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

STI 2017, Paris, 8 September 2017

Gunnar Sivertsen

Nordic Institute for Studies in Innovation, Research and Education, Oslo, Norway
1. Background:
   - The proliferation of PRFS in Europe

2. Analytics:
   - Evaluation-based versus indicator-based PRFS

3. Ethics:
   - How to engage with instruments for New Public Management?

4. Advice:
   - Bibliometrics for evaluation-based PRFS: Ten useful guidelines
   - Bibliometrics for indicator-based PRFS: Ten problems and considerations

5. Conclusion: The relevance for a network of indicator designers (ENID)
1. Background:
   - The proliferation of PRFS in Europe

2. Analytics:
   - Evaluation-based versus indicator-based PRFS

3. Ethics:
   - How to engage with instruments for New Public Management?

4. Advice:
   - Bibliometrics for evaluation-based PRFS: Ten useful guidelines
   - Bibliometrics for indicator-based PRFS: Ten problems and considerations

5. Conclusion: The relevance for a network of indicator designers (ENID)
PRFS before 2003
Red: Evaluation-based, Blue: Indicator-based
Geuna & Martin, Minerva 2003
PRFS before 2010
Red: Evaluation-based, Blue: Indicator-based
Hicks, Research Policy, 2012
PRFS before 2016

Red: Evaluation-based, Blue: Indicator-based
Jonkers & Zacharewicz, European Commision, 2016
Outline
PRFS = Performance-based Research Funding Systems for institutions

1. Background:
   - The proliferation of PRFS in Europe

2. Analytics:
   - Evaluation-based versus indicator-based PRFS

3. Ethics:
   - How to engage with instruments for New Public Management?

4. Advice:
   - Bibliometrics for evaluation-based PRFS: Ten useful guidelines
   - Bibliometrics for indicator-based PRFS: Ten problems and considerations

5. Conclusion: The relevance for a network of indicator designers (ENID)
Four examples: United Kingdom

Red: Evaluation-based; Purposes: Research evaluation and institutional funding

Seven major research assessments since 1986.

Originally, evaluation was the method and funding allocation was the purpose. Now, the method has become an even more important purpose.

*The Metric Tide* report (Wilsdon et al., 2015): “Metrics should support, not supplant, expert judgement.”
Four examples: The Netherlands
Purpose: Research evaluation

The indicator-based funding model for the universities does NOT include research performance.

The research evaluation exercises do NOT influence institutional funding.

This country does NOT have a PRFS.
Four examples: Sweden 2009-2014

Blue: Indicator-based Purpose: Institutional funding

PRFS reallocation a small portion of institutional funding based two indicators:

1) External revenues
2) Publications and citations in Web of Science
Four examples: Sweden in 2014: A report to the government

Red: Evaluation-based Purposes: Research evaluation and institutional funding

A UK-inspired model was designed by the Research Council and presented to the government.
Four examples: Sweden in 2016: FOKUS not implemented

Blue: Indicator-based Purpose: Institutional funding

PRFS reallocation a small portion of institutional funding based two indicators:

1) External revenues
2) Publications and citations in Web of Science
**Four examples:** Sweden in 2016: The report was not implemented

**Blue:** Indicator-based  
**Purpose:** Institutional funding

A UK-inspired model was designed by the Research Council and presented to the government.
Four examples: Norway since the 1990’s

Red: Evaluation-based Purpose: Research evaluation, not funding

National research assessments inspired by the UK are performed by intervals.

Like in the Netherlands, they do not influence funding.
The indicator-based model reallocates a small portion of the funding.

It is based on several indicators. Most of them represent education, some represent research.

The publication indicator reallocates 1.6 per cent
The two purposes of PRFS may have different emphasis: Research evaluation and funding.

Hicks (2012) defines PRFS as related to both purposes. They are “national systems of research output evaluation used to distribute research funding to universities.”
1. Background:
   - The proliferation of PRFS in Europe

2. Analytics:
   - Evaluation-based versus indicator-based PRFS

3. Ethics:
   - How to engage with instruments for New Public Management?

4. Advice:
   - Bibliometrics for evaluation-based PRFS: Ten useful guidelines
   - Bibliometrics for indicator-based PRFS: Ten problems and considerations

5. Conclusion: The relevance for a network of indicator designers (ENID)
Best practice or mutual learning?

The Metric Tide report (Wilsdon et al., 2015): “Metrics should support, not supplant, expert judgement.”
Best practice or mutual learning?

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. Country differences in the design of a PRFS and its motivations should be expected and respected.

*The Metric Tide* report (Wilsdon et al., 2015): “Metrics should support, not supplant, expert judgement.”
The motivations differ among countries and are context-specific. They cannot be reduced to a general wish to implement NPM. 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. Country differences in the design of a PRFS and its motivations should be expected and respected.
Effects of PRFS depend on their design and implementation
Outline
PRFS = Performance-based Research Funding Systems for institutions

1. Background:
   - The proliferation of PRFS in Europe

2. Analytics:
   - Evaluation-based versus indicator-based PRFS

3. Ethics:
   - How to engage with instruments for New Public Management?

4. Advice:
   - Bibliometrics for evaluation-based PRFS: Ten useful guidelines
   - Bibliometrics for indicator-based PRFS: Ten problems and considerations

5. Conclusion: The relevance for a network of indicator designers (ENID)
The “Leiden Manifesto”
Bibliometrics for evaluation-based PRFS: Ten useful guidelines

1. Quantitative indicators cannot replace the judgment of expert assessors, but they can be used to help support them.

2. Evaluation of research activity has to adapt to the mission and objectives of the institution, individual or group being evaluated.

3. Indicators need to be developed that reflect the impact of research activities locally and regionally, and those that are developed in languages other than English.

4. The data collection and analysis processes have to be open, transparent and simple.

5. Those evaluated have to be able to verify the analysis of the indicators being used for the evaluation and, if they disagree, request re-evaluation.

6. The differences existing in terms of impact in different fields of research have to be taken into account when producing indicators.

7. Individual evaluation of researchers has to be based on qualitative assessment of their portfolio. Indicators cannot be used without taking into account the researcher’s context.

8. False precision and misplaced concreteness must be avoided.

9. The effects of certain indicators as incentives for certain activities and disincentives for others must be taken into account.

10. The indicators have to be reviewed and updated regularly.
1. Quantitative indicators cannot replace the judgment of expert assessors, but they can be used to help support them.

2. Evaluation of research activity has to adapt to the mission and objectives of the institution, individual or group being evaluated.

3. Indicators need to be developed that reflect the impact of research activities locally and regionally, and those that are developed in languages other than English.

4. The data collection and analysis processes have to be open, transparent and simple.

5. Those evaluated have to be able to verify the analysis of the indicators being used for the evaluation and, if they disagree, request re-evaluation.

6. The differences existing in terms of impact in different fields of research have to be taken into account when producing indicators.

7. Individual evaluation of researchers has to be based on qualitative assessment of their portfolio. Indicators cannot be used without taking into account the researcher’s context.

8. False precision and misplaced concreteness must be avoided.

9. The effects of certain indicators as incentives for certain activities and disincentives for others must be taken into account.

10. The indicators have to be reviewed and updated regularly.
1. **Background:**
   - The proliferation of PRFS in Europe

2. **Analytics:**
   - Evaluation-based versus indicator-based PRFS

3. **Ethics:**
   - How to engage with instruments for New Public Management?

4. **Advice:**
   - Bibliometrics for evaluation-based PRFS: Ten useful guidelines
   - Bibliometrics for indicator-based PRFS: Ten problems and considerations

5. **Conclusion:** The relevance for a network of indicator designers (ENID)
Bibliometrics for indicator-based PRFS:
Ten problems and considerations: 1

- Bibliometrics is not ‘objective’.
- The idea that one indicator based on one particular data source might solve the problem or promote the best research should be regarded as subjective as long as it is not tested and discussed.
- The available data sources and indicators for bibliometrics may be problematic already at the outset if the aim is to give a balanced representation of performances in all areas of research and all types of research organizations.
- Main consideration: To avoid power games in the design process.

Background:
Bibliometrics for indicator-based PRFS: Ten problems and considerations: 2

- The viable solution will probably not come from a single desktop.
- Design the PRFS in dialogue between the funder and the funded organizations.
- Represent all areas of research in the process.
- Make use of (commercially) independent bibliometric expertise
- **Main consideration:** Participation and transparency.
The effects of a PRFS may be strong even with a limited economic influence.

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 and indicators define what counts.

Available data sources such as WoS or Scopus represent definitions and delimitations.

If data are not defined by a chosen data source, an explicit definition is needed along with a set of reporting instructions and some monitoring of the reporting practices.

Main considerations: data quality; disinterested data production; incentives for internationalisation; costs; comprehensiveness; balanced representation of all fields; the representation of national language publishing.
Indicators must be defined.

Both publication, collaboration and citation indicators may be 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.
Field normalization or balances are needed.

The well-established field normalization methods for citation indicators need to be supplemented with a balanced representation of productivity across fields.

This can be solved in the design of the indicators, or with balances in the funding formula itself, or by separating field-specific streams of funding.

Main considerations: Institutions with different profiles of specialization, e.g. a technical university versus a general university, need to be treated equally. The funding mechanism should be acceptable across fields.
Counting methods can create biases.

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

Main considerations: to balance the indicators across subfields with different co-authorship practices, and to incentivize collaboration without stimulating the inclusion of authors with minimal contributions.
Weighing of publication types can be necessary.

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.

Main considerations: to balance the indicators across subfields with different publication practices, and to incentivize a favourable development of those practices.
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.

Main considerations: 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.
Bibliometric indicators designed for the *macro level* (institutions) are often not adequate at the level of research groups or individuals.

Applications at the *micro level* should instead follow the ten principles of the *Leiden Manifesto*. 
Outline
PRFS = Performance-based Research Funding Systems for institutions

1. Background:
   - The proliferation of PRFS in Europe

2. Analytics:
   - Evaluation-based versus indicator-based PRFS

3. Ethics:
   - How to engage with instruments for New Public Management?

4. Advice:
   - Bibliometrics for evaluation-based PRFS: Ten useful guidelines
   - Bibliometrics for indicator-based PRFS: Ten problems and considerations

5. Conclusion: The relevance for a network of indicator designers (ENID)
A bibliometric indicator
with a balanced representation of all fields

Gunnar Sivertsen

gunnar.sivertsen@nifi.no
Nordic Institute for Studies in Innovation, Research and Education (NIFU)
P.O. Box 2815 Tøyen, N-0608 Oslo, Norway

Abstract
As research in progress, we present two studies aimed at redesigning the bibliometric indicator of the “Norwegian Model” as response to an evaluation in 2013. The indicator is supposed to give a balanced representation of all fields, also those that are constructed as “peripheral” in traditional bibliometrics because of limited coverage in databases. The first study deals with balancing between different field-dependent co-authorship practices in the indicator, the other with the possible addition of a measurement of citation impact that could be applicable across all fields.

Keywords
Bibliometric indicators; productivity; citation impact; co-authorship; fractionalization; publication patterns; evaluation; the Norwegian model.

Submission type: Research in progress paper.
Relevant track: Data infrastructure for research metrics.
Conclusion: The relevance for a network of indicator designers (ENID)

- PRFS designs are embedded in national contexts and policies
- They may do harm, or they may have other effects, depending on how they are designed and implemented
- A manifesto does not seem possible in this case
- Reflexivity, criticism and academic distance to practice is already abundant
- Engagement of independent experts is needed