RESEARCH ARTICLE

Research Synthesis Methods **WILEY**

Identifying and managing problematic trials: A research integrity assessment tool for randomized controlled trials in evidence synthesis

Stephanie Weibel[1] [ORCID] | Maria Popp[1] | Stefanie Reis[1] | Nicole Skoetz[2] | Paul Garner[3] | Emma Sydenham[4]

[1]Department of Anaesthesiology, Intensive Care, Emergency and Pain Medicine, University Hospital Wuerzburg, Wuerzburg, Germany

[2]Evidence-based Oncology, Department I of Internal Medicine, Center for Integrated Oncology Aachen Bonn Cologne Duesseldorf, Faculty of Medicine and University Hospital Cologne, University of Cologne, Cologne, Germany

[3]Centre for Evidence Synthesis in Global Health, Liverpool School of Tropical Medicine, Liverpool, UK

[4]Cochrane Injuries Group, London School of Hygiene & Tropical Medicine, London, UK

Correspondence
Stephanie Weibel, Department of Anaesthesiology, Intensive Care, Emergency and Pain Medicine, University Hospital Wuerzburg, Oberduerrbacher Str. 6, Wuerzburg 97080, Germany.
Email: weibel_s@ukw.de

Abstract

Evidence synthesis findings depend on the assumption that the included studies follow good clinical practice and results are not fabricated or false. Studies which are problematic due to scientific misconduct, poor research practice, or honest error may distort evidence synthesis findings. Authors of evidence synthesis need transparent mechanisms to identify and manage problematic studies to avoid misleading findings. As evidence synthesis authors of the Cochrane COVID-19 review on ivermectin, we identified many problematic studies in terms of research integrity and regulatory compliance. Through iterative discussion, we developed a research integrity assessment (RIA) tool for randomized controlled trials for the update of this Cochrane review. In this paper, we explain the rationale and application of the RIA tool in this case study. RIA assesses six study criteria: study retraction, prospective trial registration, adequate ethics approval, author group, plausibility of methods (e.g., randomization), and plausibility of study results. RIA was used in the Cochrane review as part of the eligibility check during screening of potentially eligible studies. Problematic studies were excluded and studies with open questions were held in awaiting classification until clarified. RIA decisions were made independently by two authors and reported transparently. Using the RIA tool resulted in the exclusion of >40% of studies in the first update of the review. RIA is a complementary tool prior to assessing "Risk of Bias" aiming to establish the integrity and authenticity of studies. RIA provides a platform for urgent development of a standard approach to identifying and managing problematic studies.

KEYWORDS
COVID-19 pandemic, evidence synthesis, good clinical practice, randomized controlled trial, research integrity, systematic review

Highlights

What is already known?
• Including problematic studies in terms of research integrity in evidence syntheses can lead to misleading conclusions, and harm human health. Cochrane has published guidance to facilitate research integrity checks in the reviews it publishes. However, these checks have not routinely formed part of evidence synthesis or guideline development processes to date.

What is new?
• We developed the research integrity assessment (RIA), a tool to assess the integrity of research and adherence to good clinical practice reported in RCTs of investigational drugs for an update of a Cochrane COVID-19 review. The assessment uses signaling questions to identify problematic RCTs and is used when studies are being considered for inclusion in an evidence synthesis. Problematic studies are excluded or held in an awaiting classification category.

Potential impact for research synthesis methods readers outside the authors' field
• Ensuring research integrity of studies included in evidence synthesis is an approach that helps to provide evidence closer to the unbiased truth and improves respect of human rights in evidence synthesis.

1 | INTRODUCTION

Systematic reviews aim to identify all studies that meet the eligibility criteria for the review. Studies that challenge the principles of good clinical practice and scientific integrity can mislead and corrupt the findings of a systematic review, and hence mislead guidelines and official recommendations that use the reviews. Public health laws may be developed based on the findings of systematic reviews, which has been particularly important in responding to the COVID-19 pandemic.[1] This causes a dilemma for review authors: Ioannidis et al. pointed out that authors should become aware that false and fatally flawed trials are very common in the field of medicine and suggests toning down the confidence in their conclusions.[2] Avenell et al. found evidence that trials retracted due to misconduct distorted the evidence base including systematic reviews, meta-analyses, narrative reviews, and clinical guidelines citing such trials and concluded that many of those guidelines, systematic or other reviews would likely change their findings if the affected trial reports were removed.[3]

The COVID-19 pandemic threw these dilemmas into sharp focus. We systematically synthesized the evidence for the Cochrane review "Ivermectin for preventing and treating COVID-19."[4] During the second pandemic year and after publication of the Cochrane review, several clinical trials were retracted due to critical concerns on trustworthiness.[5–8] After their retraction, studies claiming to prove ivermectin's huge beneficial effect for treating this disease remain considered in published evidence syntheses.[9–12] Additionally, it has to be considered that even if evidence syntheses are retracted as well[13] or corrected for such distortions[14,15] the initial effect on patient demand, public beliefs, clinical practice, and research priorities was difficult to reverse. Whilst most of the retracted studies actually did not meet our inclusion criteria, we felt uneasy about the full extent of the problematic study pool investigating ivermectin for COVID-19, and in preparing the review update, we developed a tool for our Cochrane review to help us identify studies that were potentially "problematic" in relation to whether they had been fabricated, data had been altered, or were not in accordance with good clinical practice.

Mechanisms for identifying studies with fabricated or false data, or where problems with research conduct amounting to failure of research integrity, is an active area of research. A group of researchers recently reported results from a qualitative international interview study with the aim to better understand the views of research integrity experts about what might make a study problematic or untrustworthy and what warning signs could be used in a practical screening tool to identify potentially problematic studies.[16] Cochrane defines a "problematic study" as "any published or unpublished study where there are serious questions about the

trustworthiness of the data or findings, regardless of whether the study has been formally retracted. Scientific misconduct will not be the only reason that a study might be problematic; problems may result from poor research practices or honest errors."[17] Cochrane has also published guidance to facilitate research integrity checks in the reviews it publishes,[18,19] but these checks have not routinely formed part of evidence synthesis or guideline development processes to date. We drew on a few non-validated tools available at the time of updating our Cochrane review to develop our approach (such as the "REAP-PRAISED" checklist for evaluation of publication integrity[20] and the data extraction tool from the Cochrane Pregnancy and Childbirth Group[21] that addresses various aspects of scientific integrity). The latter tool was recently used for the systematic exploration of trustworthiness of published trial data in 10 RCTs from one author investigating psychological interventions for the treatment of chronic pain.[22] With this screening tool the authors identified concerns about research governance, data plausibility at baseline, the results, and apparent data duplication.

In this paper, we describe the research integrity assessment (RIA) tool to assess randomized controlled trials (RCTs) and how we have used it for the Cochrane review on ivermectin.[23] This case study is another important step for the urgent development and adoption of a standard approach to sifting out problematic studies.

2 | METHODS

2.1 | Development of the new research integrity assessment tool

We developed a tool for the assessment of research integrity of RCTs focusing on investigational medicinal products (IMPs)[24–26]. The tool was developed prior to preparing the first update of the Cochrane review "Ivermectin for preventing and treating COVID-19"[23] and is guided by specific questions of how to identify and deal with problematic studies in the context of this systematic review. Study characteristics to assure research integrity of RCTs have already been discussed and considered in various other publications and places,[18,20,21] and some characteristics are legally required, such as ethics committee approval.[27] Based on study characteristics and specific items reported in existing screening tools[20,21] and legal requirements,[27] we used iterative discussions, piloting preliminary forms, and web-conferences among our team of six authors to agree on critical and important study characteristics and handling of identified problematic studies in this systematic review. Selection of study characteristics and criteria is based on the content expertise and

subjective assessment of the authors involved. We are content experts who were either authors on the Cochrane ivermectin review (Stephanie Weibel, Maria Popp, Stefanie Reis, and Nicole Skoetz), part of the editorial process (Paul Garner), or had previously developed strategies for dealing with problematic studies in Cochrane Review Groups (Emma Sydenham). The assessment tool was developed in an Excel-based format and contains questions to critical and important criteria that help to identify problematic RCTs when deciding on inclusion into a systematic review. The tool was used in a case study of the Cochrane review update on ivermectin and it was not validated.

3 | RESULTS

3.1 | RIA: Critical and important study characteristics to assess research integrity of RCTs investigating IMPs

We achieved consensus among the authors on six domains considering critical and important study characteristics to assure research integrity of RCTs investigating IMPs based on adherence to good clinical practice and scientific integrity: Retraction notices, prospective trial registration, ethics committee approval, and written informed consent, author group, sufficient reporting and plausibility of methods (e.g., study design/randomization), and plausibility of study results. In the workflow of RIA through domains 1–6, three decisions on a study's eligibility for the evidence synthesis are possible at any hierarchical steps from 1 to 6: RCTs may be either included or excluded from the evidence synthesis, or moved to awaiting classification (Figure 1). The "awaiting classification" category is typically used in Cochrane reviews for studies about which an inclusion or exclusion decision cannot be made because insufficient information is currently available.[28] In the following paragraphs, we introduce the RIA tool based on critical and important criteria of study characteristics summarized in key domains, explain their rationale, and provide methodological guidance. Critical and important criteria of the RIA tool are summarized in Table 1. An Excel-based format of the tool with critical and important signaling questions to the domains, is available online (https://doi.org/10.5281/zenodo.7024699).

3.1.1 | Domain 1: Retracted studies or studies with published expression of concern

1. The problem: Journal editors should retract a published study if they have clear evidence that the
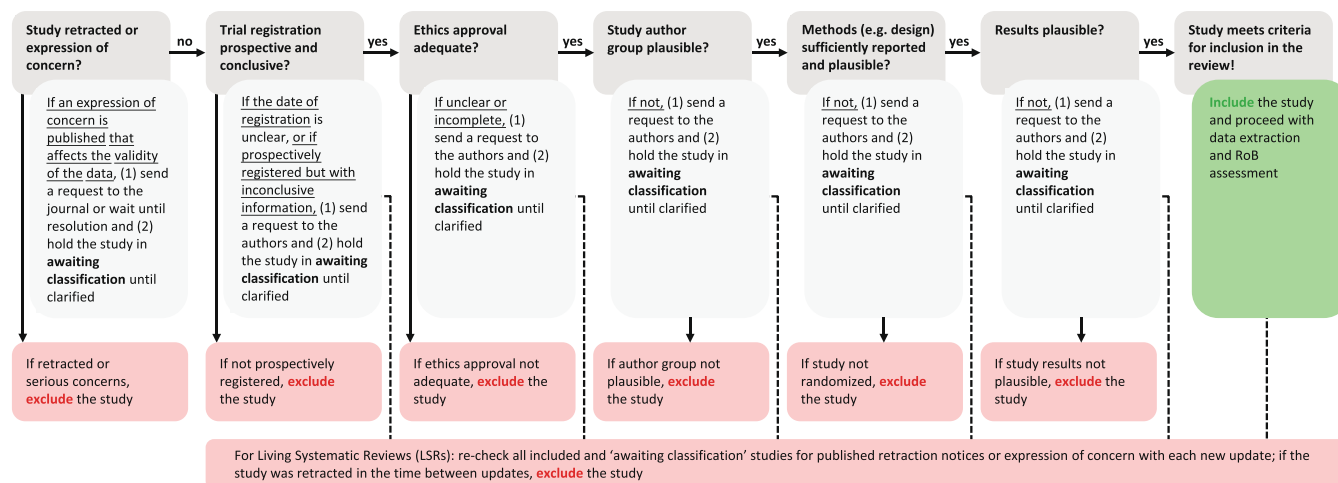
Study retracted or expression of concern?		Trial registration prospective and conclusive?		Ethics approval adequate?		Study author group plausible?		Methods (e.g. design) sufficiently reported and plausible?		Results plausible?		Study meets criteria for inclusion in the review!
	no →		yes →		yes →		yes →		yes →		yes →	
If an expression of concern is published that affects the validity of the data, (1) send a request to the journal or wait until resolution and (2) hold the study in **awaiting classification** until clarified		If the date of registration is unclear, or if prospectively registered but with inconclusive information, (1) send a request to the authors and (2) hold the study in **awaiting classification** until clarified		If unclear or incomplete, (1) send a request to the authors and (2) hold the study in **awaiting classification** until clarified		If not, (1) send a request to the authors and (2) hold the study in **awaiting classification** until clarified		If not, (1) send a request to the authors and (2) hold the study in **awaiting classification** until clarified		If not, (1) send a request to the authors and (2) hold the study in **awaiting classification** until clarified		**Include** the study and proceed with data extraction and RoB assessment
If retracted or serious concerns, **exclude** the study		If not prospectively registered, **exclude** the study		If ethics approval not adequate, **exclude** the study		If author group not plausible, **exclude** the study		If study not randomized, **exclude** the study		If study results not plausible, **exclude** the study		

For Living Systematic Reviews (LSRs): re-check all included and 'awaiting classification' studies for published retraction notices or expression of concern with each new update; if the study was retracted in the time between updates, **exclude** the study

FIGURE 1 Hierarchical work flow and decision tree of the research integrity assessment tool. Potentially eligible randomized controlled trials (RCTs) identified during screening should be assessed for research integrity hierarchically considering domains 1–6. Retraction, lack of prospective registration, lack of adequate ethical approval with informed written consent, inconsistencies in the author group and the location of the study, lack of proper randomization, and implausible study results should lead to exclusion of a RCT. Concerns with the RCT in any domain put the study in "awaiting classification" and should lead to further investigations. If no concerns appear through all domains or could be clarified, for example, in correspondence with study authors, the RCT meets criteria for inclusion in the review and can be processed further. In living systematic reviews, included RCTs and RCTs "awaiting classification" must be reassessed for retraction notices. [Colour figure can be viewed at wileyonlinelibrary.com]

findings are unreliable.[29] Consequently, retracted studies should not be included in evidence syntheses as they can distort the evidence base.[3]

2. Assessment: Cochrane has published detailed guidance on how to search for retraction notices and handle retracted studies in Cochrane reviews.[18,19] Retracted RCTs can be identified by review authors as such through a search for post-publication amendments in the systematic search for studies or on the Retraction Watch Database (https://retractiondatabase.org/RetractionSearch.aspx). Retracted RCTs should simply be excluded. An expression of concern can be published by a journal to raise awareness of a possible problem in a published study[29] and may also announce a full retraction. According to the Cochrane problematic studies guidance, review authors should take RCTs with an expression of concern depending on the nature of the concern and either exclude from the review, or put in awaiting classification category, or include in the review (e.g., for reasons that do not affect the validity of the data).[18]

3.1.2 | Domain 2: Prospective trial registration

1. The problem: The WHO declares that registration of all intervention trials is a scientific, ethical, and moral responsibility and expects that all clinical trials are prospectively registered in a WHO Registry Network approved registry.[30] The WHO regards trial registration as the publication of an internationally-agreed set of information about the design, conduct, and administration of clinical trials. These details should be published on a publicly accessible website managed by a registry conforming to WHO standards. Registries checking data as part of the registration process may lead to improvements in the quality of clinical trials by allowing identification of potential problems early in the research process.[30] One of the minimum standards set out for trial registries in the International Standards for Clinical Trial Registries (Item 2.3) is that registries must obtain written third-party confirmation of a trial's existence as part of the registration process.[31] Prospective trial registration can be a proxy for trial quality as investigators and authors of high quality trials know and follow these responsibilities, and prospectively registered studies have been shown to be at lower risk of bias.[32] The exclusion of non- and retrospectively registered RCTs in systematic reviews may shrink the study pool dramatically because there is poor compliance with trial registration[33,34] despite the fact that it may be legally required as defined in clinical trial regulations world-wide. On the other hand, this approach could bring us closer to the truth about the effectiveness of an intervention and gives us more confidence in the conclusions of our systematic reviews.[35]

TABLE 1 Critical and important criteria for a research integrity assessment of RCTs investigating IMPs for evidence syntheses.

	Domain 1: Retraction or expression of concern	Domain 2: Trial registration	Domain 3: Ethics approval	Domain 4: Author group	Domain 5: Methods	Domain 6: Results
Critical criteria	1. Retraction of the study	1. Registry number not reported 2. Not prospectively registered	1. Ethics committee approval not reported 2. Name and location of the ethics committee not reported 3. Ethics approval number not reported 4. Written informed consent not obtained	1. Inconsistency of authors' affiliations and countries the study is reported to have taken place in 2. Inconsistency in different parts of the article, for example, country and location of study conduct	1. Insufficient reporting of the study design (e.g., randomization) and unclear that the study was properly randomized	
Important criteria	1. Expression of concern published elsewhere	1. Inconsistency in details of study dates reported in the publication and in the registration documents	1. Ethics committee approval not obtained by a nationally recognized ethics committee as defined in the country's clinical trial regulations	1. Implausible number of authors for the study design (e.g., a single author article reporting a randomized control trial is unrealistic)	1. Number of participants in each group inconsistent with the reported randomization method (e.g., block randomization) 2. Baseline details not reported in sufficient detail to assess whether randomization worked properly 3. Incomplete or missing study flow diagram (optional)	1. Implausible number of patients with the condition recruited within the timeframe 2. Unrealistic response rate or loss of follow-up 3. Plagiarism 4. Excessive similarity or difference in the characteristics of the study participants between groups 5. Discrepancies between data reported in figures, tables, and text 6. Calculation errors 7. Implausible study results (e.g., massive risk reduction)

Abbreviations: IMP, investigational medicinal products; RCTs, randomized controlled trials.

2. Assessment: Whether or not a RCT has been prospectively registered can be proven in the trials register. The trial registry number should be reported in the study publication.[36] Prospective registration is defined as registration of a trial in a recognized national or international trials register before enrolment of the first participant.[31] Due to the increase in trial initiation early in the COVID-19 pandemic, a delay

between submission of a trial registration to and the actual publication on the register web site may have occurred. Extraordinary circumstances such as this pandemic, however, do not release investigators from submitting their registration prospectively or justify a total lack of registration. To avoid an unfair and unreasonable judgement, the registration's first submission date should be considered and deemed prospective if occurring before enrolment of the first study participant. In case of doubt, review authors should contact the authors for the submission date of the trial protocol and the RCT should be moved to the "awaiting classification" category. Review authors should also search for any inconsistencies in details of study dates (e.g., study start, start of recruitment, and registration date) reported in the publication and in the registration documents. In case of inconsistency, review authors should contact the investigators for clarification. Review authors should exclude non-registered and retrospectively registered RCTs.

There is no empirical evidence that inclusion of only prospectively registered RCTs in a review guarantees the inclusion of trustworthy studies only,[18] therefore, additional study criteria are critical and considered in the following.

3.1.3 | Domain 3: Adequate ethics approval

1. The problem: Adherence to ethical principles in clinical studies is compulsory to protect the dignity, rights, and welfare of research participants.[27] As such, all clinical trials involving human beings have to be reviewed by an ethics committee to ensure that the appropriate ethical standards are being upheld.[27] It is further good clinical practice that the study investigators obtain written informed consent from all participants before randomization.[37] Details on ethics approval and written informed consent are included in the WHO trial registration guidelines.[31]
2. Assessment: Review authors should check the trials registry record, the study protocol and the published study report for a published statement on whether approval from an ethics committee was granted. This statement should include the name of the ethics committee granting the approval, and an approval number.[18] A statement on whether written informed consent was obtained, or a justification for its absence, should also be included in the study records. Lack of such statements in the article does not necessarily mean that a study did not have ethics approval or did not obtain participant's written informed consent.

Therefore, if a study did not report the name of the ethics committee and/or the approval number and information on consent, review authors should send a request to the authors and the RCT should be moved to the pool of studies awaiting classification until clarified. If the authors cannot provide any component of the above, the RCT should be excluded. Moreover, it should be assured that a nationally recognized ethics committee as defined in the country's clinical trial regulations gave the approval. The ethics committee can be searched for on the WHO List of National Ethics Committees (https://apps.who.int/ethics/nationalcommittees/nec.aspx) or searching the national responsible authority's list of recognized ethics committees. Unfortunately, there is no international list of responsible authorities available and an individual web search might be necessary. An English language summary of the specific regulations for some countries can be found on the NIH Clinical Trials Regulation website (https://clinregs.niaid.nih.gov/).

3.1.4 | Domain 4: Author group

1. The problem: Authorship confers credit and has important academic, social, and financial implications, but authorship also implies responsibility and accountability for published work.[38] When authorship is abused, accountability and responsibility can be questionable and the potential for manipulated analysis and conclusions may increase. The domain "author group" is a proxy for whether it is plausible that the study has actually taken place with a focus on authors' details, number of authors, and the location where the study was conducted. There are certainly many more details regarding authorship that may be important to address, for example, author contributorship, ghost authorship, and funding details, which are currently not part of the RIA tool. However, when applying our tool, other review authors are welcome to examine studies in more extensive detail in this domain. The REAPPRAISED list examines research governance, authorship, and research conduct, and details (unfortunately without further guidance) can be found there.[20]
2. Assessment: Review authors should focus on the authors' details and the location where the study was conducted. Review authors should check whether an article is authored primarily by individuals with affiliations different from the country(ies) where the study was conducted without sufficient explanation, though information about trial sponsorship, ethics committee approval, funding, and regulatory oversight included in trial registry details can assist in evaluating such

cases. Moreover, inconsistencies in the article regarding different countries specified in different parts of the article or as compared to the trial registry entry should also flag a study as "potentially problematic." Finally, the review authors should check whether the number of authors is plausible for the study design. In its most extreme form, one single author article reported an RCT, and may indicate a fabricated study, since it is impossible for one person to have managed such a complex study design alone.[18] Li et al. suggested a low author-to-study size ratio or a number of authors less than three should trigger concern.[39] If there are concerns regarding a study's author group, a request should be sent to the authors and the RCT should be moved to the pool of studies awaiting classification until clarified. If the authors cannot justify any component or inconsistency of the above, the RCT can be excluded. However, exclusion of the RCT should be weighed against judgments from the other domains.

3.1.5 | Domain 5: Sufficient reporting and plausibility of methods (e.g., randomization methods)

1. The problem: RCTs that report very sparsely on their methods (e.g., study design/randomization) immediately raise alarms, particularly when study authors provide insufficient information to be able to make an adequate assessment of the risk of bias using the Cochrane Risk of Bias (RoB) tools 1 or 2.0 in evidence synthesis.[40] Cochrane reviews include an assessment of the risk of bias for each included study. The RoB tools 1 or 2.0 are structured into a fixed set of domains of bias, focusing on different aspects of trial design, conduct, and outcome reporting.[40] With incomplete methods the RoB assessment frequently returns to "unclear" or "some concerns" of bias which is actually misleading. Conversely, study authors who applied deficient study methods or introduced bias into their studies could avoid a poor risk of bias rating simply by underreporting. Moreover, it is important to check that the study was properly randomized. Identical numbers allocated to each group in the absence of a block approach to randomization are considered as cause for concern.[22]

2. Assessment: The most critical criteria for domain 5 is a sufficient reporting on study design methods (e.g. randomization) in the study report to consider that the study was properly randomized. The method used for the randomization must be described and the process must lead to a random allocation of the participants. The sole designation "randomized study" is not sufficient. Identical numbers of participants allocated to each group without use of a block randomization was considered as cause for concern. Baseline details must also be provided in sufficient detail to estimate whether the randomization worked. If it turns out that a trial declared as "randomized" in the article was not properly randomized, the review authors should exclude the study. Considering details on the study flow diagram such as the number of participants being randomized, receiving the intervention, and being analyzed, is optional for the RIA, as those aspects will be covered by the Cochrane RoB tool 1 and 2.0 (and missing information may be adjudicated for with a high risk of bias judgement). Therefore, with domain 5 it should be assured that only RCTs with sufficient reporting of their methods and proper randomization are to be included in the study pool of the review, and studies with insufficient details are held in awaiting classification until the authors provide further details upon request.

3.1.6 | Domain 6: Plausible results

1. The problem: There are alarming numbers of fabricated and false data or trials published each year[2,41] and this has major consequences for the entirety of the health research ecosystem and for naïve systematic reviewers who assume that the published studies are real. How can systematic reviewers identify and deal with false or fabricated data and trials when looking at published articles? In-depth checks of baseline details and individual participant data (IPD) are time consuming and require specific statistical training—and both may be limited for most systematic reviewers. Therefore, we achieved expert consensus on a few criteria, which are used already in existing screening tools,[20,21,42] to be warnings for implausibility of reported results in RCTs.

2. Assessment: Review authors should assess for plausibility (1) the number of patients recruited within the timeframe with the condition; (2) the response rate or number of participants lost to follow-up; (3) excessive similarity or difference in the characteristics of the study participants between groups; and (4) results that could be implausible (e.g., massive risk reduction, unexpected outlier data, and unusual frequency of a rare outcome). Furthermore, review authors should note any data error (e.g., number of participants or events that did not add up), calculation error, and discrepancies between data reported in figures, tables, and text. In addition, when multiple reports of the

RCT are available, review authors should check for overlap in text and data between published articles by the same or different authors without explanation.

When working through those criteria, and inconsistencies or implausible data are identified, the review authors should send an information request to the study authors to allow for comments and clarification. Until resolution, the RCT should be held in awaiting classification. This domain should be handled with care as for the above mentioned complexity and should not lead to exclusion of a trial without clear evidence.

3.2 | When should the RIA be used during evidence synthesis?

It is important to assess RCTs that pass the PIC(O) (participants, intervention, comparator, and [outcomes]) eligibility screening as early as possible for research integrity. Exclusion of problematic RCTs cleans the whole study pool, not only the estimated effects in meta-analyses and conclusions thereof, but also qualitative analyses, summaries of baseline characteristics, and conclusions regarding evidence gaps. Therefore, early exclusion of problematic RCTs or movement to the awaiting classification category of inconclusive RCTs is preferred.

3.3 | How should the RIA be used during evidence synthesis?

In the hierarchical workflow of RIA through domains 1–6, three decisions on a study's eligibility are possible at any steps from 1 to 6: RCTs may be either included, excluded, or moved to awaiting classification (Figure 1). Whenever it is concluded from the decision on one domain that a RCT has to be excluded, the following domains no longer apply, can be omitted and do not have to be answered. The first three domains, for example, retraction, prospective registration, and ethics approval, can frequently be answered with certainty and direct decisions without sending an author request (e.g., for retraction, retrospective registration, and verbal informed consent). Retraction, lack of prospective registration, lack of adequate ethics approval with informed written consent, should lead to exclusion of a RCT. For domains 4–6, a definite decision on a study's research integrity is much more difficult to find and correspondence with the study authors can be necessary. Decisions can be subjective. For performance of the RIA, it has to be kept in mind that RCTs should be included only if there is no obvious doubt regarding any of the critical criteria included in the tool. However, inconsistencies in the author group and the location of the study, lack of proper randomization, or implausible study results are sometimes difficult to answer with certainty. We suggest to be careful in excluding a RCT for a single domain concern regarding domains 4–6, especially if there are no other concerns raised when using this tool—it should apply in dubio pro reo. If there is any inconsistency, insufficient information, or serious concerns, requests should be sent to the authors. Each study authors must have the opportunity to respond and clarify the questions. When authors do not respond or questions remain unanswered, the RCT should be held in the awaiting classification category when the evidence synthesis is published. Systematic reviews in a living mode have to re-evaluate all included RCTs and RCTs held in awaiting classification for published retraction notices or expression of concern for each update.

For the RIA, similar to a standard systematic review approach, all documents of a study should be used for the assessment (e.g., all reports of a study including full text publication, registration, and protocol).

The Excel version of the tool is based on critical and important signaling questions to the domains and includes columns to summarize a conclusion for each domain and an overall conclusion, which justifies the decision on research integrity and eligibility (https://doi.org/10.5281/zenodo.7024699). The tool offers a transparent way to document what the review authors have done (e.g., correspondence with the trial authors) and their judgements. For transparency, the table should be published as a supplement to the systematic review or deposited in an online repository with permanent digital object identifier (doi). The consequence of the RIA on the study pool should also be documented in the PRISMA flow diagram of the review using a new reason for exclusion "failed research integrity assessment."

3.4 | Who should apply RIA?

Each study should be independently assessed by two review authors and discrepancies should be resolved by discussion. It is important that the review authors team consists of researchers who have knowledge in clinical trial design and regulation, systematic review methodology, and clinical content expertise. Checks for plausibility of results (domain 6) require some knowledge of the topic area of the RCT.

3.5 | Working example: Cochrane ivermectin review

For the Cochrane ivermectin review, we evaluated 25 RCTs in our first review update[23]: 14 included studies

TABLE 2 Research integrity assessment of RCTs applied to the Cochrane ivermectin review update.

	Domain 1: Retraction or expression of concern	Domain 2: Trial registration	Domain 3: Ethics approval	Domain 4: Author group	Domain 5: Methods	Domain 6: Results
Number of studies evaluated	25	24 (−1)	15 (−9)	15 (−0)	15 (−0)	14 (−1)
Result	1 retracted study (awaiting classification 1st review); identified via retraction watch database and post-publication amendment	9 studies without prospective trial registration (6 included and 1 awaiting classification 1st review and 2 updated search); 4 non- and 5 retrospectively registered	1 study posted results on a trials register without reporting ethics approval and number (awaiting classification 1st review); all studies with written informed consent	2 studies posted results on a trials register and could not be assessed (1 awaiting classification 1st review, 1 updated search)	1 non-randomized study (included 1st review); 3 study authors did not report/clarify sufficient detail on randomization (1 awaiting classification 1st review and 2 updated search)	3 study authors did not clarify questions on study results (1 awaiting classification 1st review and 2 updated search)
Decision/ Action	1. 1 study excluded 2. 24 studies without concerns	1. 9 studies excluded 2. 15 studies without concerns	1. no study excluded 2. 1 study awaiting classification 3. 14 studies without concerns	1. no study excluded 2. 2 (+1) study awaiting classification 3. 13 studies without concerns	1. 1 study excluded 2. 3 (+1) study awaiting classification 3. 11 studies without concerns	1. no study excluded 2. 3 (+0) study awaiting classification 3. 11 studies without concerns

Note: A total of 25 RCTs were evaluated (14 included studies and 3 studies with results awaiting classification from the original review version, and 8 studies with results identified by the updated search). The assessment took place hierarchically from domains 1 to 6 for all studies. Whenever it is concluded from the decision on one domain that a study has to be excluded, the following domains were not answered.

Abbreviation: RCTs, randomized controlled trials.

and 3 studies with results awaiting classification from the original review version[4] were re-evaluated; and 8 studies with results identified by the updated search were evaluated. In the review update, we excluded 11 out of 25 assessed studies, moved 3 to awaiting classification, and included 11 studies meeting all criteria for inclusion into the review update (Table 2, Supplementary information S1). The most frequent reason for exclusion was a lack of prospective trial registration in 9 studies; 1 study was retracted and another turned out not to be a randomized trial. The RIA excel sheet was published as a supplement to the updated Cochrane review and deposited in an online repository with permanent digital object identifier.[43] Changes to the RIA tool according to peer review of this method paper which appeared after the Cochrane review had been published are listed in Supplementary information S2.

4 | DISCUSSION

Whilst we know that problematic studies create challenges, the number of problematic studies and their influence has not been clearly quantified.[3,44] The example of studies included in, then later excluded from, the Cochrane ivermectin review demonstrates the need to manage problematic studies. Every study included in a systematic review has an influence on the results and conclusions of the review and this has far-reaching consequences for the population treated on the basis of the review.[45] As the number of systematic reviews being part of or being considered in evidence ecosystems continues to grow, we believe ensuring research integrity of studies included in the study pool of a systematic review is an approach that provides evidence closer to the unbiased truth and respects human rights. We hope the RIA tool, consisting of six domains to assess the research integrity of RCTs included in systematic reviews, will be viewed as a new transparent option to include the concept of research integrity in evidence synthesis as part of the eligibility screening and serves as a platform for urgent developments in this direction.

We believe the RIA tool represents a considerable expansion over previous efforts. The strength of our tool is that it was developed in accordance with recent Cochrane guidance on identifying problematic trials[17,18]

using an international team of (Cochrane) editors and content experts in systematic reviews methodology, clinical trial regulations, and evidence ecosystems. Using this tool for evaluation of research integrity of potentially eligible RCTs should be performed as part of the review's eligibility screening. Documentation of reviewers' decisions on in- and exclusion will be transparent for evidence users increasing systematic reviewers' accountability. We view that we need to implement an integrity assessment process quickly, but see this tool as the first pilot version that can and should be adapted, modified, tested, and validated over time. We have already launched a meta-epidemiological study to assess the impact of RIA on a larger pool of interventional RCTs included in systematic reviews on COVID-19.[46]

This RIA tool is not a "Risk of Bias" assessment. Critical aspects of research integrity assessed with the RIA are not considered in established RoB assessments such as the Cochrane RoB 1 or 2.0 tool. RIA aims to establish the integrity and authenticity of the studies. Cochrane's research integrity department points out that the Cochrane RoB tool operates on the basis that the data are true.[47] Research integrity of the trials is acknowledged as being a separate issue that should be handled prior to RoB assessment.

There may be several limitations to our approach. The example presented here describes the development of the RIA tool used in the Cochrane review ivermectin for COVID-19. Therefore, this tool was used in one case study and has not yet been validated. In addition, using this tool on trials in standard evidence synthesis work, where data are used from trials conducted in any country at any point in time, will not be as straightforward as in this example. Depending on the clinical question, in some cases, the majority of trials may have been conducted before trial registries even came into existence. Using this tool on trials is challenging in these circumstances, as there may not be any public records of trial governance, and alternative approaches may need to be developed, for example, including older studies in secondary analysis only. A larger group of experts with different backgrounds and a Delphi panel process could have led to a different final set of critical domains. Future application of this tool may lead to refinement or adjustment of some domains, for example, for different scenarios with intervention studies which are not IMPs. Another area in need of development for the RIA tool is the number of key criteria that may be needed to warrant study exclusion. Identification and handling of false and fabricated data or trials by in-depth checks of baseline details and IPD requires specific statistical training—and both may be limited for most systematic reviewers. Statistical experts for IPD analysis and

detection of fabricated data may not be well represented by our sample of experts. Therefore, the checking domain for plausibility of study results may still be under development and may change when new methods for identification of false and fabricated data become available. Research in this area is still in its infancy. Bordewijk et al. recently published a scoping review on methods to assess research misconduct in health research and concluded that tools to investigate are rudimentary and labor-intensive, and automatic tools and routine validation of these methods is needed.[42] We plan to apply the RIA tool to a larger sample of systematic reviews to verify performance. With these findings, we plan to further develop the tool within the next 3 years in a Delphi process with a subsequent update.

We believe that this tool and the ivermectin review example has important teaching capabilities to raise awareness of the concept of research integrity for systematic reviewers and clinical study investigators. Clinical trial regulations are designed to ensure patient safety and must be adhered to. It is unfortunate that research was identified in the context of the ivermectin review that does not meet the required standard, and the only way forward is to improve access to Good Clinical Practice training for everyone involved in clinical trials, health research, and evidence synthesis worldwide, with the aim of achieving full compliance with the regulations. We hope that research integrity criteria will lead to a rethinking so that the results of lawfully conducted studies are given prominence rather than treating all data equally in evidence synthesis. In addition, we hope that ensuring research integrity of studies included in the study pool of a systematic review is an approach that helps to provide evidence closer to the unbiased truth and improves respect of human rights in evidence synthesis. It is a moral duty that problematic studies are not included in systematic reviews used to inform guidelines or as information for health professionals or the public.

AUTHOR CONTRIBUTIONS

The research integrity assessment tool was developed prior to the preparation of the first update of the Cochrane review "Ivermectin for preventing and treating COVID-19" and addresses specific questions of how to identify and deal with problematic studies in the context of this systematic review. All authors of this article drafted the tool. Stephanie Weibel, Maria Popp, and Stefanie Reis applied the tool to studies eligible for the Cochrane review. Stephanie Weibel wrote the first version of the manuscript and all authors edited. Stephanie Weibel is the guarantor of the article.

CONFLICT OF INTEREST

The authors declare no conflicts of interest. The tool was developed by a group of content experts who were either Cochrane review authors (Stephanie Weibel, Maria Popp, Stefanie Reis, and Nicole Skoetz) and/or otherwise involved in Cochrane (Stephanie Weibel, Nicole Skoetz, Paul Garner, and Emma Sydenham). Stephanie Weibel is Content Editor of Cochrane Anaesthesia. Nicole Skoetz is joint Co-ordinating Editor of Cochrane Haematology. Paul Garner is Co-ordinating Editor of the Cochrane Infectious Diseases Group and responsible for assuring quality of reviews. Emma Sydenham is Co-ordinating Editor of the Cochrane Injuries Group and was a member of Cochrane's Scientific Misconduct Policy Advisory Group.

DATA AVAILABILITY STATEMENT

All data relevant to the study are included in the article, uploaded as supporting information, or available in the Cochrane ivermectin review with link to a public, open access repository.

ORCID

Stephanie Weibel https://orcid.org/0000-0003-1900-2535

REFERENCES

1. Soares-Weiser K, Lasserson T, Jorgensen KJ, et al. Policy makers must act on incomplete evidence in responding to COVID-19. *Cochrane Database Syst Rev*. 2020;(11):ED000149. doi:10.1002/14651858.ED000149

2. Ioannidis JPA. Hundreds of thousands of zombie randomised trials circulate among us. *Anaesthesia*. 2021;76(4):444-447. doi:10.1111/anae.15297

3. Avenell A, Stewart F, Grey A, Gamble G, Bolland M. An investigation into the impact and implications of published papers from retracted research: systematic search of affected literature. *BMJ Open*. 2019;9(10):e031909. doi:10.1136/bmjopen-2019-031909

4. Popp M, Stegemann M, Metzendorf MI, et al. Ivermectin for preventing and treating COVID-19. *Cochrane Database Syst Rev*. 2021;(7):CD015017. doi:10.1002/14651858.CD015017.pub2

5. Samaha AA, Mouawia H, Fawaz M, et al. Effects of a single dose of ivermectin on viral and clinical outcomes in asymptomatic SARS-CoV-2 infected subjects: a pilot clinical trial in Lebanon. *Viruses*. 2021;13(6):989. doi:10.3390/v13060989

6. Samaha AA, Mouawia H, Fawaz M, et al. Effects of a single dose of ivermectin on viral and clinical outcomes in asymptomatic SARS-CoV-2 infected subjects: a pilot clinical trial in Lebanon. *Viruses*. 2021;13(11):2154. doi:10.3390/v13112154

7. Elgazzar A, Eltaweel A, Youssef SA, Hany B, Hafez M, Moussa H. Efficacy and safety of ivermectin for treatment and prophylaxis of COVID-19 pandemic (preprint). Accessed April 5, 2021. https://www.researchsquare.com/article/rs-100956/v3.

8. Lawrence J. Why was a major study on ivermectin for COVID-19 just retracted? Accessed August 6, 2021. https://grftr.news/why-was-a-major-study-on-ivermectin-for-covid-19-just-retracted/.

9. Bryant A, Lawrie TA, Dowswell T, et al. Ivermectin for prevention and treatment of COVID-19 infection: a systematic review, meta-analysis, and trial sequential analysis to inform clinical guidelines. *Am J Ther*. 2021;28(4):e434-e460. doi:10.1097/MJT.0000000000001402

10. Izcovich A, Peiris S, Ragusa M, et al. Bias as a source of inconsistency in ivermectin trials for COVID-19: a systematic review. ivermectin's suggested benefits are mainly based on potentially biased results. *J Clin Epidemiol*. 2021;144:43-55. doi:10.1016/j.jclinepi.2021.12.018

11. Kory P, Meduri GU, Varon J, Iglesias J, Marik PE. Review of the emerging evidence demonstrating the efficacy of ivermectin in the prophylaxis and treatment of COVID-19. *Am J Ther*. 2021;28(3):e299-e318. doi:10.1097/MJT.0000000000001377

12. Hill A, Garratt A, Levi J, et al. Meta-analysis of randomized trials of ivermectin to treat SARS-CoV-2 infection. *Open Forum Infect Dis*. 2021;8(11):ofab358. doi:10.1093/ofid/ofab358

13. Hill A, Garratt A, Levi J, et al. Meta-analysis of randomized trials of ivermectin to treat SARS-CoV-2 infection. *Open Forum Infect Dis*. 2022;9(3):ofac056. doi:10.1093/ofid/ofac056

14. Hill A, Mirchandani M, Pilkington V. Ivermectin for COVID-19: addressing potential bias and medical fraud. *Open Forum Infect Dis*. 2022;9(2):ofab645. doi:10.1093/ofid/ofab645

15. Bryant A, Lawrie TA, Fordham EJ. Ivermectin for prevention and treatment of COVID-19 infection: a systematic review, meta-analysis, and trial sequential analysis to inform clinical guidelines. *Am J Ther*. 2021;28(5):e573-e576. doi:10.1097/MJT.0000000000001442

16. Parker L, Boughton S, Lawrence R, Bero L. Somebody might actually make up data': an empirical qualitative study on how to screen for fake research. SSRN 2022; 10.2139/ssrn.4071671.

17. Cochrane Library. Cochrane policy for managing potentially problematic studies. Accessed October 11, 2021. https://www.cochranelibrary.com/cdsr/editorial-policies#problematic-studies.

18. Cochrane Library. Cochrane policy for managing potentially problematic studies: implementation guidance. Accessed October 11, 2021. https://documentation.cochrane.org/display/EPPR/Policy+for+managing+potentially+problematic+studies%3A+implementation+guidance.

19. Lefebvre C, Glanville J, Briscoe S, et al. Technical supplement to chapter 4: Searching for and selecting studies. In: JPT H, Thomas J, Chandler J, et al., eds. *Cochrane Handbook for Systematic Reviews of Interventions. Version 6.3.* Cochrane; 2022 https://training.cochrane.org/technical-supplement-chapter-4-searching-and-selecting-studies-v63

20. Grey A, Bolland MJ, Avenell A, Klein AA, Gunsalus CK. Check for publication integrity before misconduct. *Nature.* 2020; 577(7789):167-169. doi:10.1038/d41586-019-03959-6

21. Cochrane. Cochrane Pregnancy and Childbirth Group data extraction form_v_2.2_-_13_july_2021_1. Accessed October 11, 2021. https://pregnancy.cochrane.org/author-resources-new-reviews.

22. O'Connell NE, Moore RA, Stewart G, et al. Investigating the veracity of a sample of divergent published trial data in spinal pain. *Pain.* 2022. doi:10.1097/j.pain.0000000000002659. Epub ahead of print.

23. Popp M, Reis S, Schießer S, et al. Ivermectin for preventing and treating COVID-19. *Cochrane Database Syst Rev.* 2022;(6): C015017. doi:10.1002/14651858.CD015017.pub3

24. European Commision. Definition of investigational medicinal products (IMPs), definition of non investigational medicinal products (NIMPs). Accessed October 11, 2021. https://www.erasme.ulb.ac.be/sites/default/files/files/documents/2016/draft_guidance_definition_imp_2006_07_27_2.pdf.

25. Investigational Medicinal Product Dossiers (IMPD). Investigational medicinal product dossiers. Accessed October 11, 2021. https://www.imp-dossier.eu/.

26. Medicines and Healthcare Products Regulatory Agency. Medicines and healthcare products regulatory agency: is it a clinical trial of a medicinal product? Accessed October 11, 2021. https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/949145/Algorithm_Clean__1_.pdf.

27. WHO. WHO - ensuring ethical standards and procedures for research with human beings. Accessed October 14, 2021. https://www.who.int/activities/ensuring-ethical-standards-and-procedures-for-research-with-human-beings.

28. Lefebvre CGJ, Briscoe S, Featherstone R, et al. Chapter 4: Searching for and selecting studies. In: Higgins JPT, Chandler J, Cumpston M, Li T, Page MJ, Welch VA, eds. *Cochrane Handbook for Systematic Reviews of Interventions* Version 63 (updated February 2022). Cochrane; 2022.

29. COPE. Retraction guidelines - English. Version 2: November 2019. Accessed August 6, 2021. https://publicationethics.org/node/19896.

30. International Clinical Trials Registry Platform (ICTRP). International Clinical Trials Registry Platform (ICTRP). Accessed august 6, 2021. https://www.who.int/clinical-trials-registry-platform.

31. WHO. *International Standards for Clinical Trial Registries.* 2nd ed. World Health Organization; 2021 http://apps.who.int/iris/bitstream/handle/10665/274994/9789241514743-eng.pdf

32. Lindsley K, Fusco N, Li T, Scholten R, Hooft L. Clinical trial registration was associated with lower risk of bias compared with non-registered trials among trials included in systematic reviews. *J Clin Epidemiol.* 2022;145:164-173. doi:10.1016/j.jclinepi.2022.01.012

33. Lindsley K, Fusco N, Teeuw H, Mooij E, Scholten R, Hooft L. Poor compliance of clinical trial registration among trials included in systematic reviews: a cohort study. *J Clin Epidemiol.* 2021;132:79-87. doi:10.1016/j.jclinepi.2020.12.016

34. Riemer M, Kranke P, Helf A, et al. Trial registration and selective outcome reporting in 585 clinical trials investigating drugs for prevention of postoperative nausea and vomiting. *BMC Anesthesiol.* 2021;21(1):249. doi:10.1186/s12871-021-01464-w

35. Harvey LA, Dijkers MP. Should trials that are highly vulnerable to bias be excluded from systematic reviews? *Spinal Cord.* 2019;57(9):715-716. doi:10.1038/s41393-019-0340-y

36. Schulz KF, Altman DG, Moher D, Group C. CONSORT 2010 statement: updated guidelines for reporting parallel group randomised trials. *BMJ.* 2010;340:c332. doi:10.1136/bmj.c332

37. Good Clinical Practice Network. Informed consent of trial subjects. Accessed October 14, 2021. https://ichgcp.net/publications/informed-consent-of-trial-subjects.

38. ICMJE. Defining the role of authors and contributors. Accessed October 12, 2021. https://www.icmje.org/recommendations/browse/roles-and-responsibilities/defining-the-role-of-authors-and-contributors.html.

39. Li W, Bordewijk EM, Mol BW. Assessing research misconduct in randomized controlled trials. *Obstet Gynecol.* 2021;138(3): 338-347. doi:10.1097/aog.0000000000004513

40. Higgins JPT, Page MJ, Elbers RG, Sterne JAC. Chapter 8: Assessing risk of bias in a randomized trial. In: Higgins JPT, Chandler J, Cumpston M, Li T, Page MJ, Welch VA, eds. *Cochrane Handbook for Systematic Reviews of Interventions* Version 62 (updated February 2021). Cochrane; 2021.

41. Carlisle JB. False individual patient data and zombie randomised controlled trials submitted to Anaesthesia. *Anaesthesia.* 2021;76(4):472-479. doi:10.1111/anae.15263

42. Bordewijk EM, Li W, van Eekelen R, et al. Methods to assess research misconduct in health-related research: a scoping review. *J Clin Epidemiol.* 2021;136:189-202. doi:10.1016/j.jclinepi.2021.05.012

43. Weibel S, Popp M, Reis S. Supplementary File_Ivermectin_Research Integrity Assessment (Version 1). 10.5281/zenodo.6630618.

44. Fanelli D, Wong J, Moher D. What difference might retractions make? An estimate of the potential epistemic cost of retractions on meta-analyses. *Account Res.* 2021;14:1-18. doi:10.1080/08989621.2021.1947810

45. The Guardian. Huge study supporting ivermectin as Covid treatment withdrawn over ethical concerns. The Guardian. Accessed July 15, 2021. https://www.theguardian.com/science/2021/jul/16/huge-study-supporting-ivermectin-as-covid-treatment-withdrawn-over-ethical-concerns.

46. Weibel S, Bora A-M, Meybohm P, et al. Impact of a research integrity assessment (RIA) of randomized controlled trials on interventional COVID-19 systematic reviews: protocol for a meta-epidemiological study. *OSF.* 2022. doi:10.17605/OSF.IO/NBRHX

47. Boughton SL, Wilkinson J, Bero L. When beauty is but skin deep: dealing with problematic studies in systematic reviews. *Cochrane Database Syst Rev.* 2021;6:ED000152. doi:10.1002/14651858.ED000152

SUPPORTING INFORMATION

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