteria: (i) focusing on the effect of treatments for Parkinson's disease; (ii) including only randomized controlled trials (RCTs); (iii) presenting at least one meta-analysis; and (iv) being published in English. All types of treatments, whether pharmacological or non-pharmacological, were considered eligible. Narrative reviews, network meta-analyses, overviews of SRs, protocols of SRs, and animal studies were excluded. For SRs with updated versions, the most recent version was included.

Literature search

We searched 4 databases (Cochrane Database of Systematic Reviews, MEDLINE, Embase, and PsycINFO) for potentially eligible SRs published between January 2016 and December 2021. SRs published before 2016 were excluded because out-of-date SRs were considered to be less relevant in decision-making.⁷ Detailed search strategies are listed in [Appendix 1](#). Searches were limited to the English language. Validated filters with maximized specificity for systematic reviews were applied to searches on MEDLINE, Embase, and PsycINFO.^{8–10}

Literature selection and data extraction

All retrieved citations were imported into Endnote, version X9. Titles and abstracts of retrieved citations were screened against the eligibility criteria after deduplication. Following that, the full texts of potentially eligible studies were retrieved for additional evaluation. Uncertainties were resolved by discussion with another senior reviewer (CZ). A third reviewer was consulted if disagreements persisted (VC). The bibliographical characteristics of the included SRs were extracted based on a published, pre-designed questionnaire.^{11–13} Details can be found in [Appendix 2](#).

Methodological quality assessment

The methodological quality of the included SRs was evaluated with the AMSTAR-2 tool.⁶ The 16-item AMSTAR-2 is a validated tool for assessing the methodological quality of SRs. It has been applied to evaluate the methodological quality of SRs in many different diseases, such as chronic obstructive pulmonary disease,¹⁴ osteoporosis,¹¹ mental disorders,¹⁵ and hypertension.¹² Among its 16 items, 7 (items 2, 4, 7, 9, 11, 13, and 15) were considered critical items. The remaining 9 items were considered non-critical (items 1, 3, 5, 6, 8, 10, 12, 14, and 16).

Judgements were classed as either "Yes" or "No" for items 1, 3, 5, 6, and 10–16. For items 2, 4, and 7–9, an option of "Yes", "Partial yes" or "No" could be selected. According to the different judgements of these 16 items, SRs have different numbers of critical flaws and non-critical weaknesses. The overall methodological quality of SRs was classified into 4 levels (high, moderate, low, or critically low) based on the number of critical flaws and non-critical weaknesses they presented.⁶

The appraisal process was conducted by another author (YZ) after discussion with 2 senior authors (CZ and VC). A separate author (FH) randomly sampled 10% of the included SRs and conducted an independent, blinded appraisal using the AMSTAR-2 to assess accuracy. Discrepancies were resolved by discussion between the 2 appraisers, and adjustment of appraisal results were made to all included SRs using a final, standardized rating procedure agreed by all authors ([Appendix 3](#)).

Data analysis

Results on bibliographical characteristics and overall methodological quality are presented descriptively as frequencies with percentages, or as medians with ranges, according to the data distribution. The Kruskal–Wallis test was used to analyze overall methodological quality across different bibliographical characteristics. *P*-values < .05 on the Kruskal–Wallis test were considered to indicate statistical significance.

Ordinal logistic regression was used to analyze the association between bibliographical characteristics and the overall methodological quality of SRs. *P*-values > .1 on the Pearson chi-squared and deviance test indicate good model fit.

Associations between bibliographical characteristics and individual AMSTAR-2 items were further analyzed using binary logistic regression (for items 1, 3, 5, 6, and 10–16) or multinomial logistic regression (for items 2, 4, and 7–9). For binary logistic regression, *P*-values > .1 on the Hosmer–Lemeshow test indicate good model fit. For multinomial logistic regression, *P*-values < .05 on the likelihood ratio test reflect good model fit. In all logistic regression analyses, the independent variables included 7 bibliographical characteristics: being a Cochrane review, being an updated review, publication year, number of review authors, type of treatments, location of the corresponding author, and journal impact factor (IF) before the year the SR was published. An adjusted odd ratio (AOR) with 95% confidence interval (95% CI) was calculated to measure the magnitude of the association between bibliographical characteristics and quality. All analyses were carried out using the Statistical Package for Social Sciences (SPSS), version 27.

Results

Literature selection

The literature search yielded a total of 4294 citations. After deduplication, the titles and abstracts of 2810 studies were screened. The full texts of 634 citations were further

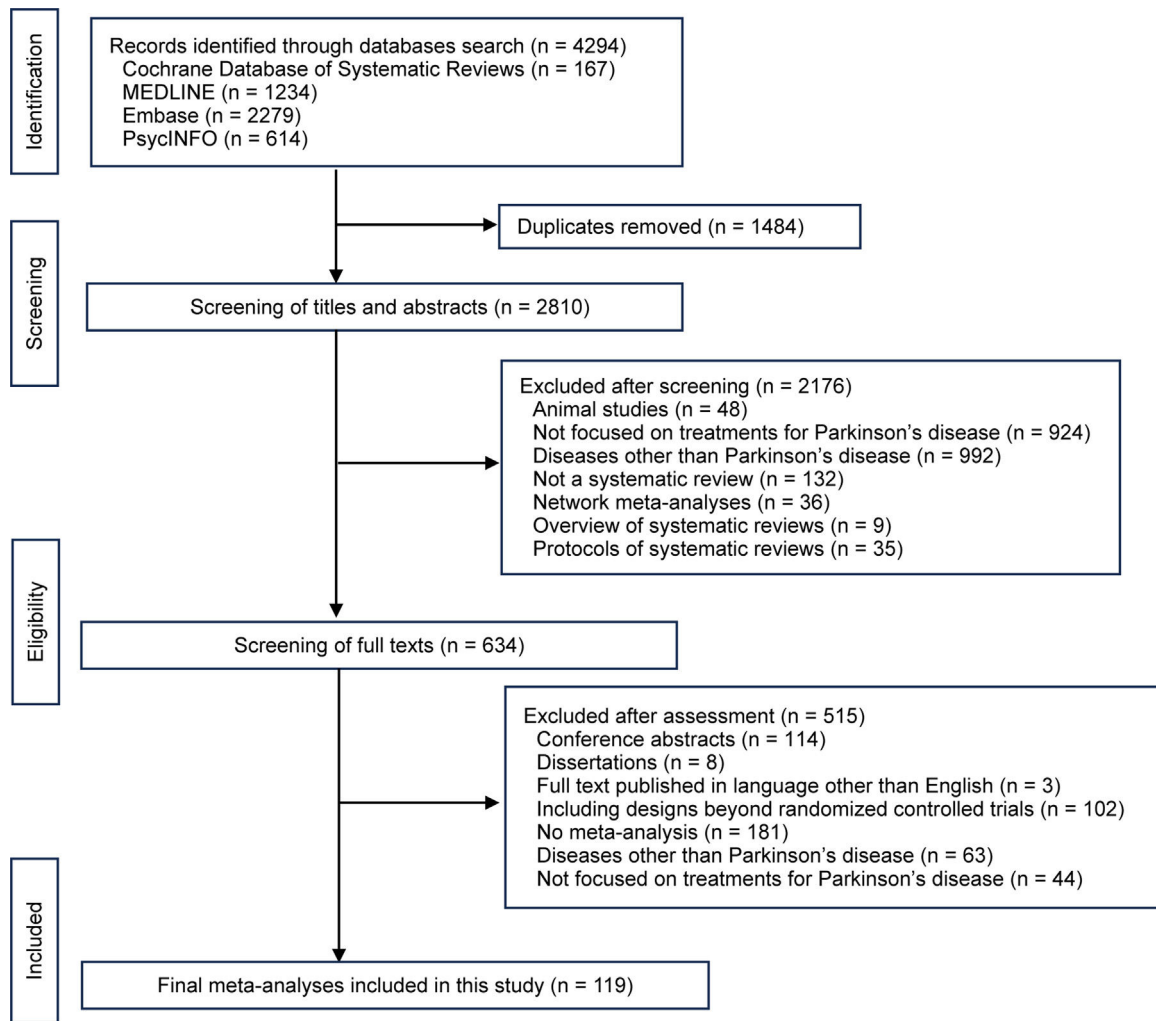


Figure 1 Flowchart of literature selection for systematic reviews on treatments for Parkinson's disease.

assessed for eligibility. The excluded SRs and the rationale for exclusion are shown in [Appendix 4](#). Finally, a total of 119 SRs were included and appraised. Details of the literature selection process are presented in [Fig. 1](#). The included SRs are listed in [Appendix 5](#).

Bibliographical characteristics of the included systematic reviews

The 119 SRs included 1648 primary studies with a total of 137 789 participants. Three (2.5%) SRs were Cochrane reviews and 7 (5.9%) were an update of a previous SR. The median publication year was 2020. The IF of the published journals ranged from 0 to 8.679, with a median of 2.776. The number of review authors of the SRs ranged from 1 to 15, with a median of 6. Seventy-eight (65.5%) corresponding authors were from Asia, 26 (21.9%) were from Europe, 10 (8.4%) were from the Americas, 4 (3.4%) were from Africa, and one (0.8%) was from Oceania. More than half (59.7%) focused on pharmacological treatments.

Thirty-seven (31.1%) SRs were funded by institutions or organizations in Asia, while 37 (31.1%) received no funding, and 29 (24.4%) did not report whether they received fund-

ing. All SRs included only RCTs. Only 77 (64.7%) reported adverse treatment effects. One hundred and sixteen (97.5%) SRs included a PRISMA-like flow diagram, and 115 (97.5%) reported at least one tool for assessing the methodological quality of primary studies.

All SRs included literature searches on English-language databases, yet only 43 (36.1%) conducted searches on non-English databases. More than half (58.0%) included primary studies in English and other languages. Twenty-one SRs (17.6%) did not report whether language restrictions were applied during literature selection. Only 2 SRs (1.7%) did not report any details of the literature search. Over half of the SRs (64.7%) used the Cochrane Risk of Bias Tool for critical appraisal, followed by the PEDro Scale (16.0%) and the Jadad Scale (6.7%). Details on bibliographical characteristics are shown in [Table 1](#).

Methodological quality

Performance on the 16 AMSTAR-2 domains

The included SRs performed poorly across most of the non-critical domains. Only 2 non-critical items were satisfied by more than 80% of SRs: ninety-six (80.7%) included PICO

Table 1 Bibliographical characteristics of the 119 included systematic reviews on treatments for Parkinson's disease.

Bibliographical characteristic	Results ^a
<i>Cochrane review</i>	3 (2.5)
<i>Update of a previous SR</i>	7 (5.9)
<i>Publication year, median (range)</i>	2020 (2016–2021)
<i>Publication journal impact factor, median (range)</i>	2.776 (0–8.679)
<i>Number of review authors, median (range)</i>	6 (1–15)
<i>Location of corresponding author</i>	
Europe	26 (21.9)
America	10 (8.4)
Asia	78 (65.5)
Oceania	1 (0.8)
Africa	4 (3.4)
<i>Type of treatment</i>	
Pharmacological	71 (59.7)
Non-pharmacological	47 (39.5)
Both	1 (0.8)
<i>Number of included primary studies</i>	
Total	1648
Median (range)	10 (2–191)
<i>Number of participants included in primary studies</i>	
Total	137 789
Median (range)	680 (67–7998)
<i>SRs reporting adverse treatment effects</i>	77 (64.7)
<i>Funding location of the SR</i>	
Europe	11 (9.2)
America	4 (3.4)
Asia	37 (31.1)
Oceania	1 (0.8)
Not reported	29 (24.4)
No funding support	37 (31.1)
<i>SRs that searched English databases</i>	119 (100)
<i>SRs that searched non-English databases</i>	43 (36.1)
<i>Reported years covered by literature search</i>	
Yes, reported both starting and ending years	67 (56.3)
Partially, only reported ending years	39 (32.8)
Not mentioned	13 (10.9)
<i>Search terms</i>	
Topics/free text/keywords/MeSH	44 (37.0)
Full Boolean	72 (60.5)
Readers are referred elsewhere for full search strategy	1 (0.8)
No search terms	2 (1.7)
<i>Language of included primary studies in SRs</i>	
English only	29 (24.4)
English and other languages	69 (58.0)
Not reported	21 (17.6)
<i>Tools for assessing quality of primary studies</i>	
Cochrane Risk of Bias Tool	77 (64.7)
Jadad scale	8 (6.7)
EPHPP	1 (0.8)
PEDro Scale	19 (16.0)
Other tools	3 (2.5)
More than one tool	7 (5.9)
No quality assessment	4 (3.4)
<i>Included a PRISMA-like flow diagram</i>	116 (97.5)

EPHPP: effective public health practice project: quality assessment tool for quantitative studies; MeSH: Medical Subject Headings; PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-analyses; SR: systematic review.

^a Values are expressed as *n* (%) or median (range). Percentages were calculated using the total number of the categories as the denominator.

Table 2 Performance on the AMSTAR-2 items for the 119 included systematic reviews on treatments for Parkinson's disease.

AMSTAR-2 items	Yes (%)	Partial yes (%)	No (%)
1. Did the research questions and inclusion criteria for the review include the components of PICO?	96 (80.7)	NA	23 (19.3)
2. Did the report of the review contain an explicit statement that the review methods were established prior to the conduct of the review and did the report justify any significant deviations from the protocol? ^a	20 (16.8)	7 (5.9)	92 (77.3)
3. Did the review authors explain their selection of the study designs for inclusion in the review?	5 (4.2)	NA	114 (95.8)
4. Did the review authors use a comprehensive literature search strategy? ^a	3 (2.5)	112 (94.1)	4 (3.4)
5. Did the review authors perform study selection in duplicate?	93 (78.2)	NA	26 (21.8)
6. Did the review authors perform data extraction in duplicate?	90 (75.6)	NA	29 (24.4)
7. Did the review authors provide a list of excluded studies and justify the exclusions? ^a	11 (9.2)	3 (2.5)	105 (88.3)
8. Did the review authors describe the included studies in adequate detail?	18 (15.1)	96 (80.7)	5 (4.2)
9. Did the review authors use a satisfactory technique for assessing RoB in individual studies that were included in the review? ^a	78 (65.6)	28 (23.5)	13 (10.9)
10. Did the review authors report on the sources of funding for the studies included in the review?	4 (3.4)	NA	115 (96.6)
11. If meta-analysis was performed did the review authors use appropriate methods for statistical combination of results? ^a	49 (41.2)	NA	70 (58.8)
12. If meta-analysis was performed, did the review authors assess the potential impact of RoB in individual studies on the results of the meta-analysis or other evidence synthesis?	40 (33.6)	NA	79 (66.4)
13. Did the review authors account for RoB in individual studies when interpreting/discussing the results of the review? ^a	77 (64.7)	NA	42 (35.3)
14. Did the review authors provide a satisfactory explanation for, and discussion of, any heterogeneity observed in the results of the review?	92 (77.3)	NA	27 (22.7)
15. If they performed quantitative synthesis did the review authors carry out an adequate investigation of publication bias (small study bias) and discuss its likely impact on the results of the review? ^a	54 (45.4)	NA	65 (54.6)
16. Did the review authors report any potential sources of conflict of interest, including any funding they received for conducting the review?	106 (89.1)	NA	13 (10.9)

AMSTAR-2: A Measurement Tool to Assess Systematic Reviews 2; NA: not applicable; PICO: patients, intervention, comparison, and outcomes; RoB: risk of bias.

^a Critical domain.

components (population, intervention, control group, and outcome) in their research questions and inclusion criteria (item 1); and 106 (89.1%) reported potential sources of conflict of interest, including funding they received for conducting the review (item 16).

The performances of included SRs across 3 critical domains were disappointing, with less than 20% satisfying the criteria. Twenty (16.8%) SRs contained an explicit statement that the review methods were established prior to the conduct of the review and justified any significant deviations from the protocol (item 2). Eleven (9.2%) SRs provided a list of excluded studies and justified the exclusions (item 7). Only 3 (2.5%) SRs applied a comprehensive literature search strategy (item 4).

The remaining 4 critical domains had relatively better performances, with more than 40% of SRs considered as meeting the requirements. Seventy-eight (65.6%) SRs used a satisfactory technique for assessing the risk of bias in individual studies (item 9). Seventy-seven (64.7%) SRs accounted for the risk of bias in individual studies when interpreting or discussing the results (item 13). Forty-nine (55.4%) SRs used appropriate methods for the statistical combination of results (item 11). Fifty-four (45.4%) SRs conducted an ade-

quate investigation of publication bias and discussed the associated impact (item 15). Details of the methodological quality of included SRs are shown in Table 2.

Overall methodological quality

Of the 119 included SRs, only one (0.8%) presented high overall methodological quality. Four (3.4%) and 7 (5.9%) presented moderate and low overall methodological quality, respectively. The remaining 107 (89.9%) SRs were of critically low overall methodological quality.

In terms of overall methodological quality, Kruskal–Wallis tests showed that Cochrane reviews ($P < .001$) performed better than their counterparts. The Spearman rank correlation coefficient also showed that SRs published in higher IF journals ($r_s = 0.27$; $P = .003$) presented higher overall methodological quality. No significant associations were identified between overall methodological quality and being an update of a previous SR, publication year, location of the corresponding author, type of treatment, searching non-English databases, reporting of adverse effects, funding location, year of coverage, reporting of search terms, publication language restriction,

risk-of-bias assessment tools used, or the inclusion of a PRISMA-like flow diagram. Details are shown in [Table 3](#).

Results from ordinal logistic regression analysis of the association between bibliographical characteristics of SRs and the overall methodological quality of SRs are shown in [Appendix 6](#). The *P*-value of the Pearson Chi-square test was $<.1$, indicating that the model fit is poor.

Factors associated with adherence to individual AMSTAR-2 items

Binary logistic regression analyses showed that SRs published in journals with a higher IF presented better performance in adopting appropriate methods for the statistical combination of results in the meta-analysis (item 11: AOR: 1.29; 95% CI, 1.06–1.81). SRs on pharmacological treatments showed poorer performance than SRs on non-pharmacological treatments in presenting the research question with PICO components (item 1: AOR: 0.31; 95% CI, 0.10–0.98) and study selection in duplicate (item 5: AOR: 0.13; 95% CI, 0.04–0.40). SRs with corresponding authors from Asia showed better performance than those from Europe in data extraction in duplicate (item 6: AOR: 3.25; 95% CI, 1.10–9.64), but poorer performance in adopting appropriate methods for statistical combination of results in the meta-analysis (item 11: AOR: 0.22; 95% CI, 0.08–0.66). SRs with more review authors were more likely to account for the risk of bias when interpreting results (item 13: AOR: 1.22; 95% CI, 1.03–1.45). The *P*-values of all Hosmer–Lemeshow tests were $>.1$, indicating good model fit for all binary logistic regression analyses. Details are shown in [Table 4a](#).

Multinomial logistic regression analyses showed that SRs published in journals with a higher IF had a higher probability of explicitly stating that the review methods were established prior to the conduct of the review, and justifying any significant deviations from the protocol (item 2: AOR: 1.87; 95% CI, 1.27–2.77). SRs on pharmacological treatments were less likely to satisfy the following criteria when compared to those on non-pharmacological treatments: (i) containing an explicit statement that the review methods were established prior to the conduct of the review and justifying any significant deviations from the protocol (item 2: AOR: 1.87; 95% CI, 1.27–2.77), and (ii) describing the included studies in adequate detail (item 8: AOR: 0.03; 95% CI, 0.001–0.69). The *P*-values for all likelihood ratio tests were $<.05$, indicating good model fit for all multinomial logistic regression analyses. Details are shown in [Table 4b](#). Results from the remaining binary logistic regression analyses (items 3, 10, 12, 14, 15, and 16) and multinomial logistic regression analyses (items 4, 7, and 9) are not reported due to poor model fit.

Discussion

Summary of results

In this cross-sectional study, the methodological quality of 119 SRs on Parkinson's disease treatments published between 2016 and 2021 was evaluated. The majority of the appraised SRs (89.9%) were judged to be of critically low quality, indicating that the methodological quality of cur-

rent SRs on Parkinson's disease treatments is unsatisfactory. Three critical domains (items 2, 4, and 7) and 3 non-critical domains (items 3, 8, and 10) were fulfilled by fewer than 20% of SRs.

Cochrane reviews and SRs published in journals with higher IF were associated with higher overall methodological quality. The first phenomenon may be attributable to the high editorial requirement of the Cochrane Collaboration, which consists of peer-reviewing of SR protocols.¹⁶ This process serves as a gatekeeper, ensuring the rigor of Cochrane reviews. The second observation on the positive association between IF and rigor echoes previous findings,¹⁷ but meanwhile SRs published in journals with higher IFs had poor performance in describing the included studies in adequate detail (item 8). This is probably due to journals' word limits, as well as the lack of availability of online appendices. Compared to corresponding authors from European institutions, SRs with corresponding authors from Asia performed better for data extraction in duplicate (item 6) but poorer for using appropriate statistical methods (item 11). Regardless of location, efforts are still needed to improve the methodological rigor of future SRs on treatments for Parkinson's disease.

Compared to similar studies, the proportion of Parkinson's disease SRs with high or moderate overall methodological quality (4.2%) is similar to those on osteoporosis treatments (4.0%),¹¹ but is substantially lower than SRs on asthmatic treatments (15.4%)¹⁸ and osteoarthritis interventions (9.0%).¹⁹ In terms of performance for critical items, SRs on treatments for Parkinson's disease are inferior to SRs on asthma, osteoporosis, and osteoarthritis interventions in (i) providing a list of excluded studies with justifications (item 7), and (ii) using appropriate methods for statistical combination of results (item 11).

According to the Cochrane Handbook, all SRs should attempt to report evidence on harm.²⁰ A thorough synthesis of evidence on adverse effects is particularly important in contexts in which the risk/benefit ratio can have a major influence on clinical or policy decision-making.⁷ However, in our study, treatment adverse effects were not reported by one-third of the included SRs. The lack of attempt to summarize evidence of harm may be attributed to poor reporting among primary studies. The use of unstandardized adverse effects terminology poses a major challenge for systematic reviewers in appropriately synthesizing adverse event data.²⁰ For example, the number and type of possible adverse effects, as well as how adverse events are defined, ascertained, analyzed, and reported, can vary greatly between primary studies.²⁰ Addressing the root cause of suboptimal adverse events description among RCTs is a prerequisite for improving the lack of harm reporting in SRs.

Recommendations for future systematic reviews

Establishing and registering protocols a priori

Systematic reviewers should develop and register their protocols a priori; however, in this study only 16.8% of SRs met this requirement. By improving the transparency of SR topics and methods via protocols, unnecessary duplication of SRs can be avoided.⁷ With detailed SR protocols published

Table 3 Overall methodological quality of the 119 included systematic reviews on treatments for Parkinson's disease, according to different bibliographical characteristics.

Characteristics	High	Moderate	Low	Critically low	P
<i>Total</i>	1 (0.8)	4 (3.4)	7 (5.9)	107 (89.9)	
<i>Cochrane review</i>					<.001*
Yes	1 (33.3)	2 (66.7)	0 (0.0)	0 (0.0)	
No	0 (0.0)	2 (1.7)	7 (6.0)	107 (92.2)	
<i>Update of a previous SR</i>					.619
Yes	1 (14.3)	0 (0.0)	0 (0.0)	6 (85.7)	
No	0 (0.0)	4 (3.6)	7 (6.3)	101 (90.2)	
<i>Publication year</i>					.765
2016–2018	1 (2.2)	1 (2.2)	3 (6.7)	40 (88.9)	
2019–2021	0 (0.0)	3 (4.1)	4 (5.4)	67 (90.5)	
<i>Location of corresponding author</i>					.281
Europe	1 (3.8)	3 (11.5)	1 (3.8)	21 (80.8)	
America	0 (0.0)	1 (10.0)	0 (0.0)	9 (90)	
Asia	0 (0.0)	0 (0.0)	5 (6.4)	73 (93.6)	
Oceania	0 (0.0)	0 (0.0)	0 (0.0)	1 (100.0)	
Africa	0 (0.0)	0 (0.0)	1 (25.0)	3 (75.0)	
<i>Type of treatment</i>					.502
Pharmacological	1 (1.4)	3 (4.2)	5 (7.0)	62 (87.3)	
Non-pharmacological	0 (0.0)	1 (2.1)	2 (4.3)	44 (93.6)	
Both	0 (0.0)	0 (0.0)	0 (0.0)	1 (100.0)	
<i>SRs reporting adverse treatment effects</i>					.309
Yes	1 (1.3)	3 (3.9)	2 (2.6)	71 (92.2)	
No	0 (0.0)	1 (2.4)	5 (11.9)	36 (85.7)	
<i>Funding location of the SR</i>					.141
Europe	1 (9.1)	1 (9.1)	1 (9.1)	8 (72.7)	
America	0 (0.0)	1 (25.0)	0 (0.0)	3 (75.0)	
Asia	0 (0.0)	0 (0.0)	1 (2.7)	36 (97.3)	
Oceania	0 (0.0)	0 (0.0)	0 (0.0)	1 (100.0)	
Not reported	0 (0.0)	0 (0.0)	2 (6.9)	27 (93.1)	
No funding support	0 (0.0)	2 (5.4)	3 (8.1)	32 (86.5)	
<i>SRs that searched non-English databases</i>					.357
Yes	0 (0.0)	0 (0.0)	3 (7.0)	40 (93.0)	
No	1 (1.3)	4 (5.3)	4 (5.3)	67 (88.2)	
<i>Reported years covered by literature search</i>					.112
Yes	1 (1.5)	4 (6.0)	5 (7.5)	57 (85.1)	
Partial yes	0 (0.0)	0 (0.0)	2 (5.1)	37 (94.9)	
No	0 (0.0)	0 (0.0)	0 (0.0)	13 (100.0)	
<i>Search terms</i>					.139
Topic/free text/keyword/MeSH	0 (0.0)	0 (0.0)	1 (2.3)	43 (97.7)	
Full Boolean	1 (1.4)	4 (5.6)	6 (8.3)	61 (84.7)	
Readers are referred elsewhere for full search strategy	0 (0.0)	0 (0.0)	0 (0.0)	1 (100.0)	
No search term	0 (0.0)	0 (0.0)	0 (0.0)	2 (100.0)	
<i>Language of included primary studies in SRs</i>					.784
English only	0 (0.0)	1 (3.4)	1 (3.4)	27 (93.1)	
English and other languages	1 (1.4)	2 (2.9)	5 (7.2)	61 (88.4)	
Not reported	0 (0.0)	1 (4.8)	1 (4.8)	19 (90.5)	
<i>Tools for assessing quality of primary studies</i>					.381
Cochrane risk of bias	1 (1.3)	2 (2.6)	4 (5.2)	70 (90.9)	
Jadad scale	0 (0.0)	0 (0.0)	0 (0.0)	8 (100.0)	
PEDro scale	0 (0.0)	2 (10.5)	2 (10.5)	15 (78.9)	
Other tools	0 (0.0)	0 (0.0)	1 (25.0)	3 (75.0)	
More than one tool	0 (0.0)	0 (0.0)	0 (0.0)	7 (100.0)	
No quality assessment	0 (0.0)	0 (0.0)	0 (0.0)	4 (100.0)	
<i>Included a PRISMA-like flow diagram</i>					.559
Yes	1 (0.9)	4 (3.4)	7 (6.0)	104 (89.7)	
No	0 (0.0)	0 (0.0)	0 (0.0)	3 (100.0)	

MeSH: Medical Subject Headings; PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-analyses; SR: systematic review.

* P-value in the Kruskal–Wallis test was <.05.

Table 4a Association between bibliographical characteristics and performance on individual AMSTAR-2 items of systematic reviews on treatments for Parkinson's disease.

AMSTAR-2 item (dependent variable)	Predictors	AOR (95% CI)	P
<i>a) Binary logistic regression analyses*</i>			
1. Did the research questions and inclusion criteria for the review include the components of PICO?	Pharmacological treatment ^a	0.31 (0.10–0.98)	.046
5. Did the review authors perform study selection in duplicate?	Pharmacological treatment ^a	0.13 (0.04–0.40)	<.001
6. Did the review authors perform data extraction in duplicate?	Corresponding author was from Asia ^b	3.25 (1.10–9.64)	.034
11. If meta-analysis was performed did the review authors use appropriate methods for statistical combination of results?	Journal impact factor	1.29 (1.06–1.81)	.018
	Corresponding author was from Asia ^b	0.22 (0.08–0.66)	.007
13. Did the review authors account for RoB in individual studies when interpreting/discussing the results of the review?	Number of review authors	1.22 (1.03–1.45)	.024

AMSTAR-2: Assessing the Methodological Quality of Systematic Reviews; AOR: adjusted odds ratio; CI: confidence interval; PICO: patients, intervention, comparison, and outcomes; RoB: risk of bias.

^a Systematic reviews of non-pharmacological treatment were used as the reference.

^b Systematic reviews with corresponding authors from Europe were used as the reference.

* The *P*-values of all Hosmer–Lemeshow tests were >.1, indicating good model fit for all binary logistic regression analyses.

in open-access journals, potential errors may be identified through peer review, and be resolved before the review is conducted. In addition, discrepancies between the protocol and the final SR publication may be evaluated, reducing the possibility of selective outcome reporting and other methodological deviations.^{21,22}

Conducting a comprehensive literature search

In this study, only 2.5% of SRs fulfilled the criterion of conducting a comprehensive literature search. The majority of SRs concentrated solely on database searches, omitting reference lists, trial registries, gray literature, and consultation with subject matter experts. Consequently, relevant unpublished articles may have been overlooked. It is well known that such a lack of comprehensiveness may cause publication bias, skewing estimates of effectiveness.²³ A previous study showed that an incomprehensive literature search, especially one lacking a search of the gray literature, can result in an overestimation of treatment effect of nearly 12%.²⁴ It is also worth noting that only 36.1% of SRs searched non-English databases. Meanwhile, language-related selection criteria for primary studies were not reported in 55.7% of SRs, raising uncertainty as to whether non-English publications were considered eligible. Potential language bias may result in over- or underestimation of pooled treatment effects.²⁵

Providing a list of excluded studies with justifications for exclusion

To ensure the transparency and reproducibility of an SR, a list of potentially relevant but excluded primary studies, with justifications for exclusion, should be provided.^{26,27} In this study, only 9.2% of the included SRs adhered to this recommendation. As the study selection process inevitably involves some degree of subjectivity, reviewers should explicitly state the reasons for exclusion. This practice will assist readers in understanding whether such exclusion would impact the generalizability of the SR results, and in locating excluded studies that may inform their current clinical decision-making process.⁶

Transparent reporting on the sources of funding for the included studies

The results and conclusions of sponsored clinical trials and SRs may be biased toward the sponsored interventions, especially among studies of pharmacological interventions.^{28–30} Conflicts of interest in an SR are typically caused by 2 potential sources: the included primary studies, and the SR itself. These 2 sources are distinguished as 2 separate items in the AMSTAR-2 tool (items 10 and 16).⁶ The SRs appraised in this study generally performed well in declaring any conflicts of interest of the SR itself (89.1% satisfied item 16). However, only 3.4% reported sources of funding among the included primary studies (item 10).

In this study, when compared to SRs on non-pharmacological treatments, SRs on pharmacological treatments performed worse on comprehensively reporting research questions and inclusion criteria (item 1), establishing protocols a priori (item 2), conducting study selection in duplicate (item 5), and in describing the included trials in adequate detail (item 8). Similar to the well acknowledged observations on RCTs,³¹ the trustworthiness of pharmacological treatments SRs may be jeopardized due to methodological flaws, despite the availability of funding support from commercial sources.

Limitations

This cross-sectional study only evaluated SRs published in English, which may limit the representativeness of our sample. Our evaluation was only based on the information reported in published texts. Therefore, the appraisal results depend mainly on how comprehensively the SRs reported their methodology. Reporting quality of SRs is often limited by journals' restrictions on article length, poor adherence to reporting guidelines such as PRISMA, and the lack of availability of online appendices. These factors may contribute to unclear reporting among SR publications, compromising the accuracy of our assessment of methodological quality.³²

Table 4b Association between bibliographical characteristics and performance on individual AMSTAR-2 items of systematic reviews on treatments for Parkinson's disease. AMSTAR-2 item (dependent variable)

Predictors	Yes ^a		Partial yes ^a	
	AOR (95% CI)	P	AOR (95% CI)	P

b) Multinomial logistic regression analyses^{}*

2. Did the report of the review contain an explicit statement that the review methods were established prior to the conduct of the review and did the report justify any significant deviations from the protocol?	Journal impact factor	1.87 (1.27–2.77)	.002	1.74 (1.07–2.84)	.026
	Pharmacological treatment ^b	0.15 (0.03–0.88)	.036	0.55 (0.07–4.20)	.567
8. Did the review authors describe the included studies in adequate detail?	Journal impact factor	0.68 (0.37–1.26)	.221	0.55 (0.32–0.93)	.026
	Pharmacological treatment ^b	0.03 (0.001–0.69)	.029	0.55 (0.05–6.29)	.628

AMSTAR-2: Assessing the Methodological Quality of Systematic Reviews; AOR: adjusted odds ratio; CI: confidence interval.

^a Each AMSTAR-2 item could be judged as "Yes," "Partial yes," and "No." In multinomial logistic regression analyses, the reference category of dependent variable is "No."

^b Systematic reviews of non-pharmacological treatment were used as the reference.

^{*} The P-values of all likelihood ratio tests were < .05, indicating good model fit for all multinomial logistic regression analyses.

Conclusions

This study showed that the methodological quality of SRs on treatments for Parkinson's disease treatments is far from satisfactory. To enhance their methodological quality, future SRs should establish and register protocols a priori, conduct comprehensive literature searches, provide a list of excluded studies with justifications for exclusion, and report sources of funding among the included primary studies.

Ethics approval and consent to participate

All data for this study were obtained from existing publications and ethical approval was not required for this study.

Funding

This work was supported by the National Natural Science Foundation of China (81973709) and the High-level Talents Introduction Plan from Central South University (502045003). The funders had no role in considering the study design or in the collection, analysis, and interpretation of the data, manuscript drafting, or decision to submit the article for publication.

Authors' contributions

Study design: VC and CZ; Literature search and literature selection: YZ, YL, and CZ; Data extraction, methodological quality assessments: YZ, FH, and CZ; Data analysis: YZ, VC, and CZ; Data interpretation: YZ, VC, and CZ; Tables, Figures, and Appendix preparation: YZ; Manuscript drafting: YZ, CZ, and VC; Critical revision of the manuscript: FH, IW, CM, and XY.

Consent for publication

Not applicable.

Conflict of interests

The authors declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Availability of data and materials

The data that support the findings of this study are available on request from the corresponding author.

Acknowledgements

Not applicable.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at [doi:10.1016/j.nrl.2023.03.003](https://doi.org/10.1016/j.nrl.2023.03.003).

References

- Homayoun H. Parkinson disease. *Ann Intern Med*. 2018;169:ITC33–48.
- Dorsey ER, Elbaz A, Nichols E, Abbasi N, Abd-Allah F, Abdelalim A, et al. Global, regional, and national burden of Parkinson's disease, 1990–2016: a systematic analysis for the Global Burden of Disease Study 2016. *Lancet Neurol*. 2018;17:939–53.
- Boiles A. Parkinson's disease economic burden on patients, families and the federal government is \$52 billion, doubling previous estimates. 2019;vol. 1:2022. Available from: <https://www.michaeljfox.org/publication/parkinsons-disease-economic-burden-patients-families-and-federal-government-52-billion> [accessed 1. 8. 24].
- Murad MH, Asi N, Alsawas M, Alahdab F. EBHC pyramid 5.0 for accessing preappraised evidence and guidance. *Evid Based Med*. 2016;21:125–7.
- Mulrow CD, Cook DJ, Davidoff F. Systematic reviews: critical links in the great chain of evidence. *Ann Intern Med*. 1997;126:389–91.
- 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;358:j4008.
- Higgins JPT, Thomas J, Chandler J, Cumpston M, Li T, Page MJ, Welch VA (editors). *Cochrane handbook for systematic reviews of interventions* version 6.3. Available from: www.training.cochrane.org/handbook [updated February 2022, accessed 1. 8. 24].
- Unit MHIR. Search filters for MEDLINE in Ovid syntax and the PubMed translation: McMaster health information research unit. Available from: https://hiru.mcmaster.ca/hiru/HIRU_Hedges_MEDLINE_Strategies.aspx [accessed 1. 8. 24].
- Unit MHIR. Search strategies for EMBASE in Ovid syntax: McMaster health information research unit. Available from: https://hiru.mcmaster.ca/hiru/HIRU_Hedges_EMBASE_Strategies.aspx [accessed 1. 8. 24].
- Unit MHIR. Search strategies for PsycINFO in Ovid syntax: McMaster health information research unit. Available from: https://hiru.mcmaster.ca/hiru/HIRU_Hedges_PsycINFO_Strategies.aspx [accessed 1. 8. 24].
- Tsoi AKN, Ho LTF, Wu IXY, Wong CHL, Ho RST, Lim JYY, et al. Methodological quality of systematic reviews on treatments for osteoporosis: a cross-sectional study. *Bone*. 2020;139:115541.
- Wu XY, Du XJ, Ho RS, Lee CC, Yip BH, Wong MC, et al. Characteristics and methodological quality of meta-analyses on hypertension treatments – a cross-sectional study. *J Clin Hypertens (Greenwich)*. 2017;19:137–42.
- Wu XY, Lam VC, Yu YF, Ho RS, Feng Y, Wong CH, et al. Epidemiological characteristics and methodological quality of meta-analyses on diabetes mellitus treatment: a systematic review. *Eur J Endocrinol*. 2016;175:353–60.
- Ho RS, Wu X, Yuan J, Liu S, Lai X, Wong SY, et al. Methodological quality of meta-analyses on treatments for chronic obstructive pulmonary disease: a cross-sectional study using the AMSTAR (Assessing the Methodological Quality of Systematic Reviews) tool. *NPJ Prim Care Respir Med*. 2015;25:14102.
- Firth J, Teasdale SB, Allott K, Siskind D, Marx W, Cotter J, et al. The efficacy and safety of nutrient supplements in the treatment of mental disorders: a meta-review of meta-analyses of randomized controlled trials. *World Psychiatry*. 2019;18:308–24.
- Reviews. London UKTCC. Information for authors submitting to Cochrane. Available from: <https://www.cochranelibrary.com/about/author-information> [accessed 1. 8. 24].
- Mackinnon S, Drozdowska BA, Hamilton M, Noel-Storr AH, McShane R, Quinn T. Are methodological quality and completeness of reporting associated with citation-based measures of publication impact? A secondary analysis of a systematic review of dementia biomarker studies. *BMJ Open*. 2018;8:e020331.
- Wu IXY, Deng Y, Wang H, Chen Y, Wong CHL, Chung VCH. Methodological quality of systematic reviews and meta-analysis on asthma treatments. A cross-sectional study. *Ann Am Thorac Soc*. 2020;17:949–57.
- Wu IX, Wang H, Zhu L, Chen Y, Wong CH, Mao C, et al. Methodological quality of systematic reviews on interventions for osteoarthritis: a cross-sectional study. *Therap Adv Musculoskelet Dis*. 2020;12, 1759720X20959967.
- Peryer G, Golder S, Junqueira DR, Vohra S, Loke YK, Group CAEM. Adverse effects. *Cochrane Handbook Syst Rev Interv*. 2019:493–505.
- Editors PLM. Best practice in systematic reviews: the importance of protocols and registration. *PLoS Med*. 2011;8:e1001009.
- Dissemination CfRa. About PROSPERO York, United Kingdom: centre for reviews and dissemination. Available from: <https://www.crd.york.ac.uk/prosperto/#aboutpage> [accessed 1. 8. 24].
- Hart B, Lundh A, Bero L. Effect of reporting bias on meta-analyses of drug trials: reanalysis of meta-analyses. *BMJ*. 2012;344:d7202.
- McAuley L, Pham B, Tugwell P, Moher D. Does the inclusion of grey literature influence estimates of intervention effectiveness reported in meta-analyses? *Lancet*. 2000;356:1228–31.
- Morrison A, Polisena J, Huserneau D, Moulton K, Clark M, Fiander M, et al. The effect of English-language restriction on systematic review-based meta-analyses: a systematic review of empirical studies. *Int J Technol Assess Health Care*. 2012;28:138–44.
- Faggion CM Jr. Critical appraisal of AMSTAR: challenges, limitations, and potential solutions from the perspective of an assessor. *BMC Med Res Methodol*. 2015;15:1–5.
- Wegewitz U, Weikert B, Fishta A, Jacobs A, Pieper D. Resuming the discussion of AMSTAR: what can (should) be made better? *BMC Med Res Methodol*. 2016;16:1–8.
- Kjaergard LL, Als-Nielsen B. Association between competing interests and authors' conclusions: epidemiological study of randomised clinical trials published in the BMJ. *BMJ*. 2002;325:249.
- Lexchin J, Bero LA, Djulbegovic B, Clark O. Pharmaceutical industry sponsorship and research outcome and quality: systematic review. *BMJ*. 2003;326:1167–70.
- Lundh A, Lexchin J, Mintzes B, Schroll JB, Bero L. Industry sponsorship and research outcome: systematic review with meta-analysis. *Intensive Care Med*. 2018;44:1603–12.
- Chopra SS. MSJAMA: industry funding of clinical trials: benefit or bias? *JAMA*. 2003;290:113–4.
- 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.