Feature Identification of Retracted Papers Based on Postprints Questioned on PubPeer

Authors: Lin Yuan, Lin Fangyu, Zhang Zhaoyun, Ding Kun, Lin Hongfei

Date: 2024-10-06T00:00:00+00:00

Abstract

Objective: This study aims to more scientifically and accurately identify and predict the characteristics of retracted papers, thereby better correcting research misconduct, reducing academic fraud, and protecting the fairness of science.

Methods: Based on data from the PubPeer platform regarding questioned papers, this study establishes a dataset of 1,792 retracted papers with complete data after data processing and screening, analyzes the natural and comment attributes of these retracted papers, and performs feature identification and accuracy detection on retracted papers by constructing the PubCancel model.

Results: The PubCancel model constructed in this paper accurately and effectively mines and identifies features of retracted papers, with the accuracy of retracted paper status verification reaching 98.24%98.24%. This method not only has important practical significance for evaluating paper quality but also provides researchers with a method for quickly assessing paper credibility.

Conclusion: Studying the situation of retracted papers among questioned papers on the PubPeer platform is of great significance to academic early warning research. Applying the retracted paper feature identification model can help journals and editors detect anomalous papers in a timely manner, standardize paper publication and promptly retract problematic papers, and strengthen the construction of research integrity.

Full Text

Preamble

Identifying Characteristics of Retracted Papers: A Study Based on Papers Questioned on the PubPeer Platform

Lin Yuan[1]), Lin Fangyu[1]), Zhang Zhaoyun[1]), Ding Kun[1]), Lin Hongfei[2])

[1]) School of Public Administration, Dalian University of Technology, No. 2 Linggong Road, Ganjingzi District, Dalian, Liaoning 116024
[2]) School of Computer Science and Technology, Dalian University of Technology, No. 2 Linggong Road, Ganjingzi District, Dalian, Liaoning 116024

Abstract:
[**Purpose**] This study aims to more scientifically and accurately identify and predict the characteristics of retracted papers, thereby better correcting research misconduct, reducing academic fraud, and safeguarding scientific fairness.
[**Methods**] Based on data from papers questioned on the PubPeer platform, we established a dataset of 1,792 retracted papers with complete data after processing and screening. We analyzed the natural attributes and comment attributes of these retracted papers and constructed a PubCancel model to identify their features and test its accuracy.
[**Results**] The PubCancel model developed in this study accurately and effectively mines and identifies the characteristics of retracted papers, achieving a 98.24% accuracy rate in verifying retraction status. This method not only holds significant practical value for evaluating paper quality but also provides researchers with a rapid means to assess paper credibility.
[**Conclusions**] Investigating retracted papers among those questioned on the PubPeer platform is of great significance for academic early warning research. Applying a feature recognition model for retracted papers can help journals and editors promptly identify problematic papers, standardize publication processes, withdraw problematic papers in a timely manner, and strengthen research integrity construction.

Keywords: Retracted papers; PubPeer; Feature recognition; Random forest; PubCancel

Since the 18th National Congress of the Communist Party of China, the implementation of the strategy for revitalizing the country through science and education, along with advances in science and technology, has led to a continuous increase in China's total research output, with the volume of research papers ranking at the international forefront. However, in recent years, rapid technological progress and the swift development of information technology and artificial intelligence have profoundly transformed research paradigms, and the number of paper retractions has consequently increased. To purify the innovation ecosystem, foster a sound research atmosphere, and maintain research norms, the State Council issued the "Opinions on Further Strengthening Research Integrity Construction" [1] to guide researchers and institutions in conducting scientific research in a standardized manner. Retracting problematic papers is an effective measure for maintaining research integrity and purifying the academic environment. As public attention to academic misconduct has grown and the number of retracted papers has increased, research on retractions has become a hot topic among scholars both domestically and internationally, focusing primarily on the analysis of retraction characteristics. Current studies on retraction characteristics mainly concentrate on data from major journal indexing platforms such as ScienceDirect [2], WoS [3-5], PubMed [6,7], RetractionWatch [8], CNKI [9], and Scopus [10]. Recent trends show increasing use of multiple databases, such as Xu Hongping's [11] study combining CNKI, Wanfang, and VIP databases to examine retractions in Chinese academic journals, and Yuan Zihan [12] and Ma Linling et al. [13] combining RW and WoS databases to study retractions in specific journals.

These studies have explored retraction characteristics across multiple databases, analyzing basic features (time, country, journal, discipline, etc.), citation characteristics, retraction lags, and reasons for retraction, providing a basis for journals to effectively identify retracted papers. However, current research primarily focuses on journal indexing platforms, which typically contain large paper volumes but low retraction proportions, resulting in high data contingency and bias. Moreover, such research is retrospective, only exploring data after retraction and failing to provide effective academic early warning. In contrast, PubPeer, as a renowned "academic whistleblowing platform," plays a crucial role in discovering and identifying problematic papers and warning against academic misconduct through its comment content. This forum questions published papers through post-publication peer review, and its data on questioned papers can effectively mine retraction characteristics, provide early warnings for academic misconduct, and offer new research perspectives for identifying academic misconduct. Among Chinese scholars, only Chu Jingshen [14] and Lin Yuan et al. [15,16] have studied PubPeer papers, focusing on the overall situation of questioned papers rather than retractions. However, do all papers questioned on PubPeer eventually get retracted? This question has practical significance for understanding the internal mechanisms of retractions and predicting which problematic papers will be retracted. Therefore, to address this question, this study conducts statistical analysis of retraction characteristic indicators on the PubPeer platform, constructs a PubCancel model to reveal how different indicators of questioned papers influence retraction, aiming to help the academic community more comprehensively understand paper retraction phenomena, more timely and effectively detect problematic papers, achieve academic early warning effects, reduce retraction lag, maintain research integrity, and purify the academic environment.

1.1 Research Hypothesis

Based on the potential value of PubPeer for retraction research, this study first proposes a hypothesis: The proportion of retracted papers among questioned papers is significantly higher than the retraction rate among papers in ordinary databases. This study only obtains the proportion of retracted papers from the questioned paper database, while the retraction rate for ordinary papers is based on previous research. For example, Liu Hong et al. [2] studied retractions in the ScienceDirect database, which contained over 10 million papers before 2011, with only 543 retractions between 1992-2010, accounting for less than one ten-thousandth. Chen Chaomei et al. [6] used the PubMed database to calculate annual retraction ratios (number of retraction notices/number of new papers added that year) from 2001-2011, with the lowest ratio being 0.00005 in 2001 and the highest 0.00046 in 2011. Qin Huijuan [17] analyzed papers indexed by SCI, examining publication and retraction numbers for five countries (USA, China, Germany, Japan, India), finding retraction rates per 10,000 papers of 0.99, 3.05, 1.13, 1.23, and 2.60 respectively. Liu Qinghai [8] used RetractionWatch to analyze retraction ratios for major publishing countries, with the Netherlands (3.06/10,000), India (2.76/10,000), and the USA (2.09/10,000) having the highest rates. Fan Shujie et al. [18] retrieved data from the Web of Science (WoS) Core Collection, analyzing retractions in SCI-E papers, finding retraction rates between 0.63-2.20 per 10,000 papers from 2008-2017. Zhou Zhixin et al. [19] compared retraction characteristics between China and the USA based on WoS data, finding retraction-to-publication ratios (total retractions/total indexed papers) of 4.1‰ (2,119/5,136,315) for China and 0.9‰ (1,541/15,715,138) for the USA before March 31, 2020. Regarding questioned papers, this study collected information on 25,518 papers from PubPeer's public data up to December 2022, finding only 10,229 papers with complete data (including retraction status) after preprocessing. After data cleaning and screening, 1,792 papers were found to be currently retracted, representing a retraction rate of 17.52%. The highest retraction rate among existing journal indexing platforms is WoS's 4.1‰ for China. This demonstrates that the proportion of retracted papers among questioned papers is significantly higher than retraction rates in ordinary databases, making the study of retractions among questioned papers crucial for standardizing post-publication review mechanisms and academic early warning research.

1.2 Data Sources and Processing

PubPeer is an open post-publication peer review forum where authors must have a paper indexed by PubMed as first or corresponding author to register. Registered scholars can comment on published papers with criticism, questions, or suggestions, but comments require review before publication. PubMed is a database providing biomedical literature search and abstracts, so questioned papers on PubPeer concentrate in biomedicine [16]. Known as the "disciplinary committee" in academia, PubPeer, along with RetractionWatch, is considered a "dynamic duo" for academic whistleblowing, having uncovered many fraudulent papers and becoming an important channel for identifying problematic papers. The scholarly opinions and emotional evaluations embedded in its comments demonstrate significant value for fraudulent paper research and retraction identification.

This study selected PubPeer as the data source. By December 2022, we obtained relevant information on 25,518 papers from PubPeer's public data, including pubpeer_{id}, title, number of question comments, journal, first author, institution, discipline, comments, retraction status, and publication date of the last comment. The raw dataset contained natural attributes (publication year, author, discipline, journal, institution) and comment features (last comment date, comment count, comment content). After removing rows with NA values, we obtained a total dataset of 10,229 papers. We retrieved paper status information from PubPeer pages, marking papers with "RETRACTED" status as retracted papers. The initial data were then classified into two categories: retracted papers (1,792) and non-retracted papers (8,437). Given the diversity of retracted papers, we extracted the 1,792 retracted papers to form a retracted paper dataset for studying their ontological and PubPeer features. The 8,437 non-retracted papers were extracted to form a non-retracted paper dataset, which was combined with the retracted dataset for comment feature analysis and model accuracy testing.

1.3 Research Methods

To more comprehensively understand paper retraction phenomena and the internal mechanisms of retracted papers on PubPeer, this study collects data on retracted papers from the platform, conducts statistical analysis of natural and comment attributes, mines differential features between retracted and non-retracted papers, validates these features through backward stepwise selection to obtain an optimal subset, and uses a random forest model for feature identification. Finally, targeted recommendations are proposed to improve management of retracted papers. The specific methods are introduced below:

This study uses PubPeer as the data source and processes the aforementioned 1,792 retracted papers through data encoding. For non-numeric natural attribute information such as discipline names, journal names, and institution names, we use the BERT model to convert them into word vectors. For PubPeer comment content, to understand whether a paper is being questioned and the severity of questioning, and to determine whether subjective text leans positive or negative, we employ attention mechanism-based sentiment polarity analysis to convert comment text into sentiment scores between 0-1 for subsequent feature representation and identification. First, through analysis of natural and comment attributes, we select 11 indicator features: publication year, discipline, journal, journal retraction index, institution, institution retraction index, author, number of question comments, comment sentiment score, and year of last comment publication. The most relevant variables to the response variable are selected to form a subset. Second, backward stepwise selection is adopted, starting with a full model containing all 11 variables and iteratively removing the least important variable that most adversely affects model fit, to determine the most effective number of features for the PubCancel model using random forest algorithm for importance ranking, followed by accuracy testing.

Random Forest Model: Random forest is a decision tree model based on loss function minimization with three construction steps: attribute selection, decision tree generation, and pruning. When establishing a decision tree, for each split point, a random sample of m predictors is selected from all p predictors as candidate variables, and the split can only use these m variables. Resampling occurs at each split point, typically with m $\sqrt{p}$, meaning the number of predictors considered at each split approximates the square root of the total number of predictors.

This study uses the Gini Index as the evaluation metric to measure feature importance. VIM represents variable importance measures, and GI represents the Gini Index. Assuming J features and decision trees with C categories, the steps to calculate the Gini Index score for each feature (the average change in node splitting impurity across all trees) are as follows.

The Gini Index for node q of tree i is calculated by formula (1):

$$GI_q = 1 - \sum_{c=1}^{C} p_{qc}^2$$

where C represents the number of categories, and p_{qc} represents the proportion of category c in node q, i.e., the probability that two randomly drawn samples from node q have different class labels.

The importance of feature X_j at node q of tree i is the change in Gini Index before and after splitting, as shown in formula (2):

$$VIM_{jq}^{(i)} = GI_q - GI_l^{(i)} - GI_r^{(i)}$$

where GI_l(i) and GI_r(i) represent the Gini Indices of the two new nodes after splitting.

If feature X_j appears at nodes in set Q in decision tree i, its importance in tree i is given by formula (3):

$$VIM_j^{(i)} = \sum_{q \in Q} VIM_{jq}^{(i)}$$

Assuming the random forest contains I trees, formula (4) gives:

$$VIM_j = \frac{1}{I} \sum_{i=1}^{I} VIM_j^{(i)}$$

Finally, all importance scores are normalized to obtain feature importance, as shown in formula (5):

$$VIM_j^{(norm)} = \frac{VIM_j}{\sum_{j=1}^{J} VIM_j}$$

Feature importance reveals each feature's impact on prediction results, aiding feature selection and model optimization.

This study implements the random forest model using Python's scikit-learn library, using the most effective feature set from backward stepwise selection as input, outputting feature importance after prediction, and calculating accuracy.

2 Retracted Paper Feature Identification

Based on the constructed retracted paper dataset, this section analyzes the publication years, disciplinary and journal types, and research entities of 1,792 retracted papers, conducting in-depth mining of these features to explore potential causes of retraction. Specific quantitative analyses are presented below.

2.1.1 Temporal Feature Analysis

Based on the publication years and quantities of retracted papers, we plotted the annual distribution of retractions. [Figure 1: see original paper] shows that the earliest retracted paper was published in 1975, with annual retraction records since 1997. Before 2002, annual retraction numbers were stable at fewer than 4 papers. Since 2002, retraction numbers have grown dramatically, particularly from 2015-2019, with extremely significant increases peaking at 533 retractions in 2019. After 2019, retraction numbers show a declining trend, possibly because papers published after 2019 have shorter publication histories and retractions exhibit time lags. In summary, retraction publication times are not uniformly distributed, showing an "upward trend" or "prominent period." Based on these temporal differences, we identify publication year as a feature that can help predict retractions.

2.1.2 Disciplinary Feature Analysis

To examine relative retraction tendencies across disciplines, we constructed a discipline retraction index (number of retracted papers in a discipline/total papers published in that discipline) and plotted the disciplinary distribution based on both retraction numbers and indices. [Figure 2: see original paper] shows that PubPeer questioned papers involve 21 disciplines. The top three disciplines with retraction numbers exceeding one-tenth of the total are Clinical Medicine, Molecular Biology & Genetics, and Pharmacology & Toxicology. Disciplines with higher retraction indices are primarily Mathematics, Pharmacology & Toxicology, Molecular Biology & Genetics, and Clinical Medicine. Three clusters emerge: Cluster one includes Mathematics, Psychology, and Space Science, with fewer retractions but higher retraction indices. Mathematics shows the highest index, with only 31 retracted papers, 18 of which came from the same journal and were retracted for overlapping with another article. Cluster two is the "biomedical" group, with high retraction numbers and indices, particularly Pharmacology & Toxicology, Molecular Biology & Genetics, and Clinical Medicine. This concentration occurs because PubPeer users must have PubMed-indexed papers to register, and PubMed focuses on biomedicine. Cluster three comprises remaining disciplines with low retraction numbers and indices. These disciplinary differences suggest that discipline affiliation can help predict retractions, making it an important identification feature.

2.1.3 Journal Feature Analysis

Based on journal IDs (assigned sequentially from 0 by retraction count), retraction numbers, and cumulative percentages, we plotted the journal distribution. [Figure 3: see original paper] shows that PubPeer retractions came from 571 journals. Most journals had few retractions: 506 journals (88.6%) had fewer than 5 retractions, while only 39 journals (6.8%) had more than 10 retractions. These 39 journals accounted for 1,415 retractions (59.9% of the total). The top three journals had 191, 159, and 133 retractions respectively. Investigation revealed that high-retraction journals are primarily "predatory journals" with large volumes, high fees, and easy acceptance, or a few journals with long-term lax peer review or regulatory loopholes. This distribution shows significant differences, with a few high-retraction journals coexisting with many low-retraction ones, suggesting that journal identity can help predict retractions.

2.1.4 Institutional Feature Analysis

Based on first author institution IDs (assigned sequentially by retraction count), retraction numbers, and cumulative percentages, we plotted the institutional distribution. [Figure 4: see original paper] shows that PubPeer retractions came from 1,063 institutions. While more widely distributed than journals, the trend is similar: most institutions had few retractions, with only 27 institutions (2.5%) having more than 10 retractions. These 27 institutions accounted for 758 retractions (32.1% of the total). High-retraction institutions are primarily medical institutions (university-affiliated hospitals or provincial/military hospitals) and well-known comprehensive universities. This institutional concentration suggests that institution identity can help predict retractions.

2.1.5 Author Feature Analysis

We analyzed first authors of retracted papers, plotting the distribution of authors by retraction count, where point (x,y) indicates y authors had x retracted papers. [Figure 5: see original paper] shows that 1,935 first authors were involved, with 1,771 (91.5%) being "first-time offenders" who had only one retraction. One hundred sixty-four scholars had more than one retraction, with only 17 authors (less than 1%) having four or more retractions. This reveals two phenomena: widespread "first-time offenders" in terms of author numbers, but concentration among "repeat offenders" in terms of paper counts. These differences suggest that author identity is an important feature for identifying retractions.

2.2.1 Comment Quantity Feature Analysis

Statistical analysis of comment quantities for the 1,792 retracted papers revealed that only 2 papers had unusually high comment counts (285 and 156 comments) compared to others (within 30 comments). After removing these outliers, we found that retraction comment counts cluster densely around 3, with an average of 3.32 comments per retracted paper. For comparison, non-retracted questioned papers averaged 2.56 comments. This difference suggests that higher comment counts can help predict retractions, making comment quantity an important feature.

2.2.2 Comment Sentiment Score Feature Analysis

Sentiment analysis of PubPeer comments produced box plots for retracted and non-retracted papers, where higher sentiment scores indicate more negative emotions. [Figure 6: see original paper] shows that retracted papers had sentiment scores in [0.7626, 0.9998], while non-retracted papers scored in [0.5057, 0.9998]. The intervals differ significantly, with retracted papers showing higher overall sentiment scores and more negative questioning comments. Papers with sentiment scores below 0.76 are less likely to be retracted. Tracking sentiment scores can provide early warnings, making comment sentiment an important feature.

2.3.1 PubCancel Model Construction

Sections 2.1 and 2.2 identified 11 features characterizing retracted papers. We constructed the PubCancel model to analyze these features, obtain importance rankings, and conduct accuracy tests. The model workflow has two parts, shown in [Figure 7: see original paper].

Part One: Backward stepwise selection validates and selects features based on input features and the dependent variable (whether a paper is retracted), identifying the most informative feature subset.

Part Two: Random forest model produces importance rankings for the selected feature subset and conducts accuracy tests, training a classifier to identify retraction status for new papers.

2.3.2 Feature Mining

Based on the above feature analysis, we mined 11 feature variables for retracted papers. The results show that using 5 features yields 98.27% accuracy, while using all 11 features yields 98.24%—nearly identical. However, comparing recall rates reveals 95.27% for the 5-feature model versus 96.36% for the full model. Given our research goal of comprehensively identifying retracted papers, we selected the full 11-variable model. This confirms that the most effective feature subset contains all 11 variables, validating our mined features.

2.3.3 Feature Importance Analysis

Based on the selected features, we used random forest to determine importance. With 11 predictors, we set m=3 (approximately $\sqrt{11}$). We converted 128-dimensional vectors into numpy arrays and saved them in the dataset. The total dataset of 10,229 papers (1,792 retracted and 8,437 non-retracted) was randomly split into 10 parts, with a 7:3 training-testing ratio. The trained random forest classifier produced the importance rankings shown in .

** Feature Importance Ranking**

The top 6 features all relate to discipline, institution, and journal, indicating these three factors are most important for identifying retractions, reflecting academic attention to these entities. The year of last comment has low importance (0.01418784), likely because it only reflects the most recent comment time rather than comprehensive comment patterns. Comment quantity also shows low importance, suggesting retractions relate more to comment content and quality than quantity. Comment sentiment score is the least important feature, indicating weak predictive power—perhaps because PubPeer comments are inherently negative, making specific content more influential than sentiment. Though less important, these features still contribute to prediction, requiring consideration of multiple features rather than relying on any single one.

2.3.4 Accuracy Testing

Using the feature division from 2.3.2, we trained a random forest classifier based on feature importance. Output metrics include overall accuracy, retraction recall (for retraction status=1), and non-retraction recall (for status=0), achieving 98.24%, 96.36%, and 98.65% respectively. The PubCancel model demonstrates high accuracy (98.24%) with high retraction recall (96.36%), capturing most retracted papers, and high non-retraction recall (98.65%), correctly predicting non-retracted status. The model performs well on both metrics, providing a viable solution for retraction status prediction.

3.1 Conclusions

Based on PubPeer data, this study collected information on questioned papers from an open platform, mining features including publication year, discipline, journal, institution, author, comment count, sentiment score, and last comment year to reveal characteristics of retracted papers among questioned ones. Using random forest, we built the PubCancel model to rank feature importance and predict retraction status. Results demonstrate that the identified natural and comment features are effective, improving prediction performance and proving the validity and feasibility of this new feature combination.

3.2.1 Journal Perspective

Since retracted papers concentrate in a few high-retraction journals and many low-retraction ones, implications for journals have two angles. For high-retraction journals: strengthen internal quality control and review rigor to improve publication quality. First, enhance pre-publication supervision: (1) implement plagiarism detection during initial review, potentially using large model technology to examine text, images, and data; (2) improve peer review quality through double-blind or multiple reviews; (3) strengthen final editorial review by requiring authors to submit code, data sources, and original figures to verify originality. Second, strengthen post-publication review to reduce retraction lag: (1) encourage data and code sharing to enable verification and reduce fraud risk; (2) establish academic early warning mechanisms using platforms like PubPeer to track and verify questioned papers promptly; (3) establish sound retraction procedures with accountability for authors, reviewers, and editors. For low-retraction journals: maintain review standards and quality to enhance reputation, (1) promptly address problematic papers and issue standardized retraction notices as examples and warnings; (2) strengthen collaboration with academic institutions through conferences and special issues to attract high-quality papers.

3.2.2 Institutional Perspective

Since retracted papers mainly come from well-known comprehensive universities and medical institutions (university-affiliated hospitals or provincial/military hospitals), institutional implications have two angles. For multi-disciplinary institutions like comprehensive universities: (1) strengthen research ethics education to improve integrity awareness among students and faculty; (2) raise paper quality standards with stricter requirements and penalties for non-compliance; (3) promote interdisciplinary exchange to enhance research complementarity and innovation. For specialized institutions including medical schools, research institutes, and hospitals: (1) increase investment in human, material, and financial resources to strengthen discipline construction and talent development; (2) promote academic cooperation to enhance research quality and impact; (3) strengthen paper writing standards through training by experts; (4) enhance research ethics education through case demonstrations to improve integrity awareness and reduce retraction risk.

As technology, information technology, and AI develop rapidly, retraction incidents increase. Retracted papers exhibit characteristic patterns in natural attributes (temporal, disciplinary, journal, institutional, author) and comment attributes (quantity, sentiment). Deep learning models can better identify these features to warn against academic misconduct. This study proposes targeted recommendations from journal and institutional perspectives to detect problematic papers more timely and effectively, maintaining research integrity and purifying the academic environment.

Limitations include: (1) dataset constraints and subjective indicator selection, such as not considering factors like author background or funding that may affect retraction status; (2) platform data limitations preventing dynamic real-time detection models. Future research should: (1) combine PubPeer questioned paper data with retraction notices to deeply analyze retraction reasons and propose more targeted recommendations; (2) collect retraction features from multiple platforms to build an academic paper risk detection platform for timely prevention and retraction of misconduct.

References

[1] General Office of the State Council, General Office of the Central Committee of the Communist Party of China. *Opinions on Further Strengthening Research Integrity Construction* [EB/OL]. [2018-05-30]. https://www.gov.cn/zhengce/2018-05/30/content_{5294886}.htm.

[2] Liu Hong, Hu Xinhe. Empirical Analysis of Retracted Papers in International Academic Journals—A Case Study of ScienceDirect Database [J]. *Chinese Journal of Scientific and Technical Periodicals*, 2011, 22(6): 848-852.

[3] 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.

[4] Fu Zhongjing. Distribution Characteristics of Retracted Papers in International Journals and Implications for Chinese Journal Publishing [J]. *Chinese Journal of Scientific and Technical Periodicals*, 2016, 27(9).

[5] Jin Zihan, Liao Anlan, Zhou Zhixin. Multi-dimensional Analysis of Reasons for Retraction of Chinese Scholars' Papers in International Journals [J]. *Chinese Journal of Scientific and Technical Periodicals*, 2023, 34(2).

[6] CHEN C, HU Z, MILBANK J, et al. A visual analytic study of retracted articles in scientific literature [J]. *J. Assoc. Inf. Sci. Technol.*, 2013, 64: 234-253.

[7] 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.

[8] Liu Qinghai. Analysis of Retraction of Chinese Scholars' Papers in International Journals—Based on RetractionWatch Website [J]. *Chinese Journal of Scientific and Technical Periodicals*, 2016, 27(4): 339-345.

[9] 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.

[10] 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.

[11] Xu Hongping. Characteristic Analysis of Retracted Papers in Chinese Academic Journal Databases [J]. *Chinese Journal of Scientific and Technical Periodicals*, 2023, 34(4): 519-529.

[12] Yuan Zihan, Jin Tong. Characteristic Analysis and Implications of Retracted Papers in High-impact International Scientific Journals—A Case Study

of Cell, Nature, and Science [J]. *Chinese Journal of Scientific and Technical Periodicals*, 2024, 35(2): 216-225.

[13] Ma Linling, Feng Lingzi, Yuan Junpeng, et al. Analysis and Reflection on Retraction of Papers in Chinese SCI/SSCI-indexed Journals [J]. *Chinese Journal of Scientific and Technical Periodicals*, 2023, 34(5): 584-592.

[14] Chu Jingshen. Analysis and Reflection on Comments Related to Chinese Academic Papers on PubPeer [J]. *Chinese Journal of Scientific and Technical Periodicals*, 2022, 33(11): 1499-1506.

[15] Lin Yuan, He Junyao, Jiang Chunlin, et al. Research on Academic Early Warning Based on PubPeer [J]. *Journal of Modern Information*, 2021, 41(7): 160-167.

[16] Lin Yuan, He Junyao, Liu Shengbo, et al. Research on Intelligent Early Warning of Academic Journals Based on PubPeer [J]. *Journal of Intelligence*, 2022, 41(9): 184-191.

[17] Qin Huijuan. Research on Retracted Papers Based on SCI [D]. Shanghai: East China Normal University, 2013.

[18] Fan Shujie, Fu Xiaoxia. Analysis and Reflection on Retraction of Chinese Authors' Papers in SCI-E (2008-2017) [J]. *Acta Editologica*, 2019, 31(1): 51-55.

[19] Zhou Zhixin, Wang Zaixing. Comparative Analysis of Characteristics of Retracted Papers Between China and the USA Based on WoS Database [J]. *Journal of Intelligence*, 2020, 39(11): 165-.

Author Contribution Statement:
Lin Yuan: Proposed research direction and designed paper framework.
Lin Fangyu: Collected data, conducted experiments, and drafted the paper.
Zhang Zhaoyun: Participated in writing and revising the paper.
Ding Kun: Revised and reviewed the paper.
Lin Hongfei: Reviewed the paper.

Corresponding Author: Zhang Zhaoyun, Master's student, E-mail: 1434810898@qq.com.

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