

Tackling the flood of paper mill manuscripts: a journal perspective

Stefanie Heck, Anneke Birkenstock, Manu Jain Goyal, Konstantina Falida, Franca Bianchini, Christoph Plass

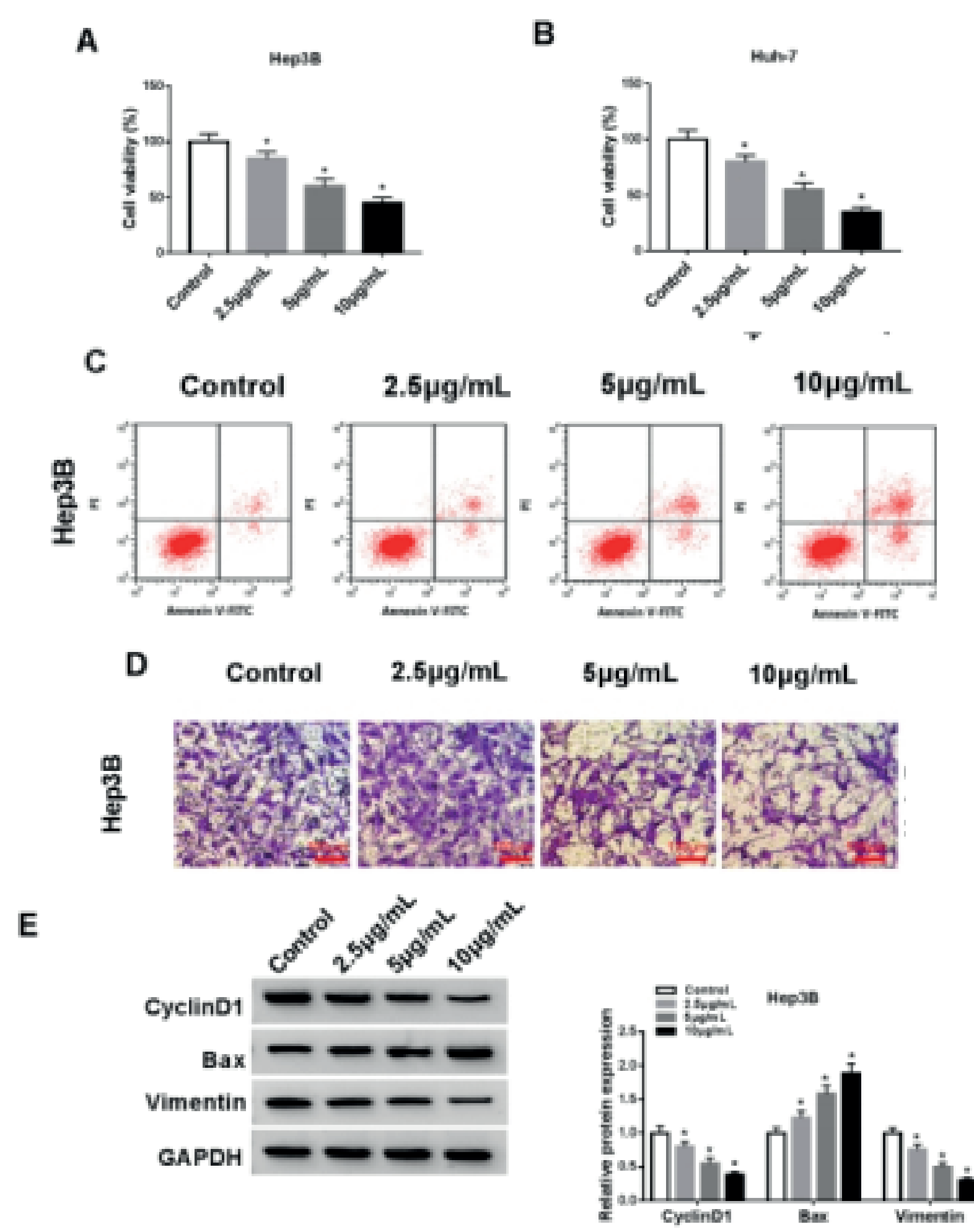


International Journal of Cancer (IJC), 69120 Heidelberg, Germany



“Typical” features of paper mill papers⁽¹⁾

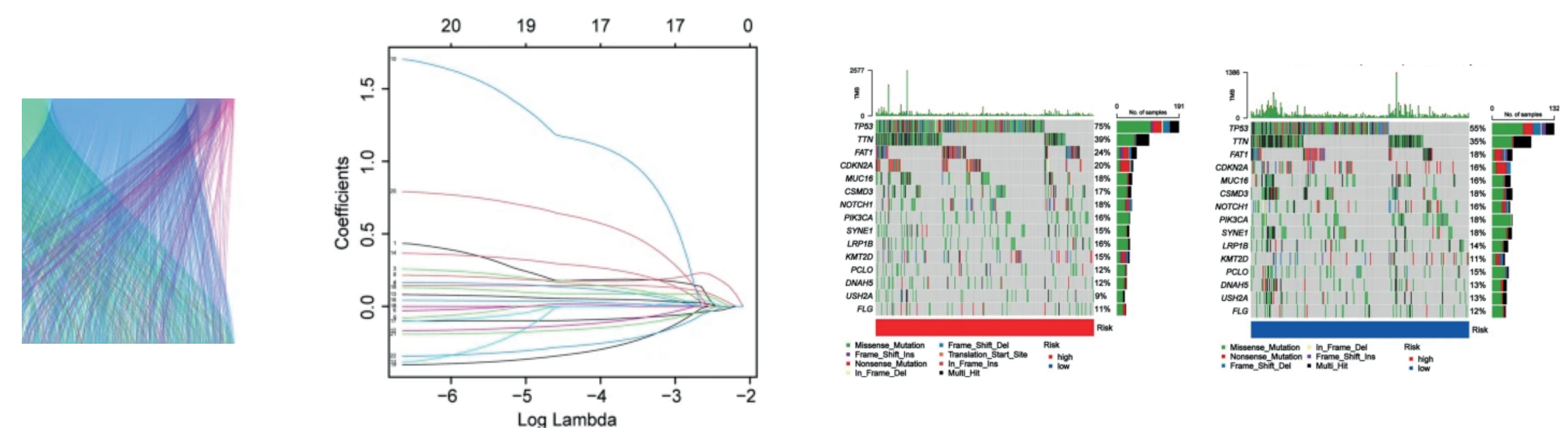
- Not scientifically sound studies
- Similarly structured titles
- Non-institutional email addresses
- Recycled text and layout similarities
- Falsified or fabricated data from “stock images”
- No valid cell line reports⁽²⁾
- Reluctance to share data



Each feature individually might not be problematic but taken together should raise concerns and could be part of a pattern

“Non-typical” features of paper mill papers

- Bioinformatics studies with similar patterns/layouts/images
- Images different from “typical” paper mill images



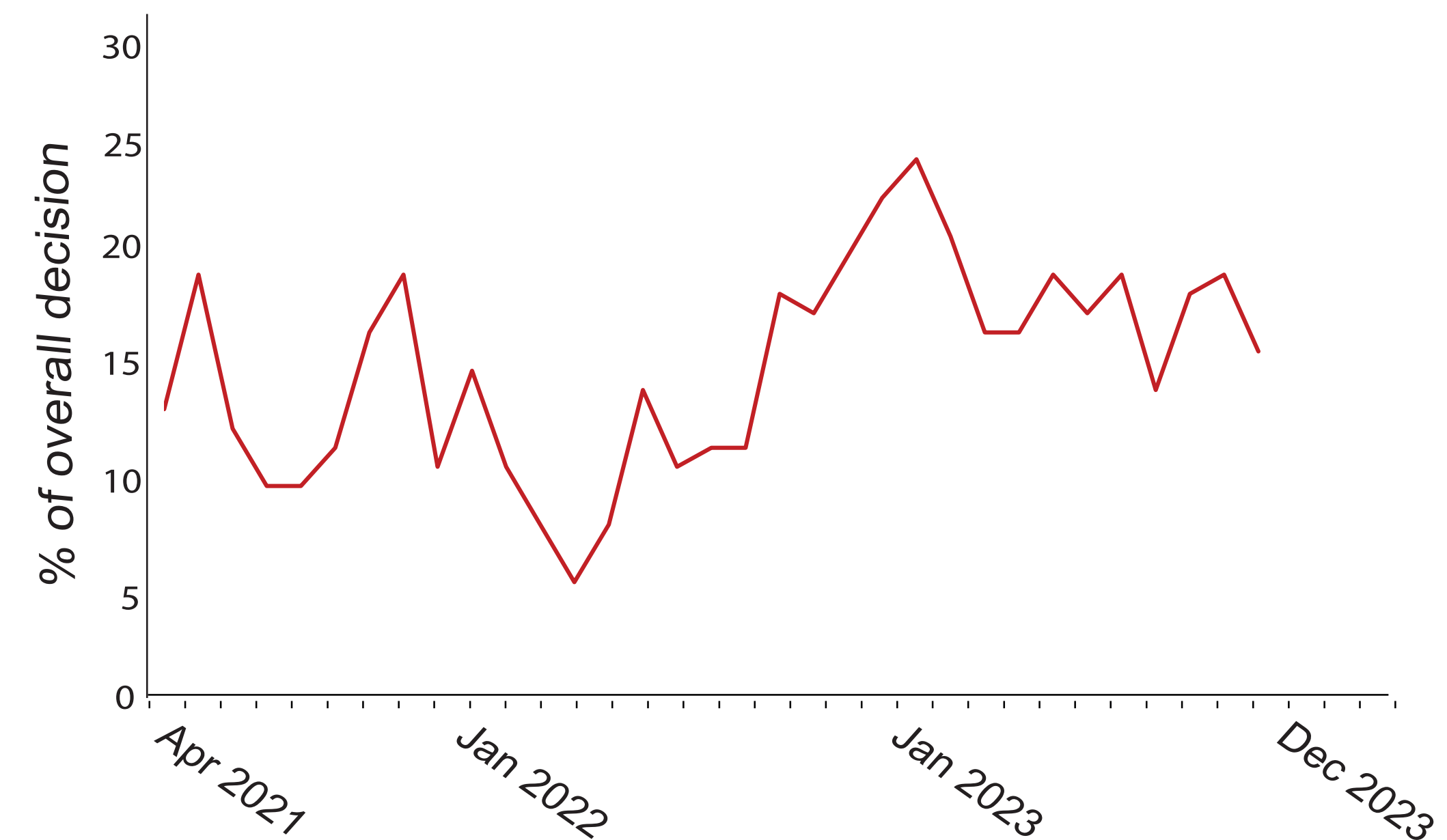
- Also epidemiological papers e.g. Meta-analyses, Mendelian randomization studies
- Sharing typical features of paper mill papers (e.g. non-institutional emails, flagged authors, text overlaps)

Quality control checks to screen out suspicious papers

- Information about authors (Flagged authors, PubPeer comments)
- Image checking and assessment of raw data
- Cell line authentication
- Plagiarism check
- Data sharing
- Sequencing coverage and quality statistics
- Large Language Models/AI (e.g. ChatGPT) usage

Suspicious papers received between 2021-2024

- High variation over a year, but overall no decrease
- 5-25% of papers rejected due to suspicious features



Since the end of 2022 papers with “non-typical” features of paper mills were also included

Identification of suspicious papers published in IJC

Search criteria:

- Limited to molecular biology content
- Suspicious figures, titles, emails, affiliations
- Authors flagged due to previous problems
- Proof of image problems using imagetwin software

Results:

Year	Published papers	Suspicious papers	Image issues identified
2014	27	7	4
2015	16	5	3
2016	15	6	3
2017	45	15	6
2018	53	19	3
2019	60	19	6
2020	18	8	0
2021	10	4	0
2022	21	3	0

Summary and follow-up

- A peak was observed in problematic manuscripts published in IJC between 2017 and 2019
- In seven cases images were reused among publications
- From 2020 screening of suspicious papers showed success
- Follow-up is ongoing, two papers have been retracted

Recommendations

- Critical monitoring of new submissions
- Careful cleaning up of the published record
- Continuous update of quality control measures to counteract rapidly adjusting paper mills

References:

- (1) Heck, Stefanie et al. “Fake data, paper mills, and their authors: The International Journal of Cancer reacts to this threat to scientific integrity.” International journal of cancer vol. 149,3 (2021): 492-493. doi:10.1002/ijc.33604
 (2) Souren, Nicole Y et al. “Cell line authentication: a necessity for reproducible biomedical research. ” The EMBO journal vol. 41,14 (2022): e111307. doi:10.15252/embj.2022111