



**Retracted papers originating from paper mills: a cross-sectional analysis.**

Journal:	<i>BMJ</i>
Manuscript ID	BMJ-2022-071517
Article Type:	Original research
Date Submitted by the Author:	21-Apr-2022
Complete List of Authors:	Candal-Pedreira, Cristina; University of Santiago de Compostela Ross, Joseph; Yale School of Medicine, Internal Medicine Ruano-Ravina, Alberto; Universidade de Santiago de Compostela, Preventive Medicine and Public Health Egilman, David; Brown University Warren Alpert Medical School, Family Medicine Fernandez, Esteve; Instituto Catalán de Oncología, Tobacco Control Unit Pérez-Rios, Mónica; Universidade de Santiago de Compostela, Preventive Medicine and Public Health
Keywords:	Ethics, medical, Epidemiology, Medicine, Public health

SCHOLARONE™  
Manuscripts

## 1 **Retracted papers originating from paper mills: a cross-sectional analysis.**

---

2  
3  
4  
5  
6  
7 **Authors** Cristina Candal-Pedreira <sup>1,2</sup>, Joseph S. Ross <sup>3,4,5</sup>, Alberto Ruano-  
8 Ravina <sup>1,2,6\*</sup> (0000-0001-9927-7453), David S. Egilman <sup>7</sup>, Esteve  
9 Fernández <sup>8,9</sup>, Mónica Pérez-Ríos <sup>1,2,6</sup>.

10  
11 **Affiliations** (1) Preventive Medicine and Public Health, University of Santiago  
12 de Compostela, Santiago de Compostela, Galicia, Spain.  
13  
14 (2) Health Research Institute of Santiago de Compostela (*Instituto de*  
15 *Investigación Sanitaria de Santiago de Compostela-IDIS*), Santiago de  
16 Compostela, Galicia, Spain.  
17  
18 (3) Section of General Internal Medicine and National Clinician  
19 Scholars Program, Yale School of Medicine, New Haven,  
20 Connecticut, USA.  
21  
22 (4) Department of Health Policy and Management, Yale University  
23 School of Public Health, New Haven, Connecticut, USA.  
24  
25 (5) Center for Outcomes Research and Evaluation, Yale-New Haven  
26 Hospital, New Haven, Connecticut, USA.  
27  
28 (6) Consortium for Biomedical Research in Epidemiology and Public  
29 Health (*CIBER en Epidemiología y Salud Pública-CIBERESP*),  
30 Madrid, Spain.  
31  
32 (7) Family Medicine Department, Alpert School of Medicine, Brown  
33 University, Providence, Rhode Island, USA.  
34  
35 (8) Tobacco Control Unit, WHO Collaborating Centre for Tobacco  
36 Control, Institut Català d'Oncologia-ICO, L'Hospitalet de  
37 Llobregat, Catalonia, Spain.  
38  
39 (9) Consortium for Biomedical Research in Respiratory Diseases  
40 (*CIBER en Enfermedades Respiratorias-CIBERES*), Madrid, Spain.  
41  
42  
43  
44  
45  
46  
47  
48  
49

50 **\*Corresponding** Professor Alberto Ruano-Ravina  
51 **Author** Preventive Medicine and Public Health, School of Medicine  
52 San Francisco Street s/n.,  
53 Universidad de Santiago de Compostela,  
54 15782 Santiago de Compostela, Spain.  
55 Tel: +34-981-581237. Fax: +34-981-572282  
56 E-mail: [alberto.ruano@usc.es](mailto:alberto.ruano@usc.es)  
57  
58  
59  
60

### 3 **Word count**

1  
2  
3 4 Abstract: 283  
4  
5

6 5 Text: 2,779  
7

8  
9 6 Tables: 4  
10

11 7 Figures: 3  
12

13  
14  
15 8 References: 21  
16  
17

18 9 **Author contribution**  
19

20  
21 10 **CCP:** Conceptualization, Methodology, Data Curation, Formal Analysis, Original draft  
22  
23 11 preparation. **ARR:** Conceptualization, Methodology, Review and Editing, Supervision.  
24  
25 12 **JSR:** Methodology, Review and Editing. **DSE:** Methodology, Review and Editing. **EF:**  
26  
27 13 Methodology, Review and Editing. **MPR:** Conceptualization, Methodology, Review  
28  
29 14 and Editing.  
30  
31

32  
33 15 The corresponding author attests that all listed authors meet authorship criteria and that  
34  
35 16 no others meeting the criteria have been omitted.  
36  
37

38 17 **Ethics approval**  
39

40  
41 18 Because this study used publicly available materials and did not involve human  
42  
43 19 subjects, human subjects' ethics committee approval was not required.  
44  
45

46 20 **Transparency declaration**  
47

48 21 The lead author affirms that this manuscript is an honest, accurate, and transparent  
49  
50 22 account of the study being reported; that no important aspects of the study have been  
51  
52 23 omitted; and that any discrepancies from the study as planned (and, if relevant,  
53  
54 24 registered) have been explained.  
55  
56

57 25 **Financial Disclosure statement**  
58  
59  
60

1  
2  
3 26 This work is part of the research conducting to the PhD degree of Cristina Candal-  
4  
5 27 Pedreira, who has received a PFIS fellowship (reference number FI21/00149) from the  
6  
7 28 Health Institute Carlos III (ISCIII). The funders had no role in study design, data  
8  
9  
10 29 collection and analysis, decision to publish, or preparation of the manuscript.  
11  
12 30 **Potential Competing interests:** All authors have completed the ICMJE uniform  
13  
14 31 disclosure form at <http://www.icmje.org/disclosure-of-interest/> and declare:  
15  
16 32 Dr. Ross currently receives research support through Yale University from Johnson and  
17  
18 33 Johnson to develop methods of clinical trial data sharing, from the Medical Device  
19  
20 34 Innovation Consortium as part of the National Evaluation System for Health  
21  
22 35 Technology (NEST), from the Food and Drug Administration for the Yale-Mayo Clinic  
23  
24 36 Center for Excellence in Regulatory Science and Innovation (CERSI) program  
25  
26 37 (U01FD005938), from the Agency for Healthcare Research and Quality  
27  
28 38 (R01HS022882), from the National Heart, Lung and Blood Institute of the National  
29  
30 39 Institutes of Health (NIH) (R01HS025164, R01HL144644), and from the Laura and  
31  
32 40 John Arnold Foundation to establish the Good Pharma Scorecard at Bioethics  
33  
34 41 International; in addition, Dr. Ross is an expert witness at the request of Relator's  
35  
36 42 attorneys, the Greene Law Firm, in a qui tam suit alleging violations of the False Claims  
37  
38 43 Act and Anti-Kickback Statute against Biogen Inc. Dr. Egilman serves as an expert  
39  
40 44 witness in asbestos, talc, opioid and concussion litigation at the request of injured  
41  
42 45 persons. No other authors declared any potential competing interests.  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

47 **Summary box****What is already known on this topic**

Evidence regarding paper mills organizations and articles produced by them is scarce as it is a new phenomenon. It is necessary to generate information about the characteristics of paper-mill articles in order to identify and retract them, thus allowing the scientific literature to be corrected.

**What this study adds**

To our knowledge, this is the first study analysing the evolution of paper-mill papers, their characteristics and their visibility in the scientific community. We have observed that retractions of paper-mill papers are increasing in frequency and that some of them are highly cited papers, with the potential consequences that this entails.

48

## 49 **Retracted papers originating from paper mills: a cross-sectional analysis.**

### 50 **Abstract**

51 **Introduction:** “Paper mills” are for-profit organisations which engage in large-scale  
52 production of manuscripts that are subsequently sold to researchers. This study aimed to  
53 describe retracted papers originating from paper mills, including their characteristics,  
54 visibility and impact, and the journals in which they were published, along with trends  
55 in retractions over time.

56 **Methods:** In this cross-sectional study, we identified all retracted paper-mill papers  
57 from 1/1/2013 to 27/9/2021 from the Retraction Watch database. We collected  
58 information relating to the characteristics of the journals involved, as well as of the  
59 papers retracted. We used descriptive statistics to characterize the sample and analyse  
60 the trend of retracted paper-mill papers over time.

61 **Results:** We identified 622 retracted paper-mill papers. Retracted paper-mill papers  
62 rose from 0.04 to 5.6 per 100,000 publications from 2013 to 2019. The first paper mill  
63 retraction was in 2017; by 2021, paper mill retractions accounted for 469 (19.8%) of  
64 2,374 total article retractions. Overall, retracted paper-mill papers were mostly  
65 published in journals of the second JCR quartile (51.9%) and listed 4 to 6 authors  
66 (56.3%). The first authors of all 622 (100%) paper-mill retractions came from Chinese  
67 institutions and 515 (82.8%) listed a hospital as primary affiliation. Fifteen journals  
68 accounted for 487 (78.3%) of paper-mill retractions, with a single journal accounting for  
69 132 (21.2%). Nearly all (n=604, 97.1%) paper mill retractions had received at least 1  
70 citation, with a median of 10 (interquartile range: 5-20) citations received.

71 **Conclusions:** Papers retracted for originating from paper mills are increasing in  
72 frequency, posing a problem for the research enterprise. Retracted paper-mill papers

1  
2  
3 73 most commonly originate from China and are published in a small number of journals.  
4

5 74 New mechanisms are needed to detect and avoid this new type of misconduct.  
6

7 75 **Key words:** paper mills, retraction, research integrity, scientific misconduct, publication  
8

9  
10 76 ethics.  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

Confidential: For Review Only

## 77 **Introduction**

78 Scientific misconduct is the most common cause of retraction of biomedical papers  
79 (1,2), and it includes plagiarism, fabrication and falsification of data and/or images.  
80 Fraudulent papers have negative consequences for the scientific community and the  
81 general public, engendering a lack of trust in science, false claims of drug or device  
82 efficacy and unjustified academic promotion, among other problems. Moreover,  
83 misconduct encompasses other unethical practices, often difficult to detect, such as  
84 undeclared conflicts of interest, authorship issues, and duplicated publication (3).  
85 As scientific papers evolve and modernise, new types of misconduct and fraud emerge.  
86 One example of this is the so-called “paper mill” phenomenon. In the sphere of  
87 scientific publication, the term “paper mill” refers to for-profit organisations which  
88 engage in the large-scale production and sale of papers to researchers, academics and  
89 students who wish to or have to publish in peer reviewed journals, both national and  
90 international. It has been observed that many paper-mill papers included fabricated data  
91 (4). We refer to this process as Ghost Fabrication to distinguish it from ghost writing.  
92 According to the Committee on Publication Ethics (COPE), these organisations prepare  
93 manuscripts and engage in their sale. In some cases, they sell the authorship prior to  
94 publication, they then handle the submission and the peer review process. In others,  
95 these organizations sell the authorship after the manuscript has been accepted for  
96 publication in a legitimate scientific journal. In the latter scenario, the organization  
97 proceeds to include the author/authors who bought the authorship on the list of named  
98 authors, which amounts to a (sometimes total) change in authorship (4). In addition to  
99 selling the authorship of scientific papers, these organisations offer other services,  
100 ranging from making available or fabricating a database on which a study can be based,  
101 to falsifying a journal peer review so as to enable a paper to be published more easily



1  
2  
3 102 (5). Recently, paper mills have broadened their service portfolio, by offering citations to  
4  
5 103 papers already published by researchers on their own studies (6). Some of these  
6  
7 104 organisations claim to have links with scientific journals, thereby ensuring publication  
8  
9 105 of the manufactured manuscript (7,8).

10  
11  
12 106 Paper-mill papers are a growing problem with important potential consequences, since  
13  
14 107 they amount to systematic manipulation of the scientific publication process, as well as  
15  
16 108 dissemination of false results. In addition, they artificially inflate researchers'  
17  
18 109 curriculums without merit and diminish trust in the scientific enterprise. The main  
19  
20 110 problem which these types of publications pose for editors and reviewers of scientific  
21  
22 111 journals is the difficulty of identifying them through the peer-review process, since they  
23  
24 112 appear to be perfectly legitimate papers. Analysis of images included in a manuscript  
25  
26 113 has been identified as one of the possible strategies for detecting milled papers, since  
27  
28 114 the great majority of such images tend to be manipulated and/or duplicated (9). That  
29  
30 115 said, however, few papers contain images that allow for scrutiny, and currently there is  
31  
32 116 no software capable of detecting image-duplication and/or manipulation, thus leaving  
33  
34 117 this task to editors and reviewers.

35  
36  
37  
38  
39 118 This type of fraud has already given rise to various retractions, and Retraction Watch,  
40  
41 119 an organization that dates from 2010, maintains a database of retracted articles that now  
42  
43 120 includes "paper mill" as a reason for retraction (10). As a relatively novel phenomenon,  
44  
45 121 the *modus operandi* and characteristics of these paper mills are not very well known,  
46  
47 122 though Retraction Watch has recently published the results of its research into how the  
48  
49 123 best-known paper mill in Russia operates (11). Even so, little is known about what types  
50  
51 124 of authors use the services of paper mills, in what types of journals they publish, on  
52  
53 125 which fields, and whether there might be differences in terms of the prestige of the  
54  
55 126 journals in which they publish, based on their impact factor.  
56  
57  
58  
59  
60

1  
2  
3 127 Thus, our objective was to analyse the recent trend in papers retracted for originating  
4  
5 128 from paper mills; to characterise the papers retracted for this reason, along with the  
6  
7 129 journals in which they were published; and to analyse their impact and visibility by  
8  
9  
10 130 reference to the number of citations received.

## 11 12 13 131 **Methods**

### 14 15 16 132 **Study design and data-collection**

17  
18  
19 133 We conducted a cross-sectional analysis of all papers retracted for being paper-mill  
20  
21 134 papers, from 1 January 2013, the year of publication of the first paper mill identified,  
22  
23 135 until 27 September 2021. These papers were identified via the Retraction Watch  
24  
25  
26 136 database (10), using the filter “Reason for retraction” and choosing the option “Paper  
27  
28 137 mill”. We included all papers retracted for this reason and excluded those bearing an  
29  
30 138 “expression of concern”, where scientific misconduct had not been confirmed.

31  
32  
33 139 All the variables of interest were collected and stored in a purpose-designed database.

34  
35 140 To conduct this study, the following three main data sources were used: Web of  
36  
37 141 Science, Journal Citation Reports (both belonging to Clarivate Analytics), and the  
38  
39 142 Retraction Watch database.

#### 40 41 42 143 **Retraction Watch database**

43  
44  
45 144 Retraction Watch tracks scientific publications that have been retracted and aggregates  
46  
47 145 them in a publicly available database, including different variables of interest extracted  
48  
49 146 by their staff. This database includes more than 30,000 retractions and expressions of  
50  
51 147 concern. The Retraction Watch database is publicly available in  
52  
53  
54 148 [www.retractiondatabase.org](http://www.retractiondatabase.org) (10).

55  
56 149 We sourced the total number of papers retracted for any reason per year and the total  
57  
58 150 number retracted for originating from paper mills per year. For every paper retracted for  
59  
60

1  
2  
3 151 being a paper-mill paper, the following data were collected: title of paper; number of  
4  
5 152 authors; first author's country; first author's institution; type of institution of first author  
6  
7 153 (hospital, university, research centre); and paper's date of publication and date of  
8  
9 154 retraction.

### 12 155 **Web of Science (WOS)**

15 156 We retrieved the total number of papers published per year across the study period. For  
16  
17 157 every paper included, total citations received from date of publication until 15 October  
18  
19 158 2021 were collected.

### 22 159 **Journal Citation Reports (JCR)**

24 160 We gathered data on the journal each paper was published and its characteristics such as  
25  
26 161 its name, JCR impact factor, JCR category, relative position (JCR quartile) and  
27  
28 162 publication modality (Open Access or not). In the case where the journal was included  
29  
30 163 in more than one category, we chose the most favourable according to the Journal  
31  
32 164 Impact Factor.

35 165 We consulted the full text of the papers included to record information related to the  
36  
37 166 characteristics of the paper, such as the date of submission and publication, authors'  
38  
39 167 statement of funding and conflicts of interest.

### 43 168 **Statistical analysis**

45 169 We performed a descriptive analysis of the characteristics of the retracted papers  
46  
47 170 included, by reference to the variables of interest, with the continuous variables being  
48  
49 171 expressed as median and interquartile range, and the categorical variables as absolute  
50  
51 172 and relative frequency.

53 173 First, we analysed the trend in paper-mill papers. We then calculated rate per 100,000,  
54  
55 174 of paper-mill papers published over the total number of papers published for each year  
56  
57 175 of the study period. Additionally, we calculated the percentage of papers retracted that

1  
2  
3 176 were manufactured in paper mills per year, with respect to the total number of  
4  
5 177 retractions per year, in order to ascertain the percentage which paper-mill retractions  
6  
7 178 represent with respect to retractions for any other reason.  
8  
9  
10 179 We described the distribution of these types of papers by JCR category of the journal in  
11  
12 180 which they were published. We created a ranking of journals and publishers based on  
13  
14 181 the number of retracted paper-mill papers they published during the study period.  
15  
16 182 We calculated the time elapsed between the paper's submission and publication and the  
17  
18 183 time elapsed between the paper's publication and retraction, in days. Analysis of the  
19  
20 184 times elapsed between submission and publication, and between publication and  
21  
22 185 retraction, were stratified by quartile of the journal in which the paper was published.  
23  
24 186 Similarly, the analysis of the citations received by the papers included is also shown,  
25  
26 187 both overall and stratified by quartile.  
27  
28 188 All statistical analyses were performed using the Stata v.17.0 computer software  
29  
30 189 programme.  
31  
32  
33  
34

### 35 **Ethical aspects**

36  
37  
38 191 Because this study used publicly available materials and did not involve human  
39  
40 192 subjects, human subjects' ethics committee approval was not required.  
41  
42

### 43 **Results**

44  
45  
46 194 We identified 622 retractions of paper-mill papers from the Retraction Watch database  
47  
48 195 that fulfilled the pre-defined inclusion criteria. Figure 1 shows the number of paper-mill  
49  
50 196 papers published with respect to the total number of papers published per year. The first  
51  
52 197 identified paper-mill paper was published in 2013, and the first retraction took place in  
53  
54 198 2017.  
55  
56

57  
58 199 *Figure 1. Percentage, in rate per 100,000, of paper-mill papers published per year with respect to total publications.*  
59  
60

200 Since 2013, the proportion of published paper-mill papers in the scientific literature has  
 201 increased, from 0.04 per 100,000 in 2013 to its peak of 5.6 per 100,000 in 2019. From  
 202 2020 onwards, the number of these papers decreased in comparison with the total  
 203 number of papers published. The percentage of paper mill retractions to all-cause  
 204 retractions was low until 2021, the year in which paper mill retractions accounted for  
 205 469 (19.8%) of the 2,374 retractions (Fig 2).

206 *Figure 2. Percentage of paper-mill retractions with respect to total retractions.*

207 Table 1 shows the main characteristics of retracted paper-mill papers. Over half of these  
 208 papers had 4 to 6 authors (56.3%); the first authors of all 622 (100%) paper mill  
 209 retractions came from Chinese institutions; and in 515 (82.8%) papers the first author  
 210 was affiliated with a hospital. The papers were mainly published in journals of the  
 211 second JCR quartile (51.9%) that were assigned to the JCR category of Biochemistry and  
 212 Molecular Biology (22.5%).

213 *Table 1. Main characteristics of papers retracted for originating from paper mills.*

<b>Variable</b>	<b>n (%)</b>
<b>Number of authors</b>	
1-3	124 (19.9%)
4-6	350 (56.3%)
>6	148 (23.8%)
<b>Author's country</b>	
China	616 (99.2%)
China, Germany	1 (0.2%)
China, Canada	1 (0.2%)
China, USA.	3 (0.5%)
<b>First author's affiliation</b>	
Hospital	515 (82.8%)
Hospital and university	54 (8.7%)
University	41 (6.6%)
Other	12 (1.9%)
<b>JCR quartile of journal in which paper was published</b>	
Q1	93 (14.9%)
Q2	323 (51.9%)

Q3	150 (24.1%)
Q4	17 (2.7%)
No IF	40 (6.4%)
<b>JCR category</b>	
Biochemistry and molecular biology	135 (21.7%)
Pharmacology and pharmacy	103 (16.5%)
Oncology	77 (12.4%)
Chemistry	68 (10.9%)
Other	215 (34.7%)
Not indexed in JCR	24 (3.9%)

214 Q: quartile; JCR: Journal Citation Reports; IF: impact factor

215 Of the 622 papers, 298 (47.9%) included a funding statement, and of these, 195 (65.4%)  
 216 reported to have received external funding. Furthermore, 497 (79.9%) of papers  
 217 included a declaration of the authors' conflicts of interests.

218 Fifteen scientific journals published a total of 487 (78.3%) of all papers retracted for  
 219 being paper-mill papers, and 132 (21.2%) were published in a single journal, the Journal  
 220 of Cellular Biochemistry. Most of the journals in which these types of papers were  
 221 published were Open Access journals (Table 2): 175 (28.1%) papers were published in  
 222 journals belonging to the Wiley publishing group, and 88 (14.2%) in those belonging to  
 223 Spandidos Publications (Table 3).

224 *Table 2. Journals in which papers retracted for originating from paper mills were published, according to whether or*  
 225 *not they were Open Access.*

Journal	Quartile	Open access	Number of papers retracted for originating from paper mills (%)
Journal of Cellular Biochemistry	2	No	132 (21.2%)
RSC Advances	2	Yes	68 (10.3%)
European Review for Medical and Pharmacy	2	No	59 (9.5%)
Journal of Cellular Physiology	1	No	32 (5.1%)
Artificial Cells, Nanomedicine, and Biotechnology	1	Yes	31 (5.0%)

Molecular Medicine Reports	3	No	28 (4.5%)
Bioscience Reports	3	Yes	21 (3.4%)
Oncology Reports	3	No	20 (3.2%)
Experimental and Therapeutic Medicine	3	Yes	16 (2.6%)
International Journal of Immunopathology and Pharmacology	2	Yes	16 (2.6%)
Oncology Letters	4	Yes	15 (2.4%)
Cellular Physiology and Biochemistry	No IF	Yes	14 (2.3%)
<b>Other journals</b>			170 (27,3%)

226

227 *Table 3. Publishing houses of the journals in which papers retracted for originating from paper mills were published.*

<b>Publishing house</b>	<b>Number of papers retracted for originating from paper mills (%)</b>
Wiley	175 (28.1%)
Spandidos	88 (14.2%)
Royal Society of Chemistry (RSC)	70 (11.3%)
Verduci Editore	59 (9.5%)
Taylor and Francis	38 (6.1%)
SAGE Publications	31 (5.0%)
Elsevier	22 (3.5%)
Taylor and Francis - Dove Press	22 (3.5%)
Portland Press	21 (3.4%)
Springer	17 (2.7%)
<b>Other publishing houses</b>	79 (12,7%)

228 The time elapsed between the manuscript's submission to the journal and its publication

229 varied according to journal quartile (Table 4), from a median of 115 days (IQR, 26-

230 728), 128 (IQR, 0-724), and 163 (IQR, 14-1,943) among journals of the first, second

231 and third quartiles, respectively, and a median of 358 (IQR, 96-596) days in fourth

232 quartile journals. Likewise, the time between publication and retraction varied; it was

233 shorter in journals of the first and second quartiles, and longer in journals of the third

234 and fourth quartiles.

235 *Table 4. Times of publication and retraction of papers retracted for originating from paper mills, both overall and by*  
236 *quartile of journal in which they were published.*

#### **Time elapsed between submission and publication (in days)\***

	Overall	Q1	Q2	Q3	Q4	IF no.
<i>Median</i>	140	115	128	163	358	219
<i>(range)</i>	(0-3,876)	(26-728)	(0-724)	(14-1,943)	(96-596)	(87-3,876)

#### Time elapsed between publication and retraction (in days)

	Overall	Q1	Q2	Q3	Q4	IF no.
<i>Median</i>	898.5	706	897	1,291.5	1,767	887
<i>(range)</i>	(0-2,930)	(0-2,183)	(0-2,567)	(46-2,930)	(676-2,411)	(0-1,888)

\*Missing values: 99. IF: Impact factor. Q: quartile

237 While 604 (97.1%) of retracted paper-mill papers received at least one citation, papers  
 238 published in fourth quartile journals received a higher number of citations, as shown in  
 239 Figure 3. The median number of citations received by retracted paper-mill papers from  
 240 the date of publication was 10 (interquartile range: 5-20), with the total ranging from 0  
 241 to 110 citations.

242 *Figure 3. Citations received by papers retracted for originating from paper mills, by quartile of journal in which they*  
 243 *were published.*

## 244 Discussion

245 This cross-sectional analysis of all papers retracted for originating from paper mills  
 246 before September 2021 identified from the Retraction Watch database suggests that  
 247 these paper mill retractions are increasing in frequency. Nearly all authors of these  
 248 papers came from China and were predominantly affiliated with hospitals. The median  
 249 time for retraction of a paper-mill paper was close to two years, and increased with the  
 250 ranking of the journal in which it was published, so that the higher the quartile, the  
 251 longer the period until retraction. Furthermore, this study illustrated the impact and  
 252 visibility of these retracted papers, as some were highly cited, with the potential  
 253 consequences that this entails. To our knowledge, this is the first study to analyse  
 254 growing phenomenon of paper mill retractions and their characteristics  
 255 Our findings suggest that paper-mill papers increased between 2017 and 2019, when 4  
 256 to 6 that were eventually retracted for this reason were published per 100,000



1  
2  
3 257 publications. In 2020, the number of retracted paper-mill papers published in the  
4  
5 258 scientific literature fell sharply. This decrease may have occurred for a number of  
6  
7 259 reasons. First, it is likely that papers published in 2020 that may eventually be identified  
8  
9 260 for retraction have simply not yet been discovered or were discovered after our database  
10  
11 261 search at the end of September 2021. Second, as a result of investigations initiated in  
12  
13 262 early 2020 by a number of editors and researchers (12), the scientific community  
14  
15 263 became aware of the problem and guidelines were published to help editors identify  
16  
17 264 such papers (4). Even though these guidelines do not enable a paper-mill paper to be  
18  
19 265 unequivocally identified, they do make it possible to screen papers and identify those  
20  
21 266 originating from paper mills. Hence, numbers may have dropped because scientific  
22  
23 267 journals might improve methods for their identification during editorial review and peer  
24  
25 268 review, thereby preventing their publication. Third, the increased attention being paid to  
26  
27 269 this type of fraud may also have deterred authors from engaging the services of paper  
28  
29 270 mills, in light of the consequences of scientific fraud, specially in some countries such  
30  
31 271 as China (13). Then again, there it is possible that increased exposure has caused paper-  
32  
33 272 mill papers to change their mode of operation, thus hindering detection (9).  
34  
35 273 Although this phenomenon is relatively new, particularly in the West, it should be borne  
36  
37 274 in mind that for some years now the use of these types of organisations has been  
38  
39 275 widespread in other countries, such as China (11,14). China encouraged its researchers  
40  
41 276 to publish papers in return for money and promotion (15). Furthermore, medical  
42  
43 277 students at Chinese universities are required to produce a scientific paper in order to  
44  
45 278 graduate (14). In fact, these organisations openly advertise their services on the Internet  
46  
47 279 and by maintaining a presence on university campuses not only in China but also in  
48  
49 280 other countries, such as Russia (8,14).  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

1  
2  
3 281 Perhaps it is not surprising that the majority of articles retracted for being paper-mill  
4  
5 282 papers come from that same country. These results are in line with the findings of other  
6  
7 283 researchers and editors of scientific journals, though there have also been reports of  
8  
9 284 paper-mill papers in other countries, such as Iran or Russia (8,12,16). Indeed,  
10  
11 285 revelations surfaced recently about the activity of the largest paper-mill organisation in  
12  
13 286 Russia, “International Publisher” LLC (8,11). While it is estimated that this paper mill  
14  
15 287 has published approximately 1,000 papers, its own website announces that more than  
16  
17 288 5,000 authors have bought the co-authorship of at least one paper (8).  
18  
19 289 With the aim of preventing and detecting scientific misconduct, some countries already  
20  
21 290 have offices and specific bodies that address aspects relating to scientific integrity, but  
22  
23 291 many others do not have structures of this type (17). Fanelli et al (18) concluded that  
24  
25 292 countries that have no body or policies governing scientific misconduct incur a higher  
26  
27 293 risk of producing fraudulent papers. Countries such as Denmark, Sweden and China  
28  
29 294 have passed laws against scientific fraud. Ironically China has the most severe penalties  
30  
31 295 for research fraud. The lack of consequences that scientific misconduct has historically  
32  
33 296 had in this country may have played an important role in the increase in unethical  
34  
35 297 behaviour, including the use of paper mills (14). In 2018, following a number of  
36  
37 298 scandals in China, the law against scientific fraud was strengthened by imposing  
38  
39 299 sanctions that go beyond the purely academic and occupational sphere (19). This  
40  
41 300 tougher approach appears to have started yielding results, and in December 2021 more  
42  
43 301 than 300 researchers were reportedly penalised for scientific misconduct. Among other  
44  
45 302 things, the penalties included revocation of academic degrees and cancellation of  
46  
47 303 promotions (20). Since practically all paper-mill papers come from China, these recent  
48  
49 304 penalties policy may have played a role in the reduction in the number of this papers  
50  
51  
52  
53  
54  
55  
56  
57  
58 305 since 2020 onwards.  
59  
60

1  
2  
3 306 This study has limitations. First, retractions of paper-mill papers continue over time.  
4  
5 307 Because of this, our investigation will need to be updated over time as the conclusions  
6  
7 308 could well vary as the list of retractions grows. The main strength of this study is the  
8  
9  
10 309 use of the Retraction Watch database to identify retracted paper-mill papers, since it is  
11  
12 310 the principal and gold-standard source for aggregated information on retracted articles.  
13  
14 311 In conclusion, paper-mill papers identified and retracted to date likely represent only the  
15  
16 312 tip of the iceberg, as there are probably thousands of these papers that have been  
17  
18 313 published in the scientific literature not yet identified nor retracted. Currently, some  
19  
20 314 editors of international scientific journals began to systematically identify and retract  
21  
22 315 paper-mill papers which gave rise to mass retractions (21,22). The rise of paper mills is  
23  
24 316 a new ethical problem in research and, more specifically, in publication ethics. Not only  
25  
26 317 does it entail the sale of authorship, but these types of papers have also been observed to  
27  
28 318 contain fabricated and manipulated data and images, thus disseminating false results in  
29  
30 319 scientific literature. Different activities must be implemented to prevent the use of these  
31  
32 320 types of organisations, beginning with improved education in ethics and scientific  
33  
34 321 integrity for editorial committees of scientific journals, students and researchers.  
35  
36  
37  
38  
39

## 40 **Acknowledgments**

41  
42  
43 323 We would like to thank Retraction Watch for making their data publicly available and  
44  
45 324 Ivan Oransky and Alison Abritis for their constructive comments on the manuscript.  
46  
47 325 Both Dr. Oransky and Ms. Abritis receiving funding from Arnold Ventures to support  
48  
49 326 their work at Retraction Watch.  
50  
51

## 52 **References**

- 53  
54  
55  
56 328 1. Campos-Varela I, Ruano-Raviña A. Misconduct as the main cause for retraction.  
57  
58 329 A descriptive study of retracted publications and their authors. Gac Sanit

- 1  
2  
3 330 [Internet]. 2019 Jul 1 [cited 2021 Sep 14];33(4):356–60. Available from:  
4  
5 331 <https://pubmed.ncbi.nlm.nih.gov/29776690/>  
6  
7  
8 332 2. Fang FC, Steen RG, Casadevall A. Misconduct accounts for the majority of  
9  
10 333 retracted scientific publications. *Proc Natl Acad Sci U S A*. 2012 Oct  
11  
12 334 16;109(42):17028–33.  
13  
14  
15 335 3. International Committee of Medical Journal Editors. Recommendations for the  
16  
17 336 conduct, reporting, editing and publication of scholarly work in medical journals.  
18  
19 337 2018 [cited 2021 Nov 2]; Available from: [www.equator-network.org](http://www.equator-network.org).  
20  
21  
22  
23 338 4. COPE Forum 4 September 2020: paper mills. COPE: Committee on Publication  
24  
25 339 Ethics [Internet]. [cited 2021 Nov 7]. Available from:  
26  
27 340 [https://publicationethics.org/resources/forum-discussions/publishing-](https://publicationethics.org/resources/forum-discussions/publishing-manipulation-paper-mills)  
28  
29 341 [manipulation-paper-mills](https://publicationethics.org/resources/forum-discussions/publishing-manipulation-paper-mills)  
30  
31  
32  
33 342 5. Rivera H, da Silva JAT. Retractions, Fake Peer Reviews, and Paper Mills. *J*  
34  
35 343 *Korean Med Sci* [Internet]. 2021 [cited 2021 Oct 22];36(24):1–7. Available from:  
36  
37 344 <https://pubmed.ncbi.nlm.nih.gov/34155837/>  
38  
39  
40  
41 345 6. Christopher J. The raw truth about paper mills. *FEBS Lett* [Internet]. 2021 Jul 27  
42  
43 346 [cited 2021 Jul 16];595(13):1751–7. Available from:  
44  
45 347 <https://onlinelibrary.wiley.com/doi/10.1002/1873-3468.14143>  
46  
47  
48 348 7. Zhang, Wen and Pan Y. The Gray Industry Chain of Article Writing Business:  
49  
50 349 Starting at 2,000 and Up to 150,000 for an SCI-indexed Article. *China News*.  
51  
52 350 2015;9–17.  
53  
54  
55  
56 351 8. Abalkina A. Publication and collaboration anomalies in academic papers  
57  
58 352 originating from a paper mill: evidence from Russia. 2021;(January 2021).  
59  
60

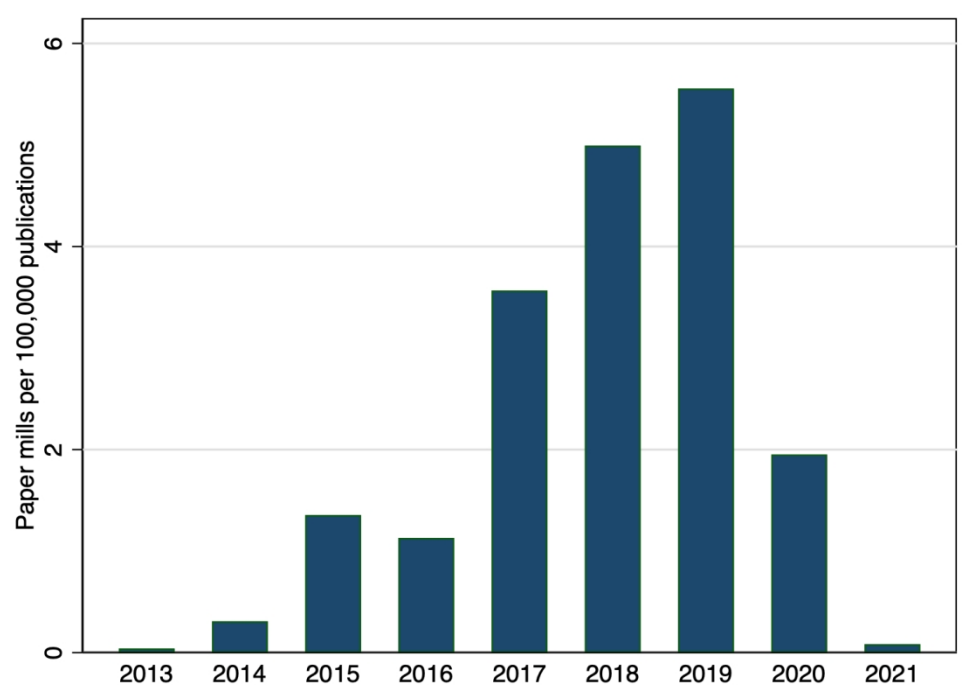
- 1  
2  
3 353 Available from: <http://arxiv.org/abs/2112.13322>  
4  
5  
6 354 9. Byrne JA, Christopher J. Digital magic, or the dark arts of the 21st century—how  
7  
8 355 can journals and peer reviewers detect manuscripts and publications from paper  
9  
10 356 mills? *FEBS Lett* [Internet]. 2020 Feb 1 [cited 2021 Oct 22];594(4):583–9.  
11  
12 357 Available from: <https://pubmed.ncbi.nlm.nih.gov/32067229/>  
13  
14  
15  
16 358 10. The Retraction Watch Database [Internet]. New York: The Center for Scientific  
17  
18 359 Integrity; 2018 [cited 2022 Apr 14]. Available from:  
19  
20 360 <https://retractionwatch.com/retraction-watch-database-user-guide/>  
21  
22  
23 361 11. Revealed: The inner workings of a paper mill – Retraction Watch [Internet].  
24  
25 362 [cited 2021 Dec 21]. Available from:  
26  
27 363 [https://retractionwatch.com/2021/12/20/revealed-the-inner-workings-of-a-paper-](https://retractionwatch.com/2021/12/20/revealed-the-inner-workings-of-a-paper-mill/)  
28  
29 364 [mill/](https://retractionwatch.com/2021/12/20/revealed-the-inner-workings-of-a-paper-mill/)  
30  
31  
32  
33 365 12. Else H, Van Noorden R. The fight against fake-paper factories that churn out  
34  
35 366 sham science. *Nature*. 2021 Mar 1;591(7851):516–9.  
36  
37  
38 367 13. Cyranoski D. China introduces ‘social’ punishments for scientific misconduct.  
39  
40 368 *Nature* [Internet]. 2018 Dec 1 [cited 2021 Dec 15];564(7736):312. Available  
41  
42 369 from: [https://click.endnote.com/viewer?doi=10.1038%2Fd41586-018-07740-](https://click.endnote.com/viewer?doi=10.1038%2Fd41586-018-07740-z&token=WzM1NTE0MjMsIjEwLjEwMzgvZDQxNTg2LTAxOC0wNzc0MC16Il0.vcvos_Aby5lPiwGn-bv7WYAWwyE)  
43  
44 370 [z&token=WzM1NTE0MjMsIjEwLjEwMzgvZDQxNTg2LTAxOC0wNzc0MC1](https://click.endnote.com/viewer?doi=10.1038%2Fd41586-018-07740-z&token=WzM1NTE0MjMsIjEwLjEwMzgvZDQxNTg2LTAxOC0wNzc0MC16Il0.vcvos_Aby5lPiwGn-bv7WYAWwyE)  
45  
46 371 [6Il0.vcvos\\_Aby5lPiwGn-bv7WYAWwyE](https://click.endnote.com/viewer?doi=10.1038%2Fd41586-018-07740-z&token=WzM1NTE0MjMsIjEwLjEwMzgvZDQxNTg2LTAxOC0wNzc0MC16Il0.vcvos_Aby5lPiwGn-bv7WYAWwyE)  
47  
48  
49  
50 372 14. Liu X, Chen X. Journal retractions some unique features of research misconduct  
51  
52 373 in China. *J Sch Publ*. 2018;49(3):305–19.  
53  
54  
55  
56 374 15. Quan W, Chen B, Shu F. Publish or impoverish: An investigation of the  
57  
58 375 monetary reward system of science in China (1999-2016). *Aslib J Inf Manag*.

- 1  
2  
3 376 2017;69(5):486–502.  
4  
5  
6 377 16. Stone R. A shady market in scientific papers mars Iran’s rise in science. *Science*  
7  
8 378 (80- ) [Internet]. 2016 Sep 14 [cited 2022 Jan 12]; Available from:  
9  
10 379 [https://www.science.org/content/article/shady-market-scientific-papers-mars-](https://www.science.org/content/article/shady-market-scientific-papers-mars-iran-s-rise-science)  
11  
12 [iran-s-rise-science](https://www.science.org/content/article/shady-market-scientific-papers-mars-iran-s-rise-science)  
13  
14  
15  
16 381 17. Candal-Pedreira C, Ruano-Ravina A, Pérez-Ríos M. Should the European Union  
17  
18 382 have an office of research integrity? [Internet]. *European Journal of Internal*  
19  
20 383 *Medicine. Eur J Intern Med*; 2021 [cited 2021 Sep 16]. Available from:  
21  
22 384 <https://pubmed.ncbi.nlm.nih.gov/34362610/>  
23  
24  
25  
26 385 18. Fanelli D, Costas R, Larivière V. Misconduct Policies, Academic Culture and  
27  
28 386 Career Stage, Not Gender or Pressures to Publish, Affect Scientific Integrity.  
29  
30 387 *PLoS One* [Internet]. 2015 Jun 17 [cited 2021 Dec 15];10(6):e0127556.  
31  
32 388 Available from:  
33  
34 389 <https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0127556>  
35  
36  
37  
38 390 19. China sets a strong example on how to address scientific fraud editorial. *Nature*  
39  
40 391 [Internet]. 2018 Jun 14 [cited 2021 Dec 15];558(7709):162. Available from:  
41  
42 392 [https://click.endnote.com/viewer?doi=10.1038%2Fd41586-018-05417-](https://click.endnote.com/viewer?doi=10.1038%2Fd41586-018-05417-1&token=WzM1NTE0MjMsIjEwLjEwMzgvZDQxNTg2LTAxOC0wNTQxNy0xIl0.fNn3uxU_RP12x7NywzIR-fRLhw4)  
43  
44 393 [1&token=WzM1NTE0MjMsIjEwLjEwMzgvZDQxNTg2LTAxOC0wNTQxNy0](https://click.endnote.com/viewer?doi=10.1038%2Fd41586-018-05417-1&token=WzM1NTE0MjMsIjEwLjEwMzgvZDQxNTg2LTAxOC0wNTQxNy0xIl0.fNn3uxU_RP12x7NywzIR-fRLhw4)  
45  
46 394 [xIl0.fNn3uxU\\_RP12x7NywzIR-fRLhw4](https://click.endnote.com/viewer?doi=10.1038%2Fd41586-018-05417-1&token=WzM1NTE0MjMsIjEwLjEwMzgvZDQxNTg2LTAxOC0wNTQxNy0xIl0.fNn3uxU_RP12x7NywzIR-fRLhw4)  
47  
48  
49  
50 395 20. Ministry of Science and Technology of the People’s Republic of China. Notice  
51  
52 396 from the Ministry of Science and Technology on the recent investigation and  
53  
54 397 handling of thesis falsification [Internet]. Available from:  
55  
56 398 [http://www.most.gov.cn/zxgz/kycxjs/kycxgzdt/202112/t20211201\\_178285.html](http://www.most.gov.cn/zxgz/kycxjs/kycxgzdt/202112/t20211201_178285.html)  
57  
58  
59  
60 399 21. Publisher retracts nearly 80 articles over three days – Retraction Watch [Internet].

- 1  
2  
3 400 [cited 2021 Dec 21]. Available from:  
4  
5 401 [https://retractionwatch.com/2021/11/14/publisher-retracts-nearly-80-articles-](https://retractionwatch.com/2021/11/14/publisher-retracts-nearly-80-articles-over-three-days/)  
6  
7 402 [over-three-days/](https://retractionwatch.com/2021/11/14/publisher-retracts-nearly-80-articles-over-three-days/)  
8  
9  
10  
11 403 22. Journal retracts 122 papers at once – Retraction Watch [Internet]. [cited 2021  
12  
13 404 Dec 21]. Available from: [https://retractionwatch.com/2021/12/15/journal-](https://retractionwatch.com/2021/12/15/journal-retracts-122-papers-at-once/#more-123768)  
14  
15 405 [retracts-122-papers-at-once/#more-123768](https://retractionwatch.com/2021/12/15/journal-retracts-122-papers-at-once/#more-123768)  
16  
17  
18 406

Confidential: For Review Only

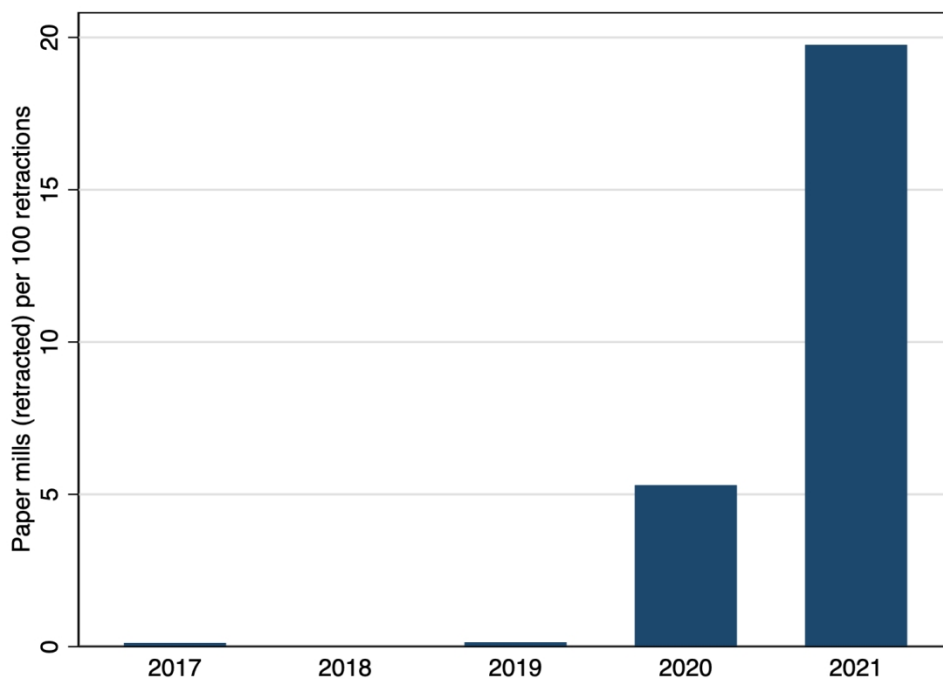
1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60



Percentage, in rate per 100,000, of paper-mill papers published per year with respect to total publications  
276x201mm (144 x 144 DPI)

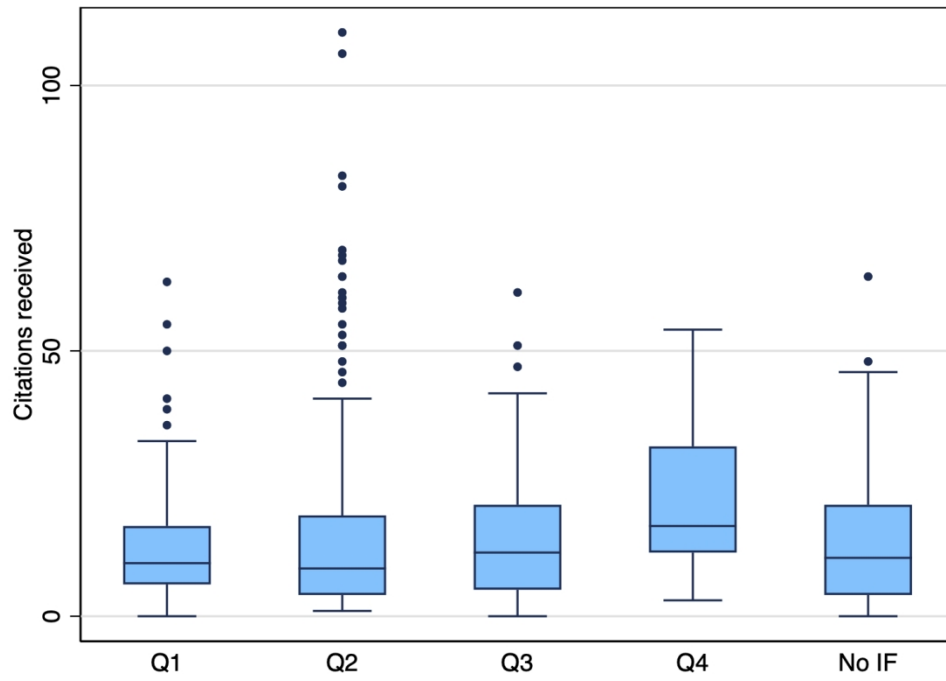


1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60



Percentage of paper-mill retractions with respect to total retractions.

276x201mm (144 x 144 DPI)



Citations received by papers retracted for originating from paper mills, by quartile of journal in which they were published.

276x201mm (144 x 144 DPI)