

## Post-publication peer review and the identification of methodological and reporting issues in COVID-19 trials: a qualitative study

Mauricia Davidson <sup>(i)</sup>, <sup>1</sup> Christoffer Bruun Korfitsen <sup>(i)</sup>, <sup>2,3</sup> Carolina Riveros, <sup>1,4,5</sup> Anna Chaimani, <sup>1,4</sup> Isabelle Boutron <sup>1,4,5</sup>

#### 10.1136/bmjebm-2024-113068

#### Abstract

► Additional supplemental material is published online only. To view, please visit the journal online (https:// doi.org/10.1136/bmjebm-2024-113068).

<sup>1</sup>Université Paris Cité and Université Sorbonne Paris Nord, Inserm, INRAE, Centre for Research in Epidemiology and Statistics (CRESS), Paris, Île-de-France, France <sup>2</sup>Open Patient Data Explorative Network (OPEN), Odense University Hospital, Odense, Denmark <sup>3</sup>Cochrane Denmark & Centre for Evidence-Based Medicine Odense (CEBMO), Department of Clinical Research, University of Southern Denmark, Odense, Denmark <sup>4</sup>Cochrane Centre France, Paris, Île-de-France, France <sup>5</sup>Centre d'Epidémiologie Clinique, AP-HP, Hôpital Hôtel Dieu, Paris, Île-de-France, France

Correspondence to: Dr Mauricia Davidson; mauricia.davidson@gmail. com

Check for updates

© Author(s) (or their employer(s)) 2025. Re-use permitted under CC BY-NC. No commercial re-use. See rights and permissions. Published by BMJ Group.

To cite: Davidson M, Korfitsen CB, Riveros C, *et al. BMJ Evidence-Based Medicine* Epub ahead of print: [please include Day Month Year]. doi:10.1136/ bmjebm-2024-113068

# **Objectives** We aimed to determine to what extent systematic reviewers and post-preprint and post-publication peer review identified methodological and reporting issues in COVID-19 trials that could be easily resolved by the authors. **Design** Qualitative study.

Data sources COVID-NMA living systematic review (covid-nma.com), PubPeer, medRxiv, Research Square, SSRN.

Methods We considered randomised controlled trials (RCTs) in COVID-NMA that evaluated pharmacological treatments for COVID-19 and retrieved systematic reviewers' assessments of the risk of bias and outcome reporting bias. We also searched for commentary data on PubPeer and preprint servers up to 6 November 2023. We employed qualitative content analysis to develop themes and domains of methodological and reporting issues identified by commenters.

Results We identified 500 eligible RCTs. Systematic reviewers identified methodological and reporting issues in 446 (89%) RCTs. In 391 (78%) RCTs, the issues could be easily resolved by the trial authors; issues included incomplete reporting (49%), selection of the reported results (52%) and no access to the pre-specified plan (25%). Alternatively, 74 (15%) RCTs had received at least one comment on PubPeer or preprint servers, totalling 348 comments. In 46 (9%) RCTs, the issues identified by post-preprint and postpublication peer review comments could be easily resolved by the trial authors; the issues were related to incomplete reporting (6%), errors (5%), statistical analysis (3%), inconsistent reporting of methods and analyses (2%), spin (2%), selection of the reported results (1%) and no access to the raw data/pre-specified plan (1%).

**Conclusions** Without changing their process, systematic reviewers identified issues in most RCTs that could be easily resolved by the trial authors; however, the lack of an established author feedback mechanism represents a wasted opportunity for facilitating improvement and enhancing the overall manuscript quality. On the other hand, despite the existing feedback loop to authors present in post-publication peer review, it demonstrated limited effectiveness in identifying methodological and reporting issues.

## WHAT IS ALREADY KNOWN ON THIS TOPIC

⇒ Despite its central role in ensuring rigorous research dissemination, a typical peer review process has limitations; however systematic reviewer assessments and postpublication peer review can identify key issues in trials, even facilitating potential editorial action.

#### WHAT THIS STUDY ADDS

- ⇒ Through risk of bias and outcome reporting bias assessments, systematic reviewers identified methodological and reporting issues in the majority of trials that could be easily resolved by trial authors.
- ⇒ Post-publication peer review is underutilised and poorly identified key issues in research quality.

#### HOW THIS STUDY MIGHT AFFECT RESEARCH, PRACTICE OR POLICY

- ⇒ Direct engagement between systematic reviewers and trial authors is a missed opportunity that should be addressed to supplement formal peer review.
- ⇒ Encouraging a culture within the research community that values postpublication peer review is essential for maximising its effectiveness.

#### Background

Peer review is regarded as the cornerstone of rigorous research. The usual peer-review process begins when a manuscript is submitted to an academic journal for publication.<sup>1</sup> A journal editor then assigns independent researchers to assess the quality of the manuscript. In turn, the independent researchers produce a report that aids the editor in deciding whether to publish or reject the submission, or request further revisions prior to acceptance or rejection.<sup>2 3</sup> While individual journal policies vary, acknowledging that the peer-review process has a few limitations is important. The process is generally slow and is often compounded by difficulties in identifying reviewers, who may

Protected by copyright, including for uses related to text and data mining, AI training, and similar technologies.

Protected by copyright, including for uses related to text and data mining, AI training, and similar technologies.

not thoroughly address issues, such as incomplete or biased reporting.  $^{\rm 4-7}$ 

Recognising the need for new methodologies in research evaluation in contrast to the formal journal-managed pre-publication peer review process, alternative approaches have been implemented or proposed.<sup>8–11</sup> Systematic reviews, particularly living systematic reviews, could provide a valuable avenue for detecting important methodological and reporting issues, such as incomplete or selective reporting of results; however, the time that lapsed between the trial publication and the review is a critical factor that warrants consideration.<sup>12</sup> Establishing a feedback loop between authors and systematic reviewers could facilitate timely alerts to authors, provide an opportunity to correct these issues, and ultimately, enhance the quality of research dissemination.

Furthermore, in the dynamic landscape of scientific communication, post-publication peer review (PPPR) platforms, such as PubPeer, have been developed. PPPR allows a wider audience to provide feedback on published work with ongoing assessments and improvements to study findings.<sup>13</sup> <sup>14</sup> Researchers using these platforms can raise community awareness of flaws in published research, prompt critical discussions, and, in some cases, cause major editorial actions, like retractions and expressions of concern.<sup>15</sup> <sup>16</sup>

The COVID-19 pandemic reshaped scientific communication and triggered an exponential increase in the number of published articles, driven by the urgency to communicate research findings.<sup>17</sup> This surge in articles shortened the peer review process and resulted in the widespread use of preprints for rapid dissemination.<sup>18</sup> <sup>19</sup> PubPeer and similar platforms were actively used during this period, and major preprint servers, such as medRxiv, facilitated open commentary on study methods and results, which made it possible to improve the manuscripts prior to their formal peer review and publication in an academic journal. Large-scale living systematic reviews, such as the COVID-NMA living systematic review, were implemented and enabled systematic reviewers to highlight and identify specific issues.<sup>20</sup>

Therefore, using a sample of trials included in the COVID-NMA living systematic review, we aimed to determine (1) to what extent systematic reviewers identified methodological and reporting issues in COVID-19 trials that could be easily resolved by authors, and (2) to what extent post-preprint and post-publication peerreview identified methodological and reporting issues in COVID-19 trials and to describe whether these issues could be easily resolved by authors.

#### Methods

We conducted a qualitative study of COVID-19 preprints and peerreviewed journal articles in the COVID-NMA living systematic review. Our protocol is available on the Open Science Framework (https://osf.io/j32df/).

#### Data source and search

We used data from the COVID-NMA living systematic review ( www.covid-nma.com), hereafter, referred to as COVID-NMA.<sup>20</sup> COVID-NMA was a living systematic review of interventions for the prevention and treatment of COVID-19. It was built from a comprehensive search of two validated secondary sources to identify eligible randomised controlled trials (RCTs): the Epistemonikos L-OVE COVID-19 platform<sup>21</sup> and the Cochrane COVID-19 Study Register.<sup>22</sup> The Retraction Watch database<sup>23</sup> was also searched to identify and remove retracted trials from the review. Screening and data extraction were performed by pairs of researchers, independently and in duplicate, with disagreements resolved through consensus and a third reviewer, when necessary. Data were extracted from preprints, all preprint updates, peer-reviewed journal articles, and all available documentation (eg, online supplemental material).<sup>24</sup> See online supplemental methods S1 and for more details on the study's methodology, search strategy and the scope of the COVID-NMA. The COVID-NMA living mapping and synthesis concluded in August 2023.

#### Study selection

We included all RCTs that evaluated pharmacological treatments for patients with COVID-19 and were available as preprints or journal articles. The last search date for any treatment RCT was 14 December 2022. Dates for individual treatment comparisons are detailed in online supplemental additional file 1.

We excluded RCTs that evaluated non-pharmacological treatments, preventive interventions (eg, personal protective equipment and movement control strategies), vaccines, and supportive treatments for patients admitted to intensive care units. We also excluded cluster RCTs and RCT results only reported in their trial registry or in a conference abstract.

#### Identification of issues by systematic reviewers

As part of the COVID-NMA protocol, two systematic reviewers, independently and with consensus, assessed each RCT included in the review for risk of bias (RoB) using the Cochrane RoB 2 tool<sup>25</sup> and outcome reporting bias (ORB)<sup>26 27</sup> in 14 pre-specified outcomes (such as clinical improvement, incidence of viral negative conversion, WHO clinical progression score of level 7 or above, all-cause mortality, hospitalisation or death (in an outpatient setting), incidence of any adverse events and incidence of serious adverse events). Systematic reviewers provided detailed justification for each RoB assessment. If an RCT did not report such outcomes, RoB could not be assessed. Details of the review outcomes, as well as RoB and ORB assessment rules are provided in online supplemental additional file 1. One researcher (MD) retrieved all the RoB justifications reported by COVID-NMA systematic reviewers for all domains and rated as 'some concerns' or 'high' RoB for the pre-specified outcomes; they also identified methodological and reporting issues that could be easily resolved by the trial authors. Additionally, MD retrieved ORB assessments for all the pre-specified outcomes.

'Easily resolvable issues' refers to methodological and reporting deficiencies in clinical trials that can be addressed by the trial authors during the peer review stage of a manuscript. They are considered 'easily resolvable' because they do not require additional data collection or major changes to the study design but rather improvements in the clarity, transparency or completeness of the information presented.

The issues that were identified through the living systematic review and that could be easily resolved by the trial authors included:

► Incomplete reporting—considered when there was no or little information on the allocation sequence generation; allocation concealment; blinding status of participants, care providers and outcome assessors; participant crossover and/or administration of co-interventions of interest (antivirals, corticosteroids, biologics) per arm during the trial (assessed only in unblinded studies); number of participants randomised per arm; number of participants analysed per arm for the review pre-specified outcomes; and the reasons for, or proportions of, missing data per arm. Information on this issue of incomplete reporting was retrieved from RoB assessments.

- Selection of the reported results-considered in cases of missing or added evidence.
  - Missing evidence that is, the outcomes were planned in the clinical trial protocol, statistical analysis plan, or trial registry; however, the results were not available for inclusion in the synthesis, (probably) because the p value, magnitude or direction of the results was considered unfavourable by the study investigators. Information on this issue was retrieved from the ORB assessments.
  - Added evidence that is, the study results were available for inclusion in the synthesis but not planned to be analysed in the clinical trial protocol, statistical analysis plan, or trial registry. Information on this issue was retrieved from RoB and ORB assessments.
- No access to the pre-specified plan-considered when there was no pre-specified clinical trial protocol, statistical analysis plan, or trial registry available for assessment, regardless of whether study results were available for inclusion in the synthesis. Information on this issue was retrieved from RoB and ORB assessments.

MD also retrieved the general trial data reported by COVID-NMA systematic reviewers: first author, publication source (preprint or journal name), publication date and full-text links.

### Identification of issues by post-preprint and post-publication peer review

One researcher (MD) systematically searched PubPeer using the digital object identifiers (DOIs) of eligible RCTs to aggregate all available comments. Commentary data published from 2020 onwards were retrieved from medRxiv using the Disqus application programming interface (API) (disqus.com/api/docs/) and R code<sup>19 28</sup>; these were then cross-referenced with the DOIs of the eligible RCTs. A manual commentary data search was conducted on the Research Square and Social Science Research Network (SSRN) preprint platforms using trial DOIs. Reports that received at least one comment were included. For preprints, only the first version was considered. The last search date for the commentary data was 6 November 2023.

We collected post-preprint and PPPR commentary data using qualitative content analysis to inductively develop themes and domains. Two researchers (MD, CBK) used 20 PubPeer comments to identify themes/domains of the issues addressed by the commenters. The two researchers (MD, CBK) then met to reach consensus on the domains to be included in a data extraction form, along with a senior researcher (IB). The researchers used this initial set of domains to extract data, independently and in duplicate, in groups of 20 comments with consensus in the case of disagreements. Two researchers (MD, CR) extracted the commentary data from the preprint servers in the same manner. Finally, one researcher (MD) identified subdomains for the 'study design' domain. During the data extraction process, newly identified domains were documented and discussed with IB for continuous fine-tuning. All researchers had a minimum of 3 years of training in clinical epidemiology, particularly trial methodology. Of note, we did not independently confirm the validity of the issues raised in the comments. Information was collected on all the comments, such as the comment source (PubPeer, preprint server [medRxiv, Research Square, SSRN]) and the publication date of the comment. Notably, we could not find the exact date of PubPeer comment posts, only the month and year. Therefore, we assigned the first day of the given month (eg, a comment dated May 2022 was extracted as 1 May 2022) during data collection. Information on whether any changes had been made to the original report (ie,



Figure 1 Flowchart of included RCTs. ORB, outcome reporting bias; PPPR, post-publication peer review; RCTs, randomised controlled trials; RoB, risk of bias.

erratum or retraction) was also retrieved. When available, data on the commenters' name, affiliation, specific requests (ie, erratum or retraction), actions (ie, conducted a specific check or reanalysis, commented the erratum/retraction notice, or published a commentary) and whether the trial author addressed the comment were collected. Finally, whether the issues identified could be easily resolved by the trial authors were assessed.

#### **Data synthesis**

Frequencies and percentages were calculated for the categorical variables, while medians with IQRs were calculated for the continuous variables. The extracted qualitative data were coded using thematic analysis and grouped to develop domains. We used R software<sup>29</sup> with the *tidyverse*<sup>30</sup> package for all analyses.

#### **Results**

#### Characteristics of the eligible RCTs

Of the 569 records of treatment RCTs identified in the database search, 494 records reporting 500 RCTs met the eligibility criteria (figure 1). Four platform or multi-cohort trials consisting of two to four individual trials were reported in four manuscripts. Overall, the median sample size of the RCTs was 123 (IQR, 62–353) participants; 65% were prospectively registered, and 47% received industry or mixed funding (table 1).

#### Systematic reviewer assessments

Of the 500 RCTs, systematic reviewers identified methodological and reporting issues in 446 (89%) RCT reports; in 391 (78%) RCT reports, issues could be easily resolved by the trial authors (figure 2). In 247 (49%) RCT reports, these issues were attributed to incomplete reporting, that is, they included no or not enough information on allocation sequence generation (2%), allocation concealment (24%), blinding details (6%), participant crossover and/or balance in the administration of co-interventions of interest per arm (30%), number of trial participants randomised or analysed per arm (1%), and the reasons for and/or proportions of missing data per arm, if any (8%). Systematic reviewers also identified issues in the selection of reported results in 261 (52%) RCT reports due to missing evidence (9%) or added evidence (48%). In

Table 1 Characteristics of eligible RCTs		
Characteristics	Total RCTs n=500 (%)	
Sample size, median (IQR)	123 (62–353)	
Publication type		
Preprint	70 (14)	
Preprint and subsequent journal article	150 (30)	
Journal article	280 (56)	
Registration timing		
Prospective	327 (65)	
Retrospective	141 (28)	
Not reported/unclear	32 (6)	
Funding type		
Industry/mixed	236 (47)	
Other	264 (53)	
Preprint post/article publication*		
<6 months	69 (14)	
6–12 months	101 (20)	
>12 months	330 (66)	
Major editorial action		
Erratum	8 (2)	
· · · · · · · · · · · · · · · · · · ·		

\*Relative to March 2020 (ie, start of the pandemic). Percentages may not add up due to rounding or shared characteristics.

Other: public/no funding/not reported/unclear. Mixed, industry and public funding; RCT, randomised controlled trial.

97 (25%) RCT reports, systematic reviewers identified that there was no access to the pre-specified plan (ie, protocol, statistical analysis plan, and/or registry). Notably, systematic reviewers rated 33 (7%) RCTs as 'low' RoB and no evidence of ORB; therefore, we considered that no issues were identified in those RCTs. Complete RoB assessments were not conducted for 21 (4%) RCTs due to lack of review pre-specified outcomes reported in these RCTs.

#### Post-preprint and PPPR

Among the 500 RCTs, 74 (15%) received at least one comment on either preprint servers or PubPeer for 348 retrieved comments in total (table 2, online supplemental figure S1). Three RCTs had both post-preprint and PPPR comments, that is, comments on the preprint and the subsequent published journal article. One report presented findings from four RCTs. The median number of comments per RCT report was 2 (IQR, 1-5; max, 51), the median word count was 64 (IQR, 28-135; max, 3569), the median delay between preprint post and comment post was 10 (IQR, 2-65) days, and the median delay between journal article publication and comment post was 29 (IQR, 0-106) days. Of the 74 RCTs (71 RCT reports) with at least one comment, 26 (35%) had commentary data posted to PubPeer, and 53 (72%) had commentary data posted on preprint servers, mainly medRxiv (40 RCTs, 54%). 23 comments from 18 (25%) RCT reports were structured as a traditional peer review report. Trial authors responded directly to 12 original comments on seven (9%) RCTs, and most satisfied the issues raised in the original comment.

#### Feasibility of issue resolution

We coded the following methodological and reporting issues identified by the commenters: incomplete reporting, selection of the reported result, result applicability, statistical analysis, error, sample size, spin, study design, conflicts of interest, ethics, fraud, inconsistent reporting of methods and analysesand no access to the raw data/pre-specified plan. Next, we determined whether these issues could be easily resolved by trial authors using the classification detailed in table 3.

Of the 500 RCTs, 46 (9%) with post-preprint and PPPR comments identified methodological and reporting issues that could be easily resolved by the trial authors (figure 3). These issues involved incomplete reporting (29 RCTs, 56 comments), errors (23 RCTs, 28 comments), statistical analysis (14 RCTs, 24



Figure 2 RCTs with resolvable issues identified by systematic reviewers (78%). RCTs, randomised controlled trials.

Table 2 Characteristics of post-preprint and PPPR commen	ts
Characteristics	N=74* (%)
Number of comments (comments)	348
Number of comments per RCT report† (comments)	2 (1-5)
Word count† (words)	64 (28–135)
Delay† between preprint post and comment post (days)	10 (2–65)
Delay <sup>†</sup> between journal article publication and comment post (days)	29 (0–106)
Publication type‡ (n)	
Preprint	55 (74)
Journal article	19 (26)
Comment source (n)	
PubPeer	26§ (35)
medRxiv	40§ (55)
Research Square	10 (14)
SSRN	3§ (4)
Methodological or reporting issues identified (n)	
Study design	34 (46)
Incomplete reporting	29 (39)
Error	23 (31)
Sample size	17 (23)
Statistical analysis	14 (19)
Results applicability	13 (18)
Inconsistent reporting of methods and analyses	12 (16)
Spin	10 (14)
Ethics	9 (12)
No access to raw data/pre-specified plan	5 (7)
Selection of the reported result	3 (4)
Conflict of interest	3 (4)
Fraud	2 (3)
Other	18 (24)
Commenter actions (n)	
Conducted a specific check	23 (31)
Conducted a reanalysis	3 (4)
Commented the erratum/retraction note	2 (3)
Published a commentary	1 (1)
Commenter requests (n)	
Response from author	12 (16)
Erratum	3 (4)
Retraction	0
Response received (n)	
From author	7 (9)
Satisfied needs of the original comment	5 (7)
*3 RCTs had both post-preprint and PPPR comments.	
tAt time of comment retrieval.	
§8 RCTs had comments on two platforms.	
N, number of RCTs; PPPR, post-publication peer review; RCT, randomised trial; SSRN, Social Science Research Network.	controlled

comments), inconsistent reporting of methods and analyses (12 RCTs, 15 comments), spin (10 RCTs, 13 comments), no access to the raw data/pre-specified plan (5 RCTs, 5 comments), and selection of the reported results (3 RCTs, 7 comments). Eight (2%) RCT reports had an erratum to the final publication. At least one of the reasons provided by the editors for the errata of 3 RCT reports was addressed in post-preprint and PPPR comments.

#### Discussion

Our study describes the methodological and reporting issues in COVID-19 trials identified by systematic reviewers and in postpreprint and PPPR. We analysed 500 RCTs and found that the issues identified in systematic reviewer assessments in 391 (78%) RCTs could be easily resolved by the trial authors. Alternatively, post-preprint and PPPR comments identified issues in 46 (9%) RCTs that could be easily resolved by the trial authors.

Earlier studies have analysed post-preprint and PPPR. Carneiro et al studied 1921 comments on 1037 preprints and observed that critical comments addressed interpretation, methodological design, analysis, reporting, data sharing and ethics.<sup>28</sup> They concluded that comments posted on preprint servers evaluate content comparable to that examined in formal peer review. Ortega et al analysed a sample of 39985 PubPeer comments in 24779 publications in 2019 and 2020 and found that 72% reported an element of fraud, with these comments sparking the most discussion and having a longer delay in posting.<sup>31</sup> They also found issues related to a lack of information (2%), honest errors (2%) and methodological flaws (8%). Additionally, in a cross-sectional study of 1983 preprints that received single comments on the bioRxiv platform before September 2019, Malički et al noted that over two-thirds of the comments did not originate from the preprint authors, with some comments being categorised as 'issue detected' (10%) and 'asking for raw data or code' (3%).32 Notably, they found that 11% of author comments explicitly encouraged others to provide feedback, with one comment expressing a preference for revising the preprint rather than making changes to the journal article.<sup>32</sup> To our knowledge, no other study has identified methodological and reporting issues that could be easily resolved by trial authors nor related these issues to those identified in systematic reviewer assessments.

In general, the overall reporting quality of COVID-19 trials has been found to have significant shortcomings. Kapp *et al* found that both preprints and peer-reviewed publications of COVID-19 RCTs lacked transparency and completeness in reporting, with issues such as inconsistent outcome reporting and inadequate descriptions of harms persisting even after peer review.<sup>4</sup> Other studies similarly noted poor adherence to the CONSORT guidelines, with average reporting rates around 54%, emphasising

Table 3 Feasibility of issue	resolution		
Issues that could be easily resolved by authors	Resolution	Issues that could not be easily resolved by the authors	Resolution
Incomplete reporting	Clearer or further explanations of the study's methods or better reporting of the study's findings	Study design	Not feasible during the peer review stage but should be presented as a limitation in the discussion section of the manuscript, with a sufficient explanation.
Selection of the reported results	Reporting all outcomes pre-specified in the clinical trial protocol, statistical analysis plan, and/or trial register $% \left( {{{\bf{r}}_{\rm{s}}}} \right)$	Sample size	
No access to the raw data/pre- specified plan	Providing access to the raw data, protocol, and/or statistical analysis plan	Results applicability	
Inconsistent reporting of methods and analyses	Consistent descriptions of methods and analyses between the pre-specified plan and all versions of the reports	Conflict of interest	
Statistical analysis	Rerunning the appropriate analyses	Ethics	
Error	Correcting typographical errors or miscalculations	Fraud	
Spin	Accurate and neutral interpretation of the study's results in the abstract and conclusions		



Figure 3 RCTs with issues identified by post-preprint and post-publication peer review. PICO, Population, Intervention, Comparator, Outcome; RCTs, randomised controlled trials.

deficiencies in critical areas like allocation concealment, blinding and sample size estimation.<sup>33 34</sup> Quinn *et al* compared COVID-19 papers with contemporaneous non-COVID-19 papers and found that COVID-19 research had a higher risk of bias and poorer adherence to reporting guidelines.<sup>35</sup>

#### Implications for research

Our findings have several important implications. Incorporating feedback from alternative and informal peer review sources, when duly acknowledged by the authors, can serve as a valuable supplement to formal peer review processes and enhance a manuscript's overall quality. First, by following the usual iterative process of living systematic reviews, which involves continuous evidence synthesis with a detailed assessment of RoB and ORB for each new RCT, systematic reviewers can identify key issues that could be communicated back to the authors to be resolved. In our sample, these issues were identified in 78% of the RCTs. Therefore, the absence of a direct link between reviewers and authors is a missed opportunity, because systematic reviewers should play a role in peer review.

Discussions on this disconnect between evidence generation and synthesis have been raised.<sup>36</sup> They highlight that interactions between trialists and systematic reviewers are often limited to data requests for unreported outcomes or methodological details, with little focus on providing feedback to improve ongoing or future trials. Similarly, trialists rarely use systematic reviews to guide decisions on comparators, sample size, or outcomes, which could enhance inclusion in meta-analyses. After completing their trials, they also often fail to share results with reviewers, hindering updates to existing reviews. A reinforced link between trialists and systematic reviewers should be a major objective in implementing this cycle of improvement.

Second, proponents of PPPR stress that it plays a role in identifying methodological and reporting issues and in improving scholarly publishing. However, given that our study showed that post-preprint and PPPR comments identified issues in only 9% of RCTs, further development of these platforms is warranted to maximise their effectiveness. Incentivising and fostering a culture within the research community that values PPPR is essential. For example, journals could consider employing a grace period after publication wherein important comments prompt additional revisions by the authors. Another approach is to integrate post-preprint and PPPR into the research workflow. Journals could require authors to post their preprint when submitting their manuscript and, within the peer review timeline, address all post-preprint peer review comments. In this study, the median time from preprint posting to comment was 10 days (IQR, 2–65), aligning well with typical peer review delays. Before acceptance, editors could ensure that authors have adequately addressed both formal and post-preprint peer review feedback. This integration could improve research quality and encourage broader scientific engagement.

Furthermore, post-preprint and PPPR platforms could develop systems for better tracking and categorising commentary to enhance usability for all stakeholders, and partner with journals to ensure significant critiques are addressed in final publications. Readers should be critical consumers of preprints and published research, paying attention to issues raised during PPPR, and, of course, participate in constructive commentary to help improve the scientific record.

PPPR, which actively identify irregularities in published data or expose potential research fraud, are often seen as lacking accountability and are labelled as engaging in vigilantism when performed anonymously without formal discourse.<sup>37</sup> A centralised mechanism for coordination and oversight is, therefore, necessary to avoid discriminative and unethical behaviour.

#### Strengths and limitations

RoB and ORB assessment data were retrieved from a large living systematic review (COVID-NMA), which implemented a robust assessment strategy, whereby assessments were performed independently and in duplicate by pairs of researchers, and disagreements were resolved by consensus. The researchers participated in a comprehensive training programme with a team of experts, and quality control of the data was performed regularly by an external group. Furthermore, both post-preprint and PubPeer comments were considered for a diverse exploration of the landscape, and rigorous methodological coding procedures were incorporated to enrich the data via thematic analysis.

However, some limitations of our study must be acknowledged. First, we focused solely on COVID-19 trials, so our results may not be generalisable to post-preprint and PPPR comments outside the context of the pandemic. One study found that COVID-19 preprints had higher levels of engagement and received more comments than non-COVID-19 preprints.<sup>19</sup> Second, our study was constrained by decisions related to living reviews; systematic reviewer assessments were only available for reviewdefined outcomes. However, these outcomes were chosen because of their clinical relevance and included both safety and efficacy endpoints. Finally, most post-preprint and PPPR comments were anonymous; therefore, we could not assess the commenters' expertise in research methodology or investigate their potential conflicts of interest. However, our aim was not to exhaustively verify the validity of the issues highlighted in the comments. Furthermore, anonymity has been linked to increased participation in PPPR, with Lapinski finding that PubPeer, a platform that allows anonymous contributions, received over 37 000 comments on 3300 publications from 2012 to 2015.<sup>38</sup> This exceeded PubMed Commons' 4000 mandatory onymous contributions on the same publications during the same period.

#### **Future work**

Future studies could investigate methods for integrating systematic reviewers' assessments into a structured feedback loop for authors, and even editors. Research could explore how to design standardised feedback templates that streamline communication between reviewers and authors, perhaps via testing digital platforms that incorporate these assessments into the revision process. Additionally, studies could examine incentives to encourage systematic reviewers, authors and editors to this process, ensuring that it is effective and sustainable. Surveys and interviews with authors, reviewers and editors could also address how to incentivise researchers to participate in PPPR and how to seamlessly integrate these processes into existing research workflows.

#### Conclusions

The majority of COVID-19 RCTs had easily resolvable issues identified through RoB and ORB assessments. Systematic reviewers are well placed to improve the quality of manuscripts; however, it is a wasted opportunity, considering that a feedback loop with the trial authors has not been established and acted on. Alternatively, the impact of post-preprint and PPPR in identifying methodological and reporting issues remains limited. Expanding its reach and leveraging the existing feedback loop to authors is imperative to optimise its effectiveness.

#### X Christoffer Bruun Korfitsen @BruunKorfitsen

Acknowledgements The authors would like to thank Elise Diard (Centre d'Epidémiologie Clinique, CRESS, INSERM U1153, Hôtel-Dieu [AP-HP], Cochrane France) for the project's data visualisation work as well as her work on the COVID-NMA website and extraction tool development. We would also like to thank all the members of the COVID-NMA consortium.

**Contributors** MD, CBK, AC and IB conceived and designed the study. MD, CBK and CR were involved in the acquisition of the data. MD conducted the analyses. All the authors were involved in data interpretation. MD drafted the manuscript. All the authors critically reviewed the manuscript. All the authors read and approved the final version of the manuscript.MD is the guarantor, had the final responsibility for the decision to submit for publication and accepts full responsibility for the work and the conduct of the study. MD attests that all listed authors meet authorship criteria and that no others meeting the criteria have been omitted.

Funding The authors did not receive specific funding for the study. MD received a PhD fellowship from the Université Paris Cité. CBK received a PhD grant from The Independent Research Fund Denmark (grant no. 1030-00317B). Data were generated in the context of the COVID-NMA initiative, which received funding from Université Paris Cité, Assistance Publique Hôpitaux de Paris (APHP), Inserm, Cochrane, France (Ministry of Health), French Ministry of Higher Education and Research, Agence Nationale de la Recherche (ANR), and the WHO.

Competing interests None declared.

Patient and public involvement Patients and/or the public were not involved in the design, or conduct, or reporting, or dissemination plans of this research.

Patient consent for publication Not applicable.

Ethics approval Not applicable.

**Provenance and peer review** Not commissioned; externally peer reviewed.

Data availability statement Data are available in a public, open access repository. The datasets generated and/or analysed during the current study are available on the Open Science Framework (https://osf.io/j32df/).

#### **Original research**

Supplemental material This content has been supplied by the author(s). It has not been vetted by BMJ Publishing Group Limited (BMJ) and may not have been peer-reviewed. Any opinions or recommendations discussed are solely those of the author(s) and are not endorsed by BMJ. BMJ disclaims all liability and responsibility arising from any reliance placed on the content. Where the content includes any translated material, BMJ does not warrant the accuracy and reliability of the translations (including but not limited to local regulations, clinical guidelines, terminology, drug names and drug dosages), and is not responsible for any error and/or omissions arising from translation and adaptation or otherwise.

**Open access** This is an open access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited, appropriate credit is given, any changes made indicated, and the use is non-commercial. See: http://creativecommons.org/licenses/by-nc/4.0/.

#### ORCID iDs

Mauricia Davidson http://orcid.org/0000-0002-9991-8035 Christoffer Bruun Korfitsen http://orcid.org/0000-0003-2159-7662

#### References

- 1 Spier R. The history of the peer-review process. *Trends Biotechnol* 2002;20:357–8.
- 2 Glonti K, Cauchi D, Cobo E, *et al.* A scoping review on the roles and tasks of peer reviewers in the manuscript review process in biomedical journals. *BMC Med* 2019;17:118.
- 3 Bornmann L. Scientific peer review. *Annual Review Info Sci & Tec* 2011;45:197–245.
- 4 Kapp P, Esmail L, Ghosn L, et al. Transparency and reporting characteristics of COVID-19 randomized controlled trials. BMC Med 2022;20:363.
- 5 Lazarus C, Haneef R, Ravaud P, et al. Peer reviewers identified spin in manuscripts of nonrandomized studies assessing therapeutic interventions, but their impact on spin in abstract conclusions was limited. J Clin Epidemiol 2016;77:44–51.
- 6 Hopewell S, Collins GS, Boutron I, et al. Impact of peer review on reports of randomised trials published in open peer review journals: retrospective before and after study. BMJ 2014;349:g4145.
- 7 Candal-Pedreira C, Rey-Brandariz J, Varela-Lema L, *et al.* Challenges in peer review: how to guarantee the quality and transparency of the editorial process in scientific journals. *Anales de Pediatría (English Edition)* 2023;99:54–9.
- 8 Tennant JP, Ross-Hellauer T. The limitations to our understanding of peer review. *Res Integr Peer Rev* 2020;5:6.
- 9 Hardwicke TE, Thibault RT, Kosie JE, *et al.* Post-publication critique at top-ranked journals across scientific disciplines: a cross-sectional assessment of policies and practice. *R Soc Open Sci* 2022;9:220139.
- Hunter J. Post-publication peer review: opening up scientific conversation. Front Comput Neurosci 2012;6:63.
- Bordignon F. Self-correction of science: a comparative study of negative citations and post-publication peer review. *Scientometrics* 2020;124:1225–39.
- 12 Hardwicke TE, Serghiou S, Janiaud P, et al. Calibrating the Scientific Ecosystem Through Meta-Research. Annu Rev Stat Appl 2020;7:11–37.
- 13 PubPeer. Search publications and join the conversation. Available: https:// blog.pubpeer.com/ [Accessed 21 Jan 2024].
- 14 O'Sullivan L, Ma L, Doran P. An Overview of Post-Publication Peer Review. *SAR* 2021;3:6.

- 15 Gao W, Ge S, Sun J. RETRACTED: Ailanthone exerts anticancer effect by up-regulating miR-148a expression in MDA-MB-231 breast cancer cells and inhibiting proliferation, migration and invasion. *Biomed Pharmacother* 2019;109:1062–9.
- 16 Marcus AA. Journals acknowledge that a critical 'reader' has a name: Elisabeth BiK [Retraction Watch]. Available: https://retractionwatch.com/ 2022/04/08/journals-acknowledge-that-a-critical-reader-has-a-nameelisabeth-bik/ [Accessed 19 May 2023].
- 17 Dimensions. Digital science. 2018. Available: https://app.dimensions.ai [Accessed 19 Nov 2024].
- 18 Horbach SPJM. No time for that now! Qualitative changes in manuscript peer review during the Covid-19 pandemic. *Res Eval* 2021;30:231–9.
- 19 Fraser N, Brierley L, Dey G, et al. The evolving role of preprints in the dissemination of COVID-19 research and their impact on the science communication landscape. *PLoS Biol* 2021;19:e3000959.
- 20 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;173:1015–7.
- 21 Living overview of the evidence (LOVE). Available: https://app.iloveevidence. com/loves/5e6fdb9669c00e4ac072701d [Accessed 19 May 2021].
- 22 "Cochrane covid-19 study register," cochrane covid-19 study register. Available: https://covid-19.cochrane.org/?sf=publishedDate&sd=desc [Accessed 19 May 2021].
- 23 Retracted coronavirus (covid-19) papers [Retraction Watch]. Available: https://retractionwatch.com/retracted-coronavirus-covid-19-papers/ [Accessed 01 Jun 2023].
- 24 Cabanac G, Oikonomidi T, Boutron I. Day-to-day discovery of preprintpublication links. *Scientometrics* 2021;126:5285–304.
- 25 Sterne JAC, Savović J, Page MJ, et al. RoB 2: a revised tool for assessing risk of bias in randomised trials. BMJ 2019;366:14898.
- 26 Cochrane handbook for systematic reviews of interventions. Available: https://training.cochrane.org/handbook [Accessed 25 May 2023].
- 27 Kirkham JJ, Altman DG, Chan A-W, et al. Outcome reporting bias in trials: a methodological approach for assessment and adjustment in systematic reviews. BMJ 2018;362:k3802.
- 28 Carneiro CFD, da Costa GG, Neves K, et al. Characterization of Comments About bioRxiv and medRxiv Preprints. JAMA Netw Open 2023;6:e2331410.
- 29 R Core Team. R: a language and environment for statistical computing. r foundation for statistical computing. Vienna, Austria, 2022. Available: https://www.r-project.org/
- 30 Wickham H, Averick M, Bryan J, et al. Welcome to the Tidyverse. JOSS 2019;4:1686.
- 31 Ortega JL. Classification and analysis of PubPeer comments: How a web journal club is used. Asso for Info Science & Tech 2022;73:655–70.
- 32 Malički M, Costello J, Alperin JP, *et al.* Analysis of single comments left for bioRxiv preprints till September 2019. *Biochem med (Online)* 2021;31:177–84.
- 33 Yin Y, Shi F, Zhang Y, et al. Evaluation of reporting quality of randomized controlled trials in patients with COVID-19 using the CONSORT statement. *PLoS One* 2021;16:e0257093.
- 34 Grüßer L, Eißing C, Kowark A, et al. Poor reporting quality of randomized controlled trials comparing treatments of COVID-19-A retrospective cross-sectional study on the first year of publications. *PLoS One* 2023;18:e0292860.
- 35 Quinn TJ, Burton JK, Carter B, *et al.* Following the science? Comparison of methodological and reporting quality of covid-19 and other research from the first wave of the pandemic. *BMC Med* 2021;19:46.
- 36 Boutron I, Créquit P, Williams H, et al. Future of evidence ecosystem series: 1. Introduction Evidence synthesis ecosystem needs dramatic change. J Clin Epidemiol 2020;123:135–42.
- 37 Yeo-Teh NSL, Tang BL. Post-publication Peer Review with an Intention to Uncover Data/Result Irregularities and Potential Research Misconduct in Scientific Research: Vigilantism or Volunteerism? Sci Eng Ethics 2023;29:24.
- 38 Lapinski PS. Information access and scholarly communication in postpublication peer review online social networks. 2021. Available: https:// dash.harvard.edu/handle/1/37367697 [Accessed 23 Jan 2024].