Przegląd Prawa Konstytucyjnego
-----ISSN 2082-1212-----
DOI 10.15804/ppk.2020.03.06
-----Nr 3 (55)/2020-----

Gunnar Sivertsen[1]

Problems and Considerations in the Design of Bibliometric Indicators for National Performance-Based Research Funding Systems

Keywords: Performance based research funding systems; performance indicators; bibliomet-rics; research evaluation; funding allocation; institutional funding; universities; incentives; REF
Słowa kluczowe: krajowe systemów finansowania oparte na wynikach; wskaźniki efektyw-ności; bibliometria; ocena badań; alokacja finansowania; finansowanie instytucjonalne; uni-wersytety; zachęty; REF (Research Excellence Framework)

Abstract

This article presents an overview of ten specific problems and considerations that are typically involved in designs of bibliometric indicators for national performance-based research funding systems (PRFS). While any such system must be understood and respected on the background of different national contexts, mutual learning across countries can inspire improvements. The paper is partly based on experiences from a Mutual Learning Exercise (MLE) on Performance Based Funding Systems which was organized by the European Commission in 2016–17and involved fourteen European countries, partly on experiences from advising a few other countries in developing such systems. A framework for understanding country differences in the design of PRFS is presented first, followed by a presentation of the five specific problems and considerations that are typically involved in designs of bibliometric indicators for such system. The article concludes with an overview of how Norway's PRFS has dealt with the same five problems.

[1] ORCID ID: 0000-0003-1020-3189, Professor, Nordic Institute for Studies in Innovation, Research and Education, Norway. E-mail: gunnar.sivertsen@nifu.no.

Streszczenie

Problemy i zagadnienia związane z projektowaniem bibliometrycznych wskaźników dla krajowych systemów finansowania badań opartych na wynikach

Artykuł obejmuje przegląd dziesięciu konkretnych problemów i zagadnień, które zazwyczaj pojawiają się przy projektowaniu bibliometrycznych wskaźników dla krajowych systemów finansowania opartych na wynikach (z ang. „performance-based research funding systems" -PRFS). Każdy taki system należy analizować w specyficznym kontekście krajowym, ale wzajemne uczenie się może inspirować do wprowadzania usprawnień. Artykuł powstał częściowo w oparciu o doświadczenia z warsztatów wzajemnego uczenia się (MLE) dotyczących systemów finansowania opartych na wynikach. Warsztaty dla czternastu krajów europejskich zorganizowała w latach 2016–2017 Komisja Europejska. Poza tym podstawą do napisania artykułu są doświadczenia związane z doradzaniem kilku krajom przy opracowywaniu takich systemów. Pierwsza część dotyczy różnic przy projektowaniu PRFS pomiędzy państwami, następnie przedstawiono pięć konkretnych problemów i zagadnień, które zazwyczaj wiążą się z projektowaniem wskaźników bibliometrycznych dla takiego systemu. Na koniec znajduje się opis tego, jak te pięć problemów rozwiązano w norweskim systemie PRFS.

*

I. A framework for understanding country differences in the design of PRFS

Most European countries have introduced performance-based research funding systems (PRFS) for institutional funding. An increasing trend is evident when comparing three overviews of the situation at different times[2]. Countries can be divided into three categories based on their use of bibliometrics in PRFS:

1. The funding allocation is based on research evaluation. The evaluation is organized at intervals of several years and based on expert

[2] A. Geuna, B.R. Martin, *University research evaluation and funding: an international comparison*, „Minerva" 2003, No. 41, pp. 277–344; D. Hicks, *Performance-based university research funding systems*, „Research Policy" 2012, No. 41, pp. 251–261; K. Jonkers, T. Zacharewicz, *Research Performance Based Funding Systems: A Comparative Assessment*, Brussels 2016.

panels applying peer review. Bibliometrics may be used to inform the panels. Examples of countries in this category are Italy, Lithuania, Portugal and United Kingdom.

2. The funding allocation is based on a set of indicators that represent research activities, and, in some countries, other activities as well. Bibliometrics is part of the set of indicators. The indicators are used annually and directly in the funding formula. Examples of countries in this category are Croatia, the Czech Republic, Flanders (Belgium), Denmark, Estonia, Finland, Norway, Poland, and Slovakia.
3. As in category B, but bibliometrics is not part of the set of indicators. Examples of countries in this category are Austria and the Netherlands.

Regarding the use of bibliometrics, a main distinction can be made between informing peer review with bibliometrics (category A) and direct use of bibliometrics in the formula (category B). We will limit the focus to these two alternatives in the following.

The two main purposes of a PRFS, research evaluation and funding allocation, can be difficult to distinguish. Hicks defines PRFS as related to both purposes; they are „national systems of research output evaluation used to distribute research funding to universities" [3]. One of the two purposes can be more relevant than the other for understanding the design of the PRFS. A few examples will demonstrate this.

United Kingdom has the best known PRFS in category A: The Research Excellence Framework (REF). It started in 1986 with peer review of individual performances as the chosen method for funding allocation, which initially was the main purpose of the exercises. Growing constraints on public funding and the prevailing political ideology resulted in policies aimed at greater accountability and selectivity[4]. Gradually method has become the more important purpose. The national research assessment exercise is now inextricably linked to UK research culture and policy, and the PRFS is viewed as a research assessment system more than as a funding allocation mechanism. From this perspective, the Metric Tide report[5], an independent review

[3] D. Hicks, *Performance-based university …*, p. 160.

[4] A. Geuna, B.R. Martin, *University research…*, pp. 277–344.

[5] J. Wilsdon et al., *The Metric Tide: The Independent Review of the Role of Metrics in Research Assessment and Management*, 2015. DOI: 10.13140/RG.2.1.4929.1363.

on the use of metrics in research evaluation, convincingly concludes that it is currently not feasible to assess research quality within the REF using quantitative indicators alone. Peer review is needed. The Metric Tide review also warns that the use of indicators may lead to strategic behaviour and gaming. One of the main recommendations is that[6].

Metrics should support, not supplant, expert judgement. Peer review is not perfect, but it is the least worst form of academic governance we have, and should remain the primary basis for assessing research papers, proposals and individuals, and for national assessment exercises like the REF.

This recommendation could also be interpreted as a formulation of best practice for other countries, particularly since it is aligned with the first of the ten principles of the Leiden Manifesto for Research Metrics: "Quantitative evaluation should support qualitative, expert assessment"[7]. The implication would then be that most other countries ought to change their PRFS. The trend, however, seems to go in another direction. The adoption of the UK model in Italy in 2003 has led to a semi-metric solution that differs considerably from the REF[8]. A few years ago, Sweden designed a UK-inspired model for resource allocation based on expert panels, FOKUS[9]. Sweden decided not to implement it, mostly for reasons of cost but also because the universities are concerned about their institutional autonomy and want to organize research evaluations themselves[10]. The understanding in Sweden is now that the purpose of research evaluation must be achieved by other means than the PRFS. The largest universities, e.g. Uppsala University[11], perform expert-based self-evaluations by their own initiative.

---

[6] J. Wilsdon et al., *The Metric Tide…*, p. xvi.

[7] D. Hicks, P. Wouters, L. Waltman, S. de Rijcke, I. Rafols, *Bibliometrics: The Leiden Manifesto for Research Metrics*, "Nature" 2015, vol. 520, pp. 429–431.

[8] A. Geuna, M. Piolatto, *Research assessment in the UK and Italy: Costly and difficult, but probably worth it (at least for a while)*, "Research Policy" 2016, No. 45, pp. 260–271.

[9] Swedish Research Council, *Research Quality Evaluation in Sweden – Fokus*, Stockholm 2015.

[10] Swedish Government, *Kunskap i samverkan – för samhällets utmaningar och stärkt konkurrenskraft*, Stockholm 2016.

[11] A. Malmberg, Å. Kettis, C. Maandi, *KoF17. Quality and Renewal 2017. Research Environment Evaluation at Uppsala University*, Uppsala 2017.

Norway has both an indicator-based PRFS[12] and a UK-inspired national research assessment exercise[13]. The latter is not used for funding allocation, only for learning. The purpose is to provide recommendations on how to increase the quality and efficiency of research. Norway's PRFS, on the other hand, is designed for the other purposes that typically may motivate such systems: increased transparency of the criteria for funding, enhancing the element of competition in the public funding system, and the need for accountability coupled to increased institutional autonomy[14].

It seems that an indicator-based PRFS can be defendable for the purpose of accountancy and resource allocation (summative evaluation), while a peer review-based evaluation system is more appropriate for the purpose of learning for improvement (formative evaluation).

PRFS need to be examined in their national contexts to understand their motivations and design. While research is mostly international, research funding is mostly national. Much of the institutional funding comes from taxpayers and is determined by democratic decisions. Hence, country differences in the design of a PRFS and its motivations should be expected and respected. However, inspiration and improvement can be created from mutual learning across countries.

II. Five typical problems and considerations

From September 2016 to September 2017, the author was engaged as an expert advisor in a Mutual Learning Exercise (MLE) on Performance Based Funding Systems organized by the European Commission[15]. Fourteen countries participated with government representatives: Armenia, Austria, Croatia, Cyprus, Czech Republic, Estonia, Italy, Moldova, Norway, Portugal, Slovenia, Spain, Sweden, and Turkey. (Some of these countries had not established PRFS

[12] G. Sivertsen, *Unique, but still best practice? The Research Excellence Framework (REF) from an international perspective*, "Palgrave Communications" 2017, No. 3, p. 17078.

[13] A. Geuna, B.R. Martin, *University research…*, pp. 277–344.

[14] K. Jonkers, T. Zacharewicz, *Research Performance Based…*, p. 10.

[15] K. Debackere, E. Arnold, G. Sivertsen, J. Spaapen, D. Sturn, *Performance-based Research Funding Systems*, Brussels 2017.

yet but were participating to learn from the other countries.) The author has also had the opportunity – through invited talks or more long-term engagements – to directly advise the development of research evaluation and funding systems in other countries including Flanders (Belgium), China, Denmark, Finland, Iceland, and United Kingdom.

The typical problems and considerations that arise in the design of direct use of bibliometrics for PRFS are for a large part different from those connected with bibliometrics for informing panels in research assessment. As an example, the problem with evaluating interdisciplinary research is inherent in assessments by disciplinary panels, but not as urgent in bibliometrics, where there are methods to come around it. On the other hand, disciplinary panels have less need to consider field differences in publication and citation practices, while bibliometric indicators certainly need to do so. Based on observations so far, below is a list of typical problems and considerations in the design of bibliometric indicators for direct use in a PRFS.

Involvement of the funder, the academic communities and their institutions, and bibliometric expertise in the design, implementation and organization of the bibliometric indicators. A well-organized representation of all three groups will be needed. Main considerations: Balanced influence, legitimacy, flexible development for learning by experience.

The relative influence of the bibliometric indicators within the total PRFS, and the relative influence of the PRFS on the total institutional revenues. The bibliometric indicators are most often part of a set of performance indicators in the PRFS, and the PRFS will not be the only source of funding or revenues. Main considerations: The perceived importance of the bibliometric indicators and their effects as incentives will partly depend on their economic influence, partly on other incentives in the research system, by which they can be strengthened.

Data sources, definitions and delimitations. Some countries use only WoS or Scopus, others add other sources, and yet other countries construct national databases to cover the research output from the institutions more comprehensively. The types of publications included in the indicators must be defined, and the set of included publication channels must be delimited. (A chosen data source, e.g. WoS, already represents a definition and delimitation.) National databases created for the purpose need

an explicit definition, a set of reporting instructions and some monitoring of the reporting practices. Main considerations: data quality; disinterested data production; incentives for internationalization; costs; comprehensiveness; balanced representation of all fields; the representation of national language publishing.

Indicators. They need to be balanced across fields with regard to counting methods, weighing of publication types and ranking of publication channels. Citations need to be field-normalized. The main considerations are whether indicators are available, valid and comparable across all fields. Balance is needed because institutions have different profiles of specialization, e.g. a technical university versus a general university, and that the funding mechanism should be legitimate across fields.

Limitations of macro indicators. Bibliometric indicators designed for the macro level (institutions) are often not adequate at the level of research groups or individuals. Main consideration: To avoid abuse or negative influence at the level of the individual researcher.

III. Solutions in the Norwegian model

A presentation of the Norwegian model is available in the Polish language[16]. An updated and extended version[17] was published in English in a special issue of a Chinese journal with seven articles about the implementation of or inspiration from the Norwegian model in seven countries. From these articles, one can learn that neither the model itself nor its implementations are uncontroversial. There is no perfect model. However, I will conclude by giving a short overview of how we have sought to solve the typical problems and considerations in the Norwegian model.

Design, implementation and organization. This has all been funded by The Norwegian Ministry of Education and Research. But the design of the indicator and its database was the responsibility of me as an expert in collabo-

[16] G. Sivertsen, *Finansowanie oparte na publikacjach – Model norweski*, „Nauka i Szkolnictwo Wyższe" 2017, nr 1(49), pp. 47–59.

[17] G. Sivertsen, *The Norwegian Model in Norway*, "Journal of Data and Information Science" 2018, No. 4, pp. 2–18.

ration with and under supervision of Universities Norway representing all Norwegian higher education institutions[18]. This organization, in collaboration with other organizations representing the institute and hospital sector, has the responsibility for the maintenance and further development of the indicator and its database. The responsibility is carried out by a National Publishing Board with representatives at the level of deans from all types of institutions and major research areas. The Board collaborates with national disciplinary panels in all fields[19].

Economic influence. The publication indicator reallocates less than 2% of the total expenses in the Higher Education Sector. One publication point represents less than 3,000 Euro. Still, the publication indicator receives a lot of attention from the researchers, much more attention than is given other and more consequential parts of the funding system. A reason might be that this indicator can be influenced directly by the researchers themselves, and that the indicators resonates with other incentives in the research system.

Data sources, definitions and delimitations. The data for the Norwegian model is collected in a shared national database called Cristin and delimited by a definition which all areas of research contributed to develop and agree on. According to this definition, a scholarly publication must: (1) present new insight (2) in a scholarly format that allows the research findings to be verified and/or used in new research activity (3) in a language and with a distribution that makes the publication accessible for a relevant audience of researchers, and (4) in a publication channel (journal, series, book publisher) which represents authors from several institutions and organizes independent peer review of manuscripts before publication. While the first two requirements of the definition demand originality and scholarly format in the publication itself, the third and fourth requirement are supported by a dynamic register of approved scholarly publication channels[20]. Publication channels representing authors from mainly one institution are not included.

---

[18] G. Sivertsen, *The Norwegian …*, p. 4.

[19] More information about the organization of the indicator can be found at the webpage of The Norwegian Publication Indicator: https://npi.nsd.no (10.03.2020).

[20] The register has an interactive webpage in Norwegian and English at: http://dbh.nsd.uib.no/kanaler (10.03.2020).

Indicators. Norway uses only one indicator, the Norwegian Publication Indicator. It has demonstrated balance between publication patterns across fields, including between languages[21] and different co-authorship practices[22].

Limitations of macro indicators. As with other indicators of this kind, there is both use and abuse of it in local management of individual researchers[23]. This is a concern that has been followed up by the National Publishing Board in Norway in a document comparable to the Leiden Manifesto[24] and the DORA declaration[25].

I end with these solutions only as an example. In general, different solutions need to be respected between countries because the conditions and traditions for public funding of research are different. The landscape of organizations and their missions in research also differ among countries. However, it can be valuable to exchange information and experiences, to have mutual learning, with focus on certain problems and considerations in the design of bibliometric indicators for national performance-based research funding systems. In this article, I suggested five aspects that could be given attention both in national design processes and in international mutual learning processes.

Literature

Aagaard K., *How incentives trickle down: Local use of a national bibliometric indicator system*, "Science and Public Policy" 2015, No. 5.

Debackere K., Arnold E., Sivertsen G., Spaapen J., Sturn D., *Performance-based Research Funding Systems*, Brussels 2017.

Geuna A., Martin B.R., *University research evaluation and funding: an international comparison*, "Minerva" 2003, No. 41.

[21] G. Sivertsen, *Balanced multilingualism in research*, "BiD: textos universitaris de biblioteconomia i documentació" 2018, No. 40, http://dx.doi.org/10.1344/BiD2018.40.25 (10.03.2020).

[22] G. Sivertsen, R. Rousseau, L. Zhang, *Measuring Scientific Production with Modified Fractional Counting*, "Journal of Informetrics" 2019, No. 2, pp. 679–694.

[23] K. Aagaard, *How incentives trickle down: Local use of a national bibliometric indicator system*, "Science and Public Policy" 2015, No. 5, pp. 725–737.

[24] D. Hicks et al., *Bibliometrics: The Leiden Manifesto for Research Metrics*, 2015.

[25] Available at: sfdora.org (10.03.2020).

Geuna A., Piolatto M., *Research assessment in the UK and Italy: Costly and difficult, but probably worth it (at least for a while)*, "Research Policy" 2016, No. 45.

Hicks D. et al., *Bibliometrics: The Leiden Manifesto for Research Metrics*, 2015.

Hicks D., *Performance-based university research funding systems*, "Research Policy" 2012, No. 41.

Hicks D., Wouters P., Waltman L., de Rijcke S., Rafols I., *Bibliometrics: The Leiden Manifesto for Research Metrics*, "Nature" 2015, vol. 520.

Jonkers K., Zacharewicz T., *Research Performance Based Funding Systems: A Comparative Assessment*, Brussels 2016.

Malmberg A., Kettis Å., Maandi C., *KoF17. Quality and Renewal 2017. Research Environment Evaluation at Uppsala University*, Uppsala 2017.

Sivertsen G., *Balanced multilingualism in research*, "BiD: textos universitaris de biblioteconomia i documentació" 2018, No. 40, http://dx.doi.org/10.1344/BiD2018.40.25.

Sivertsen G., *Finansowanie oparte na publikacjach – Model norweski*, „Nauka i Szkolnictwo Wyższe" 2017, nr 1(49).

Sivertsen G., Rousseau R., Zhang L., *Measuring Scientific Production with Modified Fractional Counting*, "Journal of Informetrics" 2019, No. 2.

Sivertsen G., *The Norwegian Model in Norway*, "Journal of Data and Information Science" 2018, No. 4.

Sivertsen G., *Unique, but still best practice? The Research Excellence Framework (REF) from an international perspective*, "Palgrave Communications" 2017, No. 3.

Swedish Government, *Kunskap i samverkan – för samhällets utmaningar och stärkt konkurrenskraft*, Stockholm 2016.

Swedish Research Council, *Research Quality Evaluation in Sweden – Fokus*, Stockholm 2015.

Wilsdon J. et al., *The Metric Tide: The Independent Review of the Role of Metrics in Research Assessment and Management*, 2015, DOI: 10.13140/RG.2.1.4929.1363.