



# How good is a model based on bibliometric indicators in predicting the final decisions made by peers?



Elizabeth S. Vieira<sup>a,b</sup>, José A.S. Cabral<sup>c</sup>, José A.N.F. Gomes<sup>a,\*</sup>

<sup>a</sup> REQUIMTE/Departamento de Química e Bioquímica, Faculdade de Ciências, Universidade do Porto, Rua do Campo Alegre, 687, 4169-007 Porto, Portugal

<sup>b</sup> Departamento Engenharia Industrial e Gestão, Faculdade de Engenharia, Universidade do Porto, Rua Dr. Roberto Frias, s/n, 4200-465 Porto, Portugal

<sup>c</sup> INESC-TEC, Faculdade de Engenharia, Universidade do Porto, Rua Dr. Roberto Frias, s/n, 4200-465 Porto, Portugal

## ARTICLE INFO

### Article history:

Received 15 October 2013

Received in revised form 16 January 2014

Accepted 23 January 2014

Available online 19 February 2014

### Keywords:

Peer-review

Bibliometric indicators

Predictive power

Auxiliary instrument

## ABSTRACT

This paper shows how bibliometric models can be used to assist peers in selecting candidates for academic openings.

Several studies have demonstrated that a relationship exists between results from peer-review evaluations and results obtained with certain bibliometric indicators. However, very little has been done to analyse the predictive power of models based on bibliometric indicators. Indicators with high predictive power will be seen as good instruments to support peer evaluations. The goal of this study is to assess the predictive power of a model based on bibliometric indicators for the results of academic openings at the level of *Associado* and *Catedrático* at Portuguese universities. Our results suggest that the model can predict the results of peer-review at this level with a reasonable degree of accuracy. This predictive power is better when only the scientific performance is assessed by peers.

© 2014 Elsevier Ltd. All rights reserved.

## 1. Introduction

Peer-review is a process that uses a set of experts, considered qualified individuals for a given field, to perform a review. These experts formulate a set of qualitative judgments related to the object under assessment. Peer-review can be applied in several contexts: (1) it is frequently used by the academic community in internal evaluations; (2) it is systematically used by the editors of journals to evaluate submitted manuscripts; (3) applicants for academic or research positions are normally selected by a special committee of experts and (4) doctoral theses are submitted to a jury of experts. In the case of manuscripts submitted for publication, peer-review is used to improve the quality of the manuscripts by detecting weaknesses and errors. The feedback given by peers is used by the author to revise and improve the work. It is usually assumed that evaluation considers the originality and the contribution of the work for the advancement of knowledge in the scientific community. Similar evaluations occur in cases 3 and 4 mentioned above. In internal evaluations, peer-review is used for decisions regarding promotions. In this situation not only the aspects stated above are evaluated, but other parameters are also assessed depending on the final purpose. Peer-review has a long history and is well accepted by the scientific community, despite its limitations. In this sense, it seems correct to say that the final judgments made by peers are considered trustworthy.

\* Corresponding author. Tel.: +351 220402507; fax: +351 220402659.

E-mail addresses: [elizabeth.vieira@fc.up.pt](mailto:elizabeth.vieira@fc.up.pt) (E.S. Vieira), [jacabral@fe.up.pt](mailto:jacabral@fe.up.pt) (J.A.S. Cabral), [jfgomes@fc.up.pt](mailto:jfgomes@fc.up.pt) (J.A.N.F. Gomes).

Several criticisms have been raised against this methodology. Studies to assess the validity and to design strategies for improving the peer-review process are not very common. The absence of agreement among peers (or reliability) when they are asked to assess the same proposal is the main weakness of this methodology. This was analysed by [Hodgson \(1997\)](#) using the proposals submitted simultaneously to the Heart Stroke Foundation (HSF) and the Medical Research Council of Canada (MRC). Both agencies use peer-review to undertake evaluations and Pearson's correlation obtained between the scores was 0.592. The two agencies made the same decision for 72.5% of the proposals. The correlation between the ratings given by two independent peers to the same proposal out a set of proposals was also analysed using the Australian Research Council (ARC) database containing around 3000 proposals. These proposals were evaluated by more than 6000 external peers who were asked to rate the proposals on the quality of the project and the quality of the proponent researchers. The authors found reliabilities ranging between 0.15 and 0.53 ([Jayasinghe, Marsh, & Bond, 2001](#); [Jayasinghe, Marsh, & Bond, 2003](#); [Jayasinghe, Marsh, & Bond, 2006](#)). [Reale, Barbara, and Costantini \(2007\)](#) studied the reliability of the peer judgments in four disciplines (biology, chemistry, economics and humanities) at the Valutazione Triennale della Ricerca (VTR). The authors calculated Spearman's coefficient to evaluate the level of agreement between two peers when assessing the same research output. They found that Spearman's coefficient ranged between 0.25 in chemistry and 0.46 in economics.

A similar study was done by [Cicchetti \(1991\)](#), but in this case the reliability was studied using manuscript submissions to journals. The author studied the reliability considering the different characteristics of the disciplines. Reliability was analysed for general and diffuse disciplines and for specific and well-defined disciplines. He found that when disciplines are general and diffuse there is more agreement on rejection of documents than on acceptance. The opposite behaviour was observed when disciplines are specific and well-defined. [Bornmann, Mutz, and Daniel \(2010\)](#) catalog the results from more than 40 studies covering the reliability of journal peer-review.

[Wood, Roberts, and Howell \(2004\)](#) studied the reliability of the peer-review process at the UK Academy for Information System (UKAIS) conference. The authors found low levels of reliability.

Reliability was also studied for the selection of doctoral and post-doctoral applicants at the Boehringer Ingelheim Fonds (BIF). The author found that in 76% of the cases peers agreed on the decision to accept or reject an applicant ([Bornmann & Daniel, 2005](#)).

Up to this point, only works studying reliability have been discussed, but other studies have analysed potential biases in the peer-review process. In particular, features such as gender, institutional affiliation, academic title and nationality among others have been studied for both peers and research applicants. At the ARC, the fact that applicants are allowed to propose their own peers has been studied to determine whether this procedure introduces a potential bias. [Marsh, Bonds, and Jayasinghe \(2007\)](#) showed that the ratings given by peers nominated by the research applicants were higher than those given by the peers nominated by the funding panel.

[Marsh, Jayasinghe, and Bond \(2008\)](#) stated that peers from Australia tend to give lower ratings than peers from other nationalities, but part of this difference can be explained by the bias introduced by the applicants appointment. They verified that peers nominated by the applicants are more likely to come from other world regions (not Australia) than peers nominated by the panel. Even controlling for this aspect the authors observed that peers from Australia give lower ratings.

The studies looking at institutional affiliation as a source of bias suggest different findings. At the ARC the authors found that high prestige universities were more successful ([Marsh et al., 2007](#)). [Reale et al. \(2007\)](#) looking at the VTR, found there was no bias associated with the prestige of the institution.

[Jayasinghe et al. \(2003\)](#) showed that the academic title has a positive effect on the final rating attributed by peers at the ARC. They found that to be a Professor in the sciences has a significant and positive effect. In social sciences and humanities being a Professor is not significant, although they found Professorial status to interact significantly with university status. The study made by [Reale et al. \(2007\)](#) showed an opposite scenario at the VTR. The authors found no association between the academic level of the applicants and peer judgments. For all scientific fields analysed (chemistry, biology, humanities and economics) they found insignificant *p*-values.

[Jayasinghe et al. \(2003\)](#) observed that gender is not a potential bias in peer-review processes. In the same study they also showed that age was not significant in sciences, but was significant in social sciences and humanities.

Studies looking to see if peer judgments are influenced by the particular characteristics of the applicants and peers do not allow a general conclusion to be reached regarding a source of bias in the methodology. However, we can say that potential biases exist, but the way they influence the final rating depends on several factors, such as the scientific area and the scientific culture of the system that is being evaluated.

Another important issue in peer-review of grant proposals is the number of proposals assessed by peers. [Jayasinghe et al. \(2003\)](#) showed that when a peer is asked to evaluate three or more proposals the results are more reliable and valid. Indeed, when peers assess several proposals more references are available to draw judgments based on the originality, quality and contribution of the proposals to the advancement of knowledge.

In addition to these criticisms associated with peer-review there are other disadvantages of the methodology:

- a) Time and implementation costs of peer-review methodology are very high. When applied on a large scale, e.g., an institutional or national level, it may be very expensive or impossible to implement. As a consequence, only the most significant research outputs are selected for evaluation. In the case of universities it may be difficult to select the best research outputs. In several cases, the selection could be based on the prestige and the position of the authors rather than the real

Download English Version:

<https://daneshyari.com/en/article/10358394>

Download Persian Version:

<https://daneshyari.com/article/10358394>

[Daneshyari.com](https://daneshyari.com)