

Can scientific journals be classified based on their 'citation profiles'?

AUTHORS:

Sayed-Amir Marashi¹
Amir Pandi²
Hossein Shariati³
Hossein Zamani-Nasab¹
Narges Damavandi¹
Mahshid Heidari¹
Salma Sohrabi-Jahromi¹
Arvand Asghari¹
Saba Aslani¹
Narjes Nakhaee¹
Mohammad Hossein Moteallehi-Ardakani¹

AFFILIATIONS:

¹Department of Biotechnology, College of Science, University of Tehran, Tehran, Iran

²School of Biology, College of Science, University of Tehran, Tehran, Iran

³School of Mathematics, Statistics and Computer Science, College of Science, University of Tehran, Tehran, Iran

CORRESPONDENCE TO:

Sayed-Amir Marashi

EMAIL:

marashi@ut.ac.ir

POSTAL ADDRESS:

Department of Biotechnology, University of Tehran, Shfiel Str. 13, Qods Str., Enghelab Ave., Tehran 1417614411, Islamic Republic of Iran

DATES:

Received: 25 Apr. 2014

Revised: 21 Oct. 2014

Accepted: 16 Jan. 2015

KEYWORDS:

scientific journals; journal classification; journal type; citation analysis; citation profiles

HOW TO CITE:

Marashi S-A, Pandi A, Shariati H, Zamani-Nasab H, Damavandi N, Heidari M, et al. Can scientific journals be classified based on their 'citation profiles'? *S Afr J Sci.* 2015;111(3/4), Art. #2014-0147, 3 pages. <http://dx.doi.org/10.17159/sajs.2015/20140147>

© 2015. The Author(s).
Published under a Creative Commons Attribution Licence.

Classification of scientific publications is of great importance in biomedical research evaluation. However, accurate classification of research publications is challenging and normally is performed in a rather subjective way. In the present paper, we propose to classify biomedical publications into superfamilies, by analysing their citation profiles, i.e. the location of citations in the structure of citing articles. Such a classification may help authors to find the appropriate biomedical journal for publication, may make journal comparisons more rational, and may even help planners to better track the consequences of their policies on biomedical research.

Introduction

Medical and biomedical research evaluation based on citation analysis has attracted much attention during the last decades.¹⁻⁵ Different aspects of citation analysis (in biomedical research evaluation) have been studied extensively in the literature.⁶⁻¹⁰

In research evaluation, classification of scientific publications is of great importance.¹¹⁻¹³ For researchers in academia it is important to publish their results in relevant journals to guarantee visibility. In Journal Citation Reports®, a list of 'subject categories' is prepared (and updated annually) to classify journals, in order to help rank journals in specialised fields. Such a classification has already found some useful applications. For example, the global map of science based on subject categories has been developed,^{14,15} and can help to better understand scientific collaborations. Comparative assessment of the 'quality' of research in different subject categories^{12,16,17} is important in sustainable development of science and technology.

For research institutes, planners and policymakers, it might be necessary to monitor subject categories of researchers. For example, facilitation and promotion of emerging interdisciplinary and multidisciplinary research may be an objective of science policies.^{18,19} There also is a great need for reliable methods for evaluating researchers, because the number of citations (in both scholarly journals²⁰ and online resources²¹) do not reflect the real impact of scientific papers.

All classification approaches are based on the central assumption that 'objects' in the same field/category have related features. For classification of research articles, similarity of words or citation patterns can be used as the features.

Classification of scholarly publications can be done from different viewpoints. Typically, classifications are done based on subject categories.²² Lewison and Paraje²³, on the other hand, suggested that biomedical journals be classified according to their approach to biomedicine, namely, clinical or basic.

Voos and Dagaev²⁴ proposed that the location at which citations appear within the citing publications (e.g. 'Introduction' and 'Methods') influences the meaning and the relevance of citations. This was later shown for larger article data sets when Maričić et al.²⁵ and also Bornmann and Daniel²⁶ studied the relationship between location of citation within citing articles and the frequency of citations.

In the present study, we classify journals into four groups based on the scope of their articles: protocol, methodology, descriptive and theoretical. Then, we show that journals in each group show a fairly similar citation pattern. We therefore propose that journals may be classified based on their citation patterns.

Materials and methods

We chose 17 (bio)medical and bioinformatics journals (Table 1). We assumed that all these journals belonged to one of the possible four groups (protocol, methodology, descriptive and theoretical). Citations to the articles published in each journal in 2011 and 2012 were found in Scopus. Then, full-text articles were downloaded, if possible, and the citation profiles (i.e. the sections in which the citations appeared) were analysed manually. Only those citations which appeared in standard sections of an article (Introduction, Methods, Results, Discussion²⁵) were considered. It should be noted that multiple citations in a specific section were counted once only.

Results

From the analysis of the 17 journals, 1472 citations were detected in the citing articles. Altogether, 818 citations appeared in one of the four standard sections of an article. We computed the percentage of citations in each of the four sections.

Figure 1a shows the citation profiles of 'protocol' journals. As expected, most of the citations occur in the 'Materials and methods' section. Journals which are devoted to the introduction of new protocols or software are naturally cited by end users who apply these protocols and tools for practical purposes.

There is a subtle difference between 'protocol' and 'methodology' journals: 'protocol' journals focus on describing the procedures to be followed, whereas 'methodology' journals also describe in detail the scientific reasons behind the presented methodology. Consequently, it can be expected that articles in methodology journals are

not only cited in the 'Materials and methods' of citing articles, but also in all other sections. Figure 1b confirms that such a trend is observed in these journals.

Table 1: List of journals in the present study and the categories in which they were classified

| Journal category | Name of journal | Total number of citations |
|------------------|------------------------------------------------------------|---------------------------|
| Protocol | <i>WIREs Computational Statistics</i> | 39 |
| | <i>Algorithms for Molecular Biology</i> | 75 |
| | <i>Journal of Visualized Experiments</i> | 70 |
| | <i>Source Code for Biology and Medicine</i> | 31 |
| Methodology | <i>Bioinformatics</i> | 123 |
| | <i>BioTechniques</i> | 234 |
| | <i>Evolutionary Bioinformatics</i> | 37 |
| | <i>Computational Biology and Chemistry</i> | 111 |
| | <i>Journal of Bioinformatics and Computational Biology</i> | 113 |
| Descriptive | <i>EXCLI Journal</i> | 133 |
| | <i>MEDICC Review</i> | 41 |
| | <i>Archives of Iranian Medicine</i> | 189 |
| | <i>Journal of Negative Results in BioMedicine</i> | 19 |
| Theoretical | <i>Acta Biotheoretica</i> | 28 |
| | <i>Medical Hypotheses</i> | 117 |
| | <i>Biology and Philosophy</i> | 86 |
| | <i>Journal of the History of Biology</i> | 26 |

'Descriptive' journals include those journals which mainly discuss experimental or clinical findings. Figure 1c shows that citations to the articles published in these journals mainly occur in the 'Introduction' and 'Discussion' sections of citing papers, followed by 'Methods' and 'Results' sections.

Finally, there are journals which focus on the theoretical aspects of science, including philosophical and historical issues. There are even journals which are devoted to presenting novel (and often radical) hypotheses. These journals are not expected to be cited in the 'Methods' or 'Results' sections of citing articles, which is reflected in the citation profile of these journals (Figure 1d).

Altogether, we observe that, depending on their approach to science, different journals show distinctive citation patterns. Figure 1e shows the average citation profile in each of the four categories.

Discussion

Different aspects of (bio)medical sciences are investigated, including clinical studies, molecular and biochemical medicine, diagnostic methods, traditional medicine, bioethics and even computational modelling in biomedical sciences. The first natural consequence of citation profile analysis is that journals, including biomedical journals, can be classified into superfamilies. Such superfamilies are based on the journal's approach to science, not necessarily the journal subject. Classification of journals may show which journals have the same approach to science, and therefore provide a framework for 'intra-superfamily' comparison of publications.

It is well known that citations can be interpreted based on many factors²⁷, including in which section of the citing paper the citation appears^{24,25,28}.

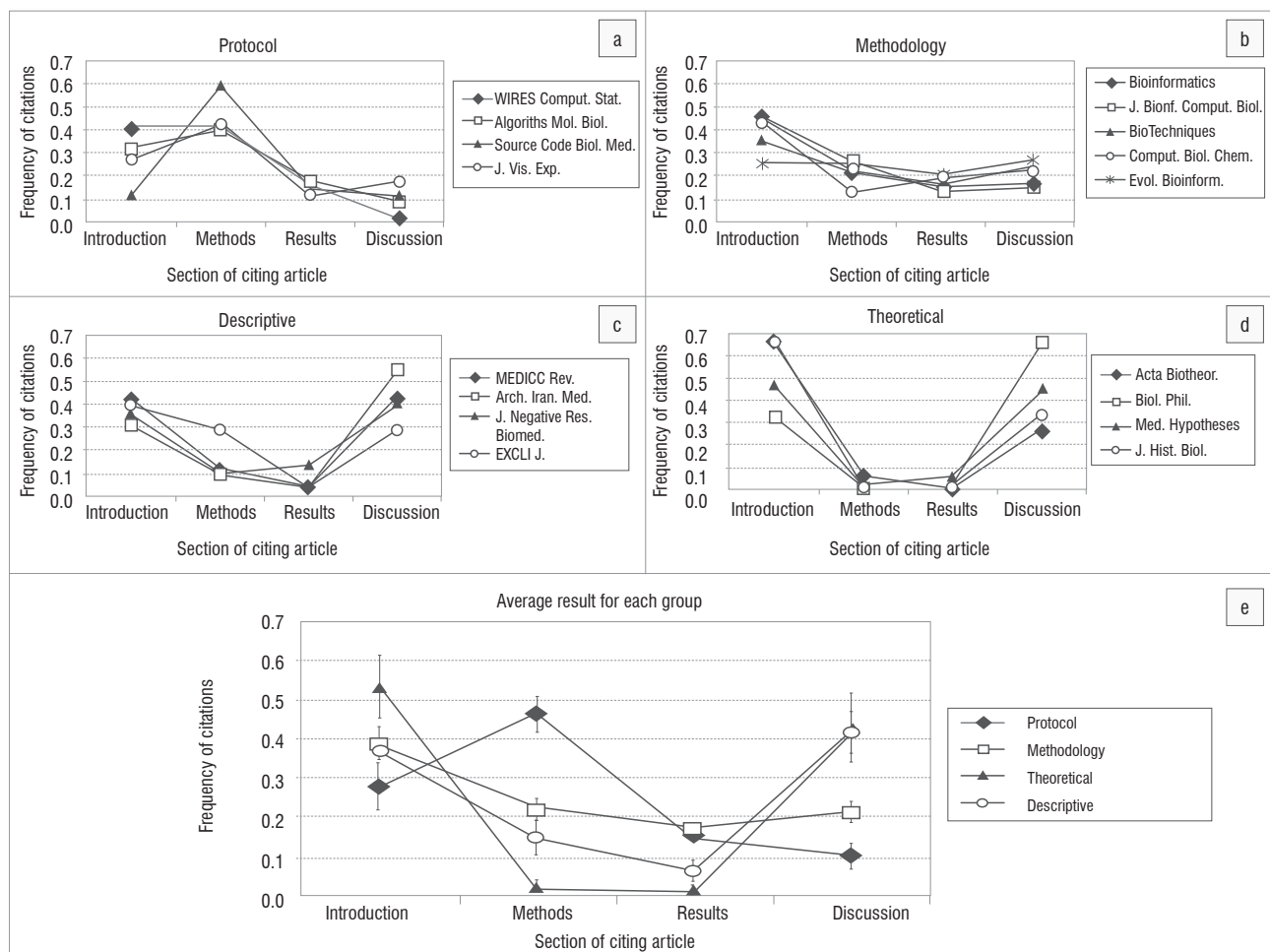


Figure 1: (a) Citation profile of protocol journals; (b) citation profile of methodology journals; (c) citation profile of descriptive journals; (d) citation profile of theoretical journals; and (e) average citation profile of each journal type.

However, citation-based measures of research evaluation do not take into account the differences among citations. We suggest that it is possible to create a context-based measure which takes into account the location of a citation in the citing article, and assign different 'usefulness' weights to the citations. For example, if the applicability of the methods is taken into account, one may give more weight to the citations which appear in the Methods section compared with the citations that appear in the Introduction section. On the other hand, for analysing groundbreaking and paradigm-changing papers, one may assign more weight to the citations in the Introduction section.

In the present study, we manually assigned the 17 journals into the four categories. For future work, one can consider a larger data set of journals and use statistical methods to automatically find the clusters, in order to show if the same four categories are found. A possible drawback of this approach is the potential difficulty of the computations. Current tools for citation analysis, e.g. Web of Science and Scopus, analyse the citations independently of the citation context.^{25,26,28} Analysing the citation profiles manually is rather inconvenient; nevertheless, with the online availability of (most) articles in machine-readable formats (e.g. HTML or PDF), text-mining algorithms may be developed to analyse the citation profiles automatically.

A side result of the automatic analysis of citation profiles could be the detection of useful keywords. More precisely, it is possible to check what issues inside a cited paper have attracted the attention of citing authors. Such a survey could provide insights for selecting better keywords for a new manuscript, in order to attract more attention, and, in turn, more citations.

Authors' contributions

S.A.M. was responsible for the experimental and project design and also wrote the manuscript. All other authors were involved in collecting the required data.

References

1. Bird SB. Journal impact factors, h indices, and citation analyses in toxicology. *J Med Toxicol.* 2008;4:261–274. <http://dx.doi.org/10.1007/BF03161211>
2. Fang MLE. Journal rankings by citation analysis in health sciences librarianship. *Bull Med Lib Assoc.* 1989;77:205–211.
3. Garfield E, Welljams-Dorof A. Of Nobel class: A citation perspective on high impact research authors. *Theor Med.* 1992;13:117–135. <http://dx.doi.org/10.1007/BF02163625>
4. Nieminen P, Carpenter J, Rucker G, Schumacher M. The relationship between quality of research and citation frequency. *BMC Med Res Method.* 2006;6:42. <http://dx.doi.org/10.1186/1471-2288-6-42>
5. Patsopoulos NA, Analatos AA, Ioannidis JPA. Relative citation impact of various study designs in the health sciences. *J Am Med Assoc.* 2005;293:2362–2366. <http://dx.doi.org/10.1001/jama.293.19.2362>
6. Falagas ME, Alexiou VG. The top-ten in journal impact factor manipulation. *Arch Immunol Ther Exp.* 2008;56:223–226. <http://dx.doi.org/10.1007/s00005-008-0024-5>
7. Marashi S-A. On the identity of 'citers': Are papers promptly recognized by other investigators? *Med Hypoth.* 2005;65:822. <http://dx.doi.org/10.1016/j.mehy.2005.05.003>
8. Seglen PO. Why the impact factor of journals should not be used for evaluating research. *BMJ.* 1997;314:498–502. <http://dx.doi.org/10.1136/bmj.314.7079.497>
9. Garfield E. How can impact factors be improved? *BMJ.* 1996;313:411–413. <http://dx.doi.org/10.1136/bmj.313.7054.411>
10. Broadwith P. End of the road for h-index rankings. *Chem World.* 2013;10:12–13.
11. Derkatch C. Demarcating medicine's boundaries: Constituting and categorizing in the journals of the American Medical Association. *Tech Commun Quart.* 2012;21:210–229. <http://dx.doi.org/10.1080/10572252.2012.663744>
12. Dorta-González P, Dorta-González MI. Comparing journals from different fields of science and social science through a JCR subject categories normalized impact factor. *Scientometrics.* 2013;95(2):645–672. <http://dx.doi.org/10.1007/s11192-012-0929-9>
13. Glänzel W, Schubert A. A new classification scheme of science fields and subfields designed for scientometric evaluation purposes. *Scientometrics.* 2003;56:357–367. <http://dx.doi.org/10.1023/A:1022378804087>
14. Leydesdorff L, Carley S, Rafols I. Global maps of science based on the new Web-of-Science categories. *Scientometrics.* 2013;94:589–593. <http://dx.doi.org/10.1007/s11192-012-0784-8>
15. Leydesdorff L, Rafols I. A global map of science based on the ISI subject categories. *J Am Soc Inf Sci Tech.* 2009;60:348–362. <http://dx.doi.org/10.1002/asi.20967>
16. Pudovkin AI, Garfield E. Rank-normalized impact factor: A way to compare journal performance across subject categories. *Proc Am Soc Inf Sci Tech.* 2004;41:507–515. <http://dx.doi.org/10.1002/meet.1450410159>
17. Sombatsompop N, Markpin T, Yochai W, Saechiew M. An evaluation of research performance for different subject categories using Impact Factor Point Average (IFPA) index: Thailand case study. *Scientometrics.* 2005;65:293–305. <http://dx.doi.org/10.1007/s11192-005-0275-2>
18. Porter AL, Roessner JD, Heberger AE. How interdisciplinary is a given body of research? *Res Eval.* 2008;17:273–282. <http://dx.doi.org/10.3152/095820208X364553>
19. Porter AL, Youtie J. How interdisciplinary is nanotechnology? *J Nanoparticle Res.* 2009;11:1023–1041. <http://dx.doi.org/10.1007/s11051-009-9607-0>
20. Hanney SR, Home PD, Frame I, Grant J, Green P, Buxton MJ. Identifying the impact of diabetes research. *Diabetic Med.* 2006;23(2):176–184. <http://dx.doi.org/10.1111/j.1464-5491.2005.01753.x>
21. Marashi S-A, Hosseini-Nami SMA, Alishah K, Hadi M, Karimi A, Hosseini S, et al. Impact of Wikipedia on citation trends. *EXCLI J.* 2013;12:15–19.
22. Glänzel W, Schubert A, Czerwon HJ. An item-by-item subject classification of papers published in multidisciplinary and general journals using reference analysis. *Scientometrics.* 1999;44:427–439. <http://dx.doi.org/10.1007/BF02458488>
23. Lewison G, Paraje G. The classification of biomedical journals by research level. *Scientometrics.* 2004;60:145–157. <http://dx.doi.org/10.1023/B:SCIE.0000027677.79173.b8>
24. Voos H, Dagaev KS. Are all citations equal? Or, did we op. cit. your idem? *J Acad Lib.* 1976;1:19–21.
25. Maričić S, Spaventi J, Pavičić L, Pifat-Mrzljak G. Citation context versus the frequency counts of citation histories. *J Am Soc Inf Sci.* 1998;49:530–540. [http://dx.doi.org/10.1002/\(SICI\)1097-4571\(19980501\)49:6<530::AID-ASIS>3.0.CO;2-8](http://dx.doi.org/10.1002/(SICI)1097-4571(19980501)49:6<530::AID-ASIS>3.0.CO;2-8)
26. Bornmann L, Daniel HD. Functional use of frequently and infrequently cited articles in citing publications: A content analysis of citations to articles with low and high citation counts. *Eur Sci Edit.* 2008;34:35–38.
27. Nicolaisen J. Citation analysis. *Ann Rev Inf Sci Tech.* 2007;41:609–641. <http://dx.doi.org/10.1002/aris.2007.1440410120>
28. Cano V. Citation behavior: Classification, utility, and location. *J Am Soc Inf Sci.* 1989;40:284–290. [http://dx.doi.org/10.1002/\(SICI\)1097-4571\(198907\)40:4<284::AID-ASIS10>3.0.CO;2-Z](http://dx.doi.org/10.1002/(SICI)1097-4571(198907)40:4<284::AID-ASIS10>3.0.CO;2-Z)

