

Assessing the Integrity of Clinical Trials Included in Evidence Syntheses

Núñez-Núñez, María; Cano-Ibáñez, Naomi; Zamora, Javier; Bueno-Cavanillas, Aurora; Khan, Khalid Saeed; Oren, Eyal; Rapone, Biagio

DOI:

[10.3390/ijerph20126138](https://doi.org/10.3390/ijerph20126138)

License:

Creative Commons: Attribution (CC BY)

Document Version

Publisher's PDF, also known as Version of record

Citation for published version (Harvard):

Núñez-Núñez, M, Cano-Ibáñez, N, Zamora, J, Bueno-Cavanillas, A, Khan, KS, Oren, E (ed.) & Rapone, B (ed.) 2023, 'Assessing the Integrity of Clinical Trials Included in Evidence Syntheses', *International Journal of Environmental Research and Public Health*, vol. 20, no. 12, 6138. <https://doi.org/10.3390/ijerph20126138>

[Link to publication on Research at Birmingham portal](#)

General rights

Unless a licence is specified above, all rights (including copyright and moral rights) in this document are retained by the authors and/or the copyright holders. The express permission of the copyright holder must be obtained for any use of this material other than for purposes permitted by law.

- Users may freely distribute the URL that is used to identify this publication.
- Users may download and/or print one copy of the publication from the University of Birmingham research portal for the purpose of private study or non-commercial research.
- User may use extracts from the document in line with the concept of 'fair dealing' under the Copyright, Designs and Patents Act 1988 (?)
- Users may not further distribute the material nor use it for the purposes of commercial gain.

Where a licence is displayed above, please note the terms and conditions of the licence govern your use of this document.

When citing, please reference the published version.

Take down policy

While the University of Birmingham exercises care and attention in making items available there are rare occasions when an item has been uploaded in error or has been deemed to be commercially or otherwise sensitive.

If you believe that this is the case for this document, please contact UBIRA@lists.bham.ac.uk providing details and we will remove access to the work immediately and investigate.



Review

Assessing the Integrity of Clinical Trials Included in Evidence Syntheses

María Núñez-Núñez ^{1,2,3,*}, Naomi Cano-Ibáñez ^{2,3,4} , Javier Zamora ^{3,5,6}, Aurora Bueno-Cavanillas ^{2,3,4}
and Khalid Saeed Khan ^{3,4}

¹ Pharmacy Department, Clínico San Cecilio University Hospital, 18016 Granada, Spain

² Biosanitary Research Institute (Ibs. Granada), 18012 Granada, Spain

³ Consortium for Biomedical Research in Epidemiology and Public Health (CIBERESP), 28029 Madrid, Spain

⁴ Department of Preventive Medicine and Public Health, University of Granada, 18016 Granada, Spain

⁵ Department of Biostatistics, Ramón y Cajal University Hospital (IRYCIS), 28034 Madrid, Spain

⁶ Institute of Metabolism and Systems Research, University of Birmingham, Birmingham B15 2TT, UK

* Correspondence: mnunez.pharm@gmail.com

Abstract: Evidence syntheses of randomized clinical trials (RCTs) offer the highest level of scientific evidence for informing clinical practice and policy. The value of evidence synthesis itself depends on the trustworthiness of the included RCTs. The rising number of retractions and expressions of concern about the authenticity of RCTs has raised awareness about the existence of problematic studies, sometimes called “zombie” trials. Research integrity, i.e., adherence to ethical and professional standards, is a multi-dimensional concept that is incompletely evaluated for the RCTs included in current evidence syntheses. Systematic reviewers tend to rely on the editorial and peer-review system established by journals as custodians of integrity of the RCTs they synthesize. It is now well established that falsified and fabricated RCTs are slipping through. Thus, RCT integrity assessment becomes a necessary step in systematic reviews going forward, in particular because RCTs with data-related integrity concerns remain available for use in evidence syntheses. There is a need for validated tools for systematic reviewers to proactively deploy in the assessment of integrity deviations without having to wait for RCTs to be retracted by journals or expressions of concern issued. This article analyzes the issues and challenges in conducting evidence syntheses where the literature contains RCTs with possible integrity deficits. The way forward in the form of formal RCT integrity assessments in systematic reviews is proposed, and implications of this new initiative are discussed. Future directions include emphasizing ethical and professional standards, providing tailored integrity-specific training, and creating systems to promote research integrity, as improvements in RCT integrity will benefit evidence syntheses.

Keywords: research integrity; evidence synthesis; systematic reviews; clinical trials



Citation: Núñez-Núñez, M.; Cano-Ibáñez, N.; Zamora, J.; Bueno-Cavanillas, A.; Khan, K.S. Assessing the Integrity of Clinical Trials Included in Evidence Syntheses. *Int. J. Environ. Res. Public Health* **2023**, *20*, 6138. <https://doi.org/10.3390/ijerph20126138>

Academic Editors: Eyal Oren and Biagio Rapone

Received: 9 February 2023

Revised: 5 June 2023

Accepted: 12 June 2023

Published: 15 June 2023



Copyright: © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Evidence syntheses of randomized clinical trials (RCTs) offer the highest level of effectiveness evidence validity for informing clinical practice and policy [1]. They make the results of RCTs available to practitioners and allow them to reach patients through systematic reviews and clinical practice guidelines. The rising number of allegations of data fabrication and falsification in retractions and expressions of concern about questionable research practices or faulty research methodology [2–5] has raised awareness about the authenticity of RCTs. (Un)intentional errors may lead to the existence of problematic studies, sometimes called “zombie” trials [6], within the literature. It has recently been recognized that: “*Even though the process for the detection and correction of error and fraud might be fairly well established and “standardized”, such as in COPE or ICMJE guidelines, inter-journal and inter-publisher variability, including editorial responsibilities, will continue to limit the effective correction of erroneous and fraudulent literature globally*” [7]. This background has important

implications for evidence syntheses, including both systematic reviews and guidelines, that have the potential to widen the impact of faulty RCTs. This commentary describes challenges in evidence syntheses, with respect to included RCT integrity, with particular focus on data fabrication and falsification.

2. How Can We Trust Evidence Syntheses?

Evidence synthesis is a type of research method that collates all relevant studies and interprets their collective findings. There has been sustained growth in systematic review publications during the last decade (Figure 1).

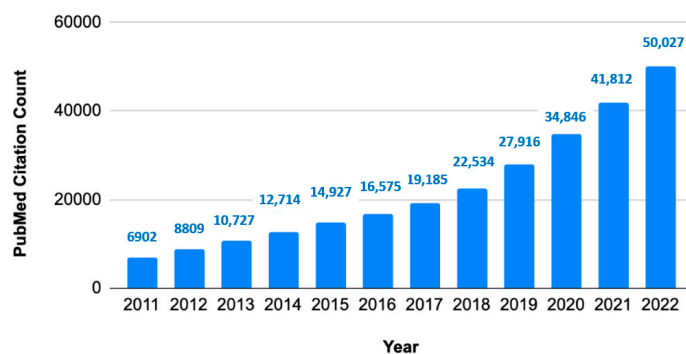


Figure 1. Growth of systematic reviews to synthesize evidence. Annual citation counts of article type systematic reviews in PubMed database.

The fundamental steps of evidence synthesis are as follows: defining the question(s); searching for relevant studies (screening and selection against defined inclusion and exclusion criteria); appraising the quality of the studies included and extracting relevant data; collating data, undertaking meta-analyses where appropriate; and interpreting the findings [1]. The approach to quality assessment of the evidence attempts to minimize the risk of bias through the use of an explicit and transparent methodology. A typical example of evidence synthesis is a systematic review of treatment effectiveness [8]. The product will be a guideline which provides evidence-based statements for clinical decision-making, practice, and policy. The critical appraisal of evidence is key to the derivation of trustworthy practice recommendations and is the essence of evidence-based medicine [9].

The evidence used in the systematic reviews may include a range of study designs, from RCTs to observational studies, including case series too. RCTs are ranked the highest in the hierarchy of evidential validity due to their unique design that randomly assigns participants into experimental or control groups to compare outcomes. Randomization targets the minimization of selection bias in generating evidence about the effectiveness of interventions. To robustly implement and report trials, researchers performing RCTs are required to undertake regular training (good clinical practice courses), and trials are required to be prospectively registered in registries such as Clinicaltrials.gov, amongst other requirements [10,11]. Therefore, whenever available, systematic reviews and guidelines endeavor to include RCTs over other designs [12,13]. To enable the appraisal of evidence syntheses, several instruments or tools exist. As examples of tools for assessing systematic reviews, two widely used tools are AMSTAR-2 and ROBIS [14,15]. Other examples of tools for practicing guidelines are RIGHT and AGREE [16,17]. These tools include domains on how studies included in the synthesis were identified, selected, appraised, and analyzed, among others. They cover the risk of bias assessment. Nevertheless, the risk of bias assessment only partially targets integrity as the two concepts are not synonymous. They do not explicitly target study integrity assessment within evidence syntheses, a topic on which this paper will focus.

3. Relationship between Primary Research Integrity and Evidence Syntheses

Research integrity is a term that captures compliance with ethical and professional standards in the conduct of scientific studies [18]. In order to elucidate the research integrity concept, we provide, in Table 1, key research integrity terms and definitions.

Table 1. Research integrity related terms and definitions.

Research Integrity Terms	Definition of Terms *
Research integrity	Undertaking research in accordance with ethical and professional principles and standards.
Integrity principles	A set of values and concepts for guiding researcher behavior.
Integrity standards	Specifications of conduct that must be adhered to when participating in or carrying out research.
Bias	Systematic error that invalidates the observed effects in trials due to flaws in methodological aspects such as failure to concealment of randomization, lack of blinding, etc. Bias is distinct from data-related integrity flaws that arise due to misconduct.
Moral Values	The set of principles and standards that differentiate “right” from “wrong”.
Bioethics	Making choices in biomedical research around what are “right” and “wrong” values and behaviors.
Anti-whistleblower	Individuals who do not report nor prevent another individual from reporting known research misconduct.
Dishonesty	Behaviors that transgress moral values and bioethical standards
Duplication	A redundant publication that recycles or borrows content from authors own previous work without citation (see also self-plagiarism).
Ethics	Decision making based on moral and bioethical principles to protect those subjects of research and wider society.
Ethical Misdemeanors	Unacceptable or improper behavior that violates formal regulations.
Expression of concern	Note issued by journal editors or publishers to make readers aware that there is a concern about the integrity of a particular published article.
Fabrication, Falsification, Plagiarism (FFP)	The unholy trinity of misconduct in education, research or scholarship. Plagiarism is not the focus of this paper.
Fabrication	Making up data, experiments, or other significant information in proposing, conducting, or reporting research and using them as if genuine.
Falsification	Forging research content, images, data, equipment, or processes in the way that they are inaccurately represented.

Table 1. *Cont.*

Research Integrity Terms	Definition of Terms *
Forgery	Forging content, images, data, equipment, or processes in the way that they are inaccurately represented.
Fraud	Any intentional act of deception in research violating research ethics.
Infringement	Breach of good practice occurring from questionable, unlawful or unethical behavior.
Irresponsible Research Practices	Practices that are regarded as unethical but fall short of being considered research misconduct
Masking	Subset of data falsification consisting of minimizing or omitting data which does not support desired conclusions or results.
Misconduct	Unethical or unprofessional behavior in research.
Negligence	Failure to follow the required standard that results in harm to a person or organization.
Plagiarism	Presenting the work of others as if it were own work without proper acknowledgment or citation of the original source. Plagiarism is not the focus of this paper.
Self-plagiarism	Auto-plagiarism, i.e., the author adds insignificant additional data or information to previously published work changing title, modifying aim of the study or recalculating results, with the omission of citation to own previous publications. Self-plagiarism is not the focus of this paper.
Questionable research practices (QRPs)	Research practices that are unethical but fall short of being considered research misconduct.
Recycle	Recycling or borrowing content from authors own previous work without citation.
Redundant Publication	A published work (or substantial sections from a published work) is/are published more than once (in the same or another language) without adequate acknowledgment of the source/cross-referencing/justification. It is also when the same (or substantially overlapping) data is presented in more than one publication without adequate cross-referencing/justification, particularly when this is done in such a way that reviewers/readers are unlikely to realise that most or all the findings have been published before.
Replication	Repeating a piece of research in order to verify and/or complement the original results.
Retraction	Withdrawing or removing a published paper from the research record because of a variety of reasons including a post-publication reassessment showing that the data or results reported are unreliable or because the paper involves research misconduct. Journals publish retraction notices and identify retracted papers in electronic databases with reasons for retraction no always clearly stated.

Table 1. *Cont.*

Research Integrity Terms	Definition of Terms *
Transgression	Breach of good practice occurring from questionable, unlawful or unethical behavior.
Transparency	Openness about activities and related decisions that affect academia and society and willingness to communicate these in a clear, accurate, timely, honest and complete manner.
Violation/Breach	Breach of responsible research practices due to questionable, unlawful or unethical behavior in the conduct, analysis and reporting of research.
Conflict of interest	Potential or perceived compromise in judgement or objectivity due to financial or personal relationships or other considerations.
Confidentiality Violation	Disclosing to others information received in confidence without prior, explicit authorization of the person to whom the information belongs.
Author's Ethical Rights	The right to vindicate the ownership of work and assure its integrity and genuine status
Authorship Abuse	Any kind of authorship attribution not based on genuine contribution.
Authorship Coercion	An authorship that is demanded rather than voluntarily awarded.
Ghost Authorship	The practice of using a non-named (merited, but not listed) author to write or prepare a text for publication.
Invented Authorship	Naming a fictitious person, a colleague or a stranger as a co-author without permission.
Unethical Authorship	Crediting a person who has not contributed to the research in authorship; excluding from authorship a genuine contributor; manipulating the sequence of authors in an unjustified and improper way; removing names of contributors in subsequent publications; using one's power to insist on being added as an author without any contribution; including an author without their permission.

* Adapted from "Glossary for Research and Academic Ethics and Integrity (<https://h2020integrity.eu/glossary>), accessed on 26 February 2018 [19], "Errata, retractions, partial retractions, corrected and republished articles, duplicate publications, comments (including author replies), updates, patient summaries, and republished (reprinted) articles policy for MEDLINE (fact sheet)" 2015 (<https://www.nlm.nih.gov/pubs/factsheets/errata.html>), accessed on 26 February 2018 [20]. COPE Forum: Expressions of concern (<https://publicationethics.org/resources/forum-discussions/expressions-of-concern>), accessed on 26 February 2018 [21] and Khan, K.S.; Zamora, J. Systematic Reviews to Support Evidence-Based Medicine, 3rd ed.; Taylor & Francis Publishing: London, UK, 2022 [1].

Focusing evidence syntheses on the best available study designs, carrying out selected study appraisal, and basing inferences on the highest quality subgroup of studies targets avoidance or minimization of the risk of bias. Current review methodology does not explicitly target study integrity. It is important to recognize that, conceptually, study quality assessment and integrity evaluation are not synonymous. Inherently, systematic reviewers tend to rely on the editorial and peer-review systems established by journals, which appear as a custodian of research integrity assessment. However, there are well-known gaps in journals' authors' instructions, editorial and peer-review evaluations, and investigation policies about post-publication allegations of scientific misconduct [22]. Study integrity assessment can be time-consuming, permitting fabricated or falsified studies to remain

usable and to be included into evidence syntheses [23]. Only 5% of systematic reviews or clinical practice guidelines have corrected or retracted their results, with respect to retractions of included studies [24]. Without actual integrity assessment, the underpinned source studies behind evidence syntheses may include those that do not comply with responsible research conduct as genuine data.

The incorporation of RCT integrity assessment in evidence syntheses is an important consideration because the number of studies with expressions of concern has been rising exponentially (Figure 2).

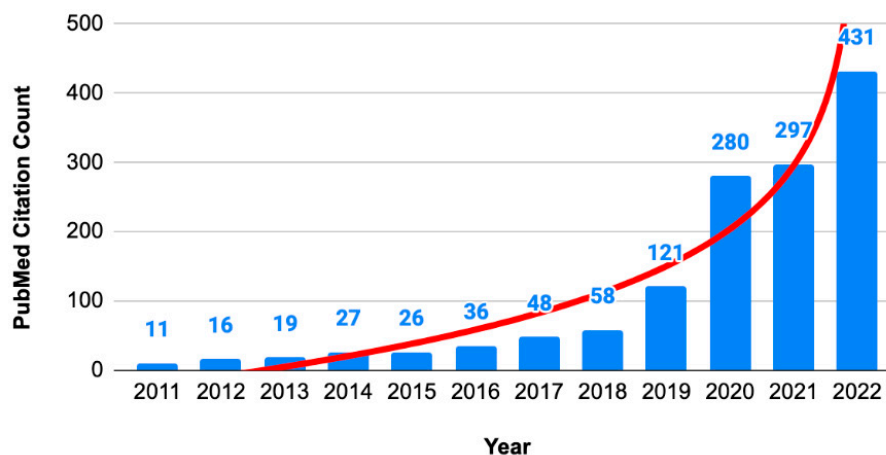


Figure 2. Numbers of articles with expressions of concern. Annual citation counts of expression of concern about articles in PubMed database.

Based on data retrieved on 2 February 2023, from the Retraction Watch Database, the top five countries with the highest number of retracted clinical studies per country are the United States, Japan, China, Germany, and India, with 299, 281, 245, 192, and 98 retractions, respectively. This data indicates that the retraction of studies is a global issue, and developed countries bear a great deal of responsibility for studies lacking integrity (Figure 3).

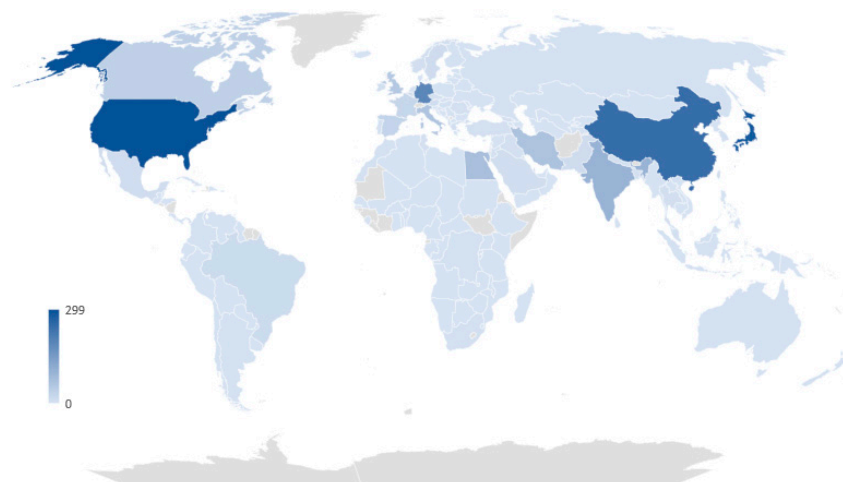


Figure 3. The number of retracted clinical studies per country based on Retraction Watch Database (<http://retractiondatabase.org>, data extracted on 2 February 2023).

These issues, in turn, raise concerns about the integrity of evidence syntheses as inclusion of studies with retractions due to issues in “data/analyses/results” lead to summary estimates in systematic reviews that depart from the studies without these issues [25].

4. Importance of Assessing the Integrity of RCTs Included in Evidence Syntheses

In this paper, we focus on evidence syntheses that deploy RCTs, as it is a high-validity study design to underpin evidence-based medicine. We focus specifically on potential data-related integrity issues in RCTs. Behind every disease prevention and treatment breakthrough, there are thousands of volunteer participants in RCTs whose data are collated in evidence syntheses. Despite the need for obtaining ethics approval, confirming informed consent, and applying independent oversight during trial conduct, RCTs are not exempt from the possibility of (un)intentional integrity deviations. The general fact that expressions of concern (Figure 2) and retractions (Figure 3) are numerous has shaken public confidence, being markedly astounding during the COVID-19 pandemic [26]. It is likely that not all retractions are the result of deliberate fraud, falsification, and fabrication. Unintentional errors, spin, or flawed techniques are bound to have played their part [27,28]. However, every RCT with integrity concerns that remains usable poses a threat to patients and public health. Therefore, systematic reviewers and guideline developers need to be vigilant about problematic or “zombie” RCTs [29].

Evidence syntheses affected by inclusion of RCTs that have integrity deficits are not difficult to find. The need for change in the attitude towards integrity assessment within reviews is highlighted by the following examples. Recently, Hill et al. retracted a meta-analysis of RCTs concerning COVID-19. The significant benefits initially observed could not be sustained after several of the included studies in the meta-analysis were withdrawn due to fraudulent data or other additional problems [30,31]. Avenell et al. reviewed the impact of the inclusion of retracted RCTs on evidence syntheses [29]. This group of retracted RCTs were published in the late nineties, and the reason for the retractions was serious misconduct, including concerns related to data integrity. RCT retractions in this case were only applied nearly two decades after publication [32]. Following the retractions, at least one of the retracted RCTs had been included in 32 evidence syntheses published. Avenell et al. judged that the conclusion of 13/32 evidence syntheses would have changed if the retracted RCTs had been excluded. Marret et al. reviewed the impact of the inclusion of another group of retracted RCTs on evidence syntheses and judged that the conclusion of 2/14 would have changed if the retracted RCTs had been excluded [33]. Habib et al., in a commentary concerning the impact of the inclusion of yet another group of retracted RCTs on evidence syntheses, indicated that the likelihood of compromise was modest with some systematic reviews that performed sensitivity analyses, noting that their conclusions were different after excluding the retracted data [34]. Fanelli et al. similarly concluded that the potential epistemic cost of retraction was modest, with emphasis on the reason for retraction as the key issue [25].

These examples confirm that a specific methodology is required to address the issue of RCT integrity in systematic reviews head-on. It is important to recognize that the purpose of this methodology ought to focus on the protection of patients and public health. The precursors of failure to comply with responsible research conduct are many, including misconduct, recklessness, carelessness, lack of training, etc. [35]. It is not the systematic reviewers' role to judge original authors' motivations; journals, employers, funders, etc. have investigative and sanctioning roles. What systematic reviewers need is a proactive attitude towards synthesizing evidence that does not harbor integrity deviations [27]. This will protect the trustworthiness of evidence syntheses and evidence-based medicine.

5. How to Incorporate RCT Integrity Assessment in Evidence Syntheses

A systematic approach is required on various evidence synthesis fronts, including but not limited to the identification of honestly conducted research in searching, assessment of integrity (separately from bias assessments) of included studies, prior planning of sensitivity analyses for integrity, and transparency in generating inferences, given the cautious possibility of compromise in underlying data (Figure 4).

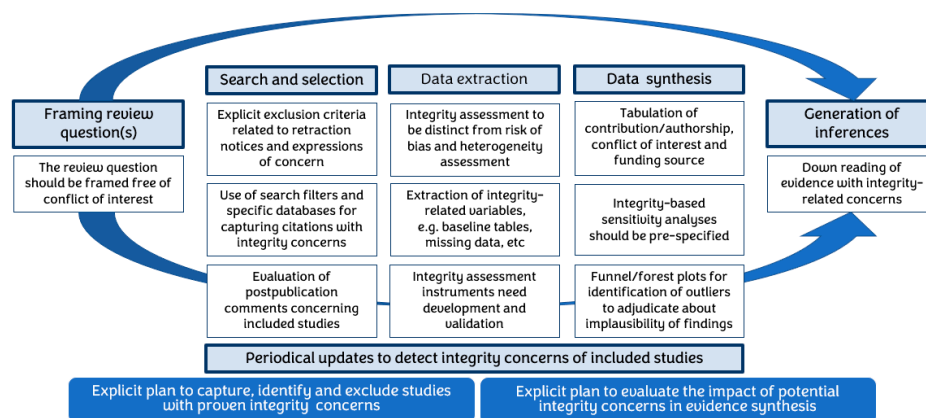


Figure 4. Infographic of the key steps during evidence synthesis to maximize the integrity of included research.

To assess the integrity of included RCTs, it is important to follow rigorous tool development methodology that has been widely applied previously in the development of study quality or risk of bias assessment tools [36,37]. Once developed, these should be fed back into revisions of current reporting guidelines [38]. Statistical analyses would need to be refocused, addressing integrity issues, e.g., funnel plot analyses may be used to inspect small studies that have implausibly large effects. In study-level published data meta-analyses, sub-group and meta-regression analyses may routinely include an integrity assessment-based variable. In individual patient data meta-analyses, statistical techniques can be applied for the detection of anomalous patterns in the underlying numerical data to check for data integrity [39–42]. These more sophisticated analyses should feed into evidence grading for the generation of judicious inferences.

6. Implications, Issues, Challenges, and Limitations

Evidence synthesis, mindful of research integrity, will need to attempt to collate all empirical evidence that fits pre-specified eligibility criteria, excluding studies with proven integrity concerns, using explicit methods to detect and quantify these concerns, evaluating their impact in planned statistical analyses and minimizing or eliminating the pollution of the inferences that may arise due to inclusion of studies with possibly compromised data. Periodical updates of reviews to detect integrity concerns of included studies should be performed (Table 2).

There are many issues to consider. A controversial aspect here is whether to include studies with expressed concerns but without proven misconduct. This controversy is not too different to the inclusion of RCTs with varying levels of risk of bias arising due to faulty randomization, lack of allocation concealment, or blinding. This is now routine in effectiveness reviews. The development, validation, and application of advanced methods that can accurately detect integrity breaches in publicly available RCTs are needed [43–45]. Handling the integrity assessment of selected studies after excluding those with confirmed integrity breaches requires further consideration in methodological development of evidence syntheses as, at present, there are no clear procedures established. We could only find one article providing a method for detecting retracted literature cited in systematic reviews and meta-analyses [46]. Unfortunately, there are no validated tools for integrity assessment yet [40]. One important aspect to highlight is that to date, there is no standard definition of the term “research integrity”. Thus, the concepts of bias, quality, validity, and integrity can be confusing for readers as well as reviewers. A precise characterization of research integrity distinct from the idea of risk of bias assessments is needed as a starting point for the required methodological developments to take place in the right direction in the future.

Table 2. Steps of evidence synthesis and key integrity related issues.

Steps of Evidence Synthesis	Integrity Related Issues
Framing question	Review question should be framed free of conflict of interest and should specify its focus on including studies with integrity
Search and selection	Explicit exclusion criteria related to retraction notices and expressions of concern about integrity should be pre-specified
	Specific retraction and integrity concern searches should be deployed, e.g., in Retraction Watch Database
	Search filters for capturing citations with integrity concerns should be developed and used
	Evaluation of post-publication comments concerning included studies should be sought and evaluated, e.g., letters to editors, Pub-Peer comments, etc.
Data extraction	Specific data extraction to permit integrity assessment, e.g., baseline tables, missing data, etc.
	Integrity assessment needs to be distinct from risk of bias and heterogeneity assessment
	Integrity assessment instruments need development and validation
	Developed integrity assessment instruments need automation
Data synthesis	Tabulation of contribution/authorship, conflict of interest and funding source, etc. related to integrity should be routine
	Integrity-based sensitivity analyses should be pre-specified
	Use of funnel plots to look for outliers should additionally be pre-specified with delineation of threshold for defining implausibly extreme results
Inference generation	Down grading of evidence with integrity concerns should be explicitly deployed in generation of recommendations
Updates	Periodical updates of reviews to detect integrity concerns of included studies and issuing correction notices

A particularly important issue that impinges on the critical appraisal of integrity is the enormous literature size and growth of publication rates. In 2015, there were about 28,100 and 6450 English-language and non-English-language science, technology, and medicine journals, respectively, growing at about 3% annually [47]. Defective studies should never get to enter circulation, but the growing volume complicates the challenge. To put a lid on the integrity-related concern that will grow with this literature expansion, automated checks will be required just as they have been used for the detection of plagiarism. This is required in part because the peer review process might not be able to cope with new ways of capturing defective literature since editors and peer reviewers would have to upgrade their knowledge/skill sets [7]. Computer sciences are being deployed for critical appraisal [48].

7. Artificial Intelligence for Integrity Assessment

The inclusion of efficient tools that will automate integrity assessment in evidence synthesis is the next methodological advance required. Review projects usually require a team of reviewers who screen and identify literature and evaluate included study quality. They will additionally need to perform integrity assessments before collating findings and

generating recommendations. Currently, reviews take up human effort and take too long to collect and evaluate the data included, undertaking double-checks to minimize errors. The review process is estimated to take, on average, 67.3 weeks (IQR 41.6) to complete an evidence synthesis, and publishing the synthesis involves 5.3 review team members (IQR 3), on average [49]. In addition to being slow, there is an inherent error rate associated with human effort, e.g., the selection process suffers a 10% error (false inclusion and false exclusion) rate [50]. Thus, it has been concluded that: *“Systematic reviews presently take much time and require large amounts of human resources. In the light of the ever-increasing volume of published studies, application of existing computing and informatics technology should be applied to decrease this time and resource burden”* [49]. The use of automation is also emphasized for integrity assessment [27] and, although infrequently used until now, it is impactful [51].

These assessments need development and validation of new instruments to enable the detection and exclusion of questionable evidence from evidence syntheses, without the need to wait for retractions. Automated detection of retractions, specifically for data-related misconduct associated with fabrication, falsification, and other types of forgery, are needed [44]. This way, integrity assessments in systematic reviews will streamline the literature correction process and may include alerts for journals to trigger investigations. The tools, once developed and validated, may also be used for improving peer-reviews, reducing the circulation of “zombie” trials. This will improve the validity of the evidence syntheses going forward and will assist in the pre- and post-publication review process in cases of allegations.

8. Current Conclusions

Evidence syntheses collating RCTs influence practice and health policies, directly impacting patient care. The investigation and retraction of RCTs with integrity concerns is a slow process. Thus, defective RCTs remain in circulation, putting patients at risk. Even after retraction, defective studies continue to be cited in systematic reviews as they are not removed from databases and their signposting is poor. Evidence syntheses fail to issue corrections even when retractions are identified. All this entails risk as patients remain exposed to interventions that are futile or even risky for their health. Evidence syntheses need to urgently upgrade their methods to incorporate integrity assessments as a routine, as outlined in Figure 4.

9. Future Direction

Research integrity, a broad concept holistically incorporating both ethical and professional standards [18], needs to be considered in evidence synthesis covering the whole range of issues inherent in responsible RCT conduct. This needs to be defined explicitly through further research. However, the illegal Tuskegee syphilis experiment from the recent past is a case in point where unethical research can permeate within the literature without comment. In a 2022 journal article [51], Tobin wrote: *“Despite 15 journal articles detailing the results, no physician published a letter criticizing the Tuskegee study. Informed consent was never sought; instead, Public Health Service researchers deceived the men into believing they were receiving expert medical care”*. These articles remain formally unretracted from the literature to this date. Note that ethics and consent standards were not covered by us in this article as we focused on data-related integrity, but it remains a key aspect demanding future research and development for its proper implementation and monitoring in RCTs. There are many articles showing deficits in informed consents in clinical studies [52–54], and this type of integrity assessment ought to be featured in evidence syntheses.

There ought to be an emphasis on prevention, metaphorically nipping the evil studies in the bud [55]. This begins with clarifying integrity-related definitions, e.g., what are questionable research practices that raise integrity concerns. Then, it would be appropriate to identify modifiable factors and barriers that may affect best practice compliance [3]. Future direction would take the above forward emphasizing routine adherence to ethical and professional standards through periodic integrity-specific training tailored to educational

environment for all stakeholders involved in the RCT research lifecycle, including but not limited to researchers, ethics committee members, funders, editors, peer reviewers, systematic reviewers, guideline makers, drug regulators, medical journalists, as well as lay readers. Integrity training in clinical trials has been recommended in a recent international multi-stakeholder consensus statement [44], with emphasis on enabling research teams from low resource settings to make contributions. The creation of solid systems backed by valid and robust instruments and methods for inculcating research integrity are urgently needed. Future evidence syntheses will directly benefit from these improvements in the integrity of the conduct, analysis, and reporting of primary RCTs collated within literature reviews.

Author Contributions: Conceptualization, K.S.K. and A.B.-C.; methodology, K.S.K. and A.B.-C.; investigation, N.C.-I. and M.N.-N.; data curation, N.C.-I. and M.N.-N.; writing—original draft preparation, K.S.K. and M.N.-N.; writing—review and editing, K.S.K., A.B.-C., N.C.-I., J.Z. and M.N.-N. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Acknowledgments: K.S.-K. is a distinguished investigator at the University of Granada funded by the Beatriz Galindo (senior modality) program of the Spanish Ministry of Education; M.N.-N. is granted by the Research institute Carlos III under the Rio Hortega training program (CM20/00074).

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Khan, K.S.; Zamora, J. *Systematic Reviews to Support Evidence-Based Medicine*, 3rd ed.; Taylor & Francis Publishing: London, UK, 2022.
2. De Vrieze, J. Large survey finds questionable research practices are common. *Science* **2021**, *373*, 265. [CrossRef] [PubMed]
3. Gopalakrishna, G.; ter Riet, G.; Vink, G.; Stoop, I.; Wicherts, J.M.; Bouter, L.M. Prevalence of questionable research practices, research misconduct and their potential explanatory factors: A survey among academic researchers in the Netherlands. *PLoS ONE* **2022**, *17*, e0263023. [CrossRef] [PubMed]
4. Steen, R.G.; Casadevall, A.; Fang, F.C. Why Has the Number of Scientific Retractions Increased? *PLoS ONE* **2013**, *8*, e68397. [CrossRef]
5. Vinkers, C.H.; Lamberink, H.J.; Tijdkink, J.K.; Heus, P.; Bouter, L.; Glasziou, P.; Moher, D.; Damen, J.A.; Hooft, L.O.W. 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*, e3001162. [CrossRef]
6. Ioannidis, J.P.A. Hundreds of thousands of zombie randomised trials circulate among us. *Anaesthesia* **2021**, *76*, 444–447. [CrossRef]
7. Teixeira da Silva, J.A. A Synthesis of the Formats for Correcting Erroneous and Fraudulent Academic Literature, and Associated Challenges. *J. Gen. Philos. Sci.* **2022**, *53*, 583–599. Available online: <https://link.springer.com/article/10.1007/s10838-022-09607-4> (accessed on 30 January 2023). [CrossRef]
8. Higgins, G. *Cochrane Handbook for Systematic Reviews of Interventions*; John Wiley & Sons: Hoboken, NJ, USA, 2011.
9. Twells, L.K. Evidence-Based Decision-Making 1: Critical Appraisal. *Methods Mol. Biol.* **2021**, *2249*, 389–404.
10. World Medical Association (WMA). Declaration on Guidelines for Continuous Quality Improvement in Healthcare. Available online: <https://www.wma.net/policies-post/wma-declaration-on-guidelines-for-continuous-quality-improvement-in-health-care/> (accessed on 9 April 2023).
11. ICH Official Web Site: ICH. Available online: <https://www.ich.org/> (accessed on 17 January 2022).
12. Bauchner, H.; Golub, R.M.; Fontanarosa, P.B. Reporting and Interpretation of Randomized Clinical Trials. *JAMA* **2019**, *322*, 732–735. [CrossRef]
13. Stavale, R.; Ferreira, G.I.; Galvão, J.A.M.; Zicker, F.; Novaes, M.R.C.G.; De Oliveira, C.; Guilhem, D. Research misconduct in health and life sciences research: A systematic review of retracted literature from Brazilian institutions. *PLoS ONE* **2019**, *14*, e0214272. [CrossRef]
14. Shea, B.J.; Reeves, B.C.; Wells, G.; Thuku, M.; Hamel, C.; Moran, J.; Moher, D.; Tugwell, P.; Welch, V.; Kristjansson, E.; et al. AMSTAR 2: A critical appraisal tool for systematic reviews that include randomised or non-randomised studies of healthcare interventions, or both. *BMJ* **2017**, *358*, j4008. [CrossRef]
15. Whiting, P.; Savović, J.; Higgins, J.P.; Caldwell, D.M.; Reeves, B.C.; Shea, B.; Davies, P.; Kleijnen, J.; Churchill, R.; ROBIS group. ROBIS: A new tool to assess risk of bias in systematic reviews was developed. *J. Clin. Epidemiol.* **2016**, *69*, 225–234. [CrossRef]

16. Molino, C.d.G.R.C.; Ribeiro, E.; Romano-Lieber, N.S.; Stein, A.T.; de Melo, D.O. Methodological quality and transparency of clinical practice guidelines for the pharmacological treatment of non-communicable diseases using the AGREE II instrument: A systematic review protocol. *Syst. Rev.* **2017**, *6*, 220. [CrossRef]
17. Chen, Y.; Yang, K.; Marušić, A.; Qaseem, A.; Meerpohl, J.J.; Flottorp, S.; Akl, E.A.; Schünemann, H.J.; Chan, E.S.Y.; Falck-Ytter, Y.; et al. A reporting tool for practice guidelines in health care: The RIGHT statement. *Ann. Intern. Med.* **2017**, *166*, 128–132. [CrossRef]
18. Steneck, N.H. Fostering integrity in research: Definitions, current knowledge, and future directions. *Sci. Eng. Ethics* **2006**, *12*, 53–74. [CrossRef]
19. H2020 INTEGRITY—Glossary. Available online: <https://h2020integrity.eu/resources/glossary/> (accessed on 27 January 2023).
20. National Library of Medicine. Errata, Retractions, and Other Linked Citations in PubMed. 2015. Available online: <http://wayback.archive-it.org/org-350/20180312141525/https://www.nlm.nih.gov/pubs/factsheets/errata.html> (accessed on 9 February 2023).
21. COPE. COPE Forum 26 February 2018: Expressions of Concern. Available online: <https://publicationethics.org/resources/forum-discussions/expressions-of-concern> (accessed on 11 April 2023).
22. Malički, M.; Jerončić, A.; Aalbersberg, I.J.J.; Bouter, L.; ter Riet, G. Systematic review and meta-analyses of studies analysing instructions to authors from 1987 to 2017. *Nat Commun.* **2021**, *12*, 5840. [CrossRef]
23. Schneider, J.; Ye, D.; Hill, A.M.; Whitehorn, A.S. Continued post-retraction citation of a fraudulent clinical trial report, 11 years after it was retracted for falsifying data. *Scientometrics* **2020**, *125*, 2877–2913. [CrossRef]
24. Kataoka, Y.; Banno, M.; Tsujimoto, Y.; Ariie, T.; Taito, S.; Suzuki, T.; Oide, S.; Furukawa, T.A. Retracted randomized controlled trials were cited and not corrected in systematic reviews and clinical practice guidelines. *J. Clin. Epidemiol.* **2022**, *150*, 90–97. [CrossRef]
25. 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.* **2022**, *29*, 442–459. [CrossRef]
26. Fleming, T.R.; Labriola, D.; Wittes, J. Conducting Clinical Research During the COVID-19 Pandemic: Protecting Scientific Integrity. *JAMA* **2020**, *324*, 33–34. [CrossRef]
27. Núñez-Núñez, M.; Andrews, J.C.; Fawzy, M.; Bueno-Cavanillas, A.; Khan, K.S. Research integrity in clinical trials: Innocent errors and spin versus scientific misconduct. *Curr. Opin. Obstet. Gynecol.* **2022**, *34*, 332–339. [CrossRef]
28. Fletcher, R.H.B.B. “Spin” in scientific writing: Scientific mischief and legal jeopardy. *Med. Law* **2007**, *26*, 511–525. [PubMed]
29. 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*, e031909. [CrossRef] [PubMed]
30. Hill, A.; Mirchandani, M.; Pilkington, V. Ivermectin for COVID-19: Addressing Potential Bias and Medical Fraud. *Open Forum Infect. Dis.* **2022**, *9*, ofab645. [CrossRef] [PubMed]
31. Hill, A.; Garratt, A.; Levi, J.; Falconer, J.; Ellis, L.; McCann, K.; Pilkington, V.; Qavi, A.; Wang, J.; Wentzel, H. Retracted: Meta-analysis of Randomized Trials of Ivermectin to Treat SARS-CoV-2 Infection. *Open Forum Infect. Dis.* **2021**, *8*, ofab358. [CrossRef] [PubMed]
32. Bolland, M.J.; Avenell, A.; Gamble, G.D.; Grey, A. Systematic review and statistical analysis of the integrity of 33 randomized controlled trials. *Neurology* **2016**, *87*, 2391–2402. [CrossRef]
33. Marret, E.; Elia, N.; Dahl, J.B.; McQuay, H.J.; Møiniche, S.; Moore, R.A.; Straube, S.; Tramèr, M.R. Susceptibility to fraud in systematic reviews: Lessons from the reuben case. *Anesthesiology* **2009**, *111*, 1279–1289. [CrossRef] [PubMed]
34. Habib, A.S.G.T. Scientific fraud: Impact of Fujii’s data on our current knowledge and practice for the management of postoperative nausea and vomiting. *Anesth. Analg.* **2013**, *116*, 520–522. [CrossRef]
35. Resnik, D.B.; Smith, E.M.; Chen, S.H.G.C. What is Recklessness in Scientific Research? The Frank Sauer Case. *Account. Res.* **2017**, *24*, 497–502. [CrossRef]
36. Sterne, J.A.C.; Savović, J.; Page, M.J.; Elbers, R.G.; Blencowe, N.S.; Boutron, I.; Cates, C.J.; Cheng, H.Y.; Corbett, M.S.; Eldridge, S.M.; et al. RoB 2: A revised tool for assessing risk of bias in randomised trials. *BMJ* **2019**, *366*, l4898. [CrossRef]
37. Jüni, P.; Witschi, A.; Bloch, R.; Egger, M. The hazards of scoring the quality of clinical trials for meta-analysis. *JAMA* **1999**, *282*, 1054–1060. [CrossRef]
38. Schulz, K.F.; Altman, D.G.; Moher, D. CONSORT 2010 statement: Updated guidelines for reporting parallel group randomized trials. *Ann. Intern. Med.* **2010**, *152*, 726–732. [CrossRef]
39. van den Bor, R.M.; Vaessen, P.W.J.; Oosterman, B.J.; Zuithoff, N.P.A.; Grobbee, D.E.; Roes, K.C.B. A computationally simple central monitoring procedure, effectively applied to empirical trial data with known fraud. *J. Clin. Epidemiol.* **2017**, *87*, 59–69. [CrossRef]
40. Pogue, J.M.; Devereaux, P.J.; Thorlund, K.; Yusuf, S. Central statistical monitoring: Detecting fraud in clinical trials. *Clin. Trials* **2013**, *10*, 225–235. [CrossRef]
41. de Viron, S.; Trotta, L.; Schumacher, H.; Hans-Juergen, L.; Höppner, S.; Young, S.; Buyse, M. Detection of Fraud in a Clinical Trial Using Unsupervised Statistical Monitoring. *Ther. Innov. Regul. Sci.* **2022**, *56*, 130–136. [CrossRef]
42. O’Kelly, M. Using statistical techniques to detect fraud: A test case. *Pharm. Stat.* **2004**, *3*, 237–246. [CrossRef]
43. Núñez-Núñez, M.; Maes-Carballo, M.; Mignini, L.E.; Chien, P.F.W.; Khalaf, Y.; Fawzy, M.; Zamora, J.; Khan, K.S.; Bueno-Cavanillas, A. Research integrity in randomized clinical trials: A scoping umbrella review. *Int. J. Gynecol. Obstet.* **2023**, 1–17. [CrossRef]
44. Khan, K.S. Cairo Consensus Group on Research Integrity. International multi-stakeholder consensus statement on clinical trial integrity. *BJOG* **2023**, 1–16. [CrossRef]
45. Khan, K.S.; Fawzy, M.; Chien, P.F.W. Integrity of randomized clinical trials: Performance of integrity tests and checklists requires assessment. *Int. J. Gynaecol. Obstet.* **2023**; in press. [CrossRef]

46. Morán, J.M.; Santillán-García, A.; Herrera-Peco, I. SCRUTATIOM: How to detect retracted literature included in systematic reviews and metaanalysis using SCOPUS© and ZOTERO©. *Gac. Sanit.* **2022**, *36*, 64–66. [[CrossRef](#)]
47. Ware, M.; Mabe, M. *The STM Report: An Overview of Scientific and Scholarly Journal Publishing Fourth Edition*; International Association of Scientific, Technical and Medical Publishers: Oxford, UK, 2015.
48. The Systematic Review Toolbox. Available online: <http://systematicreviewtools.com/software.php> (accessed on 18 January 2023).
49. Borah, R.; Brown, A.W.; Capers, P.L.; Kaiser, K.A. Analysis of the time and workers needed to conduct systematic reviews of medical interventions using data from the PROSPERO registry. *BMJ Open* **2017**, *7*, e012545. [[CrossRef](#)]
50. Wang, Z.; Nayfeh, T.; Tetzlaff, J.; O’Blenis, P.; Murad, M.H. Error rates of human reviewers during abstract screening in systematic reviews. *PLoS ONE* **2020**, *15*, e0227742. [[CrossRef](#)] [[PubMed](#)]
51. Tercero-Hidalgo, J.R.; Khan, K.S.; Bueno-Cavanillas, A.; Fernández-López, R.; Huete, J.F.; Amezcua-Prieto, C.; Zamora, J.; Fernández-Luna, J.M. Artificial intelligence in COVID-19 evidence syntheses was underutilized, but impactful: A methodological study. *J. Clin. Epidemiol.* **2022**, *148*, 124–134. [[CrossRef](#)] [[PubMed](#)]
52. Pietrzykowski, T.; Smilowska, K. The reality of informed consent: Empirical studies on patient comprehension—Systematic review. *Trials* **2021**, *22*, 57. [[CrossRef](#)] [[PubMed](#)]
53. Schellings, R.; Kessels, A.G.; ter Riet, G.; Knottnerus, J.A.; Sturmans, F. Randomized consent designs in randomized controlled trials: Systematic literature search. *Contemp. Clin. Trials* **2006**, *27*, 320–332. [[CrossRef](#)]
54. Timmermann, C.; Orzechowski, M.; Kosenko, O.; Woniak, K.; Steger, F. Informed Consent in Clinical Studies Involving Human Participants: Ethical Insights of Medical Researchers in Germany and Poland. *Front. Med.* **2022**, *19*, 901059. [[CrossRef](#)]
55. Khan, K.S. Comment on Khan: “Flawed Use of Post Publication Data Fabrication Tests”. Research Misconduct Tests: Putting Patients’ Interests First. *J. Clin. Epidemiol.* **2021**, *138*, 227. Available online: <https://pubmed.ncbi.nlm.nih.gov/34343640/> (accessed on 9 April 2023). [[CrossRef](#)]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.