

Journal Pre-proof



Research Integrity in Clinical Epidemiology: Core Concepts and Contemporary Challenges

G. Gopalakrishna

PII: S0895-4356(25)00437-8

DOI: <https://doi.org/10.1016/j.jclinepi.2025.112104>

Reference: JCE 112104

To appear in: *Journal of Clinical Epidemiology*

Received Date: 19 June 2025

Revised Date: 4 December 2025

Accepted Date: 8 December 2025

Please cite this article as: Gopalakrishna G, Research Integrity in Clinical Epidemiology: Core Concepts and Contemporary Challenges, *Journal of Clinical Epidemiology* (2026), doi: <https://doi.org/10.1016/j.jclinepi.2025.112104>.

This is a PDF of an article that has undergone enhancements after acceptance, such as the addition of a cover page and metadata, and formatting for readability. This version will undergo additional copyediting, typesetting and review before it is published in its final form. As such, this version is no longer the Accepted Manuscript, but it is not yet the definitive Version of Record; we are providing this early version to give early visibility of the article. Please note that Elsevier's sharing policy for the Published Journal Article applies to this version, see: <https://www.elsevier.com/about/policies-and-standards/sharing#4-published-journal-article>. Please also note that, during the production process, errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

© 2025 Published by Elsevier Inc.

Research Integrity in Clinical Epidemiology: Core Concepts and Contemporary Challenges

Gopalakrishna, G¹

¹ Department of Epidemiology, Faculty of Health Medicine & Life Sciences Maastricht University

Conflict of Interest: Dr Gopalakrishna is also employed by Elsevier B.V., where she leads research integrity investigations & certification efforts of journals indexed in Scopus

Abstract

Research integrity is foundational to clinical epidemiology, particularly in an increasingly transparent scientific landscape. As the field of research integrity navigates the evolving demands of open science, data transparency, and collaborative research, it must also grapple with the influence systemic challenges, such as research fairness, diversity, equity, and inclusion, have on research quality. This *Key Concepts* article provides a concise overview of research integrity for clinical epidemiologists. It summarizes key principles in research integrity and the emerging overlap with open science, research fairness, diversity, equity, and inclusion in upholding the integrity of epidemiological research. Practical guidance is provided at every stage of the research lifecycle—from preparing a research proposal to study protocol development and data collection to publication and dissemination of research findings. It addresses how these overlapping concepts demonstrate that research integrity is not merely about methodological rigor, but is a scientific imperative that requires a broader definition of research integrity to produce high-quality research that is responsible and inclusive.

Word count 162

Introduction

Clinical epidemiology plays a pivotal role in generating evidence that informs clinical decision-making, policy, and public health interventions [1]. For epidemiologists whose work directly influences patient care, clinical guidelines and public health policies, lapses in research integrity can have serious human health consequences. The field is not immune to broader concerns that have eroded trust in biomedical research, such as research misconduct and misbehavior commonly known as questionable research behaviors [2]. New threats to research integrity have emerged, such as predatory journals and irreproducibility of research findings [2]. These issues, together with a growing awareness of systemic issues in research incentives and structural inequities, have led to a global push for reform grounded in transparency, accountability, fairness, diversity, equity, and inclusion [3,4]. This shift expands the traditional focus on misconduct and questionable practices to a more holistic view that includes open science and systematic factors around access to conducting science and how it is rewarded. This article synthesizes some of the classical and emerging concepts in research integrity and offers actionable steps for epidemiologists to strengthen the rigor, openness, fairness, and inclusiveness of their research.

Core Concepts in Research Integrity

Research integrity is typically defined as the adherence to ethical principles and professional standards centered on principles of reliability, honesty, respect, and accountability [2]. Researchers, institutions, funders, and publishers collectively share responsibility for upholding these standards throughout the research lifecycle. Core concepts include falsification, fabrication, and plagiarism (FFP), as well as questionable research practices (QRPs) [2,5]. FFP is considered the most severe form of research transgression, also known as research misconduct. This constitutes falsification, defined as manipulating research materials or data such that results are misrepresented, fabrication, which is about making up data or results, and plagiarism, which is using another's ideas, data, or words without proper attribution. With the advent of artificial intelligence (AI) tools like large language models, the potential for undetected data and text manipulation has grown, raising new concerns around research misconduct, making research integrity a topic of growing importance [6,7].

Beyond FFP, questionable research practices (QRPs) are increasingly recognized as widespread and damaging. These fall in the gray zone between responsible conduct and misconduct [5,8]. In biomedical research, QRPs include, but are not limited to selective reporting, p-hacking (manipulating data to achieve statistical significance), and HARKing (hypothesizing after results are known) [5]. These practices differ by discipline and are far more common than FFP, with some studies estimating that one in two researchers frequently engage in QRPs [8]. Associative factors driving such misbehaviors span from systemic pressures around publication pressure, competitiveness in grant acquisition to poor supervision and mentoring [8]. Table 1 details some of the relevant QRPs in biomedicine at each stage of the research life cycle.

Emerging Challenges & Topics in Research Integrity

Predatory Journals & Paper Mills

A new wave of threats to research integrity comes from predatory journals and paper mills [9], which bypass quality control to publish questionable or fabricated work in exchange for payment. These entities often sell authorships and offer minimal peer review, threatening the credibility of scientific publishing. Such practices can result in misleading clinical recommendations when unverified findings are cited in guidelines. They also compromise systematic reviews and meta-analyses, distort perceptions about treatment efficacy, and damage the reputation of researchers, particularly early-career academics or those from underfunded institutions who unknowingly get co-opted into such organisations [9]. To counter these risks, the scientific community has developed resources to identify predatory practices and paper mills [9,10], although some, such as predatory journal lists, remain controversial. An added challenge to fighting this threat is the rapidly changing ways in which these unscrupulous entities operate, increasingly in the form of syndicated criminal organizations, making them harder to track and detect. While such resources are critical in raising awareness and promoting vigilance, there is an urgent need for substantive, collective action, potentially in the form of legislation against these new forms of scientific fraud akin to organized crime.

Open Science, Research Fairness & Diversity, Equity & Inclusion (DEI)

In the face of these new threats, open science offers a promising strategy to improve research integrity by fostering transparency and accountability, particularly as the concerns around failure to reproduce research, especially in biomedicine, are growing. These concerns have been attributed to lapses in transparency around detailed study protocols, opaque analysis plans, and inaccessible data and code behind one's analyses [11]. Practices such as open data and open code, protocol preregistration (i.e., sharing study protocol and full data analysis plan on open registries such as the Open Science Framework (OSF)), and adherence to FAIR (Findable, Accessible, Interoperable, Reusable) data principles enhance reproducibility and reduce bias [11]. Registered Reports are specific journal articles that involve formal peer review of the preregistered study protocol before the research is conducted, with a commitment to publication regardless of the results., have the added benefit of reducing publication bias and improving research quality as such studies go through the peer review process before they start [12]. In doing so, issues in a study design and/or analysis method are detected early.

Yet, in epidemiological research, both practices have led to much debate and are not the norm [13], especially the (pre-)registration of certain epidemiological study designs such as retrospective cohorts that use pre-existing (public) datasets [13]. In such instances, pre-registration may not offer the same protection as in prospective study designs owing to the fact that in such instances, researchers may (knowingly or unknowingly) be influenced by prior results, reducing the strength of preregistration as a safeguard against HARKing or selective reporting. Recognizing this limitation is essential in understanding an important limitation of preregistration and registered reports in epidemiological research. Other reasons for the low uptake include such practices being considered bureaucratic, time-consuming, restrictive and causing potential delay to studies. Limited incentives of journals in epidemiology (with a few

exceptions) which rarely require preregistration, mean uptake is voluntary [13]. However, the evidence that pre-registration and registered reports can be beneficial towards preventing selective outcome reporting, HARKing, publication bias and improving reproducibility of research in most instances is clear [12].

Table 1 outlines a number of these open science practices that a researcher can implement at each stage of the research lifecycle. These approaches can mitigate QRPs and the threat of fabricated research by promoting transparency and accountability and ultimately reproducibility in science.

While open science is rapidly gaining traction in responsible conduct of research through increasing global calls for research reform and conduct to include open science , it also brings with it challenges. Concerns on research integrity are not new, however the increasing use of AI in research pose added and new concerns to safeguarding the integrity of the scientific record [6,7]. Fabrication of non-existent data and results, as well as plagiarizing open datasets and previous work to support fabricated primary outcomes using AI have been reported [7]. For epidemiological research, this can mean systematic reviews based on fabricated findings, and clinical practice guidelines based on fraudulent research with serious implications for patient health. Urgent and collective effort from the scientific community is needed to ensure that the promotion of open science practices, such as open data sharing, do not lead to easier means for research fraud [7].

Another pitfall often overlooked in the push for open science is that its policies and standards are driven mostly by the West, which some argue are predicated on a narrow understanding of openness, accessibility, and objectivity [14]. Consequently, they do not effectively capture the ethos of Open Science, particularly its goal of making science more collaborative, inclusive, and socially engaged [14]. For instance, Open Access (OA) is often promoted to democratize knowledge, yet its predominant model, Gold OA requires payment of hefty article processing charges (APCs). Such practices in open access publishing reinforce global inequalities by privileging well-funded researchers.. These APCs can exclude under-resourced scholars, early-career researchers, and those from developing nations, who may face additional challenges such as predatory journals or stigmatizing fee-waiver processes [13]. Similarly, open data initiatives, while offering benefits like transparency, often overlook the contextual disparities in research environments across disciplines and regions. Factors such as data infrastructure, literacy, and funding availability shape who can effectively participate in data sharing and reuse, often reinforcing existing inequities [14,15]. Both OA and open data, when pursued without attention to broader structural dynamics, risk empowering already advantaged researchers while marginalizing others, ultimately undermining the equity and inclusive ideals of Open Science.

Under this light, there is increasing acknowledgement that research fairness, diversity, equity, and inclusion (DEI) are as essential to research integrity as methodological rigor is [4,15]. These concepts address both who benefits from research and who contributes to its creation. Persistent inequities in epidemiology—such as underrepresentation of marginalized populations in studies and disparities in authorship—highlight power imbalances both within research teams and between research teams from high-income countries (HICs) and those in low- and middle-income countries (LMICs) [4,14].

It is therefore imperative that the definition, promotion, and conduct of practices that support research integrity also address biases, privileges, and inequities that have been rooted in the system of science. Relegating these as social justice or global health concerns does injustice to the fact that epidemiology intersects with societal health and well-being in so many ways, akin to global health. Table 1 provides some tangible ways researchers can address research fairness and DEI concerns in their research through practical actions at each stage of the research process. However, it is by no means meant to be exhaustive. It is beyond this article's scope to thoroughly address the need for awareness on epistemic injustice and the coloniality of the research ecosystem, especially in fields like epidemiology, which intertwine so intricately with societal health. The essential reading list mentions scholars who have addressed these issues and their relevance to research integrity.

Conclusion

This article outlines classical and emerging dimensions of research integrity, from FFP and QRPs to emerging challenges around scientific fraud, open science, research fairness, and DEI. For clinical epidemiologists, improving research integrity should go beyond methodological rigor. It requires active engagement with values of transparency, accountability through open science practices and those that support research fairness, diversity, equity and inclusiveness throughout the research process.

(word count 1697)

References

1. Lash, T. L., VanderWeele, T. J., Haneuse, S., & Rothman, K. J. (2020). Modern epidemiology (4th ed.). Wolters Kluwer.
2. Armond, A. C. V., et al. (2022). Research integrity definitions and challenges. *Journal of Clinical Epidemiology*, 171, 111367.
3. Moher, D., Bouter, L., Kleinert, S., Glasziou, P., Sham, M. H., Barbour, V., et al. (2020). The Hong Kong Principles for assessing researchers: Fostering research integrity. *PLOS Biology*, 18(7), e3000737. <https://doi.org/10.1371/journal.pbio.3000737>
4. Horn, L., et al. (2023). The Cape Town Statement on fairness, equity and diversity in research. *Nature*, 615, 790–793. <https://doi.org/10.1038/d41586-023-00867-6>
5. Bouter, L. M., Tijdink, J., Axelsen, N., Martinson, B. C., & ter Riet, G. (2016). Ranking major and minor research misbehaviors: Results from a survey among participants of four World Conferences on Research Integrity. *Research Integrity and Peer Review*, 1(1), 17. <https://doi.org/10.1186/s41073-016-0024-5>
6. Callaway, E. (2025). The scientific publishing system is at a tipping point. *Nature*. <https://www.nature.com/articles/d41586-025-01592-0>
7. Suchak, T., Aliu, A. E., Harrison, C., Zwiggelaar, R., Geifman, N., & Spick, M. (2025). Explosion of formulaic research articles, including inappropriate study designs and false discoveries, based on the NHANES US national health

- database. *PLOS Biology*, 23(5), e3003152.
<https://doi.org/10.1371/journal.pbio.3003152>
8. Gopalakrishna, G., ter Riet, G., Vink, G., Stoop, I., Wicherts, J. M., & Bouter, L. M. (2022). Prevalence of questionable research practices, research misconduct and their potential explanatory factors: A survey among academic researchers in The Netherlands. *PLOS ONE*, 17(2), e0263023.
<https://doi.org/10.1371/journal.pone.0263023>
 9. Parker, L., et al. (2024). Paper mill challenges: Past, present, and future. *Journal of Clinical Epidemiology*, 176, 111549.
 10. City University of Hong Kong. (n.d.). *Predatory publishers*.
https://libguides.library.cityu.edu.hk/oa_gold/predatory (accessed May 14, 2025)
 11. Munafò, M., Nosek, B., Bishop, D., Button, K., Chambers, C., Percie du Sert, N., Simonsohn, U., Wagenmakers, E.-J., Ware, J. J., & Ioannidis, J. P. A. (2017). A manifesto for reproducible science. *Nature Human Behaviour*, 1, 0021.
<https://doi.org/10.1038/s41562-016-0021>
 12. Chambers, C. D., & Tzavella, L. (2022). The past, present and future of Registered Reports. *Nature Human Behaviour*, 6, 29–42.
<https://doi.org/10.1038/s41562-021-01193-7>
 13. Viergever, R. F., Karam, G., Reis, A., & Ghersi, D. (2016). Registration practices for observational studies on ClinicalTrials.gov indicated low adherence. *Journal of Clinical Epidemiology*, 70, 176–182.
<https://doi.org/10.1016/j.jclinepi.2015.09.009>
 14. Ross-Hellauer, T., Reichmann, S., Cole, N. L., Fessl, A., Klebel, T., & Pontika, N. (2022). Dynamics of cumulative advantage and threats to equity in open science: A scoping review. *Royal Society Open Science*, 9(2), 211032.
<http://doi.org/10.1098/rsos.211032>
 15. Alba, S., Lenglet, A., Verdonck, K., et al. (2020). Bridging research integrity and global health epidemiology (BRIDGE) guidelines: Explanation and elaboration. *BMJ Global Health*, 5, e003237. <https://doi.org/10.1136/bmjgh-2020-003237>

Suggestions for Further Reading

1. Committee on Publication Ethics (COPE). (2023, May). *Predatory behaviour in publication ethics*. <https://publicationethics.org/news-opinion/predatory-behaviour-publication-ethics> (Accessed May 24, 2025)

This editorial by the Committee on Publication Ethics explores the evolving definitions and manifestations of predatory behavior in scholarly publishing

2. Candal-Pedreira, C., Ross, J., Ruano-Ravina, A., Egilman, D., Fernandez, E., & Pérez-Ríos, M. (2022). Retracted papers originating from paper mills: Cross sectional study. *BMJ*, 379, e071517. <https://doi.org/10.1136/bmj-2022-071517>

This cross-sectional study investigates retracted papers linked to paper mills, revealing their significant impact on the scientific literature.

3. Kohrs, F. E., Auer, S., Bannach-Brown, A., & Fiedler, S. (2023). Eleven strategies for making reproducible research and open science training the norm at research institutions. *eLife*, 12, e89736. <https://doi.org/10.7554/eLife.89736>

This paper presents eleven institutional strategies to normalize reproducible research and open science training, promoting a culture of transparency and rigor in academia.

4. Sabel BA, Larhammar D. Reformation of science publishing: the Stockholm Declaration. *R. Soc. Open Sci.* 2025 Nov 5;12(11):251805. doi:10.1098/rsos.251805

The Stockholm Declaration warns that the rise of “paper mills” and AI-generated fake publications is eroding trust in science, and outlines set of principles and call for actions of good publishing practice to be implemented worldwide

5. Labib, K. (2025). Research integrity and research fairness: Harmonious or in conflict? *Canadian Journal of Bioethics / Revue canadienne de bioéthique*, 8(3), 44–54. <https://doi.org/10.7202/1118901ar>

This paper argues why dominant research integrity perspectives do not adequately address research fairness and, in doing so, ignore the importance of anticolonial views about research, thereby perpetuating coloniality in research.

Table 1 Overview of QRPs, Open Science and Research Fairness Practices to consider at each stage of the Research Lifecycle

Stage of Research Lifecycle	QRPs that can affect epidemiological research[^]	Open Science Practices*	Research Fairness & DEI Practices^{^^}
Research Formulation	<ul style="list-style-type: none"> • Not building on previous research • Not addressing societal relevance of research • Ignoring risks to participants, environment and/or society 	<ul style="list-style-type: none"> • Pre-register grant applications on open registers especially if the grant is funded 	<ul style="list-style-type: none"> • Plan research in partnership with local researchers • Establish knowledge gap through exploratory needs assessment involving local community and end users • Adopt practices that support fairness and equity between (HIC & LMIC) researchers, e.g. fair distribution of resources and roles in the team
Protocol Development	<ul style="list-style-type: none"> • Use of inappropriate study designs • Inappropriate attention to equipment, skills or expertise needed • Make significant changes to a study design after study start without proper disclosure • Underpowered studies 	<ul style="list-style-type: none"> • Preregister study protocols and full data analyses plans including revisions to them in open registries (e.g., ClinicalTrials.gov, Open Science Framework) • Submit study protocols and analysis plans to Registered Reports for peer review 	<ul style="list-style-type: none"> • Ensure study protocol addresses local context, needs & respects indigenous knowledge production methods • Seek ethics approval including locally • Consider equity dimensions in analyses as suggested by the PROGRESS framework**
Study Conduct & Data Collection	<ul style="list-style-type: none"> • Collecting more data than planned or stopping data collection prematurely • Inadequate handling and storage of data that is incompliant with regulations • Datasets that are not FAIR with no quality assurance 	<ul style="list-style-type: none"> • Update changes to protocol and data analysis plan in a timely manner 	<ul style="list-style-type: none"> • Use valid and reliable instruments which are locally and culturally appropriate • Collect data in a safe, and culturally appropriate manner • Select data collection staff with technical and cultural knowledge • Provide training and support
Data Management & Analysis	<ul style="list-style-type: none"> • Conducting data-driven hypotheses testing without disclosure (HARKing) 	<ul style="list-style-type: none"> • Fully annotate all data cleaning and processing steps 	<ul style="list-style-type: none"> • Define levels of anonymization and privacy protection & ensure access rights in line with national and

	<ul style="list-style-type: none"> Selectively deleting or modifying data after initial analyses Not reporting all protocol stated analyses Performing data analyses not in the protocol without disclosure p-hacking (i.e. manipulating data and statistics to provide a favorable outcome) 	<ul style="list-style-type: none"> Enable computational reproducibility by sharing full codebook on open registers Ensure datasets are as open as possible and closed as needed Adhere to FAIR principles 	international frameworks <ul style="list-style-type: none"> Data sharing should not unfairly disadvantage LMIC collaborators Address power imbalances and focus on capacity building so local researchers are not only relegated data collection roles
Reporting & Publication	<ul style="list-style-type: none"> Concealing results that contradict earlier findings or convictions Not reporting full study methods Selective citation Spreading study results over more publications than needed (salami slicing) Not declaring conflicts of interest (financial and non-financial) Adding or omitting potential authors unfairly (e.g. gift authorship) 	<ul style="list-style-type: none"> Publish open access Share manuscript as preprint on an open registry Conduct, when possible, open peer review Share full grant application, study protocols and full analyses plans in an open registry 	<ul style="list-style-type: none"> Use inclusive authorship practices that reflect real contributions (e.g., CRediT taxonomy^{&}) Report limitations honestly, including negative or null results
Dissemination (incl. post-publication)	<ul style="list-style-type: none"> Deliberately communicate findings inaccurately in the media or during presentations Not correcting or retracting publications if deemed necessary 	<ul style="list-style-type: none"> Engage in citizen science so findings are shared and communicated to the public in understandable ways 	<ul style="list-style-type: none"> Develop dissemination plans with and for affected populations and end-users Ensure collaborations result in appropriate benefit-sharing and recognition

[^]QRPs reflect a random selection adapted from reference 3; ^{*}Open Science Practices are adapted from <https://www.cos.io/open-science> accessed on 14 May 2025; ^{^^}Research Fairness & DEI Practices adapted from references 4 & 14 ^{**}O'Neill J et al *J Clin Epidemiol* 2014;67:56–64; [&]CRediT refers to Contributor Role Taxonomy; <https://credit.niso.org/> accessed on 24 May 2025

Declaration of interests

- The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.
- The authors declare the following financial interests/personal relationships which may be considered as potential competing interests:

Gowri Gopalakrishna reports a relationship with Journal of Clinical Epidemiology that includes: board membership. Given my role as Editorial Board member to the Journal of Clinical Epidemiology, I had no involvement in the peer review of this article and had no access to information regarding its peer review. Full responsibility for the editorial process for this article was delegated to another journal editor. If there are other authors, they declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Gowri Gopalakrishna is also employed by Elsevier B.V. The Netherlands. Views expressed are solely hers and do not reflect those of Elsevier B.V. The Netherlands