Response to 'How good is research really?'

Bornmann and Marx [1] analyse the use of bibliometric parameters to determine "how good the research is", advising against the use of certain scientometric indicators, such as the impact factor and the *h*-index, and proposing percentiles as an alternative. However, they do so by using debatable arguments and ignoring validation, which is the basis for any meaningful indicator. The need to validate scientometric predictors is based on general principles [2] and is especially necessary in the case of science owing to the manner in which it advances: a large volume of "normal" research supports the small fraction of "revolutionary" research that pushes scientific progress [3]. Thus, Bornmann and Marx's "good research" could refer to either "normal" or "revolutionary" research, or both [4]. If this uncertainty is not resolved, any further discussion about a reliable indicator of good research is futile.

The reliability of research indicators can be assessed in the light of one of their major applications: "to measure the effectiveness of research expenditures" [5]. To compare two institutions, we should divide the values for their research indicator by their corresponding research expenditures, and the one with the higher ratio can be said to be more cost-effective. However, for this calculation to be meaningful, the indicator must be sensitive to research that is relevant for the progress of science and that provides benefits for society, and insensitive to the large volume of "normal" results that are important for researchers but not for society at large [4]. One possibility to assess the results of a comparison is to assume that institutions obtaining Nobel Prizes are more successful in producing research that is relevant to society. Surprisingly, when applying this simple test to most conventional research indicators, the finding is that the most competitive elite universities show a low effectiveness in their use of research expenditures [4], which suggests that in these indicators the weight of "normal" research is high.

Bornmann and Marx [1] propose the proportion of the top 10% most cited papers as research indicators. Because validation is not reported, it is unknown why the top 10% is the correct percentile and not, for example, the top 5% or 20%. Yet, the reliability of this indicator can be tested by comparing universities, which is possible because it is used in the Leiden (PP_{top10%} indicator) and SCIMAGO (Exc indicator) rankings [6]. Massachusetts Institute of Technology (MIT), USA, and Complutense University of Madrid, Spain are good examples because there are previous comparisons [4]. The values for these universities are: PP_{top10%} 25.2% and 8.1% (<http://www.leidenranking.com/ranking.aspx>), and Exc 29.4% and 12.4% (<http://www.scimagoir.com/>), respectively, which gives MIT and Complutense indicator ratios of 3.1 and 2.4, respectively. Introducing the number of papers, the ratios of the transformed indicators increase to 7.1 and 5.4, respectively. Taking into account that research expenditure is ten times higher at MIT, the conclusion is that Complutense University is more effective than MIT in the use of research funding, which is not reasonable. In fact, it has been proposed that ratios for reliable research indicators should be at least 50 [4], and much higher ratios have been reported for indicators validated in terms of Nobel Prizes [7]. Similar comparisons can be made with other universities.

The research indicators proposed by Bornmann and Marx [1] join a large number of published, non-validated bibliometric indicators of research performance. Their expansion and prevalence is probably a

cause for the extended feeling among scientists that there is a lack of reliable research indicators, a feeling that is summarized in a 2011 report by the Royal Society, UK: "Better indicators are required in order to properly evaluate global science" [8]. A question recently raised by Francis Narin—an expert on science and technology analysis—relates to the same problem: "for both old and new indicators, basic validity and relevance issues remain, such as by what standard can we validate our results, and what external use can appropriately be made of them?" [9].

CONFLICT OF INTEREST

The author declares that he has no conflict of interest.

REFERENCES

1. Bornmann L, Marx W (2013) *EMBO Rep* **14**: 226–230
2. Harnad S (2009) *Scientometrics* **79**: 147–156
3. Kuhn TS (1962) *The Structure of Scientific Revolutions*. University of Chicago Press
4. Rodríguez-Navarro A (2012) *PLoS ONE* **7**: e47210
5. Garfield E, Welljams-Dorof A (1992) *Sci Public Policy* **19**: 321–327
6. Leydesdorff L, Bornmann L (2012) *Scientometrics* **92**: 781–783
7. Rodríguez-Navarro A (2011) *PLoS ONE* **6**: e20510
8. The Royal Society (2011) *Knowledge, Networks and Nations*. The Royal Society
9. Narin F (2012) *Scientometrics* **92**: 391–393

**Alonso Rodríguez-Navarro is at the Centro de Biotecnología y Genómica de Plantas, Universidad Politécnica de Madrid, Spain.
E-mail: alonso.rodriguez@upm.es**

EMBO reports (2013) **14**, 493; published online 14 May 2013; doi:10.1038/embor.2013.60

Comments to the response of Rodríguez-Navarro

In his critical response [1] to our article 'How good is research really?', Alonso Rodríguez-Navarro lists several points that we address here. First, he criticizes that we use debatable arguments for

our preference for percentiles and that we ignore the validity of the approach. Unfortunately, he does not specify which of the arguments he finds debatable. We believe that there are many good reasons