

Research funding goes metric

As UK researchers wait to find out what they will have to spend over the next five years,

Geoff Watts assesses the system for allocating funds and how it will work in future

Imagine it's handout time at the Higher Education Funding Council for England and its equivalents in Scotland, Wales, and Northern Ireland. Stacked up inside the vaults labelled "research money" sit piles of cash; outside, waiting impatiently to pocket their allotted share, a queue of vice chancellors and other university bosses. But how much will each receive? And according to what criteria?

Since 1986 the distribution of university research funds has been decided by a baroque process known as the research assessment exercise or RAE. The results of the latest,¹ due to be unveiled on 18 December, will decide who gets what during the five years from 2009. The current exercise is the sixth of these monumental and sometimes controversial feats of academic bureaucracy. It will also be the last.

The original point of rating UK academic research was to provide a measure of quality assurance. It still is, but the scheme rapidly evolved into a competition for funds—which is how most university staff now view it. A wider aim has been to promote the strength and international competitiveness of UK research by rewarding and so promoting high quality work in those institutions doing it best. Has the scheme succeeded?

Yes, according to the late Gareth Roberts, who carried out a detailed review of the last research assessment.² "All who examine the impact of the RAE upon UK research and its international reputation must, I think, agree that it has made us more focused, more self-critical and more respected across the world," he said in his report. "It has done this, in large part, by encouraging universities and colleges

to think more strategically about their research priorities." The report goes on to add that the system "has enabled funds to be concentrated in those departments best able to produce research of the highest quality. It has . . . gained the acceptance of the research community and its stakeholders."

In its report of April 2002 the Commons Science and Technology Select Committee offered a less wholehearted endorsement: "The RAE has had positive effects: it has stimulated universities into managing their research and has ensured that funds have been targeted at areas of research excellence. But it also stands accused of distorting research practice, ruining academic careers and contributing to the closure of university departments."³

Logistical challenge

The principle on which these periodic assessments are based is familiar enough: peer review. What's remarkable is the scale of the enterprise. The 2001 review, for example, saw some 70 panels of experts considering the work of almost 50 000 researchers in 2598 submissions from 173 institutions. The administrative burden of performing this heroic task is self evident and said to absorb around 3% of the funds on offer. Nor is this the only criticism of the system.

I have spoken to two senior academics who have served on panels reviewing research in medicine. Both wished to remain anonymous. One with experience of two different panels was struck by the variation between them in fairness and rigour. Both spoke of the need for more explicit assessment criteria that can be applied with greater consistency. "Some of

the panels are covering a very broad range of topics," said one of them. "It is difficult in the age of increasing specialisation for people in one area to make a reasoned judgment about what's important in another. To be honest, I think that some people have found themselves having to make judgments that lie beyond their level of competence."

The problem is particularly marked when dealing with research designed to inform policy making. Public health research, for example, is multidisciplinary, draws on findings from many different countries, and relies on a variety of methods. Getting to grips with its quality and value is a challenge for those not directly involved. A greater number of more specialised panels would ease this burden but add to the complexity and cost of the exercise.

"Applied health services research has tended to be less well regarded," according to one of the panellists. The panels' criteria have tended to focus on whether the work is original and exciting science. So finding answers to vital but unexciting questions about running a health service may end up scoring fewer "brownie points."

More generally, by creating a period of blight during the run-up to each assessment and then again while its results are awaited, the whole exercise is said to distort university planning cycles. Also, time and effort goes on trying to manipulate the system through ploys such as calculating whether it might be better to put up a small number of star performers or a larger number of staff that includes more average researchers. Recent years too have witnessed the development of a transfer market for academics, with the brightest moving





SIMON FELL/GETTY IMAGES

to more highly rated departments. Other criticisms are that the assessment undervalues interdisciplinary work, discourages risk taking, and places hurdles in the way of new universities and emerging areas of research.

Successive assessment exercises have tried to deal with some of these problems, though not always successfully. While supportive of the peer review principle, Sir Gareth's report on the 2001 assessment made a further series of recommendations.² These included a call for greater transparency, efforts to limit grade inflation, and the replacement of simple grade bands by "quality profiles" in which each submission would be rated according to the proportions of its research that fell into different grades. As Sir Gareth conceded in the preface to his report, the proposals "sacrifice simplicity for efficiency and fairness."

In the end, the Higher Education Funding Council chose to incorporate only some of the report's recommendations—quality profiles, for example. But overall both former panellists who spoke to me thought that lessons had been learnt and acted on. How well the revised system has performed we will no doubt learn in due course.

All change

In March 2006, while attempts were still being made to improve the existing assessment method, the government delivered a further and unexpected shock to the system: a radical

change in the rules. It declared that assessment would in future rely not on peer review but on statistical indicators or "metrics": potentially a simpler and cheaper method.⁴ RAE will be replaced by a different triplet: REF or research excellence framework.

Precisely which metrics the new system will use has yet to be confirmed. But the future measuring stick could be a department's total non-government research income, the quality and impact of its publications (bibliometrics)—including impact factor—the number of its postgraduates, or a combination of all these and perhaps other factors.

The prospect of abandoning peer review as the keystone of assessment provokes a mixed reaction. The Roberts report specifically backed peer review. Although some of Sir Gareth's group had initially thought that performance indicators were a satisfactory alternative, they eventually changed their minds. "Whilst we recognise that metrics may be useful in helping assessors to reach judgments on the value of research, we are now convinced that the only system which will enjoy both the confidence and the consent of the academic community is one based ultimately upon expert review. We are also convinced that only a system based ultimately upon expert judgment is sufficiently resistant to unintended behavioural consequences to prevent distorting the very nature of research activity."

Recent years have witnessed the development of a transfer market for academics, with the brightest moving to more highly rated departments

The anonymous academics agree. "Any such measure will encourage people to play games. You're dealing with very intelligent people who will find ways of manipulating the system."

The Academy of Medical Sciences seems less concerned about the use of metrics.⁵ "Although peer-review enjoys considerable support, the process is becoming increasingly burdensome and complicated." That said, it warns against reliance on a single indicator if "perverse behaviour" (that is, playing the system to your own advantage) is to be avoided.

The Medical Schools Council also seems to view metrics with equanimity, though in its consultation response to the proposed changes it does say that it would be "very concerned" if there was no peer review at all in the process.⁶ It also sounds a couple of other warnings. One concerns the increasingly interdisciplinary nature of medical research and the added difficulty of assessing work in

which medicine interacts with, for example, engineering or social science. The other is a fear that metrics will undervalue translational research. Advances in work of this kind are not necessarily published in high impact journals attracting large numbers of citations. In the current system, the council says, this is dealt with by having someone on each panel to assess the research's relevance to the NHS. It hopes that some such arrangement will continue. More recently it has learnt that the National Institute for Health Research believes that research money channelled through trusts should be included in the research excellence calculation.

One of the academics I spoke to also pointed out that if the chosen system of metrics favours basic science journals, policy related research may once again lose out. "The incentives should encourage people to do what society feels is most appropriate. Publishing papers in *Nature* is not necessarily it."

For good or ill, then, the future is metric. Bibliometric assessments have become more sophisticated over the years; and, still more encouraging, the income delivered to universities through the research assessment exercise has generally shown a good correlation with their total income from all other sources. So in principle, at any rate, there is a case to be made for a less burdensome approach to sharing out government cash.

That said, one caveat put forward by the Medical Schools Council is surely worth taking seriously. The results of any new system of metrics, it says, should be scrutinised to see how closely they correlate with those of the current research assessment exercise. Only then can researchers be expected to have confidence in the metrics by which they'll be soon be judged.

Geoff Watts freelance journalist London

geoff@scileg.freereserve.co.uk

Competing interests: None declared.

- 1 Research assessment exercise 2008. www.rae.ac.uk.
- 2 Roberts G. Joint funding bodies' review of research assessment. 2003. www.ra-review.ac.uk/reports/roberts.asp.
- 3 House of Commons. Science and Technology Select Committee. *Second report: the research assessment exercise*. 2002. www.publications.parliament.uk/pa/cm200102/cmselect/cmsctech/507/50702.htm.
- 4 HM Treasury. *Science and innovation investment framework 2004-2014: next steps*. 2006. www.hm-treasury.gov.uk/media/7/8/bud06_science_332v1.pdf.
- 5 Academy of Medical Sciences. *Response to the DfES consultation on the reform of research assessment and funding*. 2006. www.acmedsci.ac.uk/p48prid49.html#downloads.
- 6 Medical Schools Council. *Medical Schools Council response to the consultation document on the research excellence framework*. 2008. www.medschools.ac.uk/documents/MSCresponseToHEFCEconsultationREF.pdf.

Cite this as: *BMJ* 2008;337:a1805