Consultation on the design of the UK's future research assessment system: Academy of Nutrition Sciences response May 2022

Questions

Section one: purposes of research assessment

The current assessment exercise serves three primary purposes:

- inform the selective allocation of funding to HEIs for research;
- provide accountability for public investment in research; and
- provide benchmarking information.

In addition, an independent review of REF 2014, carried out by Lord Stern in 2016 identified three further purposes:

- provide an evidence base to inform strategic national priorities;
- provide an evidence base for HEIs and other bodies to inform decisions on resource allocation;
- create a performance incentive for HEIs.

The funding bodies have set out their intention to retain the link between assessment outcomes and funding, and to require any future exercise to provide accountability for public investment in research.

Questions

1. In addition to enabling the allocation of research funding and providing accountability for public investment in research, which purposes should a future UK research assessment exercise fulfil? Select all that apply

    a. **X** Provide benchmarking information
    b. **X** Provide an evidence base to inform strategic national priorities
    c. **X** Provide an evidence base for HEIs and other bodies to inform decisions on resource allocation
    d. Create a performance incentive for HEIs.

2. What, if any, additional purposes should be fulfilled by a future exercise?

Collection and analysis of demographic and career progression data according to UoA and Institution-to aid the assessment of any differences and alignment with narrative information provided in the Environment section. This would also require further guidance on actions needed to achieve progressive, inclusive, sustainable/sustained research environments.

3. Could any of the purposes be fulfilled via an alternative route? If yes, please provide further explanation.

4. Do you have any further comments to make regarding the purposes of a future research assessment system?

NO. Given the state of data gathering by more passive processes e.g. information from research funders, publishers etc. and uncertainties about usability of such information, it is difficult to envisage an alternative route that could fulfil the ambitions for universality and fairness offered by the REF approach.

Section two: setting priorities

The roundtable discussions identified a number of priorities and guiding principles that participants believed should drive the development of a future assessment exercise. Some of these principles potentially preclude or conflict with others. It is therefore important for the funding bodies to understand which principles the sector would like them to prioritise when designing a future system.

Questions

5. To what extent should the funding bodies be guided by the following considerations in developing the next assessment system? Please rank the considerations from 1 (most important) to 9 (least important)

AVGE SCORE

a. **4.3** Ability of the system to promote research with wider socio-economic impact.
b. **1.8** Comparability of assessment outcomes (across institutions, disciplines and/or assessment exercises)
c. **3.0** that the bureaucratic burden of the system is proportionate
d. **6.5** Impact of the assessment system on local/regional development
e. **5.5** Impact of the system on research culture
f. **3.3** Impact of the system on the UK research system's international standing
g. **6.0** Maintaining continuity with REF 2021
h. **9.0** Providing early confirmation of the assessment framework and guidance
i. **1.3** Robustness of assessment outcomes

RANKING = i, b, c, f, a, e, g, d, h

6. Relating to research culture, to what extent should the funding bodies be guided by the following considerations in developing the next assessment system? Please rank the considerations from 1 (most important) to 6 (least important)

   AVGE SCORE

   a. **1.5** Impact of the assessment system on research careers:
   b. **1.0** Impact of the assessment system on equality, diversity and inclusion:
   c. **3.8** Ability of the assessment system to promote collaboration (across institution sectors and/or nations)
   d. **4.3** Impact of the system on inter- and transdisciplinary research
   e. **4.5** Impact of the system on open research
   f. **4.0** Impact of the system on research integrity

   RANKING = b, a, c, f, d, e,

7. What, if any, further considerations should influence the development of a future assessment system? Please set out the considerations and indicate where they should be located in the list of priorities.

   How the system has / has not supported, encouraged and incentivised long-haul investment by HEIs in particular areas of research and changes in research culture. Some analysis is required of how HEIs have behaved and responded to previous REF exercises i.e. has REF proven to be a formative process? What evidence is there for this?

8. How can a future UK research assessment system best support a positive research culture?
   The processes (and metrics, if any) that are used should embody the key characteristics of a positive research culture. There is need for more robust capture of career development approaches and staff support infrastructures. Disincentives need to be considered to reduce the large movements of star performers that in the years leading up to the REF exercise.

   There is a strong justification for weighting the environmental component greater that 15%, if a core purpose of future research assessments is to assess and more importantly help to shape, “a positive research culture”, and sustained research excellence. Incentivisation and the development of progressive and inclusive research environments are essential for the sustainability of the discipline and the prevention of the ‘haemorrhaging’ of nutrition excellence to overseas research academic or commercial organisations. We live in uncertain times regarding future access to EU funding. The potential loss of researchers to EU organisations in order to have free access to EU funding which traditionally represents up to 50% of research funding in key food-related disciplines is a particular challenge. More robust and fit for purpose capture of staff support infrastructures and career development approaches should be developed (see response to Q16), to help promote a meaningful focus on improved research culture

Section three: identifying research excellence

The funding bodies agreed that the outcomes of the next assessment framework should continue to enable them to allocate funding based on research excellence. It is therefore important that the exercise adequately captures those elements that constitute excellent research and assesses them robustly against appropriate criteria.

Components of excellence

The REF currently assesses three elements:

- Outputs (60%)
- Impact (25%)
- Environment (15%)

Roundtable discussions suggested that a broader definition of excellence, which recognises and rewards a wider range of activities and inputs, may better support a healthy, inclusive research system. Participants recognised that some of these elements are already captured through the environment statement but called for increased weighting and/or a more structured approach to assessing elements such as open research practices and policies to support equality, diversity and inclusion.

At the same time, participants questioned how these components might be assessed robustly and consistently. It was agreed that robust indicators would be required and participants acknowledged that this may be challenging.

Questions

9. Which of the following elements should be recognised and rewarded as components of research excellence in a future assessment exercise?
   (Multiple options: 'Should be heavily weighted' – 'Should be moderately weighted' – 'Should be weighted less heavily' – 'Should not be assessed' – 'Don't know')
   a. Research inputs (e.g. research income, internal investment in research and in researchers) **Should be moderately weighted**
   b. Research process (e.g. open research practices, collaboration, following high ethical standards) **Should be moderately weighted**
   c. Outputs (e.g. journal articles, monographs, patents, software, performances, exhibitions, datasets) **Should be heavily weighted**
   d. Academic impact (contribution to the wider academic community through e.g. journal editorship, mentoring, activities that move the discipline forward) **Should be moderately weighted**
   e. Engagement beyond academia **Should be moderately weighted**
   f. Societal and economic impact **Should be heavily weighted**
   g. Other (please specify).

10. Do you have any further comments to make regarding the components of research excellence?
    If the REF is to engage in assessing and, importantly helping to shape, “a positive research culture”, then it could be argued that the weighting for Environment should be larger. However this cannot be achieved using the present narrative approach alone. Newly established and tested metrics for measuring research culture need to be developed. A more formative approach which encourages genuine culture change is required. The present approach is tokenistic and hard for assessors to distinguish different levels of achievement and / or aspirations to achieve higher level of performance between different submissions..

Assessment criteria

In assessing submissions, the REF expert panels assess three distinct elements of each submission, against the following generic criteria:

1. Outputs: The panels assess the quality of submitted research outputs in terms of their ‘originality, significance and rigour’, with reference to international research quality standards.
2. Impact: The panels assess the ‘reach and significance’ of impacts on the economy, society, culture, public policy or services, health, the environment or quality of life that were underpinned by excellent research conducted in the submitted unit.
3. Environment: The panels assess the research environment of the submitted unit in terms of its ‘vitality and sustainability’, including the approach to enabling impact from its research, and its contribution to the vitality and sustainability of the wider discipline or research base.

The detailed criteria are set out at Annex B.

Roundtable discussions revealed some concerns about the wider effects of the current REF assessment criteria. For example, the criterion of ‘originality’ in relation to outputs was seen to discourage the submission of replication studies and reviews, which are essential to driving forward high-quality research. Similarly, the criterion of ‘reach’ for impact was seen to discourage the submission of research with local impacts, despite clarifications in the Panel Criteria that reach should not be interpreted in this way.

Questions

11. Are the current REF assessment criteria for outputs clear and appropriate? (Yes/No/Don’t know)
    a. Originality **YES**
    b. Significance **YES**
    c. Rigour **YES**

12. Do you have any further comments to make regarding the criteria for assessing outputs?

YES. The Academy of Nutrition Sciences appreciates the challenges of identifying distinctive and functional criteria that are meaningful in assessing the quality of research outputs across a system involving 4 Main Panels, 34 UoAs and 140 submitting HEIs. Achievement of universality of assessment criteria is a notable feature of the REF which, along with peer review supported by metrics, needs to be retained.

Nevertheless, it is evident that assessors can sometimes struggle to reach agreement on numerical assessments for these three criteria. The Academy of Nutrition Sciences is concerned to ensure fair assessment is made of all research that is able to advance the discipline in important, measurable ways- even though not always at the highest end of novelty or originality. For many researchers the prime indicator of output quality is deemed to be the degree of originality or novelty of the output- and this often sways numerical assessment towards the top end of 4 star. However care is needed that this is not always deemed an essential criteria for determining quality. Not all important or significant research involves discovery science where novelty is the key characteristic. Rigour, especially for research areas where reproducibility is an *a priori* requirement to advance a particular field or discipline, is especially important in the study of diet-disease relationships. Observational epidemiology can provide important clues to these relationships, but RCTs (which are extremely challenging to conduct in the context of whole dietary interventions), are essential to provide the level of certainty sufficient to support whole population policy recommendations for dietary change and effective prevention of disease. Equally the assessment of significance requires carefully balanced judgement, notably for recent outputs that have potential to be paradigm-shifting or ground-breaking but are not yet recognised as such at the time of the REF assessment.

It follows that this organisation is not supportive of a metric-only based assessment of output quality. Human judgements, based on discussion within- and across- related disciplinary experts, are essential to reach fair assessment of output quality based on originality, significance and rigour.

The Academy is concerned about the potential bias within biologically-based disciplines, including nutrition, due to the increasing dominance of 4 star assessments for outputs involving very large numbers of authors. Distortion in the distribution of scores according to output-type has the potential to adversely impact on the scoring of other types of outputs with detriment to other types of experimental approaches. No satisfactory solution to this has been found in either REF 2014 or 2021.

13. Are the current REF assessment criteria for impact clear and appropriate? (Yes/No/Don't know)
    a. Reach **Don't know**

b. Significance **Don't know**

14. Do you have any further comments to make regarding the criteria for assessing impact?

The Academy welcomes the introduction and further development of the Impact component of the assessment. Impact assessment is important in nutrition, where research is often directed towards the translational pathway such as in the development of effective diet policy, new product developments based on health effects, or nutritional enhancement of agricultural outputs.
A limitation of the Impact assessment process is that, to a large extent, REF panels assess Impact against the ambitions that the institutions set out for each individual impact case. This can mean that it is challenging to assess the "true" significance of impact cases with narrowly defined / local ambitions. 'Reach' remains a debated issue within the assessment process due to inconsistencies in scoring depending on the extent of the Reach achieved at the time of assessment. This appears to be especially the case with respect to policy-based case studies in those scenarios where policy actions remain to be fulfilled in terms of measurable uptake. Government departments seem to vary in their willingness to provide supportive comments with respect to reach.

As in the case of Output assessment more guidance is needed on the apportioning of Impact scores when a large team is involved.

15. Are the current REF assessment criteria for environment clear and appropriate? (Yes/No/Don't know)

a. Vitality Yes
b. Sustainability No

16. Do you have any further comments to make regarding the criteria for assessing environment?
There is an understandable tension between efforts to increase the objectivity of information reported in Environment statements and the extent to which the proforma places institutions in a straight-jacket that limits opportunity for describing features of their research environment**.**
Nevertheless many researchers consider the present approach to the environment element of REF needs to be carefully re-evaluated with respect to the core characteristics and how they can be assessed. The current narrative approach does not work well and in its present format does not justify an increase in the score allocation to >15%. New metrics need to be developed that reflect the aspirations for open, equitable and sustainable research cultures.
Suggested improvements for a more quantitative, focussed, measurable approach include:
The 'People' section 2 is currently highly narrative in nature making the challenge of verifying and scoring different submissions very difficult given the same generic text tends to appear in most of them. The people section should receive more attention and

better metrics on career development for that institution. Currently this section consists of little more than standard hoop jumping exercises at which bigger institutions excel- but which do not reflect many staff's lived experiences. A suggestion going forward is in its place REF **randomly** select (from the HEIs submission as submitted to REF at least 1y prior to the REF deadline) a number of staff members to provide their signed off Case Studies of the 'experience' of the research environment and progression since appointment/last REF. As with Impact Cases, the number of Staff Experience cases would be dependent on the size of the submission.
The REF assessment system does not adequately acknowledge long-haul track-record and investment by HEIs in particular areas of research. The Environment section should capture the temporal view more robustly including use of a standardised approach to report progress on key deliverable detailed in the previous REF assessment.

Section four: assessment processes

Frequency

Outcomes from the REF are used to inform the allocation of block grant funding to universities. This funding method ensures a degree of research stability and independence not provided by other funding sources, because the results of research assessment are used over a prolonged period and the funding can be used as providers choose rather than being directed to particular research programmes.

Participants at the roundtables were asked to consider the frequency and sequencing of assessment exercises. Currently, the REF takes place every 5-7 years and assessment of all disciplines takes place in parallel. It has been suggested that a more regular exercise could increase its formative element and would ensure that funding based on REF outcomes more accurately reflects recent performance. However, it was noted that this must be weighed up against the potentially destabilising effect arising from the uncertainty of funding outcomes on a more regular basis.

The funding bodies recognise that views on the frequency of a future exercise will depend on the overall design of the assessment system. For example, it would not be feasible to run the exercise as it currently stands every three years without significantly increasing the burden on the sector. However, the funding bodies are keen to understand in principle whether the sector considers the availability of more current information to be more important than the stability offered by a less frequent exercise.

Questions

17. When considering the frequency of a future exercise, should the funding bodies prioritise:
    a. stability
    b. currency of information
    c. **X** both a. and b.

d. neither a. nor b.
e. Don't know.

18. Do you have any further comments to make regarding the prioritisation of stability vs. currency of information?
In most cases, research is carried out for longer-term benefit so there is little imperative to undertake the exercise too frequently. Every 5- 7 years would seem to be sufficiently frequent.

Sequencing

During discussions on the frequency of the exercise, some roundtable participants expressed some appetite for moving to a rolling exercise, sequenced by main panel or by assessment element. It was suggested that this would remove some of the perverse behaviours linked to the cyclical nature of the REF, particularly around recruitment and publishing practices. It was suggested that this would also reduce burden at an institutional level as effort would be spread across a number of years, rather than focused on a single end point. As with the frequency of the exercise, any decision to move to a rolling exercise must be weighed up the potentially destabilising effect of such a change.

Questions

19. Should a future exercise take place on a rolling basis?
f. Yes, split by main panel
g. Yes, split by assessment element (e.g. outputs, impact, environment)
h. **X** No
i. Don't know.

20. Do you have any further comments to make regarding conducting future research assessment exercises on a rolling basis?

The advantages of a rolling programme are far from clear. A rolling programme might offer logistical and resource advantages for Research England through a smaller, highly-skilled and continuously employed administrative team. However, potential advantages for HEI and for research disciplines are much less obvious. For equity reasons, there is considerable benefit in assessing all UK university research at the same time. A rolling programme has potential to disadvantage some institutions, disciplines, and possibly individuals, particularly in the early phases of such a programme. It is not obvious that a rolling programme would reduce the burden on institutions or the wider academic community. Similarly, it is not evident that either the quality of the research assessment or its utility of its findings would be improved by undertaking a rolling programme.

Granularity

A number of the changes made between REF2014 and 2021 were intended to reduce the emphasis on the individual in order to shift the focus onto the submitting unit as a whole. At the same time, an institutional-level environment statement is being piloted alongside REF 2021.

Roundtable discussions emphatically rejected a return to a more individual-focused exercise. However, views were divided on the extent to which future exercises should retain the Unit of Assessment structure. Those in favour of a more institution-focused approach frequently cited the current (perception of) disadvantage to inter- and transdisciplinary research in a discipline-based system. It was also noted that many of the issues relating to research culture and environment can only be addressed at the level of the institution.

There was, however, concern amongst others that a move to a fully institutional-level assessment would conceal the 'pockets of excellence', particularly in less research-intensive HEIs, and would make REF assessment outcomes less robust indicators of excellence. This is an important consideration for the funding bodies, given the continued link with funding.

Questions

21. At what level of granularity should research be assessed in future exercises?
    a. Individual
    b. Unit of Assessment based on disciplinary areas
    c. Unit of Assessment based on self-defined research themes
    d. Institution
    e. **X** Combination of b. and d.
    f. Combination of c. and d.
    g. Other (please specify)

22. Do you have any further comments to make regarding the granularity of assessment in a future research assessment exercise?
    Nutrition research is returned to different UoA by different institutions. In most cases, this is likely to be UoA3 or U0A 6 but nutrition research is also returned in other UoA including 1 and 2. This diversity reflects the broad nature of the discipline and the different settings in which nutrition research is undertaken in different UK institutions. This leads to two main problems for nutrition: i) it is difficult to know whether nutrition research is being assessed on the same basis in each of the UoA and ii) the totality of UK nutrition research is largely "invisible" in the REF which makes it difficult, as a discipline, to identify structural problems or to celebrate collective success.

    Self-defined research themes offer considerable opportunities to foster high quality inter-disciplinary research within and between HEI. As a discipline, nutrition is central to several themes that address e.g. food systems with a focus on environmental issues or human health. However, because such self-defined themes differ in focus, scope and disciplinary make-up, it is difficult to see how such granularity could be incorporated into the current REF processes that are largely discipline-based. One way forward would be to raise the profile of inter-disciplinary research within the REF and to strengthen the

processes for its recognition by HEIs at the submission stage and its assessment subsequently.
Moving to assessment based on self-defined research themes would be disadvantageous for equity and would increase the difficulty in ensuring that assessment was rigorous and fair

Metrics

Roundtable discussions suggest limited appetite for increasing the role of metrics in the assessment of outputs. However, there was greater support for exploring quantitative indicators in the environment section. The use of metrics in the REF has been discussed at length, most notably in the 2015 Metric Tide Report. However, discussions persist in the sector.

Questions

23. To what extent and for what purpose(s) should quantitative indicators be used in future assessment exercises? (Please select as many as apply)
    a. Move to an entirely metrics-based assessment
    b. Replace peer review with standardised metrics for:
        i. Outputs **NO**
        ii. Impact **NO**
        iii. Environment **NO**
    c. Use standardised metrics to inform peer review of:
        i. Outputs **YES;**
        ii. Impact **NO (?)**
        iii. Environment **YES;**
    d. Should not be used at all.
    e. Other (please specify)

24. Do you have any further comments to make regarding the use of metrics in a future research assessment exercise?

    Provided that the data are robust, the use of standardised metrics to inform peer review can be helpful. However, there are too many uncertainties and risks to consider moving to (or even towards) an entirely metrics-based assessment. Such metrics are likely to be least useful for recent outputs particularly those that have potential to be paradigm-shifting or ground-breaking.

Burden

The cost and bureaucratic burden of the REF are frequently cited in criticism of the exercise. Roundtable discussions identified some sources of burden specific to the current exercise (e.g. special circumstances procedures), along with the overall scale and complexity of the exercise. However, several participants stated that the bureaucracy is, to a certain extent, generated by institutions' approaches to the REF and can be difficult to distinguish from activities that would be carried out as part of business as usual or in response to requirements elsewhere in the system (e.g. by research funders). Several respondents expressed scepticism that burden would increase or diminish significantly with changes made to the exercise. It was also noted that changes may, in themselves, create additional burden for institutions regardless of their nature or intent. While UK-wide research assessment falls outside the scope of the ongoing 'Independent review of research bureaucracy' commissioned by the Department for Business, Energy & Industrial Strategy, related discussions have largely echoed the views expressed in the roundtables.

25. How might a future UK research assessment exercise ensure that the bureaucratic burden on individuals and institutions is proportionate?

Maintain, as far as possible, core elements of the current exercise and ensure that any new elements/ changes are well tested, understood (and accepted) by institutions as early as possible in the REF cycle. Minimize the changes required/ build clarity into existing requirements, and give institutions plenty of time and opportunity to establish a common understanding of what is required. Keep changes between REFs to a minimum.

Make environment statement less verbose, quantitative and based on metrics;