European Commission

# MLE on Performance-based Research Funding Systems (PRFS)

## Bibliometrics in PRFS

Thematic Report No 2


[*Month of the report– 20XX*]

Research and Innovation

**Bibliometrics in PRFS – MLE on Performance-based Research Funding Systems (PRFS)**

European Commission
Directorate-General for Research and Innovation
Directorate A – Policy Development and Coordination
Unit A4 – Analysis and monitoring of national research and innovation policies
Contact Marta Truco Calbet
E-mail marta.truco-calbet*@ec.europa.eu*
RTD-PUBLICATIONS@ec.europa.eu
European Commission
B-1049 Brussels

Manuscript completed in June 2017.

EUROPEAN COMMISSION

# MLE on Performance-based Research Funding Systems (PRFS)

## Bibliometrics in PRFS

Thematic Report No 2

Gunnar Sivertsen

Directorate-General for Research and Innovation

2017 EN

## Table of Contents



## List of Tables

## 1 INTRODUCTION

This thematic paper is written as part of the process in the Mutual Learning Exercise (MLE) in 2016-17 on Performance-based Research Funding Systems (PRFS) for institutional funding. The particular focus of this report is on **the use of bibliometrics and similar data and methods** in relation to such funding systems. As in general for this MLE, only the systems for direct funding of **universities** will be considered.

The general aim of the MLE is to support the countries in designing, implementing, evaluating and improving their PRFS. The key questions of this report are how bibliometrics is used directly or indirectly in the funding system, the indicators and data sources used, and the degree to which bibliometric indicators or bibliometric information influence funding allocation. Particular attention is dedicated to PRFS where part of the institutional funding directly depends on the performance of institutions measured through bibliometric indicators. In such cases it is important to understand and monitor expected and realised impacts as well as unintended consequences, which can depend on how and to what extent bibliometrics is used. An overview is given of eight typical questions that one should take into consideration in the design of bibliometrics for a PRFS, as well as a closer look at different solutions that are currently used in five of the countries participating in the MLE.

## 2 WHAT IS BIBLIOMETRICS?

OECD (2002) defines bibliometrics as a statistical analysis of books, articles, or other publications. Bibliometics is particularly relevant in the research domain because the publication and communication of new results in peer-reviewed scientific and scholarly publication channels is often seen as an integral and necessary part of the research process itself. Bibliometrics therefore covers the research activity in general by tracing publications and bibliographic records that it leaves behind.

Originally, bibliometrics was limited to collecting data on numbers of scientific articles and publications, classified per authors and/or institutions, fields of science, country, etc., in order to construct simple 'productivity' indicators for academic research. Subsequently, more sophisticated and multidimensional techniques were developed, based on citations in articles (and more recently also in patents). The resulting citation indexes and co-citation analyses are used both to obtain more sensitive measures of research quality and to trace the development of fields of science and of networks.

Bibliometric analyses use data on numbers and authors of scientific publications and on articles and the citations therein (as well as in patents) to measure the 'output' and 'scientific impact' of individuals/research teams, institutions, and countries. The purpose is to identify national and international networks, and to map the development of new (multi-disciplinary) fields of science and technology.

Bibliometrics is an international field of research with its own conferences and journals. The most important conferences are the annual STI ENID conferences and the biannual conferences of the International Society for Scientometrics and Informetrics (ISSI). Examples of main journals are Journal of the American Society for Information Science and Technology, Journal of Informetrics, Research Evaluation, Research Policy, and Scientometrics. Introductory courses to bibliometrics are held by CWTS at Leiden University and by the European Summer School for Scientometrics.

In recent years, bibliometric methods have been extended beyond scholarly publishing to include activity and impact in social media, particularly those media that researchers most often use. This evolving branch of bibliometrics is often called **altmetrics** or **webometrics**. As with traditional bibliometrics, the main sources of data and indicators are produced by commercial suppliers. Altmetrics is further discussed in a separate thematic report on third-mission evaluation and indicators.

Bibliometrics is mostly used for statistical and monitoring purposes with macro indicators at the level of countries and institutions. Bibliometrics is also often used to inform research evaluation or funding decisions, sometimes directly in funding formulas in PRFS. Bibliometrics also has a long tradition of being used professionally by libraries to monitor their information resources, and by scientists to study the dynamics of science itself. We find the most widespread use of bibliometrics in:

- Statistical reports and strategic documents on R&D at the national or international level
- Provided information for research assessments at the national or institutional level
- Local research management at research institutions
- PRFS, either used as indicators in the funding formula or to inform research assessment in connection to PRFS
- Informal daily practice among the researchers themselves. This has become the most widespread use of bibliometrics in recent years, particularly after the introduction of specialised social media where researchers upload and share their publications

# 3 CITATION INDICATORS IN RELATION TO PRFS

## *3.1 Why use citation indicators?*

Citation indicators are the best known and most used bibliometric indicators. If a previous publication appears in the reference list of a new publication, this reference can be taken as a 'citation' to the previous publication, thereby indicating influence or usefulness in further research. Citations are often interpreted as representing 'research quality' in policy documents. However, experts agree that they represent only one dimension of research quality, namely the influence on further research, or: *internal scientific impact*.

Citation distributions are extremely skewed. Most publications are seldom or never cited. A small proportion of all articles receive most of the citations. Highly cited are those publications that receive wide attention internationally. Because of this, citation indicators are sometimes used in relation to PRFS with the explicit aim of stimulating research quality.

## *3.2 Most used citation indicators*

The most used citation indicators in relation to PRFS:

1. Total number of citations
2. Total number of citations compared to the average in the field
3. Proportion of publications among the most cited in the world in the same field

The first of these indicators is *size-dependent*, the two other indicators are not. The third of the indicators has had increased use in recent years, perhaps because it is understood as representing 'excellence'. But again, just as with 'quality', the concept is not easy to measure in one dimension.

An interactive source of bibliometric information in which one can study the typical well-established field-normalised citation indicators by selecting particular countries and universities is the Leiden Ranking: http://www.leidenranking.com/. Technical explanations for the indicators are also given there.

Citation indicators are not very easy to construct. Citation analysis is an expert domain. Caution has to be taken because the number of citations will depend on several factors such as:

- The year of publication (time to get cited)
- The type of publication (e.g. review articles are more cited than original articles)
- The field of research – see next section.

## *3.3 Field normalisation of ciation indicators*

Citation frequencies and productivity vary much across fields, partly because publication and referencing practices are different across fields, and partly because the fields may be covered to a different degree in the database used for the measurement. Indicator 1, above, is without field normalisation and cannot be used to compare publications from different fields, which is often an implicit condition of a PRFS. Indicators 2 and 3, above, are examples of field-normalised indicators. With caution, these can be compared across fields, but a very detailed and sometimes problematic classification is often needed, e.g. distinguishing between neurosciences and clinical neurology.

## *3.4 Citation databases*

Citation databases are bibliographic databases where publications are linked whenever they refer to each other in the reference list. This method demands a recording of the full

reference list of each publication. Citation indicators are only possible if this method (citation indexing) is used.

There are two large commercial citation databases: Web of Science (WoS), which is now provided by Clarivate Analytics, and Scopus, which is provided by Elsevier. They are built on the same principles, i.e. a selection of the most important scientific journals for indexing, and recording a full range of bibliographic data (e.g. including all published author affiliations). They can be seen as competing products. The two citation databases are now available as toolboxes for local research management, i.e. SciVal from Elsevier, and InCites from Clarivate Analytics.

Google Scholar (GS) is also based on citation indexing, but here, the method is an algorithm that automatically identifies documents on the web as 'scholarly' and discovers possible links in the references. There is no restriction to certain journals, and GS is free for users. Hence, there is a widespread use of GS among researchers. However, the algorithm and the database itself are not available, and search results are unstable. GS has therefore not yet become a datasource for professional bibliometrics.

### *3.5 Not recommended: the JIF and the h-index*

Two easily available indicators are widely used in research assessment, at least informally, but are **not** recommended to be used in *research evaluation* where the purpose is to assess the quality of the researchers' contribution. Even if the purpose is only to *allocate funding* at macro level in a PRFS, one should be careful with the application because there is a risk with these indicators to allocate funding to research that is below the standards that the PRFS aims to improve.

The Journal Impact Factor (JIF) measures the average number of citations to articles in a certain journal within a certain time period. The JIF was developed by the founder of Web of Science, Eugene Garfield, for the purpose of monitoring journal selections in libraries. While the JIF, and the alternative, SNIP (based on Scopus), may be **valid indicators of journal impact**, they **cannot be used to assess the citation performance of individual publications or authors**. The reason is that average citation impact of a journal is only a weak predictor of the citation impact of individual publications in the journal because article citedness tends to be highly skewed among the publications. While a few publications receive many citations, most publications are seldom or never cited. This has been demonstrated in several bibliometric studies since the 1990's.

Nevertheless, the Journal Impact Factor is widely used for the evaluation of individual researchers and their articles. This practice has recently resulted in a series of well-organised reactions from the scientific communities. First came the *San Francisco Declaration on Research Assessment* (ascb.org/dora), which was initiated by the American Society for Cell Biology and now has more than 13,000 signees across the world. Then, published in Nature in April 2015 by experts in bibliometrics and research evaluation, came the **Leiden Manifesto for research metrics**, an annotated list of ten principles to guide research evaluation (Hicks et al., 2015). A few months later appeared **The Metric Tide** report (Wilsdon et al., 2015), which provided the Higher Education Funding Council for England with an independent review on the role of metrics in research assessment and management. All of these documents agree with Eugene Garfield, the inventor of the JIF:

> *It would be more relevant to use the actual impact (citation frequency) of individual papers in evaluating the work of individual scientists rather than using the journal impact factor as a surrogate. The latter practice is fraught with difficulties, as Seglen and others have pointed out.*

The strength of the JIF, however, is that it tends to give consistent results from year to year with regard to the performance of journals. Some countries, e.g. Flanders (Belgium), Italy and Spain, use the JIF explicitly as an indicator of **journal performance,** not of individual

performance**,** in their funding systems. The indicator is thereby not replacing more proper indicators of performance at the article level, but supplementing them with extended information. The reason for using the JIF in such contexts may be to incentivise publishing in world-leading journals.

There is, unfortunately, widespread unofficial use of the JIF in research evaluation. There are, however, many reasons for not installing the JIF in a PRFS for institutional funding. Two of them can be mentioned here: The risk of opening up for gaming of the system, and the risk of allocating funds to research performance below the standards that the PRFs aims to improve.

Another indicator often used informally is **the h-index (or Hirsch index),** which is defined as follows: A research unit has index h if h of its publications each have at least h citations and the other publications each have no more than h citations. The h-index is 3 if at least three publications have at least three citations. Perhaps because of its simplicity, the h-index has recently become the most popular bibliometric indicator among amateurs.

In practice, the h-index is almost never field-normalised and therefore contains the problems discussed above under field-normalisation. Most often, it is used with whole counts on individual author level, disregarding the contribution of other authors. The dynamics of the h-index imply that you get the highest score at the end of your career, and that the indicator may increase even after you cease to do research. The above-mentioned Leiden Manifesto warns against using the indicator in research evaluation:

> *The older you are, the higher your h-index, even in the absence of new papers. The h-index varies by field: life scientists top out at 200; physicists at 100 and social scientists at 20–30. It is database dependent: there are researchers in computer science who have an h-index of around 10 in the Web of Science but of 20–30 in Google Scholar. Reading and judging a researcher's work is much more appropriate than relying on one number. Even when comparing large numbers of researchers, an approach that considers more information about an individual's expertise, experience, activities and influence is best.*

Again, unfortunately, there is widespread unofficial use of the h-index in research evaluation. There are, however, many reasons for not installing the h-index in a PRFS for institutional funding. One of them is that it is difficult to aggregate the indicator from the individual level. Another is the risk of harming young careers in the research system. Advanced citation analysis based on actual citation counts per article is preferable.

# 4 PRODUCTIVITY INDICATORS IN RELATION TO PRFS

## *4.1 Why use productivity indicators?*

Productivity indicators reflect *research activity*. Sometimes, the aim of a PRFS using productivity indicators is just to allocate funding according to research activity. Productivity indicators in a PRFS may aim at and have the effect to *increase productivity*. They can also increase the focus on stimulating research at the level of *research management*. Using them can demand a data source with such correctness and completeness that the data can support a *better overview of and insight into* ongoing research. Such overview can *create comparisons of research profiles* if the data are properly field-classified. Depending on how the productivity indicators are constructed, they can also stimulate or disincentivise *research collaboration*.

## *4.2 Most used productivity indicators*

Just as with citation indicators, productivity indicators are not easy to construct in order to give a balanced representation of all fields and incentives to the research and publishing practices that the PRFS aims to stimulate.

Some of the basic indicators used in bibliometrics that are based on publication data only are:

1. Total number of publications

2. Total number of publications compared to input variables (resources for research)

3. Distribution of publications among fields (research profile) compared to the general distribution in the database or in other units of assessment

4. Share of publications with co-authors in certain collaborations, e.g. with other countries or with industry

5. Frequencies of co-authored publications in certain relations

Of these, the first two are most often used in relation to PRFS. The first is size-dependent, the second is not.

The three other indicators reflect properties of publication data (field of research, co-authorship) that needs to be taken into account whenever productivity indicators are constructed for the purpose of PRFS. We will explain this in the next sections.

## *4.3 How to attribute publications to collaborating authors and institutions*

Co-authorship practices differ widely across fields. The average number of authors may differ, and also the norms with regard to the sequence of authors and the importance of some positions, e.g. the first or corresponding author. As a consequence of these differences, the productivity at the level of individual researchers or at department level cannot be compared directly. This is often neglected, even if field-normalised citation indicators are applied.

In professional bibliometric studies on the macro level, e.g. when comparing countries or institutions, both publications and citations are distributed between contributing authors and affiliations by using *fractional counts*. The alternative, however, *whole counts* of publications to each contributor, may be relevant for other purposes, such as assessing individual researchers by their CV. Intermediate solutions are also possible (see an example in section 7.3.2 below)

Particularly in any *direct* use of bibliometrics in PRFS, all aspects of field-normalisation and its possible consequences need to be simulated, discussed and resolved.

### *4.4 Weighting of publication counts*

Apart from taking multiple authorships and multiple affiliations of publication into account, publications may be given different *weights* in order to obtain a *balanced representation of all fields* and to *stimulate certain research and publishing practices*. Particularly in relation to PFRS, we can see (and need to discuss) different practices, such as:

1. Publications are given different weights according to publication type.

2. The choice of whole counts versus fractional counts as described above. Intermediate solutions can occur by not fractionalising in certain relations, e.g. for co-authors abroad.

3. Weighting according to publication channels can occur in order to stimulate for example internationalisation and/or publishing in the most prestigious journals. Use of Journal Impact Factor can occur here. Ranking by other means is possible, e.g. language, database coverage (WoS, ERIH PLUS, etc.), or expert judgement.

4. Co-authorship in certain relations, e.g. internationalisation, can be incentivised by multiplying the counts by certain factors.

An example of such choises is given in section 7.3.2 below.

### *4.5 Publication databases*

Several bibliometric indicators, mainly those representing productivity, research profiles (relative composition of disciplinary fields) and collaboration in publications, do not depend on citation indexing. In these instances, other international databases such as PubMed or Latindex may be used. They can be used instead of or as a supplement to citation databases, which are, of course, publication databases as well, but less complete in their representation of the literature.

Of particular interest for bibliometrics in relation to PRFS are local or national bibliometric databases that aim to give a more complete representation of the research output. These databases are often produced by or in collaboration with the research institutions themselves. Examples in Croatia are the Croatian Scientific Bibliography (CROSBI) and Hrčak, the central portal of Croatian scientific journals. The R&D Information System in the Czech Republic is an example of a current research information system (CRIS) serving the purpose of a PRFS. The two most used CRIS systems so far that serve this purpose and have been used in bibliometric studies as well is the VABB-SHW database in Flanders, Belgium, and the Current Research Information System in Norway (CRIStin). CRIS databases may include peer-reviewed scholarly publications that are not covered by the commercial databases. Table 1 below compares the coverage of scholarly publications from Norway in Scopus and Web of Science to their coverage in CRIStin (100 per cent).

Table 1:Norway: Scholarly publications in Scopus and Web of Science to their coverage in CRIStin

| Scientific discipline | Scopus | Web of Science |
|---|---|---|
| Health sciences | 79 % | 75 % |
| Natural sciences | 84 % | 78 % |
| Engineering | 67 % | 58 % |
| Social sciences | 38 % | 22 % |
| Humanities | 30 % | 11 % |

The main differences between the CRIS system and the two main commercial databases (Scopus and Web of Science) are the limited coverage in the latter of other publications than journal articles, of scholarly publications in national languages, and of international journals in the social sciences and humanites.

## 5 THE PURPOSES OF THE PRFS AND THE USE OF BIBLIOMETRICS

### *5.1 The four major design alternatives*

The thematic report on *PRFS Design* distinguishes between four major design alternatives:

1. Peer review-based
2. Informed peer review
3. Mix of peer review & bibliometrics
4. Metrics-based

Among the fourteen countries participating in the MLE, Cyprus and Moldova do not have a PRFS, while Austria uses performance contracts. Among the eleven remaining countries, Italy, Portugal and Slovenia use informed peer review (alternative 2), Armenia, Spain[1] and Turkey use a mix of peer review and bibliometrics (alternative 3), while Croatia, Czech Republic, Estonia, Norway and Sweden have PRFS with direct use of metrics (alternative 4).

In a wider international perspective – see again the thematic report on PRFS Design – New Zealand and United Kingdom are the only countries using alternative 1. Several other countries can be placed in all of the three other categories. 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 (Geuna & Martin, 2003; Hicks, 2012; Jonkers & Zacharewicz, 2016). There are more countries using alternatives 3 or 4 than 1 or 2.

### *5.2 The two main purposes of PRFS*

Of particular interest in this thematic report on bibliometrics in PRFS is to understand why bibliometrics is given little or no role at all in some countries while other countries rely directly on bibliometrics. This variation is possible to understand if we acknowledge that PRFS have two major purposes - *research evaluation* and *funding allocaton* – which may be given different emphasis in each national context.

In principle, the two main purposes can be difficult to distinguish. Hicks (2012) defines PRFS as related to both purposes; they are "national systems of research output evaluation used to distribute research funding to universities". Nevertheless, in national contexts, one of the two purposes can be more relevant than the other for understanding the design of the PRFS. We will use the examples of four countries to illustrate this.

### *5.3 Different emphasis on the purposes – four examples*

*United Kingdom* was the first country to introduce a PRFS in 1986. Peer review of individual performances as the chosen *method* for funding allocation (Geuna and Martin, 2003). Gradually, the method has become the more important *purpose*. The national research assessment exercise is now inextricably bound up with UK research culture and policy. This is the specific context in which The Metric Tide report (Wilsdon et al., 2015) convincingly concludes that peer review is needed: "Metrics should support, not supplant, expert judgement." This recommendation, however, is not followed by the majority of other European countries in their design of PRFS.

*Sweden* recently decided between a metrics-based and a peer review-based alternative. A UK-inspired model for resource allocation based on expert panels, FOKUS (Swedish Research

[1] The Spanish Sexenio System is regarded as a PRFS in this context.

Council, 2015), was designed. but the government decided not to implement it, mostly for reasons of cost, but also because the universities are concerned about their institutional autonomy and want to organise research evaluations themselves (Swedish Government, 2016). Sweden consequently continues with its solution since 2009: A small part of the resource allocation for research is based indicators of external funding and of productivity and citation impact within Web of Science. The understanding in Sweden is now that the purpose of *research evaluation* must be achieved by other means than the PRFS. The funding model continues to be mainly about funding allocation.

*Norway* not only has a metrics-based PRFS. It also has a UK-inspired research assessment exercise (Geuna and Martin, 2003). It is not used for funding allocation. The purpose of having subject-specific evaluations at large intervals is to provide a critical review of the Norwegian research system in an international perspective, and to provide recommendations on measures to encourage increased quality and efficiency of research. Norway's PRFS, on the other hand, is designed for other purposes that also typically may motivate these systems (Jonkers and Zacharewicz, 2016): 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. Norway's PRFS includes several performance indicators representing both research and educational activities. Hence, the idea of replacing the indicators with panels performing research evaluation has never been discussed. Research evaluation takes place in a different procedure and resource allocation is the single purpose of the PRFS.

*The Netherlands* is similar to Norway in the sense that there is a research assessment exercise at certain intervals which does not influence resource allocation. The Dutch exercise is self-organised by each of the universities and coordinated on the national level by a Standard Evaluation Protocol. With this autonomous self-evaluation system in place, there is an agreement with the government that performance indicators representing research should *not* be part of the PRFS, which covers other activities. Neither the purpose of research evaluation nor the purpose of funding allocation for research is thereby taken care of by the PRFS.

### *5.4 Understanding different national contexts for PRFS design*

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 tax payers and is determined by democratic decisions. Hence, country differences in the design of a PRFS and its motivations should be expected and respected. While the above mentioned main recommendation of *The Metric Tide* report – metrics should not replace peer review – is more than adequate for the design of the PRFS in the UK, the existence of fully metrics-based PRFS in some other countries needs to be understood in the specific national contexts.

The possible adverse effects of a PRFS, or its success in increasing the performance and effeciency of a research system, may not only depend on which of the four major alternatives is chosen. The effects of a PRFS may also depend on its design and its degree of influence on the funding allocation. The following chapter will focus mainly on the design of biblometrics for *direct* use in PRFS.

## 6 TYPICAL PROBLEMS AND CONSIDERATIONS IN THE DESIGN OF BIBLIOMETRICS FOR DIRECT USE IN PRFS

### *6.1 Introduction*

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* expert 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 from the submitted material for the MLE, and on the plenary and group discussions of the MLE, below is a list of *eight typical problems and considerations* in the design of bibliometric indicators for direct use in a PRFS.

### *6.2 Eight typical problems and considerations*

1. **The relative economic 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*.

2. **Data sources.** 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. **Main considerations** behind the solutions are: data quality; disinterested data production; incentives for internationalisation; costs; comprehensiveness; balanced representation of all fields; the representation of national language publishing.

3. **Definitions and delimitations.** The types of publication included in the indicators must be defined, and the set of included publication channels must be delimited. A chosen data source, e.g. WoS, represents a definition and delimitation. The same is true for any combination of external data sources. National databases created for the purpose need an explicit definition, a set of reporting instructions and some monitoring of the reporting practices. The **considerations** behind the solutions will be much the same as in point 2 above.

4. **Indicators.** Both publication, collaboration and citation indicators are used in PRFS designs. The **main considerations** are connected to the dimensions of performances that the different indicators represent, and whether they are available and valid across all fields.

5. **Field normalisation.** The well-established field normalisation methods for citation indicators need to be supplemented with a balanced representation of productivity across fields. This may depend on how the indicators are designed – see the remaining points below. It can also be solved with balances in the funding formula itself or by separating field-specific streams of funding. The **main considerations** are that institutions with different profiles of specialisation, e.g. a technical university versus a general university, need to be treated equally, and that the funding mechanism should be acceptable across fields.

6. **Counting methods.** How to count publications with multiple authors and affiliations is an often-overlooked problem in the design of bibliometric indicators for PRFS. The

**main considerations** will be: to balance the indicators across subfields with different co-authorship practices, and to incentivise collaboration without stimulating the inclusion of authors with minimal contributions.

7. **Weighing of publication types.** If more than one type of publication (e.g. journal articles) is included in the indicators, such as peer reviewed conference papers, book chapters and monographs, these must be weighed against each other. The **main considerations** will be: to balance the indicators across subfields with different publication practices, and to incentivise a favourable development of those practices.

8. **Ranking of publication channels.** Publications of the same type may be given different weights depending on where they are published. This can be done by using journal impact factors, journal 'quartiles", the delimitation of certain respected data sources, or panel evaluation of publication channels. The **main considerations** will be: to incentivise internationalisation or publishing in certain important publication channels; to balance between research quality and research productivity; to provide legitimate incentives that do not discriminate national language publishing in the social sciences and humanities; to respect the DORA declaration.

# 7 EXAMPLES OF SOLUTIONS IN FIVE COUNTRIES

This chapter covers the different motivations, designs, implementations and experiences with the use of bibliometrics for PRFS that we find at the national level in five of the countries participating in the MLE. They have been chosen as examples to illustrate the eight typical considerations mentioned above. Four of the countries, Croatia, Estonia, Norway and Sweden, are examples of metrics-based PRFS with different designs. The fifth country, Italy, has a PRFS based on informed peer review within a national research assessment exercise.

## *7.1 Croatia*

### *7.1.1 Background and motivation*

In 2012, the Ministry of Science and Education of the Republic of Croatia started negotiations with Public Research Organisations (PROs) - which then included 7 public universities (now 8) and 25 public research institutes - about the development of a new model of multiannual funding based on performance indicators. The idea was to reduce state intervention in the decision-making processes of PROs and to shift the responsibilities for strategic planning and financing research activities to the PROs. The new model would also enable systematic monitoring of scientific activities and an allocation of financial resources which would be in accordance with the achieved results of the PROs. Performance indicators provided by a more transparent and merit-based system were agreed upon. The first PRFS was adopted on 6 June 2013 and used for funding in 2013-2015 (three year period).

The same model was used both in 2016 and 2017. Currently there is a discussion how to proceed with the model in the following years. The idea is to modify some indicators in order to address the impact of research on society and economic challenges as well as the research environment.

It is important to note that the Croatian PRFS reallocates only a small portion - EUR 6,6 million annually - of the resources made available for research. It is only meant to provide stable funding for regular research activities at the institutional level. Other direct state funding is provided for expenditures for employees (salaries), the material costs of ongoing activities of a research organisation. In addition, competitive projects are funded by the Croatian Science Foundation.

The intention of the PRFS has been to focus on a small number of measurable performance indicators and use them to keep track of the development of certain research organisations and, based on the results achieved, continue to further finance research activities to a greater or lesser extent.

The PRFS model combines four sets of performance indicators and takes into account the specificity of each field of science and different types of PROs (with shares in the funding model):

- scientific activity - 60%
- national and international competitive research projects and research mobility - 25%
- collaboration between research and business sector, as well as collaboration with the units of local and regional governance and non-governmental sector - 10%
- popularisation of science - 5%

Taking into consideration varying research costs among fields of science, the following weights have been agreed upon on the national level: Natural Sciences = 2.7; Technical Sciences = 2.5; Biomedicine and Health Sciences = 2.7; Biotechnology Sciences = 2.6; Social Sciences =1.2; Humanities = 1.2; Arts = 1.2.

The funding formula starts by calculating the base amount per full time equivalent (FTE) of an employed researcher. The base amount per FTE is calculated on the basis of the total number of FTE researchers, weights for the fields of science and the total amount of funds foreseen in the State Budget for PRFS.

The amount for an institution is then calculated on the basis of the science performance indicators and the base amount per FTE. It takes into consideration the size of institution and its field of science.[2]

In this chapter on bibliometrics, we mainly consider the first set of indicators (scientific activity) mentioned above.

### *7.1.2 Data sources and indicators*

The Croatian model uses a combination of bibliometric indicators that depends on the field of science. Overall weight of this indicator called Science activity is 60%.

Each institution gathers the data from all faculties or departments. The institution annually submits integrated lists for the indicators which are divided by the three main groups of field of science (I: natural sciences, technical sciences, biomedicine, biotechnology; II: Social sciences and humanities, and III: arts). For group I, only publications in Web of Science are included. Group II may include publications in Scopus plus selected publications according to according to the national Regulations on Selection Procedures for Academic Ranks. Regarding group III, see below.

Only original research publications should be selected from any of the approved data sources. Research organisations are given clear instructions on how to collect data on papers published in Web of Science and Scopus and how to count their citations. Citations in a year are counted regardless of the year when the cited paper was published.

Data is collected per research organisation and not per scientist, with the published address of affiliations as criterion. Research organisations are registered according to their field of science. All papers published at a research organisation are taken into account for the field of science the research organisation is registered in. Papers must be deduplicated at the institutional level. Each list must be signed by the rector of the university or the director of the public research institute and made publicly available on the institution's website.

An agreement has been reached that Web of Science defines the eligible scientific output in the STEM fields. Two indicators are given the following weights, as shown in table 2.

[2] For example, if the base amount is 100 HRK, a researcher in the biomedicine will receive 100 x 2.6= 260 HRK and a researcher in humanities will receive 100*1.2= 120 HRK. The funding formula takes into consideration science performance indicators, so depending on the performance, the institution receives more or less money per researcher (if the institution in humanities performs good they can receive much more than 120 HRK per researcher, or if it performs bad, less than that).

Table 2: Scientific activity indicators for STEM fields

| Indicator | Data Source | Weight |
|---|---|---|
| Journal article | Web of Science | 0.5 |
| Citations | Web of Science | 0.1 |

The indicators in the social sciences and humanities are based on a wider range of data sources and publication types (table 3).

Table 3: Scientific activity indicators for SSH fields

| Indicator | Data Source | Weight |
|---|---|---|
| Journal article | Web of Science | 0.12 |
| Journal article | Scopus | 0.10 |
| Research article | Regulations on Selection Procedures for Academic Ranks | 0.08 |
| International book3 | Public Research Organisation (PRO) | 0.08 |
| Book | PRO | 0.06 |
| Edited book (with article) | PRO | 0.03 |
| Textbooks, etc. | PRO | 0.03 |
| Citations | Web of Science | 0.06 |
| Citations | Scopus | 0.04 |

There is no list of international and other publishers that define the book indicators. However, it is clearly indicated in the instructions that the books should be published with leading international publishers and in one of the major world languages.

In addition to the indicators based on scientific and scholarly publications, the Croatian model has an indicator set for scientific activity in *the arts*. It includes indicators and weights for artistic production (works, performances, exhibitions), nominations and awards, artistic research projects, and contributions to congresses, evaluation committees or festivals. This part of the Croatian model is unique in the context of European indicator-based PRFS. Some

[3] In international languages and published by leading international publishers.

other countries (e.g. Flanders/Belgium and Norway) have tried, unsuccessfully so far, to create indicators for artistic research.

The set of indicators related to popularisation (weighted 5 per cent) is unique as well in metrics-based systems. Systems based on peer review, however, often allow outputs from artistic research or popularisation to be submitted.

#### *7.1.3 Effects if the bibliometrics for the PRFS*

The effects of the Croatian solutions for PRFS has not been studied, but it is currently being evaluated for a possible redesign.

#### *7.1.4 Documentation, references and links*

Domagoj Karacic, D., Miskulin, I., Serdarusic, H. (2016.) State investment in science and scientific productivity of universities. UTMS Journal of Economics, 7(1): 37–48.

http://public.mzos.hr/Default.aspx?sec=3521

http://narodne-novine.nn.hr/clanci/sluzbeni/2013_06_69_1367.html

https://www.google.hr/url?sa=t&rct=j&q=&esrc=s&source=web&cd=4&ved=0ahUKEwiD9f-KxtPTAhWFAxoKHThsC5oQFgg5MAM&url=https%3A%2F%2Frio.jrc.ec.europa.eu%2Fen%2Ffile%2F9514%2Fdownload%3Ftoken%3D-8JG6aKx&usg=AFQjCNH8uBpJ9zvMO9L5xPwR4K-S170ZnQ&sig2=SCFU_zyJeRCdgvoeiKeXJQ

http://public.mzos.hr/fgs.axd?id=24811

### *7.2 Estonia*

#### *7.2.1 Bakground and motivation*

The funding of research in Estonia is mainly based on competitive grants from the Estonian Research Council and from similar sources abroad, and from contract research. However, in 2005, Estonia introduced a direct institutional funding for research which is called the baseline funding. It represented 13 per cent of the total R&D funding in the beginning, but the share has been increasing up to 30 per cent lately.

All kinds of research organisations may apply for the baseline funding, but they have to pass an institutional evaluation every seven years to be eligible. The evaluation is organised at the national level and based on peer review by international experts.

The annual baseline funding itself is almost completely determined by performance indicators. The exception is 5 per cent of the budget, which is allocated to support research that is particularly relevant for the Estonian language, culture, history and society. The large part is allocated on the basis of three types of performance indicators:

1. 50 per cent (40 per cent until 2016) is allocated according to indicators of external revenues, in practice mainly from contract research, but also from competitive grants.

2. 40 per cent (50 per cent until 2016) is allocated according to bibliometrics and indicators of patenting.

3. 10 per cent is allocated according to the number of doctoral graduates.

With regard to the first type, the indicators of external revenues are carefully defined to represent only so-called R&D services for the private or public sector. They involve only research activities, not other services such as consulting, commissioned by the third party and for the benefit of the third party, not of the research organisation itself. Funding from EU framework programmes is also included as evidence of quality and cooperation with high level partners.

With regard to the second type, registered patent applications are weighted 2, and registered patents are weighted 3, compared to articles in international journals, which are normally weighted 1. In the following, we will concentrate on the bibliometric indicators.

### *7.2.2 Data sources and indicators*

The data sources for the bibliometric indicators are Web of Science and (more recently) Scopus. In addition, data from scholarly book publishing is selected from the Estonian Research Information System (ETIS).

Only publications count; citation indicators are not used. Publications are only attributed to the institutions if they are mentioned as an affiliation in the publication. Publications that can be attributed to more than one institution will count once for each of them. There is, however, an element of fractionalisation in that publications with more than 100 authors will be given only half of the weight, while publications with more than 1,000 authors will be given a third of the weight. Data from the preceding three years is used every year to calculate the indicators for next year's budget. The indicators are the same for all fields of research. They are shown in table 4.

Table 4: Estonia: Bibliometric indicators and their weights

| Indicator | Data Source | Weight |
|---|---|---|
| Journal article | Web of Science or Scopus | 1 |
| Monograph | ETIS | 5 |
| Article in book | ETIS | 1 |
| Two or more articles in book | ETIS | 2 |

Books will be included only if they are published by an international scholarly publisher that is listed on a specific list in ETIS (e.g. Cambridge University Press or American Mathematical Society):
https://www.etis.ee/Portal/Classifiers/PublishingHouseDetails?lang=ENG

ETIS is an information system that covers a wider range of reseach outputs than those included in the bibliometric indicators for the baseline funding. ETIS will for example record scholarly books or journal articles in the Estonian language. The reason for limiting the bibliometric indicators to internationally indexed journal articles or internationally published books is a concern about data quality and research quality. One should bear in mind, however, that research that is particularly relevant for the Estonian language, culture, history and society is funded in a separate part of the baseline funding.

### *7.2.3 Effects of the bibliometrics for the PRFS*

There have been no direct study of the effects of the PRFS in Estonia, but the impression is that it has influenced an increased focus on peformance at the universities. The allocation of funds has changed over time, also indicating that is has an effect. Several other incentives are connected to publication performance – the effects of the PRFS are difficult to isolate. But there has been an increase in the interest of patenting at the universities.

### *7.2.4 Documentation, references and links*

To some extent, The Research and Innovation Policy Monitoring Programme (TIPS Programme), commissioned by the Estonian Ministry of Education and Research, has relevant information in English:

http://www.tips.ut.ee/index.php?module=2&op=&dok_id=735

http://www.tips.ut.ee/public_funding

## *7.3 Norway*

### *7.3.1 Background and motivation*

When a PRFS for the Norwegian higher education institutions was first introduced in 2002, it covered both research and educational activities with a relatively light touch. At that time, 40 per cent of the total direct funding was reallocated according to performance indicators with 25 per cent representing education and 15 per cent representing research. Since all performance indicators were size-dependent, the distribution of funds across institutions did not change much after the implementation. Although the general aim of the PRFS was to increase the quality of all activities, the major immediate effect was to make the funding allocation more *explicit* and to provide both the funding and the funded organisations with a statistical overview of the activities.

The performance indicators representing research were indicators of external competitive funding, completed doctoral degrees and the number of FTE in academic positions. Since the latter indicator was partly dependent on educational activities and not a direct indicator of research activity, the universities and the government agreed that it needed to be replaced by research ouput indicators. The idea of using research evaluation with peer review for the same purpose was not discussed. For advice and formative purposes, Norway had already implemented a national research assessment system based on peer review by international panels, at it was already decided that this system should not be used for direct institutional funding.

Supporting the choice of direct use of bibliometrics for the PRFS was also the fact that the institutions for several years already had been reporting statistics from metadata representing scholarly publications in their local current research information systems (CRIS). Since the institutions could see that international databases such as Web of Science only partly covered their output with wide wide differences among their faculties, the ambition became to improve the data quality in the CRIS systems and develop an bibliometric indicator with a balanced representation of all fields. The Ministry of Education and Research asked the funded organisations, represented by the Norwegian Association of Higher Education Institutions, to realise this ambition. The result was the so-called 'Norwegian model', which was developed in 2003-2004 and implemented in 2005 in the budgets for 2006.

### *7.3.2 Data sources, indicators and overall design*

After the implementation of the so-called 'Norwegian Model' (Ahlgren et al., 2012; Schneider, 2009; Sivertsen, 2016) in 2005, it has later on also been adopted at the national level by Denmark (2009) and Finland (2015), partly also by Flanders, Belgium (2009) and by several Swedish universities at the local level. It has three components:

A. A complete representation in a national database of structured, verifiable and validated bibliographical records of the peer-reviewed scholarly literature in all areas of research;

B. A publication indicator with a system of weights that makes field-specific publishing traditions comparable across fields in the measurement of 'publication points' at the level of institutions;

C. A performance-based funding model which reallocates a small proportion of the annual direct institutional funding according the institutions' shares in the total of publication points.

In principle, component C is not necessary to establish components A and B. The experience is, however, that the funding models in C support the need for completeness and validation of the bibliographic data in component A. We will return to the Norwegian varianet of component C in the next section.

With regard to component A, the local CRIS systems were united to one integrated national system, CRIStin, in 2010. Around 160 institutions from the higher education sector, the independent institute sector and the hospital sector are participating. The publication indicator now serves PRFS for institutional funding of research in all three sectors. CRIStin covers all kinds of research activities and outputs and is running for several other purposes than the PRFS, such as CV's, applications to research councils, evaluations, annual reports, internal administration, bibliographies for Open Archives, links to full text, etc. The data for the publication indicator and the purpose of the PRFS si delimited by a definition of scholarly and scientific publishing. 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
4. in a publication channel (journal, series, book publisher) which represents authors from several institutions and organises 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. Suggestions for additions can be made online at any time through the same web page.

With regard to component B, the publication indicator itself, publications are given weights according to this table: In the measurement for the funding formula by the end of each year, the publications are weighted as they are counted. The intention is to balance between field specific publishing patterns, thereby making the publication output comparable across research areas and institutions that may have different research profiles. In one dimension, threemain publication types are given different weights: articles in journals and series (ISSN), articles in books (ISBN) and books (ISBN). In another dimension, publication channels are divided into two levels in order to stimulate publishing in the most prestigious and demanding publication channels within each field of research. The highest level is named 'Level 2'. It includes only the leading and most selective international journals, series and book publishers. There is also a quantitative restriction, since the publication channels selected for Level 2 can only in total represent up to 20% of the world's publications in each field. The weighting of publications by type and channel is shown in Table 5.

Table 5: Norway: weights of punlications by type and channel

| Type of publication | Level 1 (normal) | Level 2 (20 percent) |
|---|---|---|
| Article in ISSN title | 1 | 3 |
| Article in ISBN title | 0,7 | 1 |
| Book (ISBN title) | 5 | 8 |

Publication points are measured at the level of institutions, not at the level of individual researchers. Publications with multiple authors representing several institutions are counted only once. Their points are divided between the institutions by multiplying the points with the square root of the institution's proportion of the authors. This is an intermediate solution betweeh so-called fractional counts and whole counts that has been simulated to balance well between between the types of productivity in different fields and at the same time incentivise collaboration. As an extra incentive to international collaboration, a factor of 1,3 is used in the multiplication of points for publications with co-authors in other countries.

The list of journals, series and book publishers on 'Level 2' is revised annually in collaboration with national councils in each discipline or field of research. These councils propose changes to an interdisciplinary National Publishing Board, which governs the process on behalf of all institutions and has the final decision. Bibliometric statistics (world production versus national production in channels on both levels, and citation statistics for publication channels) are used as an aid in this process, but not as criteria in themselves.

### *7.3.3 Bibliometrics as a components in the funding system*

Norway is generally characterised by generous direct funding of higher education institutions as compared to external competitive funding from research councils and other sources. State external funding only accounts for 32 per cent. The rest, 68 per cent, is direct funding for teaching and research. In the paragraphs below, the 68 per cent direct funding is analysed further as a total of 100 per cent.

Among the direct funds, only 31 per cent of the direct institutional funding in Norway's higher education sector is reallocated according to performance indicators (Research PRFS and Teaching performance based). The remaining 69 per cent is block funding for education and research. Hence, the relative influence of the PRFS is only moderate in economic terms. In the next paragraph, the 31 per cent allocated through PRFS is analysed further as a total of 100 per cent.

*Within* the 31 percent based on performance assessment, the indicators representing educational activities are dominating and reallocate 85 per cent of this funding (teaching performance based). Performance indicators for research represent only 15 per cent. Within this small share, there are three indicators representing different types of external funding. These external funding indicators represent two thirds of the Research PRFS. For the remaining third, there is a *bibliometric* indicator representing research activity as measured by scholarly publishing. If measured as **a share of the total direct institutional funding** (excluding competitive external funding), the publication indicator therefore reallocates only 1,6 per cent of the total budget.

One publication point now represents around EUR 2,500, a little more than half of what it represented in 2005. Since then, the publication activity has almost doubled while the economic basis for reallocation by the indicator has remained the same.

### *7.3.4 Effects of the bibliometrics for the PRFS*

The Norwegian model has been extensively documented and evaluated, also in the international journal literature. Initiated in 2012 by the Norwegian Association of Higher Education Institutions (representing the funded organisations) in collaboration with the Ministry of Education and Research (the funding organisation), the bibliometrics for the funding model was evaluated extensively in Norway in 2013. An independent Danish team of researchers studied its design, organisation, effects, and legitimacy (Aagaard et al. 2014). As well as advising improvement and further development, the exercise provided the basis for four in-depth studies of internationally relevant questions (Aagaard 2015; Aagaard et al 2015; Schneider et al. 2015; Bloch & Schneider 2016).

The overall results of the evaluation showed an increase in productivity beyond what could be expected from increases in resources, particularly in fields and at institutions that had been less active in research before. However, the evaluation pointed at three major areas for improvement. Since 2014, the funded and funding organisations have collaborated on following up the evaluation in order to improve the model and its practices in these three areas.

One of them was the design of the indicator itself. The former indicator had a possible disincentive to collaboration (Bloch & Schneider 2016) and an imbalance in the representation of research fields (Aagaard et al 2015) which have been solved by a redesign of the indicator (the new indicator is presented above).

In 2016, the Ministry of Education and Research asked for a further redesign of the publication indicator with which it can include a citation indicator. There has been a consultation on the proposal for a solution. The results seem to imply that the publication indicator will continue as it is, and that the added citation indicator will not be used in the PRFS, but for analytical purposes only.

A second area for improvement was the fact that the indicator is also used locally (i.e. within the institutions themselves), in some contexts for purposes where it is not appropriate and can do harm (Aagaard 2015). This has been followed up establishing inter-institutional learning arenas for proper managerial use of the indicator and by agreeing on national recommendations for good conduct on the local level.

The third problem was partial lack of transparency of the process for the definition of the high-level publication channels (Level 2). This has been remedied by introducing a new interactive national portal for the indicator where all members of the committees are listed and where they explain their decisions and respond to comments and suggestions. The portal has a bilingual interface with information also in English: https://npi.nsd.no/.

### *7.3.5 Documentation, references and links*

The design of Norwegian model in 2004 and its evaluation in 2014 is documented in two publications from the Norwegian Association of Higher Education Institutions that can be downloaded at the portal of the Norwegian Publication Indicator.

The design: *A Bibliometric Model for Performance-based Budgeting of Research Institutions* (2004, in English translation): https://npi.nsd.no/dok/Vekt_pa_forskning_2004_in_english.pdf

The evaluation: Evaluation of the Norwegian publication indicator (2014, summary in English of a report in Danish):

https://npi.nsd.no/dok/eval2014/Evaluation_of_the_Norwegian_Publication_Indicator_2014_English_Summary.pdf

Updated information in English is available at the same portal:

https://npi.nsd.no/informasjon#dokumenter

Other relevant publications available in English are:

Aagaard K. (2015). How incentives trickle down: Local use of a national bibliometric indicator system. Science and Public Policy 42(5): 725-737.

Aagaard, K, Bloch CW and Schneider JW (2015). Impacts of performance-based research funding systems: The case of the Norwegian Publication Indicator. Research Evaluation 24(2): 106-117.

Ahlgren P, Colliander C and Persson O (2012). Field normalised citation rates, field normalised journal impact and Norwegian weights for allocation of university research funds. Scientometrics 92(3): 767-780.

Bloch C and Schneider JW (2016). Performance-based funding models and researcher behavior: An analysis of the influence of the Norwegian Publication level at the individual level. Research Evaluation 25(4): 371-382.

Schneider JW (2009). An outline of the bibliometric indicator used for performance-based funding of research institutions in Norway. European Political Science 8(3): 364-378.

Schneider JW, Aagaard K and Bloch CW (2015). What happens when national research funding is linked to differentiated publication counts? A comparison of the Australian and Norwegian publication-based funding models. Research Evaluation. DOI: 10.1093/reseval/rvv036

Sivertsen G (2016). Publication-Based Funding: The Norwegian Model. In: Ochsner M, Hug SE, Daniel HD (eds). Research Assessment in the Humanities: Towards Criteria and Procedures. Springer Open: Zürich, 79-90. https://link.springer.com/chapter/10.1007%2F978-3-319-29016-4_7

## *7.4 Sweden*

### *7.4.1 Background and motivation*

Sweden's PRFS is used for distribution of increased block grants between universities and university colleges. It has also been used for redistribution of between 10 and 20 per cent of the direct funding of all higher education institutions. It was implemented in 2009 on the basis of a commissioned report to the government in 2007 on institutional funding for research (SOU, 2007). The general aim of the new partly performance-based funding model was to enhance research quality while at the same time increasing the transparency of the funding allocation and the relative autonomy of the institutions in their own resource allocation. The report proposed several performance indicators among which two were implemented, one based on bibliometrics and another based on external funding. Only the biblometric indicator will be considered here.

### *7.4.2 Data source and indicator design*

Sweden's solution to bibliometrics for a PRFS is unique in Europe in the sense that the data for the indicator are not produced in collaboration with the institutions, but instead derived from (and defined by) only one specific commercial data source, Web of Science (WoS, from Clarivate Analytics). The main argument for using WoS was to give incentives to increase quality and internationalisation (Sandström & Sandström, 2008 and 2009). The bibliometric

model is operated on behalf of the government by the Swedish Research Council, using an annually updated replication database of WoS. The Swedish Research Council refines the data (disambiguation of institutional addresses, field classification, citation links) for the purpose of the PRFS and their own analytical needs.

The design of the bibliometric model for the PRFS was first proposed in an appendix to the above mentioned report (SOU, 2007) and then further developed by the same authors (Sandström & Sandström, 2008 and 2009). A short documentation in Swedish of the model as it actually works is available at the Swedish Research Council (Vetenskapsrådet, 2014).

The bibliometric indicator for the PRFS combines productivitiy and citation impact by using the number of publications and their received citations. Both are field-normalised and fractionalised according to the institution's contribution of authors to the article. The field normalisation of citations follows an advanced bibliometric procedure where actual citations are compared to the world average within the same field, year and publication type in the database.

As for the field normalisation of productivity (number of articles), there is no available standard bibliometric procedure since data representing the avaiable resources for research will be needed in addition to the bibliometric data. The solution in Sweden is to calculate a field factor, using the so-called Waring method, that estimates the average production of a researcher in each of the 34 subject fields of the model (Sandström and Sandström, 2009). The field factor takes into consideration the average productivity of Nordic or Swedish authors within the field, based on publications in the WoS database, while the Wahring method is a mathematical solution in which the same data is used to estimate the number of authors in the fields with publications not covered by the database. The latter estimates are particularly needed in the social sciences and humanities, where the output is insufficently covered by the WoS.

An alternative to the Waring method (with special treatment of the social sciences and the humanities) was initially proposed by the Swedish Research Council (Vetenskapsrådet, 2009), but has not been implemented, in spite of the Waring method being documented to suffer from flaws in the context of the Swedish PRFS (Vetenskapsrådet, 2010).

An initial simulation of the effects using WoS compared to the comprehensive coverage of scholarly publications by the Norwegian model demonstrated that the reallocation of funds would be much the same (Sandström & Sandström, 2009). A change of model would mainly affect universities with medical faculties since medicine is relatively well represented with publications and citations in WoS.

### *7.4.3 Effects of the bibliometrics for the PRFS*

The effects on the research system of the Swedish bibliometric model for the PRFS have not been studied, and the model itself has not been evaluated so far. It is generally consided a 'black box' with limited financial consequences for the institutions. The present government has even reduced its redistribution effects. However, using bibliometrics in the national PRFS has increased the local use of bibliometric information at the institutions, and a bibliometric component has been introduced in many places for internal redistribution of resources. In a recent study, the institutions were asked whether they use the Swedish or the Norwegian model for local purposes. Most of them seem to prefer the latter (Hammarfelt et al., 2016).

A commission appointed by the government will start working in May 2017 with one of its aims to design the next generation of PRFS in Sweden.

### *7.4.4 Dismissed proposals for alternatives*

Alternatives to the present bibliometric model have been discussed at two occasions in Sweden. In a commissioned report to the government, Anders Flodström (2011) proposed that the bibliometric indicator should be based on complete data produced in collaboration with the institutions. This solution would have been similar to the one initially proposed by the Research Council (Vetenskapsrådet, 2009) and to the parallel bibliometric indicators for PRFS in the other Nordic countries. Instead of following the proposal, the government assigned the Research Council to design a PRFS model based on peer review, inspired by the Research Excellence Framework in the United Kingdom. However, after the assignment resulted in a solution, the so-called FOKUS model (Swedish Research Council, 2015), the government decided not to implement it.

In 2016 the Research Council was assigned to suggest changes to improve the model (Vetenskapsrådet 2016) but these changes have not been implemented.

Hence, the WoS-based bibliometric indicator is still in place. Presently, the government is considering to extend the two existing indicators in the PRFS (WoS bibliometrics, external funding) with a third indicator representing (evaluation of) third stream activities (VINNOVA, 2016). This indicator would constitute the realisation of ideas presented also at the earlier stages (SOU, 2007; Flodström, 2011; Swedish Research Council, 2015).

### *7.4.5 Documentations, references and links*

The design of the Swedish bibliometric model is well documented in English in Sandström & Sandström (2009). A short documentation of how it actually runs is found in Vetenskapsrådet (2014). All references and links are given here:


Flodström, A. Prestationsbaserad resurstilldelning för universitet og högskolor. Rapport till Utbildningsdepartementet (U2011/7356/UH). Available at: http://www.ksla.se/wp-content/uploads/2012/03/7356-Rapport-Prestationsbaserad-resurstilldelning-f%C3%B6r-universitet-och-h%C3%B6gskolor.pdf

Hammarfelt, B., Nelhans, G., Eklund, P., & Åström, F. (2016). The heterogeneous landscape of bibliometric indicators: Evaluating models for allocating resources at Swedish universities. Research Evaluation 25(3): 292-305.

Sandström, U. & Sandsström, E. (2008). Resurser för citeringar. Rapport 2008:18 R, Högskoleverket.

Sandström, U. & Sandsström, E. (2009). The field factor: towards a metric for academic institutions. Research Evaluation, 18(3): 243-250.

SOU. (2007). Resurser för kvalitet (Resources for quality in research). SOU 2007:81. Available at: http://www.regeringen.se/rattsdokument/statens-offentliga-utredningar/2007/11/sou-200781/

Swedish Research Council (2015). Research Quality Evaluation in Sweden - Fokus: Report of a Government Commission regarding a Model for Resource Allocation to Universities and University Colleges Involving Peer Review of the Quality and Relevance of Research. Available at: https://publikationer.vr.se/wp-content/uploads/2015/05/VR1545_FOKUS_ENG_WEBB.pdf

Vetenskapsrådet. (2009). Bibliometrisk indikator som underlag för medelsfördelning. Available at: http://www.vr.se/download/18.72e6b52e1211cd0bba8800010145/bibliometrisk_indikator.pdf

Vetenskapsrådet. (2010). Kan man använda Waringmetoden för att uppskatta antalet forskare? Available at: https://publikationer.vr.se/produkt/kan-man-anvanda-waringmetoden-for-att-uppskatta-antalet-forskare/

Vetenskapsrådet. (2014). Bibliometriskt underlag för medelsfördelning. Available at: https://www.vr.se/download/18.5a947f0d145b21c1709f9d/1399897207262/Bibliometriskt+underlag+f%C3%B6r+medelsf%C3%B6rdelning_140512.pdf

Vetenskapsrådet. (2016). Regeringsuppdrag att justera indikatorerna vetenskaplig produktion och citeringar i modellen för resursfördelning. Available at: https://www.vr.se/download/18.73bacb0c15490b79b0ac5eaf/1462782033410/Svar+på+regeringsuppdrag+U2016-01350-F.pdf

Vinnova. 2016. Evaluating the role of HEI's interaction with surrounding society. VR 2016:09. Available at: http://vinnova.se/en/Publications-and-events/Publications/Products/Evaluating-the-Role-of-HEIs-Interaction-with-Surrounding-Society/

## *7.5 Italy*

### *7.5.1 Background and motivation*

Italy is characterised by a central government university funding system that relies mainly on a single grant, the Fondo di Finanziamento Ordinario (Ordinary Financing Fund, FFO), for teaching, research, and other infrastructural needs. Non-government sources such as student fees and contractual funding are becoming more important, but account for only around a quarter of the financial resources. The FFO is allocated according to a mixed model in which the largest part is based on historical data and a smaller part is based on a formula-based adjustment that takes into account performance indicators for teaching and research.

Most of the formula-based adjustment of the FFO for research is determined by a recurring research assessment exercise based on peer review of submitted publications. This part is inspired by the Research Assessment Exercise in the UK. Italy introduced it in 2006 as the Valutazione Triennale della Ricerca (VTR), which looked back at the period 2001-2003. The outcome was used by the Government since 2009 to allocate a small but growing share of public funding, starting from 2.2% of total funding in 2009 and reaching 16% in 2014. The second assessment exercise, Valutazione della Qualita della Ricerca (VQR, Evaluation of research quality), looked back at the years 2004-2010 and was published in July 2013 by a new independent agency Agenzia Nazionale per la Valutazione dell'Universita e della Ricerca (ANVUR). The third research assessment exercise (VQR 2011-14) has been started in 2015 with reference to the period 2011-14. It was published in February 2017 by ANVUR. There was a modification of the methodology between VTR 2001-2003 and the VQR 2004–2010. The most recent VQR had a similar framework as the previous one, with some adjustment and improvement.

### *7.5.2 Submitted output for peer review*

Italy has a central online system for submitting and archiving research output for the evaluation exercise. It is administered by ANVUR in collaboration with the universities and with technical support from Cineca.

To take the example of the third round of evaluation, the VQR evaluated the research outputs of all permanent scientific staff on government contracts in 96 universities and 39 research organisations. A total of 52,677 researchers submitted their best outputs (2 for each university researcher, and 3 for each scientist employed in a public research organisation). All in all, 118,036 outputs were submitted, out of which 70 per cent were journal articles and the remainder other types of research outputs such as books, book chapters, conference

proceedings (with ISBN codes); critical editions, translations, scientific comments; patents; compositions, designs, performance, work of arts, and others.

The outputs are classified by their authors in 16 research areas. ANVUR appoints a panel of experts for each research area. In the humanities and social sciences, a pure peer review system is applied with the auxilium of external (national and international) reviewers (almost 13,000 reviewers were used in the third round).

In science, technology, engineering and medicine (STEM), the same procedure is used, but in addition, bibliometric indicators are also produced by ANVUR to inform the panels. These indicators are limited to the part of the outputs that are covered by Web of Science or Scopus.

### *7.5.3 Use of bibliometrics in STEM fields*

The researchers have to indicate if they prefer to be evaluated using either WoS or Scopus. Indicators considered are the number of citations; the 5-years Impact Factor and the Article Influence Score (WOS database); the Impact per publication and the Scimago Journal Rank (Scopus database). All of the indicators are normalised and adjusted according to the state of the art in professional bibliometrics (for details, see Anfossi et al, 2016)

In principle, then, all journal articles covered by Scopus and Web of Sciences are initially assessed on the basis of variants of two indicators representing their citation frequency and the citation impact of the journal in which they are published. The measurement is translated into the five classes of merit: Excellent; Good; Fair; Acceptable; Limited. The result is validated by the experts in order to determine the final weight attributed to the publication. If the two bibliometric indicators are clearly diverging, the article is evaluated once more with traditional peer review. In any case, according to the Ministerial decree, at least 50% of the publication submitted for evaluation should be peer-evaluated.

Consequently, the system for evaluating outputs for the Italian PRFS can be characterised as based on peer review, but the burden of large-scale reviewing is somewhat relieved by the support of a systematic, centralized and professional use of bibliometrics in the STEM fields.

### *7.5.4 Metrics for third stream activities*

ANVUR is developing, on an experimental basis, a new system of metrics designed to support evaluation of third stream activities and impact of academic research. Third stream activities have been divided in two main areas, respectively involving research economic valorisation and the production of public and social goods. As for research valorisation, indicators are produced concerning intellectual property management (patents and vegetal varieties), academic entrepreneurship (spin offs), third party activities, and intermediation activities. As for the production of public and social goods, ANVUR considers indicators concerning the management of cultural activities and the cultural heritage (museums, archeological excavations and cultural heritage), clinical trials, continuous education and public engagement. Evaluation is based on peer review, informed by the aforementioned information.

### *7.5.5 Effects of the bibliometrics for the PRFS*

There has been no systematic evaluation of the Italian PFRS as such, but some studies have been published discussing possible effects or changes in the system. Two examples:

After the second round of evaluation (the VQR covering 2004-2010), Benedetto et al. (2016) found that the ratings crucially depend on language and typology of publication and on the methodology adopted for evaluation. Younger researchers and full professors tend to receive, ceteris paribus, a better evaluation. Outcomes submitted by researchers working in the Centre-North of the country usually obtain better evaluations than those in the South. Finally, public funding to university research is found to be correlated with positive ranking in the

assessment exercise, especially in natural sciences and engineering, where there is particular need of appropriate funding for conducting laboratories experiments and research.

In addition to these observations, Ancaiani et al. (2015) found a significant degree of concordance among peer review and bibliometric evaluations, supporting the choice of using both techniques in order to assess the quality of Italian research institutions. They also found that, in general, the results of the VQR (second round) were already highly informative about the existing strength and weaknesses of the Italian university research system.

### *7.5.6 Documentations, references and links*


ANVUR. (2013.) Rapporto sullo stato del sistema universitario e della ricerca. Available at: http://www.anvur.org/attachments/article/644/Rapporto%20ANVUR%202013_UNIVERSITA%20e%20RICERCA_integrale.pdf

ANVUR (2015.) Evaluation of Research Quality 2011-2014 (VQR 2011-2014). Call for participation. Available at: http://www.anvur.org/attachments/article/825/Call%20VQR_2011_2014_11nov_~.pdf

ANVUR (2016.) Rapporto sullo stato del sistema universitario e della ricerca. Available at: http://www.anvur.org/attachments/article/1045/ANVUR_Rapporto_INTEGRALE_~.pdf

Ancaiani, A., Anfossi, A.F., Barbara, A., Benedetto, S., Blasi, B., Carletti, V., Cicero, T., Ciolfi, A., Costa, F., Colizza, G., Costantini, M., di Cristina, F., Ferrara, A., Lacatena, R. M., Malgarini, M., Mazzotta, I., Nappi, C.A., Romagnosi, S., Sileoni, S. (2015.) Evaluating scientific research in Italy: The 2004-10 research evaluation exercise. Research Evaluation, 24(3): 242-255.

Anfossi, A., Ciolfi, A., Costa, F., Parisi, G., Benedetto, S. (2016.) Large-scale assessment of research outputs through a weighted combination of bibliometric indicators. Scientometrics, 107(2): 671–683.

Benedetto, S., Cicero, T., Malgarini, M. (2016). Determinants of research quality in Italian universities: Evidence from the 2004 to 2010 evaluation exercise. Research Evaluation, 25(3): 257–263.

Geuna, A., Piolatto, M. (2016). Research Research assessment in the UK and Italy: Costly and difficult, but probably worth it (at least for a while). Research Policy, 45(1): 260-271.

# 8 APPENDIX A: REFERENCES

In addition, references are given for each of the five countries in Chapter 7.

Geuna, A. & Martin B.R. (2003). University research evaluation and funding: an international comparison. Minerva, 41: 277-344.

Hicks, D. (2012.) Performance-based university research funding systems. Research Policy, 41: 251-261.

Hicks, D., Wouters, P., Waltman, L., de Rijcke, S., Rafols, I. (2015). Bibliometrics: The Leiden Manifesto for Research Metrics. Nature, 520: 429–431.

Jonkers, K. & Zacharewicz, T. (2016). Research Performance Based Funding Systems: a Comparative Assessment. EUR 27837 EN; doi:10.2791/659483.

Swedish Government. (2016). Regeringens proposition 2016/17:50. Kunskap i samverkan – för samhällets utmaningar och stärkt konkurrenskraft. Stockholm: Regeringen.

Swedish Research Council. (2015). Research Quality Evaluation in Sweden - Fokus: Report of a Government Commission regarding a Model for Resource Allocation to Universities and University Colleges Involving Peer Review of the Quality and Relevance of Research. Stockholm: Swedish Research Council.

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

## Getting in touch with the EU

IN PERSON

All over the European Union there are hundreds of Europe Direct Information Centres.
You can find the address of the centre nearest you at: http://europa.eu/contact

ON THE PHONE OR BY E-MAIL

Europe Direct is a service that answers your questions about the European Union.
You can contact this service
– by freephone: 00 800 6 7 8 9 10 11 (certain operators may charge for these calls),
– at the following standard number: +32 22999696 or
– by electronic mail via: http://europa.eu/contact

## Finding information about the EU

ONLINE

Information about the European Union in all the official languages of the EU is available on the Europa website at: http://europa.eu

EU PUBLICATIONS

You can download or order free and priced EU publications from EU Bookshop at: http://bookshop.europa.eu. Multiple copies of free publications may be obtained by contacting Europe Direct or your local information centre (see http://europa.eu/contact)

EU LAW AND RELATED DOCUMENTS

For access to legal information from the EU, including all EU law since 1951 in all the official language versions, go to EUR-Lex at: http://eur-lex.europa.eu

OPEN DATA FROM THE EU

The EU Open Data Portal (http://data.europa.eu/euodp/en/data) provides access to datasets from the EU. Data can be downloaded and reused for free, both for commercial and non-commercial purposes.

This MLE report covers the use of bibliometrics in the context of PRFS. It sets out some key principles of good practice and describes typical problems and considerations. In-depth descriptions are provided of concrete solutions that were adopted in five MLE-participating countries, i.e. Croatia, Estonia, Norway, Sweden and Italy.