INSTITUT DE FRANCE
Académie des sciences

ON THE PROPER USE OF BIBLIOMETRICS TO EVALUATE INDIVIDUAL RESEARCHERS

Summary

January 2011

REPORT OF THE ACADEMIE DES SCIENCES

On the proper use of bibliometrics to evaluate individual researchers

Summary and recommendations

The individual evaluation of researchers is an essential step in research. Because there are no international standards for the individual evaluation of full-time researchers employed by national research institutions and academic researchers employed by the Ministry of Higher Education and Research, their evaluation is subjective and suffers from numerous potential biases. The Académie has examined the use of quantitative bibliometrics, which are considered to be more objective, and has made a number of recommendations on rigorous rules that need to be followed when using bibliometrics to support qualitative evaluations. Such rules should be recognized beyond our borders, at the European level at least.

I - Importance and limits of peer evaluation

I - 1 I Importance of peer evaluation
Peer evaluation has long been the only way to evaluate researchers. It is irreplaceable for assessing the scientific contribution of a researcher in terms of original ideas, quality of work, conceptual and technological innovation, and more generally assessing the impact and dissemination of the researcher's work.

I - 2 Limits of peer evaluation
Such evaluations pose practical problems linked to the monumental effort required to examine applications in detail, which is amplified by the excessive number of evaluations requested. Furthermore, in a number of cases peer evaluation can be tainted by a certain subjectivity and made worst in some cases by the insufficient expertise of evaluators, potential conflicts of interest, group processes and favoritism. All such ethical issues need to be reported in writing by evaluators, as was suggested by the Académie in its Report of 8 July 2009 to the Minister.

In spite of such flaws, bibliometrics cannot be a substitute for qualitative peer evaluation, although experts of a particular field can use bibliometrics, with all due precautions, as a tool to help in the evaluation.

II - Basics of bibliometrics

The term bibliometrics is already a source of confusion. It does not measure a researcher's production but his citations. It is based on the calculation of various indicators (number of citations, integrated factors such as the h indicator or others) from bibliographic databases that cover all, or almost all, the scientific publications and their citations in most disciplines.

II - 1 Main indicators

Several bibliometric indicators are frequently used. The number of publications has little value because it does not take into account the quality of the publications. The total number of citations is more informative, but suffers from certain biases, in particular the exaggerated weight of one or two highly cited articles in spite of the fact that they are not necessarily more important. Integrated factors, such as the very widely used h and g indicators, may usefully complement the number of citations. Finally, the impact factor measures the journals and not researchers, but it is often taken into account to evaluate the quality of an article. This practice, widely used in some disciplines such as biology and medicine, is dangerous because many prestigious journals with a high impact factor also contain a significant percentage of publications of average quality. The fact remains, however, that the publication of an article in one of the highest-level journals represents an indicator of recognition, provided that the researcher has contributed significantly to the work in question.

It should also be noted, and it has a true importance, that there are quantitative criteria for evaluation that are not strictly speaking bibliometric, such as invited conferences, awarding of important contracts, prizes, patents and software development.

II - 2 Databases

Database quality is good and constantly improving for most disciplines, but it should be kept in mind that not all disciplines are covered (especially those in the Social and Human Sciences). Care must be taken that the persons who compute the bibliometric indices have access to the best databases in their entirety. Databases can be usefully supplemented with descriptive entries for each article referenced as is done by the *Mathreviews* database for Mathematics (bibliographic file with comments).

II - 3 Advantages and potential drawbacks of bibliometrics

Bibliometrics is seemingly easy to use and provides an evaluator with numbers, which are attractive by their simplicity and factual nature. It involves nevertheless numerous biases and it is important to mention that in order to carry out bibliometrics in an unquestionable fashion, time, rigor and some experience are necessary. It is also essential to recall that no indicator or set of indicators alone can summarize the quality of the scientific production of a researcher. Moreover, the importance of bibliometrics in some disciplines may incite researchers to adapt their publications and even their work to the journal in which they wish to publish their articles rather than engaging in original and creative research.

II - 4 Validating data

The computation of indicators can lead to many errors as evidenced by their variability in the databases. This report presents the main weaknesses of bibliometrics and how to avoid them. Ideally, as best expert of his own case, a researcher should calculate his own indicators (in the disciplines where the bases are available) before submitting them for validation by persons in charge of indicators at the level of a research institution or academic establishment. Such a

lightly structured process is clearly meant only to provide advice in order to maintain good practices and to check the integrity and the relevance of the calculations.

The idea of an identifier associated with each researcher is the procedure adopted by some databases. Researchers should also provide the review panel with the electronic *pdf* files of the main publications listed in their application so that any use of bibliometrics can be supplemented by the examination of the work itself.

II - 5 Distribution and reference values
Bibliometric indicators have no intrinsic value. They can only be understood relative to the distribution of indicator values for a particular field and by taking into account the age of the researchers concerned.

II - 6 Authorship
In some disciplines, especially in biology, the position of a researcher's name in the order of authors to a publication is of considerable importance as it reflects the personal contribution of the scientist to the work published and consequently the notoriety that he/she may gain. Significant efforts must be made when computing bibliometric indicators in order to ensure that articles from a single author are treated differently depending on the position of his/her name in the list of authors. More generally, publication lists should specify the exact contribution of each author, especially concerning the "short lists" provided by candidates. This point should be given further consideration.

III – When and how to use bibliometrics

III - 1 When should bibliometrics be used?
In the case of panels covering a single discipline where members usually know the candidates well, recourse to bibliometrics is not necessary except for a quick overview and without considering the results as a decisive element.

In the case of interdisciplinary panels, it may be useful to rely on bibliometrics to speed up the process when making a first selection among candidates, provided panel members keep in mind the considerable differences that exist between disciplines and sub-disciplines.

On the other hand, bibliometric indicators are of no value to evaluate a young scientist just at the beginning of his/her career and should not be used during the recruiting process, except when recruiting senior scientists.

III - 2 How should bibliometrics be used?
Bibliometrics should only be used in conjunction with a qualitative evaluation (except for the first round of candidate selection as mentioned above).

More generally, indicators should be adapted to take into account both the length of a career since their value increases cumulatively with age and eventual changes in productivity or thematic orientation during a career.

Indicators should not be the same or should be given a different weight depending on the objectives of the evaluation: recruitment, promotion, awarding of contracts or distinctions.

In accordance with international practice, general bibliometric data should be accompanied by a close examination of the 5, 10 or 20 best publications (depending on the field) by attaching the relevant *pdf* files.

Thus, jury members should not merely rely on the numbers provided by bibliometrics, but rather the entirety the bibliographic notes associated with the publications chosen by the candidate.

In cases where the final evaluation does not correspond to the bibliometric indices, explicit explanation for the reasons for such an outcome must be provided.

Similarly, applications for the promotion of professors should provide bibliometric data on the students trained by the candidate once they themselves become researchers.

All of these indicators should be systematically included in the candidate applications as a tool for aiding evaluation, but should certainly not be the sole criteria.

IV – Specificity by discipline

Disciplines, and even sub-disciplines, each have their own specificity in matters of publications and use of bibliometrics. This constitutes a potential major bias that should be taken into account when evaluating a researcher and should also be tied to the size of the particular scientific community. Bibliometrics does not allow a comparison of researchers from different disciplines and even sub-disciplines.

Very briefly, the practices in the various fields are as follows:

In Fundamental Mathematics, evaluators and jury members systematically refer to bibliographic databases that provide scientific information on the work of the candidates. On the other hand, they use very little, if at all, the bibliometric indices that can be extracted from these databases.

In Mechanical Sciences, Computer Sciences and Applied Mathematics, evaluations are based on the quality and the number of publications, as well as in the case of Computer Sciences on invitation to conferences and highly recognized meetings and software development, and in the case of Mechanical Sciences on an analysis of development and technology transfer activities. Because bibliometric indices do not take into account conferences or books, they are used very little in these fields. The scientific community is not against their use and accepts them only in addition to other available information, specified for example on the website of the researcher. Their use by a person outside the field of the researcher is not allowed.

In Physics, nearly all curriculum vitae show the number of citations and the *h* factor of the candidate in spite of the fact that evaluation committees do not use bibliometric data. For researchers in a homogeneous field, there seems to be a relatively good statistical correlation between bibliometric data and qualitative evaluation criteria such as profoundness, originality, relevance and productivity. Great caution is needed however.

In Astrophysics, bibliometrics is widely used to evaluate researchers for recruitment, promotion and funding. The open astrophysics data system (ADS) provides citations of publications and many researchers use it to count citations and/or determine their *h* factor. It seems that astrophysicists know how to use bibliometric indicators appropriately.

In Chemistry, although bibliometrics is not used officially, the usual indicators are taken into account quite seriously during preliminary discussions when evaluating the career or achievements of researchers who have been active for more than ten to twelve years. Due to the size of the community and the international dissemination of works, there are high quality evaluators who can use bibliometric indicators in a relevant manner.

In Biology and Medicine, bibliometrics occupies now a considerable and sometimes excessive place in terms of the value given to the journal impact factor, the number of citations and the h factor. The position of a researcher's name in the list of authors of an article must also be taken into consideration. Publication in a prestigious journal with a high impact factor has become a goal onto itself beyond the quality of the work.

In Economics, bibliometric indicators seem to help evaluators. A double conclusion can be drawn. On the one hand, bibliometric indicators (number of publications weighted for the average quality of the journals, or an indicator based directly on the number of citations) are useful especially in the case of the most productive researchers. On the other hand, although bibliometric indicators do not sufficiently discriminate between younger researchers, their use, even partially, gives them an incentive to improve their application file.

In Social and Human Sciences, the coverage of existing databases does not allow for a serious use of bibliometrics, and it does not seem that the situation will improve any time soon due to the specific publication practices in these fields (role of scholarly books, of works published in several languages, etc.). Outside of SHS disciplines or sub-disciplines that are well covered by databases, the use of bibliometric indicators for individual evaluations should not be allowed. Given the weight of university researchers in some disciplines, it is especially important that good practices be harmonized between the French national council of universities (CNU) and national centre for scientific research (CNRS).

V – Improvement of bibliometrics

Well used, bibliometrics can become a useful tool in the hands of peers. The Académie recommends that the following studies be carried out in order to improve the too often informal use of bibliometrics:

V - 1 Retrospective tests to compare the decisions actually taken by peer panels (CNRS, IUF, ERC) and the results of a purely bibliometric-based evaluation of the candidates. Similar studies previously undertaken in France by the CNRS should be consulted and further investigated as well as those carried out in other countries, in particular by their Academies.

V - 2 Studies to refine existing indicators and define relevant bibliometric indicators to use in the context of individual evaluations, where the usage of bibliometrics has appeared only relatively recently. There should be an in-depth examination of the notion of authorship. Creation of a steering committee for individual bibliometrics within the framework of the Observatoire des Sciences et Techniques (OST).

V - 3 Development of standards that discern originality, innovation, diffusion and creation of schools of thought, in particular through the history of recent major discoveries in the context of bibliometrics (Fields medals, Nobel prizes, Gold and Silver medals of the CNRS, etc.)

V - 4 Establishment of rules of good practice for the use of bibliometrics during researcher evaluation in response to a request by the national agency for higher education and research evaluation (AERES), one of the missions of which is the validation of evaluation procedures for researchers.