

How should we rate research?

The UK's current system can be improved but shouldn't be discarded

See also *News* p 994

Last month the UK chancellor, Gordon Brown, announced in his annual budget speech that the government intends to simplify the way it funds academic research. Pending a consultation exercise through 2006, the government wishes to replace the United Kingdom's unique research assessment exercise (RAE).¹ The most radical proposal is to scrap peer assessment of the quality of each university's research, a cornerstone of the exercise, and to introduce assessment based mainly on metrics—effectively performance indicators.

Possible metrics include research income, publications, citations, and numbers of research students, all of which correlated well with scores achieved in previous exercises.¹ Preparations for RAE2008 are well under-way and will proceed as planned, but the exercise will now incorporate a shadow system, using metrics alongside the panel based peer review system. It is now entirely possible that the allocation of research funding after 2008 will be based on metrics assessment rather than the historical peer review exercise.¹

Coordinated by the UK's higher education funding councils, RAE2008 will be the sixth national peer evaluation of the quality of research conducted by higher education institutions.² The exercise assesses the outputs of research by academic staff in higher education institutions. The rating for each institution is converted to a multiplication factor—which, alongside factors controlling the volume of research (such as staff numbers, numbers of research students, and peer reviewed grant expenditure), determines how much quality adjusted research funding the government gives to that institution. Along with money allocated for teaching, this funding represents the education councils' block grant to UK higher education institutions; for 2007-8 the higher education funding council will allocate £1.45bn (€2.1bn; \$2.6bn) to institutions as quality adjusted research funding.

This research quality assessment exercise is the largest attempted anywhere in the world. Most countries continue to base their funding of public research on negotiation with academic institutions (Austria, France) or on simple formulaic models factoring in numbers of staff and students (United States, Canada, most of Europe). Only Hong Kong replicates the intensity of the research assessment exercise, though some countries such as New Zealand intend to adopt similar evaluations. In Australia, quality evaluation based on performance indicators is probably the closest system to the new proposed metrics analyses.³

Evidence suggests that the research assessment exercise is having a beneficial impact on the quality and competitiveness of UK research. It has generated a cycle of quality rating: institutions rated highly receive more money and are able to do more research of high quality. This has concentrated investment in research in the UK; the proportion of total public funding to the top tenth of researchers in the UK increased from 47% in 1980-1 to 57% in 1997-8, compared with a decline in the US from 47% to 43%.⁴ The exercise has increased the quality of research perhaps through focusing the research strategies of academic institutions and providing incentives.⁴ In RAE1996, 32% of staff worked in units rated as excellent, and this rose to 55% in RAE2001. Overall, UK research tops the world league for papers per dollar expended and citations per dollar expended, and it is fourth ranked for papers per researcher.⁵ Maintaining research investment will remain essential if the UK is to retain its international profile. Delivering 9% of the world's research effort for 4.5% of the world's research expenditure in a high cost economy seems precarious, but the research assessment exercise has provided an evidence base to justify such an investment.

The results of the exercise are also used for many secondary purposes. For example, each institution's rating may influence the chance of receiving other research funding or of fellowships or studentships being awarded. Most research sponsors now request applicants to state their institution's score, and eligibility for some initiatives is limited to researchers from institutions with the highest ratings from the most recent assessment (in 2001).

RAE2008 represents the next cycle in the process but with several notable differences from earlier exercises. Staff engaged in research who are in post in the institution on the 31 October 2007 census date are invited to submit up to four research outputs published between the dates of 1 January 2001 and 31 December 2007; most are expected to be publications in peer reviewed journals, including systematic and Cochrane reviews.²

Metric data is certainly required—each institution must provide, for the period from 1 January 2001 to 31 July 2007, data on research studentships and associates and expenditure through research grants. In contrast to earlier exercises, however, and somewhat at odds with the "metric" lobby, a far greater emphasis is placed on peer review of research outputs in generating an overall score; 75% of the total score in any given assessment will be derived from publication outputs,² with the scoring

system placing more emphasis on research of international quality. Instead of the RAE2001 rating of 0 to 5*, RAE2008 will use an overall quality profile assessment of “unclassified” (below national standard), “one star” (nationally recognised research), “two star” (internationally recognised research), “three star” (internationally excellent research), and “four star” (world-leading research). In the absence of any published criteria, peer reviewers might find it difficult, however, to rate international research using the subjective concepts “recognised,” “excellent,” and “world-leading.”

Additional assessment panels have been created to provide more emphasis on important research themes such as cardiovascular disease, cancer, and neurosciences while also recognising the methodological diversity of the more applied research areas of clinical epidemiology, health services research, and primary care. These latter areas were rated below laboratory based and hospital based research in previous research assessment exercises but are particularly important to the NHS, for instance in guiding policy, identifying cost effective ways of configuring care, and providing high quality syntheses of research evidence.

In an era of perceived crisis in training academic clinicians,^{6,7} RAE2008 will be both novel and innovative in encouraging institutions to submit younger researchers (clinical lecturers, for example) and research collaborators funded in adjacent NHS trusts; in both cases only two research outputs will be required. Accompanying text will allow institutions to outline their research strategy and vibrancy of research training programmes.²

The parallel, metric based analysis will fail to take into account many of these important additions to RAE2008. The cost and the workload of the exercise will undoubtedly be further drivers down a metrics route. For RAE2001 the direct costs were over £5m, an increase of 68% above costs in RAE1996.⁸ The amount of academic time spent was illustrated by 2598 submissions from 173 institutions, most listing four papers from each of 48 022 researchers.⁵ However, despite this direct expenditure, with estimates for institutional and opportunity costs inflating the 2001 figure to over £37m, the overall cost of the research assessment exercise represented just 0.8% of the total funding allocated on the basis of the exercise.⁹ Money well spent, or

an unnecessary burden when faced with a metrics alternative? The government has announced its intent to instigate change down this route; already, related metric based approaches are being used to guide allocation of major NHS research and development funding, for example.¹⁰

Whatever the pitfalls, our entire academic base is based on the philosophy underpinning peer review and it will be of paramount importance in RAE2008 to compare the peer review and metric based systems before recommending major changes. Equally, RAE2008 should inform the international community on the optimal approach to assessing research quality.

F D Richard Hobbs *professor*

(f.d.r.hobbs@bham.ac.uk)

Department of Primary Care and General Practice, University of Birmingham, Birmingham B15 2TT

Paul M Stewart *professor*

Department of Medicine, University of Birmingham

Competing interests: FDRH was a member of HEFCE JMAC (1998-2004), the body that coordinates the development and delivery of the RAE criteria for medicine subjects. PMS chairs the University of Birmingham Medical School RAE2008 committee and is a subpanel member (panel A4) for RAE2008.

- 1 HM Treasury, Department of Trade and Industry, Department for Education and Skills, Department of Health. *Science and innovation investment framework 2004-2014: next steps*. 2006. www.hm-treasury.gov.uk/media/1E1/5E/bud06_science_332.pdf (accessed 18 Apr 2006).
- 2 RAE 2008. *Panel criteria and working methods*. January 2006. www.rae.ac.uk/pubs/2006/01/ (accessed 6 Apr 2006).
- 3 Von Tunzelmann N, Mbu EK. Changes in research assessment practices in other countries. 2003. www.ra-review.ac.uk/reports/Prac/ChangingPractices.doc (accessed 19 Apr 2006).
- 4 RA Review. *Review of research assessment. Report by Sir Gareth Roberts to the UK funding bodies*. May 2003. www.ra-review.ac.uk/reports/roberts.asp (accessed 6 Apr 2006).
- 5 Higher Education Funding Council for England. *Review of research: international comparisons of research spending and performance*. 2000. www.hefce.ac.uk/pubs/hefce/2000/00_37d.do (accessed 19 Apr 2006).
- 6 Stewart PM. Academic medicine: a faltering engine. *BMJ* 2002;324:A37-8.
- 7 Academy of Medical Sciences. *Clinical academic medicine in jeopardy, recommendations for change*. (Savill report.) June 2002. www.acmedsci.ac.uk/images/publication/pclinaca.pdf (accessed 6 Apr 2006).
- 8 HERO—Higher Education and Research Opportunities in the United Kingdom. *Research assessment exercise 2001. Manager's report—May 2002*. www.hero.ac.uk/rae/Pubs/other/2001_managers_reportfinal.doc (accessed 18 Apr 2006).
- 9 Higher Education Funding Council for England. *Better accountability for higher education: a review for the HEFCE by PA Consulting*. 2000. www.hefce.ac.uk/Pubs/HEFCE/2000/00_36.htm (accessed 6 Apr 2006).
- 10 Van Leeuwen T, Grant J. *Bibliometric analysis of highly cited publications of health research in England, 1995-2004*. 2006. (Background paper for NIHR Biomedical Research Centre applications.) www.dh.gov.uk/assetRoot/04/13/31/94/04133194.pdf (accessed 18 Apr 2006).

bmj.com

Tell us how the UK should rate research via bmj.com

Creative destruction in the NHS

The prime minister gambles that more pain will result in greater gain

When the prime minister makes two major interventions in the NHS in the space of a week, it is a sure sign that the problems of the NHS have reached the top of the government's agenda. In the second of these interventions, a speech to the New Health Network Clinician Forum, Tony Blair rehearsed the progress made in improving the performance of the NHS and argued that a “crunch point” had been reached in its reform.

The crunch point has arisen, in his words, as part of the transition from the application of “necessarily crude and blunt instruments of centralised perform-

ance management and targets” to the use of market-like incentives “to make reform self sustaining.”¹ In recognition of the challenges involved in making this transition, the prime minister's message was that he and his government intend to hold their nerve and see the process of change through, even if this means bearing the short term political costs.

To many NHS staff struggling to deal with financial deficits and to implement a long list of government priorities, Tony Blair's determination to press ahead with reform must seem difficult to comprehend. Faced with target overload and affected by reform fatigue, these

BMJ 2006;332:984-5