Assessing Predatory Journal Publishing Within Health Sciences Authors

Katie Sparks, MA, MLIS
Library Specialist, Sr., Emory University Libraries

Kimberly R. Powell, MIS
Research Impact Informationist, Woodruff Health Sciences Center Library, Emory University

Abstract

There are many resources bemoaning predatory publishing, debating the impact of the phenomenon, and presenting methods to determine predatory practices. Here, we look at the publication record within one academic health sciences institution to determine the potential local impact of the issue.

Over the 2020 fiscal year, our institution ranked in the top 20 for US National Institutes of Health (NIH) funding, receiving over $500 million. These funds represent 85% of the federal funding received overall. With these funds comes the NIH Public Access Policy which provides that all publications produced under this support be available within the PubMed Central (PMC) database. Using this convenience sample, we hypothesize institutional publications available in PMC but not MEDLINE are either out of scope despite their NIH funding, or appear in journals not meeting the rigorous inclusion criteria, including potentially predatory publishers. To test this, we created a sub-list of articles from the last 5 years that are within scope but not in MEDLINE, examined this list for publishers that may be of particular concern for predatory practices. These records were broken down by department to investigate targeting of specific fields or patterns within departments. Of the 23,743 articles assessed, only 109 were determined to be from potentially predatory publications.

This research will allow librarians to better understand institutional publishing trends and direct educational efforts towards those who may be most susceptible. This will, ideally, lead to more effective use of research funding and higher quality publications.

Introduction

Blacklists, whitelists, lists of criteria, and warnings about publishing in "predatory" journals abound in the library world (Berisha Qehaja 2020; Cook 2017; Cortegiani 2018; Cukier 2020; Dony 2020; Goodman 2018; Koerber 2020; Strielkowski 2018; Teixeira da Silva 2018; Tsigaris 2021). These references are just a fraction of the recent discourse. The current consensus seems to be that blacklists are unnecessarily restrictive, whitelists can be imperfect, and it pays to investigate any potential publishers. There are, however, varying criteria for determining whether a journal is predatory or not, and following multiple sets of guidelines can be challenging.

The issue came to our particular attention following the 2017 Statement on Article Publication Resulting from NIH Funded Research which aimed to "protect the credibility of published research" by encouraging authors to focus publication in "reputable journals" (NIH 2017). Our researchers generate several thousand publications a year as a result of US National Institutes of Health (NIH)-funded research. For the 2021 fiscal year, institutional researchers received just under $500 million in extramural funding from the NIH, ranking among the top recipients for NIH dollars in the country. As the 2017 Statement was in direct response to concerns for diminished publication quality related to predatory and other misleading or suspicious practices, there was interest in determining the potential scale of the impact among our institutional publications.

Conveniently, all publications produced from NIH-funded research efforts are required to be made available in the open access database PubMed Central (PMC), as stipulated in the 2008 NIH Public Access Policy (NIH 2008). Since 2008, articles available in PMC are provided alongside those from the flagship MEDLINE database in PubMed search results. Curated by the US National Library of Medicine, the MEDLINE database is an extensive collection of vetted biomedical journals and inclusion within MEDLINE is often seen as a hallmark for publication integrity and quality. Since the search engine and user interface known as PubMed presents results from both databases indiscriminately, articles appearing in low-quality, untrustworthy, or predatory journals can dilute the integrity of the presented scientific literature.

Across the health sciences, there has been spirited debate about whether predatory journals are a significant enough problem to be of widespread concern (Amaral 2018; Cook 2017; Eriksson 2018; Singh Chawla 2020). There is also the potential concern that a publication is not purposely predatory, it is just low-quality and therefore untrustworthy. As information professionals, we refer to all journals and publishers we would not recommend for faculty as untrustworthy because they can be impossible to distinguish from intentionally predatory.

As previously mentioned, our institution produces several thousand NIH-funded publications each year, each of which are required to appear in PMC. These publications primarily reflect the research efforts of those within our university health sciences center, which includes Schools of Medicine, Nursing, and Public Health as well as an affiliated Primate Center. By targeting institutional publications that appear in PMC, but not in journals indexed by MEDLINE, we sought to determine the potential scale of concern regarding untrustworthy or predatory publishing practices among our health science faculty. We expect the issue is not numerically significant among our institutional researchers. However, our goal is to uncover the exact extent and any existing trends that could be effectively addressed by targeted outreach to departments and individuals.

Methods

INITIAL DATA ANALYSIS

Using an internal database of publications by current faculty, we retrieved a list of 40,654 unique publications in PMC published between January 1, 2017, and December 31, 2021. The unique PMC identifiers (PMIDs) were then queried via the Entrez Direct API using the following code to request the associated Article Title, Journal Title, Journal ISSN, and MEDLINE inclusion status, where the "All-pmid.txt" file was a .txt document listing each PMID and the "PMID2017-2021.csv" file was the resulting query data.

```
efetch -input "All-pmid.txt" -db pubmed -format xml |
xtract -pattern PubmedArticle -element MedlineCitation/PMID -element ArticleTitle
-element Title -element ISSN -element MedlineCitation@Status >PMID2017-2021.csv
```

Of the results, 23,743 (58%) were properly formatted to provide for analysis without manual clean-up of missing data or erroneous delimiters. Of this subset, 2,800 articles (12%) from 959 journals were listed in PMC but not indexed in MEDLINE.

To narrow the results further, we mapped this subset to the 2020 Journal Citation Reports (JCR) journal list by ISSN, using this list as a proxy for reputable journals. Of the API-returned subset, 1,166 articles (5%) were available in PMC, but not indexed in either MEDLINE or the 2020 JCR. This set of 1,166 articles represented 429 unique journals. At this point articles classified as conference abstracts and non-English language were also removed, leaving a remaining list of 1,134 articles from 411 journals. This remaining list of journals was searched in the Directory of Open Access Journals (DOAJ) as another proxy for trustworthy publications. The final list for manual examination consisted of 467 articles from 173 journals (see Figure 1).

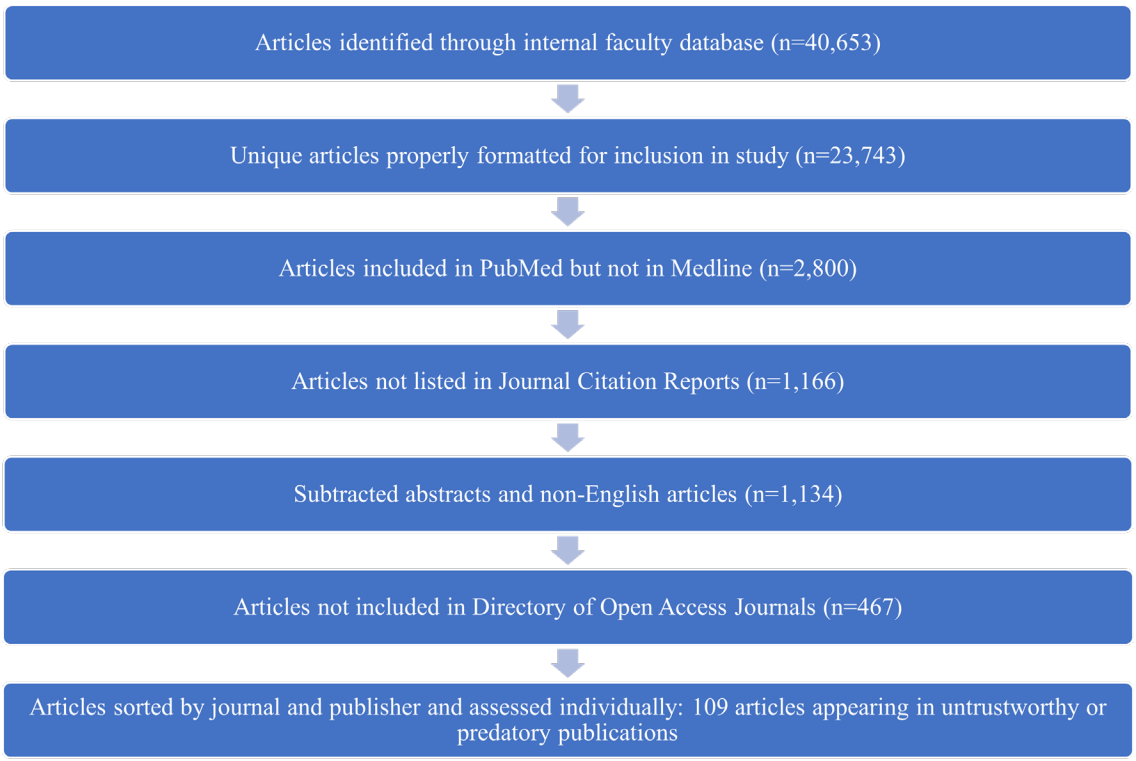

Figure 1. Visualization of search methodology

QUALITY REVIEW

The set of 173 journals that resulted from the initial analysis were assessed individually. Those from major and well-established publishers (including Elsevier, Springer, Wiley, SAGE, Thieme, etc.) were excluded from the list for further investigation, though this is not always a clear guarantee of publishing integrity (Bohannon 2013; Brembs 2018). This left 56 unique journals from 40 publishers. We then reviewed these journals and/or publisher websites to assess quality and determine trustworthiness.

Quality decisions were based on criteria provided in the literature (Cortegiani 2018; Gonzalez 2018; Soteropulos 2021), as well as the publishing resource Think Check Submit as recommended under the NIH Statement guidance. Common points of concern were:

- Poor website quality and misleading claims about indexing and impact metrics
- Lack of transparency regarding peer-review practice expectations
- Lack of statements affirming adherence to common ethical standards such as those provided by the Committee on Publication Ethics (COPE) or the International Committee of Medical Journal Editors (ICMJE)
- Charges for removal of an article from consideration or for unsolicited copy editing
- Promises of rapid (within days to weeks rather than months) or guaranteed publication.

Under these criteria we found 19 publishers representing 34 journals with 109 institutional articles that we would classify as either untrustworthy or predatory (0.46% of the total number of

assessed articles). This list represents the most conservative estimate of potentially predatory publications we consider reasonable.

Returning to our internal database of faculty publication tracking, the 109 articles were then mapped to department/school affiliations and numerical author identifiers. Articles were found to contain 26 department/school authorships from 154 unique faculty contributors. For the purposes of this assessment, authors with multiple appointments and affiliations were attributed to a single primary designation as provided within our internal dataset.

Results

A very small overall percentage of articles assessed were deemed predatory or untrustworthy (0.46%). This included 109 articles from 34 journals, from 19 publishers. In total, 154 unique authors contributed to these publications, representing 26 Health Sciences schools or departments. No individual author published more than four of the articles in this list, and only five authors published three or more articles in untrustworthy/predatory journals. There was a trend by department – five departments or schools account for 50% of the untrustworthy or predatory publications in this study – most notably our School of Medicine Department of Hematology & Medical Oncology and our School of Medicine Department of Radiology and Imaging Sciences (figure 2). Also of note, the two controversial journals *Oncotarget* and *Cureus* accounted for over 50% of institutional publications deemed of possible concern.

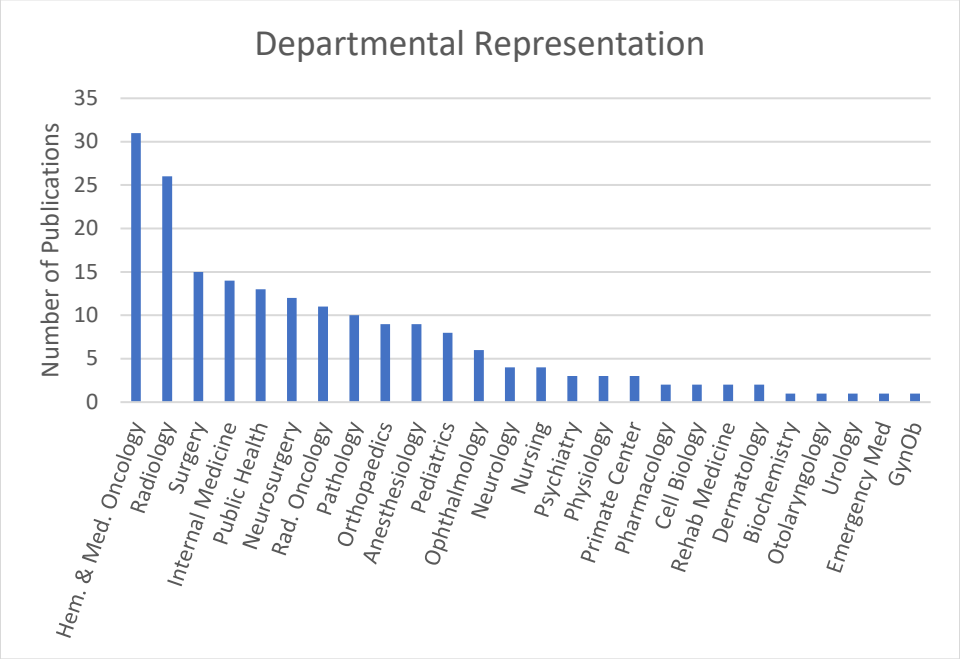

Figure 2. Articles in untrustworthy or predatory journals by department

Discussion

As expected, we found a remarkably low number of publications in potentially predatory journals, only 109 out of 23,743 articles (0.46%). While the number of untrustworthy/predatory publications are low, there are still resources and grant dollars going into these publications. Future research will attempt to quantify the overall impact of this work in terms of grant money and research subjects. And while some research indicates that articles in predatory journals are infrequently cited (Björk 2020), it would be useful to observe the trend within one institution with a reputation for high quality research, as we have available in this dataset.

Regarding limitations, the dataset we used analyzed only works listed to a contributing faculty author. This excludes staff, postdoctoral, graduate, and undergraduate students. We reason that many of the publications written by these individuals would likely be coauthored with faculty and would therefore be caught by our search, but it is worth mentioning that there is a possibility that there are non-faculty publications which could further inform our research outcomes. Similarly, significant research efforts done at this institution are not reliant on NIH funding and therefore are not required to be deposited in PMC and may therefore fall outside our current data viewpoint. Finally, we did not examine author order, rank, or seniority, nor any characteristics of our faculty co-authors. It is certainly possible that additional trends could be discovered by examining affiliations of the first, last, or corresponding authors of publications as these are often key decision makers on where a manuscript will be submitted for publication. It may also be of interest to uncover trends between junior and senior researchers.

Our findings highlight two journals that represented 50% of the controversial or predatory publications among our faculty: *Cureus* and *Oncotarget*. These use nontraditional publishing practices, and there have been arguments for and against considering each legitimate (Emerald City Journal 2017; Groneberg et al. 2018; McCook 2018). We would like to communicate with the authors of papers that were published in these journals to understand their level of awareness of the controversy surrounding these titles. We are not able to determine from our analysis if faculty have fallen prey to controversial or predatory tactics or if they have made strategic, fully informed choices when disseminating their work.

Work also remains to be done within the university to eradicate publishing in predatory journals entirely. That work is made easier knowing that there are some individuals who have published in more untrustworthy journals than others, and some departments that follow the same pattern. We can investigate the reasons behind these trends and address the issues at their roots.

References

Amaral, O. B. 2018. "All publishers are predatory - some are bigger than others." *Anais da Academia Brasileira de Ciencias* 90 (2):1643-1647. doi: 10.1590/0001-3765201820170959.

Berisha Qehaja, Albana. 2020. "Avoiding publishing in predatory journals: An evaluation algorithm." *Journal on Efficiency and Responsibility in Education and Science* 13 (3):154-163. doi: 10.7160/eriesj.2020.130305.

Björk, Bo-Christer, Sari Kanto-Karvonen, and J. Tuomas Harviainen. 2020. "How Frequently Are Articles in Predatory Open Access Journals Cited." *Publications* 8 (2). doi: 10.3390/publications8020017.

Bohannon, J. 2013. “Who’s Afraid of Peer Review?” *Science* 342 (6154):60-65. Doi: 10.1126/science.342.6154.60.

Brembs, Björn. 2018. "Prestigious Science Journals Struggle to Reach Even Average Reliability." *Frontiers in Human Neuroscience* 12. doi: 10.3389/fnhum.2018.00037.

Cook, C. 2017. "Predatory Journals: The Worst Thing in Publishing, Ever." *Journal of Orthopaedic and Sports Physical Therapy* 47 (1):1-2. doi: 10.2519/jospt.2017.0101.

Cortegiani, A., and S. L. Shafer. 2018. "’Think. Check. Submit.’ to avoid predatory publishing." *Critical Care* 22 (1):300. doi: 10.1186/s13054-018-2244-1.

Cukier, S., L. Helal, D. B. Rice, J. Pupkaite, N. Ahmadzai, M. Wilson, B. Skidmore, M. M. Lalu, and D. Moher. 2020. "Checklists to detect potential predatory biomedical journals: a systematic review." *BMC Medicine* 18 (1):104. doi: 10.1186/s12916-020-01566-1.

Dony, Christophe, Maurane Raskinet, François Renaville, Stéphanie Simon, and Paul Thirion. 2020. "How reliable and useful is Cabell's Blacklist? A data-driven analysis." *LIBER Quarterly* 30 (1). doi: 10.18352/lq.10339.

Emerald City Journal. “Some Strange Goings On at Cureus.” Emerald City Journal | Seattle Newspaper for the People by the People. Emerald City Journal, February 1, 2017. https://www.emeraldcityjournal.com/2016/08/some-strange-goings-on-at-cureus/.

Eriksson, Stefan, and Gert Helgesson. 2018. "Time to stop talking about ‘predatory journals’." *Learned Publishing* 31 (2):181-183. doi: 10.1002/leap.1135.

Gonzalez, J., M. B. Bridgeman, and E. R. Hermes-DeSantis. 2018. "Differentiating predatory scholarship: best practices in scholarly publication." *International Journal of Pharmacy Practice* 26 (1):73-76. doi: 10.1111/ijpp.12380.

Goodman, Steven N. 2018. "A quality-control test for predatory journals." *Nature* 553 (7687):155-155. doi: 10.1038/d41586-018-00403-z.

Groneberg, David A, Axel Fischer, Doris Klingelhöfer, Michael H. K Bendels, David Quarcoo, and Dörthe Brüggmann. "The Story Behind Oncotarget? A Bibliometric Analysis." Scientometrics 117, no. 3 (2018): 2195–2205.

Koerber, Amy, Jesse C. Starkey, Karin Ardon-Dryer, R. Glenn Cummins, Lyombe Eko, and Kerk F. Kee. 2020. "A qualitative content analysis of watchlists vs safelists: How do they address the issue of predatory publishing?" *The Journal of Academic Librarianship* 46 (6). doi: 10.1016/j.acalib.2020.102236.

McCook, Alison. "When a Journal Is Delisted, Authors Pay a Price." Retraction Watch, March 6, 2018. https://retractionwatch.com/2018/03/06/when-a-journal-is-delisted-authors-pay-a-price/.

NIH Office of Extramural Research. 2017. "Statement on Article Publication Resulting from NIH Funded Research." https://grants.nih.gov/grants/guide/notice-files/NOT-OD-18-011.html.

NIH Office of Extramural Research. 2008. "Revised Policy on Enhancing Public Access to Archived Publications Resulting from NIH-Funded Research" NOT-OD-08-033. https://grants.nih.gov/grants/guide/notice-files/NOT-OD-08-033.html

Singh Chawla, Dalmeet. 2020. "Predatory-journal papers have little scientific impact." *Nature.* doi: 10.1038/d41586-020-00031-6.

Soteropulos, C. E., and S. O. Poore. 2021. "Navigating the Gray of Academic Publication: Avoiding Predatory Publishers and Creating Your "Whitelist"." *Annals of Plastic Surgery* 87 (6): e171-e179. doi: 10.1097/SAP.0000000000002902.

Strielkowski, W. 2018. "Predatory Publishing: What Are the Alternatives to Beall's List?" *American Journal of Medicine* 131 (4):333-334. doi: 10.1016/j.amjmed.2017.10.054.

Teixeira da Silva, Jaime A., and Panagiotis Tsigaris. 2018. "What Value Do Journal Whitelists and Blacklists Have in Academia?" *The Journal of Academic Librarianship* 44 (6):781-792. doi: 10.1016/j.acalib.2018.09.017.

Tsigaris, Panagiotis, and Jaime A. Teixeira da Silva. 2021. "Why blacklists are not reliable: A theoretical framework." *The Journal of Academic Librarianship* 47 (1). doi: 10.1016/j.acalib.2020.102266.