

## Joint funding bodies' review of research assessment: Submission from the UK Life Sciences Committee

The UK Life Sciences Committee (UKLSC) comprises 18 leading learned societies (see appendix) and represents some 35,000 cell, molecular, and physiological life scientists. Most of those represented work in academic institutions. This submission was compiled from the responses of individual member societies to the questions in the funding councils' consultation paper, and is set out in the same format. The submission has been endorsed by all UKLSC societies.

### Summary

1. Expert review must remain the primary method of assessing the quality of research. It should be facilitated, but not replaced, by the kinds of quantitative metrics used in the current RAE.
2. There are no fundamental flaws in the RAE process, although its credibility has been seriously undermined by the inability of funding councils to reward improved performance, in many cases, following RAE 2001.
3. Institutions should have to account for all research-active staff. New procedures need to be developed to take better account of young staff as yet without a publication record, and staff with major commitments outside research. Other issues outstanding from RAE 2001, and requiring attention in any future expert review process, are listed.
4. The funding councils are encouraged to consult the research community further on the concept of alternating full expert review with a 'lighter touch' process of self-assessment.
5. All groupings applying for formula research funding should be assessed by the same process. But weaker, or developing, groups should be encouraged not to apply for expert review and to bid for funding from a small pot of money ring-fenced for developing research capacity.

### Approaches to assessment

1. *We can envisage four distinct approaches to assessment:*

- *expert review (including peer review)*
- *algorithm based entirely upon quantitative metrics*
- *self-assessment*
- *historical ratings.*

The clear majority view within UKLSC is that the primary mechanism of assessment must involve expert review, which can be supported and facilitated by quantitative metrics. Such metrics often require expert judgement in assessing their relevance and value. Self-assessment and historical ratings are unsuitable primary review mechanisms. While there will always be some criticisms of expert review, it is difficult to envisage any other fair and effective measure of research quality and quantity. The RAE process is not considered to be *fundamentally* flawed, but disillusionment arose after RAE 2001 because improved performance was in many cases not rewarded. If the RAE had occurred at a different stage of the government's Spending Review cycle satisfaction with the funding outcome may well have been different.

Any future process of expert review must address the following issues from RAE 2001:

- The perceived lack of parity between panels, despite the fact that this was recognised in advance as a potential problem and measures introduced in RAE 2001 to try to minimise it. Discrepancies between the outcomes of some panels covering biomedical subjects were reported to the funding councils and discussed in the pages of the Times Higher Education Supplement.
- The award of high grades to some departments that returned only a small percentage of staff, which may give a distorted impression of the overall strength of research in those departments. All research-active staff must be accounted for, and the assessment process needs to take better account of individual staff with other major commitments such as teaching, administration or clinical duties.
- The question of applied research. This is part of a wider issue of how universities and their staff should be rewarded for excellence in the 4 key activities of teaching, research, knowledge transfer to the community, and widening access. To what extent should applied research be judged and rewarded as part of knowledge transfer? And under what heading should the important activities for research scientists of communicating science and engaging with the public be judged?

- The historic anomalies in funding between different units of activity in the biomedical sciences need to be put right. One senior member of UKLSC estimated that if that person's School, which gained 5\* gradings in Biological Sciences and Pre-clinical units of activity, had submitted all researchers to Subjects allied to Medicine and achieved the same grades, the School would have received an additional £1.9 million per annum. The funding mechanism also means that it is more financially rewarding to submit in an overall weak rather than a strong unit of activity.
- The need to assess cross-disciplinary work more successfully. This might be aided by having a smaller number of units of activity.
- The fact that those making the submissions did not have a clear appreciation of what would be the consequences of alternative courses of action. Every effort must be made to ensure that the rules are as transparent as possible.

While stressing that expert review must remain the principal method of assessing research, UKLSC recommends that the Funding Councils should explore further with the research community the possibility of alternate rounds of assessment being a 'lighter touch' system based on self-assessment.

### **Group 1: Expert review**

2. *A variant of this system would be a combined assessment of teaching and research.*

It is difficult to envisage a system that could assess both teaching and research, but note the point made in paragraph 1 that the research assessment process needs to take account of the teaching commitment. Excellence in teaching has to be recognised and rewarded adequately if the current distortions in teaching and research are to be reduced.

3. *Suppose the funding councils have decided that they wish to retain the judgement of experts as the cornerstone of the research assessment. They are, however, willing to consider any system, however different from the 2001 RAE, so long as that condition is met. How would you advise them?*

4. *In providing your advice, you are asked to consider the following questions:*

a. *Should the assessments be prospective, retrospective or a combination of the two?*

A combination of rolling retrospective assessment and future prospective activity. Performance should be judged against forward plans submitted in the previous assessment. Fair consideration must be given to young staff with great potential, but who have not yet established international profiles.

b. *What objective data should assessors consider?*

The data most frequently stated by UKLSC respondents are: research grant income and commercial income, publication metrics as indicators of quality research output, patents, numbers of post-graduate students and indicators of successful staff training outcome, successful internal and external collaborations, facilities available, and plans for sustainability. Contributions to the broader aims of the department, by for example giving international invited lectures, receiving major prizes, or serving on research committees, should also be considered for individuals.

c. *At what level should assessments be made – individuals, groups, departments, research institutes, or higher education institutions?*

At the level of thematic groups determined by the institution.

d. *Is there an alternative to organising the assessment around subjects or thematic areas? If this is unavoidable, roughly how many should there be?*

No, thematic groups are appropriate. The number and type are difficult to judge and a sensible approach would be to consult the research community separately on whether the units of assessment used in RAE 2001 remain appropriate.

e. *What are the major strengths and weaknesses of this approach?*

These have been discussed extensively following the last RAE (eg in the report of the Commons Science and Technology Committee Inquiry). Strengths include thoroughness, fairness, relative transparency, flexibility and ability to take into account complex issues such as personal circumstances of individuals. Weaknesses include the expense for departments in time and resources in preparing submissions, some lack of parity in practices and outcomes between panels, and some concerns about how well interdisciplinary research is dealt with. The RAE has undoubtedly led to games playing that has demoralised some staff and hampered teaching, but

this is more to do with the fact that research is so important for bringing in external grant income than the mechanism by which research is assessed.

### **Group 2: Algorithm**

5. *Suppose the funding councils have decided to use an algorithm to assess research quality. The assessment must be 'automatic', leaving no room for subjective assessment. Metrics might include:*

- *measures of reputation based on surveys*
- *external research income*
- *bibliometric measures (publications or citations)*
- *research student numbers (or completions)*
- *measures of financial sustainability.*

All these are important contributory factors to an expert-review process but not by themselves sufficient to give an accurate assessment. Combining them would not make the approach more acceptable. In previous consultations external research income and bibliometric measures have been stated most frequently by biosciences researchers as possible metrics. Each has split the research community down the middle, for the reasons explained in UKLSC submissions made to HEFCE in advance of RAE 2001.

6. *Assume the councils have not, however, formed a view on what metrics should be used or how they could be combined most effectively in an algorithm. How would you advise them?*

7. *You have been asked in providing your advice to consider the following questions:*

a. *Is it, in principle, acceptable to assess research entirely on the basis of metrics? No*

b. *What metrics are available?*

See paragraph 4b

c. *Can the available metrics be combined to provide an accurate picture of the location of research strength? No*

d. *If funding were tied to the available metrics, what effects would this have upon behaviour? Would the metrics themselves continue to be reliable?*

In general terms groups would organise their activities in such a way as to maximise their ability to meet metric requirements. For example, a metric based on grant income would discourage excellent groups from working in unfashionable areas, and could discourage speculative research and risk-taking while encouraging short-term goals. If bibliometric measures exerted great influence research groups would maximise publications (salami-slicing), or number of citations (self-citation, cross-citation deals) or their impact factor (leading to the possible demise of some journals published by British Learned Societies). Impact factor of published work would dominate the university promotion system even more than it does at present, lead to the buying in of 'superstars' to the detriment of developing young researchers, and possibly lead to the further divorcing of teaching and research.

e. *What are the major strengths and weaknesses of this approach?*

The approach is less bureaucratic and can be considered objective. But used alone the various metrics have no context, the approach is inflexible, not equally fair to all thematic groupings and disciplines, and not considered reliable.

### **Group 3: Self-assessment**

8. *Suppose the funding councils have decided to pursue a self-assessment model in which institutions, departments or individuals assess themselves. A proportion of the assessments are reviewed in detail. In a self-assessment model, the assessment is made by the assessed, although its reliability may be challenged by the validators.*

9. *Assume the councils have not, however, formed a view on how the assessment should be structured and how self-assessments will be validated. How would you advise them?*

10. *In providing your advice, you are asked to consider the following questions:*

a. *What data might we require institutions to include in their self-assessments?*

Essentially those listed in paragraph 4b. There should be an internal assessment highlighting strengths and potential weaknesses of the particular thematic group, details of past and future spending, past and future research, and a future strategy.

b. *Should the assessments be prospective, retrospective or a combination of the two?*

A combination. A key requirement should be that performance is assessed against the forward plan submitted in the previous assessment.

c. *What criteria should institutions be obliged to apply to their own work. Should these be the same in each institution or each subject?*

The criteria should be similar to those used by the panel reviewers in the current RAE

d. *How might we credibly validate institutions' own assessment of their own work?*

Sampling and external peer review.

e. *Would self-assessment be more or less burdensome than expert review?*

It depends on the degree of rigour required. If self-assessment alternated with a fully expert-reviewed assessment the research community and funding bodies may be able to agree slimmed-down criteria, depending on the frequency of assessment. To comply with accountability requirements the funding councils would have to be able to insist that any groupings in which there are indications that research quality has declined are subject to full expert review.

f. *What are the major strengths and weaknesses of this approach*

It may be less expensive, but is also less detailed, less objective and less accurate than expert review and cannot completely replace it.

#### **Group 4: Historical ratings**

11. *Suppose the funding councils have decided to pursue a policy that gives each institution a rating on the basis of its historical performance and/or the value of its research infrastructure. Research would, in effect, be presumed to be strongest in those departments or institutions with the strongest track record.*

12. *The councils recognise that such an approach could only be used in conjunction with another system: there would need to be some way of identifying institutions whose performance was sharply improving or declining, even if the presumption was that the distribution of excellence would remain stable. It would also be possible to alter the share of the total pot provided for each institution on the basis of what had been achieved with the investment provided (a 'value for money' rating).*

13. *Assume you have been asked to advise on how such a system might work. In developing your advice, you have been asked to consider the following questions:*

a. *Is it acceptable to employ a system that effectively acknowledges that the distribution of research strength is likely to change very slowly?*

No. It would risk making the whole process slower and less responsive and make it more difficult for new talented research groupings to develop.

b. *What measures should be used to establish each institution's baseline ratings?*

c. *What mechanism might be used to identify failing institutions or institutions outperforming expectations? Could it involve a 'value for money' element?*

Value for money sounds attractive, but the concept could be tricky. Some work is more productive in terms of publications because of its nature, but it is not necessarily better or more valuable. Equally, some types of research are more expensive than others.

d. *What would be the likely effects upon behaviour?*

It would not encourage weaker groups and institutions to develop.

e. *What are the major strengths and weaknesses of this approach?*

The strength is that historically strong groups or institutions would be able to develop their strengths relatively unhindered. But it makes the assumption that groupings change slowly in either a positive or negative manner. The approach could delay identification of problems and putting in place measures for improvement.

#### **Group 5: Crosscutting themes**

14. *The funding councils are seeking views on the following fundamental issues regardless of the approach taken to assessment.*

- *What should an assessment of the research base be used for, other than to provide information for funding decisions?*

Management information from research assessment is bound to be used externally. It would be advantageous for the funding councils to produce and share reliable information of this kind with research councils, other funding bodies, and academic communities. But the primary aim of research assessment should be to work with institutions and departments to help them improve their research with as little top-down political intervention as possible.

- *How often should research be assessed? Should all subjects and all institutions be assessed at the same time and with the same frequency?*

Every 5-7 years. The upper figure may be preferable since it is less of a deterrent to longer-term research, and with a 5-year cycle one review is not long completed before staff have to think about the next. This detracts from the core activities of teaching and research. If expert review and self-assessment were to alternate the frequency of expert review could perhaps be increased to 8-10 years, with a smaller self-assessment after 4-5 years. With this length of time between rounds of expert review it would be important to have a mechanism to allow a department that considered it had improved since the previous assessment to request expert review at the intermediate stage.

- *What is understood by excellence in research? Is it simply intellectual excellence, or does it have other dimensions reflecting its likely impact within and beyond the research community. Are creativity and applicability measured appropriately in the existing RAE? Does excellence include a measure of the contribution to the training and development of researchers?*

Research should contribute very positively to our overall knowledge base and be useful in the sense that it either provides new insights on which future research can be based, or can be used commercially or to improve the health and well-being of the population, for instance. The RAE panels of which we are aware took pains to ensure that creativity and applicability were taken into appropriate consideration. Excellence should certainly include a measure of the contribution towards training and developing researchers.

- *Should research assessment determine the proportion of the available funding directed towards each subject? There are several ways in which 'subject pots' might be determined;*
- *Quality of UK research in the subject benchmarked against international competition*
- *Volume of research in the subject that meets a given quality threshold*
- *A strategic judgement on the importance of the area to the UK*
- *A metric based on external funding in the subject*
- *A historical distribution that aims to retain the current balance*
- *A mixture of the above*

*If the relative quality of research in different subjects is to be used as the basis for generating subject pots how is this to be assessed?*

UKLSC has no firm views on this, although the biomedical sciences would do well from such a funding distribution. There is no intrinsic reason why a particular subject in which UK research happens to be strong should receive a larger proportion of the available funding. The research councils already base some of their funding decisions on strategic factors. On the other hand the real current costs of research in different subjects need to be re-evaluated (see paragraph 1).

- *Should each institution be assessed in the same way? The 2001 RAE required all institutions, research-strong or research-weak, to submit to the same assessment. Would it be sensible to encourage weaker institutions not to enter the formal assessment process? The methodology would need to incorporate a ladder of improvement so that all researchers and institutions have the opportunity to demonstrate potential.*

All institutions applying for formula research funding should be assessed in the same way, but UKLSC considers that it would be sensible to encourage weaker, or developing groups, not to apply. As suggested by the Commons Science and Technology Committee, a separate small pot of money could be set aside for developing research capacity, to which groups not taking part in the formal expert review could bid for funding. These groups would be able to submit to expert review when they considered their research was sufficiently strong, and probably should be required to within an agreed time frame in order to show that they are making progress. The funding councils would need to consult with the research community to agree the cut-off grade for receiving funding in the present system in order that there is a guiding benchmark.

- *Should each subject or group of cognate subjects be assessed in the same way? Should each subject community be free to define the sort of assessment most appropriate for it? Should the funding councils go further in standardising assessment practice or is the current balance about right?*

Groups of cognate subjects should be assessed in the same way. The subject communities should be involved in defining assessment criteria (as they were in RAE 2001) and panels covering cognate subjects should agree in advance common methods of working. Following RAE 2001 there were considered to be discrepancies between the grades awarded by some panels for work of a similar standard. This is ascribed to:

- Cognate panels not always sharing common methods of working;
- Umbrella panels apparently not functioning effectively to ensure uniformity of outcome.

These issues need to be addressed in a future expert review process

- *How much discretion should institutions have in putting together their submissions? Institutions presently have a large degree of control over who or what is assessed and by whom. This allows institutions to trade funding for the prestige of a high rating, and is perceived to be unfair and detrimental to the careers of individuals whose work is not included. Two alternatives could provide more objective results:*
- *A more rigid system, but this risks bureaucratic inflexibility*
- *A system in which submissions are made and controlled by individuals, research groups or networks rather than by institutions. The latter could be considered unfair to institutions, whose funding would be determined by a process into which they had little input.*
- *What are the essential features of a research assessment process that encourages genuine equality of opportunity for all?*

Institutions should have discretion in interpreting a defined, agreed, and published set of criteria for assessment. As noted in paragraph 1, it should be a requirement that all research-active staff are accounted for. There needs to be further consultation on the treatment of:

- Staff working at an institution for part of the assessment period but who have left or retired;
- Staff newly arrived at an institution. Should they be able to include for the new institution only work done at that place? How should potential value to the research effort of the new institution be recognised?
- Young researchers who have not yet established a strong publication record.

In order to reduce the temptation to play games the assessment process needs to be modified in order to take better account of individual staff with other major commitments that might include teaching, administration, journal editing, book writing, service to a Learned Society or Committee or contribution to science communication. A quotient needs to be developed to delineate "fraction of time spent on research"

- *What are the most important features of an assessment process?*

*(please identify what you think are the most important characteristics of an assessment process from the list below)*

- Not a great burden
- Fair to individuals and institutions
- Informative
- Transparent
- *Another factor*
  - The research community should be consulted on the funding outcome as well as on the assessment process.

## **Appendix Members of UKLSC**

The Physiological Society  
British Biophysical Society  
Society for Endocrinology  
Biochemical Society  
Anatomical Society  
Genetics Society  
Nutrition Society  
British Toxicology Society  
Society for General Microbiology

British Society for Immunology  
British Society for Cell Biology  
British Pharmacological Society  
British Electrophoresis Society  
Society for Experimental Biology  
British Society for Developmental Biology  
British Society for Matrix Biology  
British Association for Psychopharmacology  
British Neuroscience Association