


Methodological Quality of Systematic Reviews on Treatments for Liver Cancer: A Cross-sectional Study

Billy Z Z Cheng¹, Betty H Wang², Claire Chenwen Zhong², Yuning Zhang², Fai Fai Ho ³, Vincent CH Chung²

¹Department of Administration, Shanghai Public Health Clinical Center, Fudan University, Shanghai, People's Republic of China; ²JC School of Public Health and Primary Care, Faculty of Medicine, The Chinese University of Hong Kong, Hong Kong, People's Republic of China; ³School of Chinese Medicine, Faculty of Medicine, The Chinese University of Hong Kong, Hong Kong, People's Republic of China

Correspondence: Betty H Wang, JC School of Public Health and Primary Care, Faculty of Medicine, The Chinese University of Hong Kong, Hong Kong, People's Republic of China, Email huanwang@link.cuhk.edu.hk

Background: Systematic reviews (SRs) are indispensable for presenting reliable evidence of the effectiveness of treatments. However, methodological flaws can affect their reliability and validity.

Aim: This cross-sectional study aimed to evaluate the methodological quality of SRs on liver cancer (LC) treatments and identify potential factors affecting their reliability.

Methods: A comprehensive literature search was carried out on four databases to identify eligible SRs published between January 2014 and October 2023. We appraised the methodological quality of included SRs by Assessing the Methodological Quality of Systematic Reviews 2 (AMSTAR 2) tool. Multivariable regression analysis was employed to investigate the factors influencing the methodological quality.

Results: A total of 119 SRs were included and appraised. Only one SR (0.8%) was rated as high overall quality. One (0.8%), nine (7.6%), and 108 (90.8%) were appraised as moderate, low, and critical low quality, respectively. SRs published more recently, with higher journal impact factors, or with corresponding author from Europe have better performance.

Conclusion: The methodological quality of SRs on LC treatments was unsatisfactory. Future SR authors should improve quality of SRs through registering an a priori protocol, providing explanation for selection of study designs, using a comprehensive literature search strategy, listing all excluded studies and justifying their reasons, describing the included studies in adequate detail, and reporting funding resources of primary studies.

Keywords: liver cancer, cross-sectional study, bias, meta-analysis, systematic reviews

Introduction

Liver Cancer (LC) is the sixth most common cancer in the world and the second most deadly cancer.¹ Based on the statistics from the World Cancer Research Fund International, the new cases of LC have reached 900,000 with an 8.7 per 100,000 people mortality and 20.8% 5-year relative survival rate in 2020.² The research estimates that between 2020 and 2040, there would be a 55.0% rise in new instances of LC, with 1.4 million new diagnoses predicted by 2040.³ The burden of LC is greatly increased by the rising trend in the disease's prevalence and the length of time patients remain unwell before dying, exposing global health and social care systems to a serious threat.

At present, both pharmacological and non-pharmacological therapies are used to treat LC patients. These include surgical therapies like resection, liver transplantation, which is potentially curative therapy but not suitable for the majority, ablation, transarterial embolization and radiotherapy, as well as systemic pharmacological treatments like targeted therapy, immunotherapy, and chemotherapy.⁴⁻⁸ With the availability of numerous treatment options for LC patients, it is paramount for clinicians to make sound decision based on the best available clinical evidence.

Systematic reviews (SRs) which are thought to offer the best quality of evidence about the effectiveness of specific therapy. As per the Cochrane handbook, a qualified SR with a flawless meta-analysis can identify, critically appraise, and synthesize quantitative evidence from eligible primary studies that comply with predetermined eligibility criteria.⁹ However, not every SR follows critical methodological quality.^{10,11} Methodologically flawed SRs have the potential to mislead clinical decision-makers by exaggerating or underestimating treatment effects.¹² Thus legislators, medical professionals, caregivers, and any other evidence users, should critically assess the methodological quality of SRs before implementing the results and conclusions into clinical practice.

So far, no assessment of the methodological quality of SRs in LC treatments has been conducted. Therefore, this cross-sectional study aimed to: (i) describe the bibliographic characteristics of a current sample of SRs on LC treatments; (ii) assess the methodological quality of SRs using the Assessing the Methodological Quality of Systematic Reviews 2 (AMSTAR 2) tool;¹³ and (iii) explore the relationships between methodological quality and bibliographic characteristics. Findings of this study could provide insight on how the methodological quality of the future SRs on treatments for LC could be improved.

Methods

Eligibility Criteria

In this cross-sectional study, SRs only focusing on randomized controlled trials (RCTs) and including one or more meta-analyses that summarize the effectiveness of LC treatments are eligible. SRs published in English or Chinese were taken into consideration for inclusion, and any intervention type for LC was accepted. SRs that focus mainly on the diagnosis, etiology, or risk factors of LC were excluded. Protocols, narrative reviews, network meta-analyses, and overview of SRs were excluded, as well. Only the most recent version of a single SR was included if multiple versions were discovered. We selected the most up-to-date SR with multiple versions with the remaining as [supplementary materials](#).

Literature Search, Literature Selection, Data Extraction

We searched four electronic databases, including MEDLINE (Ovid), Embase (Ovid), PsycINFO (Ovid), and the Cochrane Database of Systematic Reviews to identify potential qualified SRs published between January 2014 and October 2023. Validated search filters for SRs were applied to maximize the specificity of search on MEDLINE and Embase.^{14,15} The search strategies are presented in [Appendix 1](#). Two reviewers (ZC and YZ) screened the retrieved titles and abstracts of citations and checked the full text of SRs for determining the eligibility. After they independently collected detailed bibliographic characteristics using a pre-designed data extraction form ([Appendix 2](#)), a third reviewer (HW) was consulted in cases with persistent disagreements.

Methodological Quality Assessment

We used AMSTAR 2 tool to evaluate the methodological quality of included SRs.^{13,16} According to the published operational guidelines, AMSTAR 2 has seven critical items (items 2, 4, 7, 9, 11, 13, 15) among 16 items.¹³ SRs were assessed as “high”, “moderate”, “low”, or “critically low” in terms of overall methodological quality, depending on how well they performed in each item ([Appendix 3](#)).¹³ Two reviewers independently carried out the methodological quality assessment. Any disagreements or discrepancies were settled by senior researcher arbitration or by consensus among the authors.

Data Analysis

Data of the bibliographic characteristics and AMSTAR 2 assessment was presented as frequencies (percentages), and medians (ranges) appropriately. The Kruskal–Wallis rank test was used to compare differences between overall methodological quality of SRs and bibliographic features separately.

Binary logistic regression (for AMSTAR 2 items 1, 3, 5, 6, and 10–16) and multinomial logistic regression (items 2, 4, and 7–9) were applied to explore the associations between bibliographic characteristics and individual AMSTAR 2 items. Drawing from previous methodological research findings, the following seven bibliographical characteristics were selected as independent variables for the logistic regression analyses: (1) whether the SR was a Cochrane review, (2)

whether it was an updated SR, (3) the type of treatment that SR focus on, (4) year of publication, (5) impact factor of the journal in the year before the SR was published, (6) number of authors, and (7) location of the corresponding author.^{10,11} The model fitting of binary logistic regression was evaluated using the Hosmer–Lemeshow test, and multinomial regression analysis was evaluated using the Pearson deviance test and likelihood ratio test. The associations between AMSTAR 2 ratings and bibliographic characteristics were quantified through calculating the adjusted odds ratios (AORs) and 95% confidence intervals (CIs). Statistical significance was indicated if $P < 0.05$. All data analysis processes were conducted by IBM SPSS Statistics v. 27 software (IBM Corporation, Armonk, NY, USA).

Results

Literature Screening and Selection

We retrieved 9686 records through the literature search. After de-duplication, the remaining 7752 citations were screened based on their titles and abstracts, of which 841 full texts were downloaded for further checks. Finally, 119 SRs that met the eligibility criteria were included. The list of included studies is recorded in [Appendix 4](#). Details on literature selection are presented in [Figure 1](#).

Bibliographic Characteristics

Overall, this study included 119 SRs (1448 RCTs and 201,713 patients) published between 2014 and 2023 (median: 2019). The medians number of included RCTs and patients were eight (range: 2–65) and 1055 (range: 49–22,113), respectively. Only three (2.5%) SRs were Cochrane reviews, and around one-quarter (24.3%) were updates of previous SRs. Most of the included SRs (84%) were published in English. The impact factor of the journal in the year before the SR published was ranged from 0 to 22.682, with a median of 2.656. The median number of review authors was six (range: 2–17). Most of the corresponding authors (84%) were located at Asia, followed by Europe (9.3%), America (5.0%) and Africa (1.7%). The highest proportion of funding location of SRs was from Asia (39.5%), while nearly one third (31.9%) did not report funding relevant information. Close to two thirds of included SRs focused on pharmacological treatments (66.4%), and over 80% (82.4%) reported the intervention harms. Almost all the SRs (95.8%) included a PRISMA flowchart. All included SRs searched for English databases, but less than half of the SRs (41.2%) searched for non-English databases. More than half of SRs (55.4%) did not restrict the publication language of primary studies. Nearly two-thirds of SRs (64.7%) reported the “year of coverage” by starting and ending years. Around 75% of SRs applied Cochrane risk of bias tool for critical appraisal (74.8%), followed by Jadad scale (15.1%) ([Table 1](#)).

Overall Methodological Quality

Only one SR (0.8%) was appraised as high and moderate methodological quality, respectively. Nine SRs (7.6%) were of low quality while 108 SRs (90.8%) were judged as critically low quality. Kruskal–Wallis test results indicated that Cochrane review ($P < 0.001$), update of previous review ($P < 0.001$), SRs with corresponding author from Africa ($P = 0.001$), SRs receiving fund from institutions or organizations in Europe ($P = 0.001$), and SR reported full search strategy ($P = 0.025$) presented higher overall quality ([Table 2](#)). The Spearman’s rank correlation coefficients showed positive association between journal impact factor ($r_s = 0.290$; $P = 0.001$) and overall methodological quality.

Performance of Included SRs on Each Individual AMSTAR 2 Item

[Table 3](#) showed the performances of all included SRs across each AMSTAR 2 item. Explicitly deficient performances were found in three critical items while less than 20% of included SRs met the quality criteria. Only 15 SRs (2.6%) reported a registered protocol (item 2), four SRs (3.4%) used a comprehensive literature search strategy (item 4), and eight SRs (6.8%) provided a list of excluded studies and justify the exclusions reasons (item 7). As well as some of included SRs performed unsatisfactorily on non-critical items 3, 8, 10. Respectively, 23 review authors (19.3%) explained their selection of the study designs for inclusion in the review, 15 (12.7%) described the included studies in adequate detail, and six (5.1%) reported on the sources of funding for the studies included in the review.

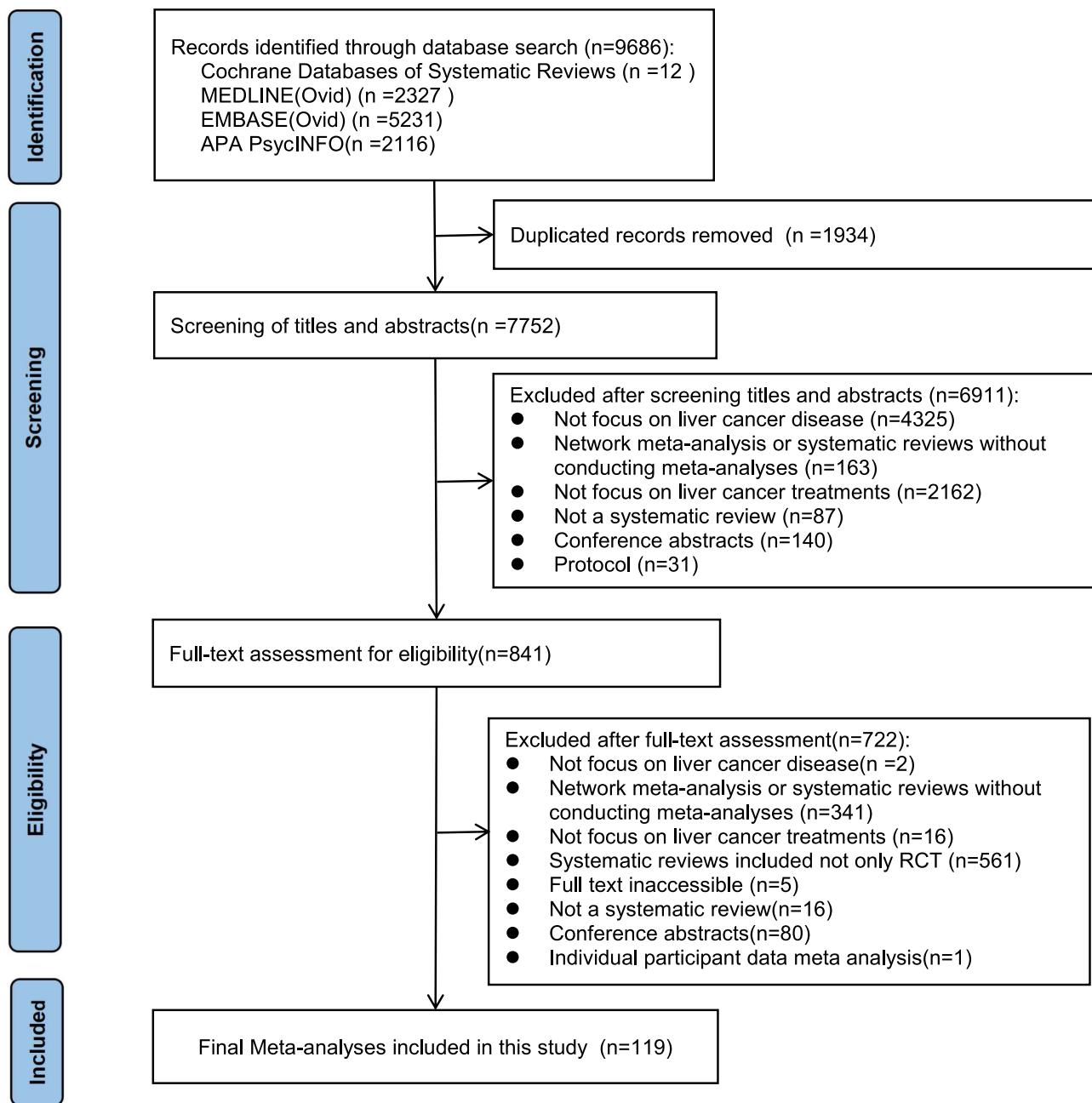


Figure 1 Flow chart of literature search and selection.

Abbreviation: RCT, randomized controlled trial.

The included SRs had excellent performances on non-critical item 1, 5, 6, and 14 for over 80% of the total fulfilling the items' requirements: (i) all SRs clearly stated PICO (population, intervention, control, outcome) in the research questions and inclusion criteria for the review (item 1); (ii) 118 SRs (99.2%) had study selection and data extraction performed in duplicate (item 5 and 6); (iii) 102 SRs (85.7%) provided a satisfactory explanation for any heterogeneity found in the results of the review. However, none of critical items reached the requirement of outstanding performances.

Factors Associated with Methodological Quality

Binary logistic regression revealed that the more recently that SR published, the better performance it had on following items: explaining inclusion criteria of the study design (AOR: 1.29; 95%CI: 1.02–1.62), applying appropriate methods

Table 1 Bibliographical Characteristics of 119 Included Systematic Reviews

| Bibliographical Characteristics | Results n (%) |
|---|----------------------|
| Cochrane SRs | |
| Yes | 3 (2.5) |
| No | 116 (97.5) |
| An update of a previous SR | |
| Yes | 29 (24.3) |
| No | 90 (75.6) |
| Publication language | |
| English | 101 (84.8) |
| Chinese | 18 (15.2) |
| Publication year | |
| 2014–2018 | 57 (47.9) |
| 2019–2023 | 62 (52.1) |
| Median year of publication, year [range] | 2019 [2014–2023] |
| Impact factor of journal (one year before SR publication), median [range] | 2.656 [0.000–22.682] |
| Number of review authors, median [range] | 6 [2–17] |
| Location of the corresponding author | |
| Asia | 100 (84.0) |
| Europe | 11 (9.3) |
| America | 6 (5.0) |
| Africa | 2 (1.7) |
| Funding location of the SR | |
| Asia | 47 (39.5) |
| Europe | 5 (4.2) |
| America | 1 (0.9) |
| Not reported funding relevant information | 38 (31.9) |
| No funding support | 28 (23.5) |
| Type of treatment | |
| Pharmacological treatments | 79 (66.4) |
| Non-pharmacological treatments | 29 (24.4) |
| Both types | 11 (9.2) |
| Median number of included RCTs [range] | 8 [2–65] |
| Median number of participants included in SRs [range] | 1055 [49–22,113] |
| SRs reporting intervention harms | |
| Yes | 98 (82.4) |
| No | 21 (17.6) |
| SRs searched for English databases | 119 (100.0) |
| SRs searched for non-English databases | |
| Yes | 49 (41.2) |
| No | 70 (58.8) |
| SRs included a PRISMA-like flow diagram | |
| Yes | 114 (95.8) |
| No | 5 (4.2) |
| Search terms reported for electronic databases | |
| No search terms reported | 2 (1.7) |
| Topics/free text/keywords/MeSH | 77 (64.7) |
| Full Boolean | 19 (16.0) |
| Readers are referred elsewhere for full search strategy | 21 (17.6) |
| Report year of coverage of literature search | |
| Yes, reported starting and ending years | 77 (64.7) |
| Partially (only reported starting year or ending year) | 28 (23.5) |
| Not mentioned | 14 (11.8) |

(Continued)

Table 1 (Continued).

| Bibliographical Characteristics | Results n (%) |
|--|---------------|
| Language restrictions of primary studies | |
| English only | 26 (21.9) |
| English and language other than English | 66 (55.4) |
| Language criteria not reported | 27 (22.7) |
| Methodological quality assessment tools | |
| Cochrane risk of bias tool 1.0 | 70 (58.8) |
| Cochrane risk of bias tool 2.0 | 19 (16.0) |
| Jadad scale | 18 (15.1) |
| More than one tool ^a | 3 (2.5) |
| Not reported the name of the assessment tool | 2 (1.7) |
| Newcastle–Ottawa Scale | 2 (1.7) |
| No quality assessment conducted | 5 (4.2) |

Notes: ^aMore than one tools refers to the SRs used both “Jadad scale” and “Delphi list” tools.

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

Table 2 Overall Methodological Quality of the Included Systematic Reviews by Bibliographical Characteristics

| Overall Methodological Quality | Critically Low n (%) | Low n (%) | Moderate n (%) | High n (%) | P-values ^b |
|---|----------------------|-----------|----------------|------------|-----------------------|
| Total | 108 (90.8) | 9 (7.6) | 1 (0.8) | 1 (0.8) | – |
| Cochrane SRs | | | | | <0.001 |
| Yes | 0 (0.0) | 2 (66.7) | 0 (0.0) | 1 (33.3) | |
| No | 108 (93.1) | 7 (6.0) | 1 (0.9) | 0 (0.0) | |
| An update of a previous SR | | | | | <0.001 |
| Yes | 22 (75.9) | 5 (16.9) | 1 (3.6) | 1 (3.6) | |
| No | 86 (95.6) | 4 (4.4) | 0 (0.0) | 0 (0.0) | |
| Publication Language | | | | | 0.114 |
| English | 90 (89.1) | 9 (8.9) | 1 (1.0) | 1 (1.0) | |
| Chinese | 18 (100.0) | 0 (0.0) | 0 (0.0) | 0 (0.0) | |
| SRs reporting intervention harms | | | | | 0.109 |
| Yes | 87 (88.8) | 9 (9.2) | 1 (1.0) | 1 (1.0) | |
| No | 21 (100.0) | 0 (0.0) | 0 (0.0) | 0 (0.0) | |
| SRs that search for non-English databases | | | | | 0.116 |
| Yes | 66 (94.3) | 3 (4.3) | 1 (1.4) | 0 (0.0) | |
| No | 42 (85.7) | 6 (12.2) | 0 (0.0) | 1 (2.1) | |
| Included a PRISMA-like flow diagram | | | | | 0.468 |
| Yes | 103 (90.3) | 9 (7.9) | 1 (0.9) | 1 (0.9) | |
| No | 5 (100.0) | 0 (0.0) | 0 (0.0) | 0 (0.0) | |
| Location of the corresponding author | | | | | 0.001 |
| Europe | 10 (90.9) | 1 (9.1) | 0 (0.0) | 0 (0.0) | |
| America | 3 (50.0) | 3 (50.0) | 0 (0.0) | 0 (0.0) | |
| Asia | 95 (95.0) | 4 (4.0) | 1 (1.0) | 0 (0.0) | |
| Africa | 0 (0.0) | 1 (50.0) | 0 (0.0) | 1 (50.0) | |
| Funding location of the SR | | | | | 0.001 |
| Europe | 1 (20.0) | 3 (60.0) | 0 (0.0) | 1 (20.0) | |
| America | 1 (100.0) | 0 (0.0) | 0 (0.0) | 0 (0.0) | |
| Asia | 44 (93.7) | 2 (4.2) | 1 (2.1) | 0 (0.0) | |
| No funding | 26 (92.9) | 2 (7.1) | 0 (0.0) | 0 (0.0) | |
| Not reported | 36 (94.7) | 2 (5.3) | 0 (0.0) | 0 (0.0) | |

(Continued)

Table 2 (Continued).

| Overall Methodological Quality | Critically Low n (%) | Low n (%) | Moderate n (%) | High n (%) | P-values ^b |
|---|----------------------|-----------|----------------|------------|-----------------------|
| Types of treatments | | | | | 0.286 |
| Non-pharmacological treatments | 25 (86.0) | 3 (10.5) | 0 (0.0) | 1 (3.5) | |
| Pharmacological treatments | 74 (93.7) | 4 (5.0) | 1 (1.3) | 0 (0.0) | |
| Both types | 9 (81.8) | 2 (18.2) | 0 (0.0) | 0 (0.0) | |
| Report year span of search | | | | | 0.343 |
| Not mentioned | 14 (100.0) | 0 (0.0) | 0 (0.0) | 0 (0.0) | |
| Partially | 26 (92.9) | 2 (7.1) | 0 (0.0) | 0 (0.0) | |
| Yes | 68 (88.3) | 7 (9.1) | 1 (1.3) | 1 (1.3) | |
| Search terms reported for electronic databases | | | | | 0.025 |
| No search term | 2 (100.0) | 0 (0.0) | 0 (0.0) | 0 (0.0) | |
| Topics/free text/keywords/MeSH | 74 (96.1) | 3 (3.9) | 0 (0.0) | 0 (0.0) | |
| Full Boolean | 16 (84.2) | 3 (15.8) | 0 (0.0) | 0 (0.0) | |
| Readers are referred elsewhere for full search strategy | 16 (76.1) | 3 (14.3) | 1 (6.3) | 1 (6.3) | |
| Language criteria of primary studies | | | | | 0.357 |
| English only | 22 (84.6) | 4 (15.4) | 0(0.0) | 0 (0.0) | |
| English and language other than English | 60 (90.9) | 4 (6.1) | 1(1.5) | 1 (1.5) | |
| Not reported | 26 (96.3) | 1 (3.7) | 0(0.0) | 0 (0.0) | |
| Methodological quality assessment tools | | | | | 0.354 |
| Cochrane risk of bias | 80 (89.9) | 7 (7.9) | 1 (1.1) | 1 (1.1) | |
| Jadad scale | 17 (94.4) | 1 (5.6) | 0 (0.0) | 0 (0.0) | |
| Others ^c | 9 (100.0) | 0 (0.0) | 0 (0.0) | 0 (0.0) | |
| More than one tool ^a | 2 (66.7) | 1 (33.3) | 0 (0.0) | 0 (0.0) | |

Notes: ^aMore than one tools refers to the SRs used both “Jadad scale” and “Delphi list” tools. ^bP-values associated with Kruskal–Wallis tests; Bold values indicate statistical significance at $P < 0.05$. ^cOthers refer to “Newcastle–Ottawa Scale”, “Names of the tools not reported” and “No quality assessment”.

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

Table 3 Results of the AMSTAR 2 Items for the Included 119 Systematic Reviews

| AMSTAR 2 Items | Yes n (%) | Partial Yes n (%) | No n (%) |
|---|-------------|-------------------|------------|
| 1. Did the research questions and inclusion criteria for the review include the components of PICO? | 119 (100.0) | N/A | 0 (0.0) |
| 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? | 15 (12.6) | 10 (8.5) | 94 (78.9) |
| 3. Did the review authors explain their selection of the study designs for inclusion in the review? | 23 (19.3) | N/A | 96 (80.7) |
| 4. *Did the review authors use a comprehensive literature search strategy? | 4 (3.4) | 112 (94.1) | 3 (2.5) |
| 5. Did the review authors perform study selection in duplicate? | 118 (99.2) | N/A | 1 (0.8) |
| 6. Did the review authors perform data extraction in duplicate? | 118 (99.2) | N/A | 1 (0.8) |
| 7. *Did the review authors provide a list of excluded studies and justify the exclusions? | 8 (6.8) | 0(0.0) | 111 (93.2) |
| 8. Did the review authors describe the included studies in adequate detail? | 15 (12.7) | 100 (84.0) | 4 (3.3) |
| 9. *Did the review authors use a satisfactory technique for assessing the RoB in individual studies that were included in the review? | 85 (71.4) | 19 (16.0) | 15 (12.6) |
| 10. Did the review authors report on the sources of funding for the studies included in the review? | 6 (5.1) | N/A | 113 (94.9) |
| 11. *If meta-analysis was performed did the review authors use appropriate methods for statistical combination of results? | 31 (26.1) | N/A | 88 (73.9) |
| 11a. For No in Q.11, Is the inappropriateness evaluated because of the use of the fixed-effect meta-analysis model for combining data? | 87 (98.9) | N/A | 1 (1.1) |

(Continued)

Table 3 (Continued).

| AMSTAR 2 Items | Yes n (%) | Partial Yes n (%) | No n (%) |
|---|------------|-------------------|-----------|
| 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? | 50 (42.1) | N/A | 69 (57.9) |
| 13. *Did the review authors account for RoB in individual studies when interpreting/discussing the results of the review? | 50 (42.1) | N/A | 69 (57.9) |
| 14. Did the review authors provide a satisfactory explanation for, and discussion of, any heterogeneity observed in the results of the review? | 102 (85.7) | N/A | 17 (14.3) |
| 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? | 87 (73.2) | N/A | 32 (26.8) |
| 15a. For No in Q.15, Is the missing of graphical or statistical test for publication bias because of less than 10 individual studies in meta-analysis? | 9 (28.1) | N/A | 23 (71.9) |
| 16. Did the review authors report any potential sources of conflict of interest, including any funding they received for conducting the review? | 95 (79.8) | N/A | 24 (20.2) |

Note: *Critical methodological items.

Abbreviations: AMSTAR 2, Assessing the Methodological Quality of Systematic Reviews 2; N/A, not applicable; PICO, patients, intervention, comparison, and outcomes; RoB, risk of bias.

into statistical combination of meta-analysis results (AOR: 1.29; 95%CI: 1.06–1.57), evaluating and interpreting the potential impact of risk of bias on results (AOR: 1.47; 95%CI: 1.23–1.76), providing a satisfactory explanation for any heterogeneity found in the reviews (AOR: 1.32; 95%CI: 1.05–1.66), and reporting the conflict of interest (AOR: 1.33; 95%CI: 1.03–1.72). Compared with non-updated reviews, updated reviews had better performance on explaining inclusion criteria of the study design (AOR: 6.93, 95%CI: 2.06–23.30). SRs conducted by Asian corresponding authors showed worse performance on using appropriate methods in statistical combination of meta-analysis results (AOR: 0.17; 95%CI: 0.04–0.78) compared to those conducted by European. SRs published in journals with higher impact factors performed better on stating the conflict of interest (AOR: 2.98; 95%CI: 1.67–5.30) (Table 4).

Multinomial logistic regression indicated that the SRs published more recently were more likely to be rated as “partial yes” (AOR: 1.46; 95%CI: 1.04–2.06) and “yes” (AOR: 1.86; 95%CI: 1.17–2.96) on registering a priori protocols. Compared with non-updated reviews, updated reviews had significantly better performance on describing included studies in details (AOR: 11.65; 95%CI: 2.38–56.92) (Table 5).

Table 4 Associations Between Bibliographical Characteristics and Methodological Quality of Individual AMSTAR 2 Domain Item: Binary Logistic Regression^a

| AMSTAR 2 Item (Dependent Variable) | Predictor | AOR (95%CI) | P-value |
|--|---|-------------------|---------|
| 3. Did the review authors explain their selection of the study designs for inclusion in the review? | Update of previous review ^b | 6.93 (2.06–23.30) | 0.002 |
| | Published more recently | 1.29 (1.02–1.62) | 0.031 |
| 11. If meta-analysis was performed did the review authors use appropriate methods for statistical combination of results? | Published more recently | 1.29 (1.06–1.57) | 0.010 |
| | Corresponding author from Asia ^c | 0.17 (0.04–0.78) | 0.023 |
| 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? | Published more recently | 1.47 (1.23–1.76) | <0.001 |
| | Published more recently | 1.47 (1.23–1.76) | <0.001 |
| 13. Did the review authors account for RoB in individual studies when interpreting/discussing the results of the review? | Published more recently | 1.47 (1.23–1.76) | <0.001 |
| | Published more recently | 1.32 (1.05–1.66) | 0.016 |
| 14. Did the review authors provide a satisfactory explanation for, and discussion of, any heterogeneity observed in the results of the review? | Published more recently | 1.33 (1.03–1.72) | 0.026 |
| | Higher journal impact factor | 2.98 (1.67–5.30) | <0.001 |
| 16. Did the review authors report any potential sources of conflict of interest, including any funding they received for conducting the review? | Published more recently | 1.33 (1.03–1.72) | 0.026 |
| | Higher journal impact factor | 2.98 (1.67–5.30) | <0.001 |

Notes: In the model, dependent variable is the rating for each item (No = 0 (reference); Yes = 1); ^aP-values for all Hosmer–Lemeshow tests were >0.1, indicating good-model-fit for all logistic regression analysis; ^bSRs that are not updates of previous reviews were used as reference group. ^cSRs led by corresponding author from Europe were used as reference group.

Abbreviations: AMSTAR 2, Assessing the Methodological Quality of Systematic Reviews 2; AOR, adjusted odds ratio; CI, confidence interval; SR, systematic review.

Table 5 Associations Between Bibliographical Characteristics and Methodological Quality of Individual AMSTAR 2 Domain Item: Multinomial Logistic Regression^a

| AMSTAR 2 Item (Dependent Variable) | Predictor | AOR (95%CI)(1/0) | P-value | AOR (95%CI)(2/0) | P-value |
|--|--|------------------|---------|--------------------|---------|
| 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? | Published more recently | 1.46 (1.04–2.06) | 0.028 | 1.86 (1.17–2.96) | 0.009 |
| 8. Did the review authors describe the included studies in adequate detail? | Update of previous review ^e | – | – | 11.65 (2.38–56.92) | 0.002 |

Notes: In the model, dependent variable is the rating for each item (No = 0 (reference); Partial yes = 1; Yes = 2); ^aP-values > 0.1 for all Pearson and deviance tests, and P < 0.05 for all likelihood ratio tests, indicating good-model-fit for all logistic regression analysis. ^eSRs that are not updates of previous reviews were used as reference group.

Abbreviations: AMSTAR 2, Assessing the Methodological Quality of Systematic Reviews 2; AOR, adjusted odds ratio; CI, confidence interval; SR, systematic review.

Discussion

Summary of Findings

This cross-sectional study appraised the methodological quality of 119 SRs on LC treatments published between 2014 and 2023, involving both pharmacological and non-pharmacological interventions as well as Cochrane and non-Cochrane reviews. Among them, only one (0.8%) SR was evaluated as high overall quality, one (0.8%) with moderate quality, 9 (7.6%) with low quality and more than 90.8% with critically low quality. Our study indicated much improvement for most of the appraised SRs, especially in registering a priori protocols of reviews, explaining the selection of the study designs, using a comprehensive literature search strategy, listing all excluded studies with reasons, describing the included studies in adequate detail, and reporting the sources of funding of primary studies.

For an individual AMSTAR 2 item, SRs published more recently performed well on registering a priori protocol, explaining the selection of study design, applying appropriate methods into statistical combination of meta-analysis results, evaluating and interpreting the impact of risk of bias in primary studies, providing an explanation for observed heterogeneity, and reporting conflict of interest. Due to the incorporation of the latest methodologies and recent evidence, which enhances the clarity and comprehensiveness of their findings, updated SRs would do better on explaining the selection of study design and describing the included studies in adequate detail.⁹ Compared to the SRs with corresponding author in Europe, SRs with Asian authors tended not to applying appropriate methods into statistical combination of meta-analysis results. Specifically, 87% of SRs with Asian authors inappropriately employed a fixed-effect model despite significant heterogeneity among the studies included in these reviews. SRs with higher journal impact factors had better performance on reporting the conflict of interest.

The progressively stringent demands for methodological quality criteria in journals could potentially account for the observed association between higher impact factors and more recent publication years with better performance on AMSTAR 2 items.^{17,18} Along with the continuous progress of scientific research, the majority of journals are increasingly focusing on the methodological quality of literature, aiming to enhance their academic prestige and influence. Consequently, they have set rigorous requirements for the articles they intend to publish in their journals, particularly those with higher impact factors and academic reputations.

Most of English journals required the authors to declare any conflict of interest before publication, while it is not compulsory for most of the Chinese journals.¹⁹ In addition, the impact factors of all 18 publications published in the Chinese language were zero because their journals were not included in the Journal Citation Reports (JCR), which may also be an explanation that SRs with higher impact factors would perform better on reporting the conflict of interest.

Recommendations for Conducting SR in the Future

Registering the Protocol of Systematic Reviews a Priori

The protocol is an essential component for a high quality SR, which presents its research plan and ensure its reproducibility and transparency.²⁰ Review authors can also be benefited from registering a protocol, because it allows

them to design and conduct a research plan carefully, reduces duplication on the same research topic, and enables other researchers to judge the validity of reviews.²⁰ Without a protocol registered prior to conducting reviews, the researchers are susceptible to selection bias caused by selective outcome gained from changing research methods on purpose.²¹ The trend shows that more SR protocols were registered on a protocol registration database, like PROSPERO (international prospective register of systematic reviews), which may support our findings that the SR was more likely to register a protocol if it was published recently.^{22,23} In Cochrane reviews it is compulsory to provide an a priori protocol, it is highly recommended for non-Cochrane researchers to register their protocols.²⁴

Providing a Rigorous Explanation for the Selection of the Study Designs

A randomized controlled trial (RCT) is proved the best primary study design for assessing the effects of treatments, because randomization can minimize the confound sourced from different groups characteristics.²⁵ Although reviewers may have a short description about the eligibility criteria of an SR, it is necessary to provide a rigorous explanation for the selection of the study designs, even if it limited the eligibility criteria of an SR to RCT only. The possibility exists that a review restricted to RCTs would give an incomplete summary of the effects of a treatment if there were no relevant RCTs or missing outcomes in available RCTs.²⁶ In our findings, only 19.3% SRs justified the study design selection with reasonable strategy. SR authors are advised to improve the rigor of the SR in terms of explaining study design.

Conducting a Comprehensive Literature Search

Conducting a comprehensive literature search is an essential process of a SR to acquire an unbiased sample of primary studies. The rigor and results of a SR would be affected by the publication bias, if review authors did not provide their search strategy with supplement, or search gray literature in some cases.²⁷ The studies with positive results of one treatment are more likely to get published than those with no significant results regardless of their scientific rigor. However, unpublished studies, including gray literature, and those consulting from experts in the field or from the reference list of published reviews, can explore new evidence and that can be helpful for reducing publication bias.²⁸ In our study, we found only four (3.4%) SRs that conducted a comprehensive literature search, while 112 SRs (94.1%) did not take into account gray literature or consulting experts. Future reviewers should allocate sufficient resources to ensure the availability of gray literature searches and access to field experts.

Listing All Excluded Studies and Justifying Their Exclusions

The list of excluded studies with rationales for exclusion could minimize the subjective influences in the literature selection, improving the transparency and reproducibility of a SR in the review process, and preventing the risk of exclusion error.¹³ This approach may alleviate issues related to restrictive eligibility criteria in certain systematic reviews, enabling evidence users to access pertinent studies that are deemed valuable for addressing specific clinical questions. In our findings, only 6.8% of SRs offered a list of excluded studies with excluded reasons. This indicates that a significant proportion of included SRs are at high risk of being exposed to conclusion bias, which could lead to exaggeration or underestimation of results.⁹ Although it is not compulsory for non-Cochrane authors, we still highly recommend that future SR authors, editors, and peer reviewers take this issue into consideration.

Describing the Included Studies in Adequate Detail and Reporting on Their Resources of Funding

While describing the primary studies included in the SR, reviewers are supposed to present their study subjects, interventions, controls, outcomes, design, analysis, settings, and funding resources of the studies. It can help evidence users evaluate external validity and applicability of the findings, facilitate the exploration of clinical heterogeneity, and judge whether the financial interest exists.¹³ Multiple investigations have also demonstrated that studies funded by commercial sponsors are more prone to producing results that favor the sponsors' products.²⁹ Unfortunately, only 12.7% SRs offered adequate details of included studies, and 5.1% SRs reported funding resources. Therefore, there is much room for improvement.

Comparisons with Similar Studies

Our results showed that the proportion of SR with high methodological quality (0.8%) is similar to those focused on osteoporosis treatments (1.0%) and atopic dermatitis treatments (1.9%), but much lower than those focused on asthma treatment (8.8%) and Alzheimer's disease (AD) treatments (3.9%).^{10,11,30,31} In terms of evaluating performance of an

individual AMSTAR 2 item, our sample had better results on reporting PICO in the research questions and inclusion criteria, conducting study selection and data extraction in duplicate, and providing an explanation for heterogeneity in the results, compared with SRs focused on AD treatments, osteoporosis treatments, atopic dermatitis treatments. However, similar to SRs focused on atopic dermatitis and osteoporosis treatments, no critical item showed satisfactory performances.

Strengths and Limitations

The strength of this study lies in using AMSTAR 2 for assessing the most up-to-date sample of SRs on LC treatments. AMSTAR 2 has been proved to be a robust tool widely applied into evaluating the methodological quality of SRs, and we utilized it to conduct critical appraisal and reported the performance of SRs on each of the AMSTAR 2 items offering recommendations in specific domains.^{16,32} Furthermore, we performed a comprehensive search to identify eligible SRs focused on LC treatments, covering pharmacological treatments, non-pharmacological interventions.

These limitations should be noted. First, appraisal results may be affected by insufficient information reported by published SR papers, which could limit the accuracy of our assessment. Second, we only included SRs published in English or Chinese, which may reduce the representativeness of our sample. Third, given the small number of SRs addressed with African corresponding authors (n=2), the finding of SRs with corresponding author from Africa presented higher overall quality may be very unstable. However, this study has tried to search literature within the latest 10 years, but only identified two eligible SRs from African corresponding authors. If manpower and material resources are sufficient, further studies are suggested to extend the search time, which might identify more relevant studies to validate our results. Fourth, this study only considered seven confounders in the regression model. We did not include other potential confounders, such as language barriers and resource differences due to time and manpower limitations. Future studies should address this by incorporating language variables and regional resource indicators to refine the associations. This cross-sectional design did not assess the concrete impact of methodological deficiencies on reported efficacy estimates. Future research specifically comparing effect sizes across SRs of varying methodological quality, or re-analyzing data accounting for deficiencies, is needed to quantify the potential distortion these deficiencies cause.

Implications

The majority of the SRs on LC treatments were of critically low methodological quality in our study. The flaws existed in the SRs would mislead readers into significantly overestimating or underestimating the effects of LC treatments. Therefore, politicians, medical professionals, and other evidence users should carefully evaluate the methodological quality of SRs before implementing the evidence, in order to guarantee well-informed policy and clinical decisions. The Cochrane Handbook and AMSTAR 2 guidelines are highly recommended to be followed by authors, editors, and reviewers to promote the scientific integrity of SRs for publication.

Conclusions

The methodological quality of SRs on LC treatments was explicitly substandard. Future SR authors are supposed to improve quality of SRs through registering the protocol of systematic reviews a priori, providing a rigorous explanation for the selection of the study designs, conducting a comprehensive literature search strategy, listing all excluded studies and justifying their exclusion reasons, describing the included studies in adequate detail, and reporting their funding resources.

Acknowledgment

This research received no funding and involved no participants.

Author Contributions

All authors made a significant contribution to the work reported, whether that is in the conception, study design, execution, acquisition of data, analysis and interpretation, or in all these areas; took part in drafting, revising or critically reviewing the article; gave final approval of the version to be published; have agreed on the journal to which the article has been submitted; and agree to be accountable for all aspects of the work.

Disclosure

The authors report no conflicts of interest in this work.

References

1. Gravitz L. Liver cancer. *Nature*. 2014;516(7529):S1–S. doi:10.1038/516S1a
2. Liver cancer statistics | world cancer research fund international [Internet]. WCRF international. Available from: <https://www.wcrf.org/cancer-trends/liver-cancer-statistics/>. Accessed September 11, 2025.
3. Runggay H, Arnold M, Ferlay J, et al. Global burden of primary liver cancer in 2020 and predictions to 2040. *J Hepatol*. 2022;77(6):1598–1606. doi:10.1016/j.jhep.2022.08.021
4. Yang JD, Hainaut P, Gores GJ, Amadou A, Plymth A, Roberts LR. A global view of hepatocellular carcinoma: trends, risk, prevention and management. *Nat Rev Gastroenterol Hepatol*. 2019;16(10):589–604. doi:10.1038/s41575-019-0186-y
5. Forner A, Reig M, Bruix J. Hepatocellular carcinoma. *Lancet*. 2018;391(10127):1301–1314. doi:10.1016/S0140-6736(18)30010-2
6. Xiang Z, Wu J, Li J, Zheng S, Wei X, Xu X. Gut microbiota modulation: a viable strategy to address medical needs in hepatocellular carcinoma and liver transplantation. *Engineering*. 2023;29:59–72.
7. Jiang C, Cai YQ, Yang JJ, et al. Radiomics in the diagnosis and treatment of hepatocellular carcinoma. *Hepatobiliary Pancreatic Dis Int*. 2023;22(4):346–351. doi:10.1016/j.hbpd.2023.03.010
8. Liu F, Chen L, Wu Q, et al. Radiomics of dynamic contrast-enhanced MRI for predicting radiation-induced hepatic toxicity after intensity modulated radiotherapy for hepatocellular carcinoma: a machine learning predictive model based on the SHAP methodology. *J Hepatocell Carcinoma*. 2025;12:999–1015. doi:10.2147/JHC.S523448
9. Higgins JPT, Thomas J, Chandler J, et al. *Cochrane Handbook for Systematic Reviews of Interventions Version 6.4*. Cochrane; 2023. Available from: www.training.cochrane.org/handbook. Accessed September 11, 2025.
10. Wu X, Deng Y, Wang H, Chen YC, Wong CA, Chung VCH. Methodological quality of systematic reviews and meta-analysis on asthma treatments: a cross-sectional study. *Ann Am Thoracic Soc*. 2020;17(8):949–957.
11. Zhong CCW, Zhao J, Wong CHL, et al. Methodological quality of systematic reviews on treatments for Alzheimer’s disease: a cross-sectional study. *Alzheimer’s Res Ther*. 2022;14(1):159.
12. Chen CY, Kang YN, Kuo KN, Glasziou P, Chen KH. Increasing retractions of meta-analyses publications for methodological flaw. *Syst Rev*. 2021;10(1). doi:10.1186/s13643-021-01822-2
13. 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(8122). doi:10.1136/bmj.j4008
14. Montori VM, Wilczynski NL, Morgan D, Haynes RB. Optimal search strategies for retrieving systematic reviews from Medline: analytical survey. *BMJ*. 2004;330(7482):68. doi:10.1136/bmj.38336.804167.47
15. Health information research unit, evidence-based health informatics search strategies for PsycINFO in ovid syntax. Hamilton, ON, Canada: McMaster University; 2016. Available from: https://hiru.mcmaster.ca/hiru/HIRU_Hedges_PsycINFO_Strategies.aspx. Accessed June 5, 2024.
16. 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
17. Ali UA, Beata TR, van Gooszen HG, et al. Journal impact factor and methodological quality of surgical randomized controlled trials: an empirical study. *Langenbecks Arch Surg*. 2017;402(7):1015–1022. doi:10.1007/s00423-017-1593-6
18. Hong JU, Kim JH, Lee KH, et al. Characteristics, trend, and methodological quality of systematic reviews and meta-analyses in nuclear medicine. *Medicine*. 2019;98(21):e15785. doi:10.1097/MD.00000000000015785
19. Zhu J, Sun J. Conflicts of interest disclosure policies among Chinese medical journals: a cross-sectional study. *PLoS One*. 2019;14(7):e0219564. doi:10.1371/journal.pone.0219564
20. Bandara W, Syed R. The role of a protocol in a systematic literature review. *J Decision Systems*. 2023;1–18.
21. Kirkham JJ, Altman DG, Williamson PR. Bias due to changes in specified outcomes during the systematic review process. *PLoS One*. 2010;5(3):e9810. doi:10.1371/journal.pone.0009810
22. Allers K, Hoffmann F, Mathes T, Pieper D. Systematic reviews with published protocols compared to those without: more effort, older search. *J Clin Epidemiol*. 2018;95:102–110. doi:10.1016/j.jclinepi.2017.12.005
23. Booth A, Clarke M, Dooley G, et al. The nuts and bolts of PROSPERO: an international prospective register of systematic reviews. *Syst Rev*. 2012;1(1). doi:10.1186/2046-4053-1-2
24. Higgins JPT, Lasserson T, Chandler J, Tovey D, Churchill R. *Methodological Expectations of Cochrane Intervention Reviews*. London: Cochrane; 2019.
25. Sibbald B, Roland M. Understanding controlled trials: why are randomised controlled trials important? *BMJ*. 1998;316(7126):201. doi:10.1136/bmj.316.7126.201
26. McKenzie JE, Brennan SE, Ryan RE, Thomson HJ, Johnston RV, Thomas J. Chapter 3: defining the criteria for including studies and how they will be grouped for the synthesis. In: Higgins JPT, Thomas J, Chandler J, et al. editors. *Cochrane Handbook for Systematic Reviews of Interventions Version 6.4*. Cochrane; 2022. Available from: www.training.cochrane.org/handbook. Accessed June 12, 2024.
27. Easterbrook PJ, Gopalan R, Berlin JA, Matthews DR. Publication bias in clinical research. *Lancet*. 1991;337(8746):867–872. doi:10.1016/0140-6736(91)90201-Y
28. Adams J, Hillier-Brown FC, Moore HJ, et al. Searching and synthesising “grey literature” and “grey information” in public health: critical reflections on three case studies. *Syst Rev*. 2016;5(1):1–11. doi:10.1186/s13643-016-0337-y
29. Lexchin J. Pharmaceutical industry sponsorship and research outcome and quality: systematic review. *BMJ*. 2003;326(7400):1167–1170. doi:10.1136/bmj.326.7400.1167
30. Tsoi AKN, Ho L, Xin I, et al. Methodological quality of systematic reviews on treatments for osteoporosis: a cross-sectional study. *Bone*. 2020;139:115541. doi:10.1016/j.bone.2020.115541

31. Ho L, Kei M, Chung C, Irene xinyin W, Mao C, Chi V. Methodological quality of systematic reviews on atopic dermatitis treatments: a cross-sectional study. *J Dermatological Treat.* 2024;35(1). doi:10.1080/09546634.2024.2343072
32. Ho L, Ke FYT, Wong CHL, et al. Low methodological quality of systematic reviews on acupuncture: a cross-sectional study. *BMC Med Res Method.* 2021;21(1). doi:10.1186/s12874-021-01437-0

Journal of Hepatocellular Carcinoma

Dovepress
Taylor & Francis Group

Publish your work in this journal

The Journal of Hepatocellular Carcinoma is an international, peer-reviewed, open access journal that offers a platform for the dissemination and study of clinical, translational and basic research findings in this rapidly developing field. Development in areas including, but not limited to, epidemiology, vaccination, hepatitis therapy, pathology and molecular tumor classification and prognostication are all considered for publication. The manuscript management system is completely online and includes a very quick and fair peer-review system, which is all easy to use. Visit <http://www.dovepress.com/testimonials.php> to read real quotes from published authors.

Submit your manuscript here: <https://www.dovepress.com/journal-of-hepatocellular-carcinoma-journal>