Comparison of Reporting and Transparency in Published Protocols and Publications in Umbrella Reviews: Scoping Review

Liang Zhao¹, MM; Caiyi Shen², BMEDSC; Ming Liu¹, MD; Jiaoyan Zhang³, MM; Luying Cheng^{3,4}, MM; Yuanyuan Li³, MM; Lanbin Yuan², BMEDSC; Junhua Zhang⁵, Prof Dr; Jinhui Tian¹, Prof Dr

¹Evidence-Based Medicine Center, School of Basic Medical Sciences, Lanzhou University, Lanzhou, China

²The First Clinical Medical College, Lanzhou University, Lanzhou, China

⁴Zigong First People's Hospital, Zigong, China

⁵Evidence-Based Medicine Center, Tianjin University of Traditional Chinese Medicine, Tianjin, China

Corresponding Author:

Jinhui Tian, Prof Dr Evidence-Based Medicine Center School of Basic Medical Sciences Lanzhou University No. 199, Donggang West Road Chengguan District Lanzhou, 730000 China Phone: 86 18059240145 Email: tjh996@163.com

Abstract

Background: Inconsistencies between a protocol and its umbrella review (UR) may mislead readers about the importance of findings or lead to false-positive results. Furthermore, not documenting and explaining inconsistencies in the UR could reduce its transparency. To our knowledge, no study has examined the methodological consistency of the protocols with their URs and assessed the transparency of the URs when generating evidence.

Objective: This study aimed to investigate the inconsistency of protocols with their URs in the methodology and assess the transparency of the URs.

Methods: We searched medical-related electronic databases from their inception to January 1, 2022. We investigated inconsistencies between protocols and their publications and transparencies in the search strategy, inclusion criteria, methods of screening and data extraction, quality assessment, and statistical analysis.

Results: We included 31 protocols and 35 publications. For the search strategy, 39 inconsistencies between the protocols and their publications were found in 26 of the 35 (74%) URs, and 16 of these inconsistencies were indicated and explained. There were 84 inconsistencies between the protocols and their URs regarding the inclusion criteria in 31 of the 35 (89%) URs, and 29 of the inconsistencies were indicated and explained. Deviations from their protocols were found in 12 of the 32 (38%) URs reporting the methods of screening, 14 of the 30 (47%) URs reporting the methods of data extraction, and 11 of the 32 (34%) URs reporting the methods for quality assessment. Of the 35 URs, 6 (17%) were inconsistent with their protocols in terms of the tools for quality assessment; one-half (3/6, 50%) of them indicated and explained the deviations. As for the statistical analysis, 31 of the 35 (89%) URs generated 61 inconsistencies between the publications and their protocols, and 16 inconsistencies were indicated and explained.

Conclusions: There was a high prevalence of inconsistencies between protocols and publications of URs, and more than one-half of the inconsistencies were not indicated and explained in the publications. Therefore, how to promote the transparency of URs will be a major part of future work.

(J Med Internet Res 2023;25:e43299) doi: 10.2196/43299



³Evidence-Based Nursing Center, School of Nursing, Lanzhou University, Lanzhou, China

KEYWORDS

umbrella reviews; protocol; publication; inconsistency; transparency

Introduction

Umbrella reviews (URs) are clusters that encompass previously published systematic reviews or meta-analyses that consider many comparisons of interventions for the management of the same disease or condition, to provide an overall examination [1]. URs not only offer the possibility to address a broad scope of issues related to a topic of interest but are also ideal to highlight where the evidence base for a question is consistent or if contradictory findings exist and to explore and detail the reasons why [2]. In recent years, URs have become increasingly influential in biomedical literature, and they represent one of the highest levels of evidence synthesis currently available [3,4].

However, since a broad synthesis of data from many systematic reviews requires a high level of subject matter expertise and methodological skills, not all published URs follow clearly described, standardized methodology, which may reduce the credibility of evidence from URs [2,4-6]. Published protocols could leverage the strength of peer review to help authors refine their study design [7]. Furthermore, they allow the tracking of inconsistencies between the protocol and its publication, and disclosed and explained inconsistencies make it possible for readers to evaluate potential bias [8-10].

Inconsistencies may exist because of conscious or subconscious manipulation to reach the desired conclusion, which misleads readers about the importance of findings or causes false-positive results [11-13]. In some cases, valid reasons may exist for modifying a protocol in the course of undertaking a study. However, deviations from the protocol are often poorly documented and explained, and this makes it impossible to tell whether the changes introduce bias, thus reducing the transparency of the evidence [14,15]. Unfortunately, inconsistencies between protocols and their publications are common in randomized controlled trials, systematic reviews, and core outcomes set in the medical field, and most inconsistencies are not disclosed and discussed in the publications, making their transparency very worrying [14-16]. Previous studies that have assessed the reporting and methodological quality of URs have suggested that the transparency of URs is deficient [17]. However, to our knowledge, no study has examined the methodological consistency of a protocol with its UR and assessed its transparency by determining whether inconsistencies were documented and explained in the UR.

The primary aim of this study was to investigate inconsistencies between a protocol and its publication when developing an UR. The secondary aim was to assess the transparency of the URs when generating evidence.

Methods

Eligibility Criteria

URs that met the following criteria were included: (1) A UR was defined as a quantitative or qualitative synthesis of

```
https://www.jmir.org/2023/1/e43299
```

medical-related systematic reviews, (2) a UR protocol and its publication along with their methodological descriptions were available, (3) URs' protocols and their publications were published in peer-reviewed journals or preprints, (4) the UR was part of a study, (5) the UR was published in English. There was no restriction on the time frame of the studies. We excluded the following studies: (1) methodological reviews of URs, scoping or rapid reviews, qualitative reviews, integrative reviews, and evidence synthesis and (2) abstracts, conference proceedings, and letters to editors.

Search Strategy

A comprehensive literature search was performed in 7 medical-related databases, including PubMed, Embase, Web of Science, The Cochrane Library, Joanna Briggs Institute (JBI) Database of Systematic Reviews and Implementation Reports, PROSPERO, and Open Science Framework (OSF), from the inception of these databases to January 1, 2022. The search strategies were developed by an expert in literature searches (JT), and the full search strategies are presented in Table S1 in Multimedia Appendix 1. We also manually searched the reference lists of included studies for potentially eligible studies.

Study Selection

We managed all records using EndNote (version X9, Clarivate Analytics) software and removed duplicates using automatic and manual screening. The study selection was divided into 4 steps. First, 2 authors (LZ and ML) independently screened all URs' protocols and publications using title, abstract, and full text, consecutively. Second, we read the full text of identified publications to find the protocols mentioned in the articles (JZ and LC). Third, for protocols that were not related to included publications in the previous step, the study names or standard abbreviations (where applicable) combined with the term "umbrella" were used to search PubMed, Embase, The Cochrane Library, and Web of Science for their publications (YL, CS, and LY). In cases where the same protocol was described in multiple publications, all related publications were included. Finally, the reference lists of included URs were manually searched for additional studies. Any disagreements were resolved through discussion between authors and, if necessary, consultation with senior authors (JZ and JT).

Data Extraction

We developed the data extraction form and revised it after piloting it on a random sample of 20% of included URs (LZ, CS, and ML). The extracted data consisted of 2 parts: general characteristics and methodological characteristics. A list of extracted items is given in Table S2 in Multimedia Appendix 1. For any observed inconsistencies between a protocol and its publication, we investigated whether the inconsistencies had been disclosed and explained in the publications and then extracted the text details of the explanation. Of the 5 authors, 2 authors (LZ and JZ) independently extracted data from the protocols and their publications, and another 3 authors (ML, CS, and LC) checked the extracted data. Conflicts were resolved

XSL•FO

by discussion between authors and, if necessary, consultation with senior authors (JZ and JT).

Inconsistency and Transparency Assessment

According to the JBI manual for URs [2], we assessed inconsistencies between protocols and their publications in 6 areas: search strategy, inclusion criteria, methods for screening, methods for data extraction, methodological quality assessment, and statistical analysis. For each area, the text details where an inconsistency existed were extracted from both the protocol and its publication. Inconsistency was defined as follows: any deviation between a protocol and its publication that altered the substance or meaning of an area or fully reported in the protocol or UR but only briefly or not reported in the UR or protocol [18,19]. The definition of consistency was the substance or meaning regarding the compared area was identical in the protocol and its publication, so changes in style, wording, tense, or abbreviations were not considered [19]. The inconsistency assessment in each area was based on an internal standard that was developed, independently pilot-tested (in 5 pairs), and revised by 2 authors (LZ and ML; Table S3 in Multimedia

Appendix 1). Transparency was defined as deviations from a protocol that were disclosed and explained in its UR [15].

Statistical Analysis

The frequency of inconsistencies and transparency were our primary outcomes. Median (IQR) was used for continuous variables, including the number of authors, journal impact factor, total number of studies included, and interval between the protocol and its publication, whereas frequency (%) was used for categorical variables. The statistical analyses were conducted with SPSS Statistics v26.0 (IBM Corp).

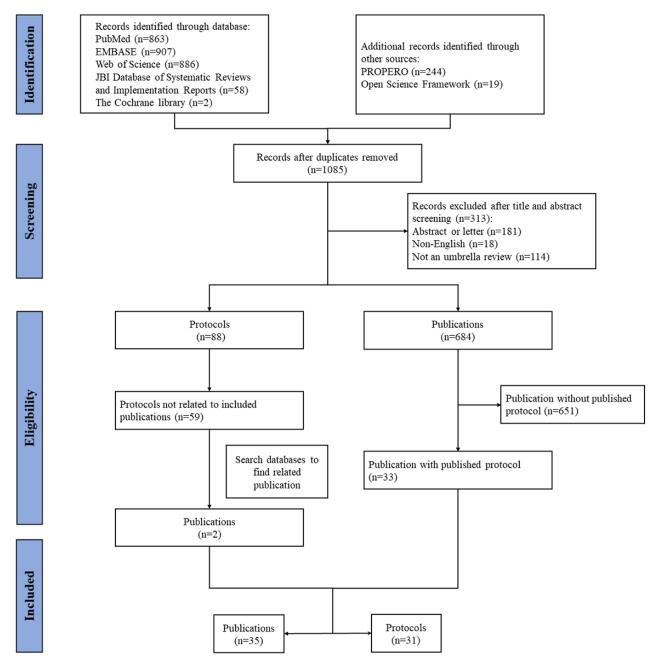
Results

Search and Selection Results

The database search yielded 2716 records, and an additional 263 records were identified through other sources. Following the removal of duplicates and title and abstract screening, 88 protocols and 684 publications were screened using the full text. In total, we identified 31 protocols and 35 publications (Figure 1). The full list of included URs can be found in Multimedia Appendix 2.



Figure 1. The flowchart of the screening process.



Characteristics of Included URs

The included URs were published between 2012 and 2021, and most of them (60/66, 91%) were published after 2016 (Multimedia Appendix 3). The interval between the protocols and their publications ranged from 1 month to 5 years and had a median interval of 2 (IQR 1-2) years. These URs were conducted in 14 countries, with Australia ranking first in both protocols and publications (Multimedia Appendix 4). Of the protocols, 30 were published in 4 journals with a median journal impact factor of 3.007 (IQR 0.000-3.136), whereas 34

publications were published in 27 journals with a median journal impact factor of 3.312 (IQR 2.329-5.926). The median number of authors of the protocols (median 6, IQR 4-8 authors) was the same as the publications (median 6, IQR 5-9 authors). Regarding registrations, 23 URs were registered in PROSPERO or OSF. Compared with the protocols, more publications indicated sources of financial support (23 vs 30) and conflicts of interests (2 vs 4). The number of reviews included in URs ranged from 1 to 114, with a median of 14 (IQR 7-36). The characteristics of the included URs are summarized in Table 1, Table S4 in Multimedia Appendix 5, and Multimedia Appendix 6.



Zhao et al

Table 1. The characteristics of the included umbrella reviews' protocols and publications.

Characteristics		Protocols (n=31)	Publications (n=35)
Journal impact factor ^a , median (IQR)		3.007 (0.000-3.136)	3.312 (2.329-5.926)
Journal impact factor, n (%)			
	0.0 to 3.0	1 (3)	6 (17)
	>3.0 to 6.0	20 (65)	14 (40)
	>6.0	0 (0)	8 (23)
	Non-SCI ^b	9 (29)	6 (17)
	Preprint	1 (3)	1 (3)
Number of authors, median (IQR)		6 (4-8)	6 (5-9)
Number of authors, n (%)			
	2-5	13 (42)	13 (37)
	6-10	15 (48)	17 (49)
	>10	3 (10)	5 (14)
Registration, n (%)			
	PROSPERO	21 (68)	23 (66)
	Open Science Framework	1 (3)	1 (3)
	Not registered	9 (29)	11 (31)
Funding, n (%)			
	Yes	18 (58)	26 (74)
	No	5 (16)	4 (11)
	Not reported	8 (26)	5 (14)
Conflicts of interest, n (%)			
	Yes	2 (6)	4 (11)
	No	29 (94)	31 (89)

^aThe journal impact factor was determined using the 2021 Journal Citation Report.

^bSCI: Science Citation Index.

Inconsistencies and Transparency in the Search Strategy

All URs described the search strategy, and 39 inconsistencies between the protocols and their publications were found in 26

of the 35 URs (74%). Of 39 inconsistencies, 16 (41%) were indicated and explained in the publications (Table 2, Table S5 in Multimedia Appendix 5, and Multimedia Appendix 6).



Table 2. The frequency and transparency of inconsistencies in 6 fields in the 35 publications included in the umbrella reviews (URs).

Items	Publications in the URs reporting each item, n	Frequency, n (%)	Transparency, n (%)
Search strategy	35	26 (74)	16 (41)
Literature sources	35	21 (60)	3 (14)
Language restrictions	34	5 (15)	4 (80)
Search time	35	13 (37)	9 (69)
Inclusion criteria	35	31 (89)	29 (35)
Participants	35	13 (37)	5 (38)
Interventions	35	16 (46)	6 (38)
Comparators	26	10 (38)	4 (40)
Outcomes	35	19 (54)	7 (37)
Type of studies	35	13 (37)	2 (15)
Other inclusion criteria	24	13 (54)	5 (38)
Methods for screening	32	12 (38)	2 (17)
Methods for data extraction	30	14 (47)	3 (21)
Quality assessment	35	13 (37)	6 (35)
Methods for quality assessment	32	11 (34)	3 (27)
Tools for quality assessment	35	6 (17)	3 (50)
Statistical analysis	35	31 (89)	16 (26)
Overlap	18	3 (17)	0 (0)
Certainty of evidence	24	13 (54)	2 (15)
Summary of findings	26	12 (46)	3 (25)
Data analysis	35	9 (26)	4 (44)
Effect size	13	11 (85)	2 (18)
Other statistical analysis	21	14 (67)	5 (36)

Of the 35 publications, 21 (60%) deviated from their protocols in the literature sources: 14 deviations occurred in the databases, and 12 URs generated deviations regarding other literature sources. Of the 21 URs, 3 (14%) indicated inconsistencies in the publications. Compared with the protocols, 1 publication expanded the range of language for included studies, and the other added language restrictions to the included studies. Although 3 publications did not describe language restrictions, their protocols did. In the 5 URs, 4 (80%) inconsistencies were disclosed in the URs. Regarding search times, 13 of the 35 (37%) publications differed from their protocols: 6 because of updates, 3 for other reasons, and the remaining 4 lacked an explanation.

Inconsistencies and Transparency in Inclusion Criteria

All URs described the inclusion criteria, and 31 of the 35 (89%) URs had inconsistencies between the protocols and their publications (Table 2, Table S6 in Multimedia Appendix 5, and Multimedia Appendix 6).

For the participants, 13 of the 35 (37%) publications were inconsistent with their protocols. In the 13 URs, 5 described the participants in detail in the protocols but not in the publications. Compared with the protocols, 6 narrowed the inclusion of the participants, and 2 expanded the inclusion of the participants.

```
https://www.jmir.org/2023/1/e43299
```

XSL•FO RenderX Of the 13 publications, 5 (38%) indicated and explained the inconsistencies.

Of the 35 URs, 16 (46%) generated the following inconsistencies between the protocols and their publications in the interventions: (1) 6 publications narrowed the scope of the interventions compared with the protocols; (2) compared with the protocols, the scope of the interventions was extended in 4 publications; (3) 3 publications were contradictory to the protocols; (4) 3 publications did not describe the interventions, while their protocols did. Of the 16 publications, 6 (38%) pointed out and interpreted the inconsistencies.

Regarding the comparators, 10 of the 26 (38%) URs showed deviations between the protocols and their publications: 4 publications increased the number of comparators compared with their protocols, 4 protocols but not their publications described the comparators, and 2 publications removed comparators that were included in the protocols. Of these 10 URs, 4 (40%) publications indicated where the inconsistencies arose.

Inconsistencies between the protocols and their publications in the outcomes were found in 19 of the 35 (54%) URs. The most common inconsistency was that the publications reduced the number of outcomes compared with their protocols (8 URs).

The second was that the publications added outcomes (4 URs). The third was that the protocols but not the publications described the outcomes (4 URs). Finally, 3 publications changed the outcomes described in the protocols. Inconsistencies were indicated in 7 of the 19 (37%) publications.

Inconsistencies between the protocols and their publications regarding the type of studies were found in 13 of the 35 (37%) URs: (1) 6 publications narrowed the inclusion of the study design compared with their protocols; (2) 5 publications expanded the inclusion of the study design, and 1 of them added inclusion of updated randomized controlled trials; (3) 2 protocols but not their publications described the type of included studies. The reasons for deviations were explained by 2 of the 13 (15%) URs.

Other inclusion criteria were reported by 24 of the 35 (69%) URs, of which 14 (58%) URs reported the requirements for included studies, 11 (46%) URs reported the quality of included studies, and only 2 (8%) URs reported the requirements of the protocol. In the other inclusion criteria, there were the following inconsistencies between the protocols and their publications in 13 of the 24 (54%) URs: 8 inconsistencies in the requirements for included studies, 5 inconsistencies in the quality of included studies, and 2 inconsistencies in the requirements of the protocol. The inconsistencies were detailed in 5 of the 13 (38%) publications.

Inconsistencies and Transparency in Screening, Quality Assessment, and Data Extraction

The following inconsistencies between the protocols and their publications were found in the methods for screening, data extraction, and quality assessment (Table 2, Table S7 in Multimedia Appendix 5, and Multimedia Appendix 6): (1) 14 inconsistencies were identified when resolving disagreements among reviewers; (2) 8 URs did not describe the methods, while their protocols did; (3) 8 inconsistencies were found in the number of reviewers; (4) 7 publications described the methods, while the protocols did not. The publications described 8 of the inconsistencies.

All URs described tools for quality assessment; the most common (13/35, 37%) was the JBI Critical Appraisal Checklist for Systematic Reviews and Research Syntheses. The tools used in the protocols were changed in 3 publications, and 3 publications did not depict tools, while their protocols did. The deviations were indicated in 3 of the 6 (50%) publications.

Inconsistencies and Transparency in Statistical Analyses

We found 62 inconsistencies between the protocols and their publications in 31 of the 35 (89%) URs, and 16 of the 62 (26%) inconsistencies were indicated and explained in the publications (Table 2, Table S8 in Multimedia Appendix 5, and Multimedia Appendix 6).

Of the 35 URs, 18 described how to deal with the overlaps of primary studies between included systematic reviews. There were inconsistencies between 3 of the 18 (17%) protocols and their publications. Of them, 2 protocols gave ways to deal with overlaps, while their publications did not implement them; 1

```
https://www.jmir.org/2023/1/e43299
```

publication added the calculation of the corrected covered area. None of the inconsistencies was indicated in the publications.

Tools for assessing the certainty of the evidence were reported by 24 URs, and 20 of these URs used the Grading of Recommendations Assessment, Development, and Evaluation (GRADE) criteria. Inconsistencies existed between 13 of the 24 (54%) URs and their protocols: 10 protocols but not their publications reported assessments of the certainty of the evidence, and 3 publications added tools for assessing the credibility of evidence that were not described in the protocols. Inconsistencies were detailed by 2 of the 13 (15%) URs.

What findings to present and how to present them were described in 26 URs: 12 of the 26 (46%) URs generated inconsistencies between the protocols and their publications, and 3 URs indicated and explained inconsistencies in the publications. Compared with the protocols, inconsistencies could be categorized as follows: (1) 5 publications differed on whether to apply the "summary of evidence" table, (2) the forms of evidence presentation were changed in 5 publications, and (3) 2 publications changed the reported elements.

All URs depicted the data analysis, and inconsistencies between the protocols and their publications were found in 9 of the 35 (26%) URs. Compared with the protocols, 3 publications removed quantitative analysis, 2 publications added quantitative analysis, 2 publications altered the rules for qualitative synthesis, 1 publication changed the quantitative analysis model, and 1 protocol did not describe the data analysis while the publication did. Of the 9 inconsistencies, 4 (44%) were detailed and explained in the publications.

How to standardize the effect sizes was reported in 13 URs, and 11 of these 13 (85%) URs had inconsistencies between the protocols and their publications: 8 URs only reported it in the publications, and 1 UR only reported it in the protocol; regarding the standardization of dichotomous data, 2 publications were inconsistent with their protocols. Of the 11 inconsistencies, 2 (18%) were indicated and explained in the publications.

Of the 35 URs, 21 described other statistical analyses, including subgroup analysis (17/35, 49%), sensitivity analysis (6/35, 17%), and publication bias or small study effects (6/35, 17%). Inconsistencies between the protocols and their publications were found in 14 of the 21 (67%) URs. Regarding subgroup analyses, 12 URs had inconsistencies with their protocols. Of the 12 URs, 9 only reported the subgroup analysis in the protocols, 2 only reported it in the publications, and 1 publication changed the subgroup analyses that were described in the protocol. Deviations between the protocols and their publications in publication bias or small study effects and sensitivity analysis were found in 4 URs and 1 UR, respectively. Of the 14 URs with inconsistencies, 5 (36%) detailed the inconsistencies in the publications.

Discussion

We identified 31 protocols and 35 publications of the URs. Inconsistencies between the protocols and their publications were found in all areas, with inconsistencies occurring most frequently in inclusion criteria and statistical analyses. In

XSL•FO

addition, less than one-half of the inconsistencies were indicated in the publications of URs, and the transparency of URs was inadequate.

Inconsistencies in Methodology

Our study showed inconsistencies in the search strategy between the protocols and their publications in 74% (26/35) of the URs, which was similar to previous result that indicated 77% of reviews changed the search strategy between the protocols and their publications [18]. Changes in the database and other literature sources between the protocols and their publications were frequent in the URs. This significantly affects the completeness and comprehensiveness of the included studies [20]. Although it has been recommended that biomedical citation databases complemented by other literature sources is the best combination for identifying comprehensive reviews, there was no guidance on which specific databases should be searched when conducting URs [21-23]. Thus, the specific combination of literature sources that could identify more available evidence should be clarified in the future. To avoid missing relevant citations, it seems like a beneficial change for the publication to expand the range of languages for included studies compared with the protocol [24]. However, we only found this change in 1 UR [25]. In addition, we also found quite a few inconsistencies in the search time frame, but it was acceptable that the main reasons were to update or avoid outdated literature.

Although URs often aim to provide a broad overview of the evidence on a topic, it is still important to be specific about the inclusion criteria [26]. These criteria provide not only a guide for the readers to clearly understand what is proposed by the authors but also a guide for the reviewers themselves to make decisions about the studies to be included in the URs [2]. More importantly, changes in inclusion criteria were much less likely to be detected as a potential source of bias [27]. In this study, 89% (31/35) of the URs had inconsistencies between their protocols and publications in the inclusion criteria, with the highest volume of inconsistencies in the outcomes, which was higher than the prevalence of inconsistencies in outcomes among the protocols and their publications of systematic reviews in a study [16] that included reviews published between 2002 and 2009. Modifications (that is, addition, removal, or reprioritization) of outcomes that have been based on prior knowledge of the results might introduce the possibility of bias into the URs, mislead clinical decision-makers, and possibly jeopardize medical safety, and adherence to the published protocol can be effective in avoiding potential bias [13,28]. Noteworthy, after the publication of the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) guidelines in 2009 [29], the prevalence of inconsistencies in outcomes was reduced in another study, which restricted it to reviews published between 2011 and 2014 [30]. However, a validated systematic reporting standard that specializes in URs is not available, which needs to be further developed.

In terms of the statistical analysis, we identified that 89% (31/35) of the URs' publications deviated from their protocols. This is similar to a finding for systematic reviews [18], in which 89% of publications were inconsistent with their protocols in the statistical analysis. One-quarter of the publications changed the

```
https://www.jmir.org/2023/1/e43299
```

data analysis described in the protocols. These changes might be a source of complexity and bias in the analysis and, worse still, lead to a significant variation in the conclusions [26,31]. Because of heterogeneity between the reviews, it might not always be feasible to conduct a quantitative synthesis in URs [32,33]. However, a narrative synthesis could become complex and open to bias if not adequately described, and there was a concern that synthesis errors at the review level could result in errors at the UR level [31]. Therefore, appropriate data analysis in URs needs to be further explored.

Over one-half of the included URs dealt with overlaps of primary studies, which was consistent with the findings by Pieper et al [34], and 12 URs did not report it in the publications. This could potentially lead to the statistical power being overestimated and thus risk producing a misleading, overly precise estimate [2,35,36]. Thus, how to deal with overlapping of primary studies in the URs should be strengthened. Although only one-third of the URs converted the results presented across the systematic reviews to one common summary statistic, 73% (8/11) of inconsistencies were shown by the fact that the differences were reported in the publications but not in the protocols. This seems like a beneficial change, as a common effect size could make the comparison more straightforward [4]. In addition, we found that the most common inconsistency between the protocol and their publication in the subgroup analysis, sensitivity analysis, and publication bias was their removal from publications. This would weaken the level of evidence of an effect and the richness of the picture of evidence [4]. These methods were not mentioned in the URs not only in our study but also in most of the guidelines [31]. Therefore, it is necessary to pay more attention to the subgroup analysis, sensitivity analysis, and publication bias in URs.

As for certainty of evidence, the most common inconsistency was that the GRADE was removed in the publications, which might lead to a diminished certainty of evidence generated by URs [2,26]. Although most of the guidelines recommended that the GRADE be used to appraise the certainty of evidence, how to apply the GRADE in URs was not yet available [31]. In addition, some criteria in the GRADE were only applicable to the level of primary studies when assessing the body of evidence, so it was inappropriate to be directly transferred to URs [37,38]. Therefore, a modification to the GRADE so it could also be applied to URs would be beneficial and practical, and many authors are already familiar with it. Furthermore, some URs [39-41] were found that added approaches to assess the certainty of evidence in the publications; it is also acceptable to include additional approaches to assess the certainty of evidence if they are objective and standardized.

Inadequate Transparency

Research transparency can speed scientific progress, can increase trust in science, and is one of the core values of science [42,43]. In this study, the transparency was inadequate; it fluctuated between 17% and 41% across the 6 areas, which was similar to the findings of previous studies [15,18]. The majority of inconsistencies between the protocols and their publications were not indicated and explained in the publications, and this made it difficult to provide definitive conclusions as to why

these changes may occur and if they introduced potential bias. The inadequate transparency of URs might have the cumulative effect of producing a distorted body of evidence with too few null effects and many false positives, exaggerating the effectiveness of programs; it could also threaten medical safety [44,45]. Therefore, how to promote the transparency of URs will be a major part of future work.

Publishing a protocol in a peer-reviewed journal is a crucial way to improve transparency, by comparing the protocol with a completed UR to detect whether unintended and undocumented changes were made [13]. However, in our study, the protocols were published in only 4 journals, while the publications were published in 27 journals. This demonstrates that the number of peer-reviewed journals that published the URs' protocols was very limited, and few protocols were published. Therefore, it is necessary to encourage more journals to accept protocols for publication. Furthermore, editors or peer reviewers could also compare publications in the URs with their protocols and check back to confirm whether the changes were indicated and explained to promote transparency.

A study published in the journal Science pointed out that systematic reporting standards also contribute to the transparency of research [42]. For example, PRISMA 2020 requires authors to describe and explain any amendments to the information provided in the protocol [46]. It provides authors with guidance to report and explain any inconsistencies. However, a validated UR-specific systematic reporting standard has not yet emerged and needs to be developed in the future. Furthermore, the current reward structure of publication fails to encourage transparent studies, which was evidenced by the greater likelihood of publication of statistically significant, novel, and theoretically tidy results than null, replicated, or perplexing results, even at the expense of transparency of research [42,44]. Therefore, greater awareness is needed that null results are as important as statistically significant results in helping others to more accurately assess the evidence base for a program.

Strengths and Limitations

To the best of our knowledge, our study was the first study to assess inconsistent reporting between the protocols and the publications of URs in the methodology. Because we comprehensively analyzed each step of the development of an UR, our findings could be considered beneficial to the production of URs. However, there were some limitations in our study. First, only URs published in English were included. Second, our study conducted only an exploratory analysis with a relatively small sample size, so some inconsistencies may be overlooked. Third, some publications depict only part of the methodology of an UR, resulting in incomplete comparisons between protocols and publications. Last, there is no standard protocol to address the analysis of inconsistencies between protocols and their publications; therefore, we created our protocol and analysis process based on previous similar studies. There is potential for bias or that other factors were not included.

Implications for Future Research and Practice

There are several areas for future development in the methodology of URs, including (1) the specific combination of literature sources that could identify more available evidence, (2) development of a validated reporting quality tool specifically for URs, (3) appropriate and adequate use of statistical analysis in URs, and (4) modification to the GRADE so that it can be applied in URs. It is also beneficial if our study can help medical editors and scientific journals draw attention to whether a UR's protocol exists and identify any unrecognized or unreasonable changes between the protocols and their publications during the editing process.

Conclusions

There was a high prevalence of inconsistencies between protocols and publications of URs, especially in the inclusion criteria and statistical analysis. More worryingly, more than one-half of URs did not indicate and explain the inconsistencies, and the transparency was inadequate. Thus, the authors of URs should be required to describe and explain any deviation from their protocols. Our study provides further evidence that published protocols allow the tracking of any changes that have taken place, to assess the transparency of URs. Therefore, authors of URs are encouraged to publish their protocols, and journals are urged to accept protocols for publication. In addition, editors or peer reviewers could compare publications of URs with their protocols and check back to confirm whether the changes were indicated and explained.

Acknowledgments

This study was funded by the Science and Technology Program of Gansu Province (20CX9ZA027, 20CX9ZA112). The funders had no role in the design and conduct of the study; management, analysis, and interpretation of the data; preparation, review, or approval of the manuscript; and decision to submit the manuscript for publication.

Authors' Contributions

JT, JZ, LZ, CS, and ML conceived and designed the research. LZ, ML, JZ, LC, YL, CS, and LY acquired the data. LZ, ML, CS, and JZ performed the statistical analysis. JT and JZ proofread the language. LZ, ML, CS, and LC interpreted the data. LZ wrote the manuscript. All authors revised the manuscript and approved the final version of the manuscript.

Conflicts of Interest

None declared.



Multimedia Appendix 1

Search strategy and extracted information. [DOCX File , 25 KB-Multimedia Appendix 1]

Multimedia Appendix 2

List of all included protocols and publications. [DOCX File , 23 KB-Multimedia Appendix 2]

Multimedia Appendix 3

Published years of the included umbrella reviews (URs)' protocols and publications (inner circle represents the UR protocol; outer circle represents the corresponding publication for the protocol; the length of the square is proportional to the year of publication).

[PNG File , 667 KB-Multimedia Appendix 3]

Multimedia Appendix 4

Countries of corresponding authors of the included URs' protocols and publications. [PNG File , 74 KB-Multimedia Appendix 4]

Multimedia Appendix 5

Information regarding the inconsistencies in the characteristics of publications compared with protocols. [DOCX File , 100 KB-Multimedia Appendix 5]

Multimedia Appendix 6

The specifics of inconsistencies between the protocol and its publication for each UR. [XLSX File (Microsoft Excel File), 26 KB-Multimedia Appendix 6]

References

- 1. Faulkner G, Fagan MJ, Lee J. Umbrella reviews (systematic review of reviews). Int Rev Sport Exerc Psychol 2021 Jun 11;15(1):73-90 [doi: 10.1080/1750984x.2021.1934888]
- 2. Aromataris E, Munn Z. JBI Manual for Evidence Synthesis. JBI. 2020. URL: <u>https://synthesismanual.jbi.global</u> [accessed 2023-06-19]
- Yang K, Zhang J, Zhao L, Cheng L, Li Y, Kang Y, et al. An umbrella review of Lianhua Qingwen combined with Western medicine for the treatment of coronavirus disease 2019. Acupunct Herb Med 2022 Sep;2(3):143-151 [doi: 10.1097/HM9.000000000000041]
- 4. Fusar-Poli P, Radua J. Ten simple rules for conducting umbrella reviews. Evid Based Ment Health 2018 Aug;21(3):95-100 [doi: 10.1136/ebmental-2018-300014] [Medline: 30006442]
- Ioannidis JPA. Integration of evidence from multiple meta-analyses: a primer on umbrella reviews, treatment networks and multiple treatments meta-analyses. CMAJ 2009 Oct 13;181(8):488-493 [FREE Full text] [doi: 10.1503/cmaj.081086] [Medline: 19654195]
- 6. Papatheodorou S. Umbrella reviews: what they are and why we need them. Eur J Epidemiol 2019 Jun;34(6):543-546 [doi: 10.1007/s10654-019-00505-6] [Medline: 30852716]
- Pavlovich MJ, Buttery S. How peer review and publication can make a good protocol even better. STAR Protoc 2021 May 14;2(2):100468 [FREE Full text] [doi: 10.1016/j.xpro.2021.100468] [Medline: 34036281]
- Munafò MR, Nosek BA, Bishop DVM, Button KS, Chambers CD, du Sert NP, et al. A manifesto for reproducible science. Nat Hum Behav 2017 Jan 10;1(1):0021 [FREE Full text] [doi: 10.1038/s41562-016-0021] [Medline: 33954258]
- Iqbal SA, Wallach JD, Khoury MJ, Schully SD, Ioannidis JPA. Reproducible research practices and transparency across the biomedical literature. PLoS Biol 2016 Jan 04;14(1):e1002333 [FREE Full text] [doi: 10.1371/journal.pbio.1002333] [Medline: 26726926]
- He Y, Li J, Li Y, Jin R, Wen Q, Li N, et al. Strengthening the quality of clinical trials of acupuncture: a guideline protocol. BMJ Open 2022 Jan 12;12(1):e053312 [FREE Full text] [doi: 10.1136/bmjopen-2021-053312] [Medline: 35022174]
- Williamson PR, Gamble C. Identification and impact of outcome selection bias in meta-analysis. Stat Med 2005 May 30;24(10):1547-1561 [doi: 10.1002/sim.2025] [Medline: 15580591]
- Schwartz LM, Woloshin S, Zheng E, Tse T, Zarin DA. ClinicalTrials.gov and Drugs@FDA: a comparison of results reporting for new drug approval trials. Ann Intern Med 2016 Sep 20;165(6):421-430 [FREE Full text] [doi: 10.7326/M15-2658] [Medline: 27294570]

RenderX

- 13. Shamseer L, Moher D, Clarke M, Ghersi D, Liberati A, Petticrew M, et al. Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015: elaboration and explanation. BMJ 2015 Jan 02;350:g7647 [FREE Full text] [doi: 10.1136/bmj.g7647] [Medline: 2555855]
- Dwan K, Altman DG, Cresswell L, Blundell M, Gamble CL, Williamson PR. Comparison of protocols and registry entries to published reports for randomised controlled trials. Cochrane Database Syst Rev 2011 Jan 19;2011(1):MR000031 [FREE Full text] [doi: 10.1002/14651858.MR000031.pub2] [Medline: 21249714]
- Liu M, Gao Y, Yuan Y, Shi S, Yang K, Lu C, et al. Inconsistency and low transparency were found between core outcome set protocol and full text publication: a comparative study. J Clin Epidemiol 2021 Mar;131:59-69 [doi: 10.1016/j.jclinepi.2020.11.009] [Medline: 33227446]
- Page MJ, McKenzie JE, Kirkham J, Dwan K, Kramer S, Green S, et al. Bias due to selective inclusion and reporting of outcomes and analyses in systematic reviews of randomised trials of healthcare interventions. Cochrane Database Syst Rev 2014 Oct 01;2014(10):MR000035 [FREE Full text] [doi: 10.1002/14651858.MR000035.pub2] [Medline: 25271098]
- 17. Li L, Tian J, Tian H, Sun R, Liu Y, Yang K. Quality and transparency of overviews of systematic reviews. J Evid Based Med 2012 Aug;5(3):166-173 [doi: 10.1111/j.1756-5391.2012.01185.x] [Medline: 23672223]
- 18. Silagy CA, Middleton P, Hopewell S. Publishing protocols of systematic reviews: comparing what was done to what was planned. JAMA 2002 Jun 05;287(21):2831-2834 [doi: <u>10.1001/jama.287.21.2831</u>] [Medline: <u>12038926</u>]
- Hu K, Zhao L, Zhou Q, Mei F, Gao Q, Chen F, et al. Inconsistencies in study eligibility criteria are common between non-Cochrane systematic reviews and their protocols registered in PROSPERO. Res Synth Methods 2021 May;12(3):394-405 [doi: <u>10.1002/jrsm.1476</u>] [Medline: <u>33522101</u>]
- Naik Y, Baker P, Ismail SA, Tillmann T, Bash K, Quantz D, et al. Going upstream an umbrella review of the macroeconomic determinants of health and health inequalities. BMC Public Health 2019 Dec 17;19(1):1678 [FREE Full text] [doi: 10.1186/s12889-019-7895-6] [Medline: 31842835]
- 21. Aromataris E, Riitano D. Constructing a search strategy and searching for evidence. A guide to the literature search for a systematic review. Am J Nurs 2014 May;114(5):49-56 [doi: 10.1097/01.NAJ.0000446779.99522.f6] [Medline: 24759479]
- Goossen K, Hess S, Lunny C, Pieper D. Database combinations to retrieve systematic reviews in overviews of reviews: a methodological study. BMC Med Res Methodol 2020 Jun 01;20(1):138 [FREE Full text] [doi: 10.1186/s12874-020-00983-3] [Medline: 32487023]
- 23. Rathbone J, Carter M, Hoffmann T, Glasziou P. A comparison of the performance of seven key bibliographic databases in identifying all relevant systematic reviews of interventions for hypertension. Syst Rev 2016 Feb 09;5:27 [FREE Full text] [doi: 10.1186/s13643-016-0197-5] [Medline: 26862061]
- 24. Pieper D, Puljak L. Language restrictions in systematic reviews should not be imposed in the search strategy but in the eligibility criteria if necessary. J Clin Epidemiol 2021 Apr;132:146-147 [doi: <u>10.1016/j.jclinepi.2020.12.027</u>] [Medline: <u>33383129</u>]
- 25. O'Malley N, Clifford AM, Conneely M, Casey B, Coote S. Effectiveness of interventions to prevent falls for people with multiple sclerosis, Parkinson's disease and stroke: an umbrella review. BMC Neurol 2021 Sep 29;21(1):378 [FREE Full text] [doi: 10.1186/s12883-021-02402-6] [Medline: 34587933]
- 26. Biondi-Zoccai G. Umbrella Reviews: Evidence Synthesis with Overviews of Reviews and Meta-Epidemiologic Studies. Cham, Switzerland: Springer International Publishing; 2016.
- 27. 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 Sep 21;358:j4008 [FREE Full text] [doi: 10.1136/bmj.j4008] [Medline: 28935701]
- 28. Pieper D, Allers K. Differences between protocols for randomized controlled trials and systematic reviews. J Clin Epidemiol 2018 Jun;98:144-145 [doi: 10.1016/j.jclinepi.2017.11.027] [Medline: 29222061]
- Moher D, Liberati A, Tetzlaff J, Altman DG, PRISMA Group. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. J Clin Epidemiol 2009 Oct;62(10):1006-1012 [FREE Full text] [doi: 10.1016/j.jclinepi.2009.06.005] [Medline: 19631508]
- Tricco AC, Cogo E, Page MJ, Polisena J, Booth A, Dwan K, et al. A third of systematic reviews changed or did not specify the primary outcome: a PROSPERO register study. J Clin Epidemiol 2016 Nov;79:46-54 [FREE Full text] [doi: 10.1016/j.jclinepi.2016.03.025] [Medline: 27079845]
- Gates M, Gates A, Guitard S, Pollock M, Hartling L. Guidance for overviews of reviews continues to accumulate, but important challenges remain: a scoping review. Syst Rev 2020 Nov 04;9(1):254 [FREE Full text] [doi: 10.1186/s13643-020-01509-0] [Medline: <u>33148319</u>]
- 32. Antonio MG, Petrovskaya O, Lau F. The state of evidence in patient portals: umbrella review. J Med Internet Res 2020 Nov 11;22(11):e23851 [FREE Full text] [doi: 10.2196/23851] [Medline: 33174851]
- Apóstolo J, Cooke R, Bobrowicz-Campos E, Santana S, Marcucci M, Cano A, et al. Predicting risk and outcomes for frail older adults: an umbrella review of frailty screening tools. JBI Database System Rev Implement Rep 2017 Apr;15(4):1154-1208 [FREE Full text] [doi: 10.11124/JBISRIR-2016-003018] [Medline: 28398987]

RenderX

- Pieper D, Antoine SL, Mathes T, Neugebauer EAM, Eikermann M. Systematic review finds overlapping reviews were not mentioned in every other overview. J Clin Epidemiol 2014 Apr;67(4):368-375 [doi: <u>10.1016/j.jclinepi.2013.11.007</u>] [Medline: <u>24581293</u>]
- 35. Neimann Rasmussen L, Montgomery P. The prevalence of and factors associated with inclusion of non-English language studies in Campbell systematic reviews: a survey and meta-epidemiological study. Syst Rev 2018 Aug 23;7(1):129 [FREE Full text] [doi: 10.1186/s13643-018-0786-6] [Medline: 30139391]
- 36. Ballard M, Montgomery P. Risk of bias in overviews of reviews: a scoping review of methodological guidance and four-item checklist. Res Synth Methods 2017 Mar;8(1):92-108 [doi: 10.1002/jrsm.1229] [Medline: 28074553]
- Jadczak AD, Makwana N, Luscombe-Marsh N, Visvanathan R, Schultz TJ. Effectiveness of exercise interventions on physical function in community-dwelling frail older people: an umbrella review of systematic reviews. JBI Database System Rev Implement Rep 2018 Mar;16(3):752-775 [doi: 10.11124/JBISRIR-2017-003551] [Medline: 29521871]
- Guyatt G, Oxman AD, Akl EA, Kunz R, Vist G, Brozek J, et al. GRADE guidelines: 1. Introduction-GRADE evidence profiles and summary of findings tables. J Clin Epidemiol 2011 Apr;64(4):383-394 [doi: <u>10.1016/j.jclinepi.2010.04.026</u>] [Medline: <u>21195583</u>]
- Moore L, Tardif PA, Lauzier F, Bérubé M, Archambault P, Lamontagne F, et al. Low-value clinical practices in adult traumatic brain injury: an umbrella review. J Neurotrauma 2020 Dec 15;37(24):2605-2615 [FREE Full text] [doi: 10.1089/neu.2020.7044] [Medline: <u>32791886</u>]
- 40. Marano L, Fusario D, Savelli V, Marrelli D, Roviello F. Robotic versus laparoscopic gastrectomy for gastric cancer: an umbrella review of systematic reviews and meta-analyses. Updates Surg 2021 Oct;73(5):1673-1689 [FREE Full text] [doi: 10.1007/s13304-021-01059-7] [Medline: 34031848]
- 41. Avşar TS, McLeod H, Jackson L. Health outcomes of smoking during pregnancy and the postpartum period: an umbrella review. BMC Pregnancy Childbirth 2021 Mar 26;21(1):254 [FREE Full text] [doi: 10.1186/s12884-021-03729-1] [Medline: 33771100]
- 42. Miguel E, Camerer C, Casey K, Cohen J, Esterling KM, Gerber A, et al. Social science. Promoting transparency in social science research. Science 2014 Jan 03;343(6166):30-31 [FREE Full text] [doi: 10.1126/science.1245317] [Medline: 24385620]
- 43. Mayo-Wilson E, Grant S, Supplee L, Kianersi S, Amin A, DeHaven A, et al. Evaluating implementation of the Transparency and Openness Promotion (TOP) guidelines: the TRUST process for rating journal policies, procedures, and practices. Res Integr Peer Rev 2021 Jun 02;6(1):9 [FREE Full text] [doi: 10.1186/s41073-021-00112-8] [Medline: 34078479]
- 44. Ioannidis JPA. Why most published research findings are false. PLoS Med 2005 Aug;2(8):e124 [FREE Full text] [doi: 10.1371/journal.pmed.0020124] [Medline: 16060722]
- 45. Humphreys M, Sanchez de la Sierra R, van der Windt P. Fishing, commitment, and communication: a proposal for comprehensive nonbinding research registration. Polit Anal 2013 Jan 04;21(1):1-20 [doi: <u>10.1093/pan/mps021</u>]
- Page MJ, McKenzie JE, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. BMJ 2021 Mar 29;372:n71 [FREE Full text] [doi: 10.1136/bmj.n71] [Medline: 33782057]

Abbreviations

GRADE: Grading of Recommendations Assessment, Development, and Evaluation
JBI: Joanna Briggs Institute
OSF: Open Science Framework
PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analyses
URs: umbrella reviews

Edited by T Leung; submitted 07.10.22; peer-reviewed by L Li, S Pranic, L Guo; comments to author 08.12.22; revised version received 15.02.23; accepted 05.05.23; published 02.08.23

<u>Please cite as:</u> Zhao L, Shen C, Liu M, Zhang J, Cheng L, Li Y, Yuan L, Zhang J, Tian J Comparison of Reporting and Transparency in Published Protocols and Publications in Umbrella Reviews: Scoping Review J Med Internet Res 2023;25:e43299 URL: <u>https://www.jmir.org/2023/1/e43299</u> doi: <u>10.2196/43299</u> PMID:

RenderX

©Liang Zhao, Caiyi Shen, Ming Liu, Jiaoyan Zhang, Luying Cheng, Yuanyuan Li, Lanbin Yuan, Junhua Zhang, Jinhui Tian. Originally published in the Journal of Medical Internet Research (https://www.jmir.org), 02.08.2023. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in the Journal of Medical Internet Research, is properly cited. The complete bibliographic information, a link to the original publication on https://www.jmir.org/, as well as this copyright and license information must be included.

