Research article

Bibliometrics as a promising tool for solving publication ethics issues☆

Vadim N. Gureyev [a,b,*], Nikolay A. Mazov [a,c]

[a] *State Public Scientific Technological Library, Siberian Branch, Russian Academy of Sciences, Novosibirsk, Russia*
[b] *Vector State Research Center of Virology and Biotechnology, Koltsovo, Novosibirsk Region, Russia*
[c] *Trofimuk Institute of Petroleum Geology and Geophysics, Siberian Branch, Russian Academy of Sciences, Novosibirsk, Russia*

ARTICLE INFO

ABSTRACT

Publication ethics principles became one of the main aspects of conducting scientific research and presenting its results. Publication ethics challenges cover a wide range of problems of varying importance that involve all participants of publication processes: authors, academic authorities, peer-reviewers, editorial board members, publishers, and funders. All stakeholders put efforts to make modern science and publication processes ethical. This goal is achieved first of all through detailed criteria of publication ethics and extensive author guidelines, as well as by increasing the level of awareness of these criteria in educational programs aimed at prophylactics of research misconduct. However, there is a need for technical facilities for detecting different cases of violation of ethical principles, and bibliometric methods are one of the most promising approaches. The paper summarizes the authors' recent studies on bibliometric perspectives for detecting plagiarism, inappropriate authorship, and official misconduct among editorial board members.

1. Introduction

Fundamental principles of publication ethics have been actively developing in the last 30 years. The last three decades were marked by a complication of carrying out scientific studies and the structure of interactions between researchers and research teams. This increased the number of authors in the byline, led to the rise of copyrights for different parts of publications, the emergence of issues of responsibilities for one or another part of the paper. Concurrently, the problem of research misconduct among scientists was emerged, primarily due to shortcomings in the system the government of science aimed at increasing the scholarly output. Some responsibility lies with academic journals because of their discriminatory policy towards young researchers forcing them to use inappropriate authorship models. These and other aspects negatively affect the quality of research papers, the spreading of predatory journals, and the use of the most dangerous issues of plagiarism, falsification, and fabrication of data.

Among others, the following ethical issues have become prevalent:

- plagiarism and self-plagiarism, especially veiled cases of translated plagiarism. The detection of incorrect borrowing in academic texts has become an international problem in the past years;

- inappropriate authorship implying the adding the persons in the byline who fail to satisfy authorship requirements or exclusion the real researchers from the byline;
- artificial increase of bibliometric indices by negotiated citations, excess of power by editorial board members in parent journals, etc.

These and other types of scholarly misconduct lead to heightened interest of researchers, librarians, and publishers in publication ethics issues. Despite significant and regular improvements and reinforced publication ethics standards, the limitation of current methods for detection of their violation should be mentioned. Therefore, additional tools for the detection of scholarly misconduct are needed, and bibliometric approaches are believed to be promising ones. As bibliometrics is usually developed in research libraries and is actively used there for variety of purposes [1, 2], we should point at an additional value of librarians' role in the assessment of analyzed issues due to their professional independence and interdisciplinary nature of their work resulting in unbiased judgment [3].

It is believed that bibliometric evaluation of publication ethics issues can result in strengthening the principles of integrity in publication processes, as well as an increase in the share of original studies. The results of detection of violation of ethical criteria are important in all

---

research areas, especially in medical sciences since any misconduct in this field directly negatively affects the quality of the healthcare system. Several notable cases of the detection of unethical behavior among medical authors were described in chapter 3 of CSE's White Paper on Promoting Integrity in Scientific Journal Publications [4].

The main goal of this paper is to demonstrate the high potential of bibliometric tools in the detection of a wide range of publication ethics issues and highlight the systematic basis of the use of bibliometrics in this field as opposite to current sporadic attempts in a bibliometric solution of mission tasks.

The rest of the article depicts three main bibliometric directions developed by the authors in the last years as applied to publication ethics issues including the detection of translated plagiarism (Section 2), inappropriate types of authorship (Section 3), and misconduct among editorial board members (Section 4).

2. Bibliometric approach for the detection of translated plagiarism in academic papers

In recent times, the detection of plagiarism in scholarly papers became an international task since it assume more and more veiled shapes in response to modern techniques of its detection. It accounted for almost half of all misconduct retractions and significantly affects non-English low-income countries as well as low-ranked journals being one of the main obstacles for their development [5]. One of the main reasons causing plagiarism (as well as some other types of violation of publication ethics principles) comprises pressure to publish in order to increase scholarly output indicators necessary to research career, proficiency testing, grants application, etc. Together with searching paths for the elimination of the causes of plagiarism phenomenon, the development of the tools for its detection is another important task.

Linguistic analysis of full texts demonstrates good results in the detection of copy and paste type of plagiarism. Recent studies engaging linguistic processing of full texts including morphological, syntax, and semantic analyses showed effectiveness even for plagiarism with significant paraphrasing [6]. At the same time, some types of plagiarism are still difficultly detected, especially translated forms.

The detection of translated plagiarism now is possible with the use of (a) expert community, (b) automated systems of machine-aided translation and neural networks [7], (c) bibliometric approaches. Expert evaluation is the most limited since it is time-consuming and requires great human resources. As for machine-aided translation enabling one to detect translated plagiarism, it can be effective in case of the predominance of open-access model due to critical requirement to full texts. Unless great volumes of research papers are inaccessible for robotic processing because of paid nature of academic papers collections, such systems will be low effective due to the limited free collections for texts comparison. Therefore, even in the case of functional capability of the detection system, a great deal of plagiarism cases will be undetected. The bibliometric approach to solving the problem seems to be the most promising and balanced.

2.1. Methodology for detecting translated plagiarism

The method is based on searching papers with identical or closely related lists of references, especially texts with a similar sequence of references so that the more recent paper may contain possible plagiarism. Such an approach implies the analysis of only cited references and their sequences that enables one to disregard texts and evade a problem of comparison of texts in different languages. Besides, it requires only generally accessible metadata including cited references, and not the full texts.

This approach was for the first time proposed by B. Gipp et al. [8, 9, 10, 11]. They developed web-service HyPlag (https://www.hyplag.org/) promoted as a prototype plagiarism detection system in addition to text-based software. The benefits of the system include text- and language-independent fingerprints (references) making it possible to detect strongly disguised cases of plagiarism. However, the prototype seems to use only free text collections, can be used for comparison of two or more texts uploaded to the system, but cannot find the original text in subscription-based collections.

With small delay we also came to the same conclusions on the high potential of citation analysis in translated plagiarism detection [12]; however, in our study, we decided to use wide capacities of bibliographic databases, e.g., Web of Science or Scopus. Their use as opposite to full-text databases may significantly enlarge the factual basis for analysis and makes an opportunity to automatically generate search queries in a bibliographic system based on cited references in the analyzed text.

Citation analysis for plagiarism detection is based on the bibliographic coupling proposed by M. Kessler [13, 14]. Two papers are considered to be closely related and bibliographically connected if their reference lists have the same items, i.e., bibliographic units. The coupling strength is expressed by the volume of identical references in two texts. The main features of bibliographic coupling are the independence of the language and the possibility to automate the process of searching closely related items. Therefore, this approach can be directly used to the detection of disguised plagiarism. In that case, the more recent paper with the same or similar reference list of older paper can be regarded as an item potentially containing plagiarism.

The methodological process of searching for an authentic source for suspicious text in bibliographic databases includes the following steps (Figure 1).
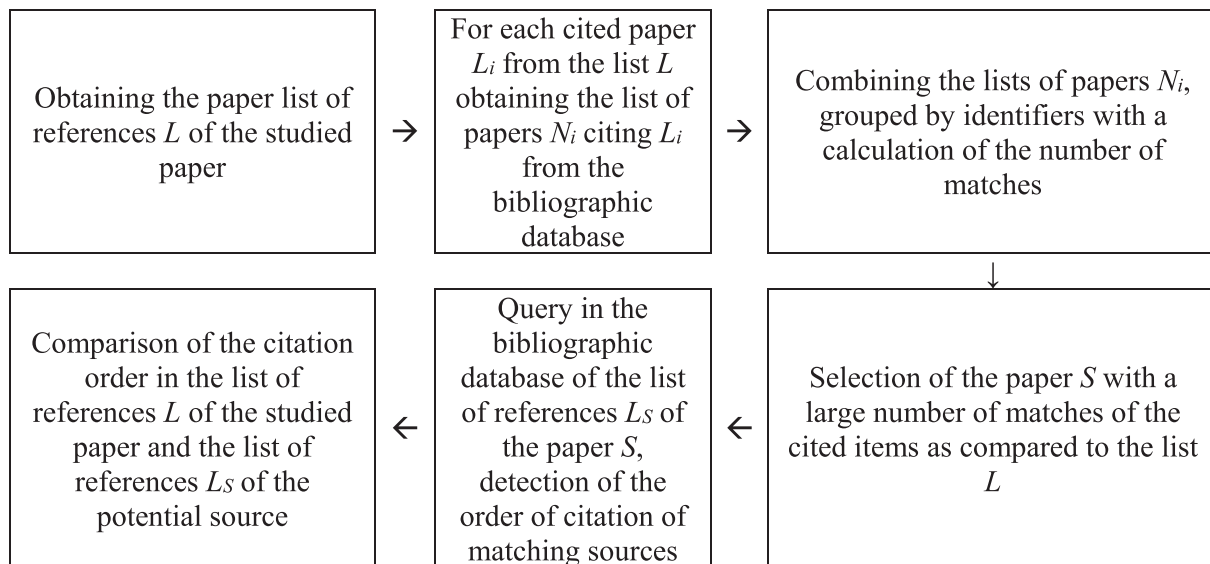
1. For each cited item from the reference list of suspicious paper, one should form a search query in a bibliographic database to receive the list of papers that also cited this item. The search query may include such metadata as authors, title, publication year, source title, page number, DOI, etc. The further data export uses publication IDs, e.g., "eid" in Scopus or "ut" in Web of Science for automated data processing in Excel.
2. After that alphabetic sorting of downloaded IDs of papers cited the same references as the suspicious paper is made with the final counting of matching.
3. Descending order of matching items is further analyzed for detecting possible plagiarism. This list can be reduced by cutting off irrelevant items (e.g., less than 10 items or 50 percent).

2.2. Study results of the detection of translated plagiarism

Figures 2 and 3 depict a specific examples of the effectiveness of the proposed algorithm permitting us to detect a paper with 20 sources cited in the same order as in the analyzed suspicious paper. An inverse search query using the identifier in the Web of Science easily permits detection of the original source. As is clear from Figure 3, the analyzed suspicious article includes only one third of the references in common with the other source, while two thirds of it were apparently written by the authors. At the same time, the total sequence of the references in the suspicious part of the paper indicates that text fragments of the two publications with common references will be also similar with a high degree of probability.

In papers [12, 15, 16] we demonstrated specific results of the proposed approach that enabled us to detect translated plagiarism cases in research articles, reviews, Ph.D. theses, monographs, and research reports in fields of Library and Information Science, Computer Sciences, and Scientometrics. The model of translated plagiarism detection based on comparison of reference lists and sequence of references demonstrated efficiency and can be automated. Algorithms of the model can be integrated into plagiarism-detection software. The wide use of such a system is believed to result in decreasing the volume of translated plagiarism and stimulate the increase of original studies.

Another possible application of bibliometric knowledge to solving publication ethics issues may include the detection of inappropriate types of authorship.

Obtaining the paper list of references L of the studied paper	$\rightarrow$	For each cited paper L_i from the list L obtaining the list of papers N_i citing L_i from the bibliographic database	$\rightarrow$	Combining the lists of papers N_i, grouped by identifiers with a calculation of the number of matches
Comparison of the citation order in the list of references L of the studied paper and the list of references L_S of the potential source	$\leftarrow$	Query in the bibliographic database of the list of references L_S of the paper S, detection of the order of citation of matching sources	$\leftarrow$	Selection of the paper S with a large number of matches of the cited items as compared to the list L

Figure 1. Search query process in a bibliographic database for searching possible original source for suspicious paper. L – reference list of the analyzed paper; L_i – each item from the reference list of the analyzed paper; N_i – list of papers from general database that also cited L_i; S – the paper from N_i list with the most similar reference list as compared to that of the analyzed paper; L_S – reference list of S paper.

3. Bibliometric approach for the detection of inappropriate types of authorship

Authorship is one of the main sources of academic "capital" of researchers. Under the conditions of competition for academic positions and funding the authorship is currently believed to be the key indicator of the academic capabilities and research potential of a scientist [17]. Changes in the approaches to the generation of knowledge, a complication of the conducted research, increase in interdisciplinarity [18],

enlarged collaborations [19] leading to hyperauthorship phenomenon [20] together with formal approaches to evaluate study results increased the need to publish for funding and career progression [21]. One of the negative consequences of such state of the art includes the spreading of inappropriate types of authorship comprising guest, gift, honorary, sold, and ghost types [17, 22].

Guest authorship involves mutual support of the authors and cases when the authors include each other in their papers in order to increase their number [17]. Gift authorship is usually seen as a sign of gratitude

	Paper's identifiers in Web of Science, citing the same sources as the suspicious paper	The sequence number of the cited publication in the list of references of the suspicious paper	
Papers with matching of only one cited source were not considered	72783300004	43	
	72848500067	44	
	72947100003	37	
	73082400040	44	
	73090100127	37	
	73225800046	61	
Identifiers of the papers citing two or more sources from the list of references of the suspicious publication were combined	73262000005	34	
	73262000005	36	

73262000005	34	36	
74160500021	14	16	
88670100008	43	12	19

Figure 2. The stages of processing of the paper lists of references where the same sources are cited as in a suspicious publication.

Identifiers of the papers in the Web of Science citing the same sources as the suspicious paper	Sequence number of the cited papers in the list of references of the suspicious paper																									
278518900002	38	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63						

Figure 3. The identifier of the paper in Web of Science with the largest number of sources (20) also present in the list of references of the analyzed suspicious paper.

and recognition from younger researchers in relation to their mentors. Honorary authorship is the most common type of unfair authorship implying an act of symbolic violence toward actual authors to force them to add an honorary author in the byline of the manuscript. It is primarily affects young researchers [23]. Sold authorship is the most significant violation of the criteria for authorship and involves the payment for a place in the byline [24]. Ghost authorship is the opposite of previous types and associated with the absence of the name of the actual author in the byline. This practice is most prevalent in the field of biomedical sciences and in pharmaceutics [22], where junior medical personnel often plays the role of ghost authors [25].

Thus, inappropriate authorship is social, ethical, and management problem being one of the types of simulation of research activity [26] resulting in a decrease of a share of quality scientific content [27]. Unethical authorship infers mismatch of some authors in the byline with international authorship criteria. There are several approaches to define the authors, as well as radical proposals to change authorship with contributors' model (CRediT) [28]. The most widely accepted principles were developed by the International Committee of Medical Journal Editors (ICMJE) [29]:

- Substantial contributions to the conception or design of the work; or the acquisition, analysis, or interpretation of data for the work; AND
- Drafting the work or revising it critically for important intellectual content; AND
- Final approval of the version to be published; AND
- Agreement to be accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved.

As a rule, the unjustified inclusion of some authors in the byline is caused by the personal advantages of some researchers for increasing scholarly output which is stipulated by the objective reasons of shortcomings of the decision-making system in the science. Thus, junior researchers are often forced by senior colleagues to include the latter in the byline [23]. As the paper is one of the main items, and funding and career progress are based to a great extent on scholarly output [30, 31], the condition of institutional pressure of university or research organization on researchers takes place leading to violation of authorship guidelines.

Current approaches to detect authorship misconduct are primarily based on the principles of publication ethics including author guidelines and detailed criteria [32, 33]. The development of methods for the detection of inappropriate authorship is carried out in three directions:

1. Editorial and publishing approach includes analysis of author's practices and experience, the definition of ethical principles, and specifying authorship criteria [4]. Providing that most issues fall into one of two categories, excluding deserving contributors or including undeserving ones, the problem is that there is no consensus on "deserving" despite ICMJE and other associations' criteria [34].
2. The sociological approach comprises the detection of the level of spreading of misconduct, its reasons, and consequences, as well as the development of prevention approaches [35, 36, 37]. One of the efficient but rarely used preventive measures includes explicit discussion about co-authoring credit before the start of the research since

collaboration practices that are viewed as routine in some fields are viewed as unethical in others [34].
3. The bibliometric approach engages the detection of deviations from the expected distribution of indicators characterizing scholarly output of a researcher [38, 39, 40].

3.1. Methodology for detecting inappropriate types of authorship

Our bibliometric approach for the detection of inappropriate authorship was described in [40, 41] and is based on searching anomalies in the distribution of several bibliometric indices of the analyzed author in different stages of his/her career. In some periods, sharp bursts of bibliometric indicators may be reasonable while in other periods they may raise suspicions in compliance with ethical principles.

Some reasons for the sharp increase of scholarly output and relevant indices may include Ph.D. thesis defense requiring several published articles; gaining the position of editorial board member implying publication of editorial materials [42] or participation in a research program with special funding. However, in some cases such as promotion to a higher career position rapid increase in the volume of bibliometric indices may be suspicious and require special analysis. In our study, we analyzed the distribution of a set of bibliometric indices in three significant career stages, i.e., Ph.D. theses defense, obtaining high academic status, and career promotion.

To detect possible cases of violation of authorship criteria, we propose publication coefficients measuring sets of indicators in equal periods before and after the event in career progress. For instance, in the case of being a head of a research organization for 6 years, we compared a set of indices for 6 years before and 6 years after the appointment at the leading position. As for Ph.D. theses defense and gaining high academic ranks, we used equal 3-years lags. Publication coefficient K_p^1 was calculated as follows:

$$K_p^1 = \frac{P_1}{P_1 + P_2}$$

where P_1 denotes the number of papers before an appointment to the leading position, P_2 is a number of papers at the time of holding an appointment. Similarly, we detected publication coefficient at the time of holding an appointment:

$$K_p^2 = \frac{P_2}{P_1 + P_2}$$

3.2. Study results of the detection of inappropriate types of authorship

Using a sample of 39 the most prominent researchers of the Siberian Branch of the Russian Academy of Sciences, we revealed that 90 percent of researchers from the analyzed sample significantly increased scholarly output and other bibliometric indicators after their appointment to leading positions (Figures 4 and 5). Only in 10 percent cases, a negative trend was detected; furthermore, the decrease in the number of papers was not as much expressed as rapid growth after the appointment.

Subject areas coefficients were calculated as that of publications coefficients. It was detected that in 74 percent of papers topic variety
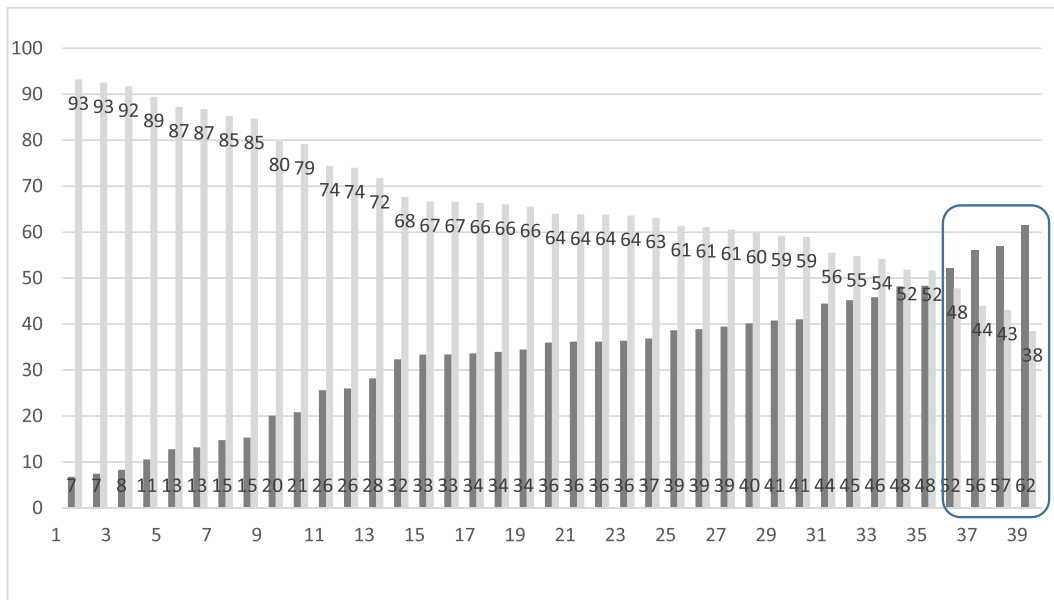
Figure 4. Publication coefficients of researchers before and after appointment to leading positions. Dark grey denotes publication coefficient before an appointment; light grey – after an appointment. Unusual cases of higher scholarly output before appointment are boxed.
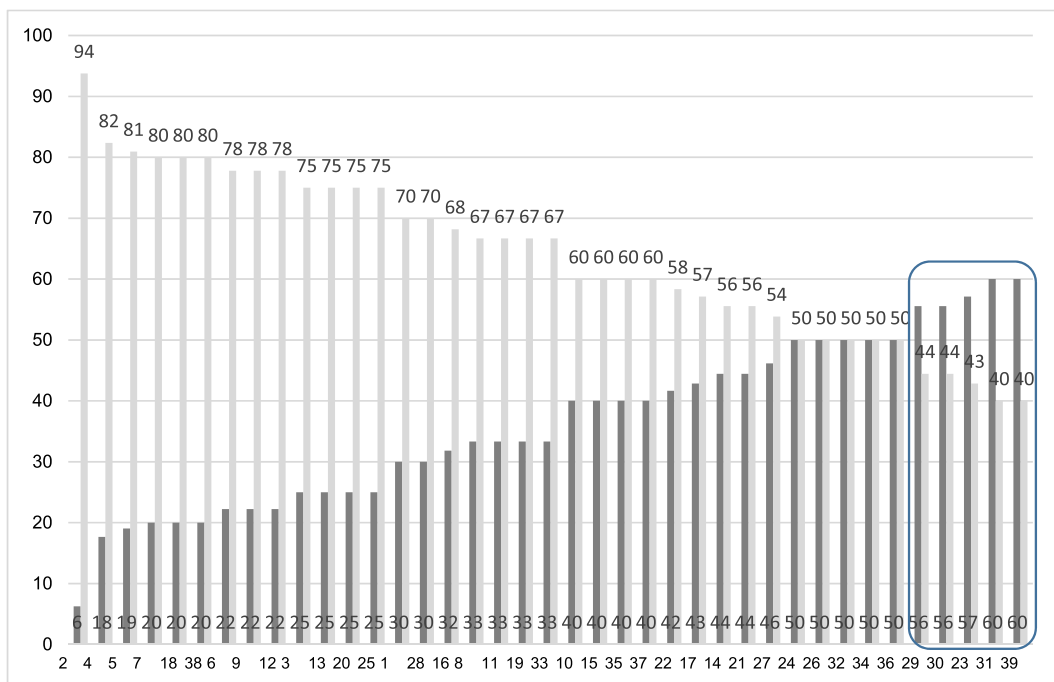
Figure 5. Dynamics pattern of subject areas in publications by scientists before and after their appointment to leading positions. Dark grey denotes subject areas coefficient before an appointment to leading positions, while light grey coefficient after an appointment. Unusual cases of wider subject distributions before an assignment are boxed.

significantly increased (Figure 5). One can hardly conceive that gaining a leading position can transform a researcher into an expert in different subject areas. In any case, it is difficult to be responsible for the contribution elements of various disciplines according to one of the authorship criteria. Finally, we analyzed the number of co-authors. Again, in 87 percent the number of co-authors significantly increased.

Considering the large administrative load of scientists holding leading positions resulting in shortening the free time for research, an increase in the number of papers is believed to be achieved exclusively by means of co-authorship. It is associated mainly with the inclusion of an executive in the byline as co-author. As mentioned above, the reasons can be different including supervision in the grants, the teaching of young scientists, the inclusion of prominent names to speed up peer review stages, etc. At the same time, in the last two decades, requirements for authorship have become stricter; thus, authors must fulfill all criteria.

Our findings enabled us to assume that the assignment of scientists to leading position frequently leads to violation of publication ethics regarding international authorship criteria since the use of guest of gift authorship seems to be rampant in some research communities. It is confirmed by a very intensive increase in the number of papers, accompanied by an increase in the list and number of co-authors and a significant enhancement of subject areas. Besides, in some organizations
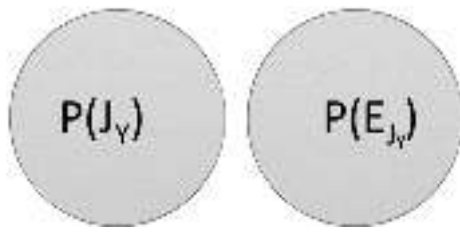
Figure 6. The absence of editorial board members' papers in parent journal.

especially medical ones we detected an unexpectedly high number of papers per year close to 100 items.

Regarding a high administrative burden, it is highly unlikely that a scientist would have enough time for publishing such a high number of papers and concurrently meet modern requirements of authorship. We have detected that sharp fluctuations in scholarly output can sometimes point to possible misconduct in publishing and fictive participation in the research. Especially we mean a sharp increase in the number of papers, significant fluctuation in the number and compound of co-authors, changes in research areas, changes in position in the byline, increase in a pool of journals with scientist's papers.

Complex analysis of changes in bibliometric indicators and their comparison with certain changes in the career path of a scientist can be used as an additional bibliometric tool to reveal common factors of the real and fictive contribution of scientists in publications. Besides, we believe that bibliometric approaches are the most promising to detect misconduct towards authorship. Mainly it is caused by wide possibilities of generation of different types of metadata in bibliographic systems: current databases enable one to export funding information, Ph.D. theses defense, affiliation history, changes in research areas, changes in a set of journals where an author was published. Thus, the formation of a detailed publication profile of each author is now possible where one may detect sharp biases from expected values of bibliometric indices that possibly point at inappropriate authorship practices.

4. Bibliometric approach for the detection of possible misconduct among editorial board members

Publication strategies of editorial board members attract great research interest due to increased specification and strictness of publication ethics criteria [43]. The importance of studying publication interactions of editorial board members with parent journals is highlighted by their high authority as compared with other authors and a high degree of responsibility for providing integrity of publication process. Therefore, editorial board members should strictly adhere to publication ethics principles. However, sometimes ethical suspicions may concern even editorial board members. For instance, they may suggest authors cite their own papers unjustified by the topic of the manuscript, redundantly submit papers to the parent journal or, oppositely, ignore parent journal publishing papers in other sources that may indicate their formal engagement in the editorial board.

4.1. Methodology for detecting misconduct among editorial board members

Although there are no uniform recommendations on the acceptable proportion of papers by editorial board members in parent journal, two cases including full absence and overrepresentation of editorial board members' papers in parent journal may point to a violation of publication ethics. The case of the absence of papers by editorial board members in parent journal can be formulated as follows (Figure 6):

$$P(J_Y) \cap P(E_{J_Y}) = \varnothing$$

where $P(J_Y)$ – papers P of the journal J in a year Y;

$P(E_{J_Y})$ – papers P by editorial board members E of the journal J in a year Y;

$P(E_{JS_Y})$ – papers P of editorial board members E of the journal J in a parent journal in a year Y.

In rare circumstances, this situation can be ethical. For instance, the journal can employ professional editors on a fee basis and charge them with a full complex of journal duties including peer review. As consequence, staff members frequently should retire from a university and stop publishing. Another case may include the editor-in-chief's decision to prohibit editorial board members to publish in a parent journal to evade any conflict of interest [44].

More often, the complete absence of editorial board members' papers may denote a formal engagement of prominent experts in the editorial board to increase the rank and the level of internationalization of the journal. Notably, the editor-in-chief does not require any activity from such members. Formal engagement of eminent researchers in editorial boards is frequently observed in predatory journals to lend them legitimacy in the opinion of potential authors [45]. This situation can be detected by bibliometric tools and should be regarded as an unethical one. Rarely the absence of papers by editorial board members in parent journal can be explained by the absence of the editorial board itself. This case was described by J. Beall [46].

The case of the overrepresented volume of papers by editorial board members in parent journal can be formulated as follows (Figure 7):

A	$P(J_Y) = P(E_{JY}), \ \text{ГПе} P(E_{JY}) = P(E_{JSY})$
B	$P(J_Y) \supset P(E_{JY}), \ \text{ГПе} P(E_{JY}) = P(E_{JSY})$
C	$P(J_Y) \subset P(E_{JY}), \ \text{ГПе} P(J_Y) = P(E_{JSY})$

Rarely such a situation can be regarded as ethical. It may be connected with the absence of alternative journals for publishing, e.g., in narrow research areas [47]. The predominance of editorial board papers in parent journal may also be explained by the case of starting a new journal. As a new title is accountably unknown for potential authors, the first several issues are usually filled in by editorial board members' papers [48].

Nonetheless, other unethical cases of overrepresented papers by editorial board members are spread to a greater extent including local journals serving the staff of a university or research organization to execute a plan for target indices of scholarly output. Two versions of the
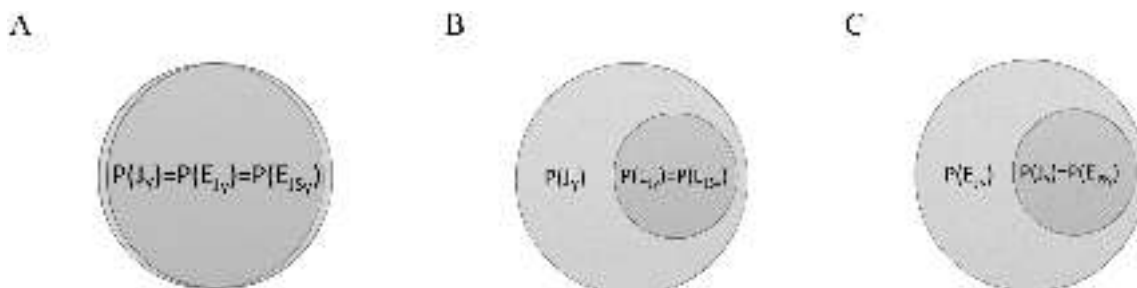

Figure 7. The overrepresented volume of editorial board papers in parent journal.

Table 1. Analyzed journals and their bibliometric indicators in the Scopus database. Summary data for 2008–2017.

Journal Title	Number of EMBs 2008–2017	Number of papers in the journal 2008–2017 $P(J_Y)$	Correction factor α	Number of EBMs papers in the journal 2008–2017 $P(E_{JS_Y})$	All EBMs papers 2008–2017 $P(E_{J_Y})$	Coefficient I
1. Izvestiya, Physics of the Solid Earth	21	843	1.219	186	488	**0.465**
2. Russian Geology and Geophysics	37	1238	1.000	301	1232	**0.244**
3. Geochemistry International	28	997	1.048	171	940	**0.191**
4. Stratigraphy and Geological Correlation	18	420	0.699	92	353	**0.182**
5. Lithology and Mineral Resources	19	351	0.539	99	338	**0.158**
6. Geology of Ore Deposits	18	577	0.940	73	566	**0.121**
7. Geotectonics	25	326	0.390	92	691	**0.052**
8. Petrology	19	339	0.540	73	817	**0.048**

situation of overrepresented papers by editorial board members in parent journal include the following cases. Editorial board members prefer to use capacities of parent journal that at the same time is open for and known among other authors. In that case, editorial board members escape from publishing in other sources and prefer to assuredly publish their results in domestic journal (Figure 7B). The other case implies the rather high authority of editorial board members publishing their results in both parent journal and other sources (Figure 7C). In this case, the quality of papers in parent journal and other sources should be compared, since editorial board members may use their official power to simplify or speed up publication processes in parent journal [49].

Bibliometric capacities for the detection of possible cases of violation of publication ethics principles by editorial board members may be used as follows. The above-mentioned extreme cases may be expressed as coefficient I that considers the volume of all papers by editorial board members over a certain period $P(E_{J_Y})$; and the volume of papers in a parent journal over a certain period $P(E_{JS_Y})$.

As all journals are distinguished by the volume of papers, issuers per year, and the volume of editorial boards, we introduce correction coefficient α that is sensitive to a share of papers in parent journal per one editorial board member. The use of correction coefficient is caused by the necessity to make equal the journals with greater issues per year (or greater number of papers) and serials with lower periodicity (or lower volume), as well as normalize journals with different volumes of editorial boards. The coefficient I can be calculated as follows:

$$I = \frac{|P(E_{JS_Y})|}{|P(E_{J_Y})|} \times \alpha$$

where $P(E_{JS_Y})$ is a number of papers by editorial board members in parent journal; $P(E_{J_Y})$ is a total number of papers by editorial board members; α is a correction coefficient accounting for the volume of journal and editorial board.

4.2. Study results of the detection of misconduct among editorial board members

For the practical application of the coefficient of publication relationships between the journal and the editorial board, eight of the most authoritative international journals on Earth sciences published in Russia were analyzed (Table 1) [50].

While all journals from the sample do not belong to the described types and demonstrate a balanced level of intersection between journal's and editorial board members' publication flows, we may suspect unethical formal engagement of editorial board members in several journals based on coefficient I and complex analyses of relevant bibliometric indices. For instance, "Geology of Ore Deposits" of all journals of this sample demonstrates the smallest intersection between the document flows of the journal and the editorial board, therefore this journal is close to type I possibly implying unethical formal work in editorial board. The lowest publication

ratio coefficients were found in "Geotectonics" and "Petrology" journals due to the imbalance of publication flows of the journal (a small number of papers) and the editorial board (a large number of papers). Both serials gravitate towards type I. The disproportion in the publication flows of the journal and the editorial board can also be explained either by the formal involvement of the editorial board members in the work in the journal. Thus, we may suspect some biases from the ethical composition of editorial boards even in the authoritative journals.

The proposed coefficient I is believed to be useful when highlighting the level of involvement of editorial board members in the development of parent journal, detecting possible cases of violation of publication ethics by abusing official power by publishing their papers in inadmissible quantity as described in [51], finding out the unethical cases of formal engagement of prominent researchers to editorial boards. Other studies carried out by the authors also demonstrate the practicability of bibliometrics in this aspect enabling us to detect both the cases of formal engagement of researchers in editorial boards [52] and the overrepresented number of papers in some serials [53].

5. Conclusion

The description of several possible techniques in this paper indicates that bibliometrics can be applied to a wide range of the detection of publication ethics issues. This is a rather new direction in the application of bibliometric tools, primarily used for the evaluation of research results. It should be noted, that different research teams have already perceived this promising direction of bibliometrics and demonstrated good results. The only limitation of such approaches includes the necessity for further and compulsory expert review of statistical bibliometric data to exclude false-positive results. Together with preventive measures to stop unethical practices using detailed guidelines for each stakeholder of publication processes and various training events to raise the awareness of publication ethics criteria, bibliometric approaches are considered as one of the additional valuable tools for increasing research integrity.

Declarations

Data availability statement

Data included in article/supplementary material/referenced in article.

Declaration of interests statement

The authors declare no conflict of interest.

Additional information

No additional information is available for this paper.

References

[1] J. Gorraiz, C. Gumpenberger, A flexible bibliometric approach for the assessment of professorial appointments, Scientometrics 105 (3) (2015) 1699–1719.

[2] G. González Alcaide, J.I. Gorraiz, Assessment of researchers through bibliometric indicators: the Area of information and library science in Spain as a case study (2001–2015), Front. Res. Metr. Anal. 3 (2018).

[3] R. Ball, D. Tunger, Bibliometric analysis – a new business area for information professionals in libraries? Scientometrics 66 (3) (2006) 561–577.

[4] CSE's white Paper on Promoting Integrity in Scientific Journal Publications, 2018 Update, Council of Science, Wheat Ridge, 2018, 81 p.

[5] S. Stretton, N.J. Bramich, J.R. Keys, J.A. Monk, J.A. Ely, C. Haley, et al., Publication misconduct and plagiarism retractions: a systematic, retrospective study, Curr. Med. Res. Opin. 28 (10) (2012) 1575–1583.

[6] I. Sochenkov, D. Zubarev, I. Tikhomirov, I. Smirnov, A. Shelmanov, R. Suvorov, et al., Exactus Like: plagiarism detection in scientific texts, in: Advances in Information Retrieval: 38th European Conference on IR Research, ECIR 2016, Padua, Italy, 20–23 March 2016, Springer International Publishing, Cham, 2016, pp. 837–840.

[7] O. Bakhteev, Y. Chekhovich, G. Gorbachev, T. Gorlenko, A. Grabovoy, K. Grashchenkov, et al., Cross-language plagiarism detection: a case study of European universities academic works, in: 7th European Conference on Academic Integrity and Plagiarism, Brno, Czechia, 9–11 June 2021, Brno, 2021, pp. 14–15.

[8] B. Gipp, N. Meuschke, C. Breitinger, Citation-based plagiarism detection: practicability on a large-scale scientific corpus, J. Asso. Infor. Sci. Technol. 65 (8) (2014) 1527–1540.

[9] B. Gipp, N. Meuschke, C. Breitinger, M. Lipinski, A. Nürnberger, Demonstration of citation pattern analysis for plagiarism detection, in: 36th International ACM SIGIR Conference on Research and Development in Information Retrieval, SIGIR 2013, Dublin, Ireland, 28 July – 1 August 2013, Dublin, 2013, pp. 1119–1120.

[10] B. Gipp, N. Meuschke, Citation pattern matching algorithms for citation-based plagiarism detection: greedy citation tiling, citation chunking and longest common citation sequence, in: Proceedings of the 11th ACM Symposium on Document Engineering (DocEng '11), Mountain, View, CA, USA, 19–22 September 2011, Mountain, View, CA, USA, ACM, 2011, pp. 1–10.

[11] N. Meuschke, Analyzing Non-textual Content Elements to Detect Academic Plagiarism, University of Konstanz, Konstanz, 2021, 282 p.

[12] V.N. Gureyev, N.A. Mazov, Citation analysis as a basis for the development of an additional module in antiplagiarism systems, Sci. Tech. Inf. Process. 40 (4) (2013) 264–267.

[13] M.M. Kessler, An experimental study of bibliographic coupling between technical papers, IEEE Trans. Inf. Theor. 9 (1) (1963) 49–51.

[14] M.M. Kessler, Comparison of the results of bibliographic coupling and analytic subject indexing, Am. Doc. 16 (3) (1965) 223–233.

[15] N.A. Mazov, V.N. Gureyev, D.V. Kosyakov, On the development of a plagiarism detection model based on citation analysis using a bibliographic database, Sci. Tech. Inf. Process. 43 (4) (2016) 236–240.

[16] N. Mazov, V. Gureyev, Study results for the detection of translated plagiarism using bibliometric databases, Sci. Tech. Librar. 12 (2017) 87–96.

[17] A. Olesen, L. Amin, Z. Mahadi, Unethical authorship practices: a qualitative study in Malaysian higher education institutions, Develop. World Bioeth. 18 (3) (2018) 271–278.

[18] W. Dang, M.D.F. McInnes, A.Z. Kielar, J.H. Hong, A comprehensive analysis of authorship in radiology journals, PLoS One 10 (9) (2015) 15.

[19] J. Youtie, B. Bozeman, Social dynamics of research collaboration: norms, practices, and ethical issues in determining co-authorship rights, Scientometrics 101 (2) (2014) 953–962.

[20] B. Cronin, Hyperauthorship: a postmodern perversion or evidence of a structural shift in scholarly communication practices? J. Am. Soc. Inf. Sci. Technol. 52 (7) (2001) 558–569.

[21] N.A. Mazov, V.N. Gureyev, Publications at any costs? Vestn. Ross. Akad. Nauk. 85 (7) (2015) 627–631.

[22] A.Y. Gasparyan, L. Ayvazyan, G.D. Kitas, Authorship problems in scholarly journals: considerations for authors, peer reviewers and editors, Rheumatol. Int. 33 (2) (2013) 277–284.

[23] A.A. Khalifa, Losing young researchers in the authorship battle, under-reported casualties, Ethics Med. Publ. Health 20 (2022) 100735.

[24] M. Hvistendahl, China's publication bazaar, Science 342 (6162) (2013) 1035–1039.

[25] J.M. Street, W.A. Rogers, M. Israel, A.J. Braunack-Mayer, Credit where credit is due? Regulation, research integrity and the attribution of authorship in the health sciences, Soc. Sci. Med. 70 (9) (2010) 1458–1465.

[26] A.V. Kuleshova, D.G. Podvoyskiy, Paradoxes of publication activity in the field of contemporary Russian science: genesis, diagnosis, trends, Monit. Publ. Opin. Econ. Soc. Chang. (4) (2018) 169–210.

[27] V.N. Gureyev, I.G. Lakizo, N.A. Mazov, Unethical authorship in scientific publications (A review of the problem), Sci. Tech. Inf. Process. 46 (4) (2019) 219–232.

[28] L. Allen, A. Brand, J. Scott, M. Altman, M. Hlava, Credit where credit is due, Nature 508 (7496) (2014) 312–313.

[29] ICMJE, Defining the Role of Authors and Contributors, 2022. Available at: http://www.icmje.org/recommendations/browse/roles-and-responsibilities/defining-the-role-of-authors-and-contributors.html#two. (Accessed 10 January 2022).

[30] D. Shaw, The prisoners' dilemmas: authorship guidelines and impact factors: between a rock and a hard place, EMBO Rep. 15 (6) (2014) 635–637.

[31] M. Hosseini, B. Gordijn, A review of the literature on ethical issues related to scientific authorship, Account. Res. 27 (5) (2020) 284–324.

[32] L. Bošnjak, A. Marušić, Prescribed practices of authorship: review of codes of ethics from professional bodies and journal guidelines across disciplines, Scientometrics 93 (3) (2012) 751–763.

[33] V. Gureev, N. Mazov, I. Lakizo, Authorship criteria and the problem of its attribution in scholarly papers, Sci. Tech. Librar. 11 (2019) 5–24.

[34] B. Bozeman, J. Youtie, Trouble in paradise: problems in academic research co-authoring, Sci. Eng. Ethics 22 (6) (2016) 1717–1743.

[35] W. Al-Herz, H. Haider, M. Al-Bahhar, A. Sadeq, Honorary authorship in biomedical journals: how common is it and why does it exist? J. Med. Ethics 40 (5) (2014) 346–348.

[36] A. Mirzazadeh, S. Navadeh, M.B. Rokni, M. Farhangniya, The prevalence of honorary and ghost authorships in iranian bio-medical journals and its associated factors, Iran. J. Publ. Health 40 (1) (2011) 15–21.

[37] R.L. Eisenberg, L.H. Ngo, A.A. Bankier, Honorary authorship in radiologic research articles: do geographic factors influence the frequency, Radiology 271 (2) (2014) 472–478.

[38] S.S. Hwang, H.H. Song, J.H. Baik, S.L. Jung, S.H. Park, K.H. Choi, et al., Researcher contributions and fulfillment of ICMJE authorship criteria: analysis of author contribution lists in research articles with multiple authors published in Radiology, Radiology 226 (1) (2003) 16–23.

[39] A.A. Rostovtsev, The Russian and international practices of identification of irresponsible journals and authors, Sci. Edit. Publ. 2 (1) (2017) 30–37.

[40] N.A. Mazov, V.N. Gureyev, Detection of inappropriate types of authorship using bibliometric approaches, in: 17th International Conference on Scientometrics and Informetrics, ISSI 2019, Rome, Italy, 2–5 September 2019, vol. 1, Edizioni Efesto, Rome, 2019, pp. 885–895.

[41] V.N. Gureyev, N.A. Mazov, A.A. Ilyichev, Career path of researchers in relation to publication ethics, Vestn. Ross. Akad. Nauk. 89 (3) (2019) 270–278.

[42] Research Metrics Guidebook, Elsevier, 2018. Available at: https://www.elsevier.com/research-intelligence/resource-library/research-metrics-guidebook. (Accessed 20 October 2021).

[43] N.A. Mazov, V.N. Gureyev, The editorial boards of scientific journals as a subject of scientometric research: a literature Review, Sci. Tech. Inf. Process. 43 (3) (2016) 144–153.

[44] S. Youk, H.S. Park, Where and what do they publish? Editors' and editorial board members' affiliated institutions and the citation counts of their endogenous publications in the field of communication, Scientometrics 120 (3) (2019) 1237–1260.

[45] J. Beall, What I learned from predatory publishers, Biochem. Med. 27 (2) (2017) 273–278.

[46] J. Beall, Criteria for Determining Predatory Open-Access Publishers, 2015. Available at: https://beallslist.net/wp-content/uploads/2019/12/criteria-2015.pdf. (Accessed 20 October 2021).

[47] J.M. Campanario, The competition for journal space among referees, editors, and other authors and its influence on journals' impact factors, J. Am. Soc. Inf. Sci. 47 (3) (1996) 184–192.

[48] O.V. Moskaleva, M.A. Akoev, Forecast of the development of Russian scientific journals: the publishers, Schol. Res. Inform. 3 (2-3) (2020) 131–154.

[49] Z. Taşkın, A. Taşkın, G. Doğan, E. Kulczycki, Editorial board member and prolific author status positively shorten publication time, in: 18th International Conference on Scientometrics and Informetrics, ISSI 2021, Leuven, Belgium, 12–15 July 2021, Leuven, 2021, pp. 1081–1089.

[50] N.A. Mazov, V.N. Gureyev, Publication relationships between editorial board members and parent journals in the context of research ethics, Sci. Tech. Inf. Process. 48 (4) (2021) 248–257.

[51] Q. Schiermeier, Self-publishing editor set to retire, Nature 456 (7221) (2008) 432.

[52] N.A. Mazov, V.N. Gureyev, Publications by editorial board of science journals as translated into bibliometric indicators (the information and library industry), Sci. Tech. Librar. 11 (2020) 33–58.

[53] N.A. Mazov, V.N. Gureyev, The state of Russian library and information sciences from the perspective of academic journals, Bibliosphere (3) (2019) 56–70.