Received: 20 January 2021 | Revised: 21 April 2021 | Accepted: 23 April 2021

DOI: 10.1111/jep.13578

ORIGINAL PAPER

Journal of Evaluation in Clinical Practice
International Journal of Public Health Policy and Health Services Research

WILEY

Methodological assessment of systematic reviews and meta-analyses on COVID-19: A meta-epidemiological study

Kristine J. Rosenberger MS[1] | Chang Xu PhD[2] | Lifeng Lin PhD[1]

[1]Department of Statistics, Florida State University, Tallahassee, Florida

[2]Department of Population Medicine, College of Medicine, Qatar University, Doha, Qatar

Correspondence
Lifeng Lin, Department of Statistics, Florida State University, 411 OSB, 117 N Woodward Avenue, Tallahassee, FL 32306.
Email: linl@stat.fsu.edu

Funding information
National Center for Advancing Translational Sciences, Grant/Award Number: UL1 TR001427; U.S. National Library of Medicine, Grant/Award Number: R01 LM012982

Abstract

Rationale, aims, and objectives: COVID-19 has caused an ongoing public health crisis. Many systematic reviews and meta-analyses have been performed to synthesize evidence for better understanding this new disease. However, some concerns have been raised about rapid COVID-19 research. This meta-epidemiological study aims to methodologically assess the current systematic reviews and meta-analyses on COVID-19.

Methods: We searched in various databases for systematic reviews with meta-analyses published between 1 January 2020 and 31 October 2020. We extracted their basic characteristics, data analyses, evidence appraisal, and assessment of publication bias and heterogeneity.

Results: We identified 295 systematic reviews on COVID-19. The median time from submission to acceptance was 33 days. Among these systematic reviews, 73.9% evaluated clinical manifestations or comorbidities of COVID-19. Stata was the most used software programme (43.39%). The odds ratio was the most used effect measure (34.24%). Moreover, 28.14% of the systematic reviews did not present evidence appraisal. Among those reporting the risk of bias results, 14.64% of studies had a high risk of bias. Egger's test was the most used method for assessing publication bias (38.31%), while 38.66% of the systematic reviews did not assess publication bias. The I^2 statistic was widely used for assessing heterogeneity (92.20%); many meta-analyses had high values of I^2. Among the meta-analyses using the random-effects model, 75.82% did not report the methods for model implementation; among those meta-analyses reporting implementation methods, the DerSimonian-Laird method was the most used one.

Conclusions: The current systematic reviews and meta-analyses on COVID-19 might suffer from low transparency, high heterogeneity, and suboptimal statistical methods. It is recommended that future systematic reviews on COVID-19 strictly follow well-developed guidelines. Sensitivity analyses may be performed to examine how the synthesized evidence might depend on different methods for appraising evidence, assessing publication bias, and implementing meta-analysis models.

KEYWORDS
COVID-19, heterogeneity, meta-analysis, publication bias, risk of bias, systematic review

1 | INTRODUCTION

Since December 2019, the pandemic caused by severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) and the associated coronavirus disease 2019 (COVID-19) has been an ongoing public health crisis.[1] As of 20 April 2021, over 141 million cases, including over 3 million deaths, have been reported to the WHO (https://covid19.who.int/). The analysis of data from individuals affected with COVID-19 is integral for understanding the clinical characteristics, disease progression, and potential treatments and outcomes. The publication time of articles related to COVID-19 has decreased by an average of 49% during the pandemic,[2] owing largely to the expedited peer-review process.[2,3] While the gravity of the COVID-19 pandemic warrants expedited efforts, there are concerns about the quality of peer reviews and resulting publications as the spread of misinformation could have harmful consequences.[4-11]

Individual studies have limited ability to summarize the current state of research; thus, many efforts have been made to conduct systematic reviews and synthesize the presently available results for policymaking during the pandemic. The evidence synthesis is achieved via meta-analyses, which play a critical role in developing new research by determining whether the proposed study is necessary and helping design the study.[12,13] Systematic reviews and meta-analyses are considered of the highest quality and can be viewed as a lens through which evidence is evaluated.[14-17]

However, systematic reviews and meta-analyses do not always provide a rationale for their methodology; if not used properly, they could produce ambiguous results and exacerbate research errors.[18-20] Even if only randomized controlled trials (RCTs) are synthesized, systematic reviews and meta-analyses themselves are essentially observational studies, which are subject to reporting bias.[21,22] It is critical to properly perform and adequately report systematic reviews and meta-analyses using rigorous methods,[23] particularly during the fast-evolving pandemic.[6,16,24-28] To standardize and improve the quality of reporting, the preferred reporting items for systematic reviews and meta-analyses (PRISMA) statement was introduced in 2009.[29] Similar guidelines, such as meta-analysis of observational studies in epidemiology (MOOSE), can be used for other specific types of research. Additionally, appraisal tools, such as a measurement tool to assess systematic reviews (AMSTAR), have been proposed to assess the methodological validity of systematic reviews.[30,31] Although the overall quality of systematic reviews has generally improved after the implementation of the reporting guidelines and appraisal tools,[32,33] there is still room for improvement.[34-36]

While rapid syntheses of existing evidence are imperative for understanding this novel disease, the quality of the current systematic reviews and meta-analyses on COVID-19 should be carefully and critically evaluated to ensure the reliability of the synthesized evidence. This meta-epidemiological study aims to summarize the state of meta-analysis research on COVID-19 and inform the conduct of future meta-analyses.

2 | METHODS

We searched in the databases CINAHL, Embase, PubMed, and PsycINFO for systematic reviews with meta-analyses published between 1 January 2020 and 31 October 2020. Of note, we did not search the Cochrane Database of Systematic Reviews (CDSR) as it was indexed in PubMed. We focused on systematic reviews with meta-analyses because they were more likely than those without meta-analyses to give intuitive conclusions for healthcare interventions. Articles only available on preprint servers and the grey literature were not considered. The search terms were 'meta-analysis' AND 'COVID-19' OR 'coronavirus' OR 'nCoV' OR 'SARS-CoV-2'. The search was restricted to English. Studies were excluded if they were duplicate publications (where only the latest versions were used in our study), letters to the editor, short correspondences, or study protocols.

From each systematic review, we extracted the journal name, the dates of submission, acceptance, online publication, and publication in issue, the total number and types of included studies, and the number of meta-analyses. A review may contain multiple meta-analyses on different outcomes or intervention comparisons. For each meta-analysis, we extracted the number of studies, outcome name and categorization, effect measure, evidence appraisal, assessment of publication bias, assessment of heterogeneity, statistical methods used for implementation, whether a prediction interval was reported, whether a meta-regression was performed, and whether a network meta-analysis was performed. These were done by the first author (K.J.R.) and were further double-checked by the last author (L.L.).

This study did not require ethical approval because all results were based on published data. This article contains several methodological terminologies for meta-analyses. Appendix A provides brief introductions to them and their references.

3 | RESULTS

We identified 295 published systematic reviews on COVID-19 from 188 journals, including a total of 7518 studies and 2609 meta-analyses. In Appendix B, Figure S1 presents the flow chart of the literature search, and Table S1 gives a brief summary of countries of origin of the 295 systematic reviews. Table 1 provides a more detailed summary of these systematic reviews regarding the types of meta-analyses and studies and the methods used for analyses. The complete information of these systematic reviews is available at https://osf.io/ahnjb/.

3.1 | Basic characteristics

Among these systematic reviews, the number of days from submission to acceptance ranged from 0 to 154 (median = 33, interquartile range [IQR] = 16-57). The number of days from acceptance to online publication ranged from 2 to 179 (median = 49, IQR = 28-81), and the number of days from acceptance to issue publication ranged from 22 to 241 (median = 83, IQR = 57-107). Figure 1 shows the month of online publication for the 278 systematic reviews that reported their publication dates. March was the first month in which systematic reviews were published online ($n = 10$), and most systematic reviews were published in July ($n = 57$).

TABLE 1 Summary of the 295 systematic reviews on COVID-19

	Count (%)
Meta-analysis type	
Clinical manifestation/comorbidity	218 (73.90%)
Diagnostic test	26 (8.81%)
Preventative intervention	4 (1.36%)
Treatment comparison	47 (15.93%)
Study type	
Case–control	40 (13.56%)
Case report	11 (3.73%)
Case series	40 (13.56%)
Cohort	227 (76.95%)
Controlled NRSI[a]	2 (0.68%)
Cross-sectional	28 (9.49%)
Non-controlled NRSI[a]	2 (0.68%)
Non-randomized controlled trial	1 (0.34%)
Randomized controlled trial	39 (13.22%)
Analysis software	
CMA	27 (9.15%)
GraphPad Prism	1 (0.34%)
JASP	2 (0.68%)
MedCalc	3 (1.02%)
Meta-Analyst	1 (0.34%)
Meta-DiSc	1 (0.34%)
MetaXL	14 (4.75%)
Network Analyst tool	1 (0.34%)
OpenMeta Analyst	12 (4.07%)
R	55 (18.64%)
RevMan	66 (22.37%)
SAS	1 (0.34%)
SPSS	4 (1.36%)
Stata	128 (43.39%)
StatsDirect	1 (0.34%)
TIBCO	1 (0.34%)
Not reported	17 (5.76%)
Meta-regression included	
No	228 (77.29%)
Yes	67 (22.71%)
Network meta-analysis included	
No	291 (98.64%)
Yes	4 (1.36%)
Prediction interval reported	
No	286 (96.98%)
Yes	9 (3.02%)
Statistical methodology	
Bayesian method	2 (0.68%)
Frequentist method	293 (99.32%)
Bayesian method: prior distribution	

(Continues)

TABLE 1 (Continued)

	Count (%)
Half-Cauchy(0,1)	1 (0.34%)
Not reported	1 (0.34%)

Abbreviation: CMA, comprehensive meta-analysis.
[a]Non-randomized studies of intervention.

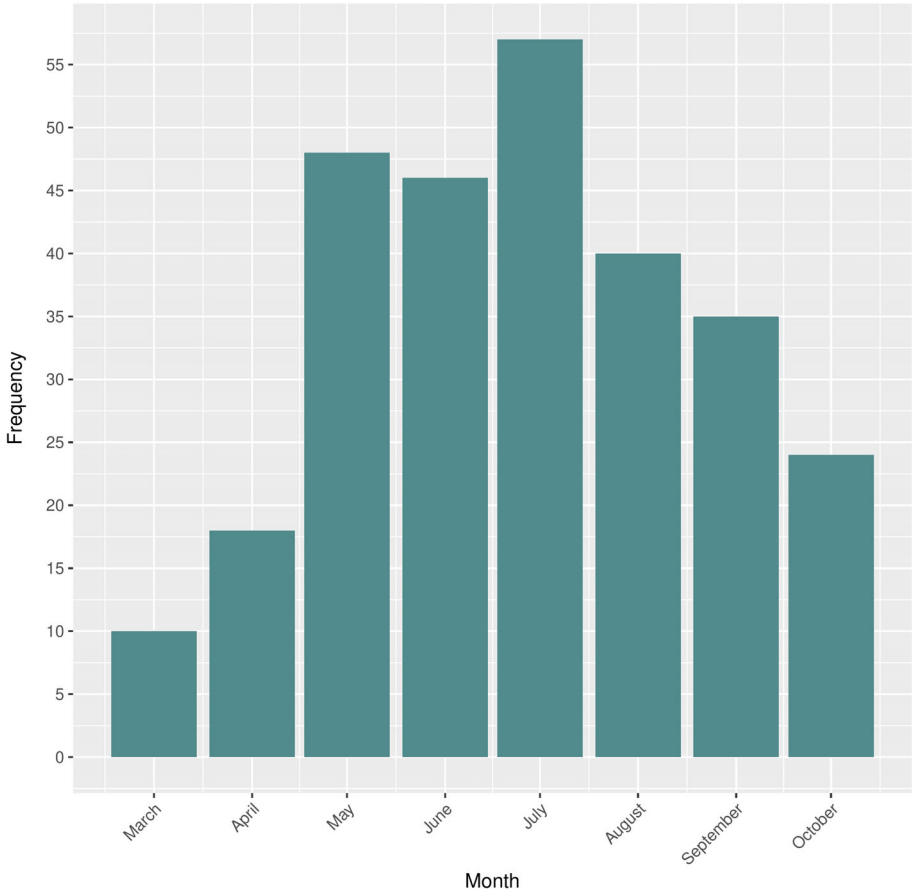

FIGURE 1 Publication time of the 295 systematic reviews

The number of meta-analyses within a systematic review ranged from 1 to 84 (median = 5, IQR = 2-10). The number of studies within a meta-analysis ranged from 2 to 189 (median = 18, IQR = 9-30).

In addition, 218 (73.90%) systematic reviews evaluated clinical manifestations or comorbidities of COVID-19, 47 (15.93%) evaluated treatment comparisons, 26 (8.81%) evaluated diagnostic tests, and 4 (1.36%) evaluated preventative interventions. The most common study type in the systematic reviews was cohort studies (n = 227, 76.95%), followed by case series and case-control studies, each included in 40 systematic reviews (13.56%). Experimental studies were used in 44 (14.92%) systematic reviews, of which 39 (13.22%) were RCTs.

3.2 | Data analyses

A total of 17 statistical software programmes were used for implementing the meta-analyses. The most frequently used software programme was Stata (n = 128, 43.39%), followed by RevMan (Review Manager, n = 66, 22.37%), R (n = 55, 18.64%), and comprehensive meta-analysis (CMA; n = 27, 9.15%). Meta-regression was included in 67 (22.71%) systematic reviews, network meta-analysis was implemented in 4 (1.36%), and 9 (3.02%) reported prediction intervals. Also, 2 (0.68%) systematic reviews used Bayesian methods, where the results were based on the posterior distributions of parameters of interest after assigning certain prior information. Among them, 1 used the half-Cauchy (0,1) prior for the heterogeneity SD, while the other did not report the prior distribution.

Table 2 lists a total of 39 effect measures specified in the systematic reviews. The odds ratio (OR), used in 101 (34.24%) systematic reviews, was the most common effect measure. Prevalence, used in 70 (23.73%) systematic reviews, was the second most common. Moreover, 26 types of effect measures were used in less than 10 systematic reviews.

TABLE 2 Effect measures specified in the 295 systematic reviews on COVID-19

Effect measure	Count (%)
Basic reproduction number	1 (0.34%)
Cluster proteins	1 (0.34%)
Diagnostic accuracy	1 (0.34%)
Diagnostic likelihood ratio	1 (0.34%)
Diagnostic odds ratio	2 (0.68%)
Event rate	3 (1.02%)
Frequency	1 (0.34%)
Hazard ratio	6 (2.03%)
Incidence	12 (4.07%)
Incidence rate	2 (0.68%)
Incubation period	1 (0.34%)
Likelihood ratio	1 (0.34%)
Mean difference	17 (5.76%)
Mean	7 (2.37%)
Meta-correlation	1 (0.34%)
Meta-median difference	1 (0.34%)
Mortality rate	1 (0.34%)
Odds ratio	101 (34.24%)
Positive rate	2 (0.68%)
Prevalence	70 (23.73%)
Prevalence ratio	1 (0.34%)
Proportion	10 (3.39%)
R_0	2 (0.68%)
Rate	3 (1.02%)
Rate difference	3 (1.02%)
Rate ratio	3 (1.02%)
Relative ratio	1 (0.34%)
Relative risk	12 (4.07%)
Reproduction number	1 (0.34%)
Risk	1 (0.34%)
Risk difference	2 (0.68%)
Risk ratio	35 (11.86%)
SD	1 (0.34%)
Sensitivity	14 (4.75%)
Standardized mean difference	18 (6.10%)
Specificity	12 (4.07%)
Subnetwork ranking	1 (0.34%)
Time to event	1 (0.34%)
Weighted mean difference	23 (7.80%)

Note: The terms of effect measures were extracted from the original systematic reviews, regardless of their appropriateness.

3.3 | Evidence appraisal and assessment of publication bias

Table 3 summarizes the evidence appraisal and assessment of publication bias. A total of 35 different assessment tools were identified for appraising evidence, including those used for individual studies and overall evidence. The tools depended both on the types of meta-analyses performed and the types of studies included. We extracted the information about evidence appraisal exactly as stated in each systematic review, regardless of whether it was appropriate. There was inconsistent terminology across the systematic reviews. For example, some systematic reviews referred to the grading of recommendations, assessment, development and evaluations (GRADE) approach as the McMaster University critical appraisal tool. The Newcastle-Ottawa Scale (NOS) was implemented in 99 (33.56%) systematic reviews and was the most frequently used tool. The NOS was designed for observational studies. Many of the identified COVID-19 studies were observational; this possibly explains the relatively high proportion of systematic reviews using the NOS. Moreover, 83 (28.14%) systematic reviews did not present evidence appraisal.

Because there was often no universal guideline for what scores constituted high, moderate, and low risk of bias, the results of the risk of bias assessment could be subjective and diverse across the systematic reviews. A total of 4063 studies from 169 systematic reviews reported the results of the risk of bias assessment. Among these studies, 1863 (45.85%) were judged to have a low risk of bias, 517 (12.72%) had a low to moderate risk of bias, 1011 (24.48%) had a moderate risk of bias, 16 (0.39%) had a moderate to high risk of bias, 595 (14.64%) had a high risk of bias, and 61 (1.50%) had an unclear risk of bias.

Egger's test was used to assess publication bias in 113 (38.31%) systematic reviews, Deeks' method was used in 51 (17.29%), Begg's rank test was used in 41 (13.90%), and Harbord's test and the trim-and-fill method were each used in 3 (1.02%). Moreover, 4 (1.36%) systematic reviews did not specify the methods, and 117 (39.66%) did not include an assessment of publication bias. Of note, 39 of these systematic reviews had less than 10 studies. In this instance, the assessment of publication bias using funnel-plot-based methods (e.g., Egger's regression) was not recommended by some researchers; the assessment methods could produce large uncertainties by chance and thus would possibly not be reliable.[37] Additionally, some meta-analyses of proportions (e.g., disease prevalence) may not have clear directions of potential bias, so it may be challenging to assess publication bias.[38] Among the 178 systematic reviews that assessed publication bias, 98 (55.06%) concluded no publication bias was present, 61 (34.27%) detected publication bias, 13 (7.30%) contained insufficient studies for assessing publication bias, and 6 (3.37%) did not report the results. Of note, although many meta-analyses did not report the presence of publication bias, most methods for assessing publication bias usually had low statistical powers, particularly when the number of studies was small.[39] Therefore, some meta-analyses that claimed no publication bias could still be subject to potential bias.

3.4 | Heterogeneity

Table 4 summarizes the assessment of heterogeneity. The I^2 statistic was the most widely used method; it was included in 272 (92.20%) systematic reviews. The Q test was included in 119 (40.34%)

TABLE 3 Evidence appraisal and publication bias in the 295 systematic reviews on COVID-19

	Count (%)
Evidence appraisal: methods	
Agency for Healthcare Research and Quality (AHRQ) tool	4 (1.36%)
Appraisal tool for Cross-Sectional Studies (AXIS)	4 (1.36%)
British National Institute for Clinical Excellence	3 (1.02%)
Cochrane Risk of Bias tool for Non-Randomized Studies (RoB 2)	25 (8.47%)
Cochrane tool	1 (0.34%)
Consolidated Standards of Reporting Trials (CONSORT)	1 (0.34%)
Critical appraisal methodological index	1 (0.34%)
Grading of Recommendations, Assessment, Development and Evaluations (GRADE)	16 (5.42%)
Hoy et al[58]	2 (0.68%)
Ijaz et al[59]	1 (0.34%)
Institute of Health Economics case series methodological quality evaluation tool	3 (1.02%)
Jadad quality scoring standard	4 (1.36%)
Joanna Briggs Institute evidence summary	11 (3.73%)
Methodological index for non-randomized studies (MINORS)	7 (2.37%)
Methodological quality and synthesis of case-series and case-reports	1 (0.34%)
Mixed methods appraisal tool (MMAT)	1 (0.34%)
National Heart, Lung, and Blood Institute tool	2 (0.68%)
National Institutes of Health quality assessment tool	12 (4.07%)
Nature Publications Quality in Publication (NPQIP)	1 (0.34%)
Newcastle-Ottawa Scale (NOS)	99 (33.56%)
Non-Randomized Studies Methods Group (NRSMG)	1 (0.34%)
Oxford Center for Evidence-Based Medicine Critical Appraisal tool	3 (1.02%)
Quality Appraisal of Case Series	4 (1.36%)
Quality Assessment of Diagnostic Accuracy Studies-2 (QUADAS-2)	10 (3.39%)
Quality Assessment Tool for Observational Cohort and Cross-Sectional Studies (QAT-OC/CSS)	1 (0.34%)
Quality in Prognostic Studies (QUIPS) tool	1 (0.34%)
Risk of Bias Assessment tool for Non-Randomized Studies (RoBANS)	1 (0.34%)
Risk of Bias in Non-Randomized Studies - of Exposures (ROBINS-E)	1 (0.34%)
Risk of Bias in Non-Randomized Studies - of Interventions (ROBINS-I)	11 (3.73%)
Scottish Intercollegiate Guidelines Network	1 (0.34%)
Standards for Reporting of Diagnostic Accuracy Studies (STARD)	1 (0.34%)
Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) checklist	4 (1.36%)
Other assessment	4 (1.36%)
Not specified	1 (0.34%)
Not included	83 (28.14%)
Evidence appraisal: results (among 4063 studies from systematic reviews that reported risk of bias results)	
Low risk of bias	1863 (45.85%)
Low to moderate risk of bias	517 (12.72%)
Moderate risk of bias	1011 (24.48%)
Moderate to high risk of bias	16 (0.39%)
High risk of bias	595 (14.64%)
Unclear risk of bias	61 (1.50%)
Assessment of publication bias: methods	
Begg's rank test	41 (13.90%)
Deeks' method	51 (17.29%)

TABLE 3 (Continued)

	Count (%)
Egger's test	113 (38.31%)
Harbord's test	3 (1.02%)
Trim-and-fill method	3 (1.02%)
Not specified	4 (1.36%)
Not included	117 (39.66%)
Assessment of publication bias: results (among 178 systematic reviews that assessed publication bias)	
No publication bias	98 (55.06%)
Publication bias detected	61 (34.27%)
Not enough studies to assess	13 (7.30%)
Not reported	6 (3.37%)

TABLE 4 Assessment of heterogeneity and model type in the 295 systematic reviews on COVID-19

	Count (%)
Heterogeneity assessment[a]	
I^2 statistic	272 (92.20%)
Q test	119 (40.34%)
SROC[b] curve with 95% prediction region	1 (0.34%)
Between-study variance τ^2	13 (4.41%)
Visually evaluating forest plots	4 (1.36%)
Not included	16 (5.42%)
Meta-analysis model	
Both fixed-effect and random-effects models	9 (3.05%)
Fixed-effect model	5 (1.69%)
Fixed-effect model when $I^2 < 25\%$	2 (0.68%)
Fixed-effect model when $I^2 < 30\%$	1 (0.34%)
Fixed-effect model when $I^2 < 50\%$	58 (19.66%)
Fixed-effect model when $I^2 < 60\%$	1 (0.34%)
Quality-effects model	1 (0.34%)
Random-effects model	200 (67.79%)
Not reported	18 (6.10%)
Implementation of random-effects model[c]	
DerSimonian–Laird	55 (20.15%)
Hartung–Knapp	1 (0.37%)
Mantel–Haenszel	7 (2.56%)
Paule–Mandel	2 (0.73%)
Sidik–Jonkman	1 (0.37%)
Not reported	207 (75.82%)

[a]Assessment methods may overlap because a systematic review may use multiple methods.
[b]Summary receiver operating characteristic.
[c]Among 273 systematic reviews that performed the random-effects model.

systematic reviews, and the between-study variance τ^2 was estimated in 13 (4.41%) systematic reviews. The visual evaluation of forest plots was used in 4 (1.36%) systematic reviews; 1 (0.34%) systematic review of diagnostic tests assessed heterogeneity via the summary receiver operating characteristic (SROC) curve with a 95% prediction region. A heterogeneity assessment was not performed in 16 (5.42%) systematic reviews.

The random-effects model was used in 200 (67.79%) systematic reviews, and 9 (3.05%) used both the fixed-effect and random-effects models. Moreover, 62 (21.02%) systematic reviews used the random-effects model if the I^2 statistic surpassed a cutoff value and the fixed-effect model otherwise. Specifically, 1 (0.34%) systematic review chose the cutoff value of I^2 at 60%, 58 (19.66%) chose 50%, 1 (0.34%) chose 30%, and 2 (0.68%) chose 25%. The fixed-effect model was used in 5 (1.69%) systematic reviews, 1 (0.34%) systematic review used the quality-effects model, and 18 (6.10%) did not report the type of model used. Among the 273 systematic reviews that utilized the random-effects model, 207 (75.82%) did not report the methods used to implement the model. The DerSimonian-Laird method was used in 55 (20.15%) systematic reviews, the Mantel-Haenszel method was used in 7 (2.56%), the Paule-Mandel method was used in 2 (0.73%), and the Hartung–Knapp and Sidik–Jonkman methods were each used in 1 (0.37%).

Figure 2 presents the histogram of the I^2 statistics reported in 2235 meta-analyses among the 295 systematic reviews. Among these meta-analyses, 661 reported I^2 values of at least 90%. A considerable number of meta-analyses also reported I^2 within 0–10%, leading to a left-skewed distribution. Among those with I^2>10%, most meta-analyses tended to have larger I^2 values. The mean of the I^2 values was 61.44%, while the median was 75.55%.

4 | DISCUSSION

4.1 | Main findings

This meta-epidemiological study methodologically assessed a total of 295 systematic reviews on COVID-19. The median time between article submission and acceptance among these systematic reviews was 33 days, and the median time between acceptance and online publication was 49 days. These short time frames were evidence of

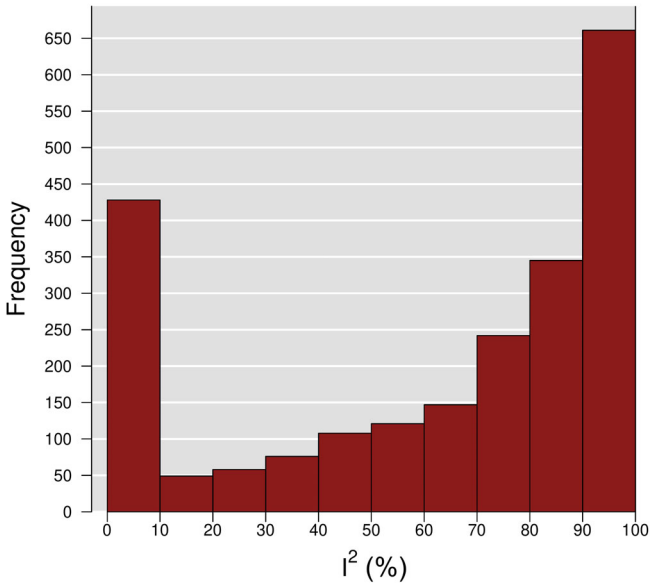

FIGURE 2 Heterogeneity measure I^2 of meta-analyses in the 295 systematic reviews

expedited peer review processes rationalized by the urgency of making information on the ongoing COVID-19 crisis accessible, even in the absence of high-quality evidence (e.g., from RCTs) that was not yet widely available. Most systematic reviews included in this study were conducted during the early stages of COVID-19 research when information on the underlying disease pathology was not widely known. This likely influenced the types of meta-analyses that have been performed and explains why most meta-analyses published by 31 October 2020 focused on clinical characteristics or comorbidities of COVID-19. The types of meta-analyses determined the types of studies included, the effect measures calculated, the assessment tool used for risk of bias, and the statistical methods used to perform the analyses.

Many systematic reviews used inconsistent terminologies and assessment criteria, especially for assessing the risk of bias, and they did not thoroughly report the methodology per the PRISMA statement. Many meta-analyses either did not present evidence appraisal or did not disclose the results if they claimed to have assessed the risk of bias. If a considerable number of studies are subject to a high risk of bias, the conclusions from meta-analyses must be interpreted cautiously.[40] Moreover, over one-third of the included systematic reviews did not assess publication bias; if studies were based on a biased sample of target populations, the meta-analyses might overestimate the effects of interventions. The lack of assessments for the risk of bias and publication bias is particularly of concern for COVID-19 research. As a new disease with many unknowns, most ongoing randomized studies on COVID-19 have relatively small sample sizes, and researchers should account for both time-sensitivity and evidence reliability.

We also found that many meta-analyses had a high level of between-study heterogeneity with I^2 greater than 90%. This may indicate that the synthesized evidence from the existing studies in these meta-analyses may not be reliably used for making decisions in future studies. Very few systematic reviews reported prediction intervals, which have been recommended as a valuable tool for appraising heterogeneity and understanding the effects of interventions in future study settings.[41,42] Given the potentially high heterogeneity presented in current meta-analyses on COVID-19, we strongly recommend that future meta-analysts report prediction intervals. Most meta-analyses using the random-effects model did not report the estimates of between-study variances, which are measures of heterogeneity as crucial as the I^2 statistic.[43]

In addition, many systematic reviews did not fully report important information about meta-analysis implementations, which may lead to a low level of transparency and reproducibility issues. For example, over 5% of the systematic reviews did not mention the software programmes used to perform statistical analyses. Among the systematic reviews that assessed publication bias, nearly 40% did not specify the methods used for the assessment. Various methods are available for assessing publication bias, and different methods could produce different results.[44,45] Among the systematic reviews that used the random-effects model, over 75% did not clearly specify how the model was implemented. Among the systematic reviews that specified implementation methods, the DerSimonian–Laird method was the most commonly used, but it has been shown that this method is inferior to several alternative methods.[46] The combination of low transparency, high heterogeneity, and suboptimal statistical methods could lead to concerns about whether the results of the current meta-analyses on COVID-19 should be trusted.

4.2 | Strengths and limitations

This study investigated a comprehensive collection of systematic reviews and meta-analyses on COVID-19 published by 31 October 2020. We have examined many methodological items that could critically affect the quality of systematic reviews and meta-analyses, including evidence appraisal, assessment of publication bias and heterogeneity, and statistical methods to implement meta-analysis models. We believe our findings could provide some timely suggestions for future meta-analysts to generate more reliable evidence for decision-making in the fast-evolving pandemic.

Nevertheless, this study has several limitations, and some further steps may be considered in the future. First, our literature search was restricted to the systematic reviews that had been published online, while many more unpublished papers on COVID-19 are available on preprint servers such as medRxiv. Many unpublished systematic reviews available on preprint servers also contributed important information about COVID-19. It may be worth investigating whether these preprints would be eventually published in peer-reviewed journals and how the conclusions might change from the preprint versions to published versions. Second, this study focused on the general topic of COVID-19 and did not distinguish different types of systematic reviews (e.g., diagnostic tests, preventive measures, drug treatments). The reporting and methodology could substantially differ across different types of systematic reviews. In the future, we expect more

information to come from RCTs, and more systematic reviews of RCTs on COVID-19 will be published to provide more reliable evidence. Third, we only summarized the effect measures used in the systematic reviews and meta-analyses (Table 2), but we did not assess the appropriateness of the effect measures. Researchers may have different opinions about the choices of effect measures, such as the mean difference versus standardized mean difference for continuous outcomes and the odds ratio versus relative risk versus risk difference for binary outcomes.[47-52] The assessment of their appropriateness should be performed on a case-by-case basis.

4.3 | Implications

The mass production of systematic reviews and meta-analyses offers many opportunities to apply the synthesized evidence to clinical practice. However, the reliability of the resulting evidence may be uncertain if the evidence synthesis used inappropriate methods and the study quality was not adequately appraised. It may be questionable if too many systematic reviews and meta-analyses are conducted to address similar or overlapped research topics, leading to research waste.[53] If the synthesized evidence in a meta-analysis is likely affected by low-quality studies, it may be of interest to identify a core set of primary studies and compare their evidence with the synthesized one. Such studies are expected to be well-designed, carefully conducted, and frequently cited in multiple systematic reviews. With more studies on COVID-19 being conducted, researchers should consider periodically updating systematic reviews by including new studies and examining the potential changes of evidence.[54-56]

Due to time sensitivity, the peer review of many published studies on COVID-19 may be insufficient, and valuable comments from reviewers may not have been fully addressed. In addition to delivering timely evidence for COVID-19, it is also crucial to safeguard the integrity of scientific findings rather than downgrading scientific rigour in the academic rush for pandemic publishing.[4,9,57] In terms of systematic reviews and meta-analyses, authors, peer reviewers, and journal editors should try their best to follow the PRISMA statement or other relevant checklists and critically assess the conduct and reporting of systematic reviews.

5 | CONCLUSION

In summary, our findings point to a need for an increase in transparency and quality of performing and reporting systematic reviews and meta-analyses on COVID-19. While the included studies play an important role in synthesizing the presently available data and providing valuable insights into the current state of COVID-19 research, it is also critical to examine their validity and reproducibility. Although the urgent need for COVID-19 research could impair the quality of systematic reviews and meta-analyses in this field, researchers might still consider several methods to remedy these concerns. For example, sensitivity analyses could be performed to examine the impact of different methods for assessing publication bias and implementing meta-analysis models on the synthesized evidence.

ACKNOWLEDGEMENTS

The authors thank two anonymous reviewers for helpful comments that have improved this article. Lifeng Lin was supported in part by the US National Institutes of Health/National Library of Medicine grant R01 LM012982 and National Institutes of Health/National Center for Advancing Translational Sciences grant UL1 TR001427. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

CONFLICT OF INTEREST

The authors declare there is no conflict of interest.

DATA AVAILABILITY STATEMENT

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

ORCID

Chang Xu https://orcid.org/0000-0002-2801-9075
Lifeng Lin https://orcid.org/0000-0002-3562-9816

REFERENCES

1. Zhu N, Zhang D, Wang W, et al. A novel coronavirus from patients with pneumonia in China, 2019. *N Engl J Med*. 2020;382(8):727-733.
2. Horbach SPJM. Pandemic publishing: medical journals strongly speed up their publication process for COVID-19. *Quantit Sci Stud*. 2020;1(3):1056-1067.
3. Fidahic M, Nujic D, Runjic R, et al. Research methodology and characteristics of journal articles with original data, preprint articles and registered clinical trial protocols about COVID-19. *BMC Med Res Methodol*. 2020;20(1):161.
4. Palayew A, Norgaard O, Safreed-Harmon K, Andersen TH, Rasmussen LN, Lazarus JV. Pandemic publishing poses a new COVID-19 challenge. *Nat Hum Behav*. 2020;4(7):666-669.
5. Ioannidis JPA. Coronavirus disease 2019: the harms of exaggerated information and non-evidence-based measures. *Eur J Clin Investig*. 2020;50(4):e13222.
6. Alexander PE, Debono VB, Mammen MJ, et al. COVID-19 coronavirus research has overall low methodological quality thus far: case in point for chloroquine/hydroxychloroquine. *J Clin Epidemiol*. 2020;123:120-126.
7. Oikonomidi T, Boutron I, Pierre O, Cabanac G, Ravaud P. Changes in evidence for studies assessing interventions for COVID-19 reported in preprints: meta-research study. *BMC Med*. 2020;18(1):402.
8. Wolkewitz M, Lambert J, von Cube M, et al. Statistical analysis of clinical COVID-19 data: a concise overview of lessons learned, common errors and how to avoid them. *Clin Epidemiol*. 2020;12:925-928.
9. Bagdasarian N, Cross GB, Fisher D. Rapid publications risk the integrity of science in the era of COVID-19. *BMC Med*. 2020;18(1):192.
10. Bellos I. A metaresearch study revealed susceptibility of Covid-19 treatment research to white hat bias: first, do no harm. *J Clin Epidemiol*. 2021;136:55-63.

11. Chin V, Ioannidis JPA, Tanner MA, Cripps S. Effect estimates of COVID-19 non-pharmaceutical interventions are non-robust and highly model-dependent. *J Clin Epidemiol*. 2021;136:96-132.
12. Borenstein M, Hedges LV, Higgins JPT, Rothstein HR. *Introduction to Meta-Analysis*. Chichester, UK: Wiley; 2009.
13. Murad MH, Nayfeh T, Urtecho Suarez M, et al. A framework for evidence synthesis programs to respond to a pandemic. *Mayo Clin Proc*. 2020;95(7):1426-1429.
14. Murad MH, Asi N, Alsawas M, Alahdab F. New evidence pyramid. *BMJ Evid-Based Med*. 2016;21(4):125-127.
15. Gurevitch J, Koricheva J, Nakagawa S, Stewart G. Meta-analysis and the science of research synthesis. *Nature*. 2018;555:175-182.
16. Haddaway NR, Akl EA, Page MJ, Welch VA, Keenan C, Lotfi T. Open synthesis and the coronavirus pandemic in 2020. *J Clin Epidemiol*. 2020;126:184-191.
17. Petkova E, Antman EM, Troxel AB. Pooling data from individual clinical trials in the COVID-19 era. *JAMA*. 2020;324(6):543-545.
18. Cote MP, Apostolakos JM, Voss A, DiVenere J, Arciero RA, Mazzocca AD. A systematic review of meta-analyses. *J Arthrosc Relat Surg*. 2016;32(3):528-537.
19. Ioannidis JPA. The mass production of redundant, misleading, and conflicted systematic reviews and meta-analyses. *Milbank Q*. 2016;94(3):485-514.
20. Hacke C, Nunan D. Discrepancies in meta-analyses answering the same clinical question were hard to explain: a meta-epidemiological study. *J Clin Epidemiol*. 2020;119:47-56.
21. Kaizar EE. Metaanalyses are observational studies: how lack of randomization impacts analysis. *Am J Gastroenterol*. 2005;100(6):1233-1236.
22. Turner EH, Matthews AM, Linardatos E, Tell RA, Rosenthal R. Selective publication of antidepressant trials and its influence on apparent efficacy. *N Engl J Med*. 2008;358(3):252-260.
23. Cook DJ, Mulrow CD, Haynes RB. Systematic reviews: synthesis of best evidence for clinical decisions. *Ann Intern Med*. 1997;126(5):376-380.
24. Akl EA, Morgan RL, Rooney AA, et al. Developing trustworthy recommendations as part of an urgent response (1–2 weeks): a GRADE concept paper. *J Clin Epidemiol*. 2021;129:1-11.
25. Macdonald H, Loder E, Abbasi K. Living systematic reviews. *BMJ*. 2020;370:m2925.
26. Page MJ, Welch VA, Haddaway NR, Karunananthan S, Maxwell LJ, Tugwell P. "One more time": why replicating some syntheses of evidence relevant to COVID-19 makes sense. *J Clin Epidemiol*. 2020;125:179-182.
27. Shokraneh F, Russell-Rose T. Lessons from COVID-19 to future evidence synthesis efforts: first living search strategy and out of date scientific publishing and indexing industry. *J Clin Epidemiol*. 2020;123:171-173.
28. Stamm TA, Andrews MR, Mosor E, et al. The methodological quality is insufficient in clinical practice guidelines in the context of COVID-19: systematic review. *J Clin Epidemiol*. 2021;135:125-135.
29. Moher D, Liberati A, Tetzlaff J, Altman DG. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *PLoS Med*. 2009;6(7):e1000097.
30. Shea BJ, Reeves BC, Wells G, et al. AMSTAR 2: a critical appraisal tool for systematic reviews that include randomised or non-randomised studies of healthcare interventions, or both. *BMJ*. 2017;358:j4008.
31. Shea BJ, Grimshaw JM, Wells GA, et al. Development of AMSTAR: a measurement tool to assess the methodological quality of systematic reviews. *BMC Med Res Methodol*. 2007;7(1):10.
32. Panic N, Leoncini E, de Belvis G, Ricciardi W, Boccia S. Evaluation of the endorsement of the preferred reporting items for systematic reviews and meta-analysis (PRISMA) statement on the quality of published systematic review and meta-analyses. *PLoS One*. 2013;8(12):e83138.
33. van der Pol CB, McInnes MDF, Petrcich W, Tunis AS, Hanna R. Is quality and completeness of reporting of systematic reviews and meta-analyses published in high impact radiology journals associated with citation rates? *PLoS One*. 2015;10(3):e0119892.
34. Wolf LA. Clinical research: the importance of meta-analysis and systematic reviews in determining appropriate practice changes. *J Emerg Nurs*. 2015;41(4):360-361.
35. Yan P, Yao L, Li H, et al. The methodological quality of robotic surgical meta-analyses needed to be improved: a cross-sectional study. *J Clin Epidemiol*. 2019;109:20-29.
36. Pieper D, Lorenz RC, Rombey T, et al. Authors should clearly report how they derived the overall rating when applying AMSTAR 2—a cross-sectional study. *J Clin Epidemiol*. 2021;129:97-103.
37. Sterne JAC, Sutton AJ, Ioannidis JPA, et al. Recommendations for examining and interpreting funnel plot asymmetry in meta-analyses of randomised controlled trials. *BMJ*. 2011;343:d4002.
38. Hunter JP, Saratzis A, Sutton AJ, Boucher RH, Sayers RD, Bown MJ. In meta-analyses of proportion studies, funnel plots were found to be an inaccurate method of assessing publication bias. *J Clin Epidemiol*. 2014;67(8):897-903.
39. Peters JL, Sutton AJ, Jones DR, Abrams KR, Rushton L. Comparison of two methods to detect publication bias in meta-analysis. *JAMA*. 2006;295(6):676-680.
40. Higgins JPT, Thomas J, Chandler J, et al. *Cochrane Handbook for Systematic Reviews of Interventions*. Chichester, UK: Wiley; 2019.
41. IntHout J, Ioannidis JPA, Rovers MM, Goeman JJ. Plea for routinely presenting prediction intervals in meta-analysis. *BMJ Open*. 2016;6(7):e010247.
42. Riley RD, Higgins JPT, Deeks JJ. Interpretation of random effects meta-analyses. *BMJ*. 2011;342:d549.
43. Schwarzer G, Schumacher M, Rücker G. Sole reliance on I^2 may mislead. *Heart*. 2017;103(18):1471-1472.
44. Lin L, Chu H, Murad MH, et al. Empirical comparison of publication bias tests in meta-analysis. *J Gen Intern Med*. 2018;33(8):1260-1267.
45. Lin L. Hybrid test for publication bias in meta-analysis. *Stat Methods Med Res*. 2020;29(10):2881-2899.
46. Langan D, Higgins JPT, Jackson D, et al. A comparison of heterogeneity variance estimators in simulated random-effects meta-analyses. *Res Synth Methods*. 2019;10(1):83-98.
47. Bond CF Jr, Wiitala WL, Richard FD. Meta-analysis of raw mean differences. *Psychol Methods*. 2003;8(4):406-418.
48. Takeshima N, Sozu T, Tajika A, Ogawa Y, Hayasaka Y, Furukawa TA. Which is more generalizable, powerful and interpretable in meta-analyses, mean difference or standardized mean difference? *BMC Med Res Methodol*. 2014;14(1):30.
49. Poole C, Shrier I, VanderWeele TJ. Is the risk difference really a more heterogeneous measure? *Epidemiology*. 2015;26(5):714-718.
50. Sonis J. Odds ratios vs risk ratios. *JAMA*. 2018;320(19):2041.
51. Bakbergenuly I, Hoaglin DC, Kulinskaya E. Pitfalls of using the risk ratio in meta-analysis. *Res Synth Methods*. 2019;10(3):398-419.
52. Doi SA, Furuya-Kanamori L, Xu C, Lin L, Chivese T, Thalib L. Questionable utility of the relative risk in clinical research: a call for change to practice. *J Clin Epidemiol*. 2020.
53. Glasziou PP, Sanders S, Hoffmann T. Waste in covid-19 research. *BMJ*. 2020;369:m1847.
54. Siemieniuk RA, Bartoszko JJ, Ge L, et al. Drug treatments for covid-19: living systematic review and network meta-analysis. *BMJ*. 2020;370:m2980.
55. Allotey J, Stallings E, Bonet M, et al. Clinical manifestations, risk factors, and maternal and perinatal outcomes of coronavirus disease 2019 in pregnancy: living systematic review and meta-analysis. *BMJ*. 2020;370:m3320.

56. Boutron I, Chaimani A, Meerpohl JJ, et al. The COVID-NMA project: building an evidence ecosystem for the COVID-19 pandemic. *Ann Intern Med*. 2020;173(12):1015-1017.
57. Garritty C, Gartlehner G, Nussbaumer-Streit B, et al. Cochrane rapid reviews methods group offers evidence-informed guidance to conduct rapid reviews. *J Clin Epidemiol*. 2021;130:13-22.
58. Hoy D, Brooks P, Woolf A, et al. Assessing risk of bias in prevalence studies: modification of an existing tool and evidence of interrater agreement. *J Clin Epidemiol*. 2012;65(9):934-939.
59. Ijaz S, Verbeek J, Seidler A, et al. Night-shift work and breast cancer—a systematic review and meta-analysis. *Scand J Work Environ Health*. 2013;39(5):431-447.

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of this article.

How to cite this article: Rosenberger KJ, Xu C, Lin L. Methodological assessment of systematic reviews and meta-analyses on COVID-19: A meta-epidemiological study. *J Eval Clin Pract*. 2021;27:1123–1133. https://doi.org/10.1111/jep.13578