COUNCIL FOR THE MATHEMATICAL SCIENCES

The Institute of Mathematics and its Applications | The London Mathematical Society | The Royal Statistical Society

Secretariat: Martin Smith, De Morgan House, 57-58 Russell Square, London WC1B 4HS
email**: cms@lms.ac.uk** tel: 020 7927 0803

J Cutshall
1E
Department for Education and Skills
Sanctuary Buildings
Great Smith Street
London
SW1P 3BT

11 October 2006

Dear Mr Cutshall

Reform of Higher Education Research Assessment and Funding

The Council for the Mathematical Sciences, comprising the Institute of Mathematics and its Applications, the London Mathematical Society and the Royal Statistical Society, is pleased to enclose its response to *Reform of Higher Education Research Assessment and Funding.* This response has been prepared by a working group consisting of representatives from the three societies as part of the Council's ongoing role of representing the interests of the mathematical sciences to government and other public bodies.

The Council believes that some reform of the present system is necessary. The current Research Assessment Exercise process is over-burdensome and the five or seven year gaps between assessments are hugely disruptive due to (a) flurries of intensive and burdensome work; (b) perverse behaviour by institutions to maximise their advantage in the exercise; (c) major discontinuities in funding following RAE judgements.

Nevertheless, the Research Assessment Exercises have resulted in judgements which have been generally accepted to be fair and accurate. Any replacement for the RAE must preserve the reliability and validity of its judgements. We do not believe that any of the methods proposed in the consultation document meet this requirement.

We provide evidence to show that the metrics proposed produce almost random results compared to RAE 2001, not only in Mathematics but in a broad range of subjects. To be acceptable, a metric must

- be a proxy for good research outputs;
- encourage behaviour which is supportive of the research function;
- be broad based in nature, seeking to measure the desired qualities in a number of different ways.

In the current state of development, we do not accept that valid and reliable judgements can be made on the basis of metrics alone. Any new system needs, therefore, to be based on peer review of research outputs, underpinned by reliable data and evidence. Such a review procedure should be performed on a regular basis with outcomes suitably smoothed over time.

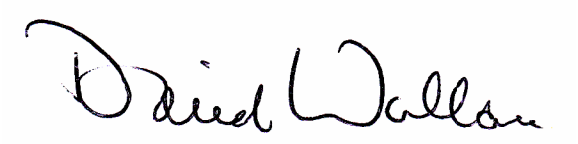
We are concerned that an approach that simply produces an assessment of 'quality' can lead to unintended outcomes, if the subsequent funding decisions are ill-considered. Even though the RAE

2001 judgements were valid, the Mathematics base in the UK was damaged by the subsequent funding decisions, particularly in some regions, resulting in "mathematical deserts".

Finally, we are concerned at the false distinction in the consultation document between STEM and non-STEM subjects. We believe that subjects lie on a continuum. Certainly, *Science and Innovation Investment Framework: Next Steps* recognised that Mathematics was different from other STEM subjects.

The Council stands ready to assist the Department in developing a replacement for the current RAE which is fit-for-purpose but which minimises the unsatisfactory features of the present system.

Yours sincerely,



Professor Sir David Wallace CBE FRS FREng

Chair, Council for the Mathematical Sciences

Question 1
Which, if any, of the RAE 2008 panels might adopt a greater or wholly metrics-based approach?

Summary

- The distinction between STEM and non-STEM subjects is unhelpful.
- Mathematics is atypical of STEM subjects in that it is people- rather than facilities-intense in the use of resources, and hence capable of being carried out in small units.
- In Mathematics, a QR measure based on inputs would effectively abolish QR funding as a distinctive entity.
- A metric based on research grant income alone is not a robust metric of Mathematics research quality.
- None of the proposed metric based models are fit for purpose for Mathematics and the relationship of results from the models with RAE 2001 outcomes is exceedingly weak.
- Any mechanism must recognise the *quality* of research, not simply *volume.*

Discussion

Our response here focuses on the Mathematical Sciences, specifically Applied Mathematics, Operational Research, Pure Mathematics and Statistics (currently represented by RAE sub-panels 20-22).

The consultation document makes the distinction between STEM (which includes Mathematics) and non-STEM subject areas. We feel that this distinction is unhelpful, particularly with respect to Mathematics. Mathematics is understandably bracketed with Science, Technology and Engineering because of the importance of Mathematics to those disciplines, and the strength of cross-disciplinary research involving Mathematical Sciences and many areas of Science, Technology and Engineering. However, research activity in the Mathematical Sciences has a different character from much of the research activity in other STEM areas. In particular, Mathematical research typically has relatively modest demands for equipment, facilities and consumables. The most valuable resource for mathematical research is the time of the researchers, and the majority of grant funding in Mathematics is used to resource this. As a result, high quality mathematical research is possible in quite small units, and by individuals and groups with little external grant funding and even without post-docs or students. The latter is made possible by the current dual support system, where QR funding supports research on a continuing basis, complementing fixed-term grants. A wholly metrics-based approach could easily result in the same criteria being applied to both types of funding, because income from research grants would essentially determine QR funds. The effect would be largely equivalent to abolishing QR funding.

The presence of support from industry or commerce in terms either of finance or endorsement can be considered positive. However, it must be noted that the absence of such support must not be taken as a negative. Whilst some applications emerge surprisingly rapidly, much mathematics research of great vitality is such a long way upstream of the eventual application that it is very rare that industry/ commerce can anticipate the potential benefit, or plan for the eventual roll out. The normal timescales of “planning for financial return” in industry/ commerce are much shorter than the timescale for the migration of benefits of the research to the workplace.

It follows that metrics which are based on research grant income alone are unlikely to provide a reliable or robust measure of research quality in the mathematical sciences. This was explicitly recognised in the paper *Science and Innovation Investment Framework: Next Steps*

(March 2006). That document states: “It is therefore not clear that a metric based on research income would fairly support excellent research in the arts and humanities, and some other subjects, such as Mathematics” (4.15). We are concerned that the current consultation document has apparently overlooked this point, by including Mathematics unconditionally with the other STEM subjects.

As evidence of the serious effect which a metric based approach would have on the QR funding in a range of STEM subjects, we present Annex 1. This plots the QR income per unit volume in the UoAs concerned as determined by each of the proposed models, against that determined by RAE 2001. The plots show the poor performance of all five models as predictors of actual QR allocations at UoA level and underline the fact that many of the discrepancies between modelled and actual allocations are very large indeed.

Whatever mechanism is adopted, it must be capable of recognising the quality of research, not simply its volume. Research on difficult problems may not lead to frequent publication, and often it will not attract research grants, but it must be assured of continuing support. In Mathematics, such work often provides the drive that leads to developmental research on a larger scale, and in due course to the widespread application of Mathematics in science and industry. This process requires support for research at all levels and stages. The current boom in the information technology sector is a case in point.

Hence, we strongly argue against any of the metrics-based models in the consultation document as a possible method for allocating funding in the Mathematical Sciences. It is possible that such an approach “might” be appropriate for certain disciplines, or that one based on a broader range of measures could be applied more generally. However, the creation of these metrics must be open to peer review, particularly by skilled academic statisticians, and should not be used to allocate funding for any discipline for which the quality of fit (between the RAE and the outcome based on metrics) falls below some threshold.

However they are based, metrics cannot and should not be used to make decisions about how to apportion funding across more than one discipline. This issue is particularly acute in the context of Full Economic Costing; the distribution of QR funding between ‘project-rich’ and ‘project-poor’ disciplines needs to be radically reviewed. In particular, different disciplines have access to vastly differing sources of external research income. This availability of funding reflects the priorities of funding bodies and is not related to quality. Furthermore, the total funding allocated to disciplines must, as at present, allow for strategic decisions to be made about national priorities in research.

This immediately invalidates models A and E. Similar arguments can be made against other possible metrics. For example, bibliometric measures are strongly influenced by differing publishing practices in different disciplines and cannot therefore be used to assess quality across more than one field.

Question 2

Have we identified all the important metrics? Bearing in mind the need to avoid increasing the overall burden of data collection on institutions, are there other indicators that we should consider?

We believe that the metrics identified are totally inadequate for determining research quality in the Mathematical Sciences. To be acceptable, a metric must

- be a proxy for good research outputs;
- encourage behaviour which is supportive of the research function;
- be broad based and plural in nature.

The best system will therefore consist of expert judgement, underpinned by reliable and valid data. Any system of the assessment of research quality must take research output into account. Similarly, the strength of PhD Training provided should remain a component of any assessment.

In contrast to these requirements, most of the metrics discussed in Section 4 of the consultation document are measures of input or volume. In many disciplines they will correlate with research output simply because a larger input will inevitably produce a greater quantity of output. In Mathematics, however, there is a poor relationship between the likely outcomes and the 2001 benchmark, as discussed in (1). Any metric must include direct assessment of output which must involve evaluation by competent persons in the field.

Appendix 2 of the consultation document lists possible output metrics. The Mathematical Sciences community has significant reservations about assessment based on bibliometric measures and it is of major concern to us that the consultation document seems to accept the reliability of bibliometric data without qualification. Studies from many countries have shown that bibliometric indicators are based on questionable assumptions, and that reliance on these indicators leads to perverse outcomes. These points are elaborated in Annex 2 to this response.

Question 3
Which of the alternative models described in this chapter do you consider to be the most suitable for STEM subjects? Are there alternative models or refinements of these models that you would want to propose?

We believe that the five alternative models described are totally inadequate for allocating research funding in Mathematical Sciences. As stated above, we believe that no model based on metrics alone is suitable for determining research funding in Mathematics. We believe that a flexible mechanism, using both data and professional expertise, is the only appropriate model.

We would not propose any model "for any discipline" without first having access to an authoritative statistical analysis which convinced us that the model was robust and produced results in close agreement with a recognised benchmark (preferably based on peer review) "at the level of institutions within disciplines".

It must be stressed that the judgement of quality is not an abstract entity; rather it is a practical reality that will determine the future health of the unit. The outcome of any judgement should be clearly linked to the funding that a unit will receive. The potential damage to a good department and the consequent creation of 'mathematical deserts' in some regions of the UK must be taken into account.

Question 4
What, in your view, would be an appropriate and workable basis for assessing and funding research in non-STEM subjects?

We feel that the STEM/non-STEM distinction is unhelpful. As stated in (1) and (3), we believe that no metric-based model should be used to allocate funding for any discipline (STEM or non-STEM) for which the quality of fit (between the RAE 2001 and the outcome based on the model) is inadequate. Our expectation is that this is likely to include the Mathematical Sciences.

We believe that an appropriate and workable basis for assessing and funding research “in all subjects” could be based primarily on available data already collected by institutions for other purposes (e.g. publications, funding, research students) but with peer review retained.

It is the current five or seven year gaps between RAEs that is hugely disruptive due to (a) flurries of intensive and burdensome work; (b) perverse behaviour by institutions to maximise their advantage in the exercise; (c) major discontinuities in funding following RAE judgements. Any replacement review procedure should therefore be performed on a regular basis with outcomes suitably smoothed over time to mitigate sudden funding discontinuities possible in the current system. However, their frequency would need to be balanced against the workload imposed on the research community. The peer review panels could be informed by carefully constructed metrics, but they would be able to moderate the resulting outputs. Although this would not result in any significant reduction in peer review activity over the current RAE system, our expectation is that the burden on researchers in institutions would be massively reduced with a huge cost saving.

Question 5
What are the possible undesirable behavioural consequences of the different models and how might the effects be mitigated?

Whatever model is adopted, individuals and institutions will adapt their behaviour with potentially undesirable consequences. This is true of the current RAE which distorts the research employment market cyclically within the seven year envelope.

Any model in which data is not moderated by informed judgement will result in the cynical pursuit of targets for their own sake. For example, the aim of producing good research becomes subordinate to the aim of publishing a certain number of papers. Researchers need to be aware of the need to justify the support they receive from public funds, but that does not imply that they should be judged on a ‘pages per annum’ basis.

The effects of game-playing can be mitigated by (a) decreasing reliance on numerical indicators and (b) allowing the panel to focus its expertise on the overall situation, rather than the details.

It follows that the greatest danger will arise from a system that is too narrowly based and provides rewards for only a small part of the spectrum of research activity. This is particularly true of the proposed models which would likely result in a massive shift in activity to grant application (and hence grant reviewing) probably at the expense of other activities such as PhD supervision, and community activity that supports or encourages research. Furthermore, the system would no longer incentivise "just doing good research".

Question 6

In principle, do you believe that a metrics-based approach for assessment of funding can be used across all institutions?

No. This question is only relevant for models A and E, which we consider to be entirely invalid.

Question 7
Should the funding bodies receive and consider institutions' research plans as part of the assessment process?

Research plans are most useful in situations where expenditure is largely focused on facilities. In Mathematics, where expenditure is mainly focused on research personnel, research plans are less relevant. QR is for maintaining a group of good researchers who are most likely to do good research. Funds should therefore support achievement and promise, rather than conformity with a so-called 'plan'.

However, some useful information could be obtained by outline plans at the departmental level but we do not believe that the funding bodies could adequately assess these plans, other than by peer review. A brief outline of the unit's thinking about future research should therefore be evaluated, but by people who are qualified to do so.

Question 8

How important do you feel it is for there to continue to be an independent assessment of UK higher education research quality for benchmarking purposes? Are there other ways in which this could be accomplished?

Section 6.4 considers "the requirement to provide international benchmarking of research quality, both to guide funders and to provide assurance to Government that funds are being allocated effectively." We acknowledge this requirement - support for research from public funds must be open to public scrutiny. However, the 'purpose' must surely also be to sustain areas of mathematical research in which the UK is well-regarded, and to stimulate some areas that are falling behind.

There is a worrying disjunction between evaluation of quality and allocation of funds which must be addressed (see our answer to Q3). We believe that the total funding for Mathematics is due for a significant increase, since Mathematics, including what used to be called 'Pure Mathematics', underpins most new technology, in particular the information and security sectors. It would be disastrous if an inappropriate funding mechanism resulted in further erosion of the Mathematics base, especially at a time when the need for a highly-educated mathematical profession is most pressing. The point about supporting high-level creative research (see our answer to Q1) is also relevant here.

We have doubts as to the reliability of most of the sources described in Section 4 in providing a measure of overall UK research quality in the Mathematical Sciences that would be robust and meaningfully calibrated. The exception, in 4.6, refers to international panels, commissioned by Research Councils, to review the health of particular subjects. In our view, these reviews provide exactly the benchmarking required, but are, of course, based on peer review.