Jointly published by Elsevier Science Ltd, Oxford and Akadémiai Kiadó, Budapest

Scientometrics, Vol. 45, No. 3 (1999) 433–437

BIBLIOMETRICS – SOME WARNINGS FROM THE NORTH

H. SKOIE
Professor and Research Director

Norwegian Institute for Studies in Research and Higher Education
Hedgehaugsveien 31, N-0352 Oslo (Norway)

1. Introduction

The The use of bibliometric indicators as assessment tools may be going far beyond their actual scope. The R&D enterprise, indeed also the universities, serve a large number of objectives in society. Bibliometrics are useful for assessing contributions to the advancement of knowledge, but much less so in assessing contributions to mission-oriented objectives of the research and they have nothing to offer in the assessment of how well academic institutions fulfil the other vital tasks that society keeps them for, i.e. training and scientific advice. To my regret, I have seen considerable misuse of bibliometrics along these lines in the Nordic countries.

2. Research evaluation in general

Research evaluation is certainly not a new activity. Assessments of research performance by review committees, funding councils, individual peers and references have a long tradition even if the terminology applied is different. What is new since the early 1980s is the more explicit demand for such evaluations on a regular and systematic basis and with sizeable programmes and entire units under scrutiny.

Increasingly we also observe some artificial or inflated 'needs' of this kind in the R&D area. Policy-makers seem in considerable doubt about research funding and what paths to follow. Does an R&D investment make any sense, solve or illuminate what is promised in a proposal? Funding boards etc. increasingly include a considerable number of user representatives or lay people representing society at large who naturally find it

0138–9130/99/US $ 15.00

hard to assess 'scientific opportunities' in a concrete setting. Furthermore, research also generates ethical dilemmas unheard of some years ago.

All in all, these developments seem to increase the demand for research evaluations, 'state of the art' reports etc. In view of the traditional autonomy of science and the self-governance of universities and research councils, this development is surprising. It is even more surprising when such activities are 'outsourced' to non-scientific bodies such as consulting firms.

A weak point in research evaluation is often what is highlighted. To judge a scientific milieu fairly, its entire professional activities should be in focus. In the evaluation of universities, their teaching activities, possibly also their transfer of knowledge contribution or scientific advice given should be evaluated in the same process as the research activities. Unfortunately, presently this is not the case in many western countries. Actually, we often see two parallel evaluation activities in university departments; of teaching and research. Most problematic is the allocation of funds to university departments solely on the basis of research criteria, as is now the case in the UK. One cannot expect academics to give teaching major attention it deserves if they are funded on the basis of their published research only. Society depends on the universities for the highest training, and the universities cannot afford widespread disenchantment with their work in this respect.

In government labs and institutes, the total activity should also be evaluated in view of the institutional objectives – be it development, monitoring, or other professional activities. The actual performance on each activity should be assessed, however, not only an overall impression or the research score if the entire unit is in focus.

In my view, the evaluation criteria applied in many research evaluations are not always appropriate for the actual research under scrutiny. Bibliometric data are often used outside their area of applicability, which is essentially basic research or an "advancement of knowledge" perspective, where publishing in international journals plays a decisive role, and bibliometrics hence make sense. Most importantly, in applied/ contract/ commissioned research, publications are in no way the only products of research, but publications data often play a decisive role in evaluation of such research, too. An EU team recently warned of this tendency: "A likely consequence is that university research funded on the basis of external criteria is, at the end of the day, evaluated on the basis of internal criteria".[1]

A major shortcoming in much evaluation work today is, as I see it, that it does not pay sufficient advance attention to the objectives for an activity, an institute, etc. The large group of university evaluators usually involved tend to be heavily biased towards academic publications. This may also strengthen the 'ivory tower' approach at

universities. For these reasons, I see too much sway over present research assessment being held by peer reviews dominated by university researchers, and overreliance on the use of bibliometrics essentially applicable for "advancement of knowledge" work.

More than fifteen years ago, our Institute was asked to write a memo on research evaluation, its possibilities and limitations.[2] We proceeded with caution, even questioning whether such activities might do more harm than good. We did, however, recommend to start systematic research evaluation to gain experience in the field. Today I must admit, however, that establishing a cumulative learning process in the area has been even harder and more disappointing than I envisioned fifteen years ago. Surprisingly many evaluation efforts in practice seem to start from scratch; the same mistakes are made time and again, including mistakes in the application of bibliometrics.

3. Bibliometrics as an academic activity

Bibliometrics has evolved much since the mid-1980s and obviously, much has been accomplished. My task today is, however, to point out weaknesses and limitations in applying bibliometrics in research evaluations. Here are a few:

- The SCI/SSCI base has a strong Anglo-American bias. Of course, to a great extent this reflects the gravity in science today. But not entirely. Accordingly, it might be of considerable interest to find out the degree of skewed distribution due to the actual composition of the journal base. To my knowledge this has not yet been done.
- Publication and citation patterns vary considerably by field and (possibly) over time. Furthermore, the data base includes only journal articles – not books, reports, etc. This is a severe limitation in the social sciences and the humanities in particular. In the latter case purely national publications often still make good sense.
- The much used 'impact factor'[3] of a journal is an average for the entire content of a journal, for a certain period of time, and the actual impact of a single article may deviate much from the average given as 'the impact factor' due to the well known skewness in scientific citation.

Hence, bibliometric findings make sense only in a macro perspective with large numbers of publications. They should not be applied to individuals and small groups. The serious researchers in the area are, of course, well aware of these limitations.

However, the real trouble is in the actual use and interpretations of bibliometrics in research evaluations, to which we shall now turn.

4. Actual use of bibliometrics

The use of bibliometrics in evaluations – including citation analysis and use of journal impact factor – has mushroomed in recent years. Bibliometric indicators certainly have some merit in assessing the huge R&D complex of modern societies, when used within their area of applicability. However, one often sees an overreliance on the data extracted, without proper attention to making the necessary reservations and bearing them in mind when interpreting the data.

Moreover, the use of bibliometric indicators is not always confined to areas where they make sense. As already pointed out, bibliometrics apply essentially to research which has "advancement of knowledge" as its rationale and in which publishing in international journals plays a meaningful and decisive role. This is by and large not the case with the heterogeneous and fast-growing area of applied/ contract/ commissioned research at the interface between university and society. For better or worse, such activities are essential features of most modern universities. The same goes for government labs and institutes, with applied research, development and monitoring as major objectives. In practice, there is an unfortunate tendency to ignore this fact in much evaluation work, quite possibly because assessing the success in relation to objectives of these activities are beyond the powers of the *bibliometric deus ex machina*. Quite simply, scientific papers are not the only successful outcome of research.

Furthermore, we may observe that what was meant as an independent peer review in practice often has come to rely on bibliometrics. The reviewers turn to (bibliometric) counting – a simpler process than forming a personal opinion of the merits of the work, and counting becomes the order of the day. Numbers seem to have a magnetic effect!

5. Possible consequences

These developments have several implications. Here we shall point to some of them:

- Research evaluation may become inadequate and unfair. At the funding end, the actual accomplishment of a research effort may not be judged according to what was intended by the effort. For example, President Nixon's Cancer Crusade obviously failed on this account, despite the fact that excellent papers were

produced by scientists funded by the programme. Society may feel increasingly let down if evaluative assessment does not measure up to the task. The EU statement already quoted may be an indication that research-internal criteria are often inappropriately applied.

- An increasing tendency towards simple counting of publications and citations etc. may have an impact on the actual behaviour of scientists. They may adjust to the new mode in funding and evaluations in agencies, concentrating even more on pub-lishing internationally and less on advice, monitoring and academic teaching while becoming more isolated in society. Thus, reliance on bibliometrics in research assessment may actually reinforce an ivory tower attitude.
- Shifts in assessment practices and techniques may also affect the distribution of power in research. R&D bureaucracies and management may favour bibliometrics over assessments given by scientific peers. The use of bibliometric data levels the ground in the ongoing 'battle' between researchers and managers in most organisations. Politicians may also be uncomfortable with the privileged role of researchers in assessing scientific competence / scientific opportunities as it stresses the de facto autonomy of science.[4]

6. Final note

Publication data/bibliometrics are useful as tools if used within their area of applicability. Quite often, non-experts overextend the interpretations which may be drawn from hard data of this kind and give them too much weight. Even scientific panels etc. are often not sufficiently critical; guidelines and warnings should be carefully spelled out in advance and sound peer review and judgment should be kept in mind. I do not oppose the use of bibliometrics, but some warnings are timely. Academic bibliometricians have a responsibility to spell out the limitations and avoid 'oversell' of their technique.

7. Reference

1. EU Commission: Research after Maastricht: An assessment, a strategy. COM (92) 682 (April 1992)
2. SKOIE, H. S. LØKKE: Evaluation of R&D; problems and methods. NIFU notat 5/84
3. SEGLEN, PER O: How representative is the journal impact factor? *Research Evaluation*, Dec. 1992
4. SKOIE, H.: Basic research – a new funding climate? *Science and Public Policy*, 2/1996