ORIGINAL ARTICLE

Association of Learned and Professional Society Publishers

(wileyonlinelibrary.com) doi: 10.1002/leap.1428

Received: 2 August 2021 | Accepted: 5 November 2021

Journals in Beall's list perform as a group less well than other open access journals indexed in Scopus but reveal large differences among publishers

Henk F. Moed[†], Carmen Lopez-Illescas[1], Vicente P. Guerrero-Bote[2*] and Felix de Moya-Anegon[3]

[1]Departamento de Biblioteconomía y Documentación, Facultad de Ciencias de la Documentación, Universidad Complutense de Madrid, Madrid, Spain

[2]SCImago Research Group, Departamento de Información y Comunicación, Facultad de Ciencias de la Documentación y la Comunicación, Universidad de Extremadura, Badajoz, Spain

[3]SCImago Research Group, Granada, Spain

ORCID:
H. F. Moed: 0000-0003-2446-905X
C. Lopez-Illescas: 0000-0001-9918-4978
V. P. Guerrero-Bote: 0000-0003-4821-9768
F. deMoya-Anegon: 0000-0002-0255-8628

*Corresponding author: Vicente Guerrero-Bote, SCImago Group, Department of Information and Communication, University of Extremadura, Badajoz, Spain.
E-mail: guerrero@unex.es

Abstract

The list of potential, possible or probable predatory scholarly open access (OA) publishers compiled by Jeffrey Beall was examined to determine the effect of their inclusion upon authors, and a possible bias against OA journals. Manually collected data from the publication archives of a sample of 250 journals from Beall publishers reveals a strong tendency towards a decline in their article output during 2012–2020. A comparison of the subset of 506 Beall journals indexed in Scopus with a benchmark set of other OA journals in Scopus with similar characteristics shows that Beall journals reveal as a group a strong decline in citation impact over the years, and reached an impact level far below that of their benchmarks. The Beall list of publishers was found to be heterogeneous in terms of bibliometric indicators but to be clearly differentiated from OA journals not included in the list. The same bibliometric comparison against comparable non-OA journals reveal similar, but less marked, differences in citation and publication growth.

Keywords: Beall's list, open access publishing, predatory journals

INTRODUCTION

What is a predatory scholarly journal?

The website of one periodical states: 'Acceptance Notification - within a week' and that of a second journal: 'If you are interested in becoming a member of the editorial team, kindly submit your Curriculum Vitae (CV) through email'. Manuscript peer review normally takes more than 1 week, and one would not expect to read invitations to join the editorial team on a website of a serious, peer reviewed journal, the launch of which has been well prepared, and that constitutes a solid pillar to the scientific-scholarly archive. Can these journals be qualified as 'predatory'?

There is not one standard definition of what constitutes a predatory publisher. But most definitions state that a predatory publisher charges a fee for the publication of a paper without providing publication services such as peer review and editing. An international group of 43 participants from various stakeholders reached the following consensus definition:

> Predatory journals and publishers are entities that prioritize self-interest at the expense of scholarship and are

[†]Deceased.
Henk F. Moed sadly passed away before publication of this article. He was a valued colleague, and will be greatly missed.

Key points

- A sample of 250 journals from publishers in Beall's list reveal a strong tendency towards a decline of their article output during 2012–2020.
- 506 Beall journals indexed in Scopus show a strong decline in citation impact and an impact level far below that of other OA and non-OA periodicals in Scopus.
- Beall's list shows a strong internal variability and bibliometric analysis raises questions about the inclusion of some publishers.
- Subscription-based journals may suffer from the same type of quality issues as periodicals captured in the Beall list do and should be analysed as well.

> characterized by false or misleading information, deviation from best editorial and publication practices, a lack of transparency, and/or the use of aggressive and indiscriminate solicitation practices (Grudniewicz et al., 2019).

Earlier studies on the Beall list and its criteria

As from 2008, Jeffrey Beall maintained a list of 'potential, possible or probable' scholarly open-access publishers. He also published a list of typical *practices* which were conducted by publishers, editors and their staff, related to editorial processes, business management, integrity, journal standards and other aspects, and which functioned as criteria to assess whether a particular journal or publisher could be qualified as 'potential, possible or probable predatory' or 'questionable' (Beall, 2017).

The number of publishers and standalone journals in this list grew rapidly during the years until 2017, the year in the beginning of which the entire content of Beall's *Scholarly open access* website was removed from his website. In the current paper, the term 'Beall journals' is used to indicate the set of journals published by publishers in Beall's list of questionable *publishers* as well as the *standalone* journals in this list.

For a comprehensive and balanced discussion on the value and limits of the Beall list the reader is referred to the paper by Macháček, V., Srholec, (Macháček & Srholec, 2021[1]). The following limitations are particularly relevant in the current article. First of all, a formal, transparent justification of Beall's judgements is lacking. This has in particular consequences for the interpretation of Beall's list of publishers. As Macháček and Srholec (2021) put it,

> Classifying an entire publishing house as predatory is a strong judgment, and it cannot be ruled out that some journals which actually apply reputable standards have been blacklisted along the way.

Moreover, the data are historical, and journals that did reveal clear characteristics of predatory behaviour may have improved their performance.

An overview of the various proposed 'frameworks' for identifying 'questionable' journals is given in Frandsen (2019). She concluded that a great diversity exists among these frameworks in format, length and content, and that they offer little information on the validation and the reliability of the applied methodologies and criteria. The current paper brings in a methodology that has proven both its merits and its limits, namely a *bibliometric* analysis (Moed, 2017). It analyses manually collected publication counts of several hundreds of journals in the Beall list, and compares the trend in the publication output and citation impact of Beall journals with other OA and non-OA periodicals indexed in Scopus, the multidisciplinary citation index published by Elsevier.

Several earlier articles applied bibliometric approaches in the study of questionable or predatory journals. In an analysis of bibliometric author profiles Xia *et* al. (2015) found that, due to economic and sociocultural conditions, authors of predatory journals are 'mostly young and inexperienced researchers from developing countries'. In a citation analysis of authors *citing* a sample of standalone journals in the Beall list, Frandsen (2017) concluded that their profile is similar to that of authors *publishing* predatory journals. Most of them are inexperienced authors from Africa, Southeast Asia or South Asia. By contrast, in a study of Beall journals in the field of Economics, Wallace and Perri (2018) found that the geographic dispersion of authorship in these journals is widespread, and that several researchers in the top 5% of most frequently publishing authors in the Economics Research Papers (RePEc) database published in predatory journals in 2015.

Analysing Web of Science, Scopus, the Directory of open access Journals (DOAJ) and other databases, Somoza-Fernández *et al.* (2016) concluded that there is 'no significant widespread presence' of journals from Beall's list in bibliographic databases. In a large study on predatory publishing, Shen and Bjork (2015) identified about 11,000 journals from publishers in the Beall list, 8000 of which were active in 2014. They concluded that 'the problems caused by predatory journals are rather limited and regional', and that 'publishing volumes in such journals will cease growing in the near future'. The current paper examines the presence of Beall journals in Scopus, and analyses trends in publication output in a sample of a few hundred Beall journals during 2012–2020.

Other recent *citation* studies of predatory journals were published by Oviedo-García (2021) on journals from the publisher MDPI based on data from Clarivate's Journal Citation Reports, and by Björk *et al.* (2020) of a sample of 250 random articles in predatory journals using citation data from Google Scholar. Our article presents a citation analysis of all Beall journals indexed in Scopus, using citation data from Scopus and comparing Beall journals with a well-defined benchmark sets of serials in Scopus.

1. The article by Macháček, V., Srholec (2021) was retracted after submission of this article. Depite this, the authors consider that the reasons for retraction do not affect the data set or other information used for this article.

Oermann *et al*. (2019) found several hundreds of citations, many of which were given in non-predatory journals, to articles published in predatory nursing journals. They argued that 'education and information may help authors and reviewers identify predatory journals'. This emphasis on the need to educate authors is also a key feature of a study by Cohen *et al*. (2019) on perspectives from authors and editors in Biomedicine. They concluded that authors publishing in predatory journals are 'alarmingly uninformed' in terms of predatory journal quality and practices. If the first pillar of the current paper is the bibliometric approach, the second is the notion that information scientists or bibliometricians should provide more information to authors about questionable journals, thus enabling authors to make informed decisions as regards the outlets they choose for publishing their papers.

Predatory and open access journals

The property that a predatory publisher charges a fee for publishing articles seems to connect predatory publishing exclusively with a particular business model of scientific publishing often denoted as the 'authors pay' or open access (OA) model, according to which the publication of articles is paid by the authors or by their institutions or sponsored by other organizations, as distinguished from a 'readers pay', subscription-based, toll or closed access (TA/CA) model. Using the perhaps somewhat suggestive term 'closed', it follows that subscription-based access is closed towards *readers*, and OA towards publishing *authors*. Beall explicitly deals with OA journals, which raises the question whether subscription based journals may be predatory or questionable as well.

Moustafa (2015) underlined that some of the criteria used by other authors to identify 'fake' (or predatory) journals are not necessarily specific to fake journals only, but they could also apply to well-established journals. In addition, subscription-based journals may have 'predatory' characteristics, and this point is picked up in the discussion section at the end of our article.

In a review of the issues raised by critics of the criteria applied by Beall and followers to identify potential, possible or probable predatory publishers and journals, Kimotho (2019) found four major issues: 'methodological flaws; Beall's bias against OA; discrimination against developing economies; and Beall's lists of predatory publishers as an onslaught to academic freedom' (Kimotho, 2019, p. 1). The bias against OA is also the main concern of a review by Krawczyk and Kulczycki (2021) who analysed how predatory journals are characterized by authors who write about such journals. They concluded that 'the overgeneralization of the flaws of some open access journals to the entire open access movement has led to unjustified prejudices among the academic community toward open access'. Our article aims to counter such prejudices by objectively comparing Beall journals to other OA periodicals, and OA to non-OA journals.

Research questions

Our paper takes the Beall list as a starting point and first addresses the question: Is there any evidence that the publication of the Beall list has had any effect at all upon the extent to which authors decided to publish papers in these journals? If there is no evidence that the Beall list directly or indirectly influenced publication practices, one might argue that participants in the publication process continued their business 'as usual', and that further research into the Beall list, although theoretically interesting, would have little practical value. We pose two empirical research questions about this issue: How did the publication output of journals in the Beall list develop during the past decade? And: how did the number of Beall journals indexed in Scopus change?

A second issue relates to the perceived negative bias towards OA journals expressed by some of the authors mentioned above. On the one hand, one may ask whether or not subscription-based journals may be predatory in the Beall sense. The fact that Beall explicitly analysed OA journals is certainly a limitation of the Beall list. If the dominant principle of the Beall list is a bias against OA rather than the lack of quality of the journals, one would expect to find that the inclusion of one journal in the list is as appropriate as any other, as long as it is OA. Do the journals in Beall's list as a group show characteristics that make them different from other OA journals that were not selected? We address this question empirically using a bibliometric analysis of OA journals indexed in Scopus. How does the publication output and citation impact of Beall journals in Scopus list compare to OA journals in Scopus that are not on the Beall list?

Finally, a third issue interprets the outcomes of the bibliometric analysis in terms of journal performance. How does the performance of journals in the Beall list compare to that of other OA journals, and how do OA journals compare to non-OA periodicals indexed in Scopus? Two bibliometric indicators are used for this analysis: the number of articles in a journal, and the citation impact of these articles. The first is assumed to reflect the willingness of potential authors to publish an article in a journal, while the second indicates the tendency of researchers to read and cite a journal's publications in their own papers.

A complicating factor in the comparison of groups of journals is that several authors dispute the inclusion of particular journals and publishers in the Beall list, especially those periodicals published by the Frontiers Research Foundation (e.g., Macháček & Srholec, 2021). Removing these journals from the study set is theoretically speaking not a neutral act, as this may have consequences for the outcomes of the research question outlined above, and is therefore methodologically incorrect. For this research, we take the position that if there were to be found any evidence at all of special characteristics of the Beall journals as a group, it should relate to the entire group, including those published by Frontiers.

Direct versus indirect effects

One could detect a direct effect of the publication of the Beall list if the very appearance of a particular journal in this list decisively convinces authors, editors and publishers not to use it as a publication outlet. A more indirect effect is that the list alerts stakeholders that certain quality issues may exist especially with

journals of particular OA publishers, and that it stimulates authors to be more critical when choosing a journal. In a purely bibliometric study it is difficult, if not impossible, to separate these two types of effects. Therefore, the discussion below speaks of effects in general, as it is not possible to further specify the type of effect. It is assumed that, generally speaking, authors base their decisions where to publish not merely upon a single external source such as the Beall list, but using their own experiences and those from their colleagues and other sources of information, such as whether a journal is indexed in important literature databases in their field.

Base assumptions on the bibliometric approach

As outlined above, we here analyse possible effects of the publication of the Beall list from a bibliometric perspective. This choice is not based on the assumption that a bibliometric approach has a preferred status over other possible approaches studying 'questionable' journals. On the contrary, we believe that other approaches may be equally valuable. However, we are bibliometricians by training, and believe that bibliometrics may provide valuable tools to obtain insight into publication practices of authors, editors and publishers, and therefore may contribute to enlighten these practices.

The use of bibliometric, citation- and publication-based indicators in the assessment of scientific-scholarly journals is the subject of intensive debate in the domains of bibliometrics and scholarly publishing (see for instance Glanzel et al., 2019). We believe that a fundamental assumption of the evaluative bibliometric approach holds that researchers are in principle able to recognize quality and consider it worth pursuing, and that they make independent judgements and decisions related to it in their daily practices. This assumption provides a justification for the search for possible traces of quality—not as direct reflections, and distorted by other factors or veiled as they may be—in these practices. If one does not accept the validity of this assumption, bibliometric indicators in any form of quality assessment seem to be meaningless. However, as counter argument against rejecting this assumption, one may argue that its rejection eventually denies that science is possible at all and can therefore jeopardize scientific progress.

Structure of this paper

The next two sections present details on the process of data collection and on the applied methodology, including the use of statistical tests. The first part of the results section presents an analysis of the trends in annual publication output of a sample of Beall journals, and in the number of Beall journals indexed in Scopus. The second part compares the publication output and citation impact of Beall journals indexed in Scopus with those measures for other OA journals and for non-OA journals in Scopus. The final section presents a discussion of the outcomes and the conclusions of the study.

DATA COLLECTION

Data on Beall journals

In an earlier study by Macháček and Srholec (2021), based on a list of publishers and standalone journals published by Beall in his blog (Beall, 2016), each publisher in the Beall list was manually uploaded to Ulrich's Periodicals Directory in 2016, and the titles of all journals were downloaded. In this way, a total number of 3275 journals were identified. Vit Machacek kindly shared this data set with us. Their data set contains 360 publishers and 301 standalone journals.

To avoid a possible temporal selection bias, it was decided to expand for the 50 top publishers in the Machacek and Srholec set the study data set with journals that were included in the Ulrich database in 2020. After this update, the total number of Beall journals in the study data set was 4038, of which 506 (12%) were indexed in at least 1 year in Scopus.

Adding information from Scopus to Beall journals

Journals in the Beall list were matched against a special data set with publication and citation counts per journal covering the time period 1996–2019, derived from SCImago Journal Rank (SJR), a database with journal indicators created by SCImago Research Group based on data from Elsevier's Scopus. The matching algorithm was based on Print- and E-ISSN codes or—if no match using ISSN was found—full journal titles. Information on journals' OA status was extracted from the Scopus Source List 2020. In this way, in the Machacek and Srholec data set 460 journals were identified as source journals in Scopus, and in the expanded data set 506 journals. About 60% of these were discontinued in Scopus, either because they were discontinued by the publisher or because Scopus stopped indexing them. For detailed information on Scopus journal coverage and evaluation the reader is referred to Scopus Content Policy (n.d.).

Creation of a study sample and collecting publication counts for Beall journals not indexed in Scopus

From the total set of 3532 journals in the Beall list *not* indexed in Scopus a random sample was created of 253 (7%) journals for which the number of articles published per year during 2012–2020 was collected manually from the journals' publication archives available via their websites. The document types included were articles, reviews and communications (collectively referred to as 'articles' in this paper). Forty-five journals (18%) could not be found in the publishers' archives and were discarded from the sample and in the calculations below. The final sample set of Beall journals not indexed in Scopus included 208 journals.

METHODOLOGY

Indicators of a journal's publication output and citation impact

The journal impact indicator used in the current study is a relative or field-normalized citation rate calculated as follows. Citations are counted in a particular (citing) year to 1-3-year-old (cited) documents published in the journal rather than to 1-2-year documents as is done in the calculation of in the standard Clarivate/ISI Journal Impact Factor. Document types included are articles, reviews, conference papers and short surveys. A field-normalized journal impact indicator denoted as RJIF is calculated by dividing a journal's citation rate by the average citation rate for all journals in the same subject categories as those assigned to the journal, using a classification in Scopus of about 300 subject categories. If a journal was assigned to multiple subject categories, a weighted average citation rate was calculated, the weights being determined to the number of subject category assignments to a journal.

There is a huge literature on the validity and usefulness of journal impact factors and related measures of citation impact of scientific journals. For an overview, the reader is referred to Larivière (2019). Although the relative or field-normalized journal citation impact used in the current paper corrects for differences in citation and publication practices among subject fields, it is an average of a distribution of citations among published papers that is in most cases highly skewed. Even journals with a low average impact can publish a few more heavily cited articles. Moreover, to some extent the value of indicators can be affected by citation manipulation. Typical examples are given by Reedijk and Moed (2008).

Benchmark approach in the Scopus-based analysis

Journals indexed in Scopus were categorized into three main groups: Beall journals; other OA journals, defined as periodicals not in the Beall list for which the Scopus Source List contained the qualification 'In DOAJ/ROAD'; and a large rest group denoted as non-OA journals. The number of journals in these three groups are around 506, 5700 and 26,000, respectively. (DOAJ, The Directory of OA Journals, is a website that hosts a community-curated list of OA journals, maintained by Infrastructure Services for OA. ROAD, the Directory of OA scholarly Resources, is a service offered by the ISSN International Centre with the support of the Communication and Information Sector of UNESCO.)

In view of the large differences in the number of journals in the three groups, it was decided not to compare statistics for the entire groups, but to create for each of the two groups of Beall journals *two* appropriate *benchmark* sets of journals from the Other OA and from the non-OA group, by selecting at random for each Beall journal a periodical with similar characteristics using the following four indicators:

- Main discipline covered—using a categorization of journals into five main disciplines (Biomedical Research; Clinical Medicine; Natural Sciences; Engineering; Social Sciences and Humanities).
- First year the journal was indexed in Scopus.
- Index of National Orientation (INO)—measured by the percentage of articles (co-)authored by researchers affiliated with the most productive country publishing in a journal (e.g., a journal's INO value of 90 means: one single country (co-) authored 90% of all articles published in this journal).
- Publication language—in terms of English versus non-English.

To find a benchmark journal for a given periodical in the Beall list, in a first step a match key was used containing all four data elements listed above. In this step, a benchmark was found for around 70% of Beall journals. The next step used a match key containing the first three elements, and found a benchmark for another 10% of journals, while the use of a key containing only the first two variables yielded another 20%.

Wilcoxon signed rank test

When comparing indicator values for journals in a study set with those in the benchmark set, a Wilcoxon Signed Rank Test was applied. Preliminary testing revealed that the underlying indicator distributions of the study and benchmark sets strongly deviated from normality. In this case the Wilcoxon test is considered a good option. This is a non-parametric statistical hypothesis test that can be used to compare the locations of two populations using a set of matched samples and does not assume that data are normally distributed.

Significance of trends

For a journal's time series of publication counts and citation impact, a growth rate was computed based on a linear regression, with the indicator as the dependent and the year as independent variable, by dividing the regression coefficient by a journal's mean annual score. It was tested whether the trend in the annual scores was significant or not, by testing whether the regression coefficient deviates significantly from zero applying a 99% confidence level. This method is assumed to give a rough, 'first order' indication of whether or not the data show a positive or negative trend.

Methodological comment on the use of Scopus data

Discrepancies between the 'reality' in the publication archives of journal publishers and its reflection in a database may to some extent affect the accuracy of the publication counts. For instance, if a journal has zero publications indexed in Scopus in a given year, one does not know whether the journal did not publish any papers in that year, or whether it did publish papers but these were not indexed in the database. It is assumed that the three journal data sets extracted from Scopus (Beall journals, other OA

and non-OA periodicals) are all affected to the same degree by this problem. Obviously, the manually collected counts are not affected by these problems, but it would be inappropriate to compare Beall journals indexed in Scopus directly with those periodicals not indexed in this database.

OVERALL TRENDS IN PUBLICATION COUNTS OF BEALL JOURNALS NOT INDEXED IN SCOPUS

Figure 1 relates to a sample of 208 Beall journals not indexed in Scopus, whose publication archives could be found at or via the publisher's website, and that published in at least 1 year during 2012–2020. It shows that 9% of journals revealed a significantly positive trend in the annual number of published articles per year during 2012–2020, and 25% – about one quarter of the journals in the sample—a negative trend; 40% of periodicals was discontinued during the time period considered. The ratio of the number of journals with a positive trend over this number for periodicals with a negative trend amounts to 0.28. This ratio will be labelled as 'Pos/Neg Trend Ratio' in the sections below. In the next section this ratio is calculated for Beall journals indexed in Scopus and compared with that of a benchmark set of *other OA* journals indexed in Scopus and a set of *non-OA* periodicals in Scopus.

BEALL JOURNALS INDEXED IN SCOPUS COMPARED WITH OTHER OA AND NON-OA JOURNALS IN SCOPUS

This section analyses the subset of Beall journals indexed in Scopus. As an introduction, the trend is analysed in the number

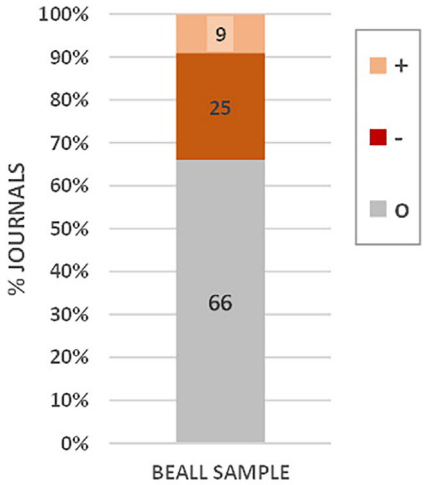

FIGURE 1 Trends in publication counts in a sample of 208 Beall journals (not indexed in Scopus) between 2012 and 2020. '+': significantly positive; '−' significantly negative; 'o': not significant at $p = 0.05$.

of Beall journals indexed in Scopus. The next two analyses compare Beall journals with other OA and with non-OA journals. The first analyses the temporal development in the number of published articles and their citation impact, while the second focuses on the level of the journals' publication output and citation impact in the last year they were indexed in Scopus.

Trends in the number of Beall journals indexed in Scopus

Beall journals analysed were indexed in Scopus in at least some years during 1996–2019, but not necessarily in all years. In fact, only 222 (44%) were still active in 2019, slightly more than 0.1% of the total number of about 20,000 journal sources active in Scopus in this year. Figure 2 shows the number of Beall journals entering or leaving Scopus in a given year during this period. About 100 Beall journals entered Scopus in 2009. As regards the number of periodicals discontinued in Scopus in a given year, 2016 was clearly a peak year, at the end of which 105 Beall journals were discontinued in Scopus.

Large differences exist among publishers. To mention a few typical examples: For Bentham Open 256 journals were found in Ulrich's database in 2016 or 2020. Eighty-three titles were identified in Scopus, 31 of which are still indexed in 2019—12% of the total number of Bentham Open journals that were found in Ulrich in 2016 or 2020. As regards the publisher Research India Publications, 190 journals were found in Ulrich in 2016 or 2020, 13 of which are in the Scopus 2020 Source List, and only 7 (4%) are still indexed in 2019.

Trends in the number of published articles and citation impact

For each journal, we tested whether the number of published articles or the citation impact showed a significantly positive or negative trend during the time period 2011–2019. Figure 3 shows all Beall journals indexed in Scopus, and the two benchmark sets containing other OA journals indexed in Scopus and non-OA journals in Scopus respectively. It presents a breakdown of journals according to two indicators: the trend in annual published articles, and the trend for field-normalized or relative citation impact. Table 1 further elaborates on the data given in Fig. 3, and presents for the total set of Beall journals the ratio of the number of journals with a significant positive trend over the number with a negative trend.

Figure 3 shows for the group of all Beall journals (indexed in Scopus) that the number of journals with a positive trend in the annual number of published articles is more than twice the number of periodicals revealing a negative trend. In fact, Table 1 reveals that the ratio amounts to 2.3. But for the benchmark set of other OA journals this ratio is higher (2.9 against 2.3), and that for non-OA periodicals slightly lower (1.9 against 2.3). In other words, there is a tendency that both the Beall journals and the two-benchmark groups increase their publication output, but this

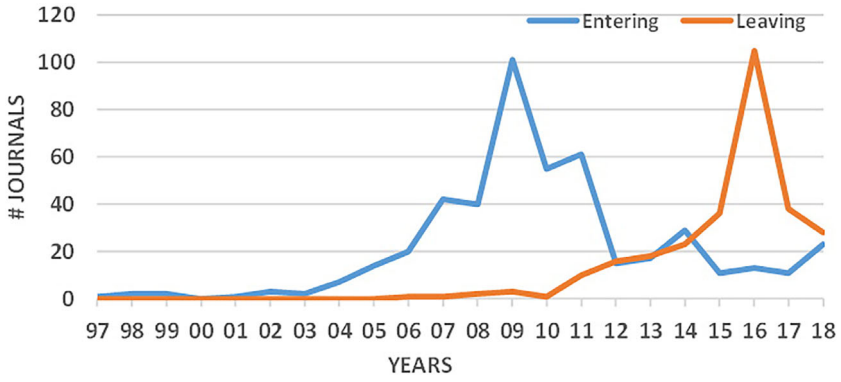

FIGURE 2 Number of Beall journals entering or leaving Scopus in a particular year.

tendency is weaker for the Beall journals than it is for the benchmark group of other OA journals indexed in Scopus.

The results with respect to citation impact show a similar pattern to that for publication counts, but differences between Beall and benchmark journals are larger. There is a tendency for Beall journals as a group to reveal a decline in their citation impact, while the other OA journals show an increase, and in the non-OA benchmark citation growth is extremely slight (1.2). Comparing all about 5700 other OA journals in Scopus with a benchmark set of non-OA periodicals, it was found that other OA journals tend to increase their article output and citation impact more often than their non-OA counterparts do, with ratios of number of increasing over declining journals of 2.9 versus 1.9 for the former indicator and 2.9 versus 1.3 for the latter. It must be noted that while the trend in annual publication counts during 2011–2019 is obviously affected by the removal of a journal from the database during this time period, the trend in citation impact is less affected by such removal, as data are beginning to be lacking to calculate the citation impact.

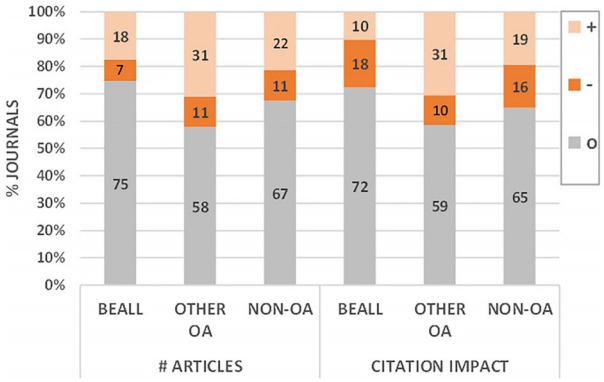

FIGURE 3 Percentage of journals with a significantly positive ('+') or negative ('−') trend or with no significant trend ('o') in two indicators: the number of published articles and the field-normalized or relative citation impact, for three journal sets: BEALL = All Beall journals indexed in Scopus; OTHER OA = Other OA journals indexed in Scopus; non-OA = Non-OA journals in Scopus.

The level of publication output and citation impact in the end year

The Wilcoxon signed rank test examines whether publication output or citation impact is different from zero between the two data sets in a pair. Table 2 shows that according to the test the differences between all pairs of data sets are statistically significant at $p < 0.001$. Thus, one must reject for each pair and each indicator the null hypothesis that the paired rank difference is symmetric around zero and therefore conclude that differences in publication output and citation impact exist between the journal sets in the various pairs.

However, in several cases the differences, though statistically significant, are small. This is especially true for the number of publications. The citation impact of journals in the Beall list and indexed in Scopus tends to be much lower than that of serials in the benchmark set of other OA journals indexed in Scopus (medians are 0.40 vs. 0.74), and also lower than it is for the benchmark set of non-OA journals in Scopus (0.40 vs. 0.49). Next, all other OA journals in Scopus tend to have a larger impact than those in the benchmark set of non-OA serials (0.58 vs. 0.47).

Bibliometric variability among publishers in the Beall list

Figure 4 presents for Beall publishers with three or more journals indexed in Scopus the average number of published articles per year, and the average relative citation impact (RJIF) over the years. It clearly illustrates the heterogeneity in terms of the key bibliometric indicators among publishers in the Beall list. For instance, for one publisher that has more than 10 journals are indexed in Scopus the average article output per journal per year

TABLE 1 Ratio per group of the number of Beall journals indexed in Scopus with a positive trend over this number with a negative trend

		Ratio pos/neg trend		
Beall set	**Indicator**	**Beall**	**Other OA**	**Non-OA**
All Beall	# Articles	2.3	2.9	1.9
	Citation impact	0.6	2.9	1.2

TABLE 2 Comparison between study sets and control groups according to a journal's number of publications and field normalized citation impact in the end year.

		Benchmark Sets							
		All other OA journals in Scopus				All non-OA journals in Scopus			
		Median				Median			
Study set	Indicator	Bench-mark set	Study set	Difference	Significance	Bench-mark set	Study set	Difference	Significance
All Beall journals indexed in Scopus ($n = 506$)	Number of articles published in end year	41.0	43.0	−0.20	***	35.0	43.0	−7.0	***
	Normalized citation impact in end year	0.74	0.40	0.23	***	0.49	0.40	0.02	***
All other OA journals in Scopus ($n = 5700$)	Number of articles published in end year					29.0	34.0	−3.0	***
	Normalized citation impact in end year					0.47	0.58	−0.08	***

Note: Difference: Difference between an indicator's value for the benchmark set minus this value for the study set. The Median Difference: median of the variable difference, not difference of the median values in two sets. The test statistic S is the sum of the ranks of the positive values minus the sum expected under the null hypothesis (which equals $n^*(n + 1)/4$, where n indicators the number of journals in a study set. ***: significant at $p < 0.001$. "End year": The last publication year of articles in journals indexed in Scopus.

is 240, while the average relative citation impact amounts to 1.4, which is above the world average for the subject fields covered by these journals. By contrast, in the bottom right border of the graph there are six publishers with fewer than 100 papers per year per journal and an average citation impact below 0.3.

DISCUSSION AND CONCLUSIONS

The observed removal from Scopus of about 100 Beall journals in peak year 2016, and the finding that one-quarter of Beall journals not indexed in Scopus shows a decline in the annual number of published articles during 2012–2020, reveals (for journals in the study set) that a declining trend and being included in Beall's list are statistically associated. It follows that empirical research into the validity of the Beall list is an important issue with a great practical relevance. But how should this association be interpreted?

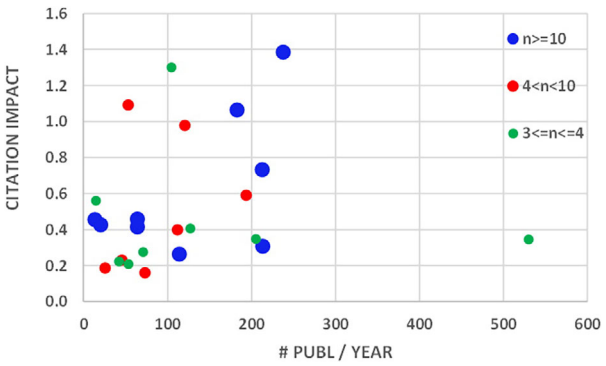

FIGURE 4 Average publication counts per year and average relative citation impact (RJIF) per published journal for major Beall publishers and derived from Scopus. *n* indicates the number of a publisher's journals indexed in Scopus.

If there is any effect at all of the Beall list, it is not necessarily a direct effect in the sense that the very appearance of a particular journal in this list by itself convinces authors or indexers not to use it for publication or indexing. It seems more plausible to hypothesize the influence of a more indirect effect according to which the list contributes to a stronger awareness of authors and indexers that certain quality issues may exist especially with journals from specific OA publishers, and that it stimulates them to be more critical when choosing a journal for publication or citation.

A comparison of Beall journals indexed in Scopus with other OA and non-OA journals indexed in this database provides bibliometric evidence that Beall journals tend to perform less well than other OA periodicals in terms of their annual article output and especially their citation impact. Other OA journals in Scopus tend to perform better than non-OA journals indexed in this database. Compared to the latter group of journals, Beall journals show statistically similar trends and levels in article output, but lower citation impacts.

However, the interpretation of the observation that journals in Beall's list as a group shows characteristics that make them different from other OA journals is not straightforward. Does this mean that the two groups of journals are essentially different in terms of performance and predatory characteristics? Or is there a confounding factor at stake, namely that one group was included in Beall's list while the other was not, and that the very appearance of a journal in Beall's list induced a bandwagon effect that made authors stop publishing in it and citing it? (The bandwagon effect is the term used to describe the tendency for people to adopt certain behaviours, styles, or attitudes simply because others are doing so https://en.wikipedia.org/wiki/Bandwagon_effect.)

As argued in the Introduction section, we defend the position that researchers are in principle able to recognize quality and consider it worth pursuing, and that they make independent judgements and decisions related to it in their daily practices. This assumption provides a justification for the interpretation that the observed differences in bibliometric outcomes between Beall journals and other OA serials should not merely be ascribed to a bandwagon effect, but also to a difference in perceived journal performance.

Obviously, these conclusions are based on partial, purely bibliometric evidence, under the assumption that journals in both sets are affected to the same extent by citation manipulation. They relate to a specific set of journals in the Beall list, namely those indexed in Scopus. Like any other scientific literature database, Scopus is not a theoretically neutral measuring device. The very inclusion or exclusion of journals in the database may affect their visibility and performance. In addition, as outlined in the Methodology section, discrepancies between the 'reality' in the publishers' archives and its reflection in a database may to some extent affect the accuracy of the publication counts.

As outlined in the introduction section, if the Beall list would reveal a dominant bias against OA journals, one would expect that the inclusion of one journal in the list is as appropriate as any other, as long as it is OA. As the results show that the journals in Beall's list as a group show bibliometric characteristics that make them different from other OA journals that were not included in this list, it is hypothesized that bias against OA is not the dominant principle of the Beall list, and that journals on the list perform less well in terms of citation impact than other OA journals in Scopus.

This does not mean that the list should not be critically analysed and updated. From a bibliometric point of view, the heterogeneity of the list is a point of great concern. A bibliometrician may ask why particular journals or publishers are on the list despite their increasing annual publication output and high citation impact during the past decade. Following Oviedo-García (2021) and Björk *et al*. (2020), the response of the scientific community as expressed in publication and citation practices can be studied from a quantitative informetric or bibliometric view point, and the outcomes may constitute a valid and useful basis—along with findings from other approaches, such as inquiries of publisher websites and of the duration of peer review processes, and qualitative interviews or questionnaires—to assess scientific-scholarly journals. Bibliometric-informetric studies may reach beyond the calculation of 'simple' publication counts or impact factors, and include for instance author retention studies or computational linguistic analyses of peer review processes (Moed, 2016).

While Beall focused on OA journals, and claimed that 'predatory publishing is just one of the consequences of gold open access' (Beall, 2013), we believe that the type of predatory behaviour Beall aimed to analyse is not a necessary consequence of Gold OA. This is fully supported by the analyses presented here. Aggregating several types of OA, and applying a journal's inclusion in the DOAJ database as the criterion to label it as OA, our study provides bibliometric evidence that OA journals tend to perform better than non-OA journals indexed in Scopus, confirming conclusions drawn in earlier papers (see for instance Bautista-Puig et al., 2020). Gold OA journals are not necessarily of poor quality.

In addition, we believe that subscription-based or Toll Access periodicals can be potential, possible or probably predatory as well. Publishing or scientific editors of subscription-based journals may also be subjected to institutional pressures to reach a certain level of published articles, for instance, to arrive at the end of a year at a number of published manuscripts planned at the beginning of the year. Although such pressures may not be directly visible in the websites of affected journals—an important source of information in the compilation of the Beall list—there is no *a priori* reason why editors of subscription-based periodicals would never promise very fast peer review or accept manuscripts without any rigorous form of peer review, two core characteristics of predatory OA journals.

If OA journals are fee-collecting directly from authors, Toll Access journals are subscription-collecting, aimed at reaching a certain number of published articles per year agreed upon in advance, either internally between a publishing editor and his/her managing director or externally between a publisher sales department and an academic library. Assessment of the quality of journals should be made for all journals, regardless of their business model. Hence, research into possible predatory behaviour of subscription-based journals should be conducted as well.

ACKNOWLEDGEMENT

The authors wish to thank Vit Machacek for sharing the data set with journal titles in Ulrich's Periodicals Database of all predatory journals in the Beall list., and the journal editor and two reviewers for their valuable comments on an earlier version of the manuscript.

CONFLICT OF INTEREST

Felix de Moya-Anegon and Vicente Guerrero-Bote are leading the Scimago Research Group who created and maintains the Scimago Journal Rank (SJR) that presents bibliometric indicators of all journals indexed in Scopus. Henk Moed was a former senior scientific advisor to Elsevier (Amsterdam, 2010–2014), scientific advisor to Scimago Research Group, and editor-in-chief of a scholarly open access journal.

REFERENCES

Bautista-Puig, N., Lopez-Illescas, C., de Moya-Anegon, F., Guerrero-Bote, V., & Moed, H. F. (2020). Do journals flipping to gold open access show an OA citation or publication advantage? *Scientometrics*, *124*, 2551–2575. https://doi.org/10.1007/s11192-020-03546-x

Beall, J. (2016). Scholarly Open Access: Critical analysis of scholarly open-access publishing (Beall's blog). Retrieved from https://scholarlyoa.com; shutdown January 2018, archived at https://archive.org/web/.

Learned Publishing published by John Wiley & Sons Ltd on behalf of ALPSP.

Beall, J. (2017). "Beall's list: Potential, possible, or probable predatory scholarly open-access publishers". Scholarly open access (last archived edition). Archived from the original on January 12, 2017.

Björk, B.-C., Kanto-Karvonen, S., & Harviainen, J. T. (2020). How frequently are articles in predatory open access journals cited. *Publication*, *8*(2), 17. https://doi.org/10.3390/publications8020017

Cohen, A. J., Patino, G., Kamal, P., Ndoye, M., Tresh, A., Mena, J., Butler, C., Washington, S., & Breyer, B. N. (2019). Perspectives from authors and editors in the biomedical disciplines on predatory journals: Survey study. *Journal of Medical Internet Research*, *21*(8), e13769. https://doi.org/10.2196/13769

Frandsen, T. F. (2017). Are predatory journals undermining the credibility of science? A bibliometric analysis of citers. *Scientometrics*, *113*, 1513–1528. https://doi.org/10.1007/s11192-017-2520-x

Frandsen, T. F. (2019). How can a questionable journal be identified: Frameworks and checklists. *Learned Publishing*, *32*(3), 221–226. https://doi.org/10.1002/leap.1230

Glanzel, W., Moed, H. F., Thelwall, M., & Schmoch, U. (Eds.) (2019). Measuring science: Basic principles and application of advanced Bibliometrics. In *Springer handbook of science and technology indicators*. Springer. https://doi.org/10.1007/978-3-030-02511-3_10

Grudniewicz, A., Moher, D., Cobey, K. D., Bryson, G. L., Cukier, S., Allen, K., Ardern, C., Balcom, L., Barros, T., Berger, M., Ciro, J. B., Cugusi, L., Donaldson, M. R., Egger, M., Graham, I. D., Hodgkinson, M., Khan, K. M., Mabizela, M., Manca, A., … Lalu, M. M. (2019). Predatory journals: No definition, no defence. *Nature*, *576*(7786), 210–212. https://doi.org/10.1038/d41586-019-03759-y

Kimotho, S. G. (2019). The storm around Beall's list: A review of issues raised by Beall's critics over his criteria of identifying predatory journals and publishers. *African Research Review*, *13*(2), 1–11. https://doi.org/10.4314/afrrev.v13i2.1

Krawczyk, F., & Kulczycki, E. (2021). How is open access accused of being predatory? The impact of Beall's lists of predatory journals on academic publishing. *The Journal of Academic Librarianship*, *47*, 102271. https://doi.org/10.1016/j.acalib.2020.102271

Larivière, V., & Sugimoto, C. R. (2019). The journal impact factor: A brief history, critique, and discussion of adverse effects. In W. Glanzel, H. F. Moed, M. Thelwall, & U. Schmoch (Eds.), *Springer handbook of science and technology indicators*. Springer. https://doi.org/10.1007/978-3-030-02511-3_10

Macháček, V., & Srholec, M. (2021). RETRACTED ARTICLE: Predatory publishing in Scopus: Evidence on cross-country differences. *Scientometrics*, *126*, 1897–1921. https://doi.org/10.1007/s11192-020-03852-4 Retraction Note at Scientometrics (2021). https://doi.org/10.1007/s11192-021-04149-w

Moed, H. F. (2016). Toward new indicators of a Journal's manuscript peer review process. *Frontiers in Research Metrics and Analytics*, *1*, 5. https://doi.org/10.3389/frma.2016.00005

Moed, H. F. (2017). *Applied evaluative Informetrics*.(ebook) (p. 312). Springer. https://doi.org/10.1007/978-3-319-60522-7

Moustafa, K. (2015). Fake journals: Not always valid ways to distinguish them. *Science and Engineering Ethics*, *21*(5), 1391–1392. https://doi.org/10.1007/s11948-014-9608-y

Oermann, M. H., Nicoll, L. H., Carter-Templeton, H., Woodward, A., Kidayi, P. L., Browning Neal, L., Edie, A. H., Ashton, K. S., Chinn, P. L., & Amarasekara, S. (2019). Citations of articles in predatory nursing journals. *Nursing Outlook*, *67*(6), 664–670. https://doi.org/10.1016/j.outlook.2019.05.001

Oviedo-García, M. A. (2021). Journal citation reports and the definition of a predatory journal: The case of the multidisciplinary digital publishing institute (MDPI). *Research Evaluation.*, *30*, 405–419. https://doi.org/10.1093/reseval/rvab020

Reedijk, J., & Moed, H. F. (2008). Is the impact of journal impact factors decreasing? *Journal of Documentation*, *64*, 183–192. https://doi.org/10.1108/00220410810858001

Scopus Content Policy (n.d.). https://www.elsevier.com/solutions/scopus/how-scopus-works/content/content-policy-and-selection

Shen, C. Y., & Bjork, B. C. (2015). 'Predatory' open access: A longitudinal study of article volumes and market characteristics. *BMC Medicine*, *13*, 15. https://doi.org/10.1186/s12916-015-0469-2

Somoza-Fernández, M., Rodríguez-Gairín, J. M., & Urbano, C. (2016). Presence of alleged predatory journals in bibliographic databases: Analysis of Beall's list. *El profesional de la información*, *25*, 5. https://doi.org/10.3145/epi.2016.sep.03

Wallace, F. H., & Perri, T. J. (2018). Economists behaving badly: Publications in predatory journals. *Scientometrics*, *115*, 749–766. https://doi.org/10.1007/s11192-018-2690-1

Xia, J., Harmon, J. L., Connolly, K. G., Donnelly, R. M., Anderson, M. R., & Howard, H. A. (2015). Who publishes in "predatory" journals? *Journal of the Association for Information Science and Technology*, *66*(7), 1406–1417. https://doi.org/10.1002/asi.23265

