BIBLIOMETRICS AND INSTITUTIONAL EVALUATION

Jane M. Russell
Centro Universitario de Investigaciones Bibliotecológicas, Universidad Nacional Autónoma de México, Mexico City, Mexico

Ronald Rousseau
Department of Industrial Sciences and Technology, KHBO, Oostende, Belgium

Keywords: Bibliometrics, Citation Measurements, Journal Impact Factors, Relative Impact Indicators

Contents

Summary

All evaluations are dependent on the availability of adequate and reliable data relating to the outcome of the activities under scrutiny. Literature-based or bibliometric indicators which quantify the production and use of bibliographic material, have been used extensively in the assessment of research performance. Their use is based on the assumption that the immediate purpose of research is to produce new knowledge and that publication is the primary form of output. Publication counts serve as an indicator of the amount of new scientific knowledge produced by researchers. The impact of this new knowledge can be measured by the number of times publications have been cited by other scientists in subsequent work. Impact, however, cannot be automatically equated with quality. A particular form of estimating the potential quality of scientific papers is to relate this to the prestige and impact levels of the journals in which these are published. These journal impact factors can also be used to compare the citation performance of research groups within specialist fields.

The validity of bibliometric indicators is much greater at the aggregate levels of research groups, university departments and research institutes and should be applied

with extreme caution when measuring or comparing the performance of individual scientists. Bibliometric indicators are not intended to replace peer review, but rather to make research visible and debatable, ensuring that experts are sufficiently informed to make sound judgements. Publication-based evaluation, however, considers purely the research aspect of institutional scientific activity and should, therefore, be seen as only a partial indicator of overall scientific performance.

The main constraint to the general validation of bibliometric techniques is the limited availability of databases and other information sources providing reliable and comprehensive raw data for analysis, particularly with regard to research carried out in developing countries. The potential of web-based electronic sources for providing comprehensive and accurate production and citation data for bibliometric analysis coupled with the capacity of the Internet to integrate information from a large number of different sources, promises to revolutionize the way indicators are constructed by eliminating many of the methodological constraints experienced today.

1. Introduction

The worldwide preoccupation with 'value for money' in science requiring the rationalization of dwindling support for scientific research, has led to increased use of quantitative data by policy makers. Indicators based on the statistical analysis of quantitative data provided by the scientific and technological literature have been used to measure scientific activity since the beginning of the 20th century. The term 'bibliometrics' was first introduced in 1969 as a substitute for statistical bibliography used up until that time to describe the field of study concerned with the application of mathematical models and statistics to research, and quantify the process of written communication. Evaluative bibliometrics is a term coined in the seventies to denote the use of bibliometric techniques, especially publication and citation analysis, in the assessment of scientific activity.

Research evaluation is not the only area of science studies where bibliometrics has a traditional role to play. These techniques are also used extensively for studying the interaction between science and technology, in the mapping of scientific fields, and for tracing the emergence of new disciplines, as well as in the development of foresight indicators for competitive advantage and strategic planning. Bibliometrics is also relevant to other fields. Economists and historians of science, for example, use bibliometric indicators to measure productivity and eminence.

Bibliometric analysis of scientific activity is based on the assumption that carrying out research and communicating the results go hand in hand. Scientific progress is attained by researchers getting together to study specific research topics, steered by the previous work of colleagues. The classic input-output model used to describe the scientific research process suggests that publications can be taken to represent the output of science. Publications, most commonly in the form of the refereed article and the scholarly monograph, are regarded as the definitive statements of the results of research projects. This production can be quantified and analyzed to determine the size and nature of the research carried out. Studies can be performed at macro level to measure global, regional, or national trends or at the micro level of institutions or groups.

Indicators of scientific research can be divided into two main groups: the input indicators such as money spent, equipment used or personnel employed while output indicators such as the literature-based indicators already mentioned, represent the results and outcomes of the research process. Indicators are either absolute or relative. Absolute indicators refer to one particular characteristic of research activity such as number of articles published, number of citations or the amount of money spent while relative indictors show the relationship between two or more aspects such as number of articles per research group or the number of citations per paper. The latter set of indicators is generally more useful in research evaluations due to their ability to establish compound relationships between inputs and outputs such as the amount of money spent per group per article or the productivity of research groups in terms of the number of articles published per group.

Bibliometric indicators are more powerful at higher levels of aggregation and are more suitable for analyzing patterns in a large set (a faculty or large research team) and less suitable for the evaluation of individuals or small research teams. Consequently, the validity of bibliometric indicators when applied to small data sets is questionable making peer review judgements imperative at this level. Whatever their level of aggregation literature based indicators should not be used by non peer policy makers who do not have the necessary background knowledge of the research area or research groups concerned. Interpretation of quantitative data must go hand in hand with qualitative assessment procedures.

At all levels of evaluation no indicator should be taken in isolation. A series of indicators representing the different facets of scientific activity should be employed. When these partial indicators converge to give a unified picture, their validity is strengthened. Some examples of these partial indicators refer both to input into the research process, such as the level of research funding, and also to impacts resulting from the research process. Examples of the latter are non-bibliometric impact indicators such as recognition in the form of prizes or invitations as keynote speakers in major international meetings.

Conceptual and methodological problems associated with finding appropriate output measures arise from the intangible nature of much of the output of basic research activities. Nonetheless, publication and citation data have proved meaningful for measuring scientific output and its impact on the course of scientific research. The number of publications that a research group produces is taken to represent their scientific production and their primary contribution to the generation of new knowledge. Contributions to scientific knowledge take the form of new facts, new hypotheses, new theories or theorems, new explanations or new synthesis of existing facts. The number of times this new information is cited by the authors of later publications measures the impact of their work on the advancement of research in their specialized field, and sometimes, even in other areas of knowledge. It is also indicative of the amount of recognition they enjoy from other members of the scientific community. The reward system theory of science implies that scientists must share their results in order to gain recognition from their peers. Furthermore, the number of publications and citations a research group receives is associated with their visibility as scientists. However, not all published scientific work is equally visible. The level of visibility depends greatly on

the place and language of publication as well as the field in question. Work that is not internationally visible will have little chance of being picked up by scientists other than those in close communication with the authors in question. The inclusion of the group's publications in international databases is also a factor affecting their visibility, particularly as these sources are used extensively for the generation of bibliometric indicators.

Over the last decade impact factors (IF) of scientific journals have gained importance in scientific work and information management, as well as in research management and policy. IF is used as an indicator of journal performance and as such has a role to play in the evaluation of research groups, institutes and even countries. Quality journals in science generally contain coherent sets of articles with respect to content as well as professional standards. This coherence stems from the fact that most journals are nowadays specialized in relatively narrow sub-disciplines and their 'gatekeepers' (editors and referees) share views on questions like relevance, validity and quality with the invisible college to which they belong.

An important consideration, therefore, in bibliometric studies are the channels used for the dissemination of research work and their coverage in widely accessible bibliographic databases. This latter point is even more important when considering impact indicators due to the fact that only one series of databases, the Citation Indexes produced by the Institute for Scientific Information (ISI) (Philadelphia, USA), is available for citation analysis and for the production of journal impact factors. This service includes only a small proportion of journals published worldwide, restricting its coverage to a few thousand highly cited, mainstream journals.

In the present study we look at one important application of bibliometric indicators, institutional research evaluation based on the analysis of the publication and citation outputs of groups of researchers. The role of journal characteristics, such as the journal impact factor, in literature based evaluations is also described. We concentrate our discussion on the natural and life sciences where bibliometric indicators have reached a higher level of development than in other areas of human knowledge. Special attention is paid to the theoretical foundations of indicator production and the different methodologies available for their construction.

2. Bibliometrics as an Evaluation Tool

With the advent of 'Big Science' bibliometric techniques found a new application in the realms of science administration as a research management and policy tool. Previously, bibliometrics had been the little known domain of librarians, sociologists and historians of science. The need for a relatively quick, easy and inexpensive alternative to peer review for evaluating research performance led to the 'discovery' of bibliometrics by science policy specialists and the emergence of a new field of study dedicated to the quantitative study of all aspects of science activity. This new field of scientometrics attracted specialists from different backgrounds, such as mathematicians, information professionals, computer scientists, psychologists, as well as researchers from the natural and medical sciences with a special interest in the study of their own disciplines. The widespread interest in this new field led to the creation in 1977 of its own journal, aptly

named *Scientometrics*, and in 1995 to the formation of its own international professional society, the International Society for Scientometrics and Informetrics. Early in the 1960s the introduction of the *Science Citation Index* (SCI) had given bibliometrics a great methodological push. Science indicators research has also been instrumental in the development of the field of scientometrics from the seventies onwards.

Apart from the theoretical and applied research aspects of the field, bibliometrics and scientometrics also give support to countless evaluation exercises performed by tenure, promotion, and awards committees all over the world, as well as by government science policy-makers. While never intended to replace peer review the adjunct of bibliometric indicators make for better-informed expert decisions with respect to budget allocations and in the definition of research agendas and strategic goals. Most bibliometric evaluations of papers, journals and institutions correlate well with peer review appraisals suggesting that bibliometric indicators are generally accordant with the intuitive notions of knowledgeable scientists, as well as with the cognitive state of the art of particular research fields. Nonetheless, rather than bibliometrics being championed as a cheap alternative to peer review, the two methods offering different viewpoints on a common problem, should be considered complimentary and, wherever possible, used concurrently, especially in small scale evaluations.

The expansion of automated bibliographic information services linked to the exponential increase in the volume of scientific literature has presented greater opportunity for the application of bibliometric indicators in research evaluation. This, in turn, has required the design and implementation of better systems design and software development for the handling of large quantities of data and the application of algorithms for the calculation of a wide range of indicators. As these indicators have become more accessible, their weaknesses and strengths have become better understood.

An important and relatively recent application of bibliometrics is in program evaluation. Mapping a field, for instance, before a program is launched, immediately after the end of the program and, perhaps, a few years later furnishes relevant information on many aspects of the field under study, such as the occurrence of cognitive and structural changes. In funding programs too, analysis of scientific publications before and after the funding period can give important insights into its effect on the generation of publishable results.

Although bibliometrics is now a routine tool in evaluations, its use still has its critics. The fact that hard techniques are applied to one important field of human activity, namely the search for new knowledge that are subject to certain social control and coercion, is frequently the basis for censure. Quantitative studies of science then are often reproved for a reputed lack of theoretical foundation. In particular, the absence of a theory of citing is frequently debated, suggesting the need for a more secure epistemological footing to support this practice. Nonetheless, the extensive body of experience gained in the application of bibliometrics in different disciplinary contexts has proved effective for the provision of reliable and useful data for science policy decisions. Interestingly enough, applied techniques, such as the mapping of science, when based on clearly formulated assumptions, have given rise to new theoretical

perceptions of the structure and development of science. Useful insights have come also from an increasingly critical user group. Given that the applied side is an important driving force in scientometrics, user feedback has undoubtedly helped to advance the field. For this reason, current research is focussed on the development of new and more powerful literature based indicators required by the user population, as well as on the advancement of fundamental aspects of the field validating it as a *bona fide* research area respected by the broader scientific community.

2.1 Role of Bibliometrics in Institutional Evaluation

In many countries stagnating expenditure on higher education coupled with a growing intake of students in many universities, limit the possibilities for research funding. Furthermore, a growing culture of accountability in research environments is forcing scientists and teachers to become more and more productive. Funds are assigned according to performance. Research evaluation and research excellence are bywords in today's academic climate.

Traditionally, assessment of scientific research has been limited to peer review during the grant awarding process or during evaluations for promotion or tenure. Today bibliometric techniques are increasingly used as an intrinsic component of a wide range of evaluation exercises. The present tendency is for institutions to be graded more on the visibility of their products then on their long-term reputation or resources.

The ability of publication and citation analysis to encompass different levels of aggregation makes it a technique ideally suited to national and institutional studies. Nonetheless, literature based indicators are appropriate only for institutional settings that reward publication and only for those activities that produce written knowledge.

The fact that the role of written knowledge is influenced by cultural and socioeconomic aspects, as well as cognitive determinants that vary between fields of science and between different institutional settings, is considered their main theoretical constraint. Some institutions, for example, recompense behavior that reinforces the reward system of the international scientific community with their own internal reward structures. Others may set their own standards and goals.

Some indicators established globally for the evaluation of scientific performance might not be adequate for a fair or realistic assessment of certain research scenarios. Scientific output indicators based on mainstream publication in international journals should not be taken as the only bibliometric indicator for the evaluation of applied research in developing countries where publication in national journals in the local language is the norm. Disciplinary considerations are paramount.

For researchers in the social sciences and humanities, monographs and books are important dissemination channels for research results. Technological research results are published mainly in congress proceedings, reports and patents, and are better represented in this type of gray literature than in mainstream journals. The output of technological and innovation research, in many cases, is not written up as such but

appears as designs, applications, models or know-how. In these instances, literature based indicators, clearly, have little meaning.

An important consideration in any exercise of institutional evaluation is that results and recommendations to policy makers should have the general acceptance of the researchers concerned. Consequently, scientists and research managers should be included in the team responsible for the planning, execution and analysis of the research activity. Without the involvement of these key players, the evaluation exercise is unlikely to receive validation by the other members of the research community.

Institutional evaluation should be a continuous process. Ideally, procedures should be in place for the systematic monitoring of research performance and other fundamental scholarly activities. To accomplish this, institutions should develop their own data-system and make it available through the local intranet. In this way information is continually available for consultation by academic staff and other internal users, as well as for providing the raw data for the periodic generation of bibliometric and other scientometric indicators required for evaluation exercises. In practice, most evaluations are focussed on the short-term, often covering only three or four years. This is understandable otherwise results span too long a period for them to be useful for science managers. Nonetheless, their ultimate value can be measured only over the medium and long term.

In institutional evaluation exercises, scientific output and impact are related to input measures, such as research expenditure and the number and categories of academic staff. When carrying out comparative studies, other factors are considered, such as differences in the institutional academic and administrative structures, educational models, etc. Consequently, before deciding upon the procedure for collecting bibliometric data it is necessary to consider the internal institutional research structure.

While research administration of many universities follows the traditional departmental structure, the increase in multi and interdisciplinary research, often organized in a program structure, has given rise to research groups formed by members of different departments. Research groups, rather than individual scientists, are today targeted for the allocation of research funds.

For this reason, the research group is the most common unit for bibliometric analysis in institutional evaluations. This in turn has produced a wave of interest in scientometric research focussed on the identification of research groups by co-author analysis and its corroboration by expert opinion. Notwithstanding, the research performance of any aggregate of scientists can be assessed using bibliometrics. This aggregate is often termed a ‘unit’ which can be taken to represent any given set or sets of scientists depending upon the objective of the evaluation.

While no absolute quantification of research performance is possible, valid and useful comparisons can be made between research groups working in the same fields. When making comparisons between groups it is essential to apply indicators to matched groups, comparing like with like, as far as possible, and to give careful thought to what the various indicators are actually measuring. It is also important to study not only the

similarities between groups but also the differences, especially those that could be directly influencing the research performance.

-
-
-

Bibliography

Cronin B. (1984). *The citation process: The role and significance of citations in scientific communication*, London: Taylor Graham. [An essay, still valid today, debating the meaning of citing practice focussing on science as a social process.]

De Bruin R. E., Kint A., Luwel M., and Moed H. F. (1993). A study of research evaluation and planning: the University of Ghent. *Research Evaluation* **3**(1), 25–41. [Provides a detailed description of how to perform a systematic evaluation of research performance.]

Egghe L. and Rousseau R. (1990). *Introduction to informetrics, Quantitative methods in library, documentation and information science.* Amsterdam: Elsevier. [Part III of this book describes the basic ISI citation tools and provides definitions of citation and citation-related concepts.]

Garfield E. (1979). *Citation indexing-its theory and application in science, technology and humanities.* New York: Wiley. [Seminal work by the 'Father' of citation indexing.]

Martin B. R., and Irvine J. (1983). Assessing basic research: Some partial indicators of scientific progress in radio astronomy. *Research Policy* **12**, 61–90. [Convergent partial indicators provide a reasonable estimate of a group's relative contribution to scientific progress.]

Moed H. F., Burger W. J. M., Frankfort,J. G., and Van Raan A. F. J. (1985). The use of bibliometric data for the measurement of university research performance. *Research Policy* **14**, 131–149. [The first complete description of the evaluation of university research groups indicating the potential of bibliometric indicators as tools for university research policy.]

Proceedings of the Biennial Conferences of the International Society for Scientometrics and Informetrics. 1987 through 1999. [Proceedings series reporting original studies in the field of Bibliometrics, Scientometrics and Informetrics, many of which are related to indicator production.]

Van Raan A. F. J., ed. (1988). *Handbook of Quantitative Studies of Science and Technology.* Amsterdam: Elsevier. [The main purpose of this compilation is to present a wide range of topics in the domain of quantitative studies, incorporating theory, methods and applications.]

Van Raan A. F. J. (1997). Scientometrics: state-of-the-art. *Scientometrics* **38**(1), 205–218. [This review emphasizes the duality of scientometrics as both a basic and applied field, in which the applied side is considered the driving force.]

Biographical Sketches

Jane Russell is a senior researcher at the University Centre for Library Research (CUIB) at the National Autonomous University of Mexico (UNAM) and professor of the postgraduate programme in Library and Information Science in the Faculty of Philosophy and Letters at the same university. Her specialist field is the production, communication and impact of research in Latin America. She is a member of the International Society for Scientometrics and Informetrics (ISSI) and of the Latin American Network for

Science and Technology Indicators (RICYT), and also of the editorial committees of the journals "*Scientometrics*" and "*Cybermetrics*".

Ronald Rousseau is professor at the KHBO, Department of Industrial Sciences and Technology, and guest professor at the Antwerp University (UA), School for Library and Information science. He has written numerous articles dealing with citation analysis, research evaluation, informetric models, information retrieval, library management, Internet-related research, applications in ecology, and undergraduate mathematics. He is a member of the American Society for Information Science and Technology, and an active member of ISSI. He has received the Prize of the Belgian Academy of Sciences, as well as the 2001 Derek de Solla Price award.