Understanding the Patterns and Magnitude of Life Science Publication Retractions in the last Four Decades using an Evidence-Based Approach

Sabuj Bhattacharyya[1], Nilanjan Chatterjee[2,3], Arvind Ramanathan[1]

[1]Institute for Stem Cell Science and Regenerative Medicine (BRIC-inStem), Bangalore, India, [2]University of Minnesota, St. Paul, USA
[3] Senckenberg Biodiversity and Climate Research Centre, Senckenberg Gesellschaft für Naturforschung, Frankfurt, Germany

Introduction

Publication Retraction : A Global Concern

- **Retractions** of **flawed published literature** is an important **correction mechanism** to **maintain academic integrity**[1]
- **~10** fold increase in publication retractions since the late 1970s with a recent serge with more than **10,000 retraction only in 2023**[2,3]
- Several studies have reported **continued citation of retracted articles** as a **preliminary source** of **scientific data** even long after their retraction[4]
- "**Retraction Watch**" database indicated **40% of retraction notices** did not indicate fraud or misconduct, instead **mentioned errors and issues with reproducibility**[5]
- **Publications retractions** has been reported to **impact the competitiveness** of the authors by **negatively** influencing the citation number (**8-9% reduction**) due to **trust deficit** and reduced success in acquiring funds[6]
- **Magnitude of retractions** significantly **varies across countries** and **subjects**[7]
- The **global research funding** for **life science** and **environment science** for the 2023 -Q3 is approx. **48 Billion USD**[8] and **high retraction** can cause **significant economic loss.**
- **Dynamics and magnitude** of **retractions** in leading scientific such as "**Life Science**" is **not well explored**.

Study Objective

To get a holistic view of the current pattern and magnitude of peer-reviewed publication retractions in the field of Life Sciences across the globe

Research Questions

- What are the **patterns of retraction across years** and **countries**?
- What are the **primary reasons for retractions**? **Does it vary** significantly **across various themes** within life science research ecosystem?
- Does the **magnitude of retraction varies** significantly **with group size of authors** (proxy for division of labour), **journals/ publishers** and their **impact factors** (proxy for quality checks)?
- Does the **retraction magnitude** is **linked with collaborative network of authors** as well as **cultural dimension of a particular country**?

Methodology

Data Source

- **Retraction data** (n= **38,405**) were obtained from **Retraction Watch Database** (http://retractiondatabase.org) which were **filtered** for **Life Science** related information
- The **impact factor of journals** were collected from **Web Of Science Master Journal list Database** (https://mjl.clarivate.com/home) for the year 2022
- **Country** wise **Cultural Dimension Da**ta as per **Hofstede 6 – dimension website** (www.hofstede-insights.com)

Study inclusion/ exclusion criteria

- Retracted **studies** which focused on **various aspect of life sciences** were **included**
- Retracted **studies** from **public health** or **without a clear focus** on **life science** were excluded
- **No specific timeline** were selected for the study

Data Arrangement

- **Retraction Watch database** had **19 categories** such as record id, title of the article, subject category, affiliated institutes, name of the journal, publisher, country, author, type of the article, publication and retraction dates, the reason for retraction etc.
- Many retracted articles were inter and transdisciplinary in nature thus, the **subjects** of the **retracted articles** has been **re categorised** in **22 categories** (e.g., biochemistry, cancer biology). Interdisciplinary studies without clear subject themes were categorised as "Other".
- **Retractions reasons** have been **recategorized** into **8 distinct classes** (e.g., ethical and compliance issues, data integrity). Reasons such as "Withdrawn (out of Date)", and "Publishing Ban" which can not be segregated under a specific section were placed under the "other" category.

Data Filtering & Analysis

- The final **dataset** was **filtered using** prog. Library package "**tidyverse**"[9], "**lubridate**"[10] and visualized using "**ggplot2**"[11] in program "R" (R Core Team 2023)
- The **author collaboration network** was analyzed using prog. library package "**statnet**"[12] in program "R".

Data availability

- Upon completion of the project **all processed data**, **R script**, graphs and plot **will be made publicly available** though "**GitHub**"

Conflict of Interest Statement

- Authors declare no Conflict of Interest

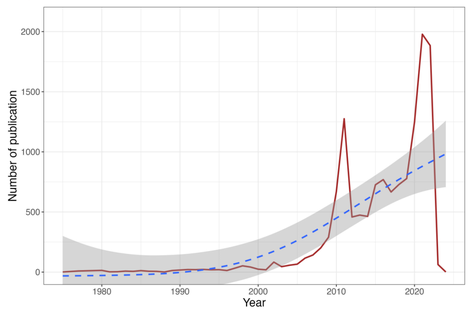
Acknowledgements

- We thank Director, BRIC-inStem providing the administrative support during the study and facilitating the first author's participation in WCRI 2024
- We thank Retraction Watch for proving us the data for the study
- We thank WCRI 2024 organizers for providing the WCRI Travel grant to SB
- This study is not sponsored by any specific funding agency or grants

Results

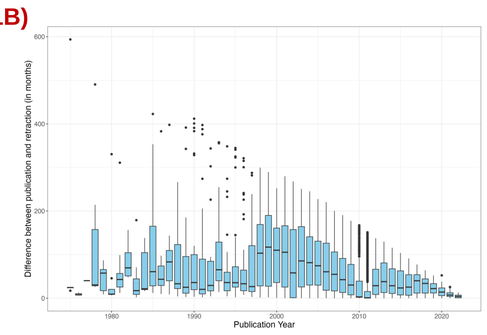
Temporal Pattern

(1A)



- A total of **13,370** life articles have been retracted till January 2023 with an **increase** of **2.5%** to **20%** since 1975.

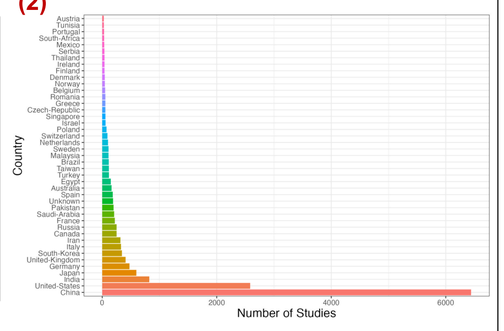
(1B)



- Articles that were **published in earlier decades** get **retracted in higher rate** in **recent past**.
- **Flawed articles** also get **retracted** much **quickly 2015 onwards**

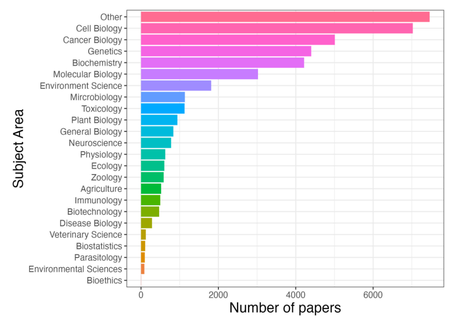
Spatial Pattern

(2)



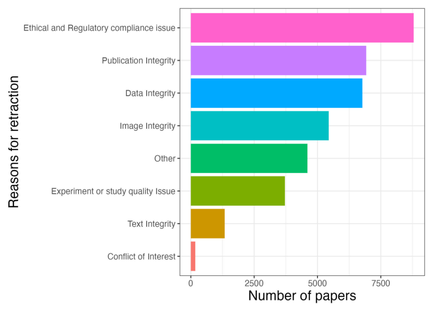
- **Highest retraction** rate was recorded in **China** (39.42%) followed by **USA** (15.81) and **India** (5.03%)
- **Retraction rate significant varied across countries** (Chi sq.=1482476, df=155, p<0.001)

(3) Retraction Pattern Vs. Subjects



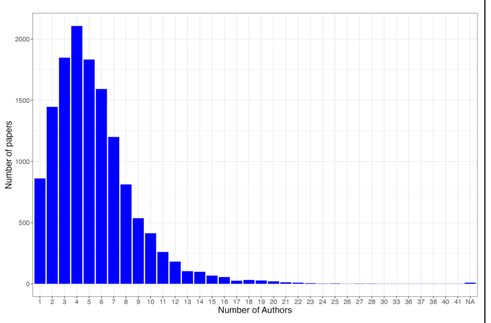
- **Inter/multi-disciplinary** life science studies (20.26%) followed by **Cell biology** (19.08%), and **Cancer Biology** (13.61%) experienced the **highest retractions**.
- **Retractions** rate **significantly varied across subjects** (Chi sq.=230051,df=134, p<0.001)

(4) Reasons for Retraction



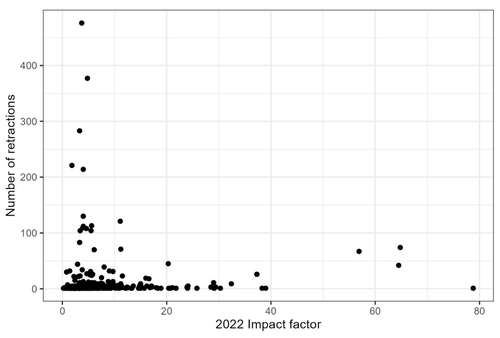
- **Ethical and Regulatory Compliance** issue (23.29%), **Publication Integrity** (18.33%), and **Data Integrity** (17.92% were primary reasons for retractions (Chi sq.=66859, df=103,p<0.001)

(5) Group Size Trend



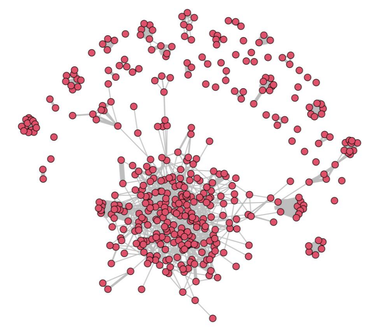
- **Articles with 3 to 5 authors** (15.53-13.51%) were **retracted more frequently than** articles with a **single author** or **more than 8 authors** (Chi sq.= 37565, df=35, p<0.001)

(6) Impact Factor Vs. Retractions



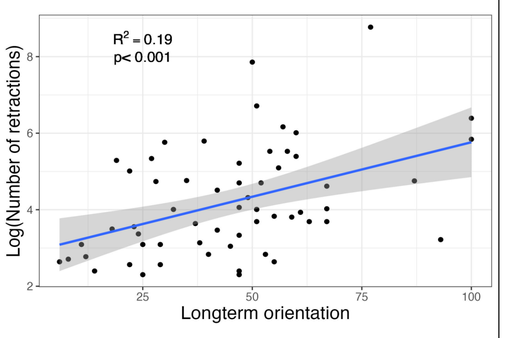
- **Life Science retractions** were **more common** among **journals with lower impact factor**

(7) Collaboration Network of Retracted Authors



- The **network of retracted authors** indicated a **higher degree of collaboration**

(8) Cultural Dimension Vs. Retractions



- **Retractions** were **positively associated** with **Long- versus Short-Term Orientation** of **Hofstede's Cultural Dimensions of countries**

Discussion

- **Awareness about academic integrity**, **use of various techniques** (e.g., use of AI-based tools) might helped journals to identify flawed publications and retract them which **resulted in increase in retractions** in the recent past.
- **Retractions** were **common** in both countries from the **global south and north** which highlights the **need for more proactive preventive measures, policy implementation** and **cooperation across countries.**
- **High retraction rates** in **multi/inter-disciplinary life science fields particularly in laboratory** and **data analysis**, **regulatory compliance** indicate a **need** for more **impactful and focused awareness** and **capacity building/ training measures**
- Unlike other studies,[13] which indicate the net size of the author group is positively related to retraction magnitude, **our study** suggested **unique quality control** and **academic integrity challenges** faced by **mid-size** (3 to 5) **author group**.
- Similar to other studies,[14] our study also revealed that the **low-impact journals experience higher burden of retractions** possibly **due to limited access to cutting-edge technology** to detect flaws during the peer review process.
- The **retracted authors** often are **part of close collaborative network** which increases the possibility of misconduct.
- Hofstede's Cultural Dimensions long-term orientation experiences **higher number of retractions** possibly due to a **lack of social obligations** and less **importance of virtues.**

Future work

The study is yet to explore influence of following factor on Life Science retraction

- **Gender** Vs **Retraction Rate**
- **Number of Authors** Vs. **Reasons for Retraction**
- **University Ranking** Vs. **Retraction Rate**
- **Extent of International Collaboration** Vs. **Retraction Rate**

References and Contact

1. Steen et a. (2013) PLoS ONE 8(7): e68397
2. Fanelli, D. (2013). PLoS medicine, 10(12), e1001563
3. Van Noorden, R. (2023). *Nature*, *624*(7992), 479-481.
4. Shuai et al. (2017). Journal of the Association for Information Science and Technology, 68(9), 2225-2236.
5. Lu et al. (2013). Scientific reports, 3(1), 3146.
6. Azoulay et al. (2017). Research Policy, 46(9), 1552-1569.
7. Oransky (2018) Science. 362(6413):395
8. Clarivate.inc 2024. https://clarivate.com/products/scientific-and-academic-research/research-funding-and-analytics/pivot-rp/funding-insights-dashboard/
9. Wickham et al (2019). *Journal of open source software*, *4*(43), 1686.
10. Grolemund & Wickham, H. (2011). *Journal of statistical software*, *40*, 1-25.
11. Wilkinson, L. (2011). ggplot2: elegant graphics for data analysis by WICKHAM, H.
12. Handcock et al.(2019). Package 'statnet'.
13. Walsh et al (2019). *Research Policy*, *48*(2), 444-461
14. Campos-Varela, et al (2021).. *Gaceta sanitaria*, *34*, 430-434.

sabujb@instem.res.in

@Sabuj_Bh

https://www.linkedin.com/in/sabuj-bhattacharyya/