# Toward an excellence-based research funding system: Evidence from Poland

Emanuel Kulczycki [a], Marcin Korzeń [b], Przemysław Korytkowski [b],*

[a] *Adam Mickiewicz University in Poznań, Poland*
[b] *West Pomeranian University of Technology in Szczecin, Poland*

A B S T R A C T

This article discusses the metrics used in the national research evaluation in Poland of the period 2009–2012. The Polish system uses mostly parametric assessments to make the evaluation more objective and independent from its peers. We have analysed data on one million research outcomes and assessment results of 962 scientific units in the period 2009–2012. Our study aims to determine how much data the research funding system needs to proceed with evaluation. We have used correlation analysis, multivariate logistic regressions models and decision trees to show which metrics of the evaluation played a major role in the final results. Our analysis revealed that many metrics taken into account in the evaluation are closely correlated. We have found that in the Polish system, not all the collected data are necessary to achieve the main goal of the system, namely the categorization of scientific units in terms of their research performance. Our findings highlight the fact that there is a high correlation between performance in terms of publications and the scientific potential of a given scientific unit. We conclude with recommendations and a suggestion of a transition from a system in which the scientific units report all their metrics to a system in which they show only the most important metrics that meet the requirements of excellence in research.

© 2017 Elsevier Ltd. All rights reserved.

## 1. Introduction

Performance-research funding systems (PRFSs) have been used as a science policy tool for the last three decades. The Research Assessment Exercise in the UK was launched in 1986. In the early 1990s, the assessment of research institutions began in Poland. Since then, many countries have introduced PRFSs and have embedded them in their national research systems (Bloch & Schneider, 2016; Hicks, 2012). The aim of developing and implementing PRFSs is to allocate funds to excellent institutions. To determine such institutions, research outcomes are evaluated *ex post*. The measurement methods which are used within such an evaluation can be divided into three categories (Aagaard, Bloch, & Schneider, 2015): peer review-based models, publication count-based models and citation-based models. The peer-review model is used in the UK's Research Excellence Framework. The other two methods, based on publication count and publication citations, are used in most other PRFSs (e.g. in the Czech Republic, Finland, Belgium/Flanders, Italy, Norway and Poland).

The performance of the scientific unit is constituted by various research outcomes, such as publications, projects, organised conferences and others. One could identify the most significant part of such a performance, but when we evaluate

---

* Corresponding author.
  *E-mail address:* pkorytkowski@zut.edu.pl (P. Korytkowski).

a scientific unit's performance, we measure all of its activities. In the PRFSs, however, such an evaluation policy did not strongly emphasize whether the outcomes were actually desirable from the funders' point of view. Contrary to the concept of performance, the concept of "excellence in research" indicates that only some parts of the performance are desirable from the funders' point of view. Most researchers and stakeholders agree that excellence should be rewarded; yet, the challenge of how to define and quantify such "excellence" remains. (Arthur, 2015; Sunkel, 2015). In the UK, the PRFSs shifted from outcomes assessment in the Research Assessment Exercise to measurement of impact and outcome excellence in the Research Excellence Framework (Chowdhury, Koya, & Philipson, 2016). This transition was conducted in the PRFS in which the evaluation had been undertaken using the peer-review model. Such a model takes for granted that in the evaluation process only the most important outcomes should be assessed.

Hicks (2012) highlights that the complexity of PRFSs has increased over time, as they are dynamic systems being expanded by the addition of new indicators, such as h-index or altmetrics. When a system becomes more complicated, its indicators and metrics can be gamed (Rijcke, de Wouters, Rushforth, Franssen, & Hammarfelt, 2016), and it becomes even harder to legitimize such a model of research evaluation. Thus, as the authors of the "Leiden Manifesto" (Hicks, Wouters, Waltman, de Rijcke, & Rafols, 2015) suggest, the indicators should be regularly scrutinized and updated. In this way, changing and improving PRFSs always face challenges such as keeping a balance between peer review and metrics, universal and specific field models, considering all activities and only important ones and redistributing funding to the best performers and motiving all performers. This is possible to achieve in a variety of ways, such as differentiated publication counts (Schneider, Aagaard, & Bloch, 2014) or expert-based and citation-based ranking of publication channels (Saarela, Kärkkäinen, Lahtonen, & Rossi, 2016).

Several countries, such as Australia, Brazil, France, Italy and Poland, have developed national journal rankings in which all the disciplines are represented (Ferrara & Bonaccorsi, 2016; Haddow & Genoni, 2010; Vanclay, 2011). These rankings serve to indicate which publications 'do count' in a national evaluation. Publications form the greatest part of evaluated research outcomes. For instance, the Italian Research Evaluation assessment for the period 2004–2010 analysed almost 184,000 publications and 1000 other research outcomes (Ancaiani et al., 2015). The Czech PRFS evaluated over 97,000 publications and 10,000 other research outcomes for the period 2005–2009 (Good, Vermeulen, Tiefenthaler, & Arnold, 2015). The Polish PRFS evaluated over 184,000 publications and 182,000 other research outcomes for the period 2009–2012.

Even though publications are the most significant component in the evaluated outcomes, there are many other scientific activities that characterize well-performing institutions and determine the level of their productivity. Among these scientific activities are granted patents, acquired projects, artistic productions, scientific degrees awarded, scholarships received, editorial engagements, research infrastructures, scientific conferences organized and others. As many studies have also shown, the chosen category of scientific activity used within the PRFS has a significant influence on evaluating research outcomes (Abramo & D'Angelo, 2015; Ancaiani et al., 2015; Good et al., 2015; Sivertsen, 2015).

The current Polish PRFS is based on lessons from previous evaluations as well as consultations with the Polish scientific community. The Polish Ministry of Science and Higher Education declares that the main objective of this system is funding distribution to scientific units according to their scientific achievements. The quality of a given scientific unit is expressed as one of four categories: A+, A, B or C. The best scientific units, the A+ category, receive much greater financing, 150% of that allocated to the A category. Scientific units with the B category receive 70% of the funds allocated to the A category units and those with C category only 40% and only for half a year. The intention of this system was to strengthen scientific performance. This aim has been achieved, among others, by reducing the number of reported publications for the four-year reporting period.

During the consultations, many of the scientific units claimed that the evaluation system should take into account all aspects of their activities. In this way, the number of PRFS parameters has been inflated in Poland. This opened space for parameter optimization and the gaming of the system. On one hand, scientists criticize the Polish system for its excessive complexity, and, on the other hand, they often demand the addition of more parameters specific to their field of study or that their importance be bolstered by increasing the number of points awarded to a parameter.

In 2015, the Polish government published a strategy entitled the *Higher Education and Science Development Programme for 2015–2030* (MNiSW, 2015), which has established goals for current science policy. In this strategy, it is explicitly stated that the aim of the research evaluation system is to identify various aspects of excellence in research. In the previous strategy published in 2008, the concept of excellence was not used: performance was the very idea on which the research evaluation was focused. Moreover, in the current strategy, one can find that "in the parametric evaluation of scientific units, the quality of research outputs (especially prestigious publications and financially rewarded implementation) should be more important than quantity (. . .) In the evaluation, there should be acknowledged only the best research outcomes, and a number of the outcomes should not be just a simple multiplicity of the full-time employment equivalents" (MNiSW 2015, p. 22). Focusing only on the best and most important outcomes is relevant from the Polish science policy point of view. Poland's publication output is below the average European Union levels; for example, the percentage of highly cited publications for Poland is 6.36%, whereas the average EU level is 12.25% (Klincewicz & Szkuta, 2016). Thus, the improvement of Polish PRFS has to face the following challenge: if we agree that 'what gets measured gets done', then we should focus only on such types of research outcomes that achieve the goals of science policy.

Our study aims to answer the following question on the assessment of scientific institutions in Poland by using the collected datasets in the evaluation process: which metrics (and data) does a performance-based research funding system actually need? We have assumed that reducing the number of metrics or parameters would allow focusing on those types