

Highlights

Two Decades of Scientific Misconduct in India: Retraction Reasons and Journal Quality among Inter-country and Intra-country Institutional Collaboration

Kiran Sharma

- The increasing number of retractions is concerning, while the decreasing duration of retractions suggests that journals are taking active measures against misconduct.
- Approximately 60% of retractions stem from private institutions. Fake peer review is recognized as the primary cause of scientific misconduct in Indian research carried out by private institutions and colleges.
- Retractions due to data integrity predominantly occur in public and medical institutions, while retractions due to plagiarism occur in journals not indexed in Scopus and conference proceedings.
- Institutional collaborations indicate that 80% of retractions originate domestically, while the remaining 20% stem from international collaborations.
- In inter-country collaborations, one-third of the total retractions originate from the top two journal quartiles, whereas in intra-country collaborations half of the total retractions originate from the top two journal quartiles, i.e. Q1 and Q2.
- In intra-country collaborations, the top journals, based on the number of retractions, fall within the Q2 and Q4 categories, whereas in inter-country collaborations, they are situated in the Q1 category

Two Decades of Scientific Misconduct in India: Retraction Reasons and Journal Quality among Inter-country and Intra-country Institutional Collaboration

Kiran Sharma^{a,*}

^a*School of Engineering and Technology, BML Munjal University, Gurugram, Haryana-122413, India*

ARTICLE INFO

Keywords:

Retracted Publication
Indian Affiliation
Institutional Collaboration
Journal Quartile
Retraction Reasons

ABSTRACT

Research stands as a pivotal factor in propelling the progress of any nation forward. However, if tainted by misconduct, it poses a significant threat to the nation's development. Recognizing the importance of genuine research and understanding the ramifications of scientific malpractice are essential for grasping a nation's trajectory of growth. This study aims to scrutinize various cases of deliberate scientific misconduct by Indian researchers. A comprehensive analysis was conducted on 3,244 retracted publications sourced from the Retraction Watch database. The upward trend in retractions is alarming, although the decreasing duration of retractions indicates proactive measures by journals against misconduct. Approximately 60% of retractions stem from private institutions, with fake peer reviews identified as the primary cause of misconduct. This trend could be attributed to incentivizing publication quantity over quality in private institutions, potentially fostering unfair publishing practices. Retractions due to data integrity issues are predominantly observed in public and medical institutions, while retractions due to plagiarism occur in conference proceedings and non-Scopus-indexed journals. Examining retractions resulting from institutional collaborations reveals that 80% originate from within the country, with the remaining 20% being international collaborations. Among inter-country collaborations, one-third of retractions come from the top two journal quartiles, whereas, in intra-country collaborations, half of the retractions stem from Q1 and Q2 journals. In intra-country collaborations, 86% of retractions are from articles and conference proceedings, while in inter-country collaborations, 87% are solely from articles. Clinical studies retracted from intra-country collaborations are mostly from Q3 and Q4 journals, whereas in inter-country collaborations, they primarily come from Q1 journals. Regarding top journals by the number of retractions in intra-country collaborations, they belong to the Q2 and Q4 categories, whereas in inter-country collaborations, they are in Q1. This underscores the fact that retractions involving international collaborations tend to occur in high-quality journals, whereas those involving domestic collaborations are more commonly associated with journals of moderate to low quality.

1. Introduction

In academia, consistent publication stands as one of the most effective ways to demonstrate one's research capabilities. Successfully publishing research not only garners attention for scholars and their affiliated institutions but also has the potential to attract increased funding and propel individuals' careers within their respective fields. Moreover, many private institutes and universities offer financial incentives for each publication, while a high volume of publications can also enhance annual appraisals. Academic institutions often assess competence based on the number of publications attributed to an individual, with administrators increasingly relying on this metric in recruitment decisions. Scholars who publish infrequently or prioritize activities that do not result in publications may find themselves at a disadvantage when seeking appraisal. These various factors collectively contribute to the phenomenon known as "Publish or perish," a term initially coined by Coolidge in 1932, which has since become an unavoidable reality (Rawat and Meena, 2014).

Whether driven by the pursuit of monetary rewards or the pressure to publish for academic advancement, these factors have led authors to resort to unethical practices such as fake peer review, data falsification, plagiarism, and false authorship (Racimo, Galtier, De Herde, Bonn, Phillips, Guillemaud and Bourguet, 2022; Receveur, Bonfanti, d'Agata, Helmstetter, Moore, Oliveira, Petit-Cailleux, Rievers Borges, Schultz, Sexton et al., 2024). Such practices have

*Corresponding author

 kiransharma1187@gmail.com (K. Sharma)

ORCID(s):

gradually eroded the quality and authenticity of research work. Subsequently, as journals and ethical authors became aware of this trend, they began retracting such unfair and false studies. Currently, the quantity of published papers facing retraction is increasing day by day (Vuong, La, Hõ, Vuong and Ho, 2020). While retractions constitute only a tiny fraction of the overall published literature, the alarming rate at which retractions are escalating raises concerns. Retracted papers not only tarnish a researcher's career (Mongeon and Larivière, 2016; Sharma and Mukherjee, 2023) but also undermine the integrity of scholarly literature itself.

In pursuit of academic progress, the majority of Indian higher education institutions expect their faculty, encompassing both academics and researchers, to produce at least one paper in a journal indexed in Scopus or Web of Science. This expectation reflects the widely acknowledged "publish or perish" phenomenon (Rawat and Meena, 2014; Lee, 2014), where the sheer quantity of publications often outweighs considerations of their quality. Addressing this concern necessitates a shift in academic evaluations to prioritize the substance of publications over the prestige of the journals where they are published (Chaddah and Lakhotia, 2018). Moreover, it underscores that retracted papers stemming from international collaborations are predominantly associated with high-quality journals, while those originating from domestic collaborations tend to be from journals of middling to low quality.

Vuong et al. presented in their study that the frequency of retraction has boomed in the past 20 years and fake peer review accounts for the major reason for retraction (Vuong et al., 2020). They also stated that the Institute of Electrical and Electronics Engineers, Elsevier, and Springer account for nearly 60% of all retracted papers globally. Bar-ilan et al. have explored the temporal characteristics of retracted articles, including time of publication, years to retract, growth of post-retraction citations over time, and social media attention for articles retracted due to ethical misconduct, scientific distortion, and administrative error (Bar-Ilan and Halevi, 2018). They found that both citation counts and Mendeley reader counts continue to grow after retraction. Tang et al. found that publications with authors from elite universities are less likely to be retracted, which is particularly true for retractions due to falsification, fabrication, and plagiarism (Tang, Hu, Sui, Yang and Cao, 2020).

The current review processes organized by journals do not effectively ensure the ethical quality of publications (Bornmann and Mungra, 2011). One approach used to address this issue is article retraction. If a published article later proves to be unfit for publication due to particularly serious reasons, it may be retracted. However, given the gravity of this action, it must be exercised cautiously to prevent innocent individuals from being adversely affected. There are uncertainties regarding the classification of retraction reasons. Previous studies have presented differing perspectives on the rise in retractions attributed to errors or misconduct (Fanelli, 2013; Fang, Steen and Casadevall, 2012; Steen, 2011; Campos-Varela and Ruano-Raviña, 2019; Fang et al., 2012; Helgesson and Eriksson, 2015). COPE suggests that journal editors should consider retracting an article in any of the following four situations (Kleinert, 2009): (i) when there is evidence that the findings are unreliable, either due to misconduct or honest error; (ii) when the same findings have already been published without proper acknowledgment (e.g., without obtaining permission or providing a reference); (iii) when the article plagiarizes another published document; or (iv) when the article is based on unethical research (Wager, Barbour, Yentis, Kleinert et al., 2009). According to these COPE guidelines, an article could be retracted due to various reasons such as honest errors, misrepresentation, data manipulation, plagiarism (including self-plagiarism), text duplication, or failure to disclose conflicts of interest (Kleinert, 2009; Wager et al., 2009).

Considerable literature has examined scientific misconduct within the Indian research landscape. Sharma et al. (Sharma, Sharma, Reza, Inampudi and Dhamija, 2023) systematically analyzed reasons for retractions and their citations in the biomedical field for papers affiliated with India. They found that plagiarism, falsification, and fabrication of data are major reasons for retractions in the biomedical field. Elango et al. (Elango, Kozak and Rajendran, 2019) similarly investigated the reasons for retractions among papers affiliated with India, identifying plagiarism as the most common cause. Furthermore, Elango analyzed biomedical literature from Indian authors, revealing that a significant portion of retracted articles were published after 2010, with plagiarism and fake data being common reasons for retractions (Elango, 2021). Additionally, more than half of the retracted articles were co-authored within the same institutions, with no repeat offenders identified. The majority of funded research retractions were attributed to fake data, whereas plagiarism was prevalent among non-funded research. Furthermore, Dhingra et al. (Dhingra and Mishra, 2014) surveyed the extent of misconduct in publication among biomedical researchers. Mukhopadhyay et al. analyzed 1376 retracted papers to identify repeated offenders, author distribution, gender distribution, and institute distribution, noting that elite educational and research institutes bear equal responsibility for retractions (Mukhopadhyay, Mukhopadhyay and Ahmed, 2023). Sabir et al. discussed misconduct in healthcare, highlighting reasons for misconduct and proposing steps to reduce it (Sabir, Kumbhare, Parate, Kumar and Das, 2015). Khurana et al. have

also shown comparative patterns of retractions among USA, China, and India on covid-19 research (Khurana, Sharma and Uddin, 2024).

The current literature on the analysis of retractions in India primarily focuses on the biomedical field. For instance, a study conducted by Elango et al. (Elango et al., 2019) examined 239 retractions from Scopus within the Indian research landscape spanning from 2005 to 2018. Mukhopadhyay et al. also analyzed 1376 retracted papers listed in Retraction Watch (Mukhopadhyay et al., 2023). In contrast, this study offers a more comprehensive analysis covering more than two decades and a larger dataset of retractions. While previous studies mainly aimed to understand the reasons behind retractions, our study delves into the reasons for retractions about journal quartiles, institutional collaboration, and publishing houses. Additionally, our study provides a detailed examination of retractions in terms of inter-country and intra-country collaborations, in correspondence with retraction reasons and journal quartiles, a perspective not previously explored in either the Indian or global literature on retractions.

1.1. Research Objectives

The study aims to analyze Indian retractions with objectives:

- Examine the year-wise trend of the number of retractions and the duration of retractions (measured in months).
- Analyze the distribution of retractions based on journal quartile, the reason for retraction, institution category, and publishing houses.
- Determine the type of institutions involved in retractions (public, private, or medical), along with the cause of retraction and its impact.
- Analyze patterns of retractions stemming from intra-country and inter-country institute collaborations, focusing on the journals involved, reasons for retraction, and journal impact.

The structure of this study is outlined as follows. Section 2 details the process of data collection, filtration, and categorization of institutions. Section 3 delves into the findings and discussion, highlighting key aspects of the retraction trajectory, journal quartile, and reasons for retraction. Section 4 offers insights into institution coverage and the distribution of author collaborations and institutional partnerships. Additionally, Section 5 presents a comparison of inter-country and intra-country institutional collaborations. Section ?? showcases the collaboration network among countries. Finally, the study concludes with corresponding considerations, limitations, and policy recommendations.

2. Methodology

2.1. Data Collection and Filtration

The primary source of data is the Retraction Watch database obtained via Crossref (Hendricks, Lammey, Ofiesh, Bilder and Pentz, 2023) in February 2024. It comprises 51,275 publications, encompassing various details such as title, author names, country affiliations, journal and publisher names, publication subjects, types, retraction specifics including reasons, publication and retraction dates, paper DOIs, and document type.

To identify publications associated with Indian affiliation, papers were filtered based on the presence of "India" within at least one affiliation, resulting in 3,415 records. To refine the dataset to unique entries, further scrutiny was conducted regarding title duplication, original paper DOI duplication, and retraction types. The dataset included four retraction types: correction, expression of concern, reinstatement, and retraction. The ultimate selection for filtering was "*Retraction*". Consequently, after these refinement steps, the dataset was narrowed down to 3,244 unique retractions.

2.2. Institute Categorization

From the initial pool of 3,244 papers, we extracted the department and institution names affiliated with India, culminating in the identification of 5,541 institutions across the 3,244 publications. These institutions underwent further classification according to their nature, including private universities/institutes, private colleges, public universities, institutes of national importance (such as IITs, IIITs, IISc, NITs, etc.) with national research centers (such as CSIR, national labs, etc.), and medical institutes (encompassing hospitals and clinics), adhering to the categorization standards outlined by the University Grants Commission, India (UGC) (<https://www.ugc.gov.in/universitydetails/university?type=ddmCMsxJZgXH2S/m0uM0KQ==>). During the data refinement stage, 102 institutes remained unidentifiable due to insufficient affiliation information or abbreviation usage, resulting in a final tally of 5,439 institutes.

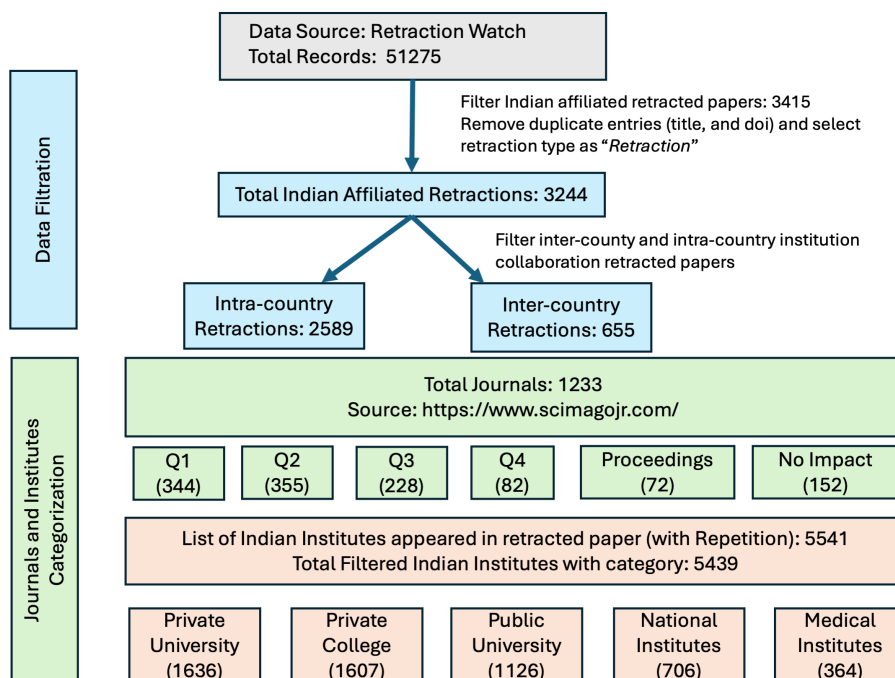


Figure 1: Data collection and filtration process of India affiliated retracted publications.

3. Results and Discussions

3.1. Retraction Trend- Yearly and Monthly

Figure 2 (a) illustrates the trajectory of retractions (red dashed line) alongside their corresponding publication dates (black solid line). There was a decline in the number of retractions in 2004, followed by a subsequent increase since 2005. The trend displays a continuous upward pattern, peaking in 2022 before declining. Among papers published during 1990-1999, there were 13 instances, with only 5 retracted before 2000, while the remaining 9 were retracted after 2007. Notably, papers published between 1990 and 1999 tend to be retracted much later, as highlighted by Elango et al. (2019). On average, retractions of papers published during 1990-1999 took approximately 15 years to occur.

In Figure 2 (b), a notable proportion of papers are retracted early, indicated by a decreasing trend. The timeline for retractions ranges from 0 to 288 months. Approximately 4.5% of papers are retracted within a month, 19.4% in six months, while 40.8% are retracted within 12 months.

3.2. Journal Quartile

Alongside the Impact Factor or Impact Index, journals are categorized into quartiles within each subject category by both Journal Citation Reports (JCR)-Clarivate (<https://clarivate.com/>) and Scimago Journal & Country Rank (SJR) (<https://www.scimagojr.com/>). A quartile represents the ranking of a journal determined by various factors such as impact factor, citation count, and indexing within a specific database (Ahmet and Aslan, 2020). JCR, a paid service, is managed by Web of Science, while SJR, an open-source platform, is maintained by Scopus. These quartiles organize journals from highest to lowest based on their impact index, dividing them into four categories: Q1, Q2, Q3, and Q4. The most esteemed journals within a subject area are typically found in the first quartile, Q1.

- *Quartile-1 (Q1)* encompasses the top 25% of journals.
- *Quartile-2 (Q2)* represents the middle-high position, ranging between 25% to 50% of journals.
- *Quartile-3 (Q3)* signifies the middle-low position, covering journals ranked between 50% to 75%.
- *Quartile-4 (Q4)* indicates the lowest position, encompassing journals ranked between 75% to 100%.

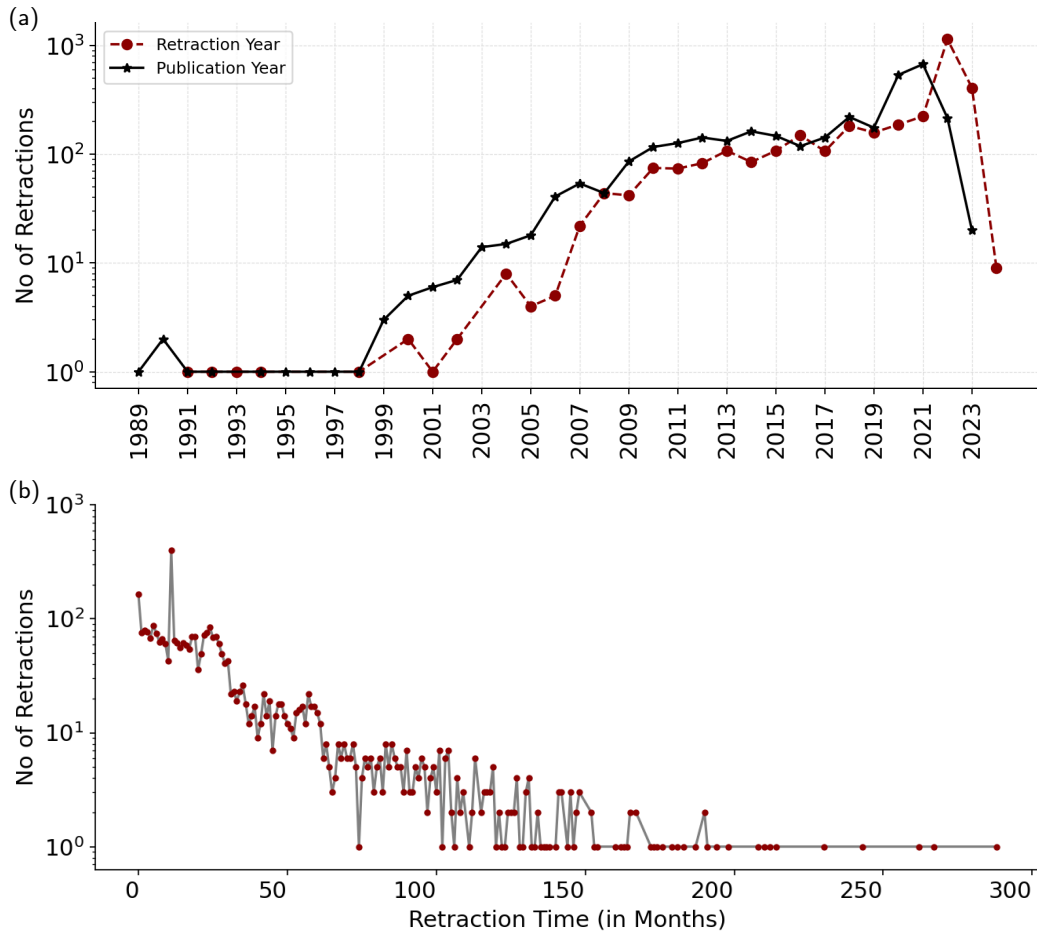


Figure 2: (a) Publication timeline of retracted papers (solid black) and retraction timeline of published papers (dotted red). (b) Retraction time in months.

Additionally, *Proceedings* refers to papers published in conference proceedings following peer review. *No Impact* denotes papers published in journals not indexed by Scopus or stored in scholarly platforms like arXiv.

Figure 3 illustrates the distribution of retractions based on journal quartiles. The highest proportion of retractions originates from Q1 journals (34%), followed by Q2 journals (23.3%). Retractions associated with Q4 journal (14.9%) are higher than Q3 (12.2%).

3.3. Retractions as per Top 10 Publishers

As illustrated in Table 1, a notable portion of retractions (28.5%) originates from Springer Publishing House, with 55.1% retracted from Q1 journals. IOP Publishing ranks second, contributing to 14.03% of retractions, with 59.8% originating from Q4 journals and 37.5% from conference proceedings. The third highest retraction rate is from Elsevier (12.7%), predominantly from Q1 journals (56.9%). Hindawi (43.1%), Taylor and Francis (36.7%), and Wiley (46.6%) predominantly experience retractions in Q2 journals. Wolters Kluwer (37.2%) primarily experiences retractions from Q3 journals, while SAGE Publications (49.5%) does so from Q4 journals. All retractions from the PLoS publishing house are associated with Q1 impact. Additionally, Hindawi (10.3%) and Wolters Kluwer (14.7%) experience significant instances of retractions with no impact. Retractions from IEEE predominantly involve proceedings (55.5%) (Vuong et al., 2020), followed by IOP Publishing (37.5%). This indicates that retractions from Q1 journals primarily originated from Springer and Elsevier, while those from Q2 journals came mainly from Hindawi, Taylor and Francis, and Wiley. Retractions from Q3 journals were associated with Wolters Kluwer, and those from Q4 journals were predominantly from SAGE Publications and IOP Publishing.

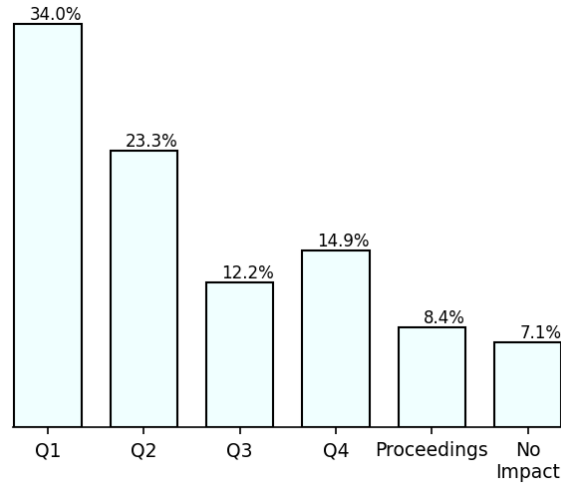


Figure 3: Number of retractions as per journal quartile.

Table 1

Number of retractions as per journal quartile for top 10 publishers.

Publisher	No of Retractions			Journal Quartile (in%)				
	Count	In%	Q1	Q2	Q3	Q4	Proceedings	No Impact
Springer	927	28.58	55.12	26.43	8.63	2.59	2.59	4.64
IOP Publishing	455	14.03	-	2.64	-	59.78	37.58	-
Elsevier	413	12.73	56.9	24.94	10.65	1.94	0.97	4.6
Hindawi	204	6.29	23.04	43.14	20.1	3.43	-	10.29
Wolters Kluwer	156	4.81	7.05	28.85	37.18	12.18	-	14.74
Taylor and Francis	136	4.19	26.47	36.76	22.06	12.5	-	2.21
IEEE	126	3.88	11.9	10.32	15.87	6.35	55.56	-
Wiley	120	3.7	27.5	46.67	19.17	-	-	6.67
SAGE Publications	99	3.05	7.07	33.33	8.08	49.49	1.01	1.01
PLoS	82	2.53	100	-	-	-	-	-

3.4. Retraction Reasons

Studies can face retraction for a variety of reasons, broadly categorized into honest error, intentional or unintentional misconduct, and other factors (Xu and Hu, 2022; Sharma et al., 2023). Honest errors typically involve mistakes in sample or data collection, irreproducibility of results, disputes over authorship attribution, redundant publication, and similar issues. On the other hand, misconduct encompasses actions such as plagiarism, undisclosed conflicts of interest, data fabrication or manipulation, ethical concerns, duplicate submissions, and related misconduct. The third category, labeled as “others”, encompasses retractions where the rationale is either not explicitly stated in the retraction notice or remains unclear to the journal editors. For example, authors may retract a paper that has been simultaneously submitted to and accepted by another journal, which constitutes misconduct for failing to inform the journal about the duplicate submission. However, if authors do not specify the reason for retraction, it would be categorized under “other”. Based on the aforementioned categories, the provided data is labeled for the following retraction reasons.

- *Fake Peer Review*: This involves recommending peer reviewers to journals upon paper submission, providing falsified contact information, and subsequently submitting fraudulent, excessively favorable peer review reports, typically resulting in fast publication (Ferguson, Marcus and Oransky, 2014; Watch).
- *Data Integrity*: This encompasses errors in data, fabricated results, data manipulation, and content generated randomly, among other issues (Chambers, Michener and Falcone, 2019).

Table 2

Number of retraction as per journal quartile and their corresponding reason of retraction.

Retraction Reasons	No of Retractions			Journal Quartile (in %)				
	Count	In %	Q1	Q2	Q3	Q4	Proceedings	No Impact
Fake Peer Review	1071	33.01	41.94	22.06	10.63	58.8	41.97	-
Multiple	589	18.16	18.84	21.93	18.48	9.94	19.34	17.75
Data Integrity	559	17.23	19.66	20.48	20.76	12.22	6.2	12.55
Plagiarism	481	14.83	10.69	16.91	21.77	6.63	15.69	32.03
Unknown	188	5.8	2.26	4.36	7.34	5.8	7.66	22.51
Duplication of Article	180	5.55	3.08	7.66	12.91	3.31	1.09	7.79
Authors Related	86	2.65	1.99	3.3	3.8	1.66	1.46	5.19
Error by Journal	49	1.51	1.09	1.98	2.53	1.45	0.73	1.3
Ethical Concerns	41	1.26	0.45	1.32	1.77	0.21	5.84	0.87
Grand Total	3244	100	1104	757	395	483	274	231

- *Plagiarism*: Plagiarism occurs when someone uses another person's intellectual work, such as texts, ideas, or results, without proper citation, implying it to be their own (Helgesson and Eriksson, 2015).
- *Duplication of Article*: This refers to the publication of the same paper in multiple journals (Alfonso, Bermejo and Segovia, 2005).
- *Authors Related*: This arises when individuals are listed as authors on a published manuscript without their participation or consent (Hosseini and Gordijn, 2020).
- *Ethical Concerns*: Ethical concerns typically arise from a failure to obtain ethical approval before conducting a study involving animal and/or human subjects (Elia, Wager and Tramèr, 2014).
- *Error by Journal*: This occurs when a journal mistakenly publishes the same paper again.
- *Multiple*: This category encompasses various reasons for retraction, such as authorship disputes, conflicts of interest, data manipulation, and copyright issues, among others.
- *Unknown*: This category includes cases where the specific reason for retraction is not known.

In Table 2, it becomes apparent that the primary reason for the majority of retractions is fake peer review, constituting 33% of the total retractions. Papers retracted from the Q1 category are mainly due to fake peer reviews (41.9%). Retractions from Q2 journals are nearly equally distributed among fake peer review (22%), data integrity (20.5%), and multiple reasons (21.9%). For Q3 retractions, the majority are due to plagiarism (21.7%) and data integrity issues (20.7%). In Q4, the majority are due to fake peer reviews (58.8%). Additionally, conference proceedings are primarily retracted due to fake peer reviews (41.9%). The second-highest cause of retraction in Q1 journals is attributed to data integrity issues (17.2%), with 19.6% of retractions originating from Q1 journals and 20.4% from Q2 journals. Plagiarism accounts for 14.8% of total retractions. Notably, 32% retraction is due to plagiarism and is from journals not listed in the Scopus, thus having no impact. Furthermore, 18.2% of retractions stem from multiple reasons, with nearly equal proportions from Q1, Q2, Q3 journals, and conference proceedings. Additionally, 12.9% of retractions from Q3 journals are attributed to article duplication. Combined, fake peer review, data integrity, and plagiarism account for approximately 65% of retractions.

3.5. Average Retraction Time

As indicated in Table 3, the shortest average retraction period is observed for papers published in conference proceedings (12.17 months), while retractions from Q2 publications took the longest time (32.8 months), followed by Q1 (30.7 months). Fake peer review, being the highest mode of retraction, took an average of 21.5 months, while data integrity ranks as the second most common reason for retraction, with the highest average retraction period of 38.8 months. Similarly, duplication of articles, despite being less frequent in retractions, also took a long time at 33.4 months. Moreover, most retractions are of the article type, with an average retraction time of 31.6 months. Conversely, conference proceedings have the shortest retraction time at 10.9 months. Reviews, articles, and clinical studies took nearly equal durations for retraction.

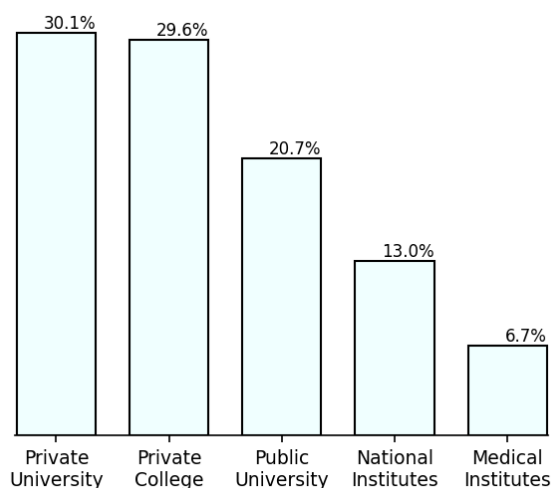
Table 3

Average time of retraction by journal quartile, retraction reasons, and document type..

Journal Quartile	Avg Retraction Time (in months)	Retraction Reasons	Avg Retraction Time (in months)	Document Type	Avg Retraction Time (in months)
Q1	30.71	Data Integrity	38.87	Review	34.14
Q2	32.85	Duplication of Article	33.47	Article	31.68
Q3	25.47	Multiple	32.64	Clinical Study	30.54
Q4	23.99	Plagiarism	27.93	Case Report	23.66
Proceedings	12.17	Error by Journal	26.67	Book Chapter	22.83
No Impact	24.55	Fake Peer Review	21.52	Others	16.93
-	-	Authors Related	21.14	Commentary	13.33
-	-	Ethical Concerns	14.65	Conference	10.99
-	-	Unknown	11.85	-	-

4. Institution Categorization and Collaboration

Figure 5 shows that approximately 60% of retractions stem from private colleges, universities, and institutes, with 33.7% originating from public universities, institutes, and research centers. A smaller proportion, about 6.7%, comes from medical institutes, hospitals, and clinics. When categorizing institutes, “National Institutes” includes all establishments of national significance such as IIT, IIIT, NIT, ISI, ISC, etc., along with research centers like CSIR and national laboratories. “Medical institutes” encompass all nationally important medical establishments such as All India Institute of Medical Sciences (AIIMS), Post Graduate Institute of Medical Education & Research (PGIMER), etc., as well as hospitals like Apollo Hospitals and Sir Ganga Ram Hospital.

**Figure 4:** Number of retractions as per institute category.

The majority of retractions occur within smaller team sizes, with only 5.6% of retractions attributed to single authors. The highest proportion of retractions is observed in teams consisting of two authors (25.1%), followed by three authors (19.9%) and four authors (17.1%). Moreover, there are individual papers with author sizes of 15, 17, 20, 21, 23, and 32. The trend depicted in Figure 5 (a) suggests that as the team size increases, the number of retractions decreases Sharma (2021). Approximately 50.6% of retractions occur when up to three authors collaborate, and this increases to 87.9% when up to six authors collaborate.

Likewise, Figure 5 (b) depicts the pattern of institute collaboration. Fewer retracted papers are affiliated with multiple institutions. The highest number of retractions (38.7%) occurs when authors are from the same institute, followed by collaborations between two institutes (, and three institutes (14.4%). Retractions involving collaborations

Table 4

Retraction reasons corresponding to affiliation category (in %). The number of institutes in each category are with repetition as same institute may appear in multiple publications.

Retraction Reasons	Private University	Private College	Public University	National Institutes	Medical Institutes
Fake Peer Review	43.28	52.77	19.8	7.93	2.75
Multiple	22.31	15.37	25.49	20.68	15.66
Data Integrity	14.24	6.97	24.25	35.41	28.85
Plagiarism	11.74	12.69	15.45	13.17	19.78
Duplication of Article	3.48	5.1	5.15	8.22	12.09
Unknown	1.71	3.98	4.62	7.37	8.79
Authors Related	1.65	1.56	2.58	3.97	6.04
Error by Journal	0.67	0.93	1.69	1.98	3.3
Ethical Concerns	0.92	0.62	0.98	1.27	2.75
Total Institutes	1636	1607	1126	706	364

among multiple institutes are minimal. Approximately 67.4% of retractions occur when up to two institutes collaborate, increasing to 92.8% when up to five institutes collaborate. In single institute retraction, majority of the retraction is from private colleges (24.5%) followed by private universities (22%). The same pattern is observe in 2 institute collaboration and so on.

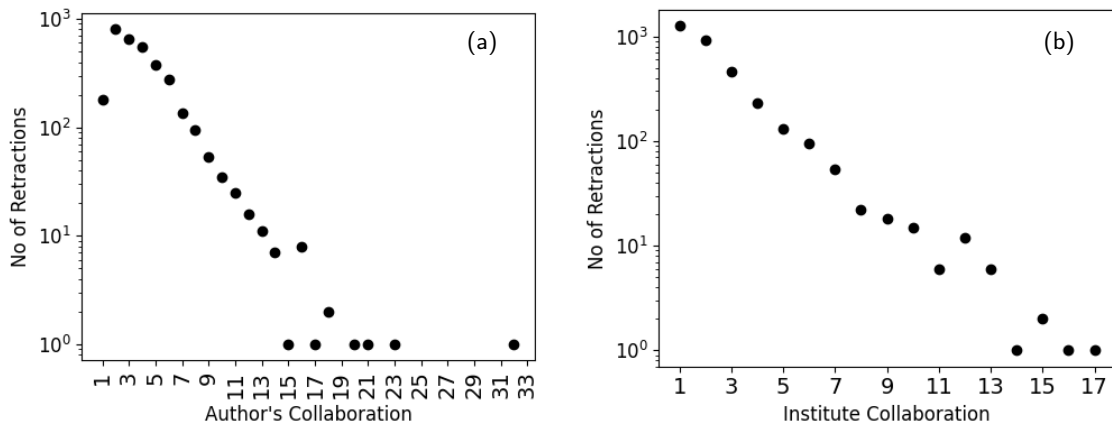


Figure 5: (a) Author's collaboration count. (b) Institute collaboration count.

4.1. Retraction Reasons Corresponding to Institution Category

As indicated in Table 4, among 1636 instances of appearances in private institutes and universities, 43.2% of retractions are attributed to fake peer review, with 22.3% attributed to multiple reasons. Similarly, out of 1607 instances of appearances in private colleges in all retracted publications, 52.7% of retractions are due to fake peer reviews, with 15.3% due to multiple reasons. Among public universities, totaling 1126 appearances, the primary cause of retractions is data integrity issues (24.2%), followed by multiple reasons at 25.4%. Similarly, in national institutes and research centers comprising 706 appearances, 35.4% of retractions are linked to data integrity, followed by 20.6% due to multiple reasons. Medical institutes experience 28.8% retractions due to data integrity and 19.7% due to plagiarism. This suggests that fake reviews are prevalent in private institutions, while data integrity issues are common in public institutions.

Table 5

Intra-country Vs. inter-country retractions as per journal quartile.

Journal Quartile	Inter-country Institute Retractions		Intra-country Institute Retractions	
	Count	In %	Count	In %
Q1	316	48.24	788	30.44
Q2	190	29.01	567	21.9
Q3	34	5.19	318	12.28
Q4	77	11.76	458	17.69
Proceedings	25	3.82	261	10.08
No Impact	13	1.98	197	7.61

5. Intra-country Vs. Inter-country Retractions

Intra-country collaborations encompass retractions wherein all authors are affiliated with Indian institutes or universities. Approximately 80% of the retracted papers exclusively feature Indian affiliations and collaborations. As depicted in Figure 6, retractions within the country display an upward trajectory from 2001 to 2004, followed by a decline in 2005. Subsequently, there is an upward trend post-2005, indicating an increase in retractions involving authors collaborating within India. Within the last five years (2019-2023), 62.6% of retractions occurred.

Conversely, in inter-country collaborations, there is an ascending trend from 2017 to 2023, with fluctuations observed before 2017. Several retractions corresponding to 2024 involve international collaboration. Notably, no retractions were reported from 2000 to 2004 and from 2006 to 2007. Within the last five years (2019-2023), 77% of retractions occurred.

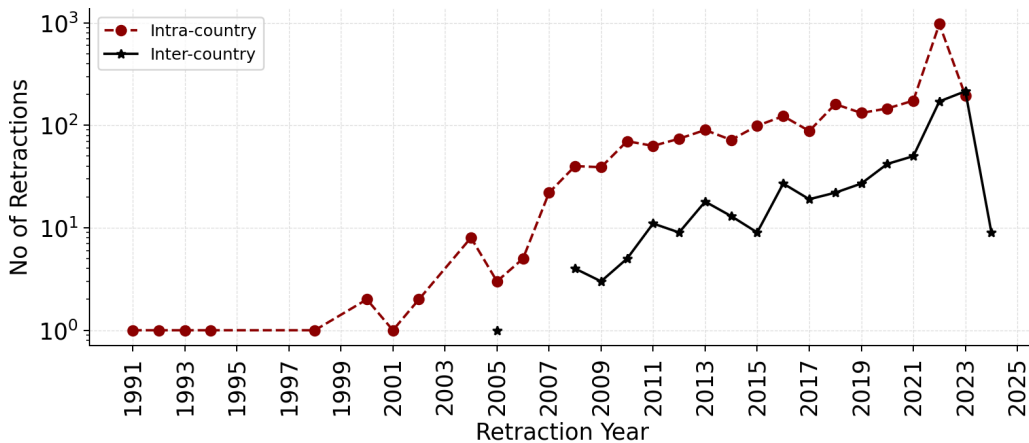


Figure 6: Retraction trend of papers with intra-country institutions collaboration (red dotted line) and inter-country institutions collaboration (black solid line).

5.1. Variation in Journal Quartile

As depicted in Table 5, retractions classified under Q1 are more prevalent in inter-country collaborations (48.2%) compared to intra-country collaborations (30.4%). Similarly, Q2 retractions are higher in inter-country collaborations (29%) than in intra-country collaborations (21.9%). Conversely, Q3 and Q4 retractions show higher frequencies in intra-country collaborations (12.3%, 17.6%) compared to inter-country collaborations (5.2%, 11.7%). In contrast to inter-country collaborations, intra-country collaborations exhibit a higher proportion of retractions in proceedings (10% vs. 3.8%). Overall, intra-country collaborations account for 52.3%—approximately half of the total retractions—from the top two journal quartiles Q1 and Q2, while in inter-country collaborations, 77.2%—roughly one-third of retractions—originate from the top two quartiles.

Table 6

Intra-country Vs. inter-country retractions as per retraction reasons.

Retraction Reasons	Inter-country Institute Retractions		Intra-country Institute Retractions	
	Count	In %	Count	In %
Fake Peer Review	156	23.82	915	35.34
Plagiarism	62	9.47	419	16.18
Data Integrity	141	21.53	418	16.15
Multiple	211	32.21	378	14.6
Unknown	20	3.05	168	6.49
Duplication of Article	24	3.66	156	6.03
Authors Related	27	4.12	59	2.28
Error by Journal	6	0.92	43	1.66
Ethical Concerns	8	1.22	33	1.27

Table 7

Intra-country: Number of retractions with respect to document type and journal quartile.

Document Type	Count	In %	Q1	Q2	Q3	Q4	Proceedings	No Impact
Article	1691	65.31	42.81	27.91	12.48	8.4	0.89	7.51
Conference	543	20.97	2.58	4.05	4.05	51.57	36.46	1.29
Review	114	4.4	18.42	25.44	24.56	8.77	1.75	21.05
Clinical Study	93	3.59	18.28	25.81	31.18	12.9	-	11.83
Case Report	65	2.51	7.69	24.62	35.38	18.46	-	13.85
Commentary	53	2.05	3.77	3.77	5.66	1.89	83.02	1.89
Book Chapter	19	0.73	-	-	-	-	10.53	89.47
Others	11	0.42	45.45	18.18	18.18	9.09	-	9.09
Total Retractions	2589	100	788	567	318	458	261	197

5.2. Variation in Retraction Reasons

Table 6 reveals that the primary reason for retraction in intra-country scenarios is fake peer review (35.3%), whereas, in inter-country cases, retractions occur due to a multitude of reasons (32.2%). plagiarism (16.2%) and data integrity issues (16.2%) constitute the secondary cause of retractions in intra-country situations, while in inter-country collaborations, fake peer review (23.8%) emerges as the secondary cause, followed by data integrity concerns (21.5%) as the third most common reason. Issues about authors are more prevalent in inter-country collaborations compared to intra-country ones, whereas article duplication is more prevalent in intra-country collaborations than in inter-country ones.

5.3. Variation in Document Type

In Table 7, within intra-country collaborative retractions, articles account for 65.3% of retractions. The second highest source of retractions is conference proceedings (20.9%), of which 51.5% are from quartile Q4 and 36.4% are listed as proceedings with no impact. Reviews contribute 4.4% overall, predominantly published in Q2 and Q3. Clinical studies and case reports fall predominantly into the Q3 and Q4 categories. Conversely, in inter-country collaborative retracted papers (refer to Table 8), articles constitute 87% of retractions, with 52.1% from Q1 journals and 29.8% from Q2. The second highest retraction source is reviews (5.8%), with the majority occurring in Q2 and Q3 journals. The majority of clinical studies are published in Q1 journals.

5.4. Variation in Publishing Journals

Table 9 displays the top 10 journals from which intra-collaborative papers were initially published and later retracted. The majority of the published journals belonged to the Q2 and Q4 categories across various publishing houses. However, the highest number of retractions (12.9%) originated from Q1 journals published by Springer. The second-highest source of retractions was conference proceedings published by IOP Publishing and listed as Q4. As mentioned earlier in Section 5.3, intra-country retractions are predominantly articles and conference proceedings, with a majority originating from *Journal of Ambient Intelligence and Humanized Computing* and *Journal of Physics: Conference*

Table 8

Inter-country: Number of retractions with respect to document type and journal quartile.

Document Type	Count	In %	Q1	Q2	Q3	Q4	Proceedings	No Impact
Article	570	87.02	52.11	29.82	11.05	2.63	-	4.39
Review	38	5.8	15.79	44.74	31.58	5.26	-	2.63
Conference	15	2.29	-	-	-	26.67	66.67	6.67
Clinical Study	13	1.98	46.15	0.53	1.3	12	-	5.88
Case Report	6	0.92	1.27	-	0.32	-	-	0.32
Commentary	5	0.76	20	-	-	20	60	-
Others	4	0.61	50	50	-	-	-	-
Book Chapter	4	0.61	-	-	-	-	-	100
Total Retractions	655	100	316	190	77	25	13	34

Table 9

Intra-country: List of top 10 journals with number of retractions, journal impact and publisher detail.

Journl Name	No of Retractions	In %	Quartile	Publisher
Journal of Ambient Intelligence and Humanized Computing	335	12.94	Q1	Springer
Journal of Physics: Conference Series	268	10.35	Q4	IOP Publishing
IOP Conference Series: Materials Science and Engineering	161	6.22	Proceedings	IOP Publishing
The International Journal of Electrical Engineering & Education	33	1.27	Q4	SAGE Publications
Cluster Computing	31	1.2	Q2	Springer
International Journal of Mechanical and Production Engineering Research and Development	31	1.2	No Impact	Trans Stellar
Wireless Personal Communications	27	1.04	Q2	Springer
PLoS One	25	0.97	Q1	PLoS
RSC Advances	23	0.89	Q2	Royal Society of Chemistry
The Journal of Biological Chemistry	21	0.81	Q4	American Society for Biochemistry and Molecular Biology

Series.

Similarly, Table 10 presents the top 10 journals from which inter-collaborative papers were initially published and subsequently retracted. The majority of the published journals were of the Q1 category across various publishing houses. However, the highest number of retractions occurred in Q1 journals, with the majority of retractions originating from Hindawi publishers as per the top 10 journals. The highest retraction rate was observed in *PLoS One*, accounting for 8.09% of retractions, followed by *Computational Intelligence and Neuroscience* with 4.5% of retractions published by Hindawi.

6. Country Collaboration Network

Indian researchers were involved in 655 retracted papers, collaborating with researchers from 93 countries. Figure 7 illustrates the country collaboration network, with India positioned as the central node, engaging in numerous collaborations, particularly with Saudi Arabia. The second most common collaboration occurred with the USA, followed by China, Ethiopia, and Malaysia, among others. Specifically, India collaborated with Saudi Arabia in 166 retracted papers (14.3%), the USA in 117 (10.1%), and China in 104 retracted papers (8.9%). Moreover, collaborations with Ethiopia (83) and Malaysia (58) jointly contributed to 12% of retractions. Additionally, notable collaborations were observed with Pakistan, South Korea, the United Kingdom, Egypt, Iraq, Australia, and Iran, highlighting India's global engagement in cases of misconduct.

Table 10

Inter-country: List of top 10 journals with number of retractions, journal impact and publisher detail.

Journl Name	No of Retractions	In %	Quartile	Publisher
PLoS One	53	8.09	Q1	PLoS
Computational Intelligence and Neuroscience	30	4.58	Q1	Hindawi
Journal of Ambient Intelligence and Humanized Computing	23	3.51	Q1	Springer
Aggression and Violent Behavior	19	2.9	Q1	Elsevier
Arabian Journal for Science and Engineering	19	2.9	Q1	Springer
Journal of Healthcare Engineering	17	2.6	Q3	Hindawi
Journal of Nanomaterials	13	1.98	Q1	Hindawi
Proceedings of the Institution of Mechanical Engineers, Part E: Journal of Process Mechanical Engineering	11	1.68	Q2	SAGE Publications
Adsorption Science & Technology	11	1.68	Q2	Hindawi
Advances in Materials Science and Engineering	10	1.53	Q2	Hindawi

Considering India's collaboration with 93 countries, a total of 1158 connections were formed. One-third (33.4%) of retractions occurred in collaboration with the top three countries: Saudi Arabia, USA, and China, while 62% of retractions originated from collaborations with the top 10 countries: Saudi Arabia, USA, China, Ethiopia, Malaysia, Pakistan, South Korea, the United Kingdom, Egypt, Iraq, Australia, and Iran.

7. Conclusion

The current investigation into retracted publications affiliated with India sheds light on valuable insights regarding the types of retractions and the institutions involved. This study analyzed 3244 retracted papers with Indian affiliations, among which Indian authors solely authored 2589, and 655 were collaborations with authors from 93 countries. The upward trend in the number of retractions and the decreasing time until retraction (measured in months) indicate that a majority of the retracted Indian articles were published after 2010. The primary reasons for retraction are fake peer review, followed by data integrity and plagiarism.

Upon examining intra- and inter-country collaboration patterns, it was observed that retractions of papers solely affiliated with India are mainly attributed to fake peer review, data integrity issues, and plagiarism. Plagiarism is primarily observed in papers published in conference proceedings or those lacking impact (not indexed in Scopus). Conversely, papers retracted with international collaboration were retracted due to various reasons.

Several private institutes and universities offer financial incentives for each publication, potentially leading to a focus on quantity over quality and encouraging unethical publishing practices. Furthermore, the types of Indian institutions and universities involved in such retractions were analyzed. It was discovered that 60% of retractions originated from private universities, institutes, or colleges, while the remaining retractions were from public universities, national institutes, research centers, or a small fraction from medical colleges or hospitals. The primary cause of retraction from these institutions was examined, revealing that fake peer review is the main cause in private institutes or colleges, whereas data integrity issues are more prominent in public institutes or universities and medical colleges.

Notably, prominent publishers such as Springer, IOP Publishing, and Elsevier account for more than half of the retractions. Overall, intra-country collaborations account for approximately half of the total retractions from the top two journal quartiles (Q1 and Q2), while in inter-country collaborations, roughly one-third of retractions originate from the top two quartiles. Additionally, 45.6% of retractions originated from the top five collaborators: Saudi Arabia, USA, China, Ethiopia, and Malaysia.

In essence, collaboration alone cannot guarantee the quality of research outcomes. It is crucial to address related actions about the researcher's adherence to scientific norms within the scientific community. Scientific misconduct closely correlates with individual awareness of scientific integrity, potentially undermining both collaboration networks and scientific integrity as highlighted by Sharma et al. (2023).

For institutions with higher retraction rates, there is a pressing need to prioritize the production of high-quality research and enhance scientific ethics through integrity training. To mitigate the adverse impact of recurrent misconduct among collaborators, it is advisable to establish a comprehensive and legally protected system for reporting scientific

Two Decades of Scientific Misconduct in India

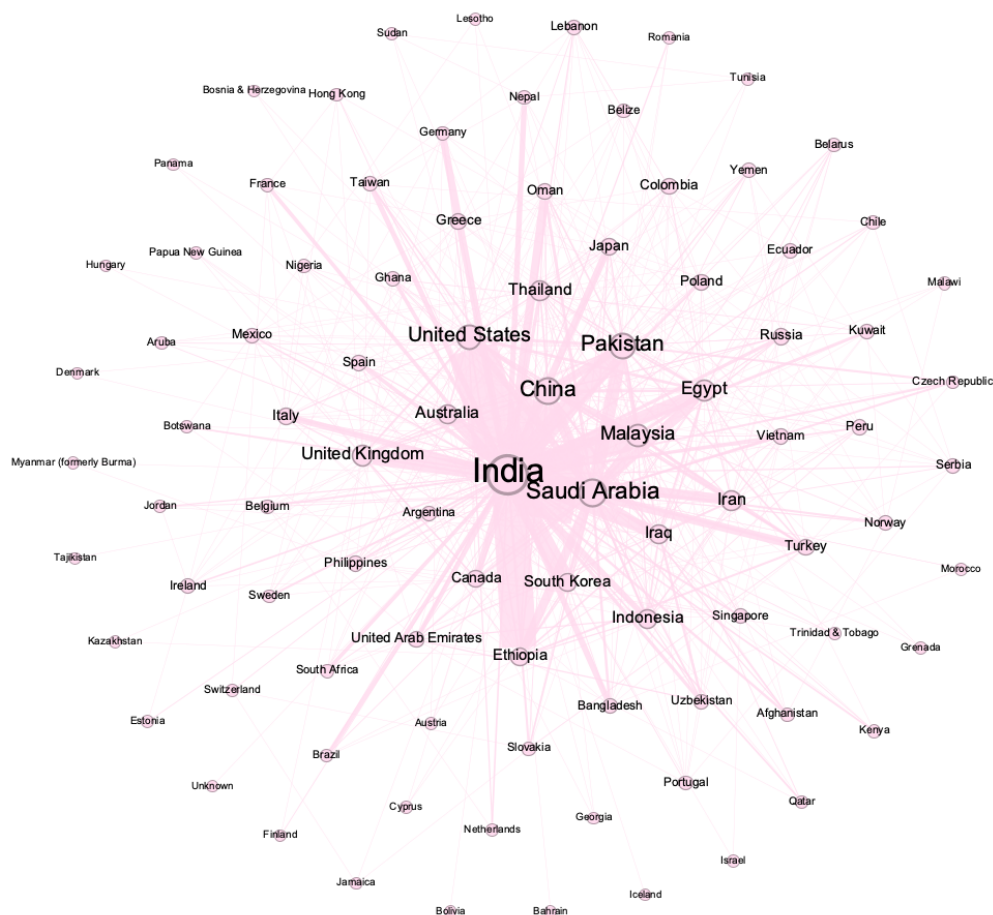


Figure 7: Country collaboration network with 94 nodes and 561 unique edges.

misconduct, along with mentorship programs for institutions or research groups (Lu, Jin, Uzzi and Jones, 2013; Davis, Riske-Morris and Diaz, 2007; Kornfeld, 2012).

Moreover, researchers themselves need to recognize the repercussions of scientific misconduct on their careers and the integrity of national scientific endeavors (Shahre and Roberts, 2020). Private institutes ought to alleviate the pressure of "publish or perish" linked to annual performance evaluations, thereby reducing the temptation to engage in fraudulent practices such as fake peer review for expedited publications or the fabrication of data to produce extensive papers. Researchers must receive training to recognize the factors contributing to plagiarism and understand its consequences. Additionally, it is vital to establish appropriate policies for governing national science, particularly those addressing issues such as fake peer review, plagiarism, and the clarification of authorship contributions. These policies facilitate transparency regarding individual research contributions, as advocated by COPE (2020).

The study has several limitations regarding data availability for addressing certain questions. Firstly, the information regarding corresponding authors is not provided, and it is unclear which author is affiliated with which institute. Consequently, determining who is responsible for the false research and who the primary contributor (first author) is becomes challenging. Additionally, the author names are abbreviated, making it difficult to discern the authors' genders. Thus, we cannot ascertain the gender distribution in the retractions or their roles in the retracted work. Similarly, information about the types of institutions the authors belong to is also lacking.

Data availability statement

The data is freely available on Crossref (<https://www.crossref.org/>).

Conflict of Interest

The author declares no conflict of interest.

References

- Ahmet, A., Aslan, A., 2020. Quartile scores of scientific journals: Meaning, importance and usage. *Acta Medica Alanya* 4, 102–108.
- Alfonso, F., Bermejo, J., Segovia, J., 2005. Duplicate or redundant publication: can we afford it? *Revista Española de Cardiología (English Edition)* 58, 601–604.
- Bar-Ilan, J., Halevi, G., 2018. Temporal characteristics of retracted articles. *Scientometrics* 116, 1771–1783.
- Bornmann, L., Mungra, P., 2011. Improving peer review in scholarly journals. *European Science Editing* 37, 41–43.
- Campos-Varela, I., Ruano-Raviña, A., 2019. Misconduct as the main cause for retraction. a descriptive study of retracted publications and their authors. *Gaceta sanitaria* 33, 356–360.
- Chaddah, P., Lakhotia, S.C., 2018. A policy statement on “dissemination and evaluation of research output in india” by the indian national science academy (new delhi) .
- Chambers, L.M., Michener, C.M., Falcone, T., 2019. Plagiarism and data falsification are the most common reasons for retracted publications in obstetrics and gynaecology. *BJOG: An International Journal of Obstetrics & Gynaecology* 126, 1134–1140.
- Davis, M.S., Riske-Morris, M., Diaz, S.R., 2007. Causal factors implicated in research misconduct: Evidence from ori case files. *Science and engineering ethics* 13, 395–414.
- Dhingra, D., Mishra, D., 2014. Publication misconduct among medical professionals in india. *Indian J Med Ethics* 11, 104–107.
- Elango, B., 2021. Retracted articles in the biomedical literature from indian authors. *Scientometrics* 126, 3965–3981.
- Elango, B., Kozak, M., Rajendran, P., 2019. Analysis of retractions in indian science. *Scientometrics* 119, 1081–1094.
- Elia, N., Wager, E., Tramèr, M.R., 2014. Fate of articles that warranted retraction due to ethical concerns: a descriptive cross-sectional study. *PLoS one* 9, e85846.
- Fanelli, D., 2013. Why growing retractions are (mostly) a good sign. *PLoS medicine* 10, e1001563.
- Fang, F.C., Steen, R.G., Casadevall, A., 2012. Misconduct accounts for the majority of retracted scientific publications. *Proceedings of the National Academy of Sciences* 109, 17028–17033.
- Ferguson, C., Marcus, A., Oransky, I., 2014. The peer-review scam. *Nature* 515, 480.
- Helgesson, G., Eriksson, S., 2015. Plagiarism in research. *Medicine, Health Care and Philosophy* 18, 91–101.
- Hendricks, G., Lammey, R., Ofiesh, L., Bilder, G., Pentz, E., 2023. Crossref acquires retraction watch data and opens it for the scientific community. URL: <https://www.crossref.org/blog/news-crossref-and-retraction-watch/>.
- Hosseini, M., Gordijn, B., 2020. A review of the literature on ethical issues related to scientific authorship. *Accountability in Research* 27, 284–324.
- Khurana, P., Sharma, K., Uddin, Z., 2024. Unraveling retraction dynamics in covid-19 research: Patterns, reasons, and implications. *arXiv/5497638*.
- Kleinert, S., 2009. Cope's retraction guidelines. *The Lancet* 374, 1876–1877.
- Kornfeld, D.S., 2012. Perspective: Research misconduct: The search for a remedy. *Academic Medicine* 87, 877–882.
- Lee, I., 2014. Publish or perish: The myth and reality of academic publishing. *Language teaching* 47, 250–261.
- Lu, S.F., Jin, G.Z., Uzzi, B., Jones, B., 2013. The retraction penalty: Evidence from the web of science. *Scientific reports* 3, 3146.
- Mongeon, P., Larivière, V., 2016. Costly collaborations: The impact of scientific fraud on co-authors' careers. *Journal of the Association for Information Science and Technology* 67, 535–542.
- Mukhopadhyay, P., Mukhopadhyay, M., Ahmed, M., 2023. Retractions in india since independence: a multifaceted analysis for 75 years through data carpentry .
- Racimo, F., Galtier, N., De Herde, V., Bonn, N., Phillips, B., Guillemaud, T., Bourguet, D., 2022. Ethical publishing: how do we get there? *Philosophy, Theory, and Practice in Biology* 14, 15.
- Rawat, S., Meena, S., 2014. Publish or perish: Where are we heading? *Journal of research in medical sciences: the official journal of Isfahan University of Medical Sciences* 19, 87.
- Receveur, A., Bonfanti, J., d'Agata, S., Helmstetter, A.J., Moore, N.A., Oliveira, B.F., Petit-Cailleux, C., Rievers Borges, E., Schultz, M., Sexton, A.N., et al., 2024. David versus goliath: Early career researchers in an unethical publishing system. *Ecology Letters* 27, e14395.
- Sabir, H., Kumbhare, S., Parate, A., Kumar, R., Das, S., 2015. Scientific misconduct: A perspective from india. *Medicine, Health Care and Philosophy* 18, 177–184.
- Shahare, M., Roberts, L.L., 2020. Historicizing the crisis of scientific misconduct in indian science. *History of Science* 58, 485–506.
- Sharma, K., 2021. Team size and retracted citations reveal the patterns of retractions from 1981 to 2020. *Scientometrics* 126, 8363–8374.
- Sharma, K., Mukherjee, S., 2023. The ripple effect of retraction on an author's collaboration network. *arXiv preprint arXiv:2310.05941* .
- Sharma, P., Sharma, B., Reza, A., Inampudi, K.K., Dhamija, R.K., 2023. A systematic review of retractions in biomedical research publications: reasons for retractions and their citations in indian affiliations. *Humanities and Social Sciences Communications* 10, 1–12.
- Steen, R.G., 2011. Retractions in the scientific literature: is the incidence of research fraud increasing? *Journal of medical ethics* 37, 249–253.
- Tang, L., Hu, G., Sui, Y., Yang, Y., Cao, C., 2020. Retraction: the “other face” of research collaboration? *Science and Engineering Ethics* 26, 1681–1708.

Two Decades of Scientific Misconduct in India

- Vuong, Q.H., La, V.P., Hõ, M.T., Vuong, T.T., Ho, M.T., 2020. Characteristics of retracted articles based on retraction data from online sources through february 2019. *Science Editing* 7, 34–44.
- Wager, E., Barbour, V., Yentis, S., Kleinert, S., et al., 2009. Retractions: guidance from the committee on publication ethics (cope). *Maturitas* 64, 201–203.
- Watch, R., . Springer-bmc retracting nearly 60 papers for fake reviews and other issues.
- Xu, S., Hu, G., 2022. A cross-disciplinary and severity-based study of author-related reasons for retraction. *Accountability in Research* 29, 512–536.