

Characteristic Analysis of Retracted Papers in Chinese Academic Journal Databases

Authors: Xu Hongping, Xu Hongping

Date: 2023-05-05T00:00:00+00:00

Abstract

Objective: To analyze problems in 404 retracted papers from CNKI, Wanfang, and VIP databases and propose targeted recommendations to provide references for governing academic misconduct. **Methods:** By retrieving 370 retraction notices from the three major Chinese journal databases, we analyzed the journal distribution, discipline type distribution, impact on journal evaluation, annual distribution of “retraction papers, retracted papers, and papers retracted in the publication year”, “average annual lag months for retraction papers and retracted papers”, retraction reasons for overall retracted papers and medical and health category retracted papers, and citation status of retracted papers. **Results:** Medical and health category retracted papers accounted for the highest proportion; the number of retracted papers did not have a very significant impact on journal impact factors or whether they were included in core journal databases; the number of retraction papers in 2020 and 2021 far exceeded the number of retracted papers, the average annual lag months for retraction papers slowly increased with fluctuations, while the average annual lag months for retracted papers gradually decreased; both overall retracted papers and medical and health category retracted papers had the highest proportion of academic misconduct as the retraction reason; in the 10 years after the use of academic misconduct detection systems, retracted papers and academic misconduct retraction reasons increased by 206 papers and 155 cases respectively compared to the 15 years before use; CNKI had the most searchable retracted papers, cited papers, and citation frequencies among the three databases, with serious post-retraction citation. **Conclusion:** To solve the problem of retracted papers, it is necessary to conduct academic ethics training for researchers, editorial departments should formulate and implement strict regulations and improve editors’ ability to identify and prevent academic misconduct papers, databases should strictly control paper entry and standardize the handling of retracted papers, and relevant departments should improve punishment mechanisms and evaluation systems, create a national manuscript tracking platform, and establish

specialized problematic paper review departments.

Full Text

Preamble

Analysis of Retracted Papers in Chinese Academic Journal Databases

Editorial Office of Journal of Sanjiang University, No. 310 Longxi Road, Yuhua District, Nanjing, Jiangsu Province, 210012

Abstract: This study analyzes 404 retracted papers identified in CNKI, Wanfang, and VIP databases to provide insights for academic misconduct governance. By examining 370 retraction notices, we investigate the journal distribution, disciplinary categories, impact on journal evaluation metrics, annual distributions of retractions, temporal lag patterns, retraction reasons, and citation patterns of retracted papers. Our findings reveal that medical and health sciences account for the highest proportion of retractions. While retraction volume does not significantly correlate with journal impact factors or core database inclusion status, the number of retractions in 2020 and 2021 substantially exceeded the number of papers published in those years that were later retracted, indicating intensified cleanup efforts. The average lag time for retractions has gradually increased, while the average lag time for retracted papers has decreased. Academic misconduct remains the primary reason for retractions across all fields. Notably, the decade following the implementation of plagiarism detection systems saw 206 more retracted papers and 155 additional misconduct-related retractions compared to the preceding 15-year period. Among the three major databases, CNKI shows the highest detection rates and citation frequencies for retracted papers, with serious post-retraction citation issues observed. To address these challenges, we recommend comprehensive academic ethics training for researchers, stricter editorial policies and enhanced misconduct identification capabilities at editorial offices, rigorous quality control and standardized retraction processing at database levels, and improved disciplinary mechanisms and evaluation systems coupled with the establishment of a national manuscript tracking platform and dedicated problematic paper review departments.

Keywords: retracted papers; papers being retracted; papers retracted in publication year; disciplinary category; average lag months per paper; retraction reasons; citation frequency

Introduction

Recent years have witnessed growing research on retracted papers in international databases such as Retraction Watch, Web of Science, PubMed, Scopus, and ScienceDirect. For instance, Tian Ruiqiang et al. identified duplication, errors, irreproducibility, and fabrication as primary retraction causes, advocating for technical controls to prevent duplication, collaborative governance and long-term institutional mechanisms to address irreproducibility and fabrication,

and open scientific data practices to effectively curb data-related misconduct [1]. Yao Changqing et al. demonstrated that retracted papers exert lasting impacts on the scientific community, misleading researchers' assessments of current research and imposing high temporal costs on scientific self-purification [2]. Domestic scholars have also examined retracted papers in Chinese databases. Zhang Qing et al., analyzing 92 retracted papers from CNKI, found concentrations in medical and health sciences, industrial technology, and natural sciences, proposing corresponding recommendations for authors, editorial offices, and domestic databases [3]. Wen Min's study of 211 retracted papers across CNKI, VIP, and Wanfang revealed substantial time lags in retractions and persistent citation of retracted papers, urging editorial offices and databases to improve post-retraction processing [4]. Yang Zhu, examining the Liang Ying and Dong Peng retraction incidents alongside 205 retraction notices from CNKI and Wanfang, identified discrepancies between stated and actual retraction reasons, attributing these distortions to factors at national, industry, journal, and author levels, and recommended standardized retraction protocols and improved academic climates [5].

However, existing studies on both international and Chinese retracted papers rarely address several critical dimensions: disciplinary distribution patterns, impacts on journal evaluation metrics, comparative analyses among "retracted papers," "papers being retracted," and "papers retracted in publication year," temporal lag comparisons, comparative retraction reasons between medical and overall papers, pre- and post-plagiarism-detection system comparisons, and cross-database comparisons of detection rates and citation patterns. Moreover, previous Chinese studies employed relatively small sample sizes. This study addresses these gaps by analyzing 404 retracted papers corresponding to 370 retraction notices from CNKI, Wanfang, and VIP databases from their inception through December 31, 2021, examining the aforementioned dimensions to enhance the deterrent effect of retractions, foster healthier academic environments, improve database quality, and promote high-quality journal development.

Data Collection

From June 6 to June 10, 2022, we conducted advanced searches across CNKI, Wanfang, and VIP databases. Search parameters included: journal scope set to "all journals," "journal articles," and "all journals" respectively; time ranges from default/unlimited/database inception to December 31, 2021; and search terms including "retraction notice," "retraction announcement," "retraction application," and "revocation statement" applied to titles, keywords, abstracts, and subjects. After preliminary screening, we excluded "informational documents from editorial offices to authors," "title-only entries," and "non-academic journal documents." We manually downloaded valid retraction notices (hereafter "retraction notices") and recorded their titles, journal names, retraction dates, and original publication dates. To ensure comprehensiveness, we performed secondary searches using additional terms such as "statement," "apol-

ogy,” “retraction apology,” “retraction,” “announcement,” “notice,” and “retraction letter.” This process yielded 370 valid retraction notices covering 404 retracted papers, representing one of the largest samples in Chinese retraction research.

Conceptual Definitions

Retracted papers refers to the number of problematic papers retracted in a given year. **Papers being retracted** denotes the number of problematic papers published in a given year that were later retracted. **Papers retracted in publication year** indicates problematic papers both published and retracted within the same year. **Average lag months for retractions** is calculated as the sum of months between publication and retraction for all papers retracted in a given year divided by the number of retracted papers that year. **Average lag months for retracted papers** represents the sum of months between publication and retraction for all papers published in a given year that were later retracted, divided by the number of such papers.

Results

2.1 Journal Distribution

The 404 retracted papers were distributed across 313 journals. Forty-two journals retracted two or more papers, while the remaining 271 journals retracted one paper each. Journals retracting exactly two papers include *Power System Technology*, *Chinese Journal of Clinical Neurosurgery*, *Film Literature*, *Chinese Journal of Biochemistry and Molecular Biology*, *Dongyue Tribune*, *Chinese Journal of Experimental Traditional Medical Formulae*, *Hubei Sports Science and Technology*, *Chinese Journal of Experimental Hematology*, *Chemical Industry and Engineering Progress*, *Chinese Journal of Tissue Engineering Research*, *Journal of Clinical Oncology*, *Cancer Research on Prevention and Treatment*, *Structure & Environment Engineering*, *Oncology*, *Human Geography*, *Automation Instrumentation*, *Journal of Shenyang Pharmaceutical University*, *Modern Distance Education*, *Medical Recapitulate*, *Plant Physiology Journal*, *Chinese Journal of Comparative Medicine*, *Chinese Journal of Spine and Spinal Cord*, and *Science China: Information Sciences*. The journal with the most retractions was *Journal of Clinical and Pathological Research*, which retracted 17 papers from a single supplement issue in 2015. The retraction notice cited vague “personal reasons” from authors, though the papers shared similar content, all originated from the Third Surgery Department of Hebei Medical University Fourth Hospital, and featured various combinations of departmental staff as authors. The second-highest retraction count was *Printing Quality & Standardization*, which retracted multiple papers by the same author published between 2013 and 2015.

2.2 Disciplinary Distribution

To characterize the disciplinary distribution of the 313 journals publishing retracted papers, we analyzed the final letters of their CN numbers, which correspond to disciplines in the *Chinese Library Classification*. The journals spanned 39 disciplines. Retracted papers predominantly appeared in medical and health sciences, light industry and handicrafts, biological sciences, chemical industry, and education journals. Medical and health sciences journals published the most retractions (168 papers, 41.58%), substantially higher than the 31.52% (29/92) reported by Zhang Qing et al. in 2014 [3] but lower than the 60.67% (128/211) found by Wen Min in 2016 [4]. Despite varying percentages across studies, the consistent predominance of medical and health retractions warrants serious attention from that field. Light industry and handicrafts ranked second (19 papers, 4.70%), including journals such as *Food Safety and Quality Detection Journal*, *Food Science*, and *Tobacco Science & Technology*. Biological sciences placed third (15 papers, 3.71%), represented by *Acta Physiologica Sinica*, *Life Sciences*, and *Chinese Journal of Biotechnology*. Chemical industry and education journals tied for fourth (13 papers each, 3.21%), with chemical industry examples including *Guangdong Chemical Industry* and *Chemical Industry and Engineering Progress*, and education examples including *Education Science Forum* and *Modern Distance Education*.

2.3 Impact of Retraction Volume on Journal Evaluation

To assess whether retraction volume affects impact factors and core database inclusion, we surveyed impact factors and database indexing status for journals retracting three or more papers. *Chinese Journal of Biochemical Drugs* and *Printing Quality & Standardization* lacked composite impact factors in 2021, indicating they were not included in impact factor statistics sources. *Film Literature*, *Mining Express*, *Chinese Journal of Biochemical Drugs*, and *Printing Quality & Standardization* were not indexed in any current core databases including CA, JST, CSCD, WJCI, EI, *Chinese Core Journals Overview*, or CSSCI. However, several high-impact journals indexed in five or more core databases also had retractions, including *Journal of Astronautics*, *Journal of Computer Research and Development*, *Chinese Journal of Biotechnology*, *Food Safety and Quality Detection Journal*, *Chinese Journal of Neuromedicine*, and *Studies in Science of Science*. Thus, when retraction volume is not excessively high, no significant correlation exists between retraction count and impact factor or core database inclusion, consistent with Bao Jingling's findings [6].

2.4 Annual Distribution

Among the 404 retracted papers, four lacked publication dates in their retraction notices and had been completely removed from databases. Figure 1 [Figure 1: see original paper] illustrates the annual distribution of retracted papers, papers being retracted, and papers retracted in publication year. The earliest retraction notice appeared in *Urban Mass Transit* in 2003, retracting a 1999 paper for

plagiarism. The oldest retracted paper in our dataset was published in 1997 in *Chinese Health Quality Management*, later than the 1992 paper retracted in 1995 identified by Zhou Zhixin [7] but much later than PubMed's earliest retraction (a 1959 paper retracted in 1966) reported by Chen Xiaoqing et al. [8].

Overall, from 1997 to 2021, all three metrics showed wave-like upward trends. Between 1997 and 2005, retraction counts were minimal (one each in 2002 and 2003, zero in other years). From 2006 to 2021, retraction peaks occurred in 2008 (21 papers), 2015 (40 papers), and 2020 (104 papers). Pre-2005 annual counts of papers being retracted remained below six, possibly because many papers were not yet database-indexed or because researchers maintained higher ethical standards then. Papers being retracted peaked in 2008 (22), 2015 (47), and 2019 (59). The 2019 peak was 2.68 times the 2008 peak, and the 2020 count (53) was only slightly lower than 2019 despite being closer to our 2022 study year, indicating that intensified scrutiny has not curbed the proliferation of problematic papers, underscoring the ongoing challenges in academic integrity enforcement.

Papers retracted in publication year were far fewer than papers being retracted. Between 1997 and 2005, only one such case occurred (in 2002). From 2006 to 2021, peaks aligned with overall retraction peaks in 2008 (10), 2015 (29), and 2020 (26), demonstrating that authorities intensified scrutiny of recently published papers during peak years.

These peaks coincided with major policy initiatives. Around the 2008 peak: the Ministry of Science and Technology's *Measures for Handling Scientific Misconduct in National Science and Technology Programs (Trial)* (2006), the Chinese Academy of Sciences' *Opinions on Strengthening Scientific Research Conduct* (2007), and the Ministry of Education's *Notice on Seriously Dealing with Academic Misconduct in Higher Education Institutions* (2009). Around the 2015 peak: the State Council's *Outline for the Construction of Social Credit System (2014-2020)* (2014), the CAST-Seven Ministries' *Five Prohibitions on Academic Paper Publication* (2015), and the Ministry of Education's *Measures for Preventing and Handling Academic Misconduct in Higher Education Institutions* (2016). Around the 2020 peak: CAST's *Self-Discipline Code for Scientific and Technological Workers* (2017), the Ministry of Industry and Information Technology and Ministry of Science and Technology's *Guiding Opinions on Handling Authors Involved in Mass Retractions* (2017), the General Office of the CPC Central Committee and State Council's *Opinions on Further Strengthening Research Integrity* (2018), the Ministry of Education's *Notice on Severely Investigating and Punishing Buying and Ghostwriting of Theses in Higher Education Institutions* (2018), the General Office of the CPC Central Committee and State Council's *Opinions on Further Promoting the Spirit of Scientists and Strengthening Work Style and Academic Atmosphere* (2019), CAST-Central Propaganda Department-Ministry of Education-Ministry of Science and Technology's *Opinions on Deepening Reform and Cultivating World-Class Scientific Journals* (2019), the 20-ministry *Rules for Investigating and Handling Research*

Integrity Cases (Trial) (2019), and the National Press and Publication Administration's *Academic Publishing Standards: Definition of Academic Misconduct in Journals* (2019). These policies reflect national commitment to combating misconduct, yet problematic papers persist, suggesting inadequate disciplinary mechanisms and evaluation systems.

Comparing the three metrics reveals that papers retracted in publication year always fall below papers being retracted. Before 2019 (except 2009, 2010, 2012, and 2016), retracted papers also numbered fewer than papers being retracted, widening the cumulative gap. The ideal scenario occurs when a year's retraction volume eliminates not only all current-year problematic papers but also accumulated historical ones, after which only current-year papers require retraction—achieving equilibrium where retracted papers = papers being retracted = papers retracted in publication year, minimizing the consequences of problematic publications. Figure 1 shows 2020 reached an unprecedented retraction peak, followed by a sharp decline in 2021 where papers being retracted equaled papers retracted in publication year. Future efforts should aim to reduce retraction volume to zero while maintaining equality between papers being retracted and papers retracted in publication year, requiring coordinated efforts from all stakeholders to accelerate historical cleanup while synchronizing current-year retractions.

2.5 Average Lag Time per Paper

To evaluate the effectiveness of retraction efforts in purifying the academic environment, we calculated two metrics: average lag months for retractions and average lag months for retracted papers. Generally, longer retraction lags increase negative impacts. Higher average lag months for retractions may indicate greater contributions to academic purification by addressing accumulated problems, while lower average lag months for retracted papers suggest more timely identification and retraction of recent problematic publications, also contributing positively. Figure 2 [Figure 2: see original paper] presents these trends.

Before 2005, only one retraction notice was published (in 2003), yielding an average lag of 44 months for that year alone. The minimum average lag for retractions occurred in 2006 (5 months). Since then, despite fluctuations, the metric has trended upward, peaking in 2012 (42 months) and remaining high in recent years (37, 30, and 36 months for 2019-2021), indicating substantial contributions from recent retraction efforts. For papers being retracted, the maximum average lag occurred in 1997 (267 months). No retractions existed for 2000 publications, leaving that year undefined. The minimum appeared in 2021 (3 months). The 25-year trend shows consistent decline, suggesting increasingly effective purification efforts. Both metrics moving in positive directions demonstrate heightened attention to retracting historical problematic papers while simultaneously addressing current-year issues, minimizing overall impact.

2.6.1 Analysis of Retraction Reasons

Following guidelines from *Journal Editor Conduct* and the national press and publication industry standard CY/T 174-2019 *Academic Publishing Standards: Definition of Academic Misconduct in Journals*, we categorized retraction reasons into three major types: editorial errors, honest errors, and academic misconduct. Academic misconduct comprised six subcategories: plagiarism (including copying), fabrication, duplicate submission, redundant publication, inappropriate authorship, and other misconduct. Among 404 retracted papers published 1997-2021, we identified 372 retraction reason records. Given the predominance of medical and health sciences papers, we also analyzed 158 reason records specific to this field [Figure 3: see original paper].

Plagiarism accounted for the most records (114, 30.65%), exceeding the 28.3% reported by Zhang Qing et al. [3], and included full-text translation from English, complete copying, and partial copying, with 38 cases in medical/health papers. Inappropriate authorship ranked second (64 records, 17.20%), encompassing unauthorized corresponding author designation, institutional name misuse, and inconsistent author consent, including 28 medical/health cases. Honest errors placed third (63 records, 16.94%), covering content/data errors and insufficient theoretical research, with 26 medical/health cases. Duplicate submission (author-caused, not editorial) contributed 53 records (14.25%). Only one record involved ghostwriting, likely due to the covert nature of such practices. No cases involved fake reviewers, as Chinese journals typically employ double-blind peer review without author-recommended reviewers. Overall, academic misconduct dominated (302 records, 81.18%), higher than the 67.4% reported by Zhang Qing et al. [3]. Medical/health paper retraction reasons mirrored overall patterns, consistent with Wen Min's findings [4].

2.6.2 Pre- and Post-Plagiarism-Detection System Comparison

In late 2008, the AMLC Management Office began offering the newly developed *Academic Misconduct Literature Detection System for Scientific Journals (AMLC)* free to editorial offices providing data to CNKI [9]. Wanfang's similarity detection system appeared on its website in March 2010 [9], and VIP's citation detection system allowed individual user registration for free self-checks starting September 2010 [9]. The system underwent development and upgrades during 2008-2010, with rising academic attention. Widespread adoption began around 2012. We therefore compared retraction reason records for 1997-2011 and 2012-2021 [Figure 4: see original paper].

The earlier period yielded 89 reason records versus 283 in the later period, with substantial increases across all categories: plagiarism (50 additional records), inappropriate authorship (42), other misconduct (32), honest errors (31), redundant publication (14), duplicate submission (13), editorial errors (8), and fabrication (4). Post-detection academic misconduct records exceeded pre-detection totals by 155. Among 344 papers with available data, 34 were published before

2012 and 310 after. Combined with Figure 1, the decade following detection system implementation (2012-2021) saw 303 retracted papers—206 more than the 197 retracted papers from the preceding 15 years (1997-2011). This increase despite detection systems demands heightened vigilance from editorial offices, databases, and management departments regarding rigorous duplication checks and data entry screening.

2.7.1 Detection Rates and Citation Patterns

Among 404 retracted papers, the three databases collectively detected 534 paper-occasions. After removing duplicates across databases, these corresponded to 302 unique papers, with 102 papers completely deleted. The databases recorded 325 citation occasions across 191 unique cited papers. The 302 detectable papers received 2,687 total citations, averaging 8.90 citations per detectable paper ($2,687/302$)—lower than the 24.9 citations per retracted paper in Web of Science (1900-2019) reported by Wang Zaixing and Zhou Zhixin [11] but higher than the 6.41 citations in recent Web of Science papers analyzed by Yao Changqing et al. [2]. The 191 cited papers averaged 14.07 citations each ($2,687/191$). Table 3 shows CNKI had the highest detection occasions (223), citation occasions (144), and citation frequency, at 2.03, 2.25, and 3.91 times VIP's respective values. Wanfang ranked second, at 1.83 times VIP for detection and citation occasions and 3.13 times for citation frequency. The severe post-retraction citation problem, particularly in CNKI and Wanfang, requires urgent attention.

2.7.2 Citation Patterns in Medical/Health Retractions

Retracted paper citations pose severe consequences, particularly in medical/health fields where erroneous data may endanger lives. We therefore examined the 168 medical/health retractions, finding 73 cited papers with 128 citation occasions totaling 878 citations. CNKI cited 60, Wanfang 44, and VIP 24, with overlapping citations across databases. Post-retraction citations (663) were 3.08 times pre-retraction citations (215), exceeding the overall ratio of 2.46 ($1,910/777$). The most post-retraction citations went to a *Chinese Journal of Spine and Spinal Cord* paper on bone cement procedures (plagiarism), followed by a *National Medical Journal of China* paper on sedative effects (redundant publication).

2.7.3 Highly Cited Medical/Health Retractions

Table 4 lists medical/health journals with papers cited over 10 times post-retraction, their total citations, and retraction reasons. Medical/health retractions exhibit more severe citation problems than overall patterns, with most papers simultaneously cited across all three databases, amplifying dissemination risks.

Conclusions

Our analysis yields seven key conclusions. First, regarding journal distribution: cases where journals simultaneously retracted entire supplement issues with multiple cross-affiliated authors from one department, or multiple papers by one author with unauthorized institutional designations, highlight the need for strengthened editorial management and researcher ethics training. Second, disciplinary distribution shows high retraction volumes in medical/health, light industry/handicrafts, biological sciences, chemical industry, and education, requiring supervisory departments—especially those overseeing high-retraction fields—to enhance governance through improved disciplinary mechanisms and evaluation systems. Third, retraction volume does not significantly affect journal impact factors or core database inclusion, enabling editors to adopt proper perspectives on retractions and proactively identify and retract problematic papers. Fourth, annual distribution patterns show 2020 and 2021 retraction volumes substantially exceeding papers being retracted, indicating rapid reduction of the historical backlog. As this backlog diminishes, governance focus must shift from retracting historical papers to addressing current-year publications, necessitating dedicated review departments for timely current-year paper scrutiny and retraction. Fifth, trends in average lag months—rising for retractions and falling for retracted papers—demonstrate policy effectiveness, yet increasing absolute retraction numbers underscore the need for ethics training and improved disciplinary mechanisms. Sixth, retraction reasons reveal academic misconduct as the primary cause for both overall and medical/health papers, with post-detection system increases of 206 retracted papers and 155 misconduct-related reasons compared to the pre-detection period. Since misconduct and honest errors originate from authors while editorial errors stem from editors, comprehensive ethics training, improved mechanisms, enhanced editorial management, and a national manuscript tracking platform are essential. Seventh, citation analysis shows CNKI leads in detection and citation metrics, followed by Wanfang, with both databases' metrics approximately double VIP's detection rates and triple its citation frequencies, demanding coordinated action across databases, editorial offices, and researchers.

Recommendations

Chinese academic journals should strengthen internal mechanisms, enhance pre-publication review, establish robust post-publication retraction systems, and improve publication quality to contribute to a healthier academic ecosystem [10]. Based on our seven conclusions, we propose recommendations targeting four stakeholder groups.

3.2.1 Academic Ethics Training for Researchers

Comprehensive ethics training should elevate researchers' moral standards. First, institutions should regularly invite authoritative experts to train

researchers on nearly 20 policy documents issued since 2006, including the *Measures for Handling Scientific Misconduct in National Science and Technology Programs (Trial)*, *Opinions on Strengthening Scientific Research Conduct*, and *Notice on Seriously Dealing with Academic Misconduct in Higher Education Institutions*. Training should incorporate academic ethics competitions to reinforce policy comprehension and foster proper ethical values. Second, experts from publishing associations should teach publication standards regarding funding acknowledgments, citations, authorship, and submission protocols, using case studies to amplify retractions' deterrent effect, cultivate ethical sensitivity, and internalize moral conduct.

3.2.2 Enhanced Editorial Office Management

3.2.2.1 Strict Institutional Policies

Our findings of redundant publications, unauthorized reprints, and mass retractions of department-affiliated supplement issues demonstrate that editorial errors also cause retractions. Editorial offices must implement strict, rational regulations: enforce the three-review system, three-proofreading process, and post-publication review with clear accountability and 奖惩 mechanisms; establish editorial team management requiring qualification exams, professional development, annual training requirements, and research output to enhance competence and prevent error-induced retractions; and create efficient review-feedback mechanisms to shorten processing cycles and prevent duplicate submissions due to excessive delays.

3.2.2.2 Improved Misconduct Identification Capabilities

Since retraction volume does not significantly affect journal metrics when modest, editors should adopt proper evaluation perspectives. As gatekeepers, editors must maintain acute observational skills and vigilance against misconduct, enhancing detection capabilities to eliminate problematic papers pre-publication. Combining plagiarism detection systems with manual review is crucial. Most misconduct can be detected through systems, but editors should perform multiple checks during initial review, final editing, and pre-printing to prevent time-lag plagiarism. Utilizing tracking systems to detect simultaneous submissions and identify agency-submitted manuscripts is essential; monitoring submission status and batch re-checking before printing effectively prevents duplicate submissions [12]. Submission systems can also identify simultaneous submissions. For covert misconduct, editors must demonstrate sharp insight, identifying agency submissions through manuscript metadata [13] and submission IP addresses [14], and detecting translated plagiarism through reference patterns and language characteristics.

3.2.3 Enhanced Database Misconduct Management

First, databases must strengthen quality control at data entry. Among 372 retraction reasons, 237 cases (plagiarism, duplicate submission, redundant publication, editorial duplicate publication) are detectable by plagiarism detection

systems. Databases should re-screen all pending entries using such systems to prevent problematic papers from entering databases. Second, databases must standardize retraction processing. Retracted papers should not be deleted from electronic archives but clearly marked to reduce harmful impacts [15]. However, our study reveals issues including: failure to publish required retraction notices, deletion of papers after notice publication, direct deletion without notice, absence of “retracted” labels/watermarks/“do not cite” warnings, and severe post-retraction citations. Databases should standardize processing by supplementing missing notices, restoring deleted papers with notices, and marking citations to retracted papers as “invalid.” Future retractions must follow uniform standards.

3.2.4 Enhanced Governance by Relevant Authorities

3.2.4.1 Improved Disciplinary Mechanisms and Evaluation Systems

Since 2006, nearly 20 authoritative documents on research integrity have been issued, proving somewhat effective in retracting historical problematic papers. However, new problematic papers have not significantly decreased. Authorities at all levels must develop stricter, more detailed, and operational disciplinary mechanisms with zero-tolerance policies, ensuring severe consequences for misconduct: reclaiming benefits and honors, revoking positions, terminating employment, and creating a public misconduct registry linked to public security and credit systems to permanently restrict offenders’ activities. However, disciplinary mechanisms alone are insufficient. Medical misconduct, for instance, stems from complex societal contexts like resource scarcity and excessive workloads. Health authorities must address these root causes while reforming evaluation systems to truly “break the five-only obsession” (the fixation on papers, titles, degrees, awards, and projects). Similarly, education authorities must balance teaching loads and research evaluation systems to enable genuine scholarship. All high-workload fields require evaluation system reforms that reconcile workload and research assessment.

3.2.4.2 Creation of a National Manuscript Tracking Platform

CNKI’s 2012 manuscript tracking system enables comparison of unpublished submissions, showing submission times, similarity ratios, submission counts, and (for participating journals) editorial office names [10], helping prevent simultaneous submissions. However, limited participation and CNKI-only coverage restrict effectiveness. We propose that research integrity authorities establish a unified national platform for all Chinese academic journals to compare unpublished manuscripts, automatically moving published works to a separate archive. Mandatory participation should be codified in policy, requiring editorial offices to screen all submissions at the first processing step, thereby preventing simultaneous submissions at the source.

3.2.4.3 Establishment of Dedicated Problematic Paper Review Departments

Problematic papers comprise historical and current-year categories. Governance involves two tasks: retracting historical papers and addressing current-year pub-

lications. Our study shows 2020 marked an unprecedented peak in historical retractions. As governance intensifies, historical cleanups will accelerate. Once historical papers are fully retracted, focus must shift entirely to current-year papers. Achieving the ideal state where current-year problematic papers are retracted within the same year would eliminate historical backlogs permanently. While difficult, striving toward this goal maximizes deterrence. We therefore recommend establishing dedicated problematic paper review departments subdivided into historical and current-year units with distinct responsibilities. When both units reach zero retractions, the ideal state will be achieved.

Limitations and Future Directions

This study's limitation lies in presenting absolute retraction counts across 39 disciplines without reflecting severity relative to publication volume. Future research will calculate retraction rates by dividing discipline-specific retraction counts by total articles published in those journals' corresponding issues, providing more precise severity comparisons to inform targeted misconduct governance. We are confident that coordinated efforts among researchers, editorial offices, databases, and authorities will improve China's retraction landscape, enhance database quality, and elevate Chinese journals' international standing.

References

- [1] Tian Ruiqiang, Yao Changqing, Pan Yuntao. Research on academic misconduct in retracted papers and countermeasures[J]. Chinese Journal of Scientific and Technical Periodicals, 2019, 30(9): 937-943.
- [2] Yao Changqing, Tian Ruiqiang, Yang Dongyu, et al. Research on retracted papers and their academic impact[J]. Chinese Journal of Scientific and Technical Periodicals, 2014, 25(5): 595-604.
- [3] Zhang Qing, Yao Changqing, Pan Yuntao, et al. Research on retracted papers in Chinese academic journals[J]. Chinese Journal of Scientific and Technical Periodicals, 2014, 25(5): 611-615.
- [4] Wen Min. Research on distribution patterns of retracted papers in Chinese literature databases[J]. Chinese Journal of Scientific and Technical Periodicals, 2016, 27(11): 1151-1155.
- [5] Yang Zhu. Distortion of retraction reasons in Chinese academic journals and reflections[J]. Acta Editologica, 2021, 33(4): 430-433.
- [6] Bao Jingling, Pan Yang, Wei Peifang, et al. Investigation and analysis of retraction reasons in international medical journals: A case study of Scopus database[J]. Acta Editologica, 2018, 30(3): 323-327.
- [7] Zhou Zhixin. Characteristic analysis and implications of retracted papers in Chinese scientific journals[J]. Chinese Journal of Scientific and Technical Periodicals, 2017, 28(11): 1065-1070.
- [8] Chen Xiaoqing, Xing Meiyuan. Bibliometric analysis of retracted papers in PubMed database[J]. Chinese Journal of Scientific and Technical Periodicals, 2016, 27(4): 352-358.

- [9] Zhang Minhao, Gao Guolong, Qian Junlong. Comparative study of domestic and foreign academic misconduct detection system platforms[J]. Chinese Journal of Scientific and Technical Periodicals, 2011, 22(4): 514-521.
- [10] Liao Anlan, Jin Zihan, Zhou Zhixin. Comparative study of retracted paper characteristics in China, Japan, India, and South Korea[J]. Chinese Journal of Scientific and Technical Periodicals, 2022, 33(08): 1147-1155.
- [11] Wang Zaixing, Zhou Zhixin. Characteristic analysis and implications of retracted papers in Web of Science database[J]. Jiangsu Science and Technology Information, 2020, 37(17): 8-10.
- [12] Guo Wei. Practice of preventing “simultaneous submission” using the “manuscript tracking” platform[J]. Acta Editologica, 2016, 28(05): 470-473.
- [13] Wang Jingzhou. Identification and prevention of agency-submitted papers[J]. Chinese Journal of Scientific and Technical Periodicals, 2018, 29(6): 557-562.
- [14] Yang Chunhua, Yao Min, Liu Na, et al. Identification, prevention, and governance of ghostwritten papers[J]. Chinese Journal of Medical Library and Information Science, 2020, 29(07): 65-68.
- [15] Wen Min. Investigation and suggestions on labeling of “retraction notices”[J]. Acta Editologica, 2020, 32(3): 278-283.

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

Source: ChinaXiv — Machine translation. Verify with original.