Annals of Epidemiology 22 (2012) 649–653

Contents lists available at SciVerse ScienceDirect

Annals of Epidemiology

journal homepage: www.annalsofepidemiology.org

Evidence-based editing: factors influencing the number of citations in a national journal

Alberto Ruano-Ravina MPH, PhD [a,b,c,*], Carlos Álvarez-Dardet MD, MPH, PhD [b,c,d]

[a] *Area of Preventive Medicine and Public Health, University of Santiago de Compostela, Spain*
[b] *CIBER de Epidemiología y Salud Pública, CIBERESP, Spain*
[c] *Associate Editor,* Gaceta Sanitaria
[d] *Public Health Research Group, University of Alicante, Alicante, Spain*

ARTICLE INFO

Article history:
Received 6 March 2012
Accepted 25 June 2012
Available online 20 July 2012

Keywords:
Publishing
Editorial policies
Journalism
Citation analyses

ABSTRACT

Purpose: Citations received by papers published within a journal serve to increase its bibliometric impact. The objective of this paper was to assess the influence of publication language, article type, number of authors, and year of publication on the citations received by papers published in *Gaceta Sanitaria*, a Spanish-language journal of public health.
Methods: The information sources were the journal website and the Web of Knowledge, of the Institute of Scientific Information. The period analyzed was from 2007 to 2010. We included original articles, brief original articles, and reviews published within that period. We extracted manually information regarding the variables analyzed and we also differentiated among total citations and self-citations. We constructed logistic regression models to analyze the probability of a *Gaceta Sanitaria* paper to be cited or not, taking into account the aforementioned independent variables. We also analyzed the probability of receiving citations from non-Spanish authors.
Results: Two hundred forty papers fulfilled the inclusion criteria. The included papers received a total of 287 citations, which became 202 when excluding self-citations. The only variable influencing the probability of being cited was the publication year. After excluding never cited papers, time since publication and review papers had the highest probabilities of being cited. Papers in English and review articles had a higher probability of citation from non-Spanish authors.
Conclusions: Publication language has no influence on the citations received by a national, non-English journal. Reviews in English have the highest probability of citation from abroad. Editors should decide how to manage this information when deciding policies to raise the bibliometric impact factor of their journals.

Introduction

The citations received by papers published in a journal serve to increase its bibliometric impact factor (BIF). Although the BIF refers only to citations received in the last 2 years, an increase in citations in other years is a mean to increase a journal's visibility and influence. Because the BIF and, therefore, the citations received are used by researchers to select the journal to which send their works [1], editors try to increase it through different mechanisms to attract the greatest number of submissions that enable the selection of the top quality investigations.

Few attempts have been made to predict the number of citations received by papers published in a specific journal. A journal BIF does not determine the citations that a published work will receive. There are factors precluding a high number of citations, such as the topic of the paper, the prestige of their authors, or the institution where the work has been done. There are other aspects that could influence the number of citations, such as article type, publication language, authors' number, or publication year. The influence of these factors can vary from journal to journal depending on their quality (measured as its relative position in Thomson-ISI classifications) and on their diffusion (local or worldwide) [2]. Scientific journals use strategies to raise their BIF, but sometimes these strategies are borderline with ethics, such as advising authors on citing works previously published in their journals [3]. Citation practices are also intentionally flawed in many occasions [4] and the use of citations received is increasing for benchmarking researchers and by funding bodies.

Journal editors have the duty to select for publication only those papers with the highest scientific standards. The problem appears

* Corresponding author. Area of Preventive Medicine and Public Health, University of Santiago de Santiago de Compostela, C/ San Francisco s/n, 15782 Santiago de Compostela, Spain.
E-mail address: alberto.ruano@usc.es (A. Ruano-Ravina).

1047-2797/$ – see front matter © 2012 Elsevier Inc. All rights reserved.
http://dx.doi.org/10.1016/j.annepidem.2012.06.104

when many papers are of high quality and there is a limit of space in journals. Should editors give priority to those papers they think will have a better citation track? And which factors would predict a higher number of citations?

The objective of this paper was to assess the influence of authors' number, article type, publication language and year of publication on the number of citations received by papers published within a Spanish journal of public health.

Methodology

Design and setting

Gaceta Sanitaria is the official journal of the Spanish Society of Public Health and Health Policy, which includes a total of 11 associations, the largest of which is the Spanish Society of Epidemiology, followed by the Health Economics Association. *Gaceta Sanitaria* was founded in 1888 as the *Gaceta Sanitaria de Barcelona*. It is published six times a year and received bibliometric impact factor for the first time in 2009 (BIF: 1.172) [5]. It publishes papers mostly in Spanish, but also in English [6].

To perform the citation analysis we used the web page of *Gaceta Sanitaria* (available from: http://www.elsevier.es/es/revistas/gaceta-sanitaria-138) and the database Web of Knowledge, from the Institute for Scientific Information (Thomson-Reuters; available from: http://www.accesowok.fecyt.es/). The webpage of *Gaceta Sanitaria* has all the summaries published since 1987, with full and free access to the contents. The title of the papers, authors' number, article type, language, and publication year were extracted manually. We searched the ISI-WoK database from April to May 2011 to obtain the total number of citations for each paper, self-citations, and citations from Spanish authors or from abroad.

As inclusion criteria we considered original papers, brief originals, and reviews published in *Gaceta Sanitaria* between 2007 and 2010. It was not possible to include papers published before 2007 because they are not indexed in ISI-WoK. We excluded *Gaceta Sanitaria* supplements and other article types different than those included.

Information retrieval and data preparation for the analysis

We searched for the published papers in *Gaceta Sanitaria* and their citations at the ISI-WoK database. Self-citations were those citing an included paper having at least one author that coauthored the "source" paper. Citations received were also categorized as authors from Spain or authors from abroad, with independency of the citing journal. A citation from abroad was one whose corresponding address was not Spanish.

Statistical analysis

We performed a descriptive analysis of the number of citations received by the included papers and used logistic regression models to analyze the probability of a paper for being cited. In the two first models, the dependent variable was to be cited or not and the independent variables were publication year, number of authors (continuous), publication type, and publication language. The difference between the two models was that in the second model self-citations were excluded.

A further analysis was performed to determine the possible predictors of a high number of citations or not. To do this, the dependent variable was cited two or more times versus one time. This model excluded self-citations and only included papers having at least one citation. The dependent variables were the same as in previous models.

A last analysis was done having as dependent variable the probability of receiving citations from non-Spanish authors. The independent variables were language and publication type, included as mentioned above. All the results are presented as odds ratios with 95% confidence intervals. To know the variability explained by the regression models we also calculated the Nagelkerke R^2 statistic for each regression. The analyses were performed with PASW 18.0 (SPSS, Inc., Chicago, IL).

Results

We included 24 issues of *Gaceta Sanitaria*, with a total of 240 papers. There were 177 original articles, 46 brief original articles, and 17 reviews. The median number of authors was 5 (range, 1–18). The median number of authors in reviews was four, and five in the other publication types. There were 23 papers in English: 14 originals, 5 brief originals, and 4 reviews. By publication year, 2007 was the year with the highest number of papers in English (n = 7), and the year with the lowest number was 2010 (n = 4). Regarding the citations number, they are referred to 238 publications because 2 were not registered in the ISI-WoK database. There were a total of 287 citations in the study period. Of them, 85 were classified as self-citations (30%). The range of citations was from 0 to 15. There were 110 papers (45.8%) with no citations; 24.4% received two or more citations, and 3.4% received five or more citations. Only two papers received more than 10 citations. When analyzing citations after excluding self-citations, the range was from 0 to 13, and 56.7% of papers did not receive any citation. There were 10.1% that received two or more citations and 3.3% received five or more citations. Table 1 shows a description of the included papers.

The variable having more influence on the citations received is the publication year, that is, time since the paper was published. Papers published in 2007 have a nine-fold greater probability of being cited versus those published in 2010. Number of authors, publication language, or publication type did not influence on the possibility of being cited (Table 2). The Nagelkerke R^2 value for this regression is 0.18. When performing the same analysis but excluding self-citations, the results do not change (Table 3). The Nagelkerke R^2 value for this regression is 0.19.

Approximately half of the papers included in the analysis have never been cited. If we exclude those papers and create a variable with two categories (one citation or two or more citations), it is observed that in addition to the publication year, publication type influences the probability of receiving two or more citations. Reviews have nearly six-fold probabilities of being cited twice or more compared with original articles (odds ratio, 5.7; 95% confidence interval, 0.9–34.5). Brief originals, on the opposite, have lower probabilities of being cited twice or more than original papers (odds ratio, 0.4; 95% confidence interval, 0.1–1.1). Language or number of authors do not influence on the number of citations received (Table 4). The Nagelkerke R^2 value for this regression is 0.19.

When the origin of citations is analyzed, language does not seem to influence on the probability of citation. When the type of article is a review, the probability of citations from abroad increases; it is 3.6 times greater (95% confidence interval, 1.2–10.4) when compared with original papers (Table 5). The Nagelkerke R^2 value for this analysis is 0.05. When analyzing reviews exclusively, we find that the mean and median citations for reviews in Spanish are 1.5 and 0 versus 6.5 and 5.5 if they are in English. The results hardly change if we exclude self-citations.

Discussion

The results obtained show that there are some factors that influence the number of citations obtained by a national scientific