Which Metric Is More Appropriate to Evaluate Researchers?

MAJID KHADEM-REZAIYAN, MALIHEH DADGAR MOGHADAM

Department of Community Medicine, Faculty of Medicine, Mashhad University of Medical Sciences, Mashhad, Iran

Iranian medical universities choose their best researchers in each field annually. The protocols of this process have been modified quite often, but the changes were not fundamental and did not lead to all-inclusive evaluation tools. The recent article in Asia Pacific Journal of Medical Toxicology, which proposes a scoring scale for evaluation of scientist's impact called "360-degree researcher evaluation score" (1), not only opens a new window for detailed evaluation of researchers' products and creations, but also provides a basic platform for promoting the research. In community medicine, there is a strict view to primarily address upstream causes of health problems while we look for solutions for downstream ones (2). It seems that this view has been considered in the designing process of this new scale, as for example, some neglected criteria which build capacity for science production (upstream causes of low science production) are taken into account.

We really appreciate the holistic view of the scale, but we believe that following suggestions would help to improve its utility:

1. It seems that the parameter of "number of downloads of articles" is not an appropriate criterion, because some journals do not report this measure. Besides, its value in scholarly communication is still under debate and it can be easily manipulated by the researcher himself (3).

2. In calculating scores for journal articles, the calculation method is based on dividing impact factor (IF) by 30. Many journals have low IFs of just 0.1 or 0.2 and this calculation can make too many decimals. However, the good news is that with this method of calculation, there is more emphasis on the design of the study than the IF of the publishing journal.

3. The measurement method is based on "ranking in institution". As some parameters only receives 1 score at maximum and there may be several staff members in one department, the acquired score for one person may be too small and calculation of the score of several persons by comparing them with each other can be difficult.

2. Cost benefit and cost effectiveness are not two distinct concepts, they are just two different approaches to a unique comparison analysis (4). Therefore, it seems that taking both of them as two different criteria for evaluation of a research will overestimate a single effect in the total score.

In "societal impact" domain:

1. Although, we declare that considering the recommended parameters in this domain for evaluation of research impacts will have enormous effects on guiding the global projects to more efficient ones, these parameters are not clearly and objectively defined. How can somebody, for example, determine the amount of increased life expectancy from a single specific research?

2. It seems that the four-fold score for international versus national conferences is a bit underestimating the value of national ones.

Taken together, the proposed scale is a well-designed protocol for considering the most important dimensions of research. However, ensuring the reliability and validity of this tool requires further studies.

REFERENCES