

An Exploratory Analysis of 4844 Withdrawn Articles and Their Retraction Notes

Catalin Toma¹, Liliana Padureanu²

¹Romedchim International S.R.L., Iasi, Romania

²Emergency Clinical Hospital "St. Spiridon", Iasi, Romania

Email: catalin.toma@gmail.com

How to cite this paper: Toma, C., & Padureanu, L. (2021). An Exploratory Analysis of 4844 Withdrawn Articles and Their Retraction Notes. *Open Journal of Social Sciences*, 9, 415-447.

<https://doi.org/10.4236/jss.2021.911028>

Received: October 18, 2021

Accepted: November 27, 2021

Published: November 30, 2021

Copyright © 2021 by author(s) and Scientific Research Publishing Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

Abstract

The objective of our study was to obtain an updated image of the dynamic of retractions and retraction notes, retraction reasons for questionable research (QRP) and publication practices (QPP), countries producing retracted articles, and the scientific impact of retractions, by studying 4844 PubMed indexed retracted articles published between 2009 and 2020 and their retraction notes. **Results:** Mistakes/inconsistent data account for 32% of total retractions, followed by images (22.5%), plagiarism (13.7%) and overlap (11.5%). There were 163 cases of duplicate submission (3.36%), 180 cases of lack of reproducibility (3.72%), 181 cases of editorial errors (3.73%), 229 cases of fabricated data (4.73%) and 350 cases of fraudulent peer review (7.22%). Journals failed to properly report the retraction in 247 cases (5.1%). Thirty countries account for 94.79% of 4844 retractions. Top five are: China (32.78%), United States (18.95%), India (7.24%), Japan (4.37%) and Italy (3.75%). The total citations number for all articles is 140,810 (Google Scholar), 96,000 (Dimensions). Average exposure time (ET) is 28.89 months. The largest ET is for image retractions (49.3 months), the lowest ET is for editorial errors (11.2 months). The impact of retracted research is higher for Spain, Sweden, United Kingdom, United States, and other nine countries and lower for Pakistan, Turkey, Malaysia, and other six countries, including China. **Conclusions:** Mistakes and data inconsistencies represent the main retraction reason; images and ethical issues show a growing trend, while plagiarism and overlap continue to represent a significant problem. There is a steady increase in QRP and QPP article withdrawals with a peak of 878 retractions issued in 2020. Retraction of articles seems to be a technology-dependent process. The number of citations of retracted articles shows a high impact of papers published by authors from certain countries. The number of retracted articles per country does not always accurately reflect the scientific impact of QRP/QPP articles. The distribution of retraction reasons shows structural problems in the

organization and quality control of scientific research, which have different images depending on geographical location, economic development, and cultural model.

Keywords

Questionable Research Practices, Questionable Publication Practices, Retracted Papers, Retraction Notes, PubMed

1. Introduction

“We publish this Note at this time because in the years that have passed since our original publication, and even now, the 1959 note has been quoted as a worthy piece of evidence. We ask to be spared the further embarrassment of having that earlier work cited in the reputable literature, and we hope we can spare other authors the labors of attempting to rationalize our aberrant data.” (Goldstein & Eastwood, 1966)

The above withdrawal note reflects an issue that continues to be relevant: a scientific article, once published, continues to exert influence and have an impact on the scientific community even if it is proven, after publication, that it contains errors that raise doubts on the conclusions, invalidating them totally or partially.

In 1756 appeared what is most likely the first documented withdrawal of scientific work (Wilfon, 1756), followed, over 150 years later, by that of Whelden (Whelden, 1927) and over time, by many others:

- unreproducible research of questionable scientific value persisted for decades before disappearing from the literature, although it was demonstrated to be wrong shortly after its publication (Deichmann & Müller-Hill, 1998);
- spectacular results were based on either a laboratory error or an attempt by a researcher to report positive preliminary results (Culliton, 1974);
- discoveries presented in an article with over 1000 citations in Google Scholar were questioned one month after publication, following an inspection that found severe problems of research methodology and that today would lead to a very rapid rejection or even would make the article impossible to publish (Maddox et al., 1988);
- for the first time, in addition to the retraction of published papers, the perpetrator of scientific fraud (the Poehlman case) was sentenced to prison (Tilden, 2010);
- hundreds of articles were withdrawn due to data fabrication, Schon (Service, 2002), Hwang (Kakuk, 2009), Boldt, Reuben, Fuji (McHugh & Yentis, 2019) being just a few examples.

Research misconduct/questionable research practices (QRP) in various forms (from 1.97% up to 72%) within a scientific community has been reported by

several studies (Baerlocher et al., 2010; Fanelli, 2009; Godecharle et al., 2018; Titus et al., 2008).

The costs generated by QRP can reach considerable amounts, conservatively evaluated, in 2010, to cost over 100 million USD only for the cases investigated by ORI in the United States (Michalek et al., 2010) or 400,000 USD/article in 2014 (Stern et al., 2014).

Indirect costs can also increase if no action is taken, such as the funding of research in which contaminated cell lines are being used (Buehring et al., 2004).

Apart from the specific situations in which the scientific value of some articles is being questioned, there is an area of questionable publication practices (QPP) in which the reasons for rejection are based instead on non-compliance with ethical standards (ethical writing) and legal regulations (copyright issues): plagiarism, overlap, authorship (Harriman & Patel, 2014; Horbach & Halffman, 2019; Roig, 2015; Scanlon, 2007). While plagiarism is rejected and considered a form of scientific misconduct, text re-use/recycling/self-plagiarism is still under debate about the quantity and type of recycled materials (Horbach & Halffman, 2019), and the decision to retract a scientific paper is mainly an editorial one.

If the mechanisms that should prevent the generation, perpetuation, and dissemination of QRP/QPP in biomedical research do not work correctly, there are situations in which harm can be caused to patients, the scientific community, research institutions, funding bodies, publishers, and scientific journals in which the results are published (Relman, 1983) (Parrish, 1999) (DeMets et al., 2017) (Selvam, 2021).

2. Background

Numerous articles have addressed the subject of retractions in recent years. Similar to papers that used other databases, those that used PubMed/Medline as their primary source of data (Nath et al., 2006) (Cokol et al., 2007) (Cokol et al., 2008) (Redman et al., 2008) (Steen, 2011) (Foo, 2011) (Wager & Williams, 2011) (Fang et al., 2012) (Decullier et al., 2013) (Singh et al., 2014) (Azoulay et al., 2015; Madlock-Brown & Eichmann, 2015) (Mongeon & Larivière, 2016) (Rosenkrantz, 2016) (Pantziarka & Meheus, 2019) (Rapani et al., 2020) have shown an increasing trend in the number of retracted papers, a diversification of retraction reasons (recently including image manipulations (Bucci, 2018)) and an increased interest of journals and publishers in correcting the scientific literature (Table 1).

Table 1. Summary of papers studying PubMed/Medline retractions.

Paper	Period	Nb of articles	Study features
(Budd et al., 1998)	1966-1997	235	Studied citations, post-retraction citations and citation types, type of citing publication, retraction reasons, time to retraction, who initiated the retraction. Main conclusion: the majority of citing articles cite the retracted papers as valid research.

Continued

(Nath et al., 2006)	1982-2002	395	The number of authors, type of study, date of publication, if retraction requested by authors. Analyzed misconduct vs mistakes: time to retraction 2 years for mistakes, 3.3 years for misconduct.
(Cokol et al., 2007)	1950-2004	596	Presents a model estimating the real number of papers that should be retracted, based on the impact factor of journals. Estimates the number of potentially retractable articles between 10,000 (best case) and 100,000 (worst case) out of 9,398,715 articles.
(Redman et al., 2008)	1995-2004	315	Analyzed pre and post-retraction citations, time to retraction, the primary reason for retraction, reproducibility as reason, and who requested the retraction. Main conclusion: pre-retraction highly cited studies continue to be highly cited after retraction. Significant differences between citations depending on retraction reason. Journal impact factor influence the number of citations.
(Steen, 2011)	2000-2010	788	Studied English language retractions: retraction reason, countries, repeat offenders, JIF, comparing fraud (data fabrication or falsification, data plagiarism) with errors (text plagiarism included). Conclusions: JIF higher for fraudulent articles, most of the retractions from USA.
(Wager & Williams, 2011)	1980-2009	312	Analyzed 312 out of 870 retractions: retraction reasons, who initiated the retraction, article types, number of authors. Conclusions: retracted articles have increased since 1980. Main retraction reasons are mistakes, overlap (redundant publication), and plagiarism, almost 90% of retractions are research articles, inconsistencies between journal retraction practices.
(Foo, 2011)	up to 2009	303	Analyzed only authors with at least two retractions (303 from 1239 PubMed records): time to retraction, citations, the number of authors. Conclusions: time to retraction decrease from 52 months (articles published before 2000) to 33 months (after 2000), almost 25% of citations received after article retraction.
(Samp et al., 2012)	2000-2011	102	Studied drug therapy retracted articles (102 out of 742 retractions): retraction reasons, time to retraction, funding sources, repeated retractions. Conclusions: mean time to retraction is 39 months, main reason is scientific misconduct (51%), 5.6 articles in average written by repeat offenders.
(Fang et al., 2012)	Up to 2012	2047	Studied retraction reasons, temporal trends, time to retraction, geographic origin, citations. Conclusions: fraud/suspected fraud/misconduct, plagiarism and redundant publication are the main retraction reasons. Mean time to retraction for all causes is 32.9 months. "USA, Germany, Japan and China account for 75% of papers retracted for fraud or suspected fraud".
(Decullier et al., 2013)	2008	250	Studied retraction reasons, author countries, time to retraction, visibility of retracted articles, retraction notes characteristics, journals impact factor. Identified mistakes, plagiarism and fraud as the main causes for retraction and articles with mistakes being published in high impact factor journals.

Continued

(Singh et al., 2014)	2004-2013	2343	PubMed retracted articles data: retraction reasons, article types, time to retraction. Conclusions: main reasons are mistakes, plagiarism, and duplicate publication. Forty-five percent of the retracted articles represent original research, 33% clinical cases and 18% reviews. A decrease over time of the period between publication and retraction was noticed.
(Madlock-Brown & Eichmann, 2015)	2003-2010	1113	This article studied reasons for retraction, citations (for 740 papers matched in the WoS database), and those who initiated the retraction. Conclusions/findings: mistakes, overlap (duplication) and plagiarism are the top 3 retraction reasons. Self cites represent 5% of the retracted article citations. Authors requested retraction in almost 40% of cases, editorial retractions account for 60%.
(Azoulay et al., 2015)	1973-2007	1104	Studies a set of PubMed retractions for retraction reasons, retraction speed, citations and the impact of these papers on related articles. Findings: retraction speed is 3.2 years; main retraction reasons are: fake data (32.7%), mistakes/errors (24.55%), duplicated publication (8.33%) and plagiarism (8.15%).
(Rosenkrantz, 2016)	1983-2015	48	The article analyzed retracted papers published in radiology journals extracted from 3685 retracted PubMed articles: time to retraction, retraction reasons, citations Findings: mistakes, duplicate publication and plagiarism were the main retraction reasons, retraction notes inconsistency between journals.
(Pantziarka & Meheus, 2019)	2000-2018	1512	Retracted oncology articles research: lifetime, citations, authors. Findings: mean lifetime is 1318 days, the median is 870 days. Mean citation number is 35.1/article (median = 14). Multiple retractions/author in 14% of articles.
(Rapani et al., 2020)	up to 2018	180	Retracted dental articles: retraction reason, type of study, citations (PubMed and Retraction Watch). Findings: mean time to retraction = 2.1 years (median 1 year), 55% of retracted articles originate from Asia, 28% from Europe and 10% from America. Mean number of authors is 4.6. Author misconduct, scientific errors and publisher related issues are the main retraction reasons.
(Bhatt, 2021)	1959-2019	6936	PubMed retractions combined with WoS information: retraction trends over time, time until retraction, trends among countries. Findings: 3.8 years average time until retraction, retraction rate per 10,000 publications increased from 0.38 in 1985 to 5.95 in 2014, China authors publish in low impact journals while authors from US are published in high impact journals.

The articles withdrawn from the scientific literature indexed in PubMed, although in a small proportion to the total volume indexed in this database, are a problem not only by question marks which they can raise on the integrity of the scientific research as a whole but also by the impact on the scientific community, which can use or invest in ideas, methods or data invalidated a few years after publication.

Taking into account the absence of recent studies, we decided to perform an

exhaustive exploratory analysis of human health-related papers withdrawn from the literature indexed in PubMed/Medline and published in the period 2009-2020 focused mainly on:

- dynamic of retracted articles and retraction notes for the period 2009-2020;
- retraction reasons with particular attention to image issues;
- countries producing QRP/QPP articles;
- the scientific impact of retracted papers;

At the beginning of this project, as far as we know, there was no exhaustive study about retracted articles indexed in PubMed (in the meantime, such a study has been published: it provides information on the evolution in time of the number of withdrawn articles and a ranking by countries without analyzing the reasons for withdrawing the articles and the number of their citations (Bhatt, 2021)).

Most articles published until now address either a limited number of withdrawn articles or less recent periods.

There is scarce information regarding the dynamics of retraction notes, the dynamic of retracted papers for image problems, and the scientific impact (citations) of withdrawn articles.

Considering these elements, we think that a snapshot of the withdrawals from the last 12 years using freely available alternative information sources could provide helpful information to those interested in the evolution and changes over time in this field.

3. Materials and Methods

3.1. Information Sources

- PubMed—PubMed (nih.gov)
- Google Scholar—citations <https://scholar.google.com>
- Dimensions—<https://dimensions.ai>
- Scopus—<https://scopus.com>

Several elements determined the selection of Pubmed as the unique source of information:

- focus on biomedical journals;
- unrestricted access;
- the existence of a dedicated keyword for withdrawn articles, Retracted Publication [PT];
- the link between the withdrawn article and the withdrawal note, Retraction of Publication [PT];
- the possibility of exporting the data in csv format in our database for the individual analysis of each article and the corresponding retraction note.

We have used Google Scholar due to its free nature and the best coverage of citations (Martín-Martín et al., 2020).

We have used Dimensions database due to free access and provision, in addition to the total number of citations, of the number of citations from the last two

years (compared to the current date).

SCOPUS (Elsevier) was used to obtain journal publishers and the journal impact factor (CiteScore) information. Because CiteScore is built based on the number of citations from the last four years, we decided to analyze a time interval that would cover three segments of 4 years each in terms of citations (2009-2020 interval, composed of 2009-2012, 2013-2016 and 2017-2020). This allows us to develop further directions of analysis of the data set obtained.

3.2. Articles Retrieval and Extracted Information

Withdrawn articles were identified in PubMed using “Retracted Publication [PT]” search without date restrictions. The data were downloaded in csv format and imported into an application developed for analysis by the author.

Data analysis period: 20.07.2020-31.05.2021

Last date of data import from PubMed: 30.01.2021

The period analyzed: 01.01.2009-31.12.2020 (taking into account the year of publication as recorded in PubMed).

The processing date was noted for each item.

Inclusion criteria: The field of study or the subject studied is related or may have an impact on human health (mentioned in the text of the article).

Exclusion criteria: The field of study or the studied subject is not related to human health (chemistry, agriculture, veterinary medicine, industrial products, ecology without mentioning in the article some implications with human health) proceedings volumes with no specific retractions mentioned, clinical practice guidelines withdrawn for updates or unspecified reasons, misclassified retractions.

Each item and associated retraction note were analyzed. The extracted data is grouped into four sections: authors and countries, retraction details (including those involved in retraction, **Table 6**), collateral damage, and citations info (**Tables 2-5**).

1) Authors and countries.

Table 2. Author and country-related variables.

1	Number of authors	Number of authors who contributed to the study.
2	International	Whether the authors are from several countries or not, or whether they are part of a research group.
3	Same organization	Whether or not the authors belong to the same organization.
4	Country of the first author	The country of the institution to which the first author is affiliated. In the case of mentioning several institutions, the first affiliation was considered.
5	Country of the other authors	The countries of the institutions to which the other authors are affiliated. In the case of mentioning several institutions, the first affiliation was considered.

2) Retraction details.

Table 3. Retraction-related variables (RN = Retraction Note; IRB = Institutional Review Board; ORI = Office of Research Integrity).

6	Exposure time (ET)	Time, measured in months, between the date of withdrawal of the article and the date of publication (latest data)
7	PMID of retraction note	The PubMed identifier of the withdrawal
8	Date of RN	Date of retraction note
9	Doi of RN	The DOI identifier of the withdrawal note
10	Paper type	Paper type (journal article, proceedings, book chapter)
11	Study type	case reports, letters/editorials, research article, review, randomized trial, meta-analysis, methods/methodology
12	Clinical	Source of data used in the article (human subjects, tissues, laboratory, big data)
13	Available online	If the entire article is online/not/paywalled
14	Publisher	The journal publisher as reported in Scopus. In case of publisher change (by a takeover of the journal), the publisher was considered from the moment of carrying out this study
15	Retraction reason	Main retraction reasons: authorship, editorial, ethics, fraud, images, mistakes and or inconsistent data, overlap, plagiarism, property or legal concerns, other, unclear
16	Retraction details	Text of retraction note (up to 2000 characters)
17	Image reasons	Details for image related retractions (copyright, duplication, ethical, fabrication, manipulation, overlap, plagiarized, unreliable, wrong)
18	Reason subcategory	Second level reasons, used when information is available in the retraction note: conflict of interest, duplicate, fabricated data, falsified information, fraudulent peer review, no raw data, no IRB approval, no patient consent, publication w/o approval, reproducibility, research misconduct, research outsourcing, other
19	Requested by/Involving	Who asked for/was involved in article retraction. Authors, publisher, editorial (editors or editorial board), institution, O.R.I., other, not specified

3) Collateral damage.

Table 4. Collateral damage variables.

20	Retraction note info	Whether or not the withdrawal note contains information related to the individual involvement of the authors
21	1st author	The first author responsible
22	Corresponding author	The corresponding author responsible
23	Authors involved	How many authors were involved
24	Innocent casualties	How many innocent authors (not responsible)

4) Citations.

Table 5. Citations-related variables.

25	Journal cite score	Cite Score in the publication year
26	G scholar citations	Google Scholar citations
27	Dimensions citations	Dimensions citations
28	Dimensions recent citations	Dimensions recent citations

5) Who was involved/requested the retraction.

Table 6. Who was involved in the retraction process?

Authors	-	One/more/all authors
Publisher	-	Journal publisher
Editor in chief	-	Editor in chief decision to withdraw the article
EB	-	Editorial board
Institution	-	Involvement of the institution where the research was conducted (investigation of research work on which the article is based);
ORI	-	Article retraction was requested after an investigation of the U.S. Office for Research Integrity

6) Definitions.

In order to define a flexible taxonomy for the main retraction reasons categories and subcategories (see **Table 7** and **Table 8**), we have considered several previously published articles (Azoulay et al., 2015; Benson, 2016; Decullier et al., 2013; Wager & Williams, 2011; Zhang et al., 2020).

7) Retraction reasons.

Table 7. Main retraction reasons.

AUTHORSHIP	<ul style="list-style-type: none"> - disputes/disagreements between authors related to the content of the article - publication of the article without the consent or knowledge of all authors - listing as authors of people who are not aware of the research - changing authors after article submission
EDITORIAL	<ul style="list-style-type: none"> - duplicate publication of an article in the same journal - other editorial errors (publication before peer review, publication of a wrong version)
ETHICS	<ul style="list-style-type: none"> - lack of patient/patients' consent - withdrawal of consent by the patient/patients after the publication of the article - lack of IRB. approval from the institution where the research was conducted/non-compliance with the protocol - failure to comply with the protocol for which the initial IRB approval was obtained - violation of other ethical norms (conflicts of interest, laboratory animals ethics) - publication of an article in which data are used without approval (non-ethical use of data)

Continued

FRAUD	<ul style="list-style-type: none"> - manipulating the peer review process by different methods (fake email addresses, fake peer review reports) - forged information (authors, grants, forged ethics agreements, forged consent) - the use in the article of unauthorized or illegally obtained information
IMAGES	<ul style="list-style-type: none"> - any situation in which one of the reasons for withdrawing the article was related to the images used in it
MISTAKES/INCONSISTENT DATA	<ul style="list-style-type: none"> - any type of error that invalidates the results in the article - fabrication of data or results
OVERLAP	<ul style="list-style-type: none"> - re-use of texts, data, or other elements of the article contained in an article published by the same author or group of authors (does not apply to images, which are analyzed separately) - publishing the same article in two or more different journals
PLAGIARISM	<ul style="list-style-type: none"> - use of any information contained in an article already published by other authors (does not apply to images, which are analyzed separately)
PROPERTY OR LEGAL CONCERNS	<ul style="list-style-type: none"> - article in which information protected by copyright or other legal means is used may entail legal liability for the author/publisher/publisher
OTHER	<ul style="list-style-type: none"> - other reasons that do not fall into those described above

8) Secondary reasons.**Table 8.** Secondary retraction reasons.

CONFLICT OF INTEREST	Conflicts of interest not declared
DUPLICATE	Duplicate publication (editorial mistake) or duplicate submission (author mistake)
FABRICATED DATA	Subcategory of main category Mistakes/Inconsistent data for articles retracted for data fabrication
FALSIFIED INFORMATION	False affiliations, false contact details, falsified documents, credentials, others
FRAUDULENT PEER REVIEW	Circumvention or hacking of peer review process
NO DATA PROVIDED	No raw data for the study were available or made available by the authors No raw data found following the institutional investigation
NO IRB APPROVAL	Nonexistent IRB approval for the study Study not respecting the protocol approved by IRB
NO PATIENT CONSENT	Patient consent incomplete, not given, or withdrawn
PUBLICATION W/O APPROVAL	No publication approval from other authors (coauthor, senior author), institution, funder, or company
REPRODUCIBILITY	Reporting by the authors of the impossibility to replicate, totally or partially, the results presented in the article Reporting, by other authors, of the impossibility to replicate the research results from the article
RESEARCH MISCONDUCT	Research misconduct was reported by one of the institutions where the research was conducted, by one of the coauthors, or by a 3rd party
RESEARCH OUTSOURCING/TEMPLATED	Research performed partially/totally by a 3rd party Templates being used for research reporting
OTHER	Various other reasons which do not fit within an existing subcategory

The data were exported from the application and analyzed in IBM SPSS (IBM SPSS Statistics for Windows 2018).

4. Results

4.1. Retracted Articles and Retraction Notes

A total of 5619 retracted papers were retrieved by a PubMed search for the period 2009-2020. Of these, 775 were excluded and 4844 analyzed. The distribution for the period 2009-2020 of the withdrawn articles and the withdrawal notes is presented in **Figure 1** and **Table 9**.

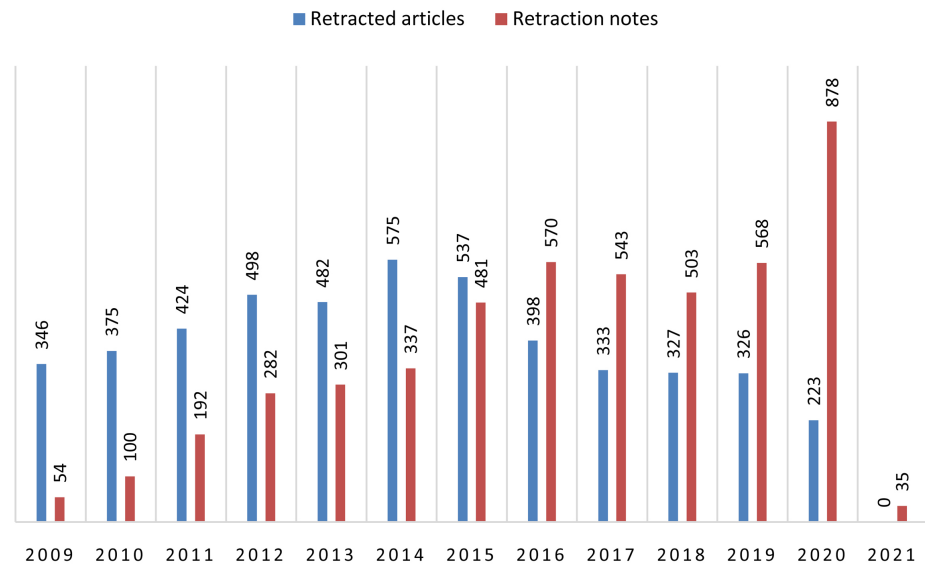


Figure 1. Retracted articles and retraction notes.

Table 9. Retracted articles and retraction notes by year (based on PubMed publication year value and PubMed reported retraction note date; 20 articles present date errors because the date of the electronic publication was far ahead of the print publication/journal date. These errors are not reflected in the ET calculations).

	Retraction note year													Total
	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021	
2009	54	52	58	32	31	32	9	24	10	14	11	16	3	346
2010	0	45	87	52	38	25	25	33	12	19	13	25	1	375
2011	0	0	50	108	52	30	37	37	26	31	31	21	1	424
2012	0	0	0	90	110	54	45	59	42	27	35	34	2	498
2013	0	0	0	0	66	118	75	53	44	33	40	50	3	482
2014	0	0	0	0	2	76	175	65	92	49	57	59	0	575
2015	0	0	0	0	0	2	112	171	113	59	38	42	0	537
2016	0	0	0	0	0	0	3	125	111	54	48	54	3	398
2017	0	0	0	0	0	0	0	3	92	103	67	64	4	333
2018	0	0	0	0	0	0	0	0	1	108	113	101	4	327
2019	0	0	0	0	1	0	0	0	1	4	114	200	6	326
2020	0	0	0	0	0	0	0	0	0	0	3	212	8	223
Total	54	97	195	282	300	337	481	570	544	501	570	878	35	4844

4.2. Retraction Reasons

Out of the 4844 analyzed retractions, 4251 (87.76%) have a unique retraction reason, and 593 (12.24%) have multiple retraction reasons (see **Table 10**).

Table 10. Retraction reasons (Column 1 displays the number of papers with a single reason as the basis for retraction. Columns 2, 3, 4 contain the number of papers with 2, 3 and 4 concurrent reasons as the basis for retraction).

	Retraction Reason	All	1	2	3	4	% of 4844 papers	
1	Mistakes/inconsistent data	1553	1247	274	31	1	32.06%	
2	Images	1088	887	188	13	-	22.46%	
3	Plagiarism	663	519	115	27	2	13.69%	
4	Overlap	556	441	94	19	2	11.48%	
5	Fraud	393	317	51	24	1	8.11%	
6	Ethics	360	196	147	17	-	7.43%	
7	Authorship	281	83	147	49	2	5.78%	
8	Unclear	247	246	1	-	-	5.1%	
9	Editor	181	176	5	-	-	3.73%	
10	Property & Legal	121	84	29	7	1	2.5%	
11	Other	63	55	7	1	-	1.30%	
		4251						

Secondary retraction reasons are presented in **Table 11**.

There were 229 instances of data fabrication, 217 in the “Mistakes/Inconsistent data” category, and 12 in other categories.

We have identified 286 instances of duplicate publication, 123 were editorial/publisher mistakes, and 163 were duplicate submissions.

Within the Images category, we found 253 instances of image overlap and 94 instances of plagiarized images. When taking into account images overlap and plagiarism, the total number of overlap articles (text and images) is 809 and the total number of plagiarism articles (text and images) is 757. Image reasons count 741 cases when image overlap and plagiarism are being moved in the overlap and plagiarism categories.

Fraudulent peer review was found in 350 instances (main category: Fraud).

The authors were unable to provide the raw data or the raw data that could not be retrieved in 293 cases.

Table 11. Secondary retraction reasons.

Secondary retraction reasons	N
Conflict of interest	31
Duplicate (submission by authors or editorial error)	286
Fabricated data	229
Falsified information	39
Fraudulent peer review	350
No data provided	293
No IRB approval	134
No patient consent	22
Publication without approval	97
Reproducibility	180
Research misconduct	47
Research outsourcing/templated	22
Other	14

4.3. Citations

Citations were counted for all retracted papers (see **Table 12**). Overall, 140,810 citations were retrieved in Google Scholar and 96,000 in Dimensions (68% ratio Dimensions/Google Scholar).

Table 12. Citations for all 4844 retracted papers.

	Mean (95% CI)	Median	SEM	SD	Range	Sum
Google Scholar	29.07 (26.43 - 31.71)	11	1.346	93.659	4482	140,810
Dimensions	19.82 (18.08 - 21.55)	7	0.884	61.547	2860	96,000

4.4. Exposure Time

Exposure time (time difference between the most recent retraction date and the earliest publication date, expressed in months) was collected for all articles included in the study (**Table 13**). The average exposure time was 28.89 months with a median value of 19 months.

Table 13. Exposure time for all 4844 retracted papers.

	Mean (95% CI)	Median	SEM	SD	Range
Exposure time (months)	28.89 (28.10 - 29.68)	19	0.403	28.052	155

4.5. Retraction Reasons, Exposure Time, Average Citations/Article, the Average Number of Authors, and Average CiteScore (Table 14)

When analyzing the exposure time, the average number of citations, the average number of authors, and the average CiteScore from the year of publication, depending on the reason for withdrawal, there are some interesting aspects:

- articles withdrawn due to images have a much higher ET than other articles, more citations, the highest number of authors, and the third-highest value of CiteScore;
- articles withdrawn for ethical reasons have the second value of ET but much fewer citations and a CiteScore below average;
- mistakes/inconsistent data have the third ET as value (but below average), citations and authors above average, and the third CiteScore as value.

Table 14. Main retraction reason categories (all instances): exposure time (ET), average Google Scholar citations (G), average Dimensions citations (D), the average number of authors (A), and the average value of CiteScore (CS).

	Retraction Reason	All	ET (95% CI)	G	D	A	CS
1	Mistakes/inconsistent data	1553	26.95 (25.66 - 28.24)	35.51	24.63	6.4	8.06
2	Images	1088	49.21 (47.27 - 51.15)	44.57	32.9	7.05	7.64
3	Plagiarism	663	24.11 (22.52 - 25.7)	27.84	16.15	4.64	3.20
4	Overlap	556	24.17 (22.18 - 26.16)	20.51	12.37	5.33	4.04
5	Fraud	393	22.32 (20.9 - 23.74)	12.87	9.24	5.59	3.49
6	Ethics	360	31.76 (28.32 - 35.21)	20.69	13.45	5.43	4.53
7	Authorship	281	17.38 (15.04 - 19.72)	11.41	7.23	5.53	3.93
8	Unclear	247	18.1 (15.31 - 20.89)	12.21	7.52	4.97	3.99
9	Editor	181	11.26 (8.7 - 13.82)	6.92	4.12	5.1	4.27
10	Property & Legal	121	12.7 (10.28 - 15.02)	10.4	6.98	4.99	4.75
11	Other	63	19.62 (12.92 - 26.32)	18.24	12.25	4.62	13.51
	All retractions (N = 4844)		28.89 (28.10 - 29.68)	29.07	19.82	5.83	6.03

The order and characteristics of retractions remain almost the same when we analyze single retraction reasons articles (**Table 15**). However, we observe two exceptions. Articles retracted for ethical reasons represent only 54% of the total, meaning that ethics is probably part of more complex retractions involving multiple reasons. Authorship displays the same pattern, with only 29% of cases representing the single retraction reason which could also mean that authorship is more frequent in multiple/complex retractions, being only one of the factors contributing to article withdrawal.

Table 15. Main retraction reasons (papers with one retraction reason only. N = 4251): exposure time (ET), average Google Scholar citations (G average), average Dimensions citations (D average), the average number of authors (Authors), and the average value of CiteScore (CS).

	Retraction Reason	Single reason	ET (95% CI)	G	D	A	CS
1	Mistakes/inconsistent data	1247	25.24 (23.89 - 26.6)	37.16	25.88	6.31	8.51
2	Images	887	52.01 (49.83 - 54.19)	47.29	34.96	7.16	7.67
3	Plagiarism	519	22.38 (20.64 - 24.12)	29.77	16.67	4.10	3.04
4	Overlap	441	22.74 (20.58 - 24.89)	20.02	11.49	5.21	3.72
5	Fraud	317	22.83 (21.29 - 24.38)	12.56	9.11	5.21	3.37
6	Unclear	246	18.14 (15.34 - 20.94)	12.26	7.55	4.96	4
7	Ethics	196	32.35 (27.56 - 37.13)	18.32	11.6	5.50	4.33
8	Editor	176	11.45 (8.83 - 14.08)	7.08	4.20	5.14	4.35
9	Property & Legal	84	13 (10.16 - 15.94)	10.35	6.95	5.51	4.98
10	Authorship	83	14.14 (10.15 - 18.14)	9.24	5.75	5.12	3.88
11	Other	55	19.36 (12.06 - 26.67)	18.98	13	3.73	14.55

4.6. Number of Authors

The average number of authors for the 4844 articles is 5.83 (5.73 - 5.93) with a median of 5, mode 4, IQR 4, and range 36. The distribution of articles according to the number of authors is displayed in **Figure 2**.

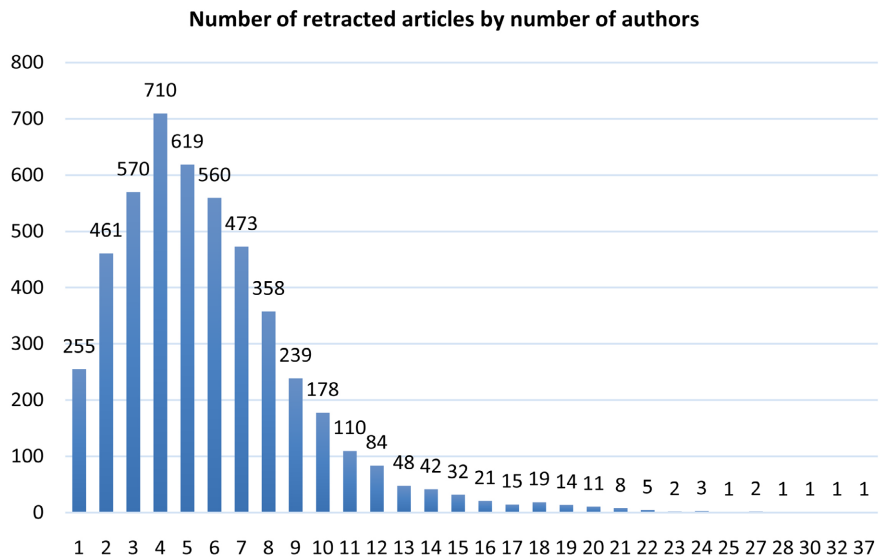


Figure 2. The number of retracted papers (n = 4844) by the number of authors.

Of the 4844 articles analyzed, 4103 (84.7%) were published by authors from the same country and 741 (15.3%) by authors from 2 or more countries. Two hundred fifty-five (5.26%) articles have only one author, 4589 (94.74%) have two or more authors.

The average number of authors (**Figure 3**) varied between 5.48 (2020) and 6.27 (2018).

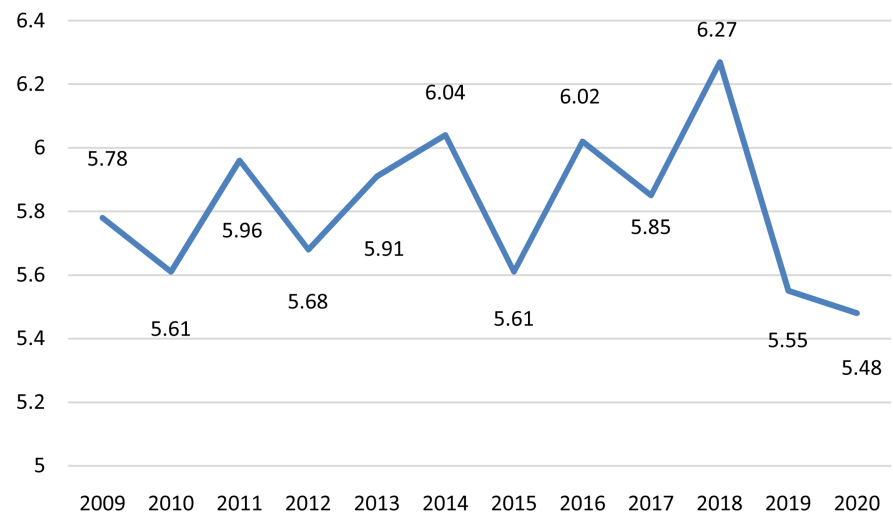


Figure 3. The average number of authors by year.

4.7. Retracted Articles and Retraction Notes by Country of the First Author

The 4844 articles had the first authors from 92 countries. The top 30 countries have 4592 retractions (94.79% of total retractions).

The number and evolution of retractions and retraction notes for the top 30 countries are presented in **Table 16**.

Table 16. Yearly distribution of retracted papers (R) & retraction notes (RN) by country of the first author.

Country & Region		T	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021
China	R	1587	42	40	42	67	146	256	267	159	142	136	160	130	
	RN		16	16	28	30	46	65	203	169	234	138	203	423	16
United States	R	918	100	107	150	118	86	79	74	53	46	39	38	28	
	RN		13	19	46	52	49	59	108	123	105	109	96	135	4
India	R	351	23	30	38	51	42	53	37	25	22	12	13	5	
	RN		2	3	21	31	36	32	32	35	22	35	48	51	3
Japan	R	212	35	34	31	26	12	14	14	14	9	11	7	5	
	RN		4	6	19	17	20	17	15	17	12	31	19	34	1
Italy	R	182	12	20	15	32	36	21	11	10	4	9	8	4	
	RN		0	4	8	15	8	24	12	13	18	20	35	23	2
South Korea	R	166	7	13	31	22	18	9	10	14	11	15	9	7	
	RN		2	2	2	46	15	8	9	15	10	13	13	31	0
Iran	R	162	6	10	6	16	16	25	31	17	11	11	9	4	
	RN		1	4	4	2	10	9	18	57	10	9	22	15	1
Germany	R	115	29	17	14	11	6	5	8	7	3	5	4	6	
	RN		1	7	21	15	8	8	11	8	12	0	9	14	1
United Kingdom	R	110	13	13	7	14	9	13	8	6	8	13	3	3	
	RN		2	4	7	10	9	8	9	11	10	14	9	17	0
Spain	R	78	6	3	5	15	6	5	13	8	3	6	2	6	
	RN		1	1	2	5	4	3	3	13	8	19	5	13	1
Canada	R	77	7	9	8	8	8	4	5	7	8	4	6	3	
	RN		0	3	2	5	7	9	5	13	6	9	5	11	2
France	R	59	3	7	3	7	7	4	2	3	9	7	5	2	
	RN		0	1	4	2	7	6	2	5	4	11	7	10	0
Egypt	R	55	3	7	5	7	4	7	3	6	2	8	3	0	
	RN		0	1	6	2	8	6	2	6	3	7	8	6	0
Australia	R	55	7	10	1	6	4	5	1	6	3	2	8	2	
	RN		3	6	4	2	3	3	3	8	6	6	5	6	0
Sweden	R	54	4	3	5	4	10	8	3	4	3	3	4	3	
	RN		1	2	1	2	3	6	5	3	5	10	8	8	0
Turkey	R	52	2	4	3	4	5	7	9	6	3	5	3	1	
	RN		1	2	1	5	4	7	8	6	6	3	4	5	0
Brazil	R	52	3	5	5	14	6	3	6	2	2	2	2	2	
	RN		0	0	0	2	8	4	1	17	6	4	5	5	0
Chinese Taipei	R	49	2	2	2	6	9	7	5	7	1	4	3	1	
	RN		0	0	1	1	4	2	4	7	11	5	5	8	1
Netherlands	R	40	4	5	12	2	2	2	1	4	3	1	3	1	
	RN		0	0	2	9	5	4	2	5	4	2	3	3	1
Saudi Arabia	R	36	1	2	5	2	6	2	3	0	3	9	2	1	
	RN		1	1	1	1	2	4	1	3	2	5	7	6	2
Singapore	R	31	4	4	3	6	2	3	2	4	2	0	1	0	
	RN		0	1	2	0	3	4	3	5	1	5	3	4	0
Switzerland	R	28	4	1	3	7	3	2	4	2	0	0	1	1	
	RN		2	0	0	1	7	5	1	1	6	3	1	1	0

Continued

Poland	R	18	0	0	0	1	2	3	2	6	2	0	0	
	RN		0	0	0	0	0	1	3	1	4	3	5	1
Pakistan	R	18	2	1	0	2	1	0	0	2	2	4	2	2
	RN		1	1	1	2	2	0	0	1	2	1	4	3
Israel	R	16	0	2	3	3	1	3	0	1	0	1	2	0
	RN		0	0	2	2	1	0	1	0	3	3	3	1
Malaysia	R	15	0	1	0	3	1	3	1	5	1	0	0	0
	RN		0	0	1	1	2	0	0	3	2	5	0	1
Greece	R	15	3	0	2	2	1	3	1	2	0	1	0	0
	RN		1	0	1	0	2	2	2	2	3	1	0	1
Russian Federation	R	14	0	0	1	2	1	4	2	1	3	0	0	0
	RN		0	0	0	0	2	0	0	0	6	1	4	1
Norway	R	14	0	1	1	1	3	2	0	1	2	2	0	1
	RN		0	0	0	1	1	0	1	0	0	3	3	5
Ireland	R	13	1	1	2	3	2	2	0	1	1	0	0	0
	RN		0	0	0	0	1	3	0	1	3	4	1	0

4.8. Retraction Reasons for Top 10 Countries, First Author Country

When we look at the proportions represented by the reasons for withdrawal, we notice notable differences between countries, probably due to different stages of development both at the institutional level and of national policies (Table 17).

The inconsistency of retraction notes affects all countries equally and editorial errors, in a not negligible percentage, seem to affect the UK and Germany.

Table 17. Retraction reasons for top 10 countries (first author affiliation considered).

Country	N	MISTAKES	IMAGES	PLAGIARISM	OVERLAP	FRAUD	ETHICS	AUTHORSHIP	UNCLEAR	EDITOR	PROPERTY	OTHER
China	1587	28.6%	23.2%	14.7%	9.3%	18.3%	5.7%	6.7%	5%	1.9%	1.4%	0.5%
United States	918	45.2%	31.3%	3.8%	7.3%	0.5%	5.4%	1.6%	3.8%	4.2%	2.8%	1.5%
India	351	17.9%	24.8%	35.3%	17.1%	1.1%	5.1%	4.3%	5.7%	2.6%	1.1%	0.3%
Japan	212	39.2%	19.8%	4.7%	16.5%	0.5%	15.6%	9%	5.7%	3.8%	2.8%	0.5%
Italy	182	12.1%	27.5%	32.4%	12.1%	0.5%	6%	3.3%	5.5%	4.4%	3.8%	2.7%
South Korea	166	33.1%	10.8%	6.6%	15.1%	16.9%	12%	4.8%	4.8%	2.4%	4.2%	0.6%
Iran	162	14.2%	4.9%	29.6%	22.2%	32.7%	4.9%	23.5%	5.6%	5.6%	1.2%	0.6%
Germany	115	36.5%	7.8%	1.7%	15.7%	0	20.9%	7%	4.3%	6.1%	5.2%	2.6%
United Kingdom	110	39.1%	20.9%	5.5%	6.4%	0	9.1%	9.1%	4.5%	8.2%	5.5%	3.6%
Spain	78	26.9%	37.2%	9%	11.5%	1.3%	3.8%	5.1%	5.1%	3.8%	6.4%	1.3%

Retraction reasons for top 10 countries, all authors countries (**Table 18**).

Table 18. Retraction reasons for top 10 countries (all author countries).

Country	N	MISTAKES	IMAGES	PLAGIARISM	OVERLAP	FRAUD	ETHICS	AUTHORSHIP	UNCLEAR	EDITOR	PROPERTY	OTHER
China	1624	28.9%	23.5%	14.4%	9.4%	18%	5.8%	6.7%	4.9%	1.8%	1.5%	0.5%
United States	1175	43.1%	31.5%	4.9%	7.6%	1.4%	6.1%	2.8%	3.7%	4.2%	2.6%	1.7%
India	384	18.2%	27.1%	33.6%	16.1%	1.3%	4.9%	4.7%	5.7%	2.3%	1.3%	0.3%
Japan	241	40.2%	20.7%	5%	14.9%	1.2%	14.1%	7.9%	5.8%	3.3%	2.5%	0.4%
Italy	227	21.6%	29.1%	26%	11%	0.9%	5.7%	2.6%	4.4%	4%	3.5%	2.2%
United Kingdom	194	43.8%	18%	6.2%	7.2%	0.5%	11.3%	8.2%	2.6%	5.2%	4.1%	4.6%
South Korea	182	32.4%	14.8%	7.7%	13.7%	15.4%	11%	4.4%	4.4%	2.2%	3.8%	0.5%
Germany	177	37.3%	16.4%	2.8%	12.4%	1.7%	16.4%	9%	2.8%	5.1%	5.1%	4%
Iran	174	14.9%	6.3%	27.6%	21.8%	31%	6.3%	23%	6.3%	5.2%	1.1%	0.6%
Canada	109	31.2%	20.2%	4.6%	15.6%	2.8%	8.3%	5.5%	2.8%	11%	3.7%	5.5%

4.9. Impact of Retracted Articles

In order to estimate the impact of retracted research, we calculated for each country the number of articles and percentage from total retractions (%A), the number of Google Scholar citations and percentage from all citations (%G), the number of Dimensions citations and percentage from all Dimensions citations (%D), the ET and CiteScore average (see **Table 19** for proposed grading).

Seven categories were defined: four marked green (positive or stationary evolution) and three red (negative evolution).

The high impact was considered any difference between %G or %D and %A greater than 25% (with steps at >25%, >50%, and >75%).

Low impact was considered any difference between %G or %D and %A less than 25%.

Values between -25% and +25% were considered stationary/neutral.

The impact assessment was made for:

- first author country of origin with (**Table 20**) and without editorial errors (**Table 21**).
- all authors countries of origin with (**Table 22**) and without editorial errors (**Table 23**).

Table 19. Impact grading.

Impact much smaller than expected (percentage of citations is smaller with more than 75% of the percentage of articles)		Minimal impact
The impact is smaller than expected (percentage of citations smaller with 50% - 75% of the article percentage)		Very low impact
Small impact (-25% - 50%)		Low impact
Stationary/Neutral (-25% - +25%)		Neutral
Increased impact (between +25 and +50%)		Increased impact
Large impact (between +50 and +75%)		Large impact
Very large impact (percentage of citations is bigger with more than 75% than the percentage of retracted articles)		Very large impact

Table 20. Retracted research impact for top 30 countries (first author, all retraction reasons). = >-25% - <+25%; = <25 up to 50%; = <50% - 75%; = <75%; = >25 up to 50%; = >50% - 75%; = > 75%; CS—CiteScore average (3933 records with available information); ET—Exposure Time average.

	Country & Region	N	%A	G	%G	Impact	D	%D	Impact	CS	ET	Authors
1	China	1587	32.76%	22,973	16.31%		17,222	17.94%		4.39	22.93	6.2
2	United States	918	18.95%	48,931	34.75%		34,019	35.44%		10	39.15	5.8
3	India	351	7.24%	13,231	9.39%		7708	8.03%		4.38	31.74	4.5
4	Japan	212	4.37%	5477	3.89%		3645	3.79%		5.67	39.24	7.1
5	Italy	182	3.75%	4821	3.43%		3283	3.42%		4.93	38.14	6.4
6	South Korea	166	3.42	2606	1.85%		1718	1.79%		4.44	21.61	5.3
7	Iran	162	3.34	3346	2.37%		2072	2.16%		3.34	23.04	5
8	Germany	115	2.2%	3100	2.2%		2044	2.13%		5.78	26.96	5.6
9	United Kingdom	110	2.27%	5461	3.88%		3893	4.05%		13.04	27.15	4.7
10	Spain	78	1.61%	7515	5.33%		5040	5.25%		8.92	33.94	6.6
11	Canada	77	1.59%	1570	1.12%		1050	1.09%		7.26	27.91	5.6
12	France	59	1.22%	1785	1.27%		1313	1.36%		10.44	24.29	6.7
13	Egypt	55	1.13%	2410	1.71%		1476	1.54%		4.64	25.8	4.2
14	Australia	55	1.13%	1018	0.72%		1050	1.09%		6.28	19.25	4.4
15	Sweden	54	1.11%	1795	1.27%		1346	1.4%		10.43	29.74	6.6
16	Brazil	52	1.07%	1994	1.42%		1277	1.33%		5.27	36.42	6.8
17	Turkey	52	1.07%	418	0.3%		201	0.21%		2.91	13.9	4.9
18	Chinese Taipei	49	1.01%	1037	0.73%		728	0.76%		8.13	35.76	6.3
19	Netherlands	40	0.83%	2043	1.45%		967	1%		8.57	26.23	5.8
20	Saudi Arabia	36	0.74%	593	0.42%		424	0.44		3.61	29.14	4.1
21	Singapore	31	0.64%	1410	1%		1027	1.07%		8.28	40.58	5.9
22	Switzerland	28	0.58%	973	0.69%		625	0.65%		13.66	25.14	6.5
23	Poland	18	0.37%	261	0.18%		190	0.2%		4.85	23.94	6.7
24	Pakistan	18	0.37%	73	0.05%		49	0.05%		3.7	7.06	5.7
25	Israel	16	0.33%	402	0.28%		285	0.29%		8.95	31.5	8.8
26	Greece	15	0.31%	236	0.17%		125	0.13%		7.35	24.8	4.5

Continued

27	Malaysia	15	0.31%	195	0.14%	● ●	125	0.13%	● ●	3.08	21.8	5.7
28	Norway	14	0.29%	377	0.27%	➔	251	0.26%	➔	4.7	40.57	7.1
29	Russian Federation	14	0.29%	211	0.15%	●	111	0.11%	● ●	4.13	35.21	5.5
30	Ireland	13	0.27%	503	0.36%	●	316	0.33%	➔	7.05	49.31	4.6
Totals/averages		4844		140,810			96,000			6.04	28.89	5.83

Table 21. Retracted research impact for top 30 countries (first author, without editorial retraction reason). ➔ = >-25% - <+25%; ● = <25 up to 50%; ● ● = <50% - 75%; ● ● ● = <75%; ● = >25 up to 50%; ● ● = >50% - 75%; ● ● ● = > 75%; CS—CiteScore average (3933 records with available information); ET—Exposure Time average.

	Country & Region	N	%A	G	%G	Impact	D	%D	Impact	CS	ET	Authors
1	China	1557	33.39%	22,726	16.28%	● ●	17,060	17.91%	●	4.4	23.17	6.2
2	United States	883	18.93%	48,717	34.90%	● ● ●	33,876	35.56%	● ● ●	10.21	40.35	5.9
3	India	342	7.33%	13,179	9.44%	●	7684	8.06%	➔	4.4	31.98	4.5
4	Japan	204	4.37%	5425	3.89%	➔	3582	3.76%	➔	5.65	40.67	7
5	Italy	174	3.73%	4774	3.42%	➔	3255	3.42%	➔	4.75	39.28	6.4
6	South Korea	162	3.47%	2586	1.85%	●	1712	1.8%	●	4.48	21.38	5.2
7	Iran	153	3.28%	3289	2.35%	●	2028	2.13%	●	3.45	23.67	5
8	Germany	108	2.32%	3081	2.21%	➔	2010	2.11%	➔	5.94	28.49	5.7
9	United Kingdom	101	2.17%	5276	3.78%	● ●	3826	4.02%	● ● ●	13.26	27.21	4.7
10	Spain	76	1.63%	7515	5.38%	● ● ●	5036	5.29%	● ● ●	9.12	34.79	6.7
11	Canada	66	1.41%	1504	1.08%	➔	1029	1.08%	➔	7.73	31.82	5.7
12	France	56	1.2%	1777	1.27%	➔	1306	1.37%	➔	10.64	25.09	6.9
13	Egypt	54	1.16%	2409	1.73%	●	1476	1.55%	●	4.64	26.22	4.2
14	Sweden	51	1.09%	1786	1.28%	➔	1335	1.4%	●	10.95	30.18	6.9
15	Australia	49	1.05%	941	0.67%	●	1021	1.07%	➔	6.72	21.16	4.5
16	Brazil	48	1.03%	1959	1.4%	●	1250	1.31%	●	5.52	37.79	6.9
17	Chinese Taipei	48	1.03%	1036	0.74%	●	728	0.76%	●	8.27	36.46	6.3
18	Turkey	48	1.03%	412	0.29%	● ●	196	0.2%	● ● ●	2.89	15.51	4.9
19	Saudi Arabia	36	0.77%	593	0.42%	●	424	0.44%	●	3.61	29.14	4.1
20	Netherlands	34	0.73%	2010	1.44%	● ● ●	955	1%	●	9.27	30	5.8
21	Singapore	31	0.66%	1410	1.01%	● ●	1027	1.08%	● ●	8.28	40.6	5.9
22	Switzerland	28	0.6%	973	0.7%	➔	625	0.66%	➔	13.66	25.14	6.5
23	Poland	18	0.38%	261	0.19%	● ●	190	0.2%	●	4.85	23.94	6.7
24	Pakistan	17	0.36%	73	0.05%	● ● ●	42	0.04%	● ● ●	3.7	7.35	5.8
25	Israel	16	0.34%	402	0.29%	➔	285	0.3%	➔	8.96	31.5	8.7
26	Greece	15	0.32%	236	0.17%	●	125	0.13%	● ●	7.36	24.8	4.5
27	Malaysia	15	0.32%	195	0.14%	● ●	125	0.13%	● ●	3.08	21.8	5.7
28	Norway	14	0.3%	377	0.27%	➔	251	0.26%	➔	4.7	40.57	7.1
29	Russian Federation	14	0.3%	211	0.15%	● ●	111	0.12%	● ●	4.14	35.21	5.5
30	Ireland	13	0.28%	503	0.36%	●	316	0.33%	➔	7.05	49.31	4.6
Totals/averages		4663		139,564			95,260			6.1	29.54	5.85

All author countries and impact of retracted papers.

Table 22. Retracted research impact for top 30 countries (all authors, all retraction reasons). → = >-25% - <+25%; ● = <25% up to 50%, ●●—<50% - 75%; ●●●—<75%; ●—>25% up to 50%; ●●—>50% - 75%; ●●●—>75%; CS—CiteScore average (3933 records with available information); ET—Exposure Time average.

	Country & Region	N	%A	G	%G	Impact	D	%D	Impact	CS	ET	Authors
1	China	1624	33.52%	24,547	17.43%	●	18,264	19.02%	●	4.54	23.12	6.3
2	United States	1175	24.26%	58,666	41.66%	●●	40,610	42.30%	●●	9.82	37.24	6.4
3	India	384	7.93%	15,163	10.77%	●	9019	9.39%	→	4.41	32.18	4.7
4	Japan	241	4.97%	8117	5.76%	→	5496	5.72%	→	6.72	39.87	7.3
5	Italy	227	4.69%	8547	6.07%	●	5770	6.01%	●	6.37	39.06	7.7
6	United Kingdom	194	4.00%	9575	6.8%	●●	7204	7.5%	●●●	12.26	28.2	6.6
7	South Korea	182	3.76%	3218	2.28%	●	2177	2.27%	●	4.85	22.96	5.5
8	Germany	177	3.65%	6449	4.58%	●	4186	4.36%	→	8.18	29.55	7.7
9	Iran	174	3.59%	3437	2.44%	●	2122	2.21%	●	3.4	22.25	5.1
10	Canada	109	2.25%	2927	2.08%	→	1733	1.80%	→	7.91	29.37	6.4
11	Spain	101	2.08%	8638	6.13%	●●●	5910	6.16%	●●●	9.32	34.53	8
12	France	96	1.98%	3881	2.76	●	2500	2.6%	●	11.08	29.6	8.2
13	Australia	77	1.59%	1367	0.97%	●	1265	1.32%	→	7.03	18.77	5.5
14	Sweden	73	1.50%	3840	2.73%	●●●	2628	2.74%	●●●	11.32	31.58	7.7
15	Netherlands	66	1.36%	3029	2.15%	●●	1598	1.66%	→	8.49	24.33	7.1
16	Egypt	63	1.3%	2492	1.77%	●	1537	1.6%	→	4.45	25.33	4.3
17	Chinese Taipei	59	1.22%	1225	0.87	●	864	0.9%	→	8.01	34.54	6.6
18	Brazil	57	1.18%	2175	1.54%	●	1402	1.46%	→	5.97	36.96	6.9
19	Turkey	55	1.13%	420	0.3%	●●	203	0.21%	●●●	3.11	13.35	5.1
20	Switzerland	50	1.03%	2183	1.55%	●●	1427	1.49%	●	15.19	25.68	8.7
21	Saudi Arabia	46	0.95%	1424	1.01%	→	793	0.83%	→	3.74	31.91	4.6
22	Singapore	39	0.8%	1904	1.35%	●●	1380	1.44%	●●●	8.2	36.62	6
23	Pakistan	28	0.61%	102	0.07%	●●●	71	0.07%	●●●	3.88	6.39	5.5
24	Norway	26	0.54%	757	0.54%	→	516	0.54%	→	6.31	35.12	8.9
25	Greece	25	0.52%	869	0.62%	→	592	0.62%	→	8.49	24.88	5.6
26	Russian Federation	24	0.49%	541	0.38%	→	302	0.31%	●	7.04	41.87	7.4
27	Israel	23	0.47%	690	0.49%	→	510	0.53%	→	11.13	30.83	9.4
28	Ireland	22	0.45%	807	0.57%	●	535	0.56%	→	10.11	38.41	7.5
29	Poland	22	0.45%	335	0.24%	●	247	0.26%	●	4.97	23.64	7.9
30	Malaysia	19	0.39%	268	0.19%	●●	180	0.19%	●●	3.63	21.42	5.8
	Totals/averages	4844		140,810			96,000			6.04	28.89	5.83

Table 23. Retracted research impact for top 30 countries (all authors, without editorial retraction reason). → = >-25% - <+25%; ● = <25% up to 50%; ●● = <50% - 75%; ●●● = <75%; ● = >25% up to 50%; ●● = >50% - 75%; ●●● = > 75%; CS—CiteScore average (3933 records with available information); ET—Exposure Time average.

	Country & Region	N	%A	G	%G	Evolution	D	%D	Evolution	CS	ET	Authors
1	China	1594	34.18%	24,300	17.41%	●	18,102	19%	●	4.55	23.36	6,3
2	United States	1130	24.23%	58,332	41.79%	●●	40,394	42.44%	●●●	10.04	38.25	6,5
3	India	375	8.04%	15,111	10.83%	●	8995	9.44%	→	4.43	32.41	4,8
4	Japan	233	4.99%	8065	5.78%	→	5433	5.7%	→	6.74	41.15	7,2
5	Italy	218	4.67%	8500	6.09%	●	5740	6.02%	●	6.29	40.17	7,8
6	United Kingdom	184	3.95%	9329	6.68%	●●	7112	7.47%	●●●	12.35	28.43	6,8
7	South Korea	178	3.81%	3198	2.29%	●	2171	2.28%	●	4.89	22.78	5,5
8	Germany	165	3.6%	6538	4.68%	●	4139	4.34%	→	8.42	30.93	7,8
9	Iran	166	3.54%	3380	2.42%	●	2078	2.18%	●	3.5	22.79	5,1
10	Spain	99	2.12%	8638	6.19%	●●●	5906	6.2%	●●●	9.48	35.2	8,2
11	Canada	97	2.08%	2861	2.05%	→	1712	1.8%	→	8.37	32.37	6,6
12	France	93	1.99%	3868	2.77%	●	2490	2.61%	●	11.3	30.49	8,3
13	Australia	71	1.52%	1290	0.92%	●	1236	1.3%	→	7.37	20.04	5,6
14	Sweden	70	1.5%	3831	2.74%	●●●	2617	2.75%	●●●	11.8	31.97	8
15	Egypt	62	1.33%	2491	1.78%	●	1537	1.61%	→	4.45	25.68	4,3
16	Netherlands	60	1.29%	2996	2.14%	●●	1586	1.66%	●	8.87	26.3	7,3
17	Chinese Taipei	58	1.24%	1224	0.88%	●	864	0.91	●	8.12	35.1	6,6
18	Brazil	52	1.12%	2136	1.53%	●	1381	1.45%	●	6.3	38.98	7,1
19	Turkey	51	1.09%	414	0.3%	●●	198	0.21%	●●●	3.17	13.76	5,1
20	Switzerland	49	1.05%	2183	1.56%	●	1425	1.5%	●	15.42	26.16	8,8
21	Saudi Arabia	46	0.99%	1424	1.02%	→	793	0.83%	→	3.74	31.91	4,6
22	Singapore	39	0.84%	1904	1.36%	●●	1380	1.45%	●●	8.19	36.62	6
23	Pakistan	27	0.58%	102	0.07%	●●●	64	0.07%	●●●	3.88	6.56	5,6
24	Greece	25	0.53%	869	0.62%	→	592	0.62%	→	8.49	24.88	5,6
25	Norway	24	0.51%	752	0.54%	→	511	0.54%	→	6.22	37.83	8,8
26	Russian Federation	24	0.51%	541	0.39%	→	302	0.32%	●	7.04	41.87	7,4
27	Israel	23	0.49%	690	0.49%	→	510	0.54%	→	11.13	30.83	9,4
28	Ireland	22	0.47%	807	0.61%	→	535	0.59%	→	10.11	36.87	7,5
29	Poland	21	0.45%	335	0.24%	●	245	0.26%	●	4.88	24.67	7,9
30	Malaysia	19	0.41%	268	0.19%	●●	180	0.19%	●●	3.63	21.42	5,8
	Totals/averages	4663		139,564			95,260			6.1	29.54	5.85

5. Discussions

5.1. Trend

The number of articles withdrawn from the biomedical literature is increasing, a fact that was constantly signaled in the articles that study this subject (Bordino et al., 2020; Budd et al., 1998; Cokol et al., 2007; Redman et al., 2008; Samp et al., 2012; Singh et al., 2014; Wager & Williams, 2011). Therefore, we analyzed the evolution over time of retractions.

We have not identified elements to signal a slowdown. On the contrary, 2020 seems to be a record year for retraction notes, 878 (18.6% of the total) being already registered in PubMed on January 31st, 2021. Almost half of the 2020 retractions (423/878) are issued for articles with the first author affiliated to a Chinese institution and 135 for authors affiliated to US institutions. Considering the period 2009-2020, five of the top 10 countries recorded the highest number of withdrawals in 2020: China, United States, India, Japan, and the United Kingdom.

The process of correcting the biomedical literature seems to be now continuous and consistent, going back ten years or more, 11% of the retraction notes appearing in 2020 and 15.8% in 2019, being for papers published in 2009-2012 (see [Table 9](#)).

5.2. Countries and International Cooperation

More than 50% of the total number of retracted articles come from China (1588/32.76%) and the United States (918/18.95%), followed by India (351/7.24%), Japan (212/4.37%) and Italy (182/3.75%). We can presume that these figures probably reflect a mix of scientific production volume and the control policies implemented at institutional and national levels. Recent results seem to support this hypothesis (Bhatt, 2021).

The top 30 countries account for 94.7% of all withdrawn articles.

The data does not radically change when considering the country of origin of any author of the article. More consistent growth when compared to first author country numbers, reflecting probably a systemic penchant for international cooperation, is recorded in the United States, United Kingdom, Germany, and Canada ([Table 23](#)). Retracted articles involving international cooperation represent a percentage of 15.3% (741 articles), articles written by a single author account for 5.26% (255 papers).

5.3. Retraction Reasons

The withdrawal of a scientific article is a complex process in which several reasons can justify the retraction decision. In our series, we have found that 12.6% of the articles had multiple retraction reasons.

Mistakes and data inconsistencies represent 32.06% (1st place), the most affected countries being United States, United Kingdom, and Japan.

There is a substantial increase of scientific articles retracted because of im-

age-related issues at a rate lower (22.46%) than the one estimated on a PubMed sample by Bucci (Bucci, 2018), but still significant. Since the paper published by Rossner in 2004 (Rossner & Yamada, 2004), much progress has been made (Irwin et al., 2012), but we can hypothesize that only the recent years' advances in image processing and analysis have generated the adoption by publishers and journals of technologies able to discover image related QRP/QPP which were not easy to identify on a large scale a couple of years ago. The dynamic of retraction notes seem to support such a hypothesis (Table 24). Countries with high image retractions percentages are the United States, Italy, and India.

Table 24. Evolution of retraction notes for images.

	Retraction note year												Total	
	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020		2021
Image retractions	1	8	16	20	37	42	57	118	120	140	198	315	16	1088

Plagiarism and overlap continue to represent a problem with more than 25% of the retraction causes (plagiarism and overlap excluding images) and more than 30% when images overlap and plagiarism is added. India, Italy, and Iran have disproportionately large percentages (around 50%) of their retractions in this category.

The 5th place is the fraud, predominantly represented by fraudulent peer review (350 out of 393). Iran (32.7%), China (18.2%) and South Korea (16.9%) have important percentages of their retractions in this category. As the top years were 2015 (118 retraction notes), 2016 (75 retraction notes), and 2017 (134 retraction notes), we have reasons to think that publishers and journals have fixed their vulnerabilities or the exploits in the publishing systems are not yet discovered.

Ethics also represent an important reason for retraction, with 7.43% of the retracted articles. Interestingly enough, countries we did not expect have in this category substantial percentages of the volume of their withdrawn articles: Germany (20.9%), Japan (15.6%), South Korea (12%). Also, only one country in the top 10 is below 5% (Iran, 4.9%), and the number of retractions for ethical reasons is continuously increasing (Table 25). This fact leads us to consider the possibility of greater attention from publishers and journals to research ethics and/or publication ethics.

Table 25. Evolution of retraction notes for ethical reasons.

	Retraction note year												Total	
	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020		2021
Ethics	1	5	21	18	20	27	23	25	33	41	56	88	2	360

Authorship is the retraction reason for 5.78% of the papers; editorial mistakes account for 3.73%, property & legal concerns 2.5%, and other reasons 1.3%. We must mention here that a relatively high number of papers (241, 5.1%) have no clear reason mentioned as the cause of retraction. Details about distribution by the top 10 countries are in [Table 17](#).

The data do not change very much when the countries of all authors are taken into account ([Table 18](#)).

5.4. Number of Authors

Several published reports on the number of authors in the scientific literature ([Gu et al., 2017](#); [Jang et al., 2016](#); [Larivière et al., 2015](#); [Sacco & Milana, 1984](#)) mention a continuous increase of average number. With an average of 5.83, a median of 5, and a mode of 4 authors per retracted article published between 2009 and 2020, we did not find an upward trend for the average number of authors. Single author papers account for 5.3% of the total, confirming the decreasing trend reported previously ([Larivière et al., 2015](#)).

5.5. Citations, Journal Impact, and Exposure Time

Citations and citations based indicators are still considered to reflect the impact and relevance of scientific work.

During the period in which the journal article is available both in search and on the journal's website without any mention of withdrawal, it is considered valid by many researchers and can influence decisions related to research projects or, in the least pessimistic scenario, can be cited in the bibliography of an article.

Not all articles are cited, but there are enough situations in which the critical analysis of the content can be diminished in intensity by the journal's prestige, author's affiliations, or other factors. Therefore, to better highlight this multi-dimensional model, we extracted citations of withdrawn articles and the journal's impact indicator (CiteScore) from the year of publication of the article (information available for the period 2011-2020). Also, the exposure time in months was calculated as described in the material and methods section for each article.

5.6. Citations

For all 4844 articles, the total number of citations in Google Scholar is 140,810 (mean = 29.07/median = 11), and the total number of citations in Dimensions database is 96,000 (19.82/7).

We found a 68% coverage of Dimensions database when reported to Google Scholar, which was slightly higher than a recent report ([Martín-Martín et al., 2020](#)).

Most cited articles were the ones retracted for image reasons (average 44.6 on Google Scholar/32.9 on Dimensions) followed by mistakes (35.5/24.6) and plagiarism (27.8/16.1). Less cited were those retracted for editorial reasons (6.9/4.1), property and legal concerns (10.4/6.9) and authorship (11.4/7.2) (see [Table 14](#)

for details).

5.7. Journal Impact (Cite Score)

There is old history and many controversies behind the indicators that were initially used to ease the purchasing decisions of academic libraries (Brodman, 1944; Gross & Gross, 1927) or to assess the quality of scientific literature (Garfield, 1955) and nowadays to give a measure of prestige for scientific journals (Garfield, 1999). Currently, there are two leading indicators for the impact of a journal: JIF (Journal Impact Factor, dominating the academic market since 1960) and CiteScore (launched in 2016) (Teixeira da Silva & Memon, 2017). We extracted the journal CiteScore (CS) score (publicly available) for all retracted articles. CS information was available for 3933 articles published between 2011 and 2020.

Average CiteScore for 3933 articles is 6.03 (5.81 - 6.26).

Articles with retractions reasons, Other', Mistakes/inconsistent data' and Images' had the largest CS average value. In contrast, articles retracted for plagiarism, fraud, overlap, or authorship were published in lower CS journals (see **Table 14** for details).

5.8. Exposure Time

Exposure time was 28.89 months in average (median = 19). Previous studies reported article lifespans between 26 months and 44 months (Decullier et al., 2013; Fang et al., 2012; Foo, 2011; Nath et al., 2006; Pantziarka & Meheus, 2019; Rosenkrantz, 2016; Samp et al., 2012) excepting the paper by Singh (Singh et al., 2014) which reports an 18 months lifespan for papers published between 2004 and 2013 (12 months for 1695 papers published between 2009 and 2013). Our findings show a moderate decrease in the article lifespan for 2009-2020 compared to the previously reported data. For 2125 the papers published between 2009 and 2013, the lifespan is 41 months, a difference (compared to Singh) which late retractions for images/other reasons could explain.

Articles retracted due to image issues have a much longer average exposure time (49.21 months) than items withdrawn for any other reason (**Table 14**). This long exposure time, together with an average CS significantly higher than the average of the whole group and an average number of authors also higher than the average (the highest of all reasons for withdrawal), could explain why this group has the highest number of citations/article (44.6 in Google Scholar, 32.9 in Dimensions).

Articles withdrawn due to errors or inconsistency of data have a CS of 8.06 (highest) and an average number of authors of 6.4 (second in size) but an average exposure time of only 27 months. In this case, Google Scholar citations are 35.9/article and 24.9/article in Dimensions.

The other causes are significantly below average exposure time (less ,Ethics' reason, with 31.7 months). All have fewer authors on average, and excepting the

reason ,Others', have a significantly lower CS.

5.9. Impact of Retracted Research

We tried, in this article, to formulate a representation of the impact(citations received) that retracted research (all retraction reasons, without editorial reasons, country of the first author, countries of all authors) for the top 30 countries (Tables 20-23).

5.9.1. First Author Only, without Editorial Reasons (Table 21)

Our findings suggest that retracted scientific papers from 9 out of the first 30 countries have a higher impact (United States, India, United Kingdom, Spain, Egypt, Brazil, Netherlands, Singapore, Ireland) while retracted research from other countries has a lower than expected impact (Pakistan, China, Turkey, Malaysia, Russian Federation, South Korea, Iran, Australia, Saudi Arabia, Poland, and Greece).

5.9.2. All Authors, without Editorial Reasons (Table 23)

High impact for 13 countries: Spain, Sweden, United States, United Kingdom, Netherlands, Singapore, Switzerland, India, Italy, Germany, France, Egypt, Brazil;

Low impact (countries & regions): Pakistan, Turkey, Malaysia, China, South Korea, Iran, Australia, Chinese Taipei, Poland.

Some outliers can bring some modifications. For example, Spain has an article with more than 4000 citations in Google Scholar. Australia has a highly cited article in Dimensions, while the same article has a small number of citations in Google Scholar. However, these do not modify the direction of the impact (Spain has a single red point instead of three, Australia loses one green point and passes to stable in only one of the four classifications).

Various factors may explain these differences in the impact of retracted research when considering the originating country: scientific tradition and strategies, funding of research, faster or easier access to better journals, scientific networking, better international cooperation, journal or institution prestige. When one or several of these factors concur in the direction of a negative impact of questionable research or publication practices, questions could be raised about the safety mechanisms of the scientific process.

The ecosystem of withdrawn articles is a complex one. The matrix in which they are framed reflects cultural differences, differences in development, and organization of the scientific research system in different countries that contribute to the shared heritage and advancement of biomedical knowledge. Our study is only a snapshot of a short period and its exploratory nature, and inherent weaknesses (lack of an unambiguous wording in some withdrawal notes, especially for ethical reasons; lack of/unclear content in over 5% of the withdrawal notes; online absence of some articles) determines us to proceed with caution in formulating conclusions.

6. Conclusion

Mistakes and inconsistent data (including data fabrication) are the main retraction reasons for articles published from 2009-2020 and indexed in PubMed.

Images (22.46%) and ethical issues (7.43%) are retraction reasons growing in recent years and count for almost 30% of all retractions.

Plagiarism and overlap still represent a significant problem (>30% of the total when images are included), especially in lower impact journals.

The number of fraudulent peer review cases (350) shows the need to strengthen these processes, making them less vulnerable to circumvention attempts. As with plagiarism and overlap, the journals affected are those with a low CiteScore.

Major publishers and biomedical journals are involved in the retrospective verification (going back at least 12 years), resulting in a steady increase in withdrawals. The process does not appear to be slowing down. Unfortunately, 5.1% inconsistent/unclear reporting of retractions together with 3.7% editorial errors signals severe deficiencies for some publishers and journals.

Withdrawal of articles seems to be a technology-dependent process (image analysis and anti-plagiarism software). Increasing the accessibility of these technologies can help combat QRP and QPP more effectively.

The number of citations of retracted articles shows a high impact of papers published by authors from certain countries, most of them scientifically advanced. This impact suggests a need for improving the verification processes at the national and institutional levels, publishers, and biomedical journals. The number of retracted articles per country does not always accurately reflect the scientific impact of QRP and QPP papers. Authors from countries whose articles have a negative scientific impact have published in journals with a high CiteScore: Switzerland (15.42), Spain (12.35), or United States (10.04) are some examples. On the opposite side, countries figuring in the top 10 but with low scientific impact published in much lower CiteScore journals (Iran (3.5), India (4.43), or China (4.45)). One may also ask whether authors' access to high-impact journals is easier for some countries and more difficult for others.

The country distribution of retraction reasons shows structural problems in the organization and quality control of scientific research, which have different images depending on geographical location, economic development, and cultural model.

Further research of retracted articles could focus on the equitable distribution of the responsibility for the authors of withdrawn articles or the early detection of policy changes that lead to better research. Other directions could be the evaluation of citation dynamics and the prevention of post-retraction citations or the implementation of more effective policies for reporting and disseminating retraction notes at the publisher or journal level. Last but not least, a stable classification would allow the differentiation between QRP and QPP articles and, therefore, the rapid rejection of invalid research and the appropriate correction of defective scientific reporting practices without the annulment of scientifically

valid results.

Conflicts of Interest

The authors declare no conflicts of interest regarding the publication of this paper.

References

- Azoulay, P., Furman, J. L., Krieger, & Murray, F. (2015). Retractions. *Review of Economics and Statistics*, *97*, 1118-1136. https://doi.org/10.1162/REST_a_00469
- Baerlocher, M. O., O'Brien, J., Newton, M., Gautam, T., & Noble, J. (2010). Data Integrity, Reliability and Fraud in Medical Research. *European Journal of Internal Medicine*, *21*, 40-45. <https://doi.org/10.1016/j.ejim.2009.11.002>
- Benson, P. J. (2016). Seven Sins in Publishing (but Who's Counting...). *Annals of the Royal College of Surgeons of England*, *98*, 1-5. <https://doi.org/10.1308/rcsann.2016.0046>
- Bhatt, B. (2021). A Multi-Perspective Analysis of Retractions in Life Sciences. *Scientometrics*, *126*, 4039-4054. <https://doi.org/10.1007/s11192-021-03907-0>
- Bordino, M., Ravizzotti, E., & Vercelli, S. (2020). Retracted Articles in Rehabilitation: Just the Tip of the Iceberg? A Bibliometric Analysis. *Archives of Physiotherapy*, *10*, Article No. 21. <https://doi.org/10.1186/s40945-020-00092-w>
- Brodman, E. (1944). Choosing Physiology Journals. *Bulletin of the Medical Library Association*, *32*, 479-483.
- Bucci, E. M. (2018). Automatic Detection of Image Manipulations in the Biomedical Literature. *Cell Death & Disease*, *9*, Article No. 400. <https://doi.org/10.1038/s41419-018-0430-3>
- Budd, J. M., Sievert, M., & Schultz, T. R. (1998). Phenomena of Retraction: Reasons for Retraction and Citations to the Publications. *JAMA*, *280*, 296-297. <https://doi.org/10.1001/jama.280.3.296>
- Buehring, G. C., Eby, E. A., & Eby, M. J. (2004). Cell Line Cross-Contamination: How Aware Are Mammalian Cell Culturists of the Problem and How to Monitor It? *In Vitro Cellular & Developmental Biology-Animal*, *40*, 211-215. [https://doi.org/10.1290/1543-706X\(2004\)40%3C211:CLCHAA%3E2.0.CO;2](https://doi.org/10.1290/1543-706X(2004)40%3C211:CLCHAA%3E2.0.CO;2)
- Cokol, M., Iossifov, I., Rodriguez-Esteban, R., & Rzhetsky, A. (2007). How Many Scientific Papers Should Be Retracted? *EMBO Reports*, *8*, 422-423. <https://doi.org/10.1038/sj.embor.7400970>
- Cokol, M., Ozbay, F., & Rodriguez-Esteban, R. (2008). Retraction Rates Are on the Rise. *EMBO Reports*, *9*, 2. <https://doi.org/10.1038/sj.embor.7401143>
- Culliton, B. J. (1974). The Sloan-Kettering Affair: A Story without a Hero. *Science*, *184*, 644-650. <https://doi.org/10.1126/science.184.4137.644>
- Decullier, E., Huot, L., Samson, G., & Maisonneuve, H. (2013). Visibility of retractions: A Cross-Sectional One-Year Study. *BMC Research Notes*, *6*, Article No. 238. <https://doi.org/10.1186/1756-0500-6-238>
- Deichmann, U., & Müller-Hill, B. (1998). The Fraud of Abderhalden's Enzymes. *Nature*, *393*, 109-111. <https://doi.org/10.1038/30090>
- DeMets, D. L., Fleming, T. R., Geller, G., & Ransohoff, D. F. (2017). Institutional Responsibility and the Flawed Genomic Biomarkers at Duke University: A Missed

- Opportunity for Transparency and Accountability. *Science and Engineering Ethics*, 23, 1199-1205. <https://doi.org/10.1007/s11948-016-9844-4>
- Fanelli, D. (2009). How Many Scientists Fabricate and Falsify Research? A Systematic Review and Meta-Analysis of Survey Data. *PLoS ONE*, 4, Article ID: e5738. <https://doi.org/10.1371/journal.pone.0005738>
- 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 of the United States of America*, 109, 17028-17033. <https://doi.org/10.1073/pnas.1212247109>
- Foo, J. Y. A. (2011). A Retrospective Analysis of the Trend of Retracted Publications in the Field of Biomedical and Life Sciences. *Science and Engineering Ethics*, 17, 459-468. <https://doi.org/10.1007/s11948-010-9212-8>
- Garfield, E. (1955). Citation Indexes for Science; A New Dimension in Documentation through Association of Ideas. *Science*, 122, 108-111. <https://doi.org/10.1126/science.122.3159.108>
- Garfield, E. (1999). Journal Impact Factor: A Brief Review. *CMAJ: Canadian Medical Association Journal = Journal De L'association Medicale Canadienne*, 161, 979-980.
- Godecharle, S., Fieuws, S., Nemery, B., & Dierickx, K. (2018). Scientists Still Behaving Badly? A Survey Within Industry and Universities. *Science and Engineering Ethics*, 24, 1697-1717. <https://doi.org/10.1007/s11948-017-9957-4>
- Goldstein, L., & Eastwood, J. M. (1966). On the Primary Site of Nuclear RNA Synthesis. A Retraction. *The Journal of Cell Biology*, 31, 195.
- Gross, P. L., & Gross, E. M. (1927). College Libraries and Chemical Education. *Science*, 66, 385-389. <https://doi.org/10.1126/science.66.1713.385>
- Gu, A., Almeida, N., Cohen, J. S., Peck, K. M., & Merrell, G. A. (2017). Progression of Authorship of Scientific Articles in *the Journal of Hand Surgery*, 1985-2015. *The Journal of Hand Surgery*, 42, 291.e1-291.e6. <https://doi.org/10.1016/j.jhsa.2017.01.005>
- Harriman, S., & Patel, J. (2014). Text Recycling: Acceptable or Misconduct? *BMC Medicine*, 12, Article No. 148. <https://doi.org/10.1186/s12916-014-0148-8>
- Horbach, S., & Halffman, W. (2019). The Extent and Causes of Academic Text Recycling or “Self-Plagiarism”. *Research Policy*, 48, 492-502. <https://doi.org/10.1016/j.respol.2017.09.004>
- Irwin, R. S., Augustyn, N., French, C. T., Rice, J., & Welch, S. J. (2012). Spread the Word about the Journal in 2012: From Impact Factor to Plagiarism and Image Falsification Detection Software. *Chest*, 141, 1-4. <https://doi.org/10.1378/chest.11-2949>
- Jang, H., Kim, K., Huh, S., & Kim, H. (2016). Increasing Number of Authors Per Paper in Korean Science and Technology Papers. *Science Editing*, 3, 80-89. <https://doi.org/10.6087/kcse.70>
- Kakuk, P. (2009). The Legacy of the Hwang Case: Research Misconduct in Biosciences. *Science and Engineering Ethics*, 15, Article No. 545. <https://doi.org/10.1007/s11948-009-9121-x>
- Larivière, V., Gingras, Y., Sugimoto, C. R., & Tsou, A. (2015). Team Size Matters: Collaboration and Scientific Impact since 1900. *Journal of the Association for Information Science and Technology*, 66, 1323-1332. <https://doi.org/10.1002/asi.23266>
- Maddox, J., Randi, J., & Stewart, W. W. (1988). “High-Dilution” Experiments a Delusion. *Nature*, 334, 287-290. <https://doi.org/10.1038/334287a0>
- Madlock-Brown, C. R., & Eichmann, D. (2015). The (Lack of) Impact of Retraction on Citation Networks. *Science and Engineering Ethics*, 21, 127-137.

- <https://doi.org/10.1007/s11948-014-9532-1>
- Martín-Martín, A., Thelwall, M., Orduna-Malea, E., & Delgado López-Cózar, E. (2020). Google Scholar, Microsoft Academic, Scopus, Dimensions, Web of Science, and OpenCitations' COCI: A Multidisciplinary Comparison of Coverage via Citations. *Scientometrics*, *126*, 871-906. <https://doi.org/10.1007/s11192-020-03690-4>
- McHugh, U. M., & Yentis, S. M. (2019). An Analysis of Retractions of Papers Authored by Scott Reuben, Joachim Boldt and Yoshitaka Fujii. *Anaesthesia*, *74*, 17-21. <https://doi.org/10.1111/anae.14414>
- Michalek, A. M., Hutson, A. D., Wicher, C. P., & Trump, D. L. (2010). The Costs and Underappreciated Consequences of Research Misconduct: A Case Study. *PLoS Medicine*, *7*, Article ID: e1000318. <https://doi.org/10.1371/journal.pmed.1000318>
- 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. <https://doi.org/10.1002/asi.23421>
- Nath, S. B., Marcus, S. C., & Druss, B. G. (2006). Retractions in the Research Literature: Misconduct or Mistakes? *The Medical Journal of Australia*, *185*, 152-154. <https://doi.org/10.5694/j.1326-5377.2006.tb00504.x>
- Pantziarka, P., & Meheus, L. (2019). Journal Retractions in Oncology: A Bibliometric Study. *Future Oncology*, *15*, 3597-3608. <https://doi.org/10.2217/fon-2019-0233>
- Parrish, D. M. (1999). Scientific Misconduct and Correcting the Scientific Literature. *Academic Medicine*, *74*, 221-230. <https://doi.org/10.1097/00001888-199903000-00009>
- Rapani, A., Lombardi, T., Berton, F., Del Lupo, V., Di Lenarda, R., & Stacchi, C. (2020). Retracted Publications and Their Citation in Dental Literature: A Systematic Review. *Clinical and Experimental Dental Research*, *6*, 383-390. <https://doi.org/10.1002/cre2.292>
- Redman, B. K., Yarandi, H. N., & Merz, J. F. (2008). Empirical Developments in Retraction. *Journal of Medical Ethics*, *34*, 807-809. <https://doi.org/10.1136/jme.2007.023069>
- Relman, A. S. (1983). Lessons from the Darsee Affair. *The New England Journal of Medicine*, *308*, 1415-1417. <https://doi.org/10.1056/NEJM198306093082311>
- Roig, M. (2015). *Avoiding Plagiarism, Self-Plagiarism, and Other Questionable Writing Practices: A Guide to Ethical Writing*. <https://ori.hhs.gov/sites/default/files/plagiarism.pdf>
- Rosenkrantz, A. B. (2016). Retracted Publications within Radiology Journals. *American Journal of Roentgenology*, *206*, 231-235. <https://doi.org/10.2214/AJR.15.15163>
- Rossner, M., & Yamada, K. M. (2004). What's in a Picture? The Temptation of Image Manipulation. *The Journal of Cell Biology*, *166*, 11-15. <https://doi.org/10.1083/jcb.200406019>
- Sacco, W. P., & Milana, S. (1984). Increase in Number of Authors Per Article in Ten APA Journals: 1960-1980. *Cognitive Therapy and Research*, *8*, 77-83. <https://doi.org/10.1007/BF01315100>
- Samp, J. C., Schumock, G. T., & Pickard, A. S. (2012). Retracted Publications in the Drug Literature. *Pharmacotherapy*, *32*, 586-595. <https://doi.org/10.1002/j.1875-9114.2012.01100.x>
- Scanlon, P. M. (2007). Song From Myself: An Anatomy of Self-Plagiarism. *Plagiarism*, *2*, 57-66. <https://quod.lib.umich.edu/cgi/p/pod/dod-idx/song-from-myself-an-anatomy-of-self-plagiarism.pdf?c=plag;idno=5240451.0002.007;format=pdf>

- Selvam, T. (2021). Plagiarism and Its Consequences. In C. Thanavathi (Ed.), *Advanced Educational Research and Statistics: Plagiarism and Its Consequences* (1st ed., Vol. 1, pp. 33-38). ESN Publications.
https://www.researchgate.net/profile/Thanavathi-C/publication/340581145_Advanced_Educational_Research_and_Statistics_Edited_Book/links/60562ebd458515e8345cdc82/Advanced-Educational-Research-and-Statistics-Edited-Book.pdf#page=45
- Service, R. F. (2002). Scientific Misconduct. Bell Labs Fires Star Physicist Found Guilty of Forging Data. *Science*, 298, 30-31. <https://doi.org/10.1126/science.298.5591.30>
- Singh, H. P., Mahendra, A., Yadav, B., Singh, H., Arora, N., & Arora, M. (2014). A Comprehensive Analysis of Articles Retracted between 2004 and 2013 from Biomedical Literature—A Call for Reforms. *Journal of Traditional and Complementary Medicine*, 4, 136-139. <https://doi.org/10.4103/2225-4110.136264>
- Steen, R. G. (2011). Retractions in the Scientific Literature: Do Authors Deliberately Commit Research Fraud? *Journal of Medical Ethics*, 37, 113-117.
<https://doi.org/10.1136/jme.2010.038125>
- Stern, A. M., Casadevall, A., Steen, R. G., & Fang, F. C. (2014). Financial Costs and Personal Consequences of Research Misconduct Resulting in Retracted Publications. *eLife*, 3, Article ID: e02956. <https://doi.org/10.7554/eLife.02956>
- Teixeira da Silva, J. A., & Memon, A. R. (2017). CiteScore: A Cite for Sore Eyes, or a Valuable, Transparent Metric? *Scientometrics*, 111, 553-556.
<https://doi.org/10.1007/s11192-017-2250-0>
- Tilden, S. J. (2010). Incarceration, Restitution, and Lifetime Debarment: Legal Consequences of Scientific Misconduct in the Eric Poehlman Case. Commentary on: “Scientific Forensics: How the Office of Research Integrity Can Assist Institutional Investigations of Research Misconduct during Oversight Review”. *Science and Engineering Ethics*, 16, 737-741. <https://doi.org/10.1007/s11948-010-9228-0>
- Titus, S. L., Wells, J. A., & Rhoades, L. J. (2008). Repairing Research Integrity. *Nature*, 453, 980-982. <https://doi.org/10.1038/453980a>
- Wager, E., & Williams, P. (2011). Why and How Do Journals Retract Articles? An Analysis of Medline Retractions 1988-2008. *Journal of Medical Ethics*, 37, 567-570.
<https://doi.org/10.1136/jme.2010.040964>
- Whelden, C. H. (1927). A Retraction. *Journal of the American Statistical Association*, 22, 229. <https://doi.org/10.1080/01621459.1927.10502959>
- Wilfon, B. (1756). CVI. A Retraction, by Mr. Benjamin Wilson, F. R. S. of His Former Opinion, Concerning the Explication of the Leyden Experiment. *Philosophical Transactions of the Royal Society of London*, 49, 682-683.
<https://doi.org/10.1098/rstl.1755.0107>
- Zhang, Q., Abraham, J., & Fu, H.-Z. (2020). Collaboration and Its Influence on Retraction Based on Retracted Publications during 1978-2017. *Scientometrics*, 125, 213-232.
<https://doi.org/10.1007/s11192-020-03636-w>