EMBO
reports

editorial

Measuring success

There is an adage that I first heard more than 30 years ago that, "everything you measure gets better". The current enthusiasm for measuring is certainly high, so we can anticipate great improvements—if the measurements have a meaning.

When considering the metrics we use to measure the quality of science, a good starting point is the ranking of journals, given their crucial role in evaluating academic research. I believe that there is, in general, a real difference between the quality of papers that are accepted by journals with high or low impact factors (IF) in the same field. But, that does not mean that all papers in high IF journals are superior to those in lower IF journals, nor is there an inevitable link between the IF of a journal and the number of times a paper it publishes will be cited (Gannon F (2000) *EMBO Rep* **1:** 293). Yet, many of the committees that evaluate scientists take a short cut—consciously or otherwise—and base their decisions on the number of papers a researcher has published in iconic journals. Clearly, more care should be taken to evaluate individual scientists, especially because an applicant's field of research or the stage of their career can have distorting effects. However, other methods of evaluation also have flaws: the Hirsch factor, for example, has difficulty measuring the quality of those at an early stage in their career with few publications. Of course, an expert, in-depth analysis of an individual's contribution to science is always preferable, but not always possible or practical; hence the use of surrogate measurements that evaluate the location of published papers rather than their content.

Moving up the organizational level from scientists to institutes, rankings take even more measures into account. Clearly, an institute with a solid output of papers accepted by high-ranking journals is probably better than one that has only occasional successes. However, rankings do not stop at the comparative measurement of publications; for example, one popular system awards points for the number of Nobel prize-winners associated with the university. This is irrelevant given that there are only a few awards made each year and that some areas, such as the social sciences or mathematics, are excluded from the prize. Moreover, the Nobel Prize is restricted to three winners—irrespective of the actual number of pioneers in a field—and there is a real difference between a single important experiment and a lifetime of making important contributions to science. Nonetheless, and not surprisingly, the entrepreneurial universities are signing up, part-time, Noble prize-winners to boost their rankings. Another ranking uses the opinion of 'leading' scientists to judge an institute; the subjectivity of such a ranking is obviously biased and open to the vagaries of a chance visit.

But, the obsession with ranking does not stop at rating institutions. The past years have seen the ranking of whole nations, with their research output measured using any of the following: the number or impact of published papers, the number of scientists and engineers per 1,000 workers, the number of Nobel prizes awarded to those working in the country, the level of investment in research, and so on. Country ranking is now a well-established sport and becomes increasingly popular as each country tries to move up the table to be more attractive to high-tech industry. But, again, these rankings are often flawed; for example, some countries make major investments in defence-related research, which skews the measurement, or the data are not normalized for the investments made per paper or the number of papers per scientist.

Notwithstanding the imperfection of the metrics, the resulting league tables are having real effects: university presidents worldwide await with trepidation the outcome of the latest scores. They know that it is easier to attract staff to a university that is moving up the ranking tables and this, inevitably, is leading to policy changes. Research areas that contribute little to the overall ranking might be closed and the appointments of new faculty members will reflect, to some extent, their potential to contribute to the university's metric success. Perversely, universities are entering a time of greater competition when co-operation might in fact be more appropriate.

Governments also watch what is happening in the league tables, which translates into funding decisions. In this way, the power of the tables becomes amplified—although in keeping with the maxim quoted at the beginning of this article, such measurements will probably improve research in the long run because they stimulate competition. Often an external wake-up call is needed to end complacency and instigate much needed changes.

There will be consequences from the rankings, not least in the career choices made by faculty and students. But the scoring of researchers, universities and countries should only be one part of the equation when deciding where to study or where to start an independent research group. Such decisions have many components, some of which are social rather than academic—as research has social and academic consequences. In fact, governments must also act to improve their country's ranking in other leagues—whether they evaluate income per person, average life span or simply the happiness of citizens. Although these are much harder to measure, they are no less important than rankings that look only at research output and quality.

Frank Gannon

This Editorial represents the personal views of Frank Gannon and not those of Science Foundation Ireland or the European Molecular Biology Organization.

doi:10.1038/embor.2008.47