



Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

Elsevier hereby grants permission to make all its COVID-19-related research that is available on the COVID-19 resource centre - including this research content - immediately available in PubMed Central and other publicly funded repositories, such as the WHO COVID database with rights for unrestricted research re-use and analyses in any form or by any means with acknowledgement of the original source. These permissions are granted for free by Elsevier for as long as the COVID-19 resource centre remains active.



ELSEVIER

Contents lists available at ScienceDirect

International Journal of Infectious Diseases

journal homepage: www.elsevier.com/locate/ijid

Randomized controlled trials in patients with COVID-19: a systematic review and critical appraisal

Kavina Kudhail^a, Jacqueline Thompson^b, Vivek Mathews^c, Breanna Morrison^b,
Karla Hemming^{b,*}

^a College of Medical and Dental Sciences, University of Birmingham, Birmingham, United Kingdom

^b Institute of Applied Health Research, University of Birmingham, Birmingham, United Kingdom

^c College of Medical and Dental Sciences, University of Birmingham, Birmingham, United Kingdom

ARTICLE INFO

Article history:

Received 12 April 2022

Revised 13 May 2022

Accepted 13 May 2022

Keywords:

COVID-19 treatments

LMICs

Randomized trials

Risk of bias

ABSTRACT

Objectives: This study aimed to describe the prevalence of risks of bias in randomized trials of therapeutic interventions for COVID-19.

Methods: Systematic review and risk of bias assessment performed by two independent reviewers of a random sample of 40 randomized trials of therapeutic interventions for moderate-severe COVID-19. We used the RoB 2.0 tool to assess the risk of bias, which evaluates bias under five domains as well as an overall assessment of each trial as high or low risk of bias.

Results: Of the 40 included trials, 19 (47%) were at high risk of bias, and this was particularly frequent in trials from low-middle income countries (11/14, 79%). Potential deviations to intended interventions (i.e., control participants accessing experimental treatments) were considered a potential source of bias in some studies (14, 35%), as was the risk due to selective reporting of results (6, 15%). The randomization process was considered at low risk of bias in most studies (34, 95%), as were missing data (36, 90%) and measurement of the outcome (35, 87%).

Conclusion: Many randomized trials evaluating COVID-19 interventions are at risk of bias, particularly those conducted in low-middle income countries. Biases are mostly due to deviations from intended interventions and partly due to the selection of reported results. The use of placebo control and publicly available protocol can mitigate many of these risks.

© 2022 The Author(s). Published by Elsevier Ltd on behalf of International Society for Infectious Diseases.

This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>)

Introduction

The COVID-19 pandemic resulted in a surge of medical research, with time-constrained circumstances leading to research conducted at accelerated rates (Mather, 2020). For example, the protocols for the platform trials, RECOVERY and Solidarity, were approved within a fraction of the conventional time (Tikkinen et al., 2020). RECOVERY and Solidarity were large multicenter studies led by experienced researchers, but many smaller, single-center trials have also been completed (Karlsen et al., 2020). Although a pandemic necessitates a fast response, it also raises the possibility of significant compromises to the methodologic quality and rigor of clinical trials (Jung et al., 2021). Poorly conducted trials, which

generate biased results, can ultimately translate into ineffective or even harmful treatments or might mean that truly effective treatments are disregarded as being ineffective (Schulz et al., 1995). Not only can such biases lead to detrimental health consequences, but the research can lead to wastage of valuable resources and is unethical (Chalmers and Glasziou, 2009; Ioannidis et al., 2014).

The effects of risks of bias in randomized trials have been well documented and researched before the pandemic (Schulz et al., 1995; Wood et al., 2008). The number of published randomized trials has been steadily increasing over the years; however, risk of bias is still high despite having improved over time (Vinkers et al., 2021). For example, allocation concealment and other biases associated with a poorly implemented randomization have improved, as has the use of trial registration (Clark et al., 2016; Savovic et al., 2012; Vinkers et al., 2021). Nonetheless, the risk of bias remains a prevalent issue, with publication in a low-impact journal and nonblinding being associated with an increased risk of bias (Vinkers et al., 2021). Risks of bias are known to be as-

* Corresponding author at: Public Health Building, University of Birmingham, B15 2TT.

E-mail address: k.hemming@bham.ac.uk (K. Hemming).

sociated with studies conducted in low-middle income countries (Panagiorou et al., 2013; Wells et al., 2021), trials funded for-profit (Hamm et al., 2010; Panagiorou et al., 2013), smaller trials (Dechartres et al., 2013), and studies with poor quality of reporting (Tikka et al., 2021).

Our objective was to identify randomized trials of COVID-19 interventions and evaluate their risk of bias using a validated assessment tool. The review was out of scope of PROSPERO, but a documented protocol was completed before the review was undertaken (Supplementary Material 1).

Methods

Inclusion and exclusion criteria

Included studies had to be published after the COVID-19 global outbreak (January 1, 2020 onward) and before the date the search was undertaken (June 18, 2021); preprints were not included. Owing to resource constraints, we determined that it was possible to undertake the detailed risk of bias review on a random sample of 40 eligible studies and limited these to only to studies published in English. Eligible studies were fully published randomized superiority trials of therapeutic interventions for COVID-19. To focus on studies only investigating treatment effectiveness rather than drug safety, we only included superiority or phase III or phase IV trials. Information on the trial phase was sought by assessing trial registration data if this was not clear from the published report. Studies with an unknown trial phase were included if they met the eligibility criteria for all other aspects.

Excluded study designs were phase I/II trials, nonrandomized trials, cluster-randomized trials, crossover trials, and noninferiority trials. Study protocols, pilot studies, preliminary trial reports, and conference abstracts were also excluded. We restricted study populations to participants with moderate/severe disease who were hospitalized (although trials that included participants with mild COVID-19 were included if they also included moderate/severe hospitalized participants). Trials investigating any therapeutic treatment of COVID-19 were included. This included traditional Chinese medicine, herbal medications, convalescent plasma, intravenous drugs, and oral medication but excluded prophylactic medication, vaccinations, or supportive therapy. No limitations were made on the basis of the type of comparator or number of study arms.

Search mechanism

We searched the World Health Organization (WHO) COVID-19 database and the Cochrane COVID-19 study register. The WHO COVID-19 database sources data from more than 30 databases, including MEDLINE and PubMed. Similarly, the Cochrane COVID-19 study register sources data from numerous databases, including Embase and Cochrane Central Register of Controlled Trials (CENTRAL). These databases have been created for finding COVID-19-related literature only; therefore, using their in-built search filters alone was deemed sufficient. Filters applied to the WHO COVID-19 database ensured that studies were clinical trials, had full text available, and were testing “therapy” interventions. Similar filters were used for conducting the Cochrane COVID-19 study register search (Supplementary Material 2).

Studies identified from the two searches were combined, duplicates were removed, and all studies underwent a title and abstract preliminary screen, followed by a full-text screen. All screening was undertaken independently by two reviewers, with a third opinion taken in cases of disagreement. As a validation step, we used the trials identified in a similar review to confirm that no

eligible studies were missed (Zhou et al., 2021). From those studies assessed to be eligible, we randomly selected a sample of 40 studies using the random number generator in Microsoft Excel to undertake the full risk of bias assessment.

We then screened each of the included full trial reports to identify any reference to study protocols or statistical analysis plans. Where none was identified, we carried out a Google search (full study title and author list) to identify any nonreferenced protocols or statistical analysis plans (these did not have to be published). We searched for trial registration documentation for each included study using any trial registration reported in the text or using Google searches.

Data collection process

Data were collected from the full trial report and any available protocol, statistical analysis plan, or trial registration. Data on study characteristics were collected by one reviewer only. Data for the risk of bias assessment were collected independently by two reviewers. After both assessments were completed, disagreements were identified, and a consensus (henceforth referred to as the joint assessment) was reached by discussion. Study reports were randomly sorted before data collection.

Data collected on general characteristics of trials

We collected the following trial characteristics: the type of intervention being tested (categorized as immune-based therapy, antiviral, corticosteroid, or other), the type of comparator (standard of care, placebo, or other intervention), the primary outcome type (classified as symptom severity, mortality, composite, or other), and the total realized sample size across all study arms (i.e., the number recruited rather than the number planned at the design stage). Additionally, we collected the following characteristics because of their previous associations with risk of bias: country of conduct (Asia, Europe, South America, Northeast Africa, United States, and multiple regions) to classify as low-middle income country or non-low-middle income country according to the 2020 gross national income (Fantom and Serajuddin, 2016); whether the trial was published in a high-impact journal (<https://libguides.anu.edu.au/medicine/journals/high-impact>); whether a trial protocol, statistical analysis plan, or trial registration were available; and whether the trial documented ethical approval and individual participant consent.

Risk of bias assessment

We used the RoB 2.0 tool to assess risk of bias (Higgins et al., 2016), which describes risks of bias under five domains (Table 1). These risks are identified by a series of signaling questions with a set of elaborations, providing extensive detail about how to answer the signal questions (Higgins et al., 2016). Reviewers independently answered each signaling question for each study in turn with full access to the supporting elaboration as an aid. Each signaling question allows an assessment of “yes,” “probably yes,” “no,” “probably no,” and “no information.” This assessment was made for what was considered the primary outcome of the study and assuming interest in the effect of assignment to the intervention (as opposed to the effect of adhering to the intervention). After the independent assessment of the signaling questions, the two independent reviewers resolved any discrepancies. This then formed the joint assessment. These joint assessments are summarized in a series of Supplementary tables, describing all responses under the signaling questions. In the tabulated results, we combine the classification of “probably yes” with “yes” and “probably no” with “no.” Following RoB 2.0, these joint responses to the signaling questions were

Table 1
Summary and description of risks of bias in randomized trials as documented in RoB 2.0.

Domain	Description
Domain 1: Risk of bias arising from the randomization process	Randomization creates (on average) an even distribution of participants across the study arms. To achieve this, the randomization must both be truly random, and it must be fully concealed at the time of recruitment. A poorly implemented randomization can manifest as notable differences in the characteristics of participants across the study arms.
Domain 2: Risk of bias due to deviations from the intended interventions	Participants should have been offered these interventions to evaluate the comparative effect of two interventions. Deviations in the uptake of this offer are usually not a concern when interest is in the effect of the intervention when some participants will naturally not want the intervention, for example, due to side effects or the way it is administered. The use of other concomitant interventions in both arms (background care) should be the same across both study arms. Problems arise when with knowledge of the participation in a study, particularly what intervention the participant is receiving, selectively lead to differences in this background care across the study arms.
Domain 3: Risk of bias due to missing outcome data	To evaluate the effect of the assignment to the intervention outcome, data are required on all randomized participants, even those who decline the offer of the intervention. Small amounts of missing outcome data, evenly distributed across study arms, are unlikely to be problematic. Problems arise when the amount of missing data is more substantial, and it differs across the study arms. Of note, bias can still arise when the proportion of missing outcome data is similar across study arms, but the characteristics of those with missing data differ across the arms.
Domain 4: Risk of bias in the measurement of the outcome	Assessment and measurement of the outcome should not be influenced or differ across the treatment arms. If those who are measuring the outcome have knowledge of the treatment arm, this may unintentionally affect the way the outcome is measured if there is any subjectivity involved in the outcome assessment.
Domain 5: Risk of bias in selection of the reported results	Clear specification of a primary outcome mitigates the problem of selecting outcomes that are apparently different, but this difference reflects a chance finding. To achieve this, the primary outcome must be clearly specified, including how it will be measured, scales that will be used, any cut-points that will be used, and when it will be measured (primary assessment time). Likewise, a clear primary analysis method should be specified, which includes how any covariates will be adjusted for in the analysis.

RoB, risk of bias.

mapped onto a risk of bias assessment for each domain, classifying trials as “low risk of bias,” “some concerns,” or “high risk of bias.” Finally, again following RoB 2.0, we created an overall study assessment of risk of bias: a study is judged at high risk of bias if it is assessed at high risk in at least one domain or some concerns for multiple domains, low risk of bias if it is assessed as low risk in all domains, and some concerns otherwise [Table 2](#), [Table 3](#), [Table 4](#).

Statistical analysis

We described the assessment of risk of bias (on the basis of the joint assessment) for all domains and signaling questions using simple descriptive statistics (numbers and percentages). For the overall assessment of bias, we summarized by the following characteristics (all identified elsewhere to be risk factors for bias): (1) those without a placebo control, (2) those without a publicly available protocol or statistical analysis plan, (3) trials with a sample size less than 150 (the median sample size in the sample), (4) low-middle income countries, and (5) non-high-impact journals.

Finally, we describe the reliability of the independent assessments by computing the percentage agreement (including raw percentage agreement and a weighted Gwet AC value) ([Gwet, 2014](#); [Wongpakaran et al., 2013](#)) between the two independent assessments for each of the five domains. Reliability was computed across an ordinal three-point scale (high risk of bias/some concerns/low risk of bias). The Gwet AC statistic was weighted with the penalization (weights) set to thirds: low penalization set to two-thirds for high-some concerns, low-some concerns, and anything-unclear; and high penalization set to one-third for high-low concerns.

Results

Searches were conducted on June 18, 2021, yielding a total of 1628 citations from the two databases. After 151 duplicated articles were removed, 1477 were assessed for title and abstract screening ([Figure 1](#)). A further 1323 were removed, leaving 154 studies for full-text screening. Of these, 69 trials were identified as eligi-

ble for inclusion (with reasons for exclusion provided in [Figure 1](#)). No additional studies were included after assessing the validation systematic review reference list. From this, a subset of 40 studies were randomly sampled to form the sampling frame (Supplementary Material 3) [Figure 2](#).

Study characteristics

Of the 40 included trials, many were conducted in Asia (18, 45%); some studies were conducted in Europe (7, 17%), South America (6, 15%), and Northeast Africa (2, 5%); one study (3%) was conducted in the United States, and some studies were conducted across multiple regions (6, 15%) ([Table 2](#)). Approximately, a third were conducted in low-middle income countries (14, 35%). The trials included a range of different therapeutic interventions: immune-based therapy (9, 23%), antiviral (8, 20%), and corticosteroids (5, 12%). The majority compared against standard of care (26, 65%), with only 25% (10) including a placebo control. A minority compared against another therapeutic intervention (4, 10%). The most commonly reported primary outcome type was symptom severity (24, 60%), with mortality being the primary outcome in seven studies (18%). Almost all studies were registered (37, 93%) and half had a publicly available protocol or statistical analysis plan (20, 50%). The average realized sample size was 150 (interquartile range: 89–379). All studies reported obtaining ethical approval and consent.

Broad assessment of risk of bias

Only 10 (25%) of the studies were assessed to be at low risk of bias, 19 (47%) were assessed at high risk of bias, and 11 (28%) were assessed as having some concerns ([Table 3](#)). Most trials were assessed as low risk of bias due to the randomization process (34, 85%), bias due to missing data (36, 90%), and bias due to measurement of the outcome (35, 87%). In contrast, fewer trials were assessed as low risk of bias due to deviations from the intended interventions (16, 40%) and due to selection of the reported result (21, 52%).

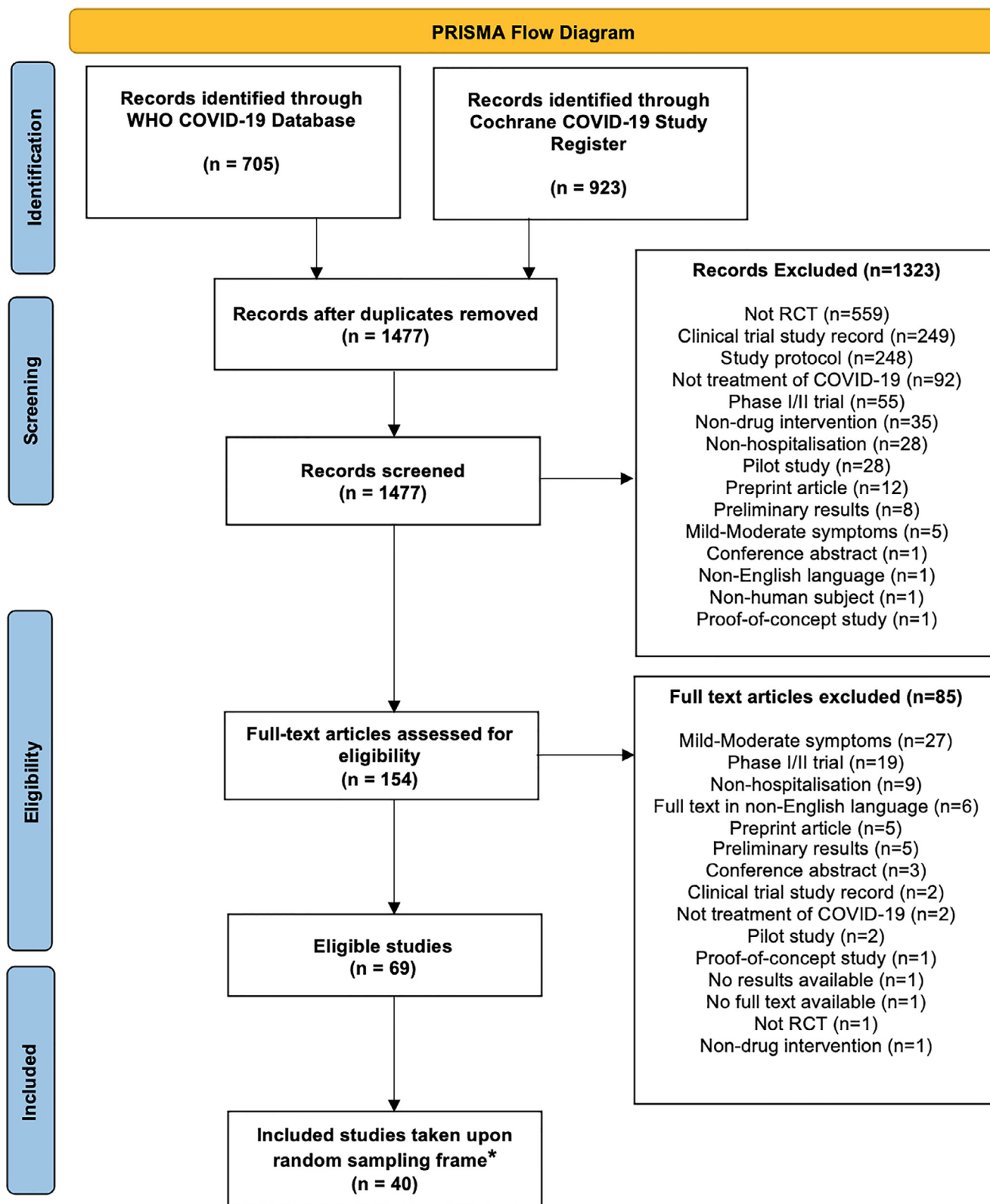


Fig. 1. Flow diagram trials identified and screened for inclusion. COVID-19, coronavirus disease 2019; PRISMA, Preferred Reporting Items for Systematic Reviews and Meta-Analyses; RCT, randomized controlled trial; WHO, World Health Organization.

Table 2
 Characteristics of trials included in review (N = 40).

Characteristic	n (%)
Journal	18 (45)
High-impact	22 (55)
Not high-impact	
Country of study conduct	18 (45)
Asia	7 (17)
Europe	6 (15)
South America	2 (5)
Northeast Africa	1 (3)
United States	6 (15)
Multiple regions	
Income of country	14 (35)
LMIC	26 (65)
Non-LMIC	
Type of intervention	9 (23)
Immune-based therapy	8 (20)
Antiviral	5 (12)
Corticosteroid	18 (45)
Other	
Comparator	26 (65)
Standard of care	10 (25)
Placebo	4 (10)
Other intervention	
Primary outcome	24 (60)
Symptom severity	7 (18)
Mortality	4 (10)
Composite	5 (12)
Other	
Prespecification documentation availability	20 (50)
Accessible protocol paper or SAP	37 (92.5)
Trial registration	3 (8)
Neither protocol paper nor trial registration	
Average study size ^a , median (IQR)	149.5 (88.8 – 378.5)

IQR, interquartile range; LMIC, low-middle income country; SAP, statistical analysis plan.

^a Numbers refer to realized numbers across all study arms as opposed to those planned in any sample size calculation for example (i.e., number of participants on whom baseline measures were taken).

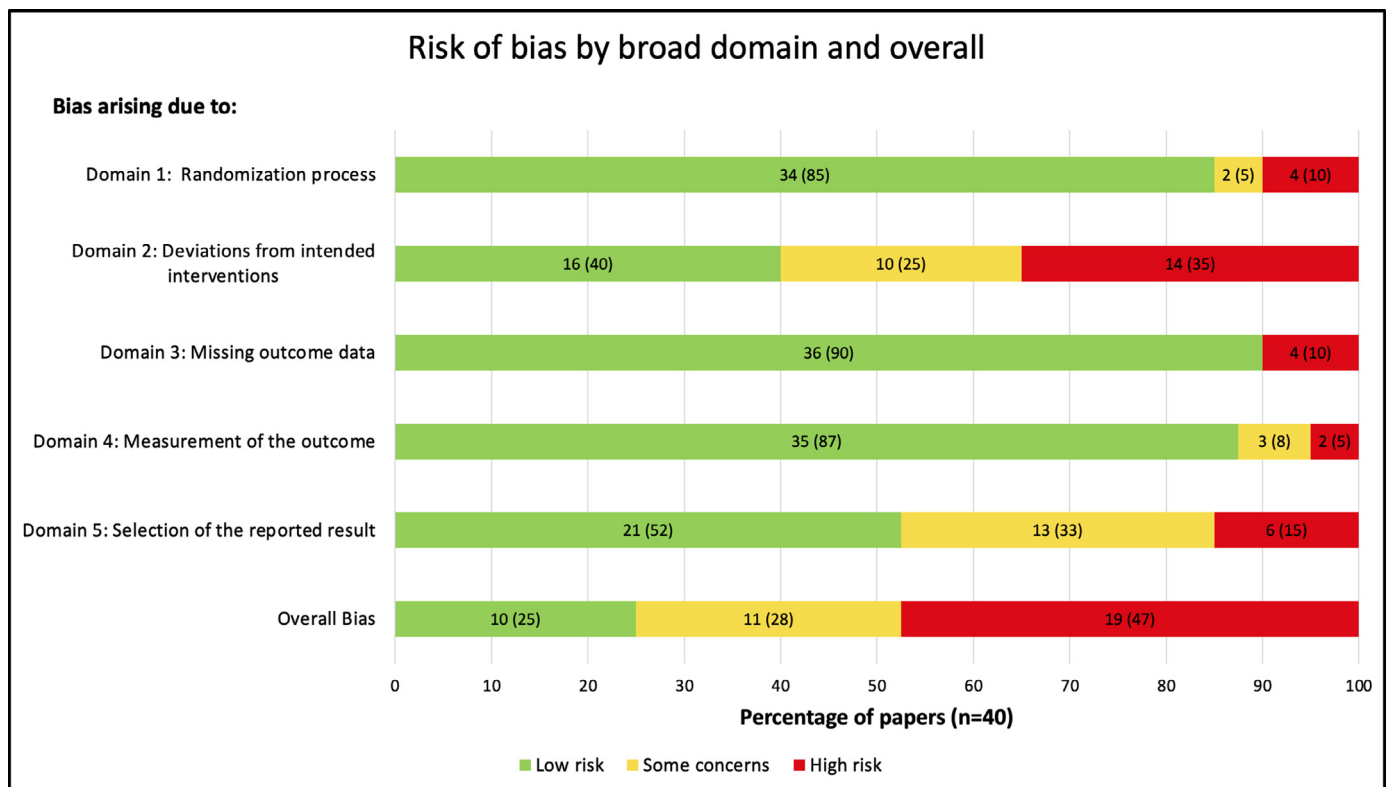


Fig. 2. Percentage of papers in each risk category across the broad domains of risk.

Table 3
Risk of bias assessment by broad domains of risk.

Domain	Level of risk	n (%) N = 40	Reliability between reviewers	
			Gwet's AC (95% CI)	% agreement
1. Bias arising from the randomization process	Low risk	34 (85)	71.5 (52.1–90.9)	67.5
	Some concerns	2 (5)		
	High risk	4 (10)		
2. Bias due to deviations from intended interventions	Low risk	16 (40)	21.3 (–8.1 to 50.6)	40.0
	Some concerns	10 (25)		
	High risk	14 (35)		
3. Bias due to missing outcome data	Low risk	36 (90)	79.3 (64.2–94.4)	75.0
	Some concerns	0 (0)		
	High risk	4 (10)		
4. Bias in measurement of the outcome	Low risk	35 (87)	91.9 (84.0–99.8)	85.0
	Some concerns	3 (8)		
	High risk	2 (5)		
5. Bias in selection of the reported results	Low risk	21 (52)	58.3 (35.2–81.2)	62.5
	Some concerns	13 (33)		
	High risk	6 (15)		
Overall risk of bias judgment ^a	Low risk	10 (25)	33.2 (5.0–61.3)	47.5
	Some concerns	11 (28)		
	High risk	19 (47)		
Number of domains at high risk	0 ^b	20 (50)		
	1	13 (33)		
	2	4 (10) ^u		
	3	3 (7)		
	4	0 (0)		
	5	0 (0)		

CI, confidence interval.

^a Overall risk of bias judgment: low risk of bias is defined as all domains at low risk of bias; some concerns is defined as at least one domain has some concerns but does not include any high risk of bias for any domain; and high risk of bias is defined as high risk of bias in at least one domain or some concerns for multiple domains.

^b Zero domains at risk includes one at low risk and two with some concerns (overall risk). Gwet's AC statistic is weighted (see Methods) to give more weight to disagreements between low and high risk compared with some concerns.

Table 4
Overall assessment of bias by selected characteristics.

Characteristic	Level of risk	n (%)
No placebo control	Low risk	4 (13)
	Some concerns	9 (30)
	High risk	17 (57)
No protocol/SAP available	Low risk	0 (0)
	Some concerns	8 (40)
	High risk	12 (60)
Sample size <150	Low risk	2 (10)
	Some concerns	6 (30)
	High risk	12 (60)
Low-middle income countries	Low risk	1 (7)
	Some concerns	2 (14)
	High risk	11 (79)
Non-high-impact journal	Low risk	0 (0)
	Some concerns	8 (36)
	High risk	14 (64)

SAP, statistical analysis plan.

Domain 1: bias arising from the randomization process

Almost all studies used a clearly reported random allocation method (37, 93%), and in most studies (32, 80%), it was clear that the allocation had been properly concealed before the consent process (Supplementary Table 1). Baseline imbalance was identified

(by the reviewers) in only a small number of trials (3, 7%), which might indicate a poorly implemented randomization process.

Domain 2: bias due to deviations from intended interventions

In many trials, the participants (27, 67%) and their carers or individuals delivering the intervention (29, 72%) were aware of the assigned interventions (Supplementary Table 2). In some trials (15, 37%), it was assessed that this knowledge could have led to deviations from the intended interventions and that these deviations could have affected the outcome (11, 27%). In some trials (3, 7%), these deviations were assessed to affect one of the intervention arms more than the other. Most trials (31, 77%) carried out an analysis to appropriately estimate the effect of the assignment to the intervention (i.e., an intention to treat analysis), and in only a few studies (5, 13%) was it assessed that any deviations to assigned interventions could have had a substantial impact.

Domain 3: bias due to missing outcome data

In most trials (32, 80%), the outcome data were available for all or nearly all of the participants (Supplementary Table 3). However, in a minority of trials (7, 17%), it was assessed that the missing outcomes might have been biased by the missing data, perhaps because the missingness might depend on the outcome (4, 10%).

Domain 4: bias in measurement of the outcome

In most studies, the method of measuring the outcome was assessed to be appropriate (37, 93%), and it was assessed that the measurement of the outcome could have differed across the study arms only occasionally (2, 5%) (Supplementary Table 4). Outcome

assessors were often aware of the allocated study arm (17, 43%), although this was not considered to have influenced outcome assessment in any studies.

Domain 5: bias in selection of the reported result

Although in more than half of the studies (24, 60%), it was assessed that the data were analyzed according to a prespecified approach, in many studies (12, 30%), this was unclear (Supplementary Table 5). In a number of studies (9, 23%), it was assessed that there were multiple possible eligible outcomes (i.e., the scale, cut-points, assessment times were not clearly defined), and in a few studies (2, 5%), it was assessed that there had been multiple possible analyses of the data; however, in many studies (17, 43%), no information on prespecification was provided.

Design features associated with increased risks of bias

When investigating whether there was a tendency for trials with certain characteristics to be at increased risk of bias (compared with the 47% classified as at-risk overall 40 trials), we identified the that likelihood of high risk of bias was 57% (17/30) in those trials without the inclusion of a placebo control, 60% (12/20) in those trials without a publicly available protocol or statistical analysis plan, 60% (12/20) in small trials (sample size <150), 79% (11/14) in those from low-middle income countries, and 64% (14/22) in those in a non-high-impact journal (Table 4).

Reliability of independent assessments

Agreement between the two independent assessments varied across the domains and was greater for the randomization process (Gwet AC: 71.5; 95% confidence interval [CI]: 52.1–90.9), missing outcome data (Gwet AC: 79.3; 95% CI: 64.2–94.4), measurement of outcome (Gwet AC: 91.9 95% CI: 84.0–99.8) but lower for risks of bias around deviations from intended interventions (Gwet AC: 21.3; 95% CI: –8.1 to 50.6), selection of the reported result (Gwet AC: 58.3; 95% CI: 35.2–81.2), and overall assessment (Gwet AC: 33.2; 95% CI: 5.0–61.3).

Discussion

Summary of findings

We identified that a substantial proportion of the randomized trials of COVID-19 interventions published over the first 18 months of the pandemic are at risk of bias. We identified that these risks are mostly due to risks around deviations from the intended intervention (e.g., control arm participants receiving the active intervention) and that risks of bias were particularly prevalent in studies conducted in low-middle income countries. Other risks of bias, such as those due to selection of the reported result or measurement of the outcome, were much lower but nonetheless still prevalent in a significant minority of the studies. Modifiable factors that will reduce these risks are using a placebo control and publicly available protocol: only half of the trials had a publicly available protocol, and only a quarter used a placebo control.

Research in context

High-quality evidence is always important, particularly in the context of a pandemic (Raynaud et al., 2021). Evidence is emerging that many COVID-19 trials are of poor quality: reviews of registered COVID-19 trials (as opposed to fully published studies) have revealed that many are not blinded and have many other features associated with poor quality (Jung et al., 2021; Karlsen et al., 2020; Mainoli et al., 2021; Mehta et al., 2020; Pundi et al., 2020;

Zhu et al., 2020). Additionally, there is evidence that many studies of traditional Chinese medicine for COVID-19 are also at risk of bias (Alexander et al., 2020; Gao et al., 2021), and where risk of bias has been investigated in the context of specific treatments (e.g., hydroxychloroquine), concerns have been raised of the lack of high-quality studies (Mazhar et al., 2020). Moreover, the BMJ living systematic review investigating treatments for COVID-19 also conducted a risk of bias assessment using a revised version of the Cochrane RoB 2.0 tool and found most studies (~60%) to have a high risk of bias (BMJ, 2021). Our findings here also support those of another risk of bias assessment using the older version of the Cochrane risk of bias tool but for which the assessment was not conducted in duplicate (Zhao et al., 2021) and are not dissimilar to assessments of trials in settings other than COVID-19 (Hamm et al., 2020; Turner et al., 2013; Vinkers et al., 2021).

The importance of blinding

The most common source of bias identified in this review was due to deviations from intended interventions. Bias due to deviations of intended interventions, also known as performance bias, is unlikely to affect fully blinded trials (Porta, 2014). This bias refers to the use of other complementary interventions differentially across study arms in a way that would not happen outside a trial context or to the use of the active intervention in the control arm. Others have argued that blinding is not desirable in pragmatic trials where the objective is to estimate effect in real-world conditions (Mansournia et al., 2017). In pragmatic trials, the interest is not in the isolated effect of the intervention but rather how it works in its proposed context (Christian et al., 2020). Arguably, the COVID-19 trials included in this review were all wanting to estimate effects in “real-world conditions.” Indeed, any non-adherence to the active intervention was not considered a deviation in our assessment of risk of bias, in line with pragmatic intent (Higgins, 2019). However, a bias can arise, even in pragmatic trials when, for example, complementary interventions (e.g., the use of other experimental concomitant therapies in patients with COVID-19) is differentially used across trial arms because of a perceived rather than real need (Henderson et al., 2007). This type of bias very possibly affected one of the recovery trials, where 17% of patients in the usual care group were given azithromycin (the active intervention) or another macrolide antibiotic (RECOVERY Collaborative group, 2021). It can thus be very important in the so-called pragmatic trials to blind therapeutic interventions to prevent these types of biases, and when blinding is not possible, reporting post-randomization treatments by arm can aid in the identification of this type of bias. In relation to this, in the situation of possible side effects due to the active intervention, controls can be very important to prevent the possibility of unblinding when side effects occur (Jensen et al., 2017).

Three large, multicenter pragmatic trials (RECOVERY, Solidarity, and ACTT) are widely hailed as exemplary designs, cutting red tape, delivering fast, and promoting collaboration (Tikkinen et al., 2020). Yet, both the RECOVERY and Solidarity trials were assessed as having some concerns in our review and at high risk of bias by the BMJ living systematic review (BMJ, 2021). For both, this assessment of high risk of bias was because of potential deviations to intended interventions, arising because of the unblinded nature of the studies. The ACTT trial, also large, pragmatic, and timely, was blinded; therefore, it is unlikely to be at risk of bias due to deviations from intended interventions (Beigel, 2020). For one drug compared in these trials, remdesivir, the (unblinded) Solidarity trial (WHO Solidarity Trial Consortium et al., 2021) and the (blinded) ACTT trial (Beigel, 2020) gave conflicting results. Direct conflict of study results, such as this, especially when one of the studies is unblinded, creates uncertainty, particularly in the case

where the unblinded trial (sample size in Solidarity ~11,000) is much larger than the blinded trial (sample size in ACTT ~1,000).

Limitations

The Cochrane risk of bias tool is a generic tool designed to be used across a range of randomized trials although it generally does not consider aspects that might relate only to certain types of trials, for example, platform trials for which there can be other specific concerns around biases (Normand, 2021). We applied the risk of bias assessment only to the primary outcome and only to ascertain the risk against the effect of assignment to the intervention. Risks of bias might differ for other outcomes; although, arguably, the primary outcome should be considered more important. Interest might reasonably be in the effect of adhering to the treatment (efficacy as opposed to effectiveness), particularly in earlier phases of treatment assessment (Hernán and Robins, 2017). Although our assessment was performed independently by two reviewers, the agreement between the independent assessments was low for those domains where risk of bias was assessed as a potential concern (deviations from intended interventions and selection of reported results). These discrepancies likely represent a real uncertainty around whether bias is or is not present. This low agreement might be considered a further indicator of potential bias or poor reporting (Hartling et al., 2013). Our assessment included only 40 of 69 eligible studies. These 40 studies represent a random sample of those eligible, and although we are not identifying risk of bias to inform treatment decision, this review can nonetheless inform on the likely risk of bias in trials of COVID-19 interventions more generally.

Conclusion

High-quality, low risk of bias randomized trials are fundamental to responding to a pandemic. Several large platform trials are exemplary in their timeliness and collaborative nature. Nonetheless, many smaller trials have been initiated. A large majority of published trials testing COVID-19 treatments are at high risk of bias, particularly those conducted in low-middle income countries. Even the small number of trials considered exemplary might be at risk of bias, mostly because of their unblinded nature. To ensure that patients receive effective treatments, future randomized trials must be designed to be at low risk of bias to avoid wastage of research and spurious findings.

Conflict of interest

The authors have no conflicts of interest to declare.

Funding

Karla Hemming is funded by a National Institute for Health and Care Research (NIHR) Senior Research Fellowship SRF-2017-10-002. This research was also partly funded by the UK NIHR Collaborations for Leadership in Applied Health Research and Care West Midlands initiative.

Ethical approval statement

No ethical approval was obtained for this study.

Acknowledgments

Christina Easter acted as an independent opinion for the second screen.

Author contributions

KH led the development of the project and wrote the first draft of the study. KK led the search process and identification of studies for inclusion, designed and developed the data collection tools, collected the data, and conducted the statistical analysis. KK, JT, BM, and KH collected the data. KH, JT, and BM acted as the independent opinion in case of disagreements during data collection. VM conducted the second study screening for eligibility. All authors commented on the draft study.

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.ijid.2022.05.034.

References

- Alexander PE, Debono VB, Mammen MJ, Iorio A, Aryal K, Deng D, et al. COVID-19 coronavirus research has overall low methodological quality thus far: case in point for chloroquine/hydroxychloroquine. *J Clin Epidemiol* 2020;123:120–6.
- Beigel JH, Tomashek KM, Dodd LE, Mehta AK, Zingman BS, Kalil AC, Hohmann E, Chu HY, Luetkemeyer A, Kline S, Lopez de Castilla D, Finberg RW, Dierberg K, Tapson V, Hsieh L, Patterson TF, Paredes R, Sweeney DA, Short WR, Touloumi G, Lye DC, Ohmagari N, Oh MD, Ruiz-Palacios GM, Benfield T, Fätkenheuer G, Kortepeter MG, Atmar RL, Creech CB, Lundgren J, Babiker AG, Pett S, Neaton JD, Burgess TH, Bonnett T, Green M, Makowski M, Osinusi A, Nayak S, Lane HC; ACTT-1 Study Group Members. Remdesivir for the Treatment of Covid-19 – Final Report. *N Engl J Med*. 2020 Nov 5;383(19):1813–1826. doi:10.1056/NEJMoa2007764. Epub 2020 Oct 8. PMID: 32445440; PMCID: PMC7262788.
- BMJ. Drug treatments for COVID-19: living systematic review and network meta-analysis. *BMJ* 2021;373:n967.
- Chalmers I, Glasziou P. Avoidable waste in the production and reporting of research evidence. *Obstet Gynecol* 2009;114:1341–5.
- Christian JB, Brouwer ES, Girman CJ, Bennett D, Davis KJ, Dreyer NA. Masking in pragmatic trials: who, what, and when to blind. *Ther Innov Regul Sci* 2020;54:431–6.
- Clark L, Fairhurst C, Torgerson DJ. Allocation concealment in randomised controlled trials: are we getting better? *BMJ* 2016;355:i5663.
- Dechartres A, Trinquart L, Boutron I, Ravaud P. Influence of trial sample size on treatment effect estimates: meta-epidemiological study. *BMJ* 2013;346:f2304.
- Gao G, Leung SW, Jia Y. Serious concerns about the RCT study quality of traditional Chinese medicine injections for treating cardiovascular diseases and COVID-19. *Cardiovasc Drugs Ther* 2021;35:191–3.
- Gwet KL. Handbook of inter-rater reliability: the definitive guide to measuring the extent of agreement among raters. 4th ed. Gaithersburg, MD: Advanced Analytics; 2014.
- Fantom NJ, Serajuddin U. World Bank's classification of countries by income. World Bank policy research working paper no. 7528. <https://openknowledge.worldbank.org/handle/10986/23628>, 2016. Accessed 18 June 2021.
- Hartling L, Hamm MP, Milne A, Vandermeer B, Santaguida PL, Ansari M, et al. Testing the risk of bias tool showed low reliability between individual reviewers and across consensus assessments of reviewer pairs. *J Clin Epidemiol* 2013;66:973–81.
- Hamm MP, Hartling L, Milne A, Tjosvold L, Vandermeer B, Thomson D, et al. A descriptive analysis of a representative sample of pediatric randomized controlled trials published in 2007. *BMC Pediatr* 2010;10:96.
- Henderson GE, Churchill LR, Davis AM, Easter MM, Grady C, Joffe S, et al. Clinical trials and medical care: defining the therapeutic misconception. *PLoS Med* 2007;4:e324.
- Hernán MA, Robins JM. Per-protocol analyses of pragmatic trials. *N Engl J Med* 2017;377:1391–8.
- Higgins JPT, Sterne JAC, Savović J, Page MJ, Hróbjartsson A, Boutron I, Reeves B, Eldridge S. A revised tool for assessing risk of bias in randomized trials. In: Chandler J, McKenzie J, Boutron I, Welch V, editors. *Cochrane Methods. Cochrane Database of Systematic Reviews* 2016, Issue 10 (Suppl 1). dx.doi.org/10.1002/14651858.CD201601. Accessed 12 June 2021.
- Ioannidis JP, Greenland S, Hlatky MA, Khoury MJ, Macleod MR, Moher D, et al. Increasing value and reducing waste in research design, conduct, and analysis. *Lancet* 2014;383:166–75.
- Jensen JS, Bielefeldt AØ, Hróbjartsson A. Active placebo control groups of pharmacological interventions were rarely used but merited serious consideration: a methodological overview. *J Clin Epidemiol* 2017;87:35–46.
- Jung RG, Di Santo P, Clifford C, Prosperi-Porta G, Skanes S, Hung A, et al. Methodological quality of COVID-19 clinical research. *Nat Commun* 2021;12:943.
- Karlsen APH, Wiberg S, Laigaard J, Pedersen C, Rokamp KZ, Mathiesen O. A systematic review of trial registry entries for randomized clinical trials investigating COVID-19 medical prevention and treatment. *PLoS One* 2020;15.
- Mainoli B, Machado T, Duarte GS, Prada L, Gonçalves N, Ferreira JJ, et al. Analysis of clinical and methodological characteristics of early COVID-19 treatment clinical trials: so much work, so many lost opportunities. *BMC Med Res Methodol* 2021;21:42.

- Mansournia MA, Higgins JPT, Sterne JAC, Hernán MA. Biases in randomized trials: a conversation between trialists and epidemiologists. *Epidemiology* 2017;28:54–9.
- Mather N. How we accelerated clinical trials in the age of coronavirus. *Nature* 2020;584:326.
- Mazhar F, Hadi MA, Kow CS, Marran AMN, Merchant HA, Hasan SS. Use of hydroxychloroquine and chloroquine in COVID-19: how good is the quality of randomized controlled trials? *Int J Infect Dis* 2020;101:107–20.
- Mehta HB, Ehrhardt S, Moore TJ, Segal JB, Alexander GC. Characteristics of registered clinical trials assessing treatments for COVID-19: a cross-sectional analysis. *BMJ Open* 2020;10.
- Normand ST. The RECOVERY Platform. *N Engl J Med* 2021;384:757–8.
- Panagiotou OA, Contopoulos-Ioannidis DG, Ioannidis JPA. Comparative effect sizes in randomised trials from less developed and more developed countries: meta-epidemiological assessment. *BMJ* 2013;346:f707.
- Porta, Miquel, ed. *A dictionary of epidemiology*. Oxford university press, 2014.
- Pundi K, Perino AC, Harrington RA, Krumholz HM, Turakhia MP. Characteristics and strength of evidence of COVID-19 studies registered on ClinicalTrials.gov. *JAMA Intern Med* 2020;180:1398–400.
- Raynaud M, Zhang H, Louis K, Goutaudier V, Wang J, Dubourg Q, et al. COVID-19-related medical research: a meta-research and critical appraisal. *BMC Med Res Methodol* 2021;21:1.
- RECOVERY Collaborative Group. Azithromycin in patients admitted to hospital with COVID-19 (RECOVERY): a randomised, controlled, open-label, platform trial. *Lancet* 2021;397:605–12.
- Savović J, Jones HE, Altman DG, Harris RJ, Jüni P, Pildal J, et al. Influence of reported study design characteristics on intervention effect estimates from randomized, controlled trials. *Ann Intern Med* 2012;157:429–38.
- Schulz KF, Chalmers I, Hayes RJ, Altman DG. Empirical evidence of bias. Dimensions of methodological quality associated with estimates of treatment effects in controlled trials. *JAMA* 1995;273:408–12.
- Tikkinen KAO, Malekzadeh R, Schlegel M, Rutanen J, Glasziou P. COVID-19 clinical trials: learning from exceptions in the research chaos. *Nat Med* 2020;26:1671–2.
- Tikka C, Verbeek J, Ijaz S, Hoving JL, Boschman J, Hulshof C, de Boer AG. Quality of reporting and risk of bias: a review of randomised trials in occupational health. *Occup Environ Med* 2021;78:691–6.
- Turner L, Boutron I, Hróbjartsson A, Altman DG, Moher D. The evolution of assessing bias in Cochrane systematic reviews of interventions: celebrating methodological contributions of the Cochrane Collaboration. *Syst Rev* 2013;2:79.
- Vinkers CH, Lamberink HJ, Tjink JK, Heus P, Bouter L, Glasziou P, et al. The methodological quality of 176,620 randomized controlled trials published between 1966 and 2018 reveals a positive trend but also an urgent need for improvement. *PLoS Biol* 2021;19.
- Solidarity Trial Consortium WHO, Pan H, Peto R, Henao-Restrepo AM, Preziosi MP, Sathiyamoorthy V, et al. Repurposed antiviral drugs for Covid-19 - interim WHO Solidarity Trial results. *N Engl J Med* 2021;384:497–511.
- Wells JC, Sharma S, Del Paggio JC, Hopman WM, Gyawali B, Mukherji D, et al. An analysis of contemporary oncology randomized clinical trials from low/middle-income vs high-income countries. *JAMA Oncol* 2021;7:379–385.
- Wood L, Egger M, Gluud LL, Schulz KF, Jüni P, Altman DG, et al. Empirical evidence of bias in treatment effect estimates in controlled trials with different interventions and outcomes: meta-epidemiological study. *BMJ* 2008;336:601–5.
- Wongpakaran N, Wongpakaran T, Wedding D, Gwet KL. A comparison of Cohen's Kappa and Gwet's AC1 when calculating inter-rater reliability coefficients: a study conducted with personality disorder samples. *BMC Med Res Methodol* 2013;13:61.
- Zhao MZ, Zhao C, Tu SS, Wei XX, Shang HC. Evaluating the methodology of studies conducted during the global COVID-19 pandemic: a systematic review of randomized controlled trials. *J Integr Med* 2021;19:317–26.
- Zhu RF, Gao YL, Robert SH, Gao JP, Yang SG, Zhu CT. Systematic review of the registered clinical trials for coronavirus disease 2019 (COVID-19). *J Transl Med* 2020;18:274.