






REVIEW

Open Access



The randomized clinical trial trustworthiness crisis

Furqan A. Butt¹ , Mohammad Fawzy² , Bassel H. Al Wattar^{3,4} , Aurora Bueno-Cavanillas¹ ,
Khalid S. Khan^{1*}  and Yacoub Khalaf⁵ 

Abstract

Background The rising number of retracted randomised clinical trials (RCTs) is a concern over their trustworthiness. In today's digital landscape electronic observational data is easily accessible for research purposes. This emerging perspective, in tandem with the growing scrutiny of RCT credibility, may steer some researchers towards favouring non-randomized studies. It is crucial to emphasize the ongoing need for robust RCTs, shedding light on the areas within trial design that require enhancements and addressing existing gaps in trial execution.

Main body Evidence-based medicine pivots on the nexus between empirical medical research and the theoretical and applied facets of clinical care. Healthcare systems regularly amass patient data, creating a vast reservoir of information. This facilitates large-scale observational studies, which may appear as potential substitutes for RCTs. These large-scale studies inherently possess biases that place them a notch below randomized evidence. Honest errors, data manipulation, lapses in professionalism, and methodological shortcomings tarnish the integrity of RCTs, compromising trust in trials. Research institutions, funding agencies, journal editors and other stakeholders have the responsibility to establish robust frameworks to prevent both deliberate and inadvertent mishandling of RCT design, conduct and analysis. Systematic reviews that collate robust RCTs are invaluable. They amalgamate superior evidence instrumental in improving patient outcomes via informed health policy decisions. For systematic reviews to continue to retain trust, validated integrity assessment tools must be developed and routinely applied. This way it will be possible to prevent false or untrustworthy research from becoming part of the recommendations based on the evidence.

Conclusion High-quality RCTs and their systematic reviews play a crucial role in acquiring valid and reliable evidence that is instrumental in improving patient outcomes. They provide vital information on healthcare effectiveness, and their trustworthiness is key to evidence-based medicine.

Keywords Randomised clinical trials, Observational studies, Evidence-based medicine, Research integrity, Evidence-based medicine, Clinical medicine

*Correspondence:

Khalid S. Khan

Profkhan@gmail.com

Full list of author information is available at the end of the article



© The Author(s) 2023. **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

Background

The term evidence-based medicine or EBM, the approach of using medical evidence to inform clinical decision-making [8], emphasizes the need to evaluate all relevant evidence thoroughly, rather than selectively choosing data that support a particular argument. EBM operates on the premise that not all evidence carries equal weight in the decision-making [48]. Medical decisions should rely on the best available evidence, while taking into account patients' values and preferences.

The trustworthiness of the evidence, i.e., the extent to which we can be confident that the research findings are firm, is central to EBM. EBM classifies controlled clinical observations as more reliable evidence than uncontrolled clinical observations, biological experiments, or individual clinician experiences [8]. Among controlled clinical studies, randomised controlled trials (RCTs) are at the top of the evidence hierarchy. However, some academicians, lobbyists, professional associations, funding organizations, and even regulators are now considering observational data, even for licensure purposes [49]. This is largely due to the ready availability of data in a digital world where every clinical episode is electronically captured in the health system [49]. This type of data, often referred to as real-world data, is different from randomised data.

This real-world data has inherent limitations of the observational design, which cannot be resolved solely by increasing the amount of data used. The size of the data merely serves to increase the statistical power; it does not necessarily enhance study validity. Therefore, abandoning RCTs will revert to the pre-EBM research era. This proposed reliance on observational data can result in false statistical inferences, downgrading the credibility of research findings and risking spurious statistical significance [1, 14, 23]. This commentary seeks to compare the value of clinical trials with observational studies, highlighting the importance of the former. We discuss the potential advantages of rigorously conducted observational research. Given the increasing concerns about RCT trustworthiness, we explore the current gaps and areas for improvement.

Main text

Importance of RCTs

A robust RCT determines efficacy for regulatory approval and identifies effective treatment options for practice. It provides insights into causal linkages, reducing the risk of bias and confounding. The specification of sample characteristics in RCT may adversely impact its generalizability to real-world situations, as the effects of treatment on people who are not part of the RCT become a matter of

judgment [11]. Surrogate endpoints may not accurately reflect the desired patient-oriented outcomes. RCTs may not be practical in rare diseases. Emergency situations may also be considered a barrier. However, the response to COVID-19 pandemic has shown that RCTs a key to global health challenges.

Conducting RCTs that are reliable, unbiased, and transparent is crucial to improving healthcare. To achieve this, researchers should ask relevant questions that are important to patients and the public. They should conduct systematic reviews of prior evidence to design new trials that use the best possible comparators [33] to evaluate the effects on core clinical outcomes [25]. Patient and public involvement is a key feature of relevant clinical trials [34]. It is important to pre-register the RCT protocol and make it publicly accessible before recruiting participants. RCTs should be conducted as planned without protocol deviation along the course of the work. Proper randomization and allocation concealment are essential for high-quality trials, yet many RCTs have shortcomings [51]. Their analysis strategy should be pre-determined before the outcomes within the trial dataset are revealed to investigators. There should be detailed reporting of funding and conflicts of interest. Recently an international multistakeholder group has issued an integrity statement specifically covering RCTs [27].

Systematic reviews to compile RCT evidence

Systematic reviews assist medical professionals in collating and evaluating evidence, offering direction for future action. Some reviews have resulted in unintended consequences such as the production of unreliable guidelines, lapses in regulation, and tardiness in removing hazardous drugs [20]. These deficiencies have resulted in elevated treatment expenses, and superfluous medical procedures, including overdiagnosis, and overtreatment. However, there has been continued progress in the methodology of systematic review to push boundaries seeking significant advancement in this area [3, 20]. Over 50,000 systematic reviews were published in 2022 [38], coming closer to the vision of organising living collations of accumulating evidence for addressing the shortcomings of scientific research and for EBM [8]. The failure in the past to do so has led to unnecessary suffering, loss of life, and the misuse of healthcare resources. The mission now is to review existing research in a methodical way to prevent studies with integrity flaws from entering the evidence syntheses [38]. While RCTs are generally considered to be the gold standard, there are still some problematic RCTs that contribute to the current evidence base for patient management. Finding a solution to this issue will require efforts in the integrity test validation [28, 38].

Hidden aspects around RCT registration

Every day, 75 RCTs are published. Given that 60% of randomized controlled trials are eventually published [42], it is estimated that around 125 trials are initiated every day. This number is 2.5 times more than the number of trial protocols that are registered on *clinicaltrials.gov* [21]. Out of all the registered trials, approximately 5% were never started [44]. Many trials are terminated early, which accounts for 12% of those posted on *clinicaltrials.gov* and 28% of those approved by research ethics committees [53]. There is a notable amount of secondary literature [9], so it is necessary to evaluate 150 to 200 trials every day [21]. When accounting for rejections and resubmissions, the total number of submissions may exceed this range. Irrelevance can be a cause of waste in RCTs, but it is difficult to measure in part because the issue is subjective.

The ongoing tolerance of wastefulness in clinical research can be traced back to the existing incentive structure, which heavily prioritizes the number of publications as a gauge of academic achievement [22]. The way of ranking researchers needs reconsideration considering the effects of the academic incentive structure. This is essential for understanding the tremendous pressure that researchers face to publish as well as for raising the standard of scientific research. There is a need for a balance as a solution and for maintaining public confidence in science [16]. Researchers may actively cheat or purposefully remove data in order to generate publications more quickly due to academic competitiveness. Journals typically publish positive findings, which adds to publication bias. Preventing instances of cheating is necessary for achieving a reduction in academic misconduct [12].

The perspectives of patients are not always considered when formulating research questions and selecting outcomes [35]. Replication of a study's findings using the same methodology is a validation process. This notion needs to be valued in RCTs, where if a new RCT found similar conclusions to similar ones before it, all should earn credibility. It is unfortunate that null results and replications, which are scientifically sound, are not valued by clinical journals. Researchers should be incentivized to submit such findings or to conduct replication RCTs to correct the scientific record. Journals considering only positive results as a *modus operandi* to accept research for publication unintentionally incentivize the generation of spurious findings and false positives [20]. RCTs, if they lack transparency and independent scrutiny including instances of failure to follow their protocol, stopping early or including ghost authorship, will lead to publication with reporting biases, overinterpretation or misinterpretation of results, uncorrected errors, and undetected fraud [20].

The importance of replication

The issue of reproducibility is a generic concern across science. It has been said that 70% of researchers have attempted to duplicate another scientist's findings but failed [5]. This important validation research concept needs to be the norm in RCTs by encouraging trialists to do replication research instead of opining against RCT findings with which they disagree. This concept of replication research was established in the early days of the Royal Society (in 1660), where experimenters would duplicate whatever assertions, they made in front of their peers. Nowadays, replication has lost its interest, partly because journals constantly look for new findings to publish in the interest of seeking more citations. The lack of publicly available data, protocols, and statistical codes has made it unfeasible to re-analyze published data. It is commonly, but perhaps erroneously, accepted that a published study's findings or results are reliable [32]. This approach encourages the acceptance of published findings blindly. If the measures used in the original study are indirect and imprecise, and examining the raw data is arduous or time-intensive, it might be more feasible to utilize the outcomes from extension or replication studies as a fresh foundation [13].

In scientific research, replication studies will likely require a re-examination of all potentially controversial research practices. This is because the size of an effect can be impacted not only by questionable practices but also by sampling errors [43]. Replication studies serve multiple purposes, including correcting for sampling errors (false-positive detection), controlling for artefacts, addressing researcher fraud, testing generalizations to non-identical study samples, and examining the same hypothesis of a prior study with a different methodology. One single replication study cannot fulfil all five purposes simultaneously [54]. The RCT replication approach needs to be developed on the logical lines outlined above.

Intentional and unintentional misuse of statistics

Statistical analyses, regardless of methodology, are susceptible to misuse. Researchers possess various freedoms when conducting data analysis and interpretation, which can lead to variations in findings [17]. Multiple conclusions can be drawn from the same research question and data, e.g., findings from identical studies may even yield contradictory results. Statistical manipulation, whether intentionally as a kind of fraud or unintentionally as a kind of illusion, may be used to produce certain results from a dataset [17]. RCTs make claims from subgroup analysis while failing to consider multiple testing and lack appropriate interaction testing [17]. Making datasets publicly available emphasizes the need for openness in scientific publishing. The International Committee of

Medical Journal Editors (ICMJE) has stated, “it is an ethical obligation to responsibly share data obtained from interventional clinical trials because participants have put themselves at risk” [46]. The ICMJE has recommended making the deidentified individual patient data (IPD) available to the public within six months of publication. This proposal sparked a discussion [19, 24, 39, 40, 50], resulting in withdrawal of the proposal followed by a reissue of recommendation requiring a data-sharing plan defined during study registration [47]. This limitation on openness in science ought to be reconsidered in the near future. This call for reconsideration is important as the current RCT integrity tests lack validation. Therefore, a triad of data sharing, integrity test validation and stop the blame game needs to run simultaneously. Irresponsible use of integrity tests can damage the reputation and career of honest researchers.

Artificial intelligence

Certain artificial intelligence (AI) solutions which include automated or semi-automated tools will likely be used as a solution towards improving RCT integrity testing. Areas where RCT validity may include concerns such as insufficient sample size, inadequate randomization with residual confounders, and poor patient selection [36]. Take as an example RCTs which need a large sample size when the effect of the treatment in question is small [45]. In conducting such RCTs, AI can be used to potentially select those patients who are more likely to meet the eligibility criteria based on the AI's predicted algorithm. Utilizing AI also has the potential for more coherent implementation of the selection of patient subgroups with more advanced disease and greater control event rates. This approach has the potential to increase the expected effect size thus reducing the sample size, maximizing the statistical power and minimizing the time duration of the RCT compared to conventional recruitment approaches in RCTs [29]. AI may also be deployed in peer review or critical appraisal for integrity testing [52]. However, this approach depends on development of accurate tests for detecting flaws in integrity first [28].

Research integrity testing

There have been attempts to investigate published research for data integrity [7, 15, 30, 31]. Integrity tests are yet to be validated. An attempt to verify three post-publication tests failed to recommend any of them as valid [18]. Investing in the comprehensive validation of RCT integrity tests is a priority to improve the trustworthiness of the literature [28]. A meta-research process is needed to address issues in RCTs, including identifying, analyzing, proposing, and assessing solutions [17]. Meta-research by definition is a systematic study of the

research itself as the object of investigation, i.e., research on research methodology and types, aiming to improve on the process of conducting research. This framework's objective is to enhance the efficiency, credibility and quality standards of the scientific ecosystem. It is a tool that researchers can use to identify potential flaws in the scientific environment, such as publication bias and poor research design, at an early stage [17]. This process involves: conceptual development and theoretical argumentation, performing empirical studies to assess the presence and acuteness of the identified issues: developing and implementing likely solutions which may include new infrastructure, policy changes or instructional program development [17]; and, assessing the effectiveness of proposed solutions, either through controlled experiments or field implementation. In theory, feedback from later phases can inform earlier stages [17]. Meta-research on RCTs should evaluate the impact of including methodologists with experience in the statistical team, as well as on the data and safety monitoring boards. These can serve in the protection of large-scale multicentre clinical trials [2]. Going ahead with research on research in this arena within RCTs is a timely scientific challenge.

The value of rigorous observational studies

In EBM, questions may be answered by different study designs. If randomisation is not feasible, observational studies can be used. This is particularly valuable when there is no link between the intervention and the outcome that can confound the effect [6]. The main limitation of observational studies is that as the intervention is not randomised, confounding becomes a probable explanation for the observed differences between the exposed and unexposed groups [37]. However, if randomisation is not ethical or feasible, observational studies become the only viable design option to help answer important clinical questions. Adjusting for the connection between the allocation of the intervention and the patient prognosis requires advanced statistical methods that make adjustments for potential confounding. A review found “no significant difference in point estimates across heterogeneity, pharmacological intervention, or propensity score adjustment subgroups” between observational studies and RCTs [4]. The propensity score is used in the statistical analysis of observational data to estimate the effect considering the covariates available in the dataset [10]. These adjustments may be adequate to disconnect the covariates from the effect estimated [41]. Mindful interpretation of the results from rigorous observational studies will help EBM. Evidence grading systems permit the upgrading of observational results when generating recommendations [26].

Conclusion

EBM is grounded in high-quality research evidence, primarily from RCTs, to generate trustworthy recommendations. However, EBM's credibility may be jeopardized if the focus shifts away from RCTs toward observational data without sound justification. While observational studies have utility, RCTs remain the gold standard for robust inference in clinical effectiveness research. Rather than abandoning RCTs, efforts must enhance their relevance, improve their reproducibility, and promote the transparency of their publications through design-specific research integrity frameworks. RCTs remain the cornerstone of progress when executed ethically and rigorously. Renewed commitment to their evidentiary value, paired with quality control enhancements and improved relevance, will enable RCTs to provide the robust evidence required for clinical practice and health policy decision-making.

Abbreviations

RCT	Randomized Control Trials
EBM	Evidence-Based Medicine
AI	Artificial Intelligence

Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s43043-023-00161-7>.

Additional file 1.

Acknowledgements

KSK is a distinguished investigator at the University of Granada funded by the Beatriz Galindo (senior modality) programme of the Spanish Ministry of Education.

Data availability statement

Not applicable.

Authors' contributions

Writing original draft: M.F.K.S.K.; Review and editing: F.A.B, M.F.K.S.K.; Supervision: Y.K, BW, K.S.K and A.B.C.

Funding

None.

Availability of data and materials

Data sharing is not applicable to this article as no datasets were generated or analysed during the current study.

Declarations

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests. ICMJE forms are attached as [supplementary material](#).

Author details

¹Department of Preventative Medicine and Public Health, University of Granada, Granada, Spain. ²IbnSina, Banon Amshaj and Qena IVF Centres, Sohag, Egypt. ³Beginnings Assisted Conception Unit, Epsom and St Helier University Hospitals, London, UK. ⁴Institute for Clinical Trials and Methodology, University College London, London, UK. ⁵Guy's & St Thomas' Hospital Foundation Trust, London, UK.

Received: 16 October 2023 Accepted: 11 December 2023

Published online: 02 January 2024

References

- Afnan MAM, Khan KS, Mol BW (2020) Generating translatable evidence to improve patient care: the contribution of human factors. *Reprod Biomed Online* 41(3):353–356. <https://doi.org/10.1016/j.rbmo.2020.04.025>
- Agarwal A, Ioannidis JPA (2019) 'PREDIMED trial of Mediterranean diet: retracted, republished, still trusted?' *BMJ*:j341. <https://doi.org/10.1136/bmj.j341>
- Al Wattar BH, Zamora J, Khan KS (2017) Informing treatment decisions through meta-analysis: to network or not? *Evid Based Med* 22(1):12–15. <https://doi.org/10.1136/ebmed-2016-110599>
- Anglemeyer A, Horvath HT, Bero L (2014) 'Healthcare outcomes assessed with observational study designs compared with those assessed in randomized trials'. *Cochrane Database Syst Rev*. 2014(4). <https://doi.org/10.1002/14651858.MR000034.pub2>. Edited by Cochrane Methodology Review Group
- Baker M (2016) 1,500 scientists lift the lid on reproducibility. *Nature* 533(7604):452–454. <https://doi.org/10.1038/533452a>
- Bosdriesz JR et al (2020) Evidence-based medicine—When observational studies are better than randomized controlled trials. *Nephrology* 25(10):737–743. <https://doi.org/10.1111/nep.13742>
- Carlisle JB, Loadsman JA (2017) Evidence for non-random sampling in randomised, controlled trials by Yuhji Saitoh. *Anaesthesia* 72(1):17–27. <https://doi.org/10.1111/anae.13650>
- Djulbegovic B, Guyatt GH (2017) Progress in evidence-based medicine: a quarter century on. *Lancet* 390(10092):415–423. [https://doi.org/10.1016/S0140-6736\(16\)31592-6](https://doi.org/10.1016/S0140-6736(16)31592-6)
- Ebrahim S et al (2016) Randomized trials are frequently fragmented in multiple secondary publications. *J Clin Epidemiol* 79:130–139. <https://doi.org/10.1016/j.jclinepi.2016.05.016>
- Ebrahim Valojerdi A, Janani L (2018) A brief guide to propensity score analysis. *Med J Islam Repub Iran*:717–720. <https://doi.org/10.14196/mjiri.32.122>
- Frieden TR (2017) Evidence for health decision making — Beyond randomized, controlled trials. *N Engl J Med* 377(5):465–475. <https://doi.org/10.1056/NEJMra1614394>. Edited by J.M. Drazen et al
- Gall T, Ioannidis JPA, Maniatis Z (2017) The credibility crisis in research: Can economics tools help? *PLoS Biol* 15(4):e2001846. <https://doi.org/10.1371/journal.pbio.2001846>
- Gelman A (2018) Don't characterize replications as successes or failures. *Behav Brain Sci* 41:e128. <https://doi.org/10.1017/S0140525X18000638>
- Glued LL (2006) Bias in clinical intervention research. *Am J Epidemiol* 163(6):493–501. <https://doi.org/10.1093/aje/kwj069>
- Grey A et al (2020) Check for publication integrity before misconduct. *Nature* 577(7789):167–169. <https://doi.org/10.1038/d41586-019-03959-6>
- Grimes DR, Bauch CT, Ioannidis JPA (2018) Modelling science trustworthiness under publish or perish pressure. *R Soc Open Sci* 5(1):171511. <https://doi.org/10.1098/rsos.171511>
- Hardwicke TE et al (2020) Calibrating the scientific ecosystem through meta-research. *Annu Rev Stat Appl* 7(1):11–37. <https://doi.org/10.1146/annurev-statistics-031219-041104>
- Hartgerink CHJ (2019) Contributions towards understanding and building sustainable science. preprint. <https://doi.org/10.31237/osf.io/4wtpc>
- Haug CJ (2017) Whose data are they anyway? Can a patient perspective advance the data-sharing debate? *N Engl J Med* 376(23):2203–2205. <https://doi.org/10.1056/NEJMp1704485>

20. Heneghan C et al (2017) Evidence based medicine manifesto for better healthcare. *BMJ* 354(8000):e2973. <https://doi.org/10.1136/bmj.g2973>
21. Ioannidis JP, Caplan AL, Dal-Ré R (2017) Outcome reporting bias in clinical trials: why monitoring matters. *BMJ* 354(8000):e2973. <https://doi.org/10.1136/bmj.g2973>
22. Ioannidis JPA (2014) Clinical trials: what a waste. *BMJ* 349(8000):e2973. <https://doi.org/10.1136/bmj.g2973>
23. Ioannidis JPA (2017) Does evidence-based hearsay determine the use of medical treatments? *Soc Sci Med* 177:256–258. <https://doi.org/10.1016/j.socscimed.2017.02.004>
24. Kalager M, Adami H-O, Bretthauer M (2016) Recognizing data generation. *N Engl J Med* 374(19):1898–1898. <https://doi.org/10.1056/NEJMc1603789>
25. Khan K, On behalf of Chief Editors of Journals participating in The CROWN Initiative listed at the end of this article (2015) The CROWN Initiative: journal editors invite researchers to develop core outcomes in women's health. *Women's Midlife Health* 1(1):4, s40695–015–0001-z. <https://doi.org/10.1186/s40695-015-0001-z>
26. Khan KS et al (2011) Making GRADE accessible: a proposal for graphic display of evidence quality assessments. *Evid Based Med* 16(3):65–69. <https://doi.org/10.1136/ebm0005>
27. Khan KS, Cairo Consensus Group on Research Integrity (2023) International multi-stakeholder consensus statement on clinical trial integrity. *BJOG* 147(1):1471–1471. <https://doi.org/10.1111/1471-0528.14711>
28. Khan KS, Fawzy M, Chien PFW (2023) Integrity of randomized clinical trials: performance of integrity tests and checklists requires assessment. *Int J Gynecol Obstet: ijgo*. 14837. <https://doi.org/10.1002/ijgo.14837>
29. Lee CS, Lee AY (2020) How artificial intelligence can transform randomized controlled trials. *Transl Vis Sci Technol* 9(2):9. <https://doi.org/10.1167/tvst.9.2.9>
30. Li W et al (2020) Integrity of randomized controlled trials: challenges and solutions. *Fertil Steril* 113(6):1113–1119. <https://doi.org/10.1016/j.fertnstert.2020.04.018>
31. Li W, Bordewijk EM, Mol BW (2021) Assessing research misconduct in randomized controlled trials. *Obstet Gynecol* 138(3):338–347. <https://doi.org/10.1097/AOG.0000000000004513>
32. MacCoun RJ (2018) Enhancing research credibility when replication is not feasible. *Behav Brain Sci* 41:e142. <https://doi.org/10.1017/S0140525X18000778>
33. Moosapour H et al (2021) The rationale behind systematic reviews in clinical medicine: a conceptual framework. *J Diabetes Metab Disord* 20(1):919–929. <https://doi.org/10.1007/s40200-021-00773-8>
34. Moss N et al (2017) Involving pregnant women, mothers and members of the public to improve the quality of women's health research. *BJOG* 124(3):362–365. <https://doi.org/10.1111/1471-0528.14419>
35. Natafji N et al (2019) Patients' early engagement in research proposal development (PEER-PD): patients guiding the proposal writing. *J Comp Eff Res* 8(6):441–453. <https://doi.org/10.2217/ceer-2018-0129>
36. Nichol AD, Bailey M, Cooper DJ (2010) Challenging issues in randomised controlled trials. *Injury* 41:S20–S23. <https://doi.org/10.1016/j.injury.2010.03.033>
37. Noordzij M et al (2009) Study designs in clinical research. *Nephron Clin Pract* 113(3):c218–c221. <https://doi.org/10.1159/000235610>
38. Núñez-Núñez M et al (2023) Assessing the integrity of clinical trials included in evidence syntheses. *Int J Environ Res Public Health* 20(12):6138. <https://doi.org/10.3390/ijerph20126138>
39. Platt R, Ramsberg J (2016) Challenges for sharing data from embedded research. *N Engl J Med* 374(19):1897–1897. <https://doi.org/10.1056/NEJMc1602016>
40. Rockhold F, Nisen P, Freeman A (2016) Data sharing at a crossroads. *N Engl J Med* 375(12):1115–1117. <https://doi.org/10.1056/NEJMp1608086>
41. Rosenbaum PR, Rubin DB (1983) The central role of the propensity score in observational studies for causal effects. *Biometrika* 70(1):41–55. <https://doi.org/10.1093/biomet/70.1.41>
42. Schmucker C et al (2014) Extent of non-publication in cohorts of studies approved by research ethics committees or included in trial registries. *PLoS ONE* 9(12):e114023. <https://doi.org/10.1371/journal.pone.0114023>
43. Simmons JP, Nelson LD, Simonsohn U (2011) False-positive psychology: undisclosed flexibility in data collection and analysis allows presenting anything as significant. *Psychol Sci* 22(11):1359–1366. <https://doi.org/10.1177/0956797611417632>
44. Stegert M et al (2016) An analysis of protocols and publications suggested that most discontinuations of clinical trials were not based on unplanned interim analyses or stopping rules. *J Clin Epidemiol* 69:152–160. <https://doi.org/10.1016/j.jclinepi.2015.05.023>
45. Stolberg HO, Norman G, Trop I (2004) Randomized controlled trials. *Am J Roentgenol* 183(6):1539–1544. <https://doi.org/10.2214/ajr.183.6.01831539>
46. Taichman DB et al (2016) Sharing clinical trial data: a proposal from the International Committee of Medical Journal Editors. *PLoS Med* 13(1):e1001950. <https://doi.org/10.1371/journal.pmed.1001950>
47. Taichman DB et al (2017) Data sharing statements for clinical trials. *BMJ* 354(8000):e2973. <https://doi.org/10.1136/bmj.g2973>
48. Tenny S, Varacallo M (2023) Evidence based medicine!. In: StatPearls. Treasure Island (FL): StatPearls Publishing. Available at: <http://www.ncbi.nlm.nih.gov/books/NBK470182/>. Accessed: 26 Sept 2023
49. Thadhani R (2006) Formal trials versus observational studies. In: Mehta A, Beck M, Sunder-Plassmann G (eds) *Fabry Disease: Perspectives from 5 Years of FOS*. Oxford PharmaGenesis, Oxford. Available at: <http://www.ncbi.nlm.nih.gov/books/NBK11597/>. Accessed: 26 Sept 2023
50. The International Consortium of Investigators for Fairness in Trial Data Sharing (2016) Toward fairness in data sharing. *N Engl J Med* 375(5):405–407. <https://doi.org/10.1056/NEJMp1605654>
51. Van Noorden R (2023) Medicine is plagued by untrustworthy clinical trials. How many studies are faked or flawed? *Nature* 619(7970):454–458. <https://doi.org/10.1038/d41586-023-02299-w>
52. Weissgerber T et al (2021) Automated screening of COVID-19 preprints: can we help authors to improve transparency and reproducibility? *Nat Med* 27(1):6–7. <https://doi.org/10.1038/s41591-020-01203-7>
53. Williams RJ et al (2015) Terminated trials in the ClinicalTrials.gov results database: evaluation of availability of primary outcome data and reasons for termination. *PLoS ONE* 10(5):e0127242. <https://doi.org/10.1371/journal.pone.0127242>
54. Zwaan RA et al (2018) Making replication mainstream. *Behav Brain Sci* 41:e120. <https://doi.org/10.1017/S0140525X17001972>

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Submit your manuscript to a SpringerOpen[®] journal and benefit from:

- Convenient online submission
- Rigorous peer review
- Open access: articles freely available online
- High visibility within the field
- Retaining the copyright to your article

Submit your next manuscript at ► [springeropen.com](https://www.springeropen.com)