

Research on Influencing Factors of Retraction Lag for Chinese Scholars' Retracted Papers Based on the Retraction Observation Database

Authors: Chen Junhui, Lihua Zhang, Chen Junhui

Date: 2024-05-06T00:00:00+00:00

Abstract

Purpose Retraction lag reflects the purification timeliness of the scientific community, and analyzing its influencing factors can provide guidance for reducing this lag. **Methods** This study collected 17,916 paper records from the Retraction Watch database (1970-2022), constructed a regression model with retraction lag as the dependent variable and 33 variables selected from policy, paper, and author dimensions as independent variables, and examined the effects of various factors on retraction lag. **Results** Research articles, publication year, image-related issues, paper mills, and meta-analyses constitute the primary factors affecting retraction lag, whereas research integrity policies show no significant impact. **Conclusion** We recommend strengthening integrity awareness regarding image issues, streamlining paper reporting mechanisms, addressing paper mill problems, and attending to integrity issues in research teams of varying sizes.

Full Text

Research on Factors Influencing the Retraction Time Lag of Chinese Scholarly Papers: Based on the Retraction Watch Database

CHEN Junhui¹, **ZHANG Lihua**¹

¹ School of Information, Shanxi University of Finance and Economics, No. 140 Wucheng Road, Xiaodian District, Taiyuan, Shanxi 030006, China

Abstract:

[Purpose] The retraction time lag reflects the purification efficiency of the scientific community. Analyzing the factors that influence this lag can provide valuable insights for reducing it. [Methods] We collected 17,916 retracted papers from the Retraction Watch Database spanning 1970-2022, using retraction

time lag as the dependent variable and selecting 33 variables from policy, paper, and author dimensions as independent variables to construct a regression model analyzing how different factors affect retraction time lag. [Findings] Research papers, publication year, image manipulation, paper mills, and meta-analyses emerged as the primary factors influencing retraction time lag, while research integrity policies showed no significant effect. [Conclusions] We recommend strengthening awareness of image integrity, streamlining paper reporting mechanisms, paying greater attention to paper mill problems, and focusing on integrity issues in research teams of varying sizes.

Keywords: retraction time lag; Retraction Watch Database; retracted papers; regression analysis

In recent years, the frequent occurrence of retractions has drawn significant attention from both the academic community and research management institutions. According to a *Nature* report [1], over ten thousand research papers were retracted globally in 2023, setting a new record. Saudi Arabia, Pakistan, Russia, and China have exhibited the highest retraction rates over the past two decades. On August 25, 2022, China's Ministry of Science and Technology and 21 other departments revised the *Regulations on the Investigation and Handling of Research Misconduct (Trial)*, further standardizing the investigation and handling of research integrity violations.

Retraction events have made retracted paper research a hot topic and frontier theme in the fields of science studies, scientometrics, and science policy. Retracted papers are defined as published articles that, upon author application or reporting by members of the scientific community, are found to contain errors or research misconduct, prompting the publishing journal to issue an official retraction statement to inform the public [2].

In retraction research, Budd J M [3] first studied retraction time lag in 1998. Retraction time lag refers to the interval between a paper's publication date and its retraction date, reflecting to some extent the purification efficiency of the scientific community [4]: the shorter the lag, the higher the purification efficiency. Prolonged retraction time lag not only sustains the influence of erroneous research but also triggers a series of domino effects [5], ultimately damaging the reputation of the entire scientific community. Multiple factors influence retraction time lag, including paper characteristics such as publication year, retraction reason, discipline type, and manuscript type, as well as external factors derived from efforts to enhance academic authority, such as national research integrity policies, increased penalties for misconduct by funding agencies and ministries, and researcher training in research integrity. Therefore, in-depth study of the factors influencing retraction time lag is crucial for understanding its patterns and implementing appropriate preventive and corrective measures.

Current research on retraction time lag remains limited, focusing primarily on average values [6-13] and influencing factors. Studies show that the average retraction time lag for Chinese SCI and SSCI journals is 963.34 days [7]. For

Science magazine, the median retraction time lag is 24 months, with a mean of 34 months [11]. Different factors exert varying influences on retraction time lag. Publication year and collaboration show negative correlations with retraction time lag [14]; more recent publications have shorter lags, while collaborative papers have longer lags than non-collaborative ones. The correlation between manuscript type and retraction time lag remains unclear [14]. Journal H-index [15] and impact factor [16] show positive correlations with retraction time lag; higher H-index and impact factor journals exhibit longer retraction time lags. Scholars disagree on the relationship between author count and retraction time lag: Madhugiri V S et al. [15] found no significant correlation, while Bar-Ilan J [17] argued that retraction time lag increases with author count. Regarding retraction reasons, papers retracted for fabrication have significantly longer time lags than those retracted for other reasons [18-19].

A review of existing literature reveals that current studies offer limited analysis of influencing factors from narrow perspectives and lack comprehensive analysis of multiple factors' combined effects on retraction time lag. This paper first examines the relationship between average retraction time lag and various factors including publication year, research integrity policy count, author count, discipline type, and retraction reasons, then establishes a regression model to comprehensively analyze the degree of influence these factors exert. Our findings can support efforts to reduce retraction time lag, improve academic community purification efficiency, and mitigate the adverse effects of retracted papers.

2. Research Design

2.1.1 Dependent Variable

This study selects retraction time lag as the dependent variable. A shorter retraction time lag indicates a smaller interval between publication and retraction, reflecting higher purification efficiency.

2.1.2 Independent Variables

We analyze the impact on retraction time lag from three dimensions: policy, paper, and author (Figure 1 [Figure 1: see original paper]).

(1) Policy Dimension

Research integrity policies play a crucial role in improving paper quality and curbing academic misconduct. Xu Hongping' s [20] research indicates that the promulgation of research integrity policies enhances academic purification effects, drawing attention to problematic papers in their early publication stages and thereby reducing retraction time lag. In the policy dimension, indicators such as policy intensity and implementation degree are difficult to quantify, so we select research integrity policy count as the explanatory variable. We therefore propose hypothesis h1: research integrity policy count is negatively correlated with retraction time lag.

(2) Paper Dimension

Paper dimension variables include publication year, discipline type, manuscript type, and retraction reason.

With technological advances, increased scholarly attention to retraction systems [21], and greater focus on retraction incidents, retraction time lag has decreased year by year, with 2020 showing a much shorter lag than 2011 [14]. We therefore propose hypothesis h2: publication year is negatively correlated with retraction time lag.

Compared with humanities and social sciences articles, natural science articles involve calculations, figures, and data, resulting in longer time costs for error correction and relatively longer retraction time lags. Different disciplines exert varying influences on retraction time lag, so we discuss each discipline's relationship with retraction time lag separately. The Retraction Watch Database categorizes disciplines into Business/Technology (B/T), Life Sciences (BLS), Environmental Science (ENV), Health Sciences (HSC), Humanities (HUM), Physical Sciences (PHY), and Social Sciences (SOC). This study adopts this classification system, proposing hypotheses h3a: Life Sciences are positively correlated with retraction time lag; h3b: Health Sciences are positively correlated; h3c: Environmental Science is negatively correlated; h3d: Humanities are negatively correlated; h3e: Social Sciences are negatively correlated; and h3f: Physical Sciences are positively correlated.

Due to differences in carrier type, dissemination scope, and accessibility, retraction time lags may vary across manuscript types. The Retraction Watch Database classifies manuscripts into 13 categories: Articles in Press, Conference Abstracts, Trade Magazines, Case Reports, Preprints, Letters, Editorials, Book Chapters, Corrections, Clinical Research, Meta-Analyses, Research Articles, and Reviews. Following Li Junqiang's [14] research, this study selects three representative manuscript types for analysis, proposing hypotheses h4a: Meta-Analyses are positively correlated with retraction time lag; h4b: Clinical Research papers are positively correlated; and h4c: Research Articles are positively correlated.

Previous studies [22-23] demonstrate correlations between retraction reasons and time lag. The Retraction Watch Database contains 96 retraction reason categories. Some scholars have used these classifications directly or established their own classification systems. Referencing Zhang Chunbo [19], Horbach S P J M [23], and Guan Qian [24], we categorize retraction reasons into 13 types: Author Issues, Image Issues, Data Issues, Text Issues, Reference Issues, Copyright Issues, Ethical Issues, Publisher Issues, Third-Party Issues, Results/Conclusions Issues, Peer Review Issues, Paper Mills, and Other Issues, with specific classification criteria shown in Table 1. We analyze several typical retraction reasons from recent years, proposing hypotheses h5a: Paper Mills are positively correlated with retraction time lag; h5b: Ethical Issues are positively correlated; h5c: Image Issues are positively correlated; h5d: Data Issues are positively correlated; and h5e: Author Issues are negatively correlated.

(3) Author Dimension

Increased author conflicts of interest, higher communication costs between journals and authors, and enhanced influence from multi-author collaborations may all extend the time interval before paper retraction. Some scholars have noted [17] that retraction time lag increases with author count, leading to hypothesis h6: author count is positively correlated with retraction time lag.

This study includes 36 explanatory variables and 1 dependent variable, as summarized in Table 2 .

2.2 Data Collection

Except for the research integrity policy count variable, all data were sourced from the Retraction Watch Database. As of October 27, 2023, the database contained 47,886 records. We selected 17,916 records with retraction dates between 1970 and 2022, author affiliations in Mainland China, Hong Kong, Macao, or Taiwan, and retraction nature marked as “retraction” as our analytical sample. We collected basic paper information including country, authors, manuscript type, retraction date, publication date, and retraction reasons.

Research integrity policy data were obtained from seven websites including the Ministry of Science and Technology and the China Research Integrity Network, three databases (Beida Fabao, Beida Fayi, and Law Star), and two monographs (*Compilation of Policies on Science and Technology Supervision, Evaluation, and Research Integrity Construction* and *Compilation of Laws, Regulations, and Documents on Research Integrity Construction*). We screened central government-issued policies, collected and deduplicated them, and excluded policies unrelated to research integrity, ultimately obtaining 93 policy records.

2.3 Data Processing

We calculated retraction time lag in days based on publication and retraction dates. For papers with multiple retraction reasons, discipline types, or manuscript types, we conducted separate statistical analyses. For example, if a paper was retracted for both data issues and image issues, we counted both reason categories with their corresponding retraction time lags. We counted research integrity policies by year. The final dataset contained 53,489 records, each including publication year, author count, manuscript type, research integrity policy count, discipline type, retraction reason, and retraction time lag.

We calculated average retraction time lag for each factor. For manuscript types, we computed average retraction time lags for all 13 categories. Since retraction reasons, manuscript types, and discipline types are categorical variables, we included them as dummy variables in the regression model. For instance, with 13 retraction reason categories, a paper retracted for data issues would be coded as 1 for Data Issues and 0 for the other 12 categories. As regression models with dummy variables require a reference category, we selected Other Issues, Conference Abstracts, and Business/Technology disciplines as references, excluding

them from the model. The final regression model included 33 explanatory variables.

2.4 Analysis Methods

2.4.1 Descriptive Statistics For continuous variables (publication year, research integrity policy count, author count), we plotted line graphs showing their relationship with average retraction time lag. For categorical variables (discipline type, retraction reason, manuscript type), we created bar charts. We conducted ANOVA tests on these categorical variables to examine significant differences in retraction time lag distributions.

2.4.2 Regression Analysis Previous retraction time lag studies have primarily used chi-square analysis, ANOVA, and correlation analysis, which can only describe whether relationships exist but cannot measure influence magnitude. Regression analysis addresses this limitation. Using regression coefficients and standardized coefficients, we measured the magnitude and direction of each factor's influence on retraction time lag.

We employed multiple linear regression to explore the degree of influence:

$$y = \beta_0 + \beta_{1x}1 + \beta_{2x}2 + \dots + \beta_{px}p + \varepsilon$$

where ε is the error term, y is the dependent variable, x represents explanatory variables, and $\beta_1, \beta_2, \dots, \beta_p$ are coefficients.

We tested for multicollinearity using Variance Inflation Factor (VIF), with results shown in Table 3. All VIF values were below 5, indicating no first-order autocorrelation and suitability for multiple linear regression.

3. Results

3.1.1 Analysis of Average Retraction Time Lag by Publication Year

Average retraction time lag by publication year is shown in Figure 2 [Figure 2: see original paper]. Following an initial short-term increase, average retraction time lag shows a stepped decline over time. The longest average lag occurred in 2000 at 4,166.82 days, while the shortest was in 2011 at 72.52 days. Between 1995-2000, average retraction time lag fluctuated considerably but generally increased. This may be attributed to the absence of research integrity policies, insufficient scholarly attention to retraction incidents, and inadequate technology for identifying problematic papers, ultimately causing rapid growth in retraction time lag.

After 2000, average retraction time lag declined in a stepped pattern, consistent with Li Junqiang's [14] findings. The period 2000-2011 saw rapid decline, partially attributable to increased central government research integrity policies, suggesting policy promulgation positively impacts academic purification

efficiency. From 2012-2022, average retraction time lag first increased then decreased. Zhang Chunbo [19] noted that the 2013 surge resulted from concentrated retractions of “repeat offenders”—authors with multiple retractions. Our data confirm this: in 2013, Hua Zhong had 45 papers retracted and Tao Liu had 72 papers retracted.

3.1.2 Analysis of Average Retraction Time Lag by Research Integrity Policy Count

Research integrity policy count versus retraction time lag is shown in Figure 3 [Figure 3: see original paper]. When policy count ranged from 1-4, average retraction time lag decreased, indicating policy promulgation negatively impacts retraction time lag and improves purification efficiency. When policy count exceeded 4, average retraction time lag slowly increased, possibly because academic misconduct has become more complex and difficult to verify, limiting policy effectiveness. Current research on this relationship remains limited, requiring further exploration of causality and mechanisms.

3.1.3 Analysis of Average Retraction Time Lag by Author Count

Average retraction time lag by author count is shown in Figure 4 [Figure 4: see original paper]. The longest average lag (1,371.98 days) occurred with 16 authors, while the shortest (136.29 days) occurred with 21 authors. Retraction time lag follows an inverted U-shaped distribution as author count increases. For 1-15 authors, average retraction time lag increases, confirming Sharma K's [25] finding that smaller teams are more prone to retractions because they typically conduct simpler, smaller-scale projects where errors are more easily identified, resulting in shorter retraction time lags. For 16-21 authors, average retraction time lag drops sharply. The increase at 22 authors is based on only one paper and thus appears incidental.

Our findings show that beyond a certain threshold, increased author count reduces average retraction time lag. Considering disciplinary differences, we calculated average retraction time lag by author count within each discipline. Life Sciences, Environmental Science, Health Sciences, and Physical Sciences showed the same inverted U-shaped pattern, while Business/Technology, Social Sciences, and Humanities showed initial decreases followed by increases. This suggests disciplinary context must be considered when examining the author count-retraction time lag relationship.

3.1.4 Analysis of Average Retraction Time Lag by Discipline

Average retraction time lag by discipline is shown in Figure 5 [Figure 5: see original paper]. Life Sciences exhibited the longest average lag at 865.99 days, followed by Health Sciences at 648 days. Environmental Science showed the shortest lag at 149.14 days. ANOVA results revealed significant differences

across disciplines ($F=1,855.27$, $p<0.001$), indicating discipline-specific effects on retraction time lag.

The longer lags in Life Sciences and Health Sciences stem from higher proportions of retractions due to data issues, image manipulation, and ethical problems. In our data, these three reasons accounted for 33% of Life Sciences retractions and 25% of Health Sciences retractions, with average time lags of 925.68 days for data issues, 1,123.82 days for image issues, and 936.83 days for ethical issues. Environmental Science's shorter lag results from 79% of retractions being due to author issues and other problems, which have shorter average lags (258.1 days for other issues, 184.05 days for author issues).

3.1.5 Analysis of Average Retraction Time Lag by Retraction Reason

Average retraction time lag by reason is shown in Figure 6 [Figure 6: see original paper]. Image issues showed the longest average lag at 1,123.82 days. Advances in image processing technology, particularly AI-manipulated images with complex data, make detection difficult, facilitating image-related misconduct. Ethical issues also showed high average lags, with 88% of ethical retractions occurring in Life Sciences and Health Sciences, suggesting insufficient attention to ethics in these fields. Data issues showed long lags due to complex calculations that are difficult for reviewers to verify. Author issues showed the shortest lag, as these typically involve voluntary retractions or author identity problems that are relatively easy and quick to verify. ANOVA results showed significant differences across reasons ($F=1,205.39$, $p<0.001$), indicating varying impacts on retraction time lag.

3.1.6 Analysis of Average Retraction Time Lag by Manuscript Type

Average retraction time lag by manuscript type is shown in Figure 7 [Figure 7: see original paper]. Research articles showed the longest average lag at 821.17 days, followed by meta-analyses and reviews (both >700 days). Conference abstracts showed the shortest lag at 106.42 days, consistent with Ghorbi A's [10] findings. Statistical analysis revealed significant differences across types ($F=1,814.25$, $p<0.001$), indicating manuscript type-specific effects.

Research articles and meta-analyses have longer lags due to their high originality and complexity, which makes replication difficult. Reviews have longer lags because their authors are often leading scholars whose authority makes them less susceptible to challenge. Conference abstracts have the shortest lags because they represent cutting-edge research with high innovation and broad attention, making misconduct easier to detect and retraction faster.

3.2 Regression Analysis Results

Regression analysis results for retraction time lag factors are shown in Table 4. The model's F-test p-value was <0.001 , indicating overall statistical significance. Hypothesis test results are summarized in Table 5.

In the policy dimension, research integrity policy count showed no significant effect in the regression model, possibly because its influence was masked by other variables. Our policy data included only centrally issued policies, which were relatively few in number, limiting our ability to accurately observe their impact on retraction time lag.

In the paper dimension, Humanities and Social Sciences had small sample sizes with high variability, preventing comprehensive assessment of their effects. Health Sciences and Physical Sciences showed negative correlations with retraction time lag, contradicting our hypotheses and suggesting potential unexamined factors requiring further investigation. Publication year, Life Sciences, meta-analyses, clinical research, paper mills, ethical issues, image issues, and data issues all showed positive correlations with retraction time lag, indicating these factors increase lag. Environmental Science and author issues showed negative correlations, reducing lag.

In the author dimension, author count positively influenced retraction time lag, consistent with Bar-Ilan J's [17] findings. This differs from our earlier descriptive analysis showing an inverted U-shaped relationship because that analysis examined author count versus *average* retraction time lag, while regression analysis examines the relationship between author count and retraction time lag itself.

Standardized coefficients reveal that research articles had the largest effect (0.521), indicating the greatest impact on retraction time lag. This suggests editors and reviewers should exercise stricter scrutiny of research articles, including verification of data sources and analysis conclusions. Publication year, image issues, paper mills, and meta-analyses had standardized coefficients between 0.1 and 0.3, indicating substantial influence after research articles. These factors warrant focused attention during journal review to improve paper quality and reduce retraction risk. Data issues (standardized coefficient = 0.098) also merit appropriate attention.

4. Recommendations

4.1 Strengthen Image Integrity Awareness and Introduce Advanced Image Detection Technology

Our findings show that image-related retractions have long time lags, concentrated in Life Sciences and Health Sciences. To address this, specialized training on proper image use should be provided to researchers in these fields, along with strict penalties for image misconduct. Currently, journal plagiarism checks focus on text rather than images. We recommend introducing advanced image content detection technology to screen new submissions for image duplication alongside text checking, which would help reduce image-related retractions and decrease retraction time lag.

4.2 Streamline Paper Reporting Mechanisms and Establish Databases for Questionable and Retracted Papers

Our findings show retraction time lag is influenced by multiple factors, and subtle misconduct may be difficult for editors and reviewers to detect alone. Public participation in reporting problematic papers should be encouraged. Journals should expand reporting channels, improve processing procedures for questionable papers, and establish dedicated research groups for further identification and verification. Upon receiving reports, immediate response mechanisms should be activated. Confirmed problematic papers should be retracted with statements issued; for unfounded reports by repeat offenders, a “reporter blacklist” could be established. Drawing from PubPeer and Retraction Watch, Chinese-language journals should create dedicated databases for questionable and retracted papers, including retracted papers, reported but unresolved cases, and unretracted problematic papers.

4.3 Address Paper Mill Problems and Improve Paper Mill Detection Mechanisms

Our findings reveal that paper mills significantly increase retraction time lag, requiring urgent attention. Paper mills pose a growing threat, manipulating scientific publishing to disseminate false results and undermine research credibility. Candal-Pedreira C et al. [26] noted that in biomedicine, nearly all paper mill-related retractions involve Chinese authors. Journals should strengthen paper mill identification through: (1) Detection tools like Papermill Alarm (<https://rapidapi.com/clear-skies-clear-skies-default/api/papermill-alarm/>), though its effectiveness is limited outside PubMed-covered fields; (2) COPE’s paper mill research report [27] identifying indicators such as similar layouts and designs across unrelated papers; and (3) Multi-journal joint reporting systems for integrity issues, as paper mill problems are too subtle for software alone and require editor/reviewer vigilance and collaboration.

4.4 Focus on Integrity Issues in Research Teams of Different Sizes

Our analysis reveals an inverted U-shaped relationship between author count and average retraction time lag, peaking at 16 authors. For teams smaller than 16 members, retraction time lag increases with author count, making misconduct harder to detect as team size grows. For larger teams (>16 authors), misconduct becomes easier to identify. This suggests journals should focus on integrity issues in medium-sized teams (around 16 authors), while also monitoring potential misconduct in small teams as they grow.

5. Conclusion

This study analyzed relationships between publication year, author count, research integrity policy count, retraction reasons, discipline types, manuscript types, and their corresponding average retraction time lags. Using regression

analysis, we examined the degree of influence each factor exerts on retraction time lag. We found retraction time lag is influenced by multiple factors, with publication year, research articles, data issues, image issues, paper mills, and meta-analyses having substantial effects. Although research integrity policies showed no significant effect in our regression model, they do influence retraction time lag in practice.

Limitations include: (1) Lack of rigorous causal inference models (e.g., difference-in-differences, propensity score matching); (2) Focus on English-language literature only, excluding Chinese-language papers due to data source limitations; and (3) Superficial policy analysis examining only policy count rather than policy intensity or synergistic effects. Future research will expand data sources, employ causal inference methods, and analyze policy content characteristics.

References

- [1] Noorden R V. More than 10,000 research papers were retracted in 2023 – a new record. [EB/OL].[2023-12-23].<https://www.nature.com/articles/d41586-023-03974-8>.
- [2] Fan S, Zhang Z. Distribution patterns of retracted papers in scientific literature databases: A case study of Web of Science [J]. *Journal of the China Society for Scientific and Technical Information*, 2014, 33(4): 375-387.
- [3] Budd J M, Sievert M E, Schultz T R. Phenomena of retraction: reasons for retraction and citations to the publications[J]. *JAMA*, 1998, 280(3): 296-297.
- [4] Zhou Z. Characteristic analysis and implications of retracted papers in Chinese scientific journals [J]. *Chinese Journal of Scientific and Technical Periodicals*, 2017, 28(11): 1065-1070.
- [5] Leta J, Araujo K, Treiber S. Citing documents of Wakefield' s retracted article: the domino effect of authors and journals[J]. *Scientometrics*, 2022, 127(12): 7333-7349.
- [6] Zhang Q, Yao C, Pan Y, et al. Research on retracted papers in Chinese academic journals [J]. *Chinese Journal of Scientific and Technical Periodicals*, 2014, 25(5): 611-615.
- [7] Ma L, Feng L, Yuan J, et al. Analysis and reflection on retraction of Chinese SCI/SSCI journal papers [J]. *Chinese Journal of Scientific and Technical Periodicals*, 2023, 34(5): 584-592.
- [8] Fu Z. Analysis of retraction reasons and academic impact of Chinese papers in PubMed database [J]. *Chinese Journal of Scientific and Technical Periodicals*, 2017, 28(4): 293-299.
- [9] Qin H. Research on retracted papers based on SCI [D]. Shanghai: East China Normal University, 2013.
- [10] Ghorbi A, Fazeli-Varzaneh M, Ghaderi-Azad E, et al. Retracted papers by Iranian authors: causes, journals, time lags, affiliations, collaborations[J]. *Scientometrics*, 2021, 126(9):
- [11] Wray K B, Andersen L E. Retractions in science[J]. *Scientometrics*, 2018,

117(3):

- [12] Panahi S, Soleimanpour S. The landscape of the characteristics, citations, scientific, technological, and altmetrics impacts of retracted papers in hematology[J]. *Accountability in Research*, 2023, 30(7): 363-378.
- [13] Feng L, Yuan J, Yang L. An observation framework for retracted publications in multiple dimensions[J]. *Scientometrics*, 2020, 125(2): 1445-1457.
- [14] Li J. Analysis of retraction time lag and influencing factors of oncology papers worldwide from 2011-2020 [J]. *Chinese Journal of Scientific and Technical Periodicals*, 2022, 33(5): 561-565.
- [15] Madhugiri V S, Nagella A B, Uppar A M. An analysis of retractions in neurosurgery and allied clinical and basic science specialties[J]. *Acta Neurochirurgica*, 2021, 163: 19-30.
- [16] Sharma P, Sharma B, Reza A, et al. A systematic review of retractions in biomedical research publications: reasons for retractions and their citations in Indian affiliations[J]. *Humanities and Social Sciences Communications*, 2023, 10(1): 1-12.
- [17] Bar-Ilan J, Halevi G. Temporal characteristics of retracted articles[J]. *Scientometrics*, 2018, 116(3): 1771-1783.
- [18] Dal-Ré R, Ayuso C. For how long and with what relevance do genetics articles retracted due to research misconduct remain active in the scientific literature[J]. *Accountability in Research*, 2021, 28(5): 280-296.
- [19] Zhang C. Research on reasons, impacts, and countermeasures for retracted papers in international journals [D]. Dalian: Dalian University of Technology, 2021.
- [20] Xu H. Characteristic analysis of retracted papers in Chinese academic journal databases [J]. *Chinese Journal of Scientific and Technical Periodicals*, 2023, 34(4):
- [21] Cui Y. Research on factors influencing the academic purification effect of journal paper retraction systems [D]. Hangzhou: Zhejiang University, 2022.
- [22] Xu Q, Fan Q. Analysis of image-related retractions by Chinese scholars: Based on Retraction Watch database [J]. *Chinese Journal of Scientific and Technical Periodicals*, 2023, 34(10): 1255-1263.
- [23] Horbach S P J M, Halfman W. The ability of different peer review procedures to flag problematic publications[J]. *Scientometrics*, 2019, 118(1): 339-373.
- [24] Guan Q, Dong K, Xia Y. Research on data management issues and countermeasures in life sciences based on retracted papers: Taking life sciences as an example [J]. *Library and Information*, 2021, (3): 47-56.
- [25] Sharma K. Team size and retracted citations reveal the patterns of retractions from 1981 to 2020[J]. *Scientometrics*, 2021, 126(10): 8363-8374.
- [26] Candal-Pedreira C, Ross J S, Ruano-Ravina A, et al. Retracted papers originating from paper mills: cross sectional study[J]. *British Medical Journal*, 2022, 379:e071517.
- [27] COPE; STM. Paper mills research report with recommendations.[EB/OL].[2023-12-23]. <https://publicationethics.org/files/paper-mills-cope-stm-research-report.pdf>.

Note: Figure translations are in progress. See original paper for figures.

Source: ChinaXiv – Machine translation. Verify with original.