Journal of Informetrics 7 (2013) 198–208

Contents lists available at SciVerse ScienceDirect

Journal of Informetrics

journal homepage: www.elsevier.com/locate/joi

The importance of accounting for the number of co-authors and their order when assessing research performance at the individual level in the life sciences

Giovanni Abramo[a,b,*], Ciriaco Andrea D'Angelo[b], Francesco Rosati[b]

[a] Institute for System Analysis and Computer Science (IASI-CNR), National Research Council of Italy, Italy
[b] Laboratory for Studies of Research and Technology Transfer, School of Engineering, Department of Management, University of Rome "Tor Vergata", Italy

ARTICLE INFO

Article history:
Received 7 September 2012
Received in revised form 17 October 2012
Accepted 8 November 2012

Keywords:
Research evaluation
Co-authorship
Fractional counting
Bibliometrics
Biology
Medicine

ABSTRACT

Accurate measurement of research productivity should take account of both the number of co-authors of every scientific work and of the different contributions of the individuals. For researchers in the life sciences, common practice is to indicate such contributions through position in the authors list. In this work, we measure the distortion introduced to bibliometric ranking lists for scientific productivity when the number of co-authors or their position in the list is ignored. The field of observation consists of all Italian university professors working in the life sciences, with scientific production examined over the period 2004–2008. The outcomes of the study lead to a recommendation against using indicators or evaluation methods that ignore the different authors' contributions to the research results.

1. Introduction

Evaluating the productive efficiency of research organizations and individual scientists is an exercise that is as important as it is delicate. The principle indicator of efficiency for almost any activity is the labor productivity, or in very simple terms, the relationship between output produced in a defined period and the hours of labor expended to produce it. As for any measurement system, that for research productivity is subject to limits and approximations, which must be duly taken into account considering the field and the intended use of the results. In particular, research activity has certain characteristics that make it notably complicated to carry out accurate and robust measurement of labor productivity. We first observe the intangible nature of the output, and also consider that such outputs are generally obtained through collaboration of various individuals, who may or may not be from the same organization or nation, and who may cooperate by contributing resources, experience and competencies that are both similar and complementary. In evaluating the scientific activity of a researcher or organization it is thus fundamental to identify the true contribution that the individual or institution has provided to the various research results in which they have from time to time participated. In the scientific fields where codification of results is primarily through publication in scientific journals, indexed in such databases as Web of Science (WoS) or Scopus, bibliometrics can be conveniently applied for large-scale evaluation of productivity. In this case, the contribution of scientists and organizations to the individual publications can be recognized through the analysis of co-authorships.[1] In the

* Corresponding author at: Dipartimento di Ingegneria dell'Impresa, Università degli Stusi di Roma "Tor Vergata", Via del Politecnico 1, 00133 Rome, Italy. Tel.: +39 06 72597362; fax: +39 06 72597362.
E-mail address: giovanni.abramo@uniroma2.it (G. Abramo).

[1] While there may be a human tendency to assume the opposite, it should be noted that the quality of publications is obviously a priori independent of the number of authors.

1751-1577/$ – see front matter © 2012 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.joi.2012.11.003

life sciences, in particular, widespread practice is for the authors to indicate the various contributions to the results of the published research by the positioning of the names in the authors list.

In this work, we propose to measure the distortions encountered in the evaluation of research productivity for single individuals in Biology and Medicine when no consideration is given to the co-authors of a research work or to their order in the list.

As much as taking account of both of these factors in comparative measurement of research productivity would seem logical, and even mandatory under the theory of production, it is not at all rare that they are partially or completely ignored. In national research evaluation exercises with peer-review techniques, this is standard practice: for example in the UK Research Assessment Exercises (RAE) and in the Italian Triennial Evaluation Exercise (VTR), the peer evaluators are only called to judge the level of excellence of the products that the researchers submit, independent of true entity of the author's contribution to their accomplishment. The same is true of the national exercises that, while conducted with bibliometric techniques, examine only a share of the entire output (see the current Research Quality Evaluation Exercise, VQR, in Italy). Even famous and widely used bibliometric performance indicators, such as the *h*-index (Hirsch, 2005) and the *g*-index (Egghe, 2006), totally ignore any consideration of the contributions of the individual authors to the scientific product. Little attention has been paid to advice from the inventors themselves, such as that from Hirsch (2005), who warned that "subfields with typically large collaborations (e.g., high-energy experiment) will exhibit larger *h* values", and further recommended that "in cases of large differences in the number of co-authors, it may be useful in comparing different individuals to normalize *h* by a factor that reflects the average number of co-authors". Little attention has also been paid to the specific corrections proposed, such as the simple division of the *h*-index by the average number of co-authors included in the Hirsch core (Batista, Campiteli, Kinouchi, & Martinez, 2006; Egghe, 2008; Schreiber, 2009, 2010), or consideration of the actual number of co-authors and the scientists' relative position in the byline (Wan, Hua, & Rousseau, 2008). In spite of the above intrinsic limits, we still see major bibliometric databases such as WoS and Scopus provide the *h*-index of every author, and it is this that scientists widely use to compare their personal performance against that of their peers, to the point that this index has now become the regulated reference threshold for access to a professorial career in Italy, both for candidates and for members of the national competition commissions (Ministerial decree 344, 4 August 2011).

In the literature, various scholars have addressed the theme of the analysis of co-authorship in evaluating scientists' research performance. Van Hooydonk (1997) pointed out that the impact of a research unit can dramatically be affected by the counting procedures. Carbone (2011) holds that "in general fractional counting is preferred because this does not increase the total weight of a single paper", and suggests that "the best way to define a fractional counting of authorship is to divide the number of citations received by each paper by the square root of the number of co-authors". As early as 1968, Zuckerman studied the patterns of name ordering in cases of multiple authorship involving Nobel laureates, and concluded that "ordering of author's names is an adaptive device which symbolizes their relative contributions to research". Based on a random selection of 5686 chemistry papers from Current Contents volumes, Vinkler (2000) observed "only a slight preference for the alphabetical listing of authors over other rankings". In a previous work, Lukovits and Vinkler (1995) suggested that co-authors should declare their individual contributions to the research as percentages, and also introduced a simple equation for calculating individual contribution scores for coauthors of multi-authored papers. More recently Verhagen, Wallace, Collins, and Thomas (2003) proposed a Quantitative Uniform Authorship Declaration (QUAD) System that permits the reader to rapidly identify who contributed what. According to Bhandari, Einhorn, Swiontkowski, and Heckman (2003) "the answer, in the tradition of scientific transparency, is for authors to decide together their individual contributions and disclose these to their readers". The author order "can reveal subtle patterns of scientific collaboration and provide insights on the nature of credit assignment among co-authors" (He, Ding, & Yan, 2012). Trueba and Guerrero (2004) proposed a formula that assigns relative values to each co-author according to their position in the list. Laurance (2006) suggests that "the individual making the greatest intellectual contribution is the lead author, followed sequentially by those making progressively lesser contributions. In addition, the final-author slot is sometimes reserved for a lab head or project initiator, who may have made little direct contribution to the paper but deserves some vague honor nonetheless". In practice, different patterns are followed in ordering the authors list, from simple alphabetical order to sequences that signal the varying importance of the contributions from individual authors, a pattern which is particularly common in the life sciences.

There is increasing agreement among bibliometricians on the desirability of taking account of co-authorship through fractional counting, though there are still differences over the most appropriate fraction to assign to each co-author.

This work is not precisely concerned with establishing the most appropriate value to assign to contributions from co-authors in the life sciences. Rather after choosing fixed, but potentially "fine-tunable", criteria to assign different weight to the various positions in the list, the objective we set is to measure the extent of the distortion in performance ranking when the number of co-authors and their order are totally ignored. In Italy, there are no fixed guide-lines establishing the order of names in the authors list for the life sciences, even though some important academic lobbying bodies have officially pronounced themselves in favor.[2] The Italian National University Council states that the medical sciences are characterized by "scientific works that are prevailingly by multiple authors, in which the first and last authors are generally the leader of

[2] http://www.cun.it/media/100033/area6.pdf, last access Oct. 17, 2012.