Review of: "Journal citation reports and the definition of a predatory journal: The case of the Multidisciplinary Digital Publishing Institute (MDPI)"

Rafael Santos[1]

1 University of Guelph

Potential competing interests: I am presently a voluntary editorial board member of journals belonging to the following publishers: MDPI; Hindawi; Frontiers; De Gruyter.

The article titled "Journal citation reports and the definition of a predatory journal: The case of the Multidisciplinary Digital Publishing Institute (MDPI)" (Oviedo-García, 2021) contains a number of poor analytical practices in terms of data collection, data processing, and data interpretation. I will outline two such issues here.

The first pertains to Table 1 in Oviedo-García (2021). Here are listed 53 MDPI journals, their respective JCR-JIF quartiles, and the names of "leading journals" with respect to the JCR category to which the MDPI's quartile corresponds. The "leading journal" is selected as the journal having the highest JIF in that category. I find this to be an inappropriate comparison for multiple reasons and offer a better alternative. One issue is that these tend to be much smaller journals than their corresponding MDPI journals. Another issue is that these tend to be journals that focus on publishing reviews (e.g., Annual Review of Plant Biology), or that use internal editorial reviews (e.g., Nature Reviews Materials), or that are generally very well regarded and subject to much different expectations and pressures than the typical scholarly journal. I suggest that a more appropriate choice would have been to compare the MDPI journals to the Q1 journal having the lowest JIF in the category; this would more likely yield a more comparable journal and one still having a healthy JIF. For example, for the MDPI journal Sustainability (2019 JIF 2.576), the pair journal would have been Springer's Environmental Science and Pollution Research (2019 JIF 3.056).

The next issue relates to the data and analysis of Table 2 in Oviedo-García (2021). In this table, the author reports self-citation data for all 53 MDPI journals and their respective "leading journals". Evidently, the values (as percentage figures) are much larger for the MDPI journals than for the other journals. The reason for this is that the calculation is not done in a reasonable way. What the author did was to divide the number of self-citations of a journal (which is the number of times that articles published in a given year have cited articles from the same journal published in either of the preceding two years) by the numbers of citations the journal received from all citing sources in the given year (for articles from the same journal published in either of the preceding two years). This type of calculation leads to figures that are much smaller for journals having large JIFs and much larger for journals having smaller JIFs. This is because it is natural for all journals to self-cite, but when journals are more widely cited, the self-citations make up a smaller fraction of the total number of citations. To prove this, I performed an analysis of five journals (MDPI's Sustainability; RSC's Energy & Environmental Science; Springer's Environmental Science and Pollution Research; Elsevier's Journal of Hazardous Materials; and

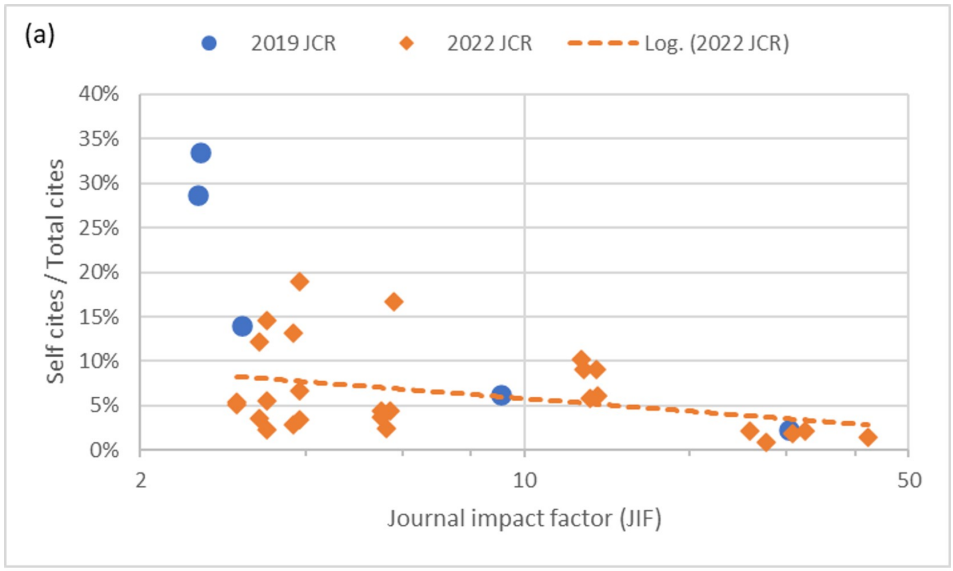
MDPI's Water) for the same 2019 JCR year as the paper used in Table 2. I then re-calculated self-citation using a different metric: dividing the number of self-citations (as aforementioned) by the number of articles the journal published in a given year. The data and analysis are available on Zenodo (https://zenodo.org/doi/10.5281/zenodo.10463873), and shown below.

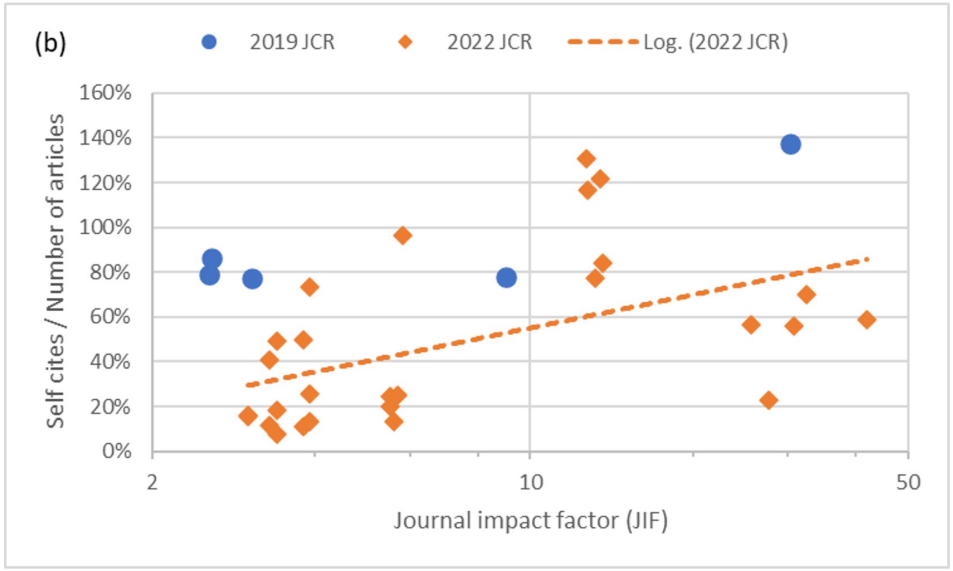
Journal	Sustainability	Energy & Environmental Science	Environmental Science and Pollution Research	Journal of Hazardous Materials	Water
Publisher	MDPI	RSC	Springer	Elsevier	MDPI
IF	2.576	30.289	3.056	9.038	2.544
IF-SC	1.711	29.598	2.628	8.467	1.814
Total Cit	18441	16114	17158	15501	7185
Self Cit	6190	368	2403	980	2062
articles	7184	268	3097	1262	2608
SelfC/TotC	33.6%	2.3%	14.0%	6.3%	28.7%
SelfC/article	86.2%	137.3%	77.6%	77.7%	79.1%

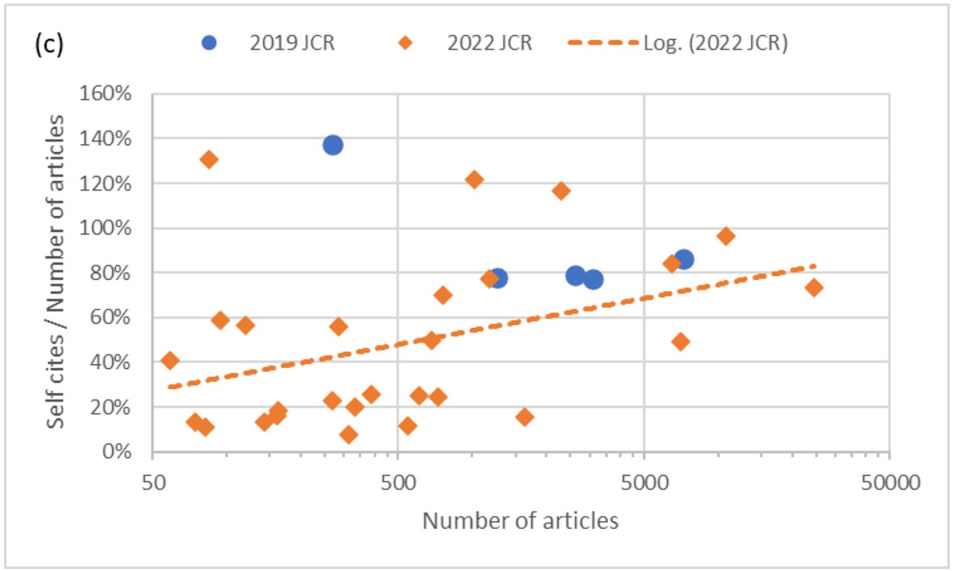
I found that this calculation metric (Self-citations/Number of Articles) gives very similar values across the four journals (from 77.6% to 86.2%), except for the RSC journal, which has a higher value of 137.3%. That is, on average, most articles cite about 1 other article published in the same journal in the preceding two years. This is a natural occurrence, since authors often choose to publish in a journal that has recently published at least one similar article, signaling a good scope fit and an editorial team familiar with the subject. It would be unnatural to expect articles not to cite other articles published in the same journal. As clearly seen in the data, even the RSC journal, having a 2019 JIF of 30.589, self-cites at a slightly higher rate than the other lower JIF journals. It is thus not possible to conclude that MDPI artificially boosts its JIF with self-citations, as their self-citation pace is equivalent to that of reputable journals from other publishers.

To expand the above analysis further and make it more contemporary, a re-analysis of the two metric ratios was performed for a larger number of journals (27) from the 2022 Journal Citation Reports (JCR) data for the Environmental Science category, which includes the 5 aforementioned journals. Other journals included in this analysis include journals with similar JIFs as those 5 journals, with JIFs ranging from as high as 42.1 (Nature Reviews Earth & Environment) to as low as 3.0 (Journal of Arid Land; Environmental Monitoring and Assessment), thus covering journals from the top of the Q1 quartile to the top of the Q3 quartile. Within this set, only the two aforementioned MDPI journals (Sustainability; Water) are from MDPI. In sub-figure (a), the analysis shows a declining trend of the self-citations/total citations metric used in Oviedo-García (2021) as a function of JIF. This confirms the previous observation that this metric artificially benefits journals with high JIF and punishes journals with low JIF, since all journals self-cite, but those which are less broadly cited will have a higher value according to this metric, irrespective of publisher. The JCR data also highlights the JIF inflation that has happened between the JCR years 2019 and 2022 (also discussed in Jones (2023)), with a shift to higher JIF values for all five of the journals analyzed using 2019 data. The highest JIF climb (+89.8%) was for Springer's Environmental Science and Pollution Research, and the lowest JIF climb (+7.3%) was for RSC's Energy & Environmental

Science.


(a)
2019 JCR
2022 JCR
Log. (2022 JCR)
Self cites / Total cites
40%
35%
30%
25%
20%
15%
10%
5%
0%
2
10
50
Journal impact factor (JIF)


(b)
2019 JCR
2022 JCR
Log. (2022 JCR)
Self cites / Number of articles
160%
140%
120%
100%
80%
60%
40%
20%
0%
2
10
50
Journal impact factor (JIF)


(c)
2019 JCR
2022 JCR
Log. (2022 JCR)
Self cites / Number of articles
160%
140%
120%
100%
80%
60%
40%
20%
0%
50
500
5000
50000
Number of articles

In sub-figure (b) above, the analysis is redone by plotting the self-citation/number of articles metric proposed in this review versus the JIF. Here, an increasing trend is observed for the 2022 JCR data, confirming the previous observation that journals with a higher JIF tend to self-cite more, given their reputation. The data is somewhat scattered, so exceptions do exist, likely linked to the focus of each journal. For example, Nature Sustainability has a broad scope and publishes more varied papers, and hence it is less likely that a new paper will cite another Nature Sustainability article. In fact, the editorial team in such a journal is likely to screen manuscripts for novelty and reject manuscripts that are too similar to recently published papers. On the other hand, the journal Biochar (Springer) focuses on a specific topic and thus is likely to have new papers cite its recent papers; hence, its metric value is the highest (131%) among the analyzed journals.

Sub-figure (c) was prepared to assess if there would be a correlation between the proposed metric and the number of articles (rather than the JIF). That is, the question was whether larger journals tend to self-cite more. The general trend says yes, though exceptions exist, most evidently the journal Biochar. Being a journal with a relatively high JIF (12.7) but publishing a relatively small number of articles (85 in two years), its self-citation rate appears high, but as aforesaid, this is due to its narrow scope. At the other end of the metric scale, as an outlier, the journal Environmental Monitoring and Assessment publishes a relatively large number of articles (1628 in two years) and self-cited infrequently (16%). This can again be linked to a journal having a broad scope and thus less likelihood of having new papers cite recently published papers from the same journal. Finally, at 73% and 49% metric values for the MDPI journals Sustainability and Water, respectively, neither of these journals appears as outliers in any of the analyses, falling within the spread of the trends that point to journals either having a high JIF or publishing a large number of articles, or having a narrow scope and self-citing more than others, understandably.

The above are only two notes, while many more shortcomings can be identified in the Oviedo-García (2021) article analysis that put into doubt much of the article's conclusions. I do not dispute that there might be areas that MPDI can improve to boost its prestige and scientific rigor, but most data that is publicly available through JCR and other sources points to much more similar patterns of peer review and editorial practices at MDPI than at various other traditional publishers and other well-established Open Access publishers. Predatory publishers exist and should be combated using reliable measures and criteria, of which there are many to use reliably.

References:

Jones, C.W. (2023). JACS Au Earns First Partial Impact Factor. JACS Au, 3(8), 2044-2045. DOI:10.1021/jacsau.3c00355

Oviedo-García, M.Á. (2021). Journal citation reports and the definition of a predatory journal: The case of the Multidisciplinary Digital Publishing Institute (MDPI). Research Evaluation, 30(3), 405-419a. DOI:10.1093/reseval/rvab020