



Research

Predatory Journals, Science Citation Index and Open Access Publications in Diagnostic Pathology

Klaus Kayser*, Stephan Borkenfeld**, Gian Kayser***

*) Charite - Berlin

**) IAT, Heidelberg, Germany

***) Institute of Surgical Pathology, Medical Center – University of Freiburg, Faculty of Medicine, University of Freiburg, Germany

Abstract

Background: Predatory behavior in open access peer reviewed scientific journals seems to be increasing. It is subject to serious discussions between editors, authors, and publishers. Herein we discuss some characteristics of electronic communication in science, the reasons for potential misuse, and present some entropy calculations.

Theoretical considerations: Publication of research articles uses visual information transfer, and seldom acoustic communication. Advantages of open access electronic journals include prompt and worldwide distribution, free access, and several free re-use rights. Disadvantages might be seen in non congruent business models of authors and publishers, which include impact of financial issues on scientific quality and easy to perform, hard to detect manipulation of published data and their sources.

Present stage: Citation Index (SI) and Journal Impact Factor (JIP) are the principal sources which are thought to guarantee the article's quality and author's prestige. Authors try to publish in and publishers try to offer journals with high SI and JIP. Independent review processes should serve to maintain the article's quality. Reviewers are invited electronically in both conventional paper printed and electronically distributed journals. They are requested to act at their earliest convenience. Reviews are as all fast electronic information distribution exposed to become manipulated, in politics, commerce, and science as well.

Entropy calculation: Electronic information exchange uses properties of virtual reality, which include the reversibility of time. They are scalable events and can be measured by an appropriate entropy concept. The entropy of three different pathology journals has been calculated for two years. It displays with a significant difference between the conventional and



the open access journals. The measured increased entropy of open access journals indicates less strict information of the reader.

Proposal and perspectives: Authors can be considered information source that should be understood by the readers (receiver). The higher the information content the more precise is the reaction of the receiver. The readers' reaction can be measured and be used to classify, modify, correct, or even erase the distributed information. The development of appropriate readers' tools would improve the quality and originality of the article and journal.

Conclusion: A readers' oriented model would be an appropriate extension of quality assurance for open access journals and articles. Such a model could be based upon the entropy concept and be a reliable measure of research quality and impact on its future development.

Keywords: [Open access publication](#), [predatory journals](#), [entropy concept](#), [journal citation index](#), [diagnosis](#).



Introduction

Scientists require worldwide distribution of their research results in order to demonstrate their capability of innovative research, effective use of their financing, acceptance by their colleagues and to promoting their professional future. They are in need to publish their data in a fast and reliable manner [1]. In the past, they wrote by themselves their reports and articles to be published, and selected the appropriate journals in close relation to their expertise.

This procedure is history for numerous researchers. Universities and large research institutions possess specialized public relation departments that take care of public information at various levels of research specificity [2-4]. In addition, several publishing companies offer 'assistance' to potential authors in design, preparation, formal furnishing and journal selection [see for example: <https://ees.elsevier.com> [5, 6]. Other commercial companies even assist in distribution of 'similar research results', most frequently in meta-analyses of biochemical findings and patients' disease or treatment. For example, the first paper reports of macromolecule A concentration in patients with disease D, and the second one of macromolecule B in patients with disease E, etc. The two research approaches use the same algorithms, statistical procedures and interpretation. The publications are, therefore, easy to transfer into each other, and a 'chain of publications' might be started that investigates automatically molecule after molecule and disease after disease similar to automated phone calls or email series [7, 8].

Remains such a publication method still ethical, or is it below any 'reliable research level', or even predatory? What are appropriate rules to judge or classify the scientific level and impact of published articles? Herein we want to apply proven tools of forecasting the development of statistically accessible systems which might assist to answer the arisen questions.

Definitions

Common tools of forecasting and judgment journals and published articles in science have been defined as follows [9]:

Science Citation Index (SCI) is a scientific citation data bank which is commercially owned by Thomson. Its larger online version Science Citation Index Expanded (SCIE) includes more than 8,800 scientific journals of approximately 150 different disciplines. Articles of the included electronic journals were automatically transferred from the journal to the data bank via XML to SCIE.

Impact Factor (Journal Impact factor, JIP) is calculated by the relationship (number of cited articles within a year) / (number of published articles in the previous two years) of the same journal.



Open access journal is an electronically distributed journal whose articles can be read worldwide via the internet without restriction, especially without payment of the reader (free open access). Usually, the authors hold the copy rights, and the publisher asks the author for a publication fee.

Libre open access includes free open access plus additional re-use rights such as creative common licenses (CC).

Quality assurance of communication in natural sciences' research is thought to be warranted by adequate authors' independent (blind or double blind) review process, proof of corresponding research protocols, citation of related published data, and consistent editorial work.

Predatory journal is an open access journal which does not fulfill its promised scientific and formal (editorial) quality assurance, or distributes faked information. Jeffrey Beall published a list of predatory journals in 2010 [10-12]. It is off line since January 17, 2017.

Principal theoretical considerations

Publication of research articles uses visual information transfer. Acoustic communication in electronic information and data exchange is seldom performed. Paper printed publications completely stay in the physical world, electronically distributed journals and articles have entered the virtual world [13, 14].

What is the most significant difference between virtual reality and the physical environment?

The most significant difference between virtual and physical reality are the reversibility of time and the principal expansion from two until to three spatial dimensions in diagnostic performance [15].

Time reversibility easily permits alterations of visual information in its origin, in contrast to acoustic information which usually includes a long period of presentation [15, 16].

Two dimensions are commonly sufficient to diagnose morphological alterations in tissue – based diagnosis, such as microscopic images, radiological findings, etc. Information of printed papers cannot be extended from two to three spatial dimensions, and cannot be modified without visible destruction in contrast to electronically distributed articles [16].

Both virtually and physically based systems of scientific communication consist of scalable events such as authors, articles, journals, reviewers, publishers, scientific fields, SCI, etc. Therefore, structure and concept of information distribution in science fulfill the prerequisites that the entropy concept might be applied to accurately measure and forecast of the systems development [17-19]. It could be applied on static and dynamic, closed and open events of information distribution [19, 20]. In addition, consideration on time reversibility could allow to accurately measure the entropy flow that can be based upon neighboring or distant located systems [21].



Specific theory, implementation and results

The preposition of a complete chaos is the simplest approach to enter the entropy concept [21]. It disregards any internal attractive or repellent forces as well as any spatial extend of included events. This approach permits the summation of calculated entropies, which can be computed to Shannon's formula $E = k \cdot S[p \cdot \ln(p)]$. The formula can be extended to

$E(MST) = kmst \cdot \sum [p(d_{ij}/dq_{(ij)}) \cdot \ln(p(d_{ij}/dq_{(ij)}))]$, if we take the distance between neighboring events into account [21].

The recently founded journal www.diagnosticpathology.eu included 3 different types of article submissions, namely conventional, interactive, and microscopically adjusted 'How do I diagnose' ones.

The journal has published within the years 2015 – 2017 a total of 60 articles which are summarized in addition with the published journals of www.diagnosticpathology.org and 'the pathology journal' in <Table 1>.

Number of published articles, and entropy concept in open access and paper printed peer reviewed journals of pathology:

Journal name	publication year	total	research	review	case report	others*
<i>Diagnosticpathology.eu</i>	2015	25	6	3	6	10
	2016	23	12	1	2	8
	2017**	9	7	1		1
<i>Diagnpathol.org</i>	2015	216	118	3	71	24
	2016	137	60	3	43	31
	2017	79	35	3	36	9
<i>Pathology Journal***</i>	2015	675	582	477	451	meeting 155
	2016	504	383	216	317	meeting 107
	2017	304	200	139	134	meeting 68

*) How do I diagnose? Interactive publication, editorial, etc.

**) January 1, 2017 – October 31, 2017

***) according to the journal's search function, probably overlapping of keywords



‘The Pathology Journal’ is originally a paper printed peer reviewed scientific journal with a high impact factor IF of approximately 6.8. www.diagnosticpathology.org is the most successfully implemented open access peer reviewed journal in pathology. It holds now an IF of approximately 2.08. www.diagnosticpathology.eu is the only completely independent and not commercially oriented peer reviewed open access journal in pathology, which was founded in 2015, and, therefore, is not yet listed and does not possess an IF.



The calculated entropy differences are presented in <Table 2>.

Entropies of www.diagnosticpathology.eu in the years

<i>article</i>	<i>2015</i>	<i>2016</i>	<i>2017</i>
research	-0.08563	-0.08486	-0.04887
review	-0.06361	-0.03408	-0.06103
case report	-0.08563	-0.05309	0
others	-0.09163	-0.09183	-0.06103
total entropy	-0.32649	-0.26387	-0.17093

Entropies of www.diagnosticpathology.org in the years

<i>article</i>	<i>2015</i>	<i>2016</i>	<i>2017</i>
research	-0.08257	-0.0904	-0.09103
review	-0.01485	-0.02092	-0.03
case report	-0.09143	-0.09093	-0.09058
others	-0.06103	-0.08406	-0.06022
total entropy	-0.24988	-0.28631	-0.27183

Entropies of The Pathology Journal in the years

<i>article</i>	<i>2015</i>	<i>2016</i>	<i>2017</i>
research	-0.09185	-0.09196	-0.09197
review	-0.08953	-0.08209	-0.08729
case report	-0.08845	-0.09076	-0.08642
others	-0.05525	-0.05903	-0.06517
total entropy	-0.32509	-0.32385	-0.33084

Entropy differences between the journals

<i>year</i>	<i>2015</i>	<i>2016</i>	<i>2017</i>
.eu-.org *	-0.07661	0.01765	0.0991
.eu-pathology*	-0.004	0.06002	0.16011
.org-pathology*	0.08779	0.04226	0.06099



*) Explanation: .eu = www.diagnosticpathology.eu
.org= www.diagnosticpathology.org
pathology = 'The Pathology Journal'

<Table 1> and <Table 2> display with some common features of these three, quite different organized journals, which can be summarized as follows:

1. The number of published articles is decreasing in all three journals during the period under consideration. The greatest decline of publications occurred for www.diagnosticpathology.org between 2015 and 2016, and for www.diagnosticpathology.eu between 2016 and 2017. 'The Pathology Journal' does not display with such a noticeable decline. However, even this scientifically and for daily practice well recognized journal published a total of 675 articles (including conference reports) in 2015, and will reach a number of only 350 articles in 2017. In other words, all journals lost a significant number of articles from 2015 to 2017.
2. The calculated entropies in relationship to 2015 until 2017 indicate, that 'The Pathology Journal' kept its good entropy ranking without any significant changes in contrast to the other two electronically distributed journals. Whereas the entropy of www.diagnosticpathology.eu started with remarkable low entropy, which increased significantly in 2017, the journal [diagnosticpathology.org](http://www.diagnosticpathology.org) could decrease its entropy for about 10%.
3. The entropy contributions of all four analyzed subjects (research, case report, review, others) remained stable in 'The Pathology Journal' and www.diagnosticpathology.org. The increase of entropy in www.diagnosticpathology.eu is induced by the missing publication of 'case reports' only. The decline of cases in the journals does not influence the calculated entropies.
4. The entropy calculations of events within a journal permit a reliable recognition of the journal's status and are a useful tool for potential corrections and modifications of its development [5, 6].

The reasons of decrease in article publication in the journals are not essentially associated with the authors' intention on information distribution in pathology. They are also not induced by changes of information quality or by implementation of journal 'thresholds' (for example changes of the behavior of reviewers, etc.). They are probably only induced by the increased number of non-listed journals which automatically reduces the number of publications per listed journal. The number of listed journal amounted to 190 journals during the investigated period and did not change in contrast to that of the identified or suggested predatory journals. On the other hand, predatory journals have rapidly increased their publication volumes from 53,000 in 2010 to an estimated 420,000 articles in 2014, published by around 8,000 active journals [22, 23].



Financial aspects

Authors aim to distribute the results of their work as fast and multiple as possible. Conventional (printed) journals are limited by their number of readers and mailing velocity. Their publishers provide the logistics and earn their money by transportation of the information from the publisher to the reader (receiver). The receiver cannot change the transmitted information and pays for its transport. The publisher's financial interest is related to the information quality and the interest of the readers: the higher the journal's quality and the interest of the receivers the higher is the transportation fee and the number of interested paying readers [7].

Open access (electronic) journals offer both prompt velocity and world wide access. They earn their money by transportation of information from the author to the publisher. The authors have to trust the publisher and enter a contract of payment prior to see its final result. The more authors submit their articles the higher is the publisher's financial benefit. Therefore, the number of readers is of no primary interest for the publisher. The only interests are the journal metrics, such as SPI or JIF because they attract authors and facilitate the funding. Consecutively, it is of no surprise that several publishers are very imaginative to create their own metrics, such as 'advanced science index', 'Eurasian Journal Science Index', or 'Universal Impact Factor' [24-26]. All in all approximately more than 150 creative CI and IF have been created [24-26].

This behavior can be compared to advertise a product by attractive packing. It usually opens the temptation to diminish the interest on its content.

The behavior of the publishers follows this idea. They try to overbid each other in offering specific services to 'polish the English', to 'permit specific assistance in writing the introduction, results, and interpretation', to prepare 'brilliant pictures', etc. The authors have to pay for these services that are offered either by the publisher of the open access journal under consideration or by specialized companies [5, 6].

In aggregate, the interest of authors and publishers of open access journals are diverse and not congruent: The publisher wants to attract as many authors as possible, because they are his source of income. The author wants that as many colleagues as ever possible will read his article. He, his administration and funding source have the suggestion that this intent is associated with the reputation of the journal, which is thought to be related to its IP and derived CI. Therefore, IP and CI are the main 'regulators' of the financial issues, and financial issues are thought to be the 'main drivers' of predatory journals and non ethical articles.

Conclusion



IP and CI can be considered the promoting key factors of predatory journals and non ethical publications: They can be considered as packing of goods which have to be sold.

Both authors and publishers are interested in high CI and IP. Authors are tempted to submit their articles to journals that promise or pretend high CI or IP, publishers are attempted to create their own specific CI or IP. In fact, the number of different CI and IP has been grown dramatically within the last five years [22, 27]. They offer the opportunity to charge more for publication and to attract funds and administration in suggesting both high quality and world wide attraction of the authors, their work, and their institutions.

The self – regulatory process, for example that a high CI and IP indexed journal will lose its reputation if it focuses on the ‘packing’ only will rarely come into action. CI and IP and similar metrics calculations are calculated after a minimum of three years, which is a long period in electronic publication. It does not influence metrics data immediately.

The distribution of faked information such as fake citation index or preparation of predatory journals cannot be avoided in electronic (open access) publication in principle, it can only be diminished. The basic reasons are the basic properties of a virtual environment, namely reversibility of time, and the underlying business model [16, 28]. Globalization promotes fast and superficial information distribution, and delays founded science and research as long as it is not adjusted to reliable and conform standards.



References

1. [Poltronieri E, Truccolo I, Di Benedetto C, Castelli M, Mazzocut M, Cognetti G. Science, institutional archives and open access: an overview and a pilot survey on the Italian cancer research institutions. J Exp Clin Cancer Res. 29: p. 168.](#)
2. [Jefferies D, McNally S, Roberts K, Wallace A, Stunden A, D'Souza S, Glew P. The importance of academic literacy for undergraduate nursing students and its relationship to future professional clinical practice: A systematic review. Nurse Educ Today. 60: p. 84-91.](#)
3. [Sarkies MN, Bowles KA, Skinner EH, Haas R, Lane H, Haines TP. The effectiveness of research implementation strategies for promoting evidence-informed policy and management decisions in healthcare: a systematic review. Implement Sci. 12\(1\): p. 132.](#)
4. [Schiff GD, Kim S, Abrams R, Cosby K, Lambert B, Elstein AS, Hasler S, Krošnjak N, Odwazny R, Wisniewski MF, McNutt RA. Diagnosing Diagnosis Errors: Lessons from a Multi-institutional Collaborative Project and Methodology\). 2005.](#)
5. ees.elsevier.com
6. <https://www.springeropen.com/journals>
7. [Karabag SF, Berggren C. Misconduct, Marginality and Editorial Practices in Management, Business and Economics Journals. PLoS One. 11\(7\): p. e0159492.](#)
8. [Rada R. A case study of a retracted systematic review on interactive health communication applications: impact on media, scientists, and patients. J Med Internet Res, 2005. 7\(2\): p. e18.](#)
9. [Neuhaus C, Daniel HD. Data sources for performing citation analysis: an overview. Journal of Documentation, 2008. 64. \(2\): p. 193-210.](#)
10. [Abad-Garcia MF, Melero R, Abadal E, González-Teruel A. Self-archiving of biomedical papers in open access repositories. Rev Neurol. 50\(7\): p. 431-40.](#)
11. [Beall J. Criteria for Determining Predatory Open-Access Publishers \(2nd edition\). Scholarly Open Access, 2012.](#)
12. [Beall J. Predatory publishers are corrupting open access. Nature, 2012. 489\(7415\): p. 179.](#)
13. [Sukhov A, Burrall B, Maverakis E. The history of open access medical publishing: a comprehensive review. Dermatol Online J. 22\(9\).](#)



14. [Tennant JP, Waldner F, Jacques DC, Masuzzo P, Collister LB, Hartgerink CHJ. The academic, economic and societal impacts of Open Access: an evidence-based review. F1000Res. 5: p. 632.](#)
15. [Hudson LN, Blagoderov V, Heaton A, Holtzhausen P, Livermore L, Price BW, van der Walt S, Smith VS. Insect: Automating the Digitization of Natural History Collections. PLoS One. 10\(11\): p. e0143402.](#)
16. [Görtler J, Kayser K, Borkenfeld S, Carvalho R, Kayser G. Cognitive Algorithms and digitized Tissue – based Diagnosis. Diagnostic Pathology, 2017. 3: p. 1.](#)
17. [Kayser K, Gabius HJ. The application of thermodynamic principles to histochemical and morphometric tissue research: principles and practical outline with focus on the glycosciences. Cell Tissue Res, 1999. 296\(3\): p. 443-55.](#)
18. [Kayser K, Görtler J, Bogovac M, Bogovac A, Goldmann T, Vollmer E, Kayser G. AI \(artificial intelligence\) in histopathology--from image analysis to automated diagnosis. Folia Histochem Cytobiol, 2009. 47\(3\): p. 355-61.](#)
19. Voss K, Süsse H. Praktische Bildverarbeitung. 1991: München, Wien: Carl Hanser Verlag.
20. Voss K. Entropie als statistisches Strukturmaß. Wiss Z Techn Univs Dresden 1970. 19: p. 1415-1419.
21. [Kayser G, Görtler J, Weis CA, Borkenfeld S, Kayser K. The application of structural entropy in tissue based diagnosis. Diagnostic Pathology, 2017. 3\(1\): p. 251.](#)
22. [Straumsheim C. Study finds huge increase in articles published by 'predatory' journals. InsideHigherED, 2015. 10\(01\).](#)
23. [Shen C, Björk B-C. 'Predatory' open access: a longitudinal study of article volumes and market characteristics. BMC Medicine, 2015. 13\(1\): p. 230.](#)
24. <https://www.journalguide.com/>
25. <https://www.wokinfo.com>
26. <http://cofactorscience.com/journal-selector>
27. [Butler D. Investigating journals: The dark side of publishing. Nature, 2013. 495\(7442\): p. 433-5.](#)
28. [Borkenfeld S, Kayser K. Specificities of Electronic Publication in Medicine. Diagnostic Pathology, 2015. 1\(3\).](#)