

March 2015

Research Excellence Framework 2014: Manager's report

Contents

	Page
1. Introduction	1
2. Background and policy development	2
2.1. Origins of the REF	2
2.2. Scoping and initial consultation (2007)	4
2.3. Bibliometrics pilot (2008-09).....	6
2.4. RAE reviews and panel feedback (2008-09)	8
2.5. Second consultation and initial decisions (2009-10)	11
2.6. Impact pilot and decisions (2010)	13
3. Units of assessment and panels	16
3.1. UOAs and panel structure	16
3.2. Recruitment of panel members	18
Criteria and guidance	21
4.1. Guidance on submissions	21
4.2. Panel criteria	26
4.3. Equality and diversity guidance.....	29
5. Submissions	31
5.1. Codes of practice review	31
5.2. Submission data requirements	32
5.3. Submission system	34
5.4. Citation data	36
5.5. Multiple submissions and classified case studies.....	37
5.6. Research income and PGR data	39
5.7. Supporting institutions	41
5.8. Survey of submission intentions	45
5.9. Summary of submissions	46
5.10. Collection of outputs.....	48
6. Audit and data verification	54
6.1. Sampling and data comparisons.....	54
6.2. Panel-instigated audits.....	58
7. Preparation for assessment	63
7.1. Recruitment of assessors and specialist advisers	63
7.2. Assessment scheduling	67

7.3. Assessment systems.....	71
8. Assessment.....	76
8.1. Distribution of submissions and outputs	76
8.2. Assessment process	81
8.3. Roles and responsibilities	85
8.4. Administrative support.....	89
8.5. Panel secretariat	93
8.6. Individual staff circumstances review.....	96
8.7. Policy support and guidance.....	98
8.8. Results and outcomes.....	100
9. Impact of the REF.....	103
9.1. Cost and regulatory impact	103
9.2. Impact on equalities, staff and privacy.....	105
9.3. Economic and sustainability impact	107
10. Project management	108
10.1. Project structure and staffing	108
10.2. Project planning and management	111
10.3. Budget	112
Annexes.....	114
Annex A: Publications	114
Annex B: Survey of submission intentions.....	116
Annex C: Project plan.....	118

1. Introduction

1. The 2014 Research Excellence Framework (REF) was managed by the REF team, based at the Higher Education Funding Council for England (HEFCE), on behalf of the four UK higher education (HE) funding bodies. The REF was developed by HEFCE, working with the other UK funding bodies, between 2007 and 2010. It was implemented from 2010 to 2014. This report provides the REF manager's account of how the REF framework was developed, and how the REF team managed the implementation of the 2014 REF exercise.

2. The first part of the report (background and policy development) describes the development of the REF until the key parameters of the assessment were decided in 2010. It is largely descriptive with some general observations; it does not provide recommendations for future policy development as this will depend on prevailing policy imperatives at the time. The policy development stage of the REF was led by the research policy team at HEFCE rather than the REF team, which was established in 2010, once the key parameters of the REF had been decided. The author of this report was heavily involved at all stages but did not take up the role of REF manager until 2010.

3. Subsequent sections of the report cover the implementation of the REF from 2010 to 2014. They include the REF team's reflections about what worked well at each stage in the exercise, and recommendations about potential improvements to be considered for the next assessment exercise. Recommendations are made by the REF team based on our experience of running the 2014 exercise; they do not necessarily represent the views of the funding bodies. Also, many of the recommendations concern refinements to the guidelines, procedures or operation of the exercise, and hence would be relevant only if a similar – rather than a substantially different – exercise is carried out in future.

4. The REF team worked very closely throughout the exercise with the institutions that made submissions to the REF; with the expert panels that carried out the assessment; and with the panel secretariat that supported the panels. While this report focuses on the work of the REF team, it also comments on issues encountered by and feedback received from institutions, the panels and the panel secretariat. These comments are provided from the perspective of the REF team and our view of how others responded to or were affected by the REF. For an account of how the REF panels conducted their work and their observations on the assessment, please see the [panel overview reports](#) and the [EDAP report](#).

5. The report that follows provides a detailed account of how the REF was managed. Those interested in a summary rather than the details can read the 'key points' at the start of each section.

Graeme Rosenberg,
REF Manager.

2. Background and policy development

2.1. Origins of the REF

Key points

- The UK system of research assessment had evolved since the first research selectivity exercise was conducted in the 1980s. As successive exercises were carried out, they became more sophisticated but also more burdensome on the HE sector.
- During 2006 – prior to the completion of the 2008 RAE – the then government decided that the RAE should be replaced by a metrics-based system for assessing and funding research, to substantially reduce the administrative burden associated with the RAE.
- Initial consultations by the government concluded that this change would be accepted by the sector so long as the metrics measured research *quality*, and that expert advice would be used wherever *quality* metrics were unavailable.
- In 2007 HEFCE was invited to develop plans for such a framework. A lengthy period of consultations, expert advice and pilot exercises followed – accompanied by a shift in the overriding policy imperatives – until the key parameters of the new system were agreed in 2010.
- In this context, HEFCE and the UK funding bodies successfully maintained stakeholder confidence in the process, through extensive dialogue and involvement of stakeholders in developing the assessment framework.

6. The UK has a highly-developed national research assessment system that forms a key part of the four HE funding bodies' performance-based funding systems. While the primary purpose of successive assessments has been to inform block-grant research funding to HEIs, they have also become increasingly important for benchmarking and reputations, for the management and strategic development of research activity, and as markers in the academic careers of individual staff. The exercises have had an underlying aim of providing a general stimulus to improving the quality of UK research, and are generally regarded as having succeeded in doing so.

7. The first national exercise was carried out in 1986 (although not at that stage used for funding), and was then refined and developed through each subsequent exercise: in 1989, 1992, 1996 (at which point it became known as the Research Assessment Exercise or RAE), 2001, 2008, and most recently the REF in 2014. At the heart of each exercise has been the notion of peer review, with the quality of research judged by the peers of those being assessed, according to criteria determined by the peers in consultation with the wider academic community.

8. Each successive exercise tended to become more sophisticated in its approach and more transparent in the criteria, than the previous one. To a large extent, this was in response to pressure from those directly involved in the assessment (that is, academics whose work is assessed, submitting institutions, those carrying out the assessment, or a combination of these).

However, with increasing sophistication and more detailed and transparent assessment criteria, the exercise has also tended to become more burdensome with each round.

9. In 2006, while the 2008 RAE was still being conducted, the then government was highly receptive to concerns that were expressed in the HE sector about the substantial burden involved in the 2008 RAE. In the context of a wider policy to reduce public sector bureaucracy, the government decided to greatly reduce the administrative burden of the RAE, by replacing it with a metrics system for assessing and funding research in England.

10. A working group co-chaired by the HEFCE Chief Executive and the Director General for Higher Education at the then Department for Education and Skills (DfES) prepared a [consultation](#) on options for the reform, proposing various combinations of metrics that could be used to drive block-grant research funding.

11. Following the consultation exercise, in December 2006, the Education Secretary [announced](#) that a new more metrics-based framework would replace the RAE after the 2008 exercise. HEFCE was invited to develop proposals for the framework, taking account of key issues raised in the consultation, and working with the other UK funding bodies if they wished to take part. These key issues were:

- that a metrics-based system should continue to involve expert advice
- that indicators or metrics should focus on research quality (rather than, say, inputs such as research income)
- that disciplinary differences would need to be recognised, and in particular, that metrics are more readily applicable to assessment in the sciences.

12. Working with the other UK funding bodies, HEFCE then started to develop the new framework through what turned out to be an extended series of consultations, pilot exercises, commissioned projects and expert advice. The ensuing development involved a shift away from the initial policy imperative to greatly reduce burden, for the reasons described below.

13. In March 2007, [HEFCE outlined its plans](#) for developing the new arrangements. At that stage, the key aims of the new framework were expressed as:

- to produce robust UK-wide indicators of research excellence for all disciplines which can be used to benchmark quality against international standards and to drive the Council's funding for research
- to provide a basis for distributing funding primarily by reference to research excellence, and to fund excellent research in all its forms wherever it is found
- to reduce significantly the administrative burden on institutions in comparison to the RAE
- to avoid creating any undesirable behavioural incentives
- to promote equality and diversity
- to provide a stable framework for our continuing support of a world-leading research base within HE.

14. The stated intention was to develop 'a single overarching framework for funding and assessment within which a differentiated approach is possible for groups of disciplines.' The framework would focus on indicators of research *quality* across all disciplines. We envisaged two

distinct processes for assessing quality – one for science-based subjects and another for all other subjects – but with the intention of producing a unified set of quality outcomes, and some common indicators across all subjects.

15. A further consideration from the outset was for the new framework to recognise quality or excellence in both curiosity-driven research, and research undertaken directly in response to the needs of research users.

2.2. Scoping and initial consultation (2007)

Key points

- Within a metrics-based assessment system, bibliometrics were considered to provide the best chance of identifying indicators of *quality*, at least for the science-based disciplines.
- However, we recognised that such indicators were both controversial and technically complex, and it was vital that from the outset we:
 - Commissioned credible experts in the field of bibliometrics to provide advice.
 - Maintained extensive formal and informal dialogue with the sector.
 - Proceeded to develop the new framework at an appropriate pace, by carrying out a scoping phase followed by a pilot exercise.
- The first consultation on the REF highlighted widespread concern about separate systems for assessing the sciences (based on metrics) and other subjects (based on peer review), and led to a commitment to develop a single unified assessment framework.

16. HEFCE envisaged that bibliometrics offered the best chance of providing a fit-for-purpose indicator of research quality, at least for the sciences. Our starting point was to commission a [scoping study](#) on the potential use of bibliometrics, by CWTS at Leiden University, widely recognised as leaders in the field. We also commissioned a study on the implications of using bibliometric indicators for [interdisciplinary research](#), by Evidence Ltd.

17. As part of the scoping study, CWTS hosted an international seminar of invited bibliometricians from several countries in Europe (some of which used bibliometrics to inform national funding systems) and Australia, to sound them out on the key issues. We also held a number of informal discussions with the sector about the scoping study, provided a special briefing to the HEFCE Board, and sponsored a national event to discuss these developments.

18. Based on the expert advice and our informal discussions with stakeholders, we took the view that bibliometric techniques had the potential to provide usable indicators of research quality for the science-based disciplines, but only if applied at the level of very broad subject groups to mitigate the problems of individual anomalies and inaccuracies in the data. At the same time, we identified considerable further work that would be required, in terms of:

- Scope: in particular, how to identify the relevant staff and/or their publications and link them to the relevant HEI, and over what timeframe.

- Data quality: especially the coverage, completeness and accuracy of the available bibliometric databases.
- Methodology: especially in terms of normalising citation rates (by field, age and other factors) and addressing self-citation of various kinds.
- Other limitations: for example, applied research and review papers.
- Potential behavioural consequences: particularly in terms of publication behaviour or citation 'clubs', and potential impact on equal opportunities.

19. Building on the scoping study and other informal discussions, in November 2007, HEFCE issued a [first consultation on the REF](#), applicable to England and Northern Ireland. The funding bodies for Scotland and Wales conducted parallel consultations. Views were sought on:

- Assessment and funding in science-based disciplines to be driven by a new bibliometric indicator of research quality, and a new light touch peer review process for the other disciplines (specific proposals were not made for this but initial comments invited).
- Where the boundaries should be drawn between the two groups, and a proposed sub-division of the science-based disciplines into six broad subject categories.
- Other indicators such as research income and research student data to be used for all disciplines, and ideas about further indicators – particularly to capture 'user value'.
- The role of expert panels in the science-based disciplines in advising on the selection and use of indicators, and in other subjects to conduct the light touch peer review process.
- Further work and a proposed timetable for developing the new framework and phasing it in as soon as practicable.
- Whether the new system would greatly reduce the burden of research assessment, and views on whether it might have a negative impact on equal opportunities.

20. Given that the reforms were controversial, bibliometrics were technically complex, and the intended timetable ambitious, it was vital throughout these developments to have extensive dialogue with the sector. As part of that dialogue, CWTS gave numerous presentations and played a large part in maintaining the confidence of the sector due to their evident integrity, thoroughness, and balanced view of the issues. It should be noted for the record that the expert bibliometricians at CWTS were rather more circumspect about the potential for using bibliometrics as the key indicator in a national assessment system, than had been proposed in the HEFCE consultation. In our discussions, CWTS maintained the view that bibliometrics were very useful within the context of peer review, but were too fraught with problems to replace it.

21. Within HEFCE at this stage (prior to pilot exercise) there was some uncertainty about whether bibliometric indicators would be able to fulfil their intended role, but it remained necessary to proceed with the proposals and care was taken to do this at an appropriate pace. Importantly, the scoping phase was to be followed by a thorough test of bibliometrics through a pilot exercise.

22. In May 2008 HEFCE announced the [outcomes of the consultation](#), which set out plans for the bibliometrics pilot exercise, and made two main modifications to the early proposals:

- The intention to move away from separate systems for the sciences and other subjects to a more unified framework (using a variable mixture of bibliometrics, other indicators and expert review as appropriate in each discipline).
- An extension to the previously announced timeframe for introducing the new framework.

2.3. Bibliometrics pilot (2008-09)

Key points

- A substantial pilot exercise was successful in making stakeholders more fully aware of what would be involved in using bibliometrics for national research assessment. It highlighted numerous shortcomings and problems with such indicators, and as a result a widespread consensus emerged that citation indicators would not be sufficiently robust to replace expert review.
- A key underlying concern about bibliometrics was the nature of the behavioural incentives that such a system might introduce. Although peer review was recognised as imperfect and burdensome, peers could be trusted to value quality above all else. It was feared that metrics, on the other hand, could encourage various strategies and behaviours to improve metric performance rather than research quality.
- Subsequent analysis of the pilot data revealed a significant gender bias in citations, which served to further reduce our and the sector's appetite for using such data in the assessment.
- An unintended consequence of the pilot exercise was to highlight limitations in HEIs' own information about their research publications, which stimulated substantial improvements in research information systems within the sector.

23. During 2008-09, HEFCE conducted a substantial [pilot exercise](#) to test and develop bibliometric indicators of research quality for use in the REF. Twenty-two UK higher education institutions took part, covering 35 units of assessment (UOAs) from the 2008 Research Assessment Exercise (RAE). Evidence Ltd were commissioned to carry out the data collection and analysis of the data; Technopolis were commissioned to gather feedback from the pilot HEIs; and we sought feedback on how the indicators might be used as an evaluation tool from the Expert Advisory Groups (discussed below). In-house capacity was also developed in HEFCE, to assist with various aspects of the analysis.

24. The pilot exercise aimed to:

- explore which subjects should use bibliometric indicators under the new framework
- test different models for including staff and/or publications in citation indicators
- test the main sources of citation data: the Web of Science (WoS) and Scopus
- develop the process for collecting bibliographic data
- develop and test methods for analysing and normalising citation data, and for producing a meaningful citation 'profile'.

25. Papers published between 2001 and 2006, and their citations up to 2007, were included in the pilot, using both of the main commercially available citation databases. Three main models for producing citation indicators were tested:

- Model 1: all papers associated with the HEI. Data were harvested from the publication databases for all 2001-06 publications according to the inclusion of an address associated with the HEI. These data were then mapped to UOAs.
- Model 2: all papers by selected staff only. This included all 2001-06 publications produced by staff that had been submitted by the HEIs to RAE 2008.
- Model 3: selected papers by selected staff only. This included only the six most highly cited papers by staff that had been submitted by HEIs to RAE 2008.

26. The pilot involved three phases:

a. Data collection. Data about institutional papers were harvested directly from WoS/Scopus for model 1. Data about staff and papers were collected from the pilot HEIs and matched to WoS/Scopus for models 2 and 3. This part of the project was carried out principally by Evidence. Evidence provided an early [report on data collection](#), which highlighted a number of technical issues that would need to be resolved. Technopolis [reported](#) that the pilot HEIs' experience of the data collection and the amount of effort involved varied widely, depending on the state of their existing research information systems (with many institutions reporting substantial gaps in their publications data). They also reported that all the HEIs had gained a deeper understanding of the universal challenge of maintaining complete, accurate and consistent bibliographic records, and the majority planned to upgrade their research information systems or procedures as a result.

b. Citation counts and normalisation. The number of citations to each paper was counted and normalised by field, year and document type. Evidence calculated the normalisation factors for WoS; HEFCE for Scopus.

c. Analysis. We produced citation indicators for the papers associated with each UOA at each pilot HEI, using each model (and some sub-variants). Advice on the interpretation of these indicators was sought from the Expert Advisory Groups (EAGs) and the pilot institutions, and a summary was published in an [interim report](#). Further details including all the indicators were published in the [final report](#).

27. The EAGs, the pilot HEIs, and the sector more widely on reading the reports, shared numerous concerns about the fitness-for-purpose of the citation indicators produced in the pilot exercise. A strong consensus emerged that citation indicators were not sufficiently robust to be used formulaically or to replace expert review in the REF. However there remained scope for citation information to be used to inform expert review. For this purpose, model 3 (citations to selected papers only) emerged as the preferred model, for reasons given in the [second report](#) by Technopolis.

28. During the course of the pilot exercise, therefore, it was accepted that the initial intention of replacing expert review (at least in the sciences) by a metrics driven system would not be

workable, and by implication that greatly reducing the burden of assessment would no longer be the primary policy goal of the reforms.

29. Some time after the pilot report, HEFCE completed further analysis on the potential implications for equality and diversity of using citation indicators. We linked the pilot data to Higher Education Statistics Agency (HESA) staff records and analysed the relationships between staff characteristics and citation rates. The [analysis](#) revealed a significant bias in the citations data, with men significantly more likely to be highly cited than women (at least in the age ranges 32-63). These findings published in 2011 served to add further caution to the use of citations, even at a secondary level, to inform peer review judgements.

2.4. RAE reviews and panel feedback (2008-09)

Key points

- In addition to the bibliometrics pilot, we commissioned evaluations of aspects of the RAE and sought expert advice to help inform developing proposals for the new framework. These included:
 - A review and analysis of the equality and diversity aspects of the RAE, which highlighted that substantial efforts would be needed to improve equality and diversity in the REF.
 - A review of the accountability burden of the RAE.
 - A series of Expert Advisory Groups, mostly comprising former RAE panellists.
- We also began to explore how the assessment could better take account of 'user valued' research. As this issue rose up the political agenda, we began discussions about how research 'impact' could be explicitly assessed as part of the framework. A review of international approaches to assessing research impact confirmed our initial view that a case study approach would be the most fruitful to pursue.
- Discussions with former RAE panellists at an early stage in developing the new framework were invaluable in understanding the panellists' appetite for the various changes being proposed, and in identifying improvements. In some areas (for example 'merging' some of the UOAs) they supported more substantial changes than we otherwise would have thought acceptable.

Recommendation

- The funding bodies should consider gathering feedback from the REF panels at an early stage, to inform changes and improvements for the next REF.

Review of accountability burden in the RAE

30. During 2008 we appointed PA Consulting, as part of a wider review of the accountability burden on higher education institutions, to identify the accountability burden associated with the 2008 RAE, and assist in identifying opportunities for the REF to reduce burden.

31. They [reported](#) in 2009 that the total sector cost to HEIs in England of RAE 2008 was approximately £47 million. The review found that the RAE was perceived as a high-cost obligation which nonetheless had multiple benefits for institutions, for example in terms of internal management and reputation, and which was proportionate to the level of funding available.

32. Some suggestions were made about how the burden might be reduced in future, and these were discussed in detail at a [workshop](#) with Pro-Vice Chancellors for research. They were generally in favour of institutions continuing to select staff for assessment, and supported some changes to simplify the exercise for institutions. These included the merging of 'smaller' UOAs, and more consistency across the UOAs. Further information about efforts to streamline the assessment is in sections 5.2 and 9.1.

Review of equality and diversity in the RAE

33. The Equality Challenge Unit commissioned the Centre for Diversity Policy Research and Practice at Oxford Brookes University to review the equality and diversity measures taken in the RAE, to help identify institutional good practice and inform our approach for the next exercise. The [report](#) concluded that substantial progress had been made, and also identified a number of areas that could be improved in future. These issues were carried forward in developing the REF, as discussed in section 4.3.

34. After the HESA staff data became available, HEFCE undertook a quantitative analysis of the selection of staff for the RAE, by staff characteristics. The [report](#) showed that staff selection to the RAE2008 was similar in composition to that seen in RAE2001. The data showed 67 per cent of men were selected compared to 48 per cent of women; and that selection rates were at similar levels for all ethnicity groups except those staff in the Black ethnic group, whose selection rate was lower. The review and the analysis highlighted that substantial effort would be needed in the next exercise to improve equality and diversity.

User valued research and impact

35. As the discussions about greater use of metrics proceeded, HEFCE also considered how the assessment could capture the value of research to users. Initially we explored how this could be addressed by building on the RAE approach, for example by extending the use of brief statements of user significance of individual outputs (that had been invited by the engineering panels in the RAE), and identifying indicators of user value. A workshop was held jointly with UUK in late 2008, including representatives from across the main subject groups. They broadly supported the aim of capturing user value in the REF, while maintaining the focus on quality. A number of challenges were identified including the multi-dimensional nature of user value and the limitations of available quantitative indicators. The general view was that the assessment would need to rely on expert judgement and that users should be more involved in this.

36. As the discussions proceeded and in the context of impending public sector budget cuts, officials in BIS and HEFCE took the view that in order to convincingly demonstrate the value of public investment in research, the REF would need to explicitly assess the value or impact of research on the economy and society. In addition, there was a strong view held in some quarters

that successive RAEs, with criteria based (almost entirely) on academic quality, had led to discouragement of research that was intended to benefit users.

37. In January 2009 the Secretary of State's [annual letter](#) to HEFCE indicated two main priorities for developing the REF: that it 'should reduce the burden on institutions and take better account of the impact research makes on the economy and society'.

38. During the course of 2009 (when it was also clear that the assessment would not be replaced by metrics), discussions began with the EAGs and informally with numerous groups about an explicit element of the framework for research 'impact'. This included a workshop with a range of research users from the public, private and third sectors, held in mid-2009.

39. Our initial thinking was that case studies would likely be the best, or perhaps only, way to explicitly assess research impact, given the intention from the outset to accommodate a wide range of types of impact, and research across all disciplines.

40. To inform thinking and wider discussions on impact we commissioned Rand Europe to survey other countries and identify existing approaches to assessing research impact, highlighting their potential applicability for national research assessment in the UK. The [report](#) confirmed the view that case studies would be the most fruitful approach to pursue, and outlined a number of key issues (such as time lags and attribution) that would need to be addressed.

Expert Advisory Groups

41. In December 2008, HEFCE (on behalf of the four UK funding bodies) established [Expert Advisory Groups](#) (EAGs) to help develop proposals for the REF. Their membership of around 100 people was drawn mainly from RAE2008 panels, and also included some nominees from the research councils, national academies and user organisations.

42. The EAGs met three times between February and June 2009 to discuss the following issues:

- experiences of the 2008 RAE and proposed key features of the REF
- the potential use of bibliometrics
- potential changes or improvements to how panels could review research outputs, including ways of reducing workload
- assessment of the research environment
- how the assessment could take account of the impact of research
- equalities issues
- the panel and UOA structure.

43. These discussions were managed in a similar way to focus groups. At each meeting, there were plenary sessions to introduce issues and breakout sessions to discuss them. Many of the breakout groups were deliberately mixed in terms of subject discipline, which proved to be very effective in reaching consensus on general issues. It also generated an understanding that a number of issues were indeed generic rather than discipline-dependent, as might otherwise have been assumed.

44. The EAGs were very valuable in gaining an understanding from the peer reviewers' perspective about potential improvements to the assessment framework, as well as their appetite for the proposed 'top-down' changes. In particular, the EAGs supported the following changes:

- more consistency across UOAs
- 'merging' some of the UOAs
- removing 'esteem' and in effect replacing it with the new impact component (which was understood to be a requirement for the new framework)
- cautious use of bibliometrics to inform but not replace peer judgement, and only in some disciplines.

45. These discussions were invaluable in providing the funding bodies with confidence that proposed changes to the assessment system would be supported or accepted by those we rely upon to carry out the assessment and, by extension, the wider academic community. The groups also raised a number of specific detailed points such as the definitions of staff eligibility and the difficulties associated with category C staff in the RAE, and operational issues that were very helpful in informing more detailed management of the REF, and to improve the quality of 'service' provided to panels in future.

2.5. Second consultation and initial decisions (2009-10)

Key points

- The second consultation set out detailed proposals for the new framework, centred on expert review of outputs as the most important element in all disciplines. Separate elements to assess research impact and the research environment were also proposed.
- The proposals were based on extensive development work and numerous discussions with stakeholders. Responses were broadly supportive and constructive, and enabled the funding bodies to take decisions on most key aspects of the framework.
- The proposal for an explicit assessment of the non-academic impact of research was controversial, as anticipated. Clear plans at the outset for a substantial impact pilot exercise were essential.
- Further decisions relating to impact and the weightings of the three elements of the assessment were taken after the impact pilot exercise; and decisions about the panel configurations were taken after further consultation.

46. In September 2009 the four UK higher education funding bodies issued a [second consultation](#) which set out detailed proposals for the new framework. The aims of the REF were revised to reflect the new emphasis on impact, and no longer emphasised reducing burden as a key aim. As the document explained, 'widespread consensus has emerged that while metrics should inform expert review they are not sufficiently robust to replace expert review. While we remain concerned to reduce the burden of assessment, we believe we have exhausted the main options for any radically different alternative approach. The REF will be driven by a process of expert review, informed by metrics.'

47. Hence the consultation stated that the REF will aim to:

- Drive up quality across the HE research base and in all forms of research.
- Support and encourage innovative and curiosity-driven research, including new approaches, new fields and interdisciplinary work.
- Reward and encourage the effective sharing, dissemination and application of research findings and the productive interchange of research staff and ideas between HEIs, business and other research users.
- Reward and encourage HEIs that deliver benefits to business, the economy and society by building on excellent research.
- Produce and publish quality assessments that are comprehensible, produced by a transparent process, and benchmarked against international standards and which identify the very best HE research wherever this is carried out.
- Support better management and sustainability of the research base.

48. Proposals were made and views sought on:

- The overall framework for assessment, involving three distinct elements (outputs, impact and environment).
- Assessing the quality of the research outputs of staff selected by the institutions, through expert review. In some disciplines, citation indicators would be provided as additional information to inform the assessment, and there would be more scope for 'double-weighting' of outputs.
- Assessing the impact of research on the economy and society, based largely on case studies, to be tested and developed through a pilot exercise.
- Assessing the research environment, through a narrative template and quantitative data.
- The configuration of panels and units of assessment (UOAs), and how to achieve appropriate levels of consistency between these.
- Measures to promote equalities and diversity.
- The timetable for implementation of the REF.
- The accountability burden implications of the REF and potential areas for streamlining the process, within the context of a peer review exercise.

49. As the proposals had been developed through extensive dialogue with many groups, the consultation responses held few surprises. We received 534 [responses](#), which overwhelmingly supported:

- block-grant research funding as part of the dual support system, to be allocated on the basis of research excellence
- research excellence to be assessed through expert review, informed by indicators
- the quality of research outputs to be the primary factor in the assessment, and the vitality of the research environment also to be a significant factor.

50. There was widespread support in principle for including an explicit assessment of impact. However, many respondents qualified their support by emphasising the need to develop a robust method for assessing impact, and suggested the weighting for this element should be lower than the proposed 25 per cent. A significant minority of responses objected to the inclusion of impact, typically for reasons described in paragraph 61.

51. As a result of the consultation, the funding bodies took decisions and [announced](#) the key features of the REF, except for:

- the configuration of panels which required further consultation
- the method for assessing impact, and the weightings between outputs, impact and environment, which depended on the outcomes of the impact pilot that followed.

52. Decisions were also taken to recruit chairs designate for the four main panels, and to establish an expert equality and diversity advisory group (EDAG) to help develop guidance on the equality aspects of the REF.

2.6. Impact pilot and decisions (2010)

Key points

- The inclusion of impact in the REF was a substantial development, potentially with much wider consequences beyond the assessment itself. A pilot exercise as well as extensive communication was essential to build stakeholder acceptance and support for the impact element of the REF.
- Initial communications responded to early fears and criticisms around the implications for curiosity-driven research, the impossibility of predicting future impact, and the need to avoid a narrow view of 'economic' impact.
- Institutions involved in the pilot were able to provide case studies rapidly, including many very impressive examples, based on research carried out prior to the 'impact agenda'. The quality of case studies in the pilot was variable, and a number of improvements to the format and guidance were identified.
- Many senior 'users' readily accepted invitations to take part on the pilot panels, but it became clear that additional efforts would be needed to engage much larger numbers of users when scaling up the exercise.
- Through the pilot and wider discussions, a broad consensus emerged that a case study approach would be workable across all disciplines. The possibility of a single approach for all disciplines had not been a forgone conclusion.
- The pilot and wider discussions also succeeded in clarifying the very broad scope of impact (well beyond economic), and identifying how the impact element would be applicable to the full range of disciplines.
- There were differing views about several specific details of the methodology, such as the inclusion of an 'impact strategy' and the minimum number of case studies required for small submissions. Such issues required judgement calls by the funding bodies, and will no doubt be revisited for the next exercise, in the light of the experience of the 2014 REF.
- The main point of debate at the end of the pilot was about the weighting of impact within the REF, and there was broad acceptance that it should be lower than the intended 25 per cent, at least for the first exercise.

53. During 2010 HEFCE, on behalf of the four funding bodies, ran a pilot exercise to test and develop the proposed approach to assessing impact in the REF. A [steering group](#) was established including people from the sector and user organisations to advise on: selection of the

disciplines and institutions to be involved; guidance for the HEIs and the panels; commissioning consultants to gather feedback; and interpretation of the outcomes.

54. The selected disciplines were:

- Clinical Medicine
- Physics
- Earth Systems and Environmental Sciences
- Social Work and Social Policy
- English Language and Literature.

55. Five panels were recruited, each with a roughly even mix of academics and research users. The academic members were drawn largely from the relevant RAE panels. To recruit the users we initially identified a range of organisations that appeared likely to cover the key types of impact in the pilot disciplines. It was useful to analyse the textual elements of the relevant RAE submissions, which frequently mentioned links with specific user organisations. We deliberately approached senior individuals from the organisations and found that many accepted the invitation very readily, although it was a little more difficult to recruit from industry than from the public or third sectors. As the pilot progressed, it became clear that substantial additional efforts would be needed to engage much larger numbers of users when scaling up the exercise, and we began to develop a specific strategy for developing user contacts, raising awareness and targeting communications for users.

56. HEIs were asked to express an interest in taking part and many did so. 29 were selected, and each was invited to submit to two panels. Guidance was developed in consultation with the steering group and the HEIs. They were invited to submit case studies (the number depended on the size of their RAE submissions), and an impact statement. Loose templates were provided, to ask for details of the impacts, the underpinning research, and how the research led to the stated impacts. The range of impacts and forms of evidence were largely up to the HEIs.

57. Institutions were able to provide case studies rapidly, and the cases were mostly based on research carried out prior to the 'impact agenda'. As it was a pilot and the timescale was short, HEIs felt able to take risks in submitting cases they were unsure about, and did not spend as much effort in preparing as they would for a real exercise. The panels found that the quality of case studies was mixed, but that there were very many impressive examples and these included types of impact that may not have been anticipated.

58. Each pilot panel reviewed and graded its submissions, and discussed its overall conclusions about the process and how it should be developed for the full exercise. The five panel chairs met on several occasions to share views and found that all five panels shared the same basic conclusions. All the panels found that the case study format enabled institutions to demonstrate impact and enabled the panels to make judgements they could be confident in – especially as they were judging cases on a scale (from 1 to 4 star), rather than trying to 'measure' the impact in absolute terms. They all found that the criteria of 'reach and significance' were appropriate. The five chairs decided therefore to write a single, collective [report](#) of the main outcome of their work – though it had not been a foregone conclusion that the five panels would agree a single set of conclusions and recommendations.

59. In parallel to the assessment by panels, Technopolis were commissioned to gather feedback and [report](#) on the lessons from the pilot HEIs' experience and views of making the submissions. Within the findings, they reported a general challenge of raising awareness amongst academics of what impact means in the context of the REF, with many apparently assuming that it meant economic impact. Like the panels, the pilot HEIs reported that case studies were an appropriate means of demonstrating impact from their research.

60. Overall the pilot concluded that the case study method would be workable. A number of recommendations were made for refining the format and guidance, some of which ought to be standard across the exercise, but some of which ought to be discipline-related and developed further by the REF panels, in consultation with their communities. The outcomes of the pilot were discussed at a series of events. The pilot panel chairs in particular played an important role in reassuring the community about some of their concerns. There appeared to be two general types of concern amongst the academic community during the course of these developments:

a. Concern (mostly from amongst the sciences) that long-term or curiosity-driven research would be disincentivised. At times it was conflated with the issue that impacts could not be predicted, and clarification was needed to distinguish the post-hoc REF assessment from the introduction by the research councils of impact pathways to accompany grant applications. It also led to the decision to have a lengthy window for research (20 years, with the option for panels to extend this to 25 years). We also took the view that curiosity driven research would not be disincentivised on the basis that research quality remained the primary factor in the assessment; and by agreeing that curiosity-driven research is indeed likely to have the most profound impacts, and therefore should feature in many of the strongest case studies.

b. Concern about the inherent difficulties of evidencing less tangible or non-quantifiable impacts, particularly on policy making and on public discourse; and the related concern that REF would privilege economic and/or readily quantifiable impacts. These concerns were expressed mostly amongst the arts, humanities and social sciences. To supplement the pilot exercise (which included English and SWSP) we ran three [workshops](#) with participants from across the arts, humanities and social sciences to explore these issues in detail. Attention was given to ensuring that REF adopted a very wide definition of impact (including for example, where research goes against the grain of prevailing policy), and also that quantitative evidence was not privileged above qualitative evidence.

61. There were also debates and discussions about some specific issues such as:

- the inclusion of an 'impact strategy' (or whether this should be part of environment)
- the minimum number of case studies required for small submissions
- whether impacts 'travel' with researchers when they move to another HEI
- the extent to which impact on teaching would count
- what kind of quality threshold should be set for the underpinning research
- how to deal with indirect links between research and impact, for example, where policy advice or public engagement is based on an academic's general standing and/or expertise, rather than their specific research. (Also, some people had difficulty with the

idea that REF would assess impact only from research, rather than other academic activity.)

62. On such issues where views were mixed in the sector, the arguments were considered and judgement calls were made by the funding bodies. These issues will no doubt resurface for discussion as the next exercise is developed, in the light of institutions' and panels' experiences of the 2014 REF.

63. Overall, by the end of the pilot there was broad acceptance through the various discussions that – with the recommended refinements arising from the pilot – the assessment of impact could be rolled out for the first full REF, but there was a very widespread view that the weighting for this first exercise should be lower than the intended 25 per cent. There was much discussion about what the appropriate weighting should be, and 20 per cent was settled on as broadly acceptable to all concerned. There was brief consideration about whether the weighting could vary by subject, but none of the subject communities appeared to have an appetite for having a lower weighting than others. In March 2011 the funding bodies [announced their decisions](#) on the key parameters for assessing impact in the REF, and on the weightings of the three components.

3. Units of assessment and panels

3.1. UOAs and panel structure

Key points

- The second consultation made proposals and sought views on the panel structure. There was widespread support for reducing the number of UOAs, primarily to enable greater consistency across the exercise, and also to reduce burden.
- Much of the proposed UOA structure was accepted, except for several disciplines that argued either for a different configuration or for a separate UOA. These were resolved through detailed dialogue with the relevant subject associations, with pragmatic compromises made in some cases.
- In July 2010, decisions were announced about the final UOA and panel structure; and about the various roles of all those to be involved on the panels (so that we could recruit panels to start meeting and develop the panel criteria during 2011).
- The total number of UOAs was reduced from 67 in the RAE to 36 in the REF; and from 15 to 4 main panels. The main panels were deliberately defined with a much stronger oversight role than in the RAE, to ensure greater consistency across UOAs. This was widely welcomed by institutions.

Recommendation

- For the next REF, the funding bodies should consider whether there might be any better alternatives for those disciplines where pragmatic compromises had been made in defining the UOAs for REF 2014.

64. The second consultation on the REF in 2009 proposed to rationalise the panel structure from 67 UOAs and 15 main panels in the RAE, to 30 UOAs and 4 main panels in the REF. This reconfiguration was proposed for several reasons, some of which had been put forward by the EAGs:

- To enable greater consistency in assessment.
- To reduce the number of (fluid) boundaries between UOAs and hence reduce the need for HEIs to make tactical decisions about what work to submit in which UOAs.
- To narrow the disparities in sub-panel workloads: in the RAE, the volume of full-time equivalent (FTE) staff submitted to each UOA had ranged from fewer than 100 to more than 3,000.

65. Specific proposals were made for 30 UOAs, based partly on advice from the EAGs (especially for the UOAs in Main Panel A). The consultation revealed widespread support for rationalising the panels, and general agreement with much of the proposed UOA structure. Mixed views or concerns were expressed about several specific UOAs that had been proposed. In these areas we entered into dialogue with the relevant subject associations, to negotiate UOAs that would be acceptable to the communities. Agreements were reached in all cases, although in some cases pragmatic compromises were involved, and some of the mergers were accepted on the basis that multiple submissions would be allowed.

66. The particular areas that required resolution were:

- Identifying a suitable configuration of engineering panels. The single proposed engineering UOA was felt by parts of the community to be too broad and unwieldy, and a solution was developed in discussion with the Royal Academy of Engineering. It involved a compromise in which only two of the three 'small' engineering UOAs from the RAE were merged to form larger UOAs. Also, in the event when submissions were made, a number of institutions decided to make a single engineering submission, and there was a large shift of material into the general engineering UOA.
- Several other 'small' (in terms of submitted volume) UOAs that we had proposed to merge argued for retaining their own distinct UOA, and their arguments were accepted even though their 'size' remained smaller than we had intended.
- A few very 'small' RAE UOAs suggested an alternative partner to merge with (such as anthropology and archaeology), or accepted the proposed merger as the best available fit, but with reservations about how cognate the resulting UOA would be and an expectation that multiple submissions would be permitted. In the event, all the sub-panels operated effectively but in the case of archaeology and to a lesser extent library and information management, there appeared to be little overlap in submitted content with many multiple submissions made, and the benefits of 'merging' were more limited.
- A number of discussions were held about the configuration of the area studies and language-based disciplines, and in the event all agreed that European studies would fit more coherently with area studies than with European languages, as had been proposed.

67. The final UOA structure as well as specifications for the roles and constitution of the main and sub-panels were [published](#) in July 2010. These decisions were required very early on – four and a half years before completion of the assessment – in order that the panels could be

recruited and could meet in a timely way to develop the assessment criteria, which in turn could be published well in advance of the submission deadline.

68. Given the agreements we had reached with the relevant subject communities, no further concerns were raised about the UOA configuration once it had been published. During the assessment, the broader UOA operated effectively and they enabled greater consistency of practice. Only occasionally, some minor concerns were raised later on about the difficulty of disaggregating the results by 'discipline' for some UOAs.

69. The role of the main panels was deliberately specified to ensure greater consistency across each family of UOAs. This included the main panels' ownership of the assessment criteria, explicit responsibilities for ensuring consistent adherence to the criteria and comparable standards of judgement, and responsibility for the results. Our intention was that by exercising greater oversight with only four main panels, the outcomes of the REF would be more consistent across UOA, addressing one of the main concerns about the RAE.

3.2. Recruitment of panel members

Key points

- The appointment of main and sub-panel chairs through an open applications process, rather than a nomination process, worked well.
- Attention was given to scaling the number of members per panel according to anticipated volumes of work, while maintaining sufficient breadth of expertise. This was largely successful, and gaps in expertise were effectively addressed later through the appointment of additional assessors.
- Careful selection of panel members was vital to gain the confidence of the academic community. After announcing the membership there were a small number of complaints about gaps or underrepresentation of particular sub-disciplines on some of the panels. These mostly related to UOAs that had 'merged' multiple RAE UOAs. Where appropriate the concerns were addressed through additional appointments.
- Efforts were made to improve the diversity and representativeness of the panels compared to the RAE, with some but limited success.

Recommendations

- Given the substantial role that many deputy sub-panel chairs played in supporting their chairs, the funding bodies should consider the timing, criteria and process for appointing deputy chairs in the next exercise.
- There is a very large list of (over 2,000) nominating bodies, who vary greatly in terms of their engagement with the REF. It would be beneficial to carry out a full review of the list and move to communicating with them electronically rather than by post in future.
- Further work is needed to improve the representativeness of the panels in future, including engaging with the nominating bodies. The funding bodies might also consider collecting demographic information about candidates at the nominations stage.

70. The [publication](#) announcing the UOA structure also set out the process for appointing panels and invited associations and organisations with an interest in research to nominate candidates for panel membership. The appointment process was designed to take place in three stages:

- Appointing the main and the sub-panel panel chairs, based on applications.
- Appointing the panel members through a nominations process during 2010 (so the panels could meet during 2011 to develop the panel criteria).
- At a later stage (after carrying out a survey of institutions' submission intentions), appointing additional assessors and specialist advisers to assist with reviewing submissions. This is discussed in section 7.1.

71. All main and sub-panel chairs, panel members and assessors were formally appointed collectively by the chief executives (or equivalent) of the four UK higher education funding bodies, on the advice of interview panels, panel chairs and/or the REF team.

72. The four main panel chairs were appointed first, through an open application and interview process in early 2010. They were appointed early on so they would be in place to advise on the selection of sub-panel chairs. During late 2010, the 36 sub-panel chairs were recruited through an open application process, rather than the nomination approach used in RAE2008. Clear criteria for appointment were laid out, and applicants were required to include statements of endorsement by subject associations and/or other organisations to demonstrate their standing in the community. Sub-panel chairs were appointed after taking advice from the main panel chairs, based on the applications and accompanying endorsements (interviews were not held). This process was effective in appointing highly qualified candidates who had the support of their communities.

74. Nominations for panel members (and assessors – though these were to be appointed later) were invited in July 2010. To ensure we would receive sufficient nominations, we first sought to compile as complete a list of nominating bodies as possible. Alongside the second consultation in 2009, we had published a list of nominating bodies based on the RAE2008 list and invited suggestions for additions and amendments. As well as making these additions, we added a number of 'user' organisations, including a range of public and third sector bodies, and the 100 businesses with the highest R&D spend in the UK. The resulting list comprised around 2,000 bodies, although in many cases we did not have contact details of specific individuals, or did not know if contact details were up to date (we relied on the bodies themselves informing us of changes, and had invited them to do so in previous consultations).

75. We invited the nominating bodies through a mailshot to nominate individuals to take part in the REF. Nominations were also openly invited from any other organisations with an interest in research, except for mission groups or individual HEIs. An online [nomination form](#) was developed to collect details of each nominated individual and their expertise and experience of peer review. A separate form was used for each stage of recruitment. We received around 4,000 nominations and provided the information to sub-panel chairs, for their first major task of considering the membership of each of their panels.

76. Chairs recommended the membership of each sub-panel according to the criteria for appointing panels (paragraph 55 of [REF 01.2010](#)). Their primary concern was to select members who would collectively provide the breadth of required expertise in the UOA, and in almost all cases the pool of nominees provided sufficient choice. The REF team provided guidance to chairs on the anticipated number of members, partly to control the overall number of appointments, but mostly to even out disparities in previous exercises in terms of the ratio of panel members to the volume of work per sub-panel. We provided a sliding scale of expected member numbers, based on the volume of FTE submitted to the 2008 RAE in each UOA. Most chairs recommended panels near to the upper end of the sliding scale. A few chairs felt it was challenging to achieve the full breadth of expertise required, within these size limits. In discussing the recommended membership with the main panels and the REF manager, some chairs with particularly diverse UOA remits, recommended additional members. After these discussions the recommended membership of each panel was put forward to the funding bodies for approval.

77. Overall this process succeeded in recruiting panels with the required breadth of expertise – with the intention of recruiting additional assessors later on to assist with the assessment – while also reducing disparities in panel sizes. Approximately 800 members were appointed during 2010 and the membership was published. We received a small number of complaints from particular discipline groups about gaps or underrepresentation of their sub-discipline. These continued for some time at a low level. Mostly, the concerns related to some of the UAOs that had combined several RAE UOAs and hence had a particularly diverse remit (such as UOA3 and UOA 28). These concerns were largely addressed through the later appointment of additional assessors, but a small number of members were also added where deemed appropriate.

78. Once the membership was approved, formal email invitations to members were sent from the REF manager. They included an equal opportunities monitoring form, login information for the panel members' website (PMW), and set out the panel member confidentiality arrangements as a condition of membership.

79. This initial recruitment of members was completed within six months. Nominations had been invited in July 2010 and we were able to arrange the first meetings of the main panels to take place in February 2011, followed by sub-panel meetings in March 2011, to input into the guidance and criteria.

80. In appointing panels the funding bodies had aimed to ensure that the overall body of members reflected the diversity of the research community. Nominating bodies and sub-panel chairs had been provided with guidance on equality and diversity issues. However, we did not collect any demographic information about candidates at the point of nomination. When selecting members, the chairs had limited knowledge – and avoided making assumptions – about candidates' personal characteristics. After appointing the members we collected demographic information from them, but were unable to assess how far they were representative of the pool of nominees.

81. With advice from the REF Equality and Diversity Advisory Group (EDAG), we analysed and published a [report](#) on how far the characteristics of the panel members represented the wider academic population. The analysis showed that some but limited progress had been made

in panel diversity, compared to the RAE. For example, the percentage of members who were female increased from 27 per cent in RAE to 33 per cent in REF, but remained considerably lower than the permanent academic population (40 per cent).

82. With advice from EDAG, some recommendations were made about how we could try to improve diversity on the panels through the second round of recruitment (for assessors), for example by specifically welcoming any further nominations of people from underrepresented groups. In the event that made little difference and EDAP advised that for the next exercise, 'the funding bodies should identify ways of more effectively mainstreaming equality and diversity considerations amongst all participants, at all stages of the recruitment process'. If continuing with a nominations process, it will be particularly important to engage with the nominating bodies to ensure effective encouragement of a representative pool of nominees. The funding bodies might consider collecting demographic information at the nominations stage in the next exercise.

Criteria and guidance

4.1. Guidance on submissions

Key points

- The guidance on submissions was published to schedule in July 2011, more than two years in advance of the submissions deadline.
- Our main efforts in developing the guidance were to:
 - Provide guidelines on impact that were clear and appropriate, for this first full-scale assessment of impact
 - Significantly strengthen the equality-related guidelines
 - Make refinements to other areas of the guidance, particularly to reduce burden on institutions where possible.
- Positive feedback was received from institutions on the quality and clarity of the guidance. As expected, we received a range of queries about various aspects of the guidance. Many related to staff circumstances, and fewer than expected related to impact (fewer than 10 per cent of queries).
- The more structured requirements for environment (using a template) were generally well received, although during the assessment panels found this part of the assessment was less evidence based than outputs and impact.

Recommendations

- Further refinements to the criteria for staff eligibility could potentially include: a more precise definition of the criteria for research assistants' eligibility; reconsideration of the eligibility of staff who hold fractional contracts but have little apparent connection to the submitting HEI; and consideration of whether there continues to be a material benefit to including category C staff in the assessment.
- As impact was assessed for the first time, the guidance on this element should be thoroughly reviewed.
- The assessment of environment could be developed further by identifying a number of specific forms of factual evidence to be included in submissions.

83. Following the publication of the [initial decisions](#) which set out the key parameters for the assessment, and the subsequent decisions on impact, the REF team began to develop detailed guidelines for institutions. From the outset we planned to, and did, publish a package of guidelines as follows:

- a. 'Guidance on submissions': This provided an overview of the assessment process and detailed guidance for HEIs in making their submissions. This included guidance on codes of practice on the fair and transparent selection of staff; the eligibility criteria for what could be submitted (in terms of staff, outputs, impacts and so on); and guidance on the data and information that HEIs would need to provide in submissions.
- b. 'Panel criteria': Developed by the main panels, this set out how the panels would assess the submissions. It also included further guidance on the textual information required in submissions, and in some areas this varied across the main panels.
- c. Supplementary procedural and technical guidance: for example relating to the data requirements, submission system and provision of research outputs.
- d. FAQs: to clarify the guidance in relation to any frequently asked questions.

84. The development of the guidance on submissions was carried out by the REF team and overseen by the REF steering group. We consulted the panels (particularly the main panels) through their early 2011 meetings, and worked closely with an expert group to develop the guidelines relating to the equality aspects.

85. For those aspects of the assessment that were broadly similar to the RAE, our starting point was to review the RAE guidance and aim to make enhancements, for example by providing further clarity, simplification, or removing unnecessary data requirements to reduce burden on institutions. Impact was entirely new and developing this aspect of the guidance involved substantial revision of the guidelines used in the impact pilot, taking on board the recommendations from the pilot panels, and consulting the REF panels. Significant attention was given to overhauling the guidance relating to equality and diversity and an expert group was convened to assist with this (discussed in section 4.3).

Staff eligibility

86. Following discussions with institutions and the panels we simplified the eligibility criteria for staff in several ways:

- For Category A (Cat A) staff, those on teaching only contracts were no longer eligible, for two reasons: as a matter of principle staff attracting public funds for research should be employed to carry out research; and to make more meaningful comparisons of selected versus eligible staff (using HESA staff data).
- Category C (Cat C) staff were defined in a much more restricted way, to avoid the difficulties experienced in the RAE with interpreting the criteria for numerous borderline cases.
- Category B and D staff were no longer eligible, as the way that income and postgraduate research (PGR) data was collected in REF made these categories redundant.

87. Overall, the refinements to the eligibility criteria for staff were beneficial. HEIs were able to apply the criteria consistently, as indicated by the very low rate of error we found when we carried out sample-based audits of staff eligibility. The changes to the Cat A staff criteria enabled HESA to publish contextual data providing an approximation of the number of eligible staff at each HEI, in each UOA.

88. Nevertheless, as staff selection is a sensitive issue, and there are many differing individual cases, some aspects of the staff eligibility criteria generated queries as HEIs prepared their submissions. (This is in addition to the numerous queries relating individual staff circumstances, discussed in section 5.7.) Some issues relating to staff eligibility were also raised subsequently by panels during the assessment. The main issues were:

- Research assistants: The criteria whereby a research assistant could be eligible for submission included an element of interpretation or judgement, and cases were queried during both the submission and assessment phases. It is possible that some institutions applied different interpretations, although no specific evidence of this emerged. For the next REF, the funding bodies should consider whether a more precise definition is possible.
- Teaching only contracts: A small number of institutions (particularly in the arts sector) initially argued against the ineligibility of staff on teaching only contracts. (Some concerns were also raised about periods of 'teaching only' work not counting for the purpose of outputs reductions.)
- Hourly paid staff: Some institutions had difficulties in calculating the FTE for such staff, and we had to clarify whether it should be done on the basis of the entire life of the contract, or just the current academic year.
- Clinical academics: There were some complexities around knock-for-knock arrangements, in that clinical academics employed by the NHS could be submitted as Cat C staff with a reduction of two outputs, whereas clinical academics employed by an HEI could be returned as Cat A staff but with no reduction in outputs for their clinical duties.
- Overseas staff on fractional contracts: During the course of the assessment several panels noted their unease about staff who had joined the HEI recently on a fractional contract, but had little apparent connection with the submitting HEI (typically holding substantive academic positions overseas). They accepted that such staff were eligible according to the criteria but suggested that in future such staff should not be eligible.
- Staff based at an overseas campus: The guidance on staff based overseas was intended to refer to those based at an overseas campus of the submitting HEI. This was not sufficiently clear in the original guidance (though was addressed through subsequent FAQs). In the event, the large majority of justifications provided for staff based overseas were unnecessary: most were either for staff carrying out fieldwork overseas on the census date, or had fractional appointments in the UK but more substantive posts overseas.

89. The tighter definition of Cat C staff was intended largely to enable the continuing submission of staff employed by embedded research council units, and NHS employees with specific research duties, but exclude those with looser relationships with the submitting HEI. This change succeeded and the difficulties experienced in the RAE of dealing with many borderline

Cat C cases were avoided. In the event, however, a total of only 304 Cat C staff were submitted, down from 2,284 in the RAE and comprising 0.5 per cent of submitted staff. This is no doubt because, in addition to the more restrictive criteria, many research council units had been integrated into HEIs (making the staff eligible as Cat A). In light of this the funding bodies might consider whether there continues to be a material benefit to including Cat C staff in future.

Outputs

90. The guidance on output eligibility was based largely on that for the RAE, with the data requirements simplified where possible (for example, instead of requiring a date of publication, this was simplified to year of publication). Nevertheless, given the large volume of outputs involved in the assessment, numerous queries were raised about eligibility issues, and some issues were identified through audit. The main issues related to:

- Timing of 'publication': After publication of the guidance a few of the REF sub-panels and some subject associations (in disciplines that typically publish 'working papers' prior to the final journal articles) raised concerns that some outputs at the cusp of the RAE and REF publication periods may not be eligible in either exercise. The guidance was amended to allow outputs to be submitted to the REF that had been 'pre-published' in 2007 and not submitted to the RAE, and subsequently published in final form during the REF period.
- Non-text outputs: Queries generally related to what exactly constituted the output to be submitted. Ambiguities arose where, for example, a series of exhibitions or performances were held (and may have evolved) over a period of time; a composition or script was published and subsequently performed with differing productions; or a website with frequently changing content.

Impact

91. As this was the first full assessment of impact, an important task was to develop guidelines that would provide sufficient clarity while at the same time being suitably broad and generic to encompass the very wide range of possible impacts, and possible links between research and impact. In developing the guidelines we sought to establish some clear boundaries to exclude only those forms of impact that were out of scope, and to establish some clear principles – rather than particular models – for linking the underpinning research and the claimed impacts. We spelled out that these principles could be satisfied by indirect as well as direct links, with particular examples given for some of the main anticipated areas that HEIs might be uncertain about – especially impact through expert advice or public engagement that drew upon, but was not directly linked to, the academics' research.

92. Following the pilot we revised the case study template, in line with the recommendations of the pilot panels, and developed other aspects of the guidance such as the formula for determining the number of case studies per submission, page limits for the impact template, and so on. Draft guidance was shared and discussed with the REF main panels at their criteria phase meetings.

93. After publishing the guidance, comparatively few queries from HEIs related to impact (around 10 per cent of queries). However, it was apparent that many HEIs made significant

efforts to raise awareness and provide training to academic staff on the impact requirements, and this was generally done by central staff who were well informed about the criteria. Although institutions were undoubtedly uncertain about the grades that their case studies would receive, they succeeded in submitting case studies that were in scope and satisfied the threshold criteria. Only 0.7 per cent of impact submissions were deemed to be unclassified by the panels, although there were a number of case studies that did not clearly provide some of the basic information requested (typically, about where the researchers were employed at the time of the research). These were resolved through audit queries and we suggest that in future the case study template could include some specific fields to collect these basic details.

94. From the perspective of the REF team, the guidance on impact appeared to be workable for institutions, allowing them to submit a very broad range of impacts from all disciplines, and the submissions enabled the panels to make judgements they were confident in. Inevitably, as this was the first exercise of this kind, the funding bodies will want to review all aspects of the impact assessment and make improvements for the next exercise.

95. Although impact in REF 2014 was successful as an assessment exercise, it appeared to us that the costs to HEIs were considerable. There were also some indications of a cottage industry forming to assist institutions in preparing case studies. Some HEIs created new posts to help prepare case studies, and several consultants began offering their services to HEIs. From their queries to the REF team our impression was that the consultants did not have any particular insights to offer over and above what institutions could gain by understanding the published documents. The panels appeared to regard the use of professional writers – where this was apparent – to have had mixed results, sometimes improving but at others detracting from the case being made.

Environment

96. Our main aim in developing the environment guidance was to produce a structured template that would elicit from HEIs the information required by panels to make their judgements. The relevant headings for the template were readily identified through discussions with the panels. The four main panels were then invited to set out in more detail (in the 'panel criteria') the particular kinds of information they wanted under each of the generic headings, and to indicate the relative weightings they would attach to each section of the template.

97. Institutions appeared to find the guidance relatively straightforward and we received few queries about this aspect of the guidance. However, during the assessment panels found that this part of the assessment (both the environment narrative, and the quantitative data discussed in section 5.6) to be less evidence based than outputs and impact. The assessment of environment could be developed further in future by identifying a number of specific forms of factual evidence to be included in submissions in a standardised way.

4.2. Panel criteria

Key points

- In parallel to development of the 'guidance on submissions' the four main panels developed and consulted on the 'panel criteria'. We, the panels and institutions all supported the aim of developing criteria that were as consistent as far as possible across the exercise.
- The main panels were effective in taking ownership of the criteria, and – especially after receiving feedback through the consultation exercise – ensured that variations in the criteria between UOAs occurred only where there was a strong discipline-specific justification.
- Additional efforts to engage with users on the impact criteria through a series of workshops were useful, both to validate the criteria and develop contacts for future involvement.

Recommendations

- Having succeeded in the 2014 REF in producing a single set of criteria for each main panel, further improvement could potentially be made in a future exercise by developing, consulting on and publishing the guidance on submissions and the panel criteria at the same time, as a single document. This could be developed by the REF team and the main panels, in consultation with subject communities. The appointment of sub-panels could potentially be deferred until a later stage.

98. The panels met in early 2011 to develop their draft criteria, for consultation in mid-2011. The panel criteria were intended to explain how the panels would operate, provide descriptions of how the star levels would be applied, and provide guidance to institutions on what the panels required or expected to be covered in the various textual elements of submissions. This included, for example, guidance on any 'additional information' that was required about outputs, types of evidence that would be suitable in case studies, and the kinds of information sought in the environment template.

Consistency

99. In the RAE, each sub-panel drafted its criteria with some level of oversight by the 15 main panels. The resulting criteria varied in many ways across the sub-panels. For the REF, a key aim was to establish as much consistency as possible, with any differences between panels justified for discipline reasons.

100. To this end we organised the criteria phase meetings in a conference format. All the sub-panels in each main panel met at the same time at the same location, with a mixture of plenary and break-out sessions for the individual sub-panels to feed into the main panel deliberations. Though organisationally challenging, this approach proved effective.

101. We provided guidance to the panels on producing their criteria, establishing the main panels' ownership of and responsibility for producing the criteria, with input from the sub-panels

(rather than sub-panels producing the criteria, with oversight from the main panels). We also established the principle – with which the main panels agreed – that the criteria ought to be consistent across the sub-panels, varying only where there was a discipline-specific justification. This principle was important both to reduce unnecessary complexity for institutions in preparing submissions, and to make the assessment outcomes fairer.

102. The guidance invited the main panels to develop the criteria and suggested a number of issues that ought to be consistent across each main panel, and those where there might be flexibility for variation across the sub-panels. As the panel discussions and the wider consultation proceeded, there was a push to even greater consistency than we had envisaged in this guidance. The consultation exercise strongly supported the principle of consistency and challenged some of the proposed variations in the draft criteria. The main panels increasingly upheld the view that consistency was essential, and variations should occur only where there was a strong and discipline-specific justification. In the final criteria, very few variations between sub-panels remained.

Impact

103. As this was the first assessment of impact – both for institutions and for the panels – particular attention was given to assisting institutions in preparing their impact submissions. The main panels were invited to indicate the types of impact and types of evidence they would anticipate in case studies, partly to indicate that impacts would be assessed by each sub-panel in ways that were appropriate to the discipline(s) concerned. Some panels had reservations about providing lists of possible impacts, fearing they would discourage institutions from submitting anything that was not included. For this reason Main Panel C initially declined to produce such lists but when this was badly received by institutions in the consultation, they did so. It was made clear by all main panels that these lists were provided as suggestions only and were not exhaustive.

104. Some users had been appointed as members to each main and each sub-panel from the outset, to assist with developing the impact criteria (with the intention of appointing more impact assessors at a later stage). We also wanted to gain feedback from users on the draft criteria but did not expect many to respond formally to the consultation document in the same way as academic associations or HEIs. We therefore arranged a series of workshops – one for each main panel – to discuss the criteria with users from a range of backgrounds. This exercise revealed general support for the inclusion of impact, and support for the very broad definition of impacted and openness to a very wide range of potential sources of evidence. The workshops also countered some of the concerns that institutions had raised about the willingness of users to provide corroboration. They were also helpful in raising awareness amongst some users and developing contacts that would prove useful when recruiting impact assessors later on.

Guidance and criteria

105. As we developed the guidance on submissions and the panel criteria, we made decisions about what requirements and guidelines should be standard across the whole exercise – in which case they belonged in the ‘guidance on submissions’ – and what requirements could vary

across the main and/or sub-panels, in which case the panels were invited to develop these as part of the 'panel criteria'. Overall this approach was effective in establishing a clear set of generic criteria and requirements that were standardised across all UOAs, with supplementary requirements that varied in appropriate ways between the main panels. We received overwhelmingly positive feedback from institutions about the quality and clarity of the published guidance and criteria.

106. Nevertheless, there were some drawbacks to the separate processes and publication of the submissions guidance and the panel criteria:

- HEIs and panels had to refer to two sets of overlapping documents to work out exactly what the requirements and guidelines were.
- The consultation on the panel criteria led to some changes to the guidance on submissions (particularly the output reductions for maternity leave, and the eligibility of certain types of outputs published at the cusp of the RAE and REF publication periods).
- It was not straightforward to decide at the outset what should be standardised and what might vary across panels. In areas this was an iterative process (for example, we had initially invited each main panel to develop guidance on double-weighted outputs but after the consultation, this was essentially standardised across the panels).

107. Given the substantial progress made in developing consistent criteria, it might be preferable in a future exercise, if practicable, to develop, consult on and publish the guidance on submissions and the panel criteria at the same time, as a single document. This could still include some variations across the main panels, but would potentially provide a more coherent set of guidelines that are more straightforward for institutions to deal with, and for panels to apply during the assessment. We suggest the funding bodies should consider the following approach to developing the guidance and criteria in future:

- At the outset, the funding bodies could establish the key parameters for assessment
- The main panels (including the sub-panel chairs and possibly deputies) could be appointed to work with the REF team to develop a single draft document setting out both the submissions guidance and the assessment criteria
- During the drafting stage the sub-panel chairs (and perhaps deputies) could consult their subject communities informally through subject associations and town meetings, to gain wider input from the subject communities (rather than holding sub-panel meetings at this stage)
- A formal consultation could then follow, including with subject associations, with resulting changes and refinements agreed by the main panels and the REF team.

108. This approach could involve deferring the appointment of sub-panels until a later stage, at which point there could be better information available about the likely range of submissions and hence expertise requirements. This approach could result in a single and more coherent document for use by both institutions and the panels, and with effective consultation on both elements. The significant costs of the criteria phase meetings could also be reduced. The main drawback of the suggested approach would be the diminished sense of involvement of the sub-panels in developing the criteria, if they are to be recruited for the assessment phase only.

4.3. Equality and diversity guidance

Key points

- Once it became clear that the REF would (like the RAE) involve selecting staff for inclusion, significant attention was given to addressing equality and diversity issues.
- A review of the RAE equality and diversity measures was invaluable in identifying the main areas that could be strengthened, namely:
 - Ensuring that all HEIs had appropriate codes of practice in place
 - Developing clear and consistent arrangements to enable staff with individual circumstances to be submitted with fewer than four outputs.
- An expert group (EDAG) was established at an early stage and was invaluable in advising us on the guidance and procedures, which succeeded in advancing the equality and diversity aspects of the exercise.
- The approach to staff circumstances was widely welcomed as a substantial improvement, although a change in the arrangements for maternity leave was required.

Recommendations

- The arrangements for individual staff circumstances succeeded but did generate high volumes of queries and audits. They should be refined or simplified in future to make the exercise more efficient.
- Given the nature of the issues, we recommend that an expert group should again be convened at an early stage to advise on the development of any equality-related guidance.

109. Once it became clear that the REF would, like the RAE, involve institutions selecting staff for inclusion in submissions, we devoted significant attention to addressing equalities issues. The review of the 2008 RAE equalities measures commissioned by the Equality Challenge Unit (ECU) was invaluable in identifying a number of areas where the measures could be strengthened for the REF.

110. In 2010 we established an expert advisory group (EDAG) to advise us on how to take forward the issues identified in the review, and to assist us in developing the relevant guidance. The key areas we worked with EDAG on were:

- a. Codes of practice on the fair and transparent selection of staff: The RAE review had highlighted that these were somewhat patchy and we discussed with EDAG how to ensure that all institutions had appropriate codes in place, well in advance of the submission deadline. On advice from EDAG we decided that institutions should be required to submit a satisfactory code as a pre-condition for making submissions to the REF.
- b. Individual staff circumstances: In discussion with EDAG we decided to standardise the arrangements for output reductions across all UOAs (in the RAE the arrangements had been developed by the panels and were somewhat variable). With EDAG we categorised the types of circumstance that would be eligible for reductions and developed guidance that aimed to provide as much certainty as possible about the applicable reductions for

each type of circumstance. 'Tariffs' were produced for some types of circumstance; others were inherently complex and we decided to deal with these through a centralised process rather than the sub-panels. In developing these arrangements there were some challenges in trying to ensure equity between different types of circumstance, and specifying how to combine different types affecting an individual.

111. The EDAG was invaluable in helping us prepare the guidance on these areas. The sector widely welcomed the approach, although during the consultation on the panel criteria, widespread objections were raised about the way the arrangements for maternity leave, which indicated that only periods of leave of at least 12 months would automatically qualify for an output reduction. This led to a quick decision by the funding bodies to change the arrangements so that any period of maternity leave, regardless of duration, would qualify. This change was universally welcomed.

112. On the advice of EDAG, the funding bodies decided to appoint a successor expert group, the Equality and Diversity Advisory Panel (EDAP), to assist with implementing these measures. EDAP reviewed the codes of practice and the complex individual staff circumstances, as discussed in sections 5.1 and 8.6, and in the [EDAP final report](#).

113. At the end of the exercise EDAP concluded that the strengthened equality measures in the REF were successful. Institutions adopted effective codes of practice and more than 16,000 staff were submitted with individual circumstances. Importantly, [analysis](#) of the results showed that research outputs by these staff had been judged by the panels as being of equal quality to outputs by all staff.

114. From an operational perspective, the individual staff circumstance arrangements did involve dealing with a large number of queries during the submissions phase (see section 5.7), and a significant number of audits during the assessment phase (see sections 6.2 and 8.6). In a future exercise, these important procedures could be made more efficient by providing further clarity on the points more frequently queried during REF 2014, the use of templates and, if possible, simplifying the arrangements.

115. Also, given the nature of the issues, we recommend that an expert group should again be convened at an early stage to advise on the development of any equality-related guidance.

5. Submissions

5.1. Codes of practice review

Key points

- As a pre-condition for making submissions to the REF, each HEI was required to submit a code of practice on staff selection, that adhered to the guidance.
- An Equality and Diversity Advisory Panel (EDAP) was appointed from amongst the REF panels and sector practitioners, to assist with implementing the equalities measures. Its first main task was to review all HEI codes of practice during 2012.
- The review of codes of practice prior to submissions was a beneficial exercise; it led to a number of improvements to HEIs' codes and provided assurance that all HEIs had an appropriate code in place.
- Equality impact assessments by HEIs were also much improved compared to the 2008 RAE, but were still variable and there is room for clearer guidance and further improvement in future.

Recommendation

- If codes of practice are reviewed in future it would be preferable to simplify the process by not allowing two submission deadlines.

116. As a pre-condition for making submissions to the REF, institutions were required to draw up and apply a code of practice on the fair and transparent selection of staff. This was intended to encourage the inclusion of all eligible staff who had conducted excellent research, and to promote equality, comply with legislation and avoid discrimination in selection processes.

117. EDAP was established from 2012 until 2014, to provide advice to the REF team, REF panels and the UK funding bodies on the implementation of the REF equality and diversity measures. Its first major task was to review all the codes and advise the funding bodies on whether they adhered to the guidance in relation to principles underpinning staff selection; communication to staff; transparency about decision making committees; training; staff circumstances arrangements; and the appeals process.

118. Some institutions were keen to have their codes approved earlier than the original deadline for submitting them (July 2012), so we set an early optional deadline (April 2012) and EDAP reviewed the codes in two batches. 58 institutions submitted codes of practice at the earlier deadline and a further 101 were submitted by the final deadline. This review in two rounds was completed successfully but it did involve some challenges in ensuring that judgments were fully consistent across the two rounds. If a similar exercise is carried out in future it would be preferable to simplify the review by not allowing two submission deadlines.

119. Codes which were judged not to adhere to the guidance in any particular respect were reported to the funding bodies, who requested revisions from the HEIs. Although a large number (nearly half) fell short of meeting some aspect of the guidance, many of these required only minor revisions. Some of the codes required more substantial revision. All institutions whose

codes required revision did so by January 2013. EDAP produced a [report](#) on its review of the codes, highlighting some common issues and points of best practice. The review of codes was a beneficial exercise; it led to a number of improvements to HEIs' codes and provided assurance that all HEIs had a satisfactory code in place prior to submissions.

120. After approval by the funding bodies, some HEIs asked whether further changes could be made to their codes. A process was put in place to allow this, with the funding bodies checking that any changes continued to adhere to the guidance. In general such revisions were adjustments to the timetables of the selection process. On making submissions, the head of institution was required to confirm adherence to their code. Final versions were collected from all HEIs and the REF team [published](#) all HEIs' codes in January 2015, after the conclusion of the assessment.

121. EDAP also wanted to consider evidence of how the codes had been applied by institutions. It asked whether institutions' Equality Impact Assessments (EIAs) could be made available and the funding bodies agreed to collect these from HEIs at the end of the submission phase. EDAP reviewed these at sector level, rather than evaluating individual HEIs' EIAs. EDAP's [final report](#) sets out its conclusions about EIAs, which need not be repeated in detail here. Essentially, EIAs had improved compared to the 2008 RAE, but were still variable and there is room for clearer guidance and further improvement in future.

5.2. Submission data requirements

Key points

- In addition to the 'guidance on submissions' the REF team needed to specify in detail the precise items of data required for each aspect of the submissions.
- With a view to reducing burden on institutions, we gave detailed consideration to aligning the REF data requirements with other data returns. In the event, some but limited opportunities for data alignment were found.
- This was done at an early stage so that work could begin to develop the submission system. In the event, it would have been beneficial to have given more attention at an early stage to specifying the data requirements more precisely, as in practice several aspects of the requirements and definitions were refined in parallel to the development of the submission system itself.

Recommendation

- The funding bodies should seek to further enhance data alignment in the next REF. This might be aided by changes to HESA cost centres, which have become more aligned with REF UOAs.

122. At an early stage – before the guidance on submissions had been developed – we initiated a 12-month project led by a secondee to HEFCE from Imperial College London, to scope the REF data requirements and identify possible options for data alignment between REF and other established mechanisms for collecting or reporting research data. The secondee held discussions with a range of other funders, agencies and institutions, to try to establish whether any common data requirements either existed or could be agreed.

123. In the event the project identified relatively limited opportunities that were practicable to pursue. They comprised:

- Aligning the definition of staff eligibility with definitions used in the HESA staff record.
- Aligning the definitions of eligible research income and PGR degrees awarded with HESA data collection; and using these data returned to HESA as far as possible to validate institutions' REF submissions.

124. The project also sought to specify the data requirements for the REF submission system that would need to be developed. This included reviewing the 2008 RAE data specification and trying to simplify the requirements or eliminating any items of data that were not expected to be of material use to the panels. Some aspects of this were inherently challenging to complete before the guidance on submissions and panel criteria were finalised, but early progress was required so that decisions could be made about how and when to start developing systems to collect submissions data, and for that work to commence in time to deliver the system.

125. The project considered the implications of the emerging data requirements for the development of REF data collection software. Consideration was given to:

- Whether the RAE system could be updated/adapted for use in REF; or whether a new system would need to be developed.
- Whether development work should be carried out in-house by HEFCE staff, or externally commissioned.

126. In the event, although a number of data requirements were similar to the RAE, the decision was taken to develop a new system for the REF, in large part because the development software used in the RAE system was no longer supported, and also because there were a number of different or new data requirements. These decisions were taken and work began in a timely manner, so that the submission system was successfully delivered on schedule. However, it would have been beneficial to have given more attention at the start of this development to fully specifying the data requirements more precisely, as in practice several aspects of the requirements and definitions were refined in parallel to the development of the submission system itself.

5.3. Submission system

Key points

- The submission system was delivered successfully, on schedule and to a high standard. HEIs' feedback was overwhelmingly positive.
- Specification of the requirements for the submission system, and project management of the systems development, were significant tasks undertaken by the REF team, working collaboratively with IT colleagues in HEFCE.
- Thorough testing of the system was successfully completed, although it would have been beneficial to have started testing earlier in the development of the system.
- A user group from institutions provided invaluable oversight of the development of the submission system, helping to ensure it met institutions' needs.
- The system was released initially in pilot form in late 2012. This was extremely helpful, allowing remaining bugs to be identified and user input to be provided.

Recommendation

- In developing the submission system, significant effort was devoted to the web interface for the submission system, with less attention devoted to the web service. In future, consideration should be given to how HEIs will be using the system to inform the balance of effort between the web interface and web service.

127. Developing the IT system by which institutions made their submissions to the REF was a major part of the REF project. Initial scoping of the system requirements was informed by a number of sources including several workshops conducted in November 2009 and a review of the RAE 2008 data collection system by PricewaterhouseCoopers, and an internal HEFCE Data Working Group (DWG). The group comprised relevant members of the REF team, IT services and the Analytical Services Group. Many of the group had been directly involved with the RAE 2008 system and were able to provide invaluable input.

128. An external Data Collection Steering Group (DCSG) was established to provide oversight of the development of the system. Its terms of reference can be summarised as follows:

- To provide direction in developing the REF submission system and monitor progress against the project plan.
- To agree a robust testing strategy and a strategy for delivering system training and guidance materials.
- To monitor and advise on key project risks and issues relating to the implementation of the submission system and its related systems.

129. DCSG included members from a number of HEIs invited to represent the different types of HEIs that would be submitting to the REF. These individuals included both IT and research support staff. The input they provided was very important in ensuring the project met institutions' needs and was planned, delivered and tested effectively. The group also included

representatives of the four UK HE funding bodies, which provided a direct reporting link to the REF steering group. DCSG met nine times during the REF project.

130. Following the initial scoping of the system requirements, a detailed user requirements document (URD) was developed in stages and considered by both DWG and DCSG. This was a significant task for members of the REF team, in translating the guidance on submissions into a detailed specification for the submission system. The IT team at HEFCE began development of the software as the first parts of the URD were developed. Several members of HEFCE's IT team devoted significant time to the development of the system, including a database manager, a web developer and a senior systems analyst. In addition a full-time software developer was recruited to lead the development.

131. Significant effort was devoted to specifying and developing the web interface for the submission system. A web service was also developed, but at a later stage with less time for testing and input from the REF policy team. For future exercises it may be beneficial to consider how the majority of HEIs will use the system when planning the resource allocated to its development. If the web interface were to be used by only a small number of HEIs a considerable amount of effort could be saved by devoting less time to the development of the interface, allowing more resource to be dedicated to the web service.

132. During the course of the software development, a testing strategy was developed, including a full schedule of scripted and unscripted in-house testing; and user testing by members of DCSG. These were completed during 2012, although it would have been beneficial to have started testing earlier in the development of the system, with testing scripts developed as the user requirements were specified.

133. The REF submission system was released in pilot form to institutions in September 2012. All HEIs were encouraged to register and use the system. The REF briefing events that had been held in May 2012 had included a walk-through of key parts of the submission system, to familiarise HEIs with the look and approach of the system in advance of its launch.

134. The pilot system was open for three months; towards the end of this period the REF team surveyed all REF contacts to seek feedback on their experience of using the submission system pilot and intentions for future use. The overwhelming majority indicated that the system was fit for purpose, and some useful feedback for potential improvements was also received. The final system went live in January 2013.

135. The submission system was load tested, prior to launch, to ensure that high levels of system performance could be delivered. The RAE system performance was used as a minimum requirement. Due to the length of time required for importing and exporting some of the larger files, a 'system queue' was implemented to ensure that other users were not adversely affected.

136. A robust disaster recovery mechanism was required for the REF system to minimise the risk of loss of data and system functionality for users. The disaster recovery plan included:

- Database mirroring with back-up of changes as they occurred stored online.
- Back-up of data changes taken every ten minutes (stored off-site, but online).

- Back-up of data taken every night (stored off-site and offline).
- Off-site back-up facilities (once a week).

137. A technical author was recruited to produce a comprehensive user guide for the submission system. The guide was built in to the submission system, cross-referring to the 'guidance on submissions' and 'panel criteria' documents where appropriate, and was also available as a stand-alone document available on the REF website. The user guide was updated during the submission period to reflect any changes to the system.

138. Telephone and email user support for the submission system was provided initially by one dedicated user-support team member. User support was expanded to two FTE as the queries increased for the later stages of the submission phase, and extended cover was provided, from 8.00 to 18.00. In addition a range of supplementary information was provided on the [submission system section](#) of the REF website, and [FAQs](#) were produced based on HEI queries.

5.4. Citation data

Key points

- Eleven of the 36 sub-panels chose to make use of citation data to inform their review of outputs, whilst recognising the limitations of the data.
- Citation data was sourced from the Scopus database and the submission system enabled HEIs to match their outputs to records in this database.
- Panels and HEIs were also provided with contextual data to assist them in interpreting citation counts, given that citations depend on the field of research and a publication's age.

139. The bibliometrics pilot exercise and subsequent consultation concluded that citation indicators were not sufficiently robust to be used as a primary indicator of quality, but that they may be useful in informing expert review in some UOAs. In developing the panel criteria, sub-panels were invited to decide whether they wished to use citation information to inform their review of outputs. Eleven of the 36 sub-panels chose to do so (sub-panels 1-9, 11 and 18). They recognised the limitations of such data and used citations as a minor indicator (as described in the 'panel criteria').

140. Through a procurement exercise, the REF team selected Elsevier to provide the citation data, to both institutions and the relevant panels. The submission system included functionality for HEIs to 'match' outputs in the relevant UOAs and retrieve the citation counts (journal articles and conference proceedings only) from the Scopus database. The process is explained [here](#). The REF team liaised regularly with Elsevier to ensure the system was operating smoothly and to discuss user support queries received by Elsevier.

141. Elsevier took a final 'snapshot' of citation counts for all matched outputs shortly after the submission deadline. These citation counts were provided to the relevant sub-panels for use in the assessment. The citation counts were not updated during the assessment phase of the REF, to avoid potential inconsistencies.

142. In addition Elsevier produced contextual data which was provided to the sub-panels to assist them in interpreting citation counts, given that citations depend partly on the field of research and a publication's age. To help institutions understand the contextual data when preparing submissions, the REF team provided analogous data in early 2013.

143. Institutions were able to contact Elsevier directly to raise queries about matches or the returned citation counts. Elsevier provided dedicated support for institutions throughout the submission phase. This support was not without teething troubles in the early months of the submission phase. However, continued dialogue between the REF team and Elsevier enabled appropriate and timely support to be provided.

144. Sub-panel 11 (Computer Science and Informatics) had intended to make use of Google Scholar data in addition to the Scopus citation data provided by the REF team. Unfortunately, following discussions with Google Scholar it was not possible to agree a process for bulk access to their citation information, and the sub-panel did not use Google Scholar to inform their assessment of outputs.

5.5. Multiple submissions and classified case studies

Key points

- In restricted circumstances, institutions were able to request that they could make multiple submissions in a particular UOA. We offered HEIs several deadlines to make these requests, to suit their internal decision making timetables.
- In total, 129 multiple submissions were made, predominantly in UOAs that had 'merged' since the RAE.
- Institutions could also request in advance that they submit impact case studies with classified content. A modest number of requests were made and were successfully accommodated.

Recommendation

- A disproportionate amount of effort was required to make arrangements for the classified case studies. It will be helpful in future to identify and appoint some panellists at the outset with high level clearance in the most likely areas of required expertise.

Multiple submission requests

145. HEIs were able to request permission from the REF manager to make multiple submissions in a UOA. Requests were approved where they fulfilled one of the following criteria:

- Where an institution involved in a joint submission wishes to make an additional individual submission in the same UOA.
- For SP28, where one submission is in Celtic Studies and the other is in Modern Languages and Linguistics.
- Where HEIs merged after 1 July 2011 and they seek permission to submit separately to all UOAs in which they wish to submit.

- Where a sub-panel considered there is a case for multiple submissions, given the nature of the disciplines covered.

146. HEIs were [invited](#) to make requests in February 2012. To suit their internal decision making processes, they were given a choice of three deadlines and were required to submit all requests at once. A template was provided to make requests. A supporting explanation was required for all cases except those returned to SP28 citing Celtic Studies. All requests were submitted via the REF extranet for security purposes.

147. Each multiple submission request was reviewed by a REF team member and the REF manager, and a recommended outcome was then proposed to the associated sub-panel chair for agreement. The decisions were made by the REF manager on the advice of sub-panel chairs. The decisions were communicated to the HEI within six weeks of each deadline, including proposed submission names where multiple submissions were agreed. Submission names were not required where only a single submission to a UOA was made.

148. A total of 76 multiple submission requests were received for a total of 158 submissions, across 43 HEIs. Only one request was rejected. Approval was sometimes dependent on the HEI providing further information. These figures compare with 49 multiple submission requests received in the RAE, of which 39 were accepted.

149. Requests were made across 13 UOAs, with 68 of the requests made where the nature of the disciplines covered were considered to warrant multiple submissions. These were mainly requested in the 10 UOAs that stated in the 'panel criteria' that they might expect requests. UOA 17 received the most requests (18) followed by UOA3 (13 requests). Only 5 requests were made to submit multiple submissions to SP28 citing Celtic Studies. Although requests were approved for 158 submissions, in the event HEIs made 129 multiple submissions in total.

Classified case study requests

150. During the development of the panel criteria, questions were raised about how case studies could be submitted in areas such as national security and defence, which would need to include some classified material in order to properly be assessed. These were expected to be important (although not common) areas of impact in some UOAs. We made special [arrangements](#) to accommodate such case studies and [invited](#) these requests at the same time as inviting multiple submission requests.

151. Requests were made for 37 case studies and were dealt with individually by the REF manager, who liaised as necessary with sub-panel chairs to identify panel members with sufficient security clearance and relevant expertise. Where such members could not be found on any of the sub-panels, plans were made to recruit additional assessors, or in one area, a specialist adviser. The process also involved frequent correspondence with the HEIs to seek clarification and update on progress. During this process a number of requests were withdrawn and a few were rejected on the grounds that security clearance was not after all required.

152. Once suitable members or assessors had been identified for each case study, arrangements were made for the case studies to be delivered directly to them; HEFCE does not have facilities to handle such material. The institutions were also required to submit, in the normal way, a redacted version that was accessible to the sub-panel. In the end 21 classified case studies were submitted. There were two main areas they covered: intelligence/cryptography and defence/atomic weapons, and also a very small number of case studies in policing/counter-terrorism and diplomatic relations. From viewing the redacted versions and in informal discussion with some of the members who reviewed these cases, it was apparent that several of the cases included only a very small amount of classified material, which made little difference to their overall assessment.

153. Overall it was worthwhile to accommodate these case studies but they required a disproportionate amount of effort to make the necessary arrangements, especially where additional panellists needed to be recruited. It will be helpful in future to identify and appoint some panellists at the outset with high level clearance in the most likely areas of required expertise.

5.6. Research income and PGR data

Key points

- It is likely that alignment of environment data with HESA income and PGR data reduced the burden for HEIs overall, although a number of institutions made requests to amend their data limits based on HESA returns.
- Panels used the environment data to inform their assessments of the relevant sections of the research environment. However, a number of sub-panels felt the utility of the data was limited as the total number of research active staff in the submitting units was unknown.

Recommendations

- To further reduce complexity and workload in future, the funding bodies should consider fuller alignment with HESA data, without inviting or allowing requests for data adjustments.
- For the next exercise the funding bodies should consider collecting the number of eligible staff as part of the submission, to allow panels to make clearer comparisons of the environment data.
- Difficulties were experienced in collecting research council income-in-kind data. Improvements could be made in future, for example through closer collaboration with the research councils; provision of the data by the research councils directly to HEIs; and early publication of a complete list of facilities that will be included.

154. Each REF submission was required to include the number of research doctoral degrees awarded in each academic year 2008-09 to 2012-13; data on the submitted unit's external research income for each academic year 2008-09 to 2012-13; and the estimated value of Research Council facility time (income-in-kind). For submission in UOAs 1-6, submissions also included the estimated value of the equivalent income-in-kind from the health research funding bodies.

155. To minimise the burden on institutions, the definitions were aligned with those used in the HESA financial and student returns, and the income and doctoral awards were not tied to submitted staff only. The total values reported, therefore, ought to have aligned with the values reported to HESA, although HEIs would still need to map their data to REF UOAs. HEIs were required to submit the data via the REF submission system, which imposed upper limits at the institution level, based on their HESA returns.

156. To support HEIs in collating this information, the REF team provided the following to institutions via the REF extranet:

- HESA data on the numbers of research doctoral degrees awarded for academic years 2008-09 to 2011-12
- HESA research income data for academic years 2008-09 to 2011-12
- Income-in-kind data compiled by the Research Councils and collated centrally by RCUK for academic years 2008-09 to 2012-13 (data supplied to end of March for AY 2012-13).
- The relevant health research funding bodies provided the equivalent income-in-kind data directly to HEIs.

157. HEIs were required to use their own records to provide data for AY 2012-13, and these were checked against HESA data later on through the REF audit process.

158. The submission system limits for the environment data were based on the data provided to HEIs. Where institutions wished to submit doctoral degrees or research income data that exceeded the limits, they were required to request a data adjustment from the REF team. Income-in-kind data adjustments were agreed directly between the institution and the relevant Research Council or health research funding body; the REF team was then informed of the outcome.

159. The initial deadline for HESA data adjustment requests was 1 June 2013. This deadline was extended by a month following several requests from institutions asking for an extension. Institutions were asked to inform the REF team by 1 July 2013 if they required an extension and for all necessary data to be provided by 31 July 2013.

160. In total 66 requests for data adjustments were made representing an increase in 368 research doctoral degrees awarded and £16.2 million of research income. Agreeing these data adjustments with HEIs required significant resource by the REF team and HEFCE's Analytical Services Group. This may be considered unnecessary resource as data submitted to HESA should have been sufficiently accurate. We would recommend that adjustments to HESA data are not accommodated or at least not invited in a future REF. As well as saving unnecessary resource this may also encourage institutions to ensure that their HESA returns are correct.

161. To obtain the Research Council income-in-kind data, the REF team worked with RCUK who collated the data from individual Research Councils. The collated data for each HEI was then made available via the REF extranet. Some of the data was provided to the REF team by financial year and therefore required manual intervention for the data to be in an appropriate form to distribute to HEIs and to set validation limits on the submission system.

162. The list of facilities to be included in the income-in-kind data was agreed between RCUK and the REF team, following which an initial tranche of income-in-kind data was provided to HEIs in summer 2012, with a final version provided in May 2013. Although some data was released in 2012, it was not until the data release in 2013 that there was a query about a facility which had not been included. It was agreed that the facility should be included and the Research Council and the REF team had to work within a limited timeframe to provide the additional data to HEIs.

163. Provision of income-in-kind data was a resource intensive activity. In future the REF team may wish to consider if it would be more appropriate for Research Councils to provide these data directly to HEIs, as the health research funding bodies did. In addition, it may be helpful at an early stage to publish a list of facilities that will be included to avoid any last minute changes.

164. The environment data was provided to the sub-panel as part of a 'standard analysis' that included data on the numbers of staff, outputs, income and doctoral awards in each submission. It included the absolute numbers as well as 'per FTE' figures for the income and doctoral awards data, to put the environment data in context with the size of individual submissions. However, since the FTE submitted did not necessarily reflect the total size of submitted units that had generated the income and doctoral awards, several panels found the environment data was of limited usefulness as direct comparisons could not be made between submissions. Instead they used the data to look at trends over time within each submission.

5.7. Supporting institutions

Key points

- Throughout the submissions phase the REF team prioritised effective communications with HEIs using a number of channels, and strived to make communications clear, timely, consistent and helpful. Institutions appreciated this approach and provided very positive feedback.
- The use of single points of contact within each HEI was very effective. Briefing events held early on for these contacts helped establish open positive relationships with the REF team and were appreciated by HEIs.
- The info@ref mailbox received a high number of enquiries with additional resource required from the REF team close to the submissions deadline.
- Queries led to the publication of numerous FAQs, some of which offered additional guidance. A range of supplementary guidance on detailed procedures, data and technical issues was also published on the web. It was therefore important that REF contacts were effective at disseminating such information as necessary within their institutions.

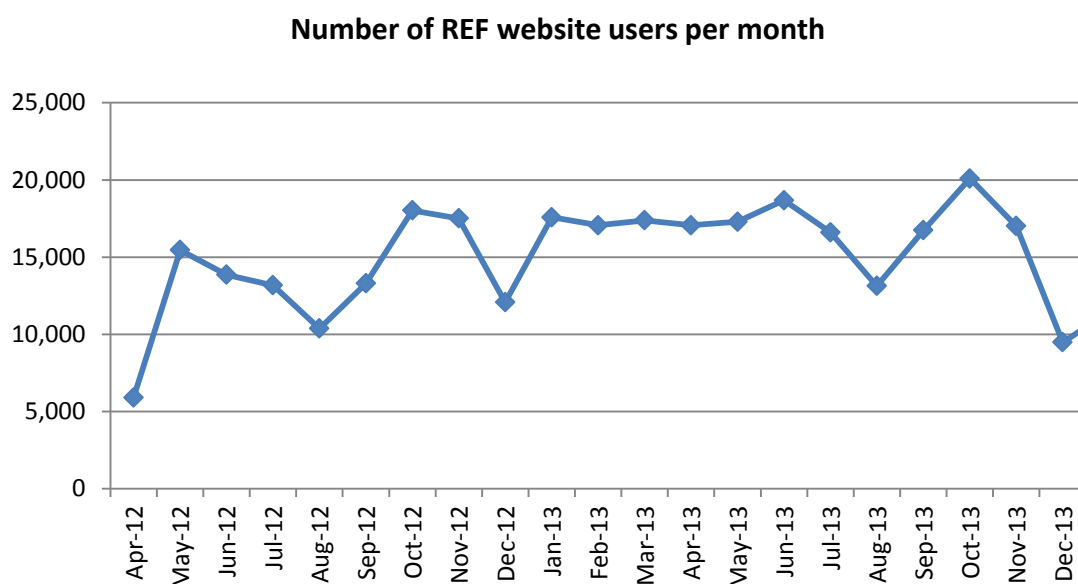
165. In late 2010 each HEI eligible to submit to the REF was invited to nominate a REF main contact and a REF technical contact. The main contact acted as the primary point of contact between the institution and the REF team on all updates and enquires relating to REF policies and guidelines. The technical contact acted as the institution's main contact with the REF team regarding the REF submission software and system support. Limiting the contact individuals at

institutions helped streamline the management of HEI communications within the REF team and enabled institutions to manage the dissemination of information effectively.

166. A REF Jiscmail list 'REF-NEWS' was also created and any interested individual was able to subscribe to this mailing list. A total of 2,100 people did so. Subscribers were notified of any REF website updates and the release of any publications.

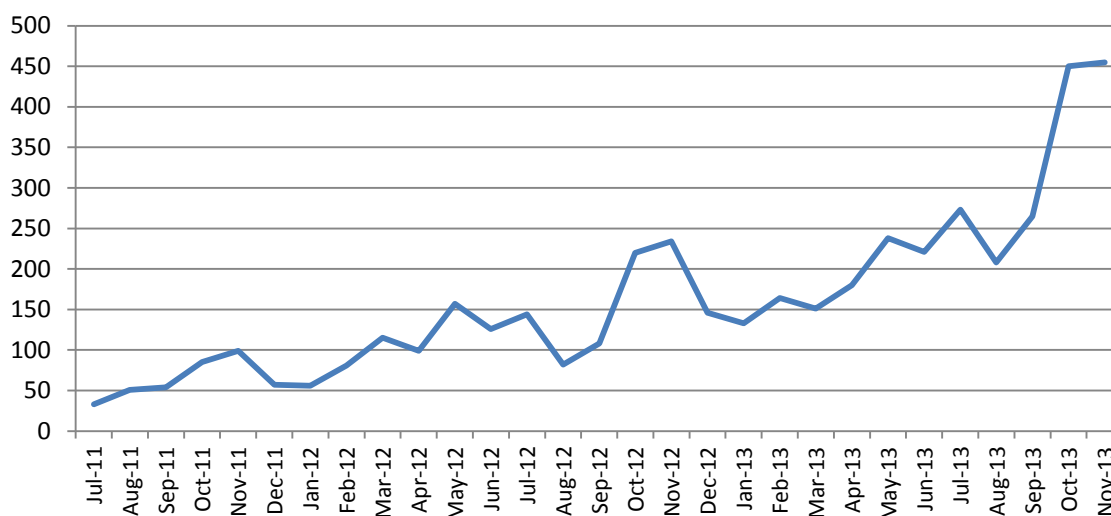
167. Numbered REF publications, supplementary guidance and FAQs were published throughout the criteria and submissions phase of the REF on the REF website (www.ref.ac.uk) and numbered publications were also made available in print. Up to 50 copies of key documents, the 'guidance on submissions' and the 'panel criteria' were sent to all submitting HEIs.

168. The REF website, separate to the HEFCE website, was launched in April 2012 and all web content was designed and published directly by the REF team. The website was used extensively and throughout the submission phase it was accessed by up to 20,000 users per month.



169. Institutions were able to contact the REF team via dedicated email support at info@ref.ac.uk where they had particular questions arising from the guidance documents. In total 4,685 email queries were received during the submissions phase of the REF. The number of queries significantly increased in the latter stages of the submissions phase with additional resource required from within the REF team to help manage these queries. In the majority of cases, all queries were responded to within one to two working days. Institutions provided very positive feedback on the quality and timeliness of responses.

Number of info@ref queries per month



170. The topics that queries related to are categorised below:

Topic of enquiry	Proportion of enquiries (%)
Outputs	16
Data collection	16
Environment (including data)	14
General administration	10
Staff (including staff circumstances)	10
Impact	9
Submissions	7
Equality and diversity	5
Nominations and panel membership	4
Early career researchers	3
Other*	6

* Other includes publications, panel configuration and roles, bibliometrics, consultation on the criteria, funding, quality profiles, and the impact pilot.

171. Output queries related to the eligibility of specific outputs and how the necessary information should be provided within the submission system. Data collection questions referred to what specific data should be provided as part of a submission. A large number of issues arose in relation to research income and PGR data (see section 5.6).

172. A large number of queries were received in relation to staff, especially individual staff circumstances. In particular, there were queries about:

- How to decide when an individual first carried out independent research and/or qualified as an early career researcher (ECR);
- How to combine multiple circumstances (including circumstances that overlapped);

- How much information was required about various circumstances (especially complex and health-related ones), and what kind of evidence might be required if a circumstance was audited;
- How to decide the appropriate reduction for complex circumstances;
- What was meant by ‘absence’ from the HE sector and academic research (for example, where an individual undertook R&D in industry, or held a non-academic post at an HEI).

173. In addition to responding to the queries individually, we periodically published a number of [FAQs](#) on the REF website to help clarify the guidance. This included some further details about the information required in relation to staff circumstances, which we relied on institutional contacts to disseminate within their institutions. The ECU also provided some fictional examples of complex circumstances, and offered advice to HEIs

174. The REF team held briefing events in autumn 2011 and in May 2012 in England, Scotland and Wales to brief institutional contacts on submission preparations. Two representatives from each institution were invited to attend an event. Sessions were held on the submissions guidance and panel criteria, the REF submission system, REF environment data, data verification, and other procedures such as multiple submission requests. The events were well received with positive feedback and helped develop very collegial working relationships between institutions and the REF team. Through the discussions, some of our intended guidance on procedures (for example the deadlines for submitting codes practice and the timing of multiple submissions requests) were adapted to suit institutions timetables, before being finalised.

175. In April 2012 we conducted an online survey of HEI main contacts to evaluate the quality of REF communications and identify any areas for improvement. 106 HEIs responded. Feedback was generally very positive. The tables below indicate the percentage of contacts who considered the overall quality of communications, and the various methods of communication, to be ‘good’ or ‘very good’.

Overall quality of communications	% good or above
Accessibility	96
Consistency	84
Clarity	81
Timeliness	78

Method of communications	% good or above
Guidance documents	95
Email alerts/updates	89
Other REF publications	88
REF briefings/events	87
Responsiveness to feedback	86
REF website	83
Responses to queries	78
FAQs	61

176. Many positive comments were made about the clarity, consistency and conciseness of key guidance documents (the guidance on submissions and panel criteria); about face-to-face contact and approachability of the REF team; speed of response to enquiries; and the frequency of email updates. Areas for improvement were identified as answering queries with more information than simply referring to the published guidance, and expanding the answers given in FAQs. We acted on these points and ensured that subsequent responses to queries and FAQs

were more expansive and helpful, or, on some topics, clarified that the issue was a matter of judgement for the HEI that the REF team could not provide any further guidance on. This received very positive anecdotal feedback, and at the end of the submissions phase, many REF contacts thanked or congratulated the team and info@ref support for the quality of our communications.

5.8. Survey of submission intentions

Key points

- In late 2012 all HEIs completed a survey about their intended submissions, which succeeded in enabling us and the panels to plan for the assessment and identify gaps in expertise that required the appointment of additional assessors.
- The volumes of staff reported in the survey proved to be reasonably accurate.

Recommendation

- The 'free text' information that was collected about impact was variable and difficult to analyse. The survey could be improved in future by requiring more succinct and structured information about impact.

177. The REF survey of submission intentions was an online survey to gather information about the submission intentions of submitting institutions, to inform the REF team's and the panels' planning and preparation for the assessment, especially to help identify panels' needs for additional assessors. The survey was not binding for institutions, but we explained its purpose and encouraged institutions to provide as accurate information as was possible at the time.

178. The survey was specified and developed in 2011 to allow the IT team to develop the online survey at a convenient time, when ample resource was available. The survey was password protected and institutions were expected to manage use of the system.

179. The survey was open for three months at the end of 2012, in parallel to the REF submission system pilot. All HEIs submitting to the REF completed the survey. Institutions were able to provide one submission intention per UOA, including the following information:

- Total volume of Cat A and Cat C staff combined (headcount).
- Total volume of Cat A staff (FTE).
- A statement outlining the areas of impact to be included in case studies.
- Information on the research specialisms to be included in the outputs:
 - i. Name of specialism(s).
 - ii. Volume of Cat A and C staff (headcount).
 - iii. Language(s) of any outputs not in English.
 - iv. Number of outputs in each language other than English.
 - v. UOA(s) to which cross-referral requests would be made.
 - vi. Number of cross-referral requests to the given UOA.

180. Reports of the data returned in the survey were designed in autumn 2012, significantly later than the original survey was specified. This involved iterations between the REF team and

IT team in order to produce reports of the data suitable for both the REF steering group and for the sub-panels.

181. Reports of the data returned in the survey were distributed to panel secretaries and sub-panel chairs in order to analyse and produce summary data for consideration at their early 2013 panel meetings, which focused on preparation for the assessment. This included an analysis of the volume of submissions, research specialisms, any special language requirements, and areas of impact expected to be submitted. Each sub-panel reviewed an analysis of expected submissions to its UOA, and on this basis made recommendations for the appointment of output assessors, impact assessors and specialist advisers (see section 7.1).

182. The analysis of the survey by the panel secretariat involved some challenges and was undertaken with varying levels of consistency and detail. More resource than expected was required to undertake this task. The way in which the data was collected added to the complexity of the analysis:

- a. Information on the volume of staff by headcount was skewed where individuals contributed to multiple research specialisms. This measure was used as collected data on the number of outputs per research specialism was thought to be inappropriate, pre-empting decisions on staff reductions in outputs.
- b. The 'free text' statement on impact areas was completed by HEIs with very varying levels of detail and helpfulness. The secretariat found it onerous and difficult to summarise/analyse this information. This data would have been easier to use if a single separate sentence had been requested for each anticipated impact case study – although institutions may not always have been able to be clear at the time of the survey.

183. The total volume of Cat A FTE staff reported in the survey was remarkably similar to the total volume that had been submitted in the 2008 RAE. We published these high level figures, partly to allay fears that institutions might be highly selective in their submissions, with adverse consequences for equality and diversity. After the actual submissions had been made, the figures reported to the survey proved to be reasonably accurate. In total HEIs overestimated the submitted FTE staff by just 4 per cent, with 52,061 FTE staff submitted compared with the 54,269 reported in the survey. Annex B provides a comparison at UOA level of the number of FTE staff actually submitted and the estimates from the survey.

5.9. Summary of submissions

Key points

- All 155 institutions completed their submissions by the deadline of 29 Nov 2013.
- The overall scale and scope of submissions was similar to that in the 2008 RAE, although the average number of outputs submitted per person reduced from 3.75 to 3.41 (due to the more widespread use of individual staff circumstances).

184. The deadline for submissions to the REF was 29 November 2013. All 155 institutions made their submissions by this deadline. (At the end of the assessment results were published for 154 institutions, as two merged during 2014.) In order to make a submission, the head of institution confirmed that the institution’s code of practice had been applied in selecting staff and that the submitted data were accurate, and agreed to the use of the information by panels and its subsequent publication.

185. The overall volume of submitted material was broadly similar to that in the 2008 RAE. The table below provides a summary comparison.

	Number of submissions	Category A FTE staff	Category A and C staff headcount	Number of outputs	Outputs per category A and C staff headcount	Number of impact case studies
REF 2014	1,911	52,060	56,033	191,150	3.41	6,975
RAE 2008	2,309	52,401	57,523	215,507	3.75	n/a
difference	-398	-341	-1,490	-24,357		

186. While the total volume of Cat A FTE staff was remarkably similar between the two exercises, there were some differences in other respects:

- Due to the ‘merging’ of UOAs there were significantly fewer (398) submissions made.
- As intended, far fewer Cat C staff were submitted.
- The total number of submitted outputs – and the average number of outputs per person – was lower in the REF, due to the more widespread use of individual staff circumstances in the REF.
- As intended, there was an increase in the number of outputs (2,851) submitted with a request for double weighting.
- Nearly 7,000 impact case studies were submitted to the REF. Of these, 296 were marked by the submitting HEI as ‘not for publication’ due to commercial or other sensitivities. These were made available to the panels for assessment but were not to be published with the rest of the submissions in 2015.

187. Institutions’ submissions ranged enormously in size and scope. The smallest submission included 1.2 Cat A FTE staff and the largest included around 450. At institutional level, submissions ranged from three FTE staff submitted by an HEI in a single UOA, up to around 2,500 FTE staff submitted by an HEI in 32 UOAs. The main panel [overview reports](#) and the published [analysis](#) provide further information about the submissions, which need not be repeated here.

188. In addition to making submissions, institutions were required to ensure that all the listed outputs were available, either electronically (through the submissions system) or in physical format. This is described in the section below.

5.10. Collection of outputs

Key points

- A significant operational aspect of the REF was to collect large volumes of research outputs, and distribute them to panel members, as efficiently as possible.
- Building on the 2008 RAE approach, and with assistance from the Publishers Licensing Society, we successfully sourced 98 per cent of journal articles and conference proceedings directly from publishers. This amounted to 80 per cent of all the submitted outputs and, as in RAE 2008, remained a very efficient method of collection.
- Institutions were encouraged to submit a range of other outputs electronically, and did so for a further 9 per cent of submitted outputs.
- The remaining 11 per cent of outputs were deposited by HEIs in physical form, representing a significant reduction in physical outputs since the RAE (from 53,525 to 21,588 in the REF). These were collected on schedule by the warehouse team.

Recommendations

- The guidance on output collection was produced by updating the equivalent RAE guidance, in discussion with a steering group. The guidance could be enhanced in future by consulting the panels at an earlier stage.
- There is scope to further increase the collection of outputs electronically, for example by requiring all book chapters to be submitted electronically. However, panel members' willingness to read the material electronically should be taken into account.
- The reports available on the submission system to inform institutions about which journal articles and conference proceedings they need to upload should be improved in future.

Preparation and guidance

189. A significant operational aspect of the REF was to collect a large volume of research outputs and distribute them to panel members as efficiently as possible. In the 2008 RAE, outputs had been collected in two ways: by sourcing journal articles directly from publishers' websites using the DOIs that institutions entered on the submission system; and requiring institutions to supply a physical copy of all other outputs. An RAE 'warehouse' facility collected and distributed the 53,525 physical outputs that were submitted.

190. During 2011 we started planning for the collection of outputs and decided to build on the RAE approach and try to collect more of the outputs electronically. We held early discussions with the Publishers Licensing Society (PLS), who proved to be very willing and helpful in working with us to repeat the RAE approach of gaining permission from publishers for us to download

outputs directly from their websites, using DOIs provided by HEIs. Based on RAE submissions, we drew up lists of relevant publishers and the PLS began raising their awareness and contacting them to gain their permission. As in the RAE, publishers generally proved to be willing, although some overseas publishers required additional efforts. The PLS employed a consultant, who had been involved in the RAE, who successfully managed contact with overseas publishers.

191. Once each publisher indicated their agreement, IT staff at HEFCE liaised with them to arrange download access to their websites, and developed the submission system software to retrieve the outputs, as in the RAE.

192. We expected that these arrangements would collect the majority of outputs, and considered whether we might be able to avoid collecting physical outputs from institutions. We held some discussions with the British Library to explore if they might be able to supply physical outputs directly to panel members. However, there proved to be significant difficulties with this idea, and in any event the British Library would not have access to non-text outputs (that are often compiled by HEIs in the form of portfolios).

193. We therefore decided that for all outputs that we would be unable to source directly from publishers, HEIs would be required either to provide an electronic copy via the submission system, or provide us with a physical output. We developed the details of this approach in discussion with DCSG, which largely represented people involved in managing institutions' REF submissions. We used the RAE guidance as the starting point and worked through each output type to decide whether HEIs ought to provide them electronically, physically, or have the option of providing them in either format.

194. For all outputs to be provided electronically by HEIs, we decided they would need to be uploaded to the submission system by the deadline of 29 November 2013 (to enable rapid distribution to panel members) in PDF format, and with some file size limits in place. For outputs to be provided physically, we decided we would need another warehouse as in the 2008 RAE. During 2013, the IT team at HEFCE began developing software to manage stock at the warehouse; the systems to support panel members were designed to allow them to place orders for physical outputs; HEFCE staff began identifying a suitable location for the warehouse; and we began recruiting a warehouse manager.

195. As we developed the approach and worked with the PLS, we put in place arrangements that would enable HEIs to supply us with outputs in the intended formats for the purposes of the REF assessment. As with RAE 2008, HEFCE and the Copyright Licensing Agency (CLA) signed a [licence agreement](#) which permitted HEIs to photocopy or to scan licensed material from paper to digital form and to provide this material to HEFCE; to submit paper copies of licensed material to HEFCE; to submit the DOIs for licensed material to HEFCE for storage on the data collection system; to download the publisher's PDF file (where the HEI is authorised to do so) to upload to the submission system; and to permit HEFCE to make such licensed material available to panel members in paper or digital form.

196. [Guidance](#) to institutions on the requirements for providing output was issued in early 2013. This included a list of the output collection formats acceptable for each output type and procedural guidance on uploading PDF files and on depositing physical outputs. The submission system data requirements also provided information on the metadata required for each output type. In developing the guidance we consulted some of the panel chairs on particular points we were unsure about. Later on, it became apparent that institutions did not always identify the appropriate 'type' for their outputs, and for portfolio-based outputs, did not always provide clear information about the research dimension of the work or included large amounts of unnecessary material. In future it will be beneficial to consult the panels on the guidance for collecting outputs, at an early stage.

197. The output types themselves were largely inherited from those used in the RAE, with the addition of 'working papers'. After submissions, it was apparent that the trend for an increasing proportion of outputs to be journal articles, observed in previous exercises, had continued. In the RAE 75 per cent of outputs had been journal articles; this increased to 80 per cent in the REF. The table below provides further details.

Output type	Number of outputs in RAE	% of RAE outputs	Number of outputs in REF	% of REF outputs	Change in % of outputs
A - Authored book	14,256	6.6%	12,873	6.7%	0.1%
B - Edited book	2,978	1.4%	2,066	1.1%	-0.3%
C - Chapter in book	18,331	8.5%	13,253	6.9%	-1.6%
D - Journal article	162,360	75.3%	153,626	80.4%	5.1%
E - Conference contribution	4,001	1.9%	2,738	1.4%	-0.4%
F – Patent	241	0.1%	111	0.1%	-0.1%
G – Software	115	0.1%	38	0.0%	0.0%
H - Website content	4,345	2.0%	157	0.1%	-1.9%
I – Performance	615	0.3%	491	0.3%	0.0%
J – Composition	736	0.3%	667	0.3%	0.0%
K – Design	454	0.2%	190	0.1%	-0.1%
L – Artefact	718	0.3%	732	0.4%	0.1%
M – Exhibition	2,691	1.2%	1,255	0.7%	-0.6%
N - Report for external body	1,276	0.6%	688	0.4%	-0.2%
O - Confidential report	58	0.0%	28	0.0%	0.0%
P - Devices and products	45	0.0%	25	0.0%	0.0%
Q - Digital or visual media	681	0.3%	487	0.3%	-0.1%
R - Scholarly edition	605	0.3%	436	0.2%	-0.1%
S - Research dataset/database	132	0.1%	79	0.0%	0.0%
T – Other	1,019	0.5%	553	0.3%	-0.2%
U - Working paper	n/a	0.0%	657	0.3%	0.3%
TOTAL	215,657		191,150		

Electronic outputs

198. Agreements from publishers were received over an extended period during 2012-13. During this period we frequently liaised with the PLS to chase up any outstanding agreements and also to identify any new publishers based on outputs that HEIs were entering on the submission system. Throughout the submissions phase, as HEIs entered DOIs (for journal articles and for conference proceedings), the system automatically downloaded and stored outputs from the publishers' websites. By the close of submissions we succeeded in directly sourcing 98 per cent of journal articles and conference proceedings for which a DOI was provided by the HEI. This amounted to 80 per cent of all outputs returned to the REF, and overall was a very efficient way of collecting the large majority of outputs.

199. HEIs were required to upload a PDF for all journal articles and conference proceedings that we were unable to source, by the submission deadline. A report was provided on the submission system which allowed HEIs to see whether the REF team had been able to source journal articles and conference proceedings directly. However, HEIs reported some difficulties with using the report and it could be improved in future. For three weeks after a DOI was entered into the submissions system the REF team attempted to source an electronic copy. Outputs not sourced in this time would appear on the report. However, if the HEI removed and re-added or changed the DOI during this three week period, the 'timer' reset for that output, causing some difficulties for institutions.

200. During the assessment it also transpired that in some cases, the DOIs provided had not retrieved the correct output. With one or two particular journals, the DOIs retrieved the 'author notes' document rather than the output, and in some cases an incorrect DOI had been provided by the HEI. These were resolved through audit, with the HEIs being asked to supply the correct PDF.

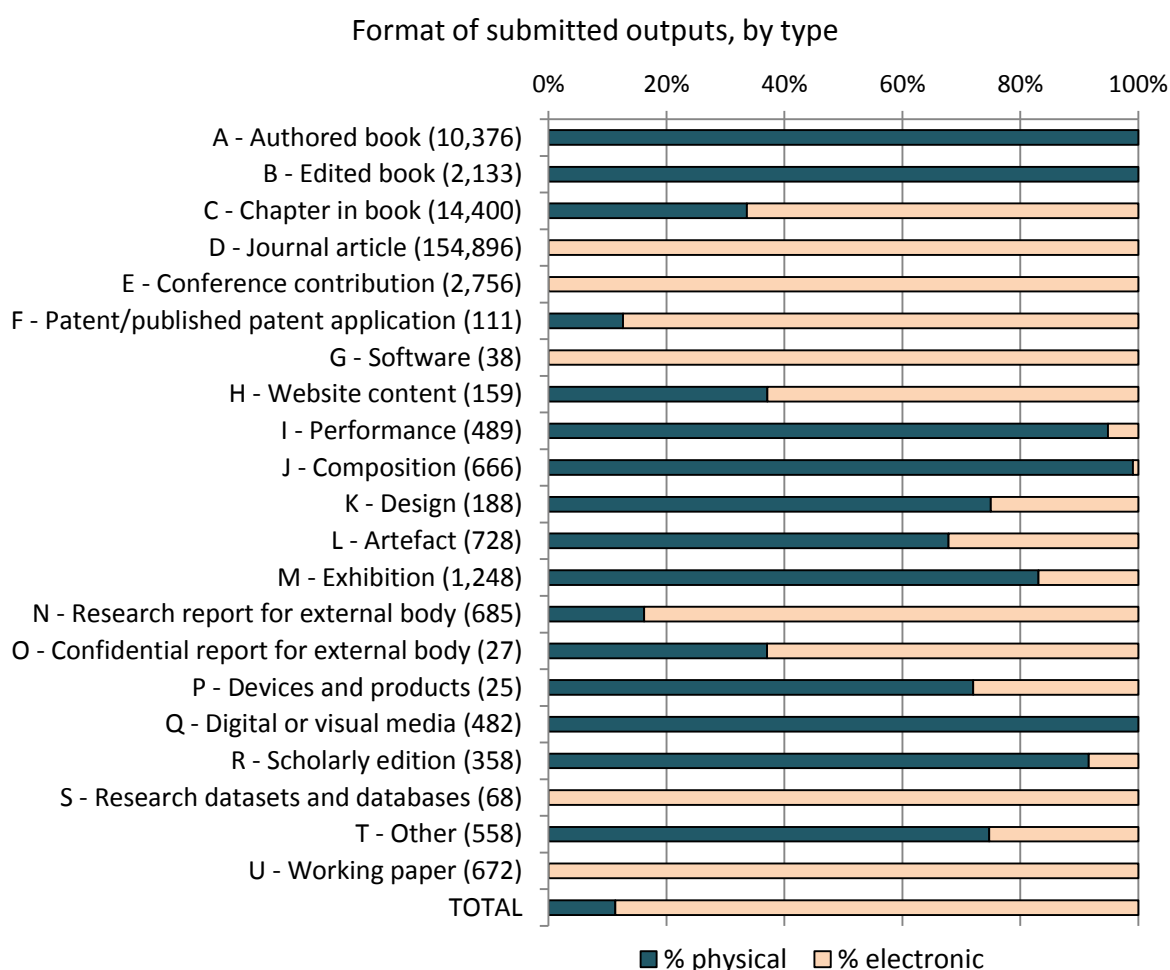
201. We required that certain output types (journal articles, conference proceedings, working papers, software and databases) could only be provided as PDFs. Some output types (authored books, edited books, and digital or visual media) were required in physical form. HEIs were encouraged – but not required – to provide the remaining output types in PDF form; and provide them physically only where it was impractical to upload a PDF.

202. In total, the REF team sourced 80 per cent of all outputs directly from publishers. HEIs uploaded a further 9 per cent of outputs in electronic form, and provided the remaining 11 per cent as physical outputs. This represents a significant shift since the 2008 RAE, in which 25 per cent of the outputs were provided in physical form. The shift improved the overall efficiency of the exercise (both in collecting outputs from institutions and in distributing them to panel members), and reduced the associated environmental impact.

203. The table below shows the proportion of each output type that was collected electronically and physically in the REF (the number of outputs for each type is indicated in brackets). In future there is scope to further increase the collection of outputs electronically. We suggest that all book chapters could be required electronically; consideration be given to increasing the electronic

provision of non-text or portfolio-based outputs; and the possibility of collecting electronic books be considered. However, the following issues would need to be taken into account:

- As indicated below, panel members made a larger than expected number of requests for us to print PDFs for them, indicating they are not always willing to read the material electronically.
- Guidance was provided about scanning quality and file sizes for PDFs that were uploaded by HEIs. During the assessment, some problems were identified by the panels in terms of inferior scan quality. There may also be burden implications for institutions scanning material that is not already available electronically.



Notes:

The number of each type of output submitted is shown in brackets.

The type of each output was identified by the submitting HEIs, and is not always accurate.

Physical outputs

204. Where HEIs could not upload a PDF of an output, and for books, HEIs were required to provide a physical copy of the output. All authored and edited books were collected in physical format, even if an e-book version was available, as format and licensing issues were found to be too restrictive to allow e-book access for panel members. In future this might change. In total

HEIs provided 21,588 physical outputs (a significant reduction compared to 53,525 returned to the RAE).

205. HEIs were required to deliver all physical outputs to a storage facility in Bristol during December 2013 and January 2014, following the close of the submission system. HEFCE was due to relocate offices in early 2015 and was able to take on the lease of its new office in 2013 for the storage and processing of submitted outputs. This site was located close to the current office where the REF team was based. HEFCE's head of facilities management, who had overseen the delivery of the RAE output storage facility, was able to take on the same role to ensure the office space was fitted out to the required specification. The facility was known as the 'warehouse', terminology inherited from the RAE, when a warehouse building had been used. In hindsight this terminology conveyed a misleading impression of the facilities available and the nature of the work, which had an impact on collection and delivery of materials, staff recruitment and external perceptions.

206. The facility was provided with a broadband voice and data connection, PCs, and telephone equipment by HEFCE's IT department. An MFD printer/copier was leased. Maintenance and support were provided by HEFCE's facilities management team, who also managed the cleaning contract.

207. A warehouse manager was recruited to oversee the set-up and operation of the facility and manage the work of further staff. The warehouse manager began work at the beginning of September 2013, based initially at HEFCE's main office, and two months were spent planning warehouse operations before the new building became available in early November. Two warehouse assistants were recruited to start work when the facility became available.

208. A delivery schedule was drawn up requiring each HEI to send their physical outputs to arrive on a particular date between Thursday 5 December and Friday 10 January. HEIs were informed of their delivery date towards the end of October. The schedule was necessary to ensure that the warehouse team were able to receive and unpack the outputs efficiently and that there was enough space to receive each delivery. As the exact number of outputs to be sent to the warehouse was not known until the close of submissions, data from the RAE was used to estimate the likely proportion of the total number of physical outputs which would be sent by each HEI. These figures were then used to draw up a schedule which would distribute the deliveries as evenly as possible. There was only a brief window between the close of submissions and the beginning of the assessment process. Those institutions sending only small numbers of outputs were given dates at the beginning of the schedule to allow more time for institutions with larger consignments to prepare.

209. Deliveries of physical outputs from institutions ranged in size from one or two books to 1,295 items. Each output received was logged into the warehouse stock management system, ROSS, which had been developed by HEFCE's IT team (based on the equivalent system used in the RAE). The output metadata used for stock control was drawn directly from the submissions system. Outputs submitted electronically did not appear on ROSS as stock items. ROSS also displayed orders and panel member information drawn from the panel members' website, to support the distribution of outputs to panel members.

210. There were queries to be resolved relating to the physical outputs sent by approximately a third of submitting HEIs. The most common issue was that of outputs sent to the warehouse which were not expected as physical outputs because a PDF had already been uploaded to the submission system. The instruction that each output could be submitted either electronically or in physical form, but not both, was not always understood, particularly in cases where it was felt that additional material was needed to support a journal article, or in cases where a portfolio contained an item available electronically. The time needed to resolve these queries took resource away from dispatching outputs at the point when demand from panel members was highest. In future it will be important to convey this guidance as clearly as possible.

6. Audit and data verification

6.1. Sampling and data comparisons

Key points

- A mixture of approaches were taken to auditing and verifying the accuracy of submissions through sampling, comparisons with other available data, and panels raising ad hoc queries. Overall this was an effective and proportionate approach.
- A sample of 2,435 staff from all HEIs was checked for eligibility, revealing a very high level of compliance with the criteria.
- Through comparing REF submissions with other available sources of data we identified some potential discrepancies and audited a total of 1,002 items. Amongst these, 195 outputs were found to be ineligible as they had been previously submitted to the RAE.

Recommendations

- The future REF team may wish to indicate more clearly in advance what forms of evidence would be acceptable to verify staff eligibility.
- In future, it may be useful for the submission system to provide a report on outputs which had been submitted to the REF2014, to avoid their resubmission.
- In future it may be possible to align the REF environment data more closely with existing HESA data returns, potentially reducing the need for checking and auditing these data.

211. Institutions were expected to have assured themselves of the accuracy of submitted information when they made submissions. To ensure that panels could make their assessments based on reliable and accurate data, and therefore uphold confidence in the outcomes, we carried out a data audit and verification process. This was adapted from the RAE approach, and similarly sought to strike an appropriate balance between rigour and avoiding undue burden on institutions.

212. All elements of submissions were subject to verification, through a combination of audit requests generated by the REF team – based on sampling and comparisons with other data –

and ad hoc queries raised by sub-panels. In early 2013 we issued [guidance](#) to provide HEIs with advance warning of the types of audits to expect, their likely timing and the nature of evidence required. Each HEI designated an audit contact that was responsible for receiving and responding to all audit queries from the REF audit team. All audit correspondence occurred via a closed secure webmail system.

213. For each audit query, HEIs were requested to provide evidence and/or further information to address the concern raised. Depending on the nature of the query, the audit response was either passed on to the relevant REF panel for consideration or reviewed by the REF team (in which case the REF manager took decisions regarding eligibility and data adjustments). In a very small number of instances, the REF Steering Group was invited to take decisions on staff eligibility.

214. Data adjustments were carried out to add, remove or replace missing or inaccurate data where it was identified via the audit procedures. Some adjustments were also made where, after the submission deadline, the HEI notified the audit team of errors or missing information in their submissions, or wished to change the redaction status of a case study. Protocols were in place to avoid making any data adjustments that would provide the institution with an unfair advantage, whether by increasing the submitted values (such as staff FTE or research income), or allowing an institution to change its intentions after the submission deadline.

Random staff sample

215. Largely following the RAE approach, we selected a sample of Cat A staff from each HEI to check their eligibility. Up to 40 staff in up to four UOAs were selected from each HEI, depending on the scale and number of the HEI's submissions. For each member of staff in the sample, HEIs were requested to provide contractual and payroll evidence to verify eligibility, and to explain their internal procedure for verifying the individual staff circumstances. There were significant differences in the types of evidence provided, depending on HEIs' HR or payroll systems and we were flexible in accepting a range of types of evidence. In future the REF team may wish to indicate more clearly in advance what forms of evidence would be acceptable.

216. In total the eligibility of 2,435 staff were checked, of which 873 staff had been submitted with individual circumstances. The process involved significant effort and the REF team temporarily drew in three members of staff from other HEFCE teams to check the evidence for each individual. Overall it revealed a very high level of compliance with the eligibility criteria, across all institutions. Only seven data adjustments were made, mostly to change the FTE values.

	No. of HEIs audited	No. of staff audited	No. of data adjustments
Random staff sample	155	2,435	7

Staff data comparison

217. We compared submitted data with other submissions and with data held externally by HESA and the CrossREF publication database. Where we identified discrepancies, we sent audit request queries to the HEIs concerned.

218. Staff data contained in REF submissions was compared across all HEIs to identify any individuals who had been returned by different HEIs with a total FTE of greater than 1. Nineteen individuals were identified involving 31 different HEIs. The HEIs were requested to revise the submitted FTE downwards to reflect the proportional contracts for the individual to sum to no greater than 1, or the audit team apportioned the FTE based on contractual evidence. Three cases were not possible to resolve in this way and were referred to the REF Steering Group to take the final decisions. These cases involved individuals taking up employment at a new HEI late in the REF period. As a result of these audits, 23 data adjustments were carried out to amend the submitted FTE, and three staff members were removed from submissions.

	No. of HEIs audited	No. of staff members audited	No. of data adjustments
Staff duplicates audit	31	19 duplicates	23

Outputs data comparison

219. Submitted data for research outputs were compared with RAE2008 data and the CrossREF publication database. Outputs were identified for audit where they had previously been submitted to the RAE2008, or where the publication year in the CrossREF database was outside the REF publication period.

220. 314 outputs were identified as having been submitted to both the REF and the RAE. Outputs submitted to the RAE were eligible for submission to the REF only in very particular circumstances, explained in the [FAQs](#). A high proportion of these outputs identified were found to be ineligible. This was due to outputs having been submitted for the same staff member by a former HEI, or outputs having been submitted to the RAE by the same HEI for a different staff member. In some cases, they were not aware that the content of the RAE submissions had been published and were available online. In future, it may be useful for the submission system to have the facility to generate a report on outputs which had been submitted to the REF2014, to reduce the incidence of ineligible outputs.

	No. of HEIs audited	No. of outputs audited	No. of ineligible outputs	No. of data adjustments
REF/RAE duplicate outputs	83	314	195	4

221. The publication dates of journal articles submitted to the REF were compared with records held by the CrossREF publication database. Outputs were matched on the basis of DOI. The comparison identified 958 outputs with a potential publication date outside of the REF publication

period. All outputs identified with a pre-2006 and 2014 CrossREF publication date were audited. Sample audits were carried out for outputs with a 2007 CrossREF publication date (as such outputs were eligible if they satisfied paragraph 43 of the 'panel criteria'). We also sampled outputs which had been flagged as pending publication.

	Publication year	No. of HEIs audited	No. of outputs audited	No. of ineligible outputs
Output publication year comparison	2001-2007	96	277	9
	2014	135	340	4

222. For both the above audits, ineligible outputs remained in the submissions and sub-panels were instructed to give these outputs a 'U' classification. A small number of data adjustments were carried out on eligible outputs (for example to add the required statement on output content within the publication period or amend bibliographic data).

Environment data comparison

223. Submitted environment data had already been validated by the REF submissions system, which imposed limits based on each HEI's returns to HESA for the years 2008-09 to 2011-12 as explained [here](#). During the audit exercise we compared submitted REF4 data against HESA data as a secondary check, both to take into account HESA data for 2012-13, and to sense check how institutions had allocated the data to REF UOAs. Audits were raised where significant discrepancies were identified for any given year or source of income; or we had difficulty in understanding how the REF data at UOA level could have related to the HESA data at cost centre level.

224. As a result of these comparisons, we sent 40 queries to 27 different HEIs. HEIs were requested to provide further information or evidence to verify that any REF data that significantly exceeded HESA data were eligible, or to explain how the data had been allocated to UOAs. Where discrepancies could not be verified, the ineligible data was removed or replaced. As a result of this audit, 29 data adjustments were made to REF submissions. Reasons for the data adjustments included where HEIs had incorrectly classified funding as research income, income reported to the REF which was not reported to HESA, and the return of income under different sources to the REF versus their HESA returns. In future it may be possible to align the REF environment data more closely with existing HESA data returns, potentially reducing the need for checking and auditing.

	No. of HEIs audited	No. of audits	No. of data adjustments
Environment data comparison	27	40	29

6.2. Panel-instigated audits

Key points

- Sub-panels raised further queries on any items they had concerns about, or to seek further information where required to make particular judgments. In total (including the audits generated by the REF team) we audited 6,635 individual items, which resulted in 1,072 data adjustments, and a further 1,936 items through 'bulk' audits.
- Overall the audit procedures were proportionate and provided the panels with confidence in the accuracy of the information they were assessing.
- The quality of institutions' data was generally high. There were some variations in the completeness of submitted data, but we did not find any institution that had repeatedly submitted inaccurate information.

Recommendations

- Many of the sub-panels' queries – and the 'bulk' audits – were due to insufficient information submitted about individual staff circumstances and the author's contribution to a co-authored output. Also, many of the impact case study queries required basic details about the staff that had carried out the underpinning research. Such queries could be reduced in future by collecting some of these data as separate fields and/or providing templates, guidance or training to institutions.
- Specific attention was given to auditing and corroborating impact case studies. This worked well although in future it may be beneficial to extend the timeframe for sub-panels' assessment of impact (typically two months in REF). If the funding bodies consider collecting all corroborating material at the point of submission, the burden implications for institutions and the difficulty of preventing supplementary evidence should be appraised.

225. In addition to the REF sample audits and data comparisons, REF sub-panels could raise queries to verify any aspect of a submission that they had concerns about, or to seek further information where it was required to make particular judgments. There were four main categories of panel-instigated audits: staff, outputs, impact and environment.

Staff

226. Over 1,000 individual queries on staff were raised by sub-panels. The large majority of these were requests for further information on individual staff circumstances, when insufficient information had been provided in the submission. This request was most commonly for ECRs (61 per cent) and also for staff submitted with part-time working as a clearly defined circumstance (25 per cent). The further information from these queries was used to inform judgements about whether the output reduction criteria had been met (see section 8.6). A number of staff queries were also raised to verify staff eligibility, particularly for individuals who the panels believed to be based overseas (7 per cent of staff queries). A small number of data adjustments for staff queries were made, largely to remove staff who did not fulfil the eligibility criteria.

227. In addition to the queries relating to individual members of staff, we found several institutions that had repeatedly failed, across their submissions, to provide sufficient information about clearly defined staff circumstances. These institutions were invited through ‘bulk’ audits to check all their clearly defined staff circumstance statements and add the necessary information to enable the sub-panels to verify that the criteria for output reductions had been met. These ‘bulk’ audits involved 988 staff who had been submitted with clearly defined circumstances. One further institution reported that it had submitted incorrect ECR dates for its staff and these were corrected through a ‘bulk’ audit.

	No. of audit queries	No. of data adjustments
Individual staff queries	1,045	19

Outputs

228. Sub-panels raised almost 1,000 queries on individual outputs. Many of them (45 per cent) requested further information about the submitted individual’s contribution to the output, especially by sub-panels within Main Panel A which had specific criteria for demonstrating author contribution. These queries were raised where the required statement on author contribution had not been provided, or the sub-panel had doubts about the author’s contribution, for example based on information available in the output itself. The outcomes of these audits were considered by panels and were used to inform their judgements about whether the criteria for author contribution had been met.

229. In addition to these queries on individual outputs, we found that several whole submissions within Main Panel A repeatedly lacked the required statements on author contribution. The institutions were invited through ‘bulk’ audits to provide the information, for a total of 751 outputs (in addition to the individual queries).

230. The second most common query regarding individual outputs was to replace an incorrect or poorly scanned PDF (29 per cent). These accounted for the majority of data adjustments regarding outputs. Other types of query included requests to obtain evidence to verify the eligibility of outputs with editions or versions which first appeared outside the publication period (11 per cent), and requests for further information or evidence to verify output authorship (3 per cent). For some practice-based outputs, there was some difficulty in determining the date at which the output first became available in the public domain. In addition to the replacement of incorrectly submitted or poorly scanned PDFs, data adjustments were also carried out to amend bibliographic data, such as changing the ‘output type’.

	No. of audit queries	No. of data adjustments
Individual output queries	991	388

Impact

231. As impact was being assessed for the first time and the evidence of external impact had been provided by the HEIs themselves, we designed a process for audit and corroboration of impact case studies that would provide sufficient confidence in the veracity of the impact claims, while not causing undue burden on institutions. The process involved auditing a sample of case studies, focusing on those that the sub-panels had concerns about, supplemented by random sampling by the REF team (where necessary) to ensure a reasonable spread of audits across UOAs and HEIs.

232. We invited sub-panels to raise queries for between 5 and 10 per cent of the case studies submitted in each UOA. We provided them with guidance on the types of query that would be appropriate to raise, for example to verify that the underpinning research met the 'threshold' criteria, or to corroborate key claims made about the impacts. Sub-panels were instructed not to seek additional information to expand on the impact claims already made within the case study; and if seeking corroboration, were limited to requesting the sources listed in the case study rather than pursuing other sources of evidence. After sub-panels had raised their queries the REF team checked how they were distributed across HEIs, and generated some additional queries for institutions that had received audits for significantly fewer than 5 per cent of their case studies.

233. In total, 594 queries were raised about impact (8.5 per cent of case studies). The most common query regarding impact was to seek corroboration of the claimed impacts (61 per cent). A significant number (24 per cent) were raised to obtain details of the key researchers and the positions held at the time of the research, as these details had not been provided in the case study. Some queries (13 per cent) requested copies of the underpinning research outputs in order to check the threshold for research quality. It should be noted that in addition to these queries raised via the REF team, panel members were able to directly access sources of corroboration that were in the public domain, and could try to access underpinning research outputs without raising an audit query. Given the frequency of missing information about who the key researchers were and their employment at the time of the research, we recommend that the case study template could be amended to include specific separate fields to capture this information.

234. The majority of the impact audit queries sought further information which the REF team passed on to the panels, to inform their decisions either about the threshold criteria, or about the veracity of the impact claims. Where a case study was deemed not to meet the threshold criteria it was awarded a 'U' grade by the panel, but the case study remained part of the submission and no data adjustment was made. A small number of data adjustments were made to replace incorrectly submitted case study PDFs.

	No. of audit queries	No. of data adjustments
Impact queries	594	9

235. The process for audit and corroboration of impact worked well overall and gave the sub-panels sufficient confidence in the veracity of the information and evidence they were assessing. Nevertheless, it did involve some complexities. Sub-panels mostly assessed impact over a period of around two months and there were some challenges for them in raising the queries early enough to receive the outcomes in time to inform their decisions. In some cases, they made decisions that were provisional, pending the outcomes of outstanding audits.

236. Some sub-panels also had differing initial expectations about the nature or extent of checking and auditing of case studies that they might carry out. For example, sub-panels within Main Panel B had initially expected to review outputs underpinning every case study, to check the quality threshold, although in the event they did not find this to be necessary. Some sub-panels suggested they would have liked access to all of the corroborating materials, and that in future this should be collected at the point of submission.

237. In the event all sub-panels applied the REF guidance and raised impact audit queries in a broadly consistent way and there was some, but relatively limited, variation in the percentage of case studies audited by each sub-panel. In future it would be beneficial to extend the time-period over which impact is assessed, partly to allow more time for audits. If the funding bodies consider collecting all the corroborating material at the point of submission, the burden implications for institutions should be appraised. Institutions might feel compelled to spend significant additional effort in compiling such material for every case study, much of which is unlikely to be needed by panels in making their judgements. In addition it would be very difficult with this approach to prevent institutions from including supplementary evidence of impact, rather than simply corroborating claims already made in the case study.

Environment

238. The panels raised a very small number of queries relating to information within the environment templates. The queries that were raised were mostly due to discrepancies identified by panels between the submitted income data and information within the environment template.

	No. of audit queries	No. of data adjustments
Environment queries	18	3

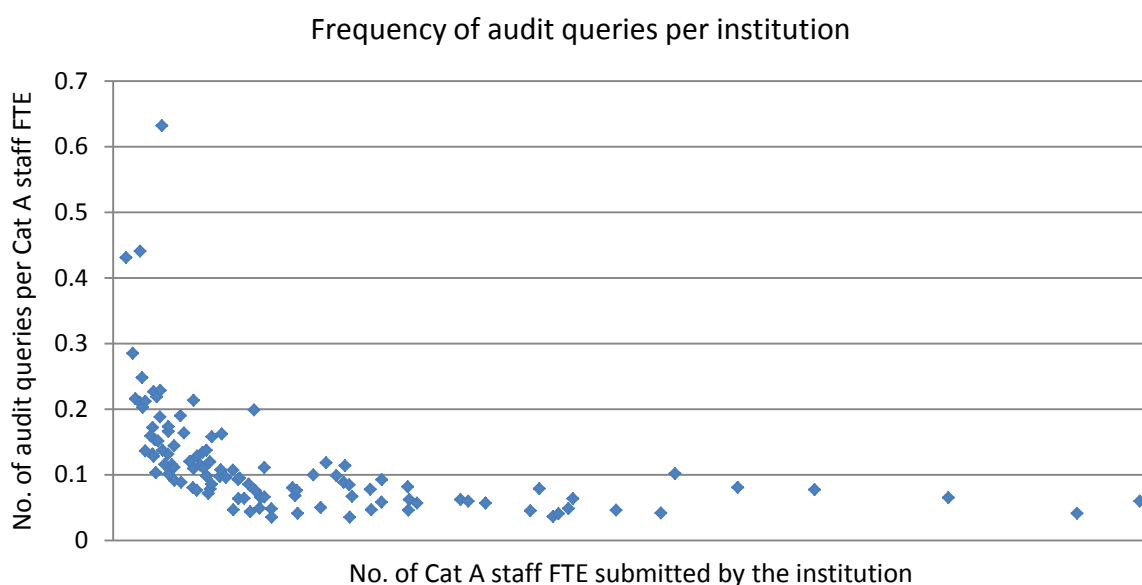
Overall data quality

239. In total, we audited 6,635 individual items (including queries generated by the REF team and those instigated by the sub-panels), which resulted in 1,072 data adjustments. In addition several institutions were 'bulk' audited to complete missing information for a total of 1,936 items.

240. The mixture of sampling, data comparisons and panel-instigated queries proved to be effective in both targeting audit queries where they were required, and ensuring the audit procedures were proportionate and avoided undue burden on institutions. The overall proportion of data that was audited was relatively small – given that the 56,069 members of staff and 191,148 outputs each comprised numerous data fields, each of which could trigger an audit –

and the audits successfully provided the sub-panels and REF team with assurance about the quality and accuracy of submitted data.

241. All institutions were audited, both through the staff sample and subsequent audits. The scale of audits varied somewhat across institutions reflecting some variations in the quality or completeness of their submitted data, but we did not find any institution that had repeatedly submitted inaccurate information. The chart below indicates the frequency of audits per HEI (excluding the staff sample audits and bulk audits), in relation to the overall size of their submissions. We reviewed the nature of the audits for those institutions with the highest frequency of audits, and found that the large majority related to insufficient rather than inaccurate information in their submissions.



Notes:

Random staff sample audits and bulk audits are excluded from this chart.

HEIs with fewer than 10 audit queries are excluded from this chart.

242. It should be noted that the overall volume of audit queries increased since RAE 2008. This is accounted for by the additional audits for impact and the large numbers of queries to request further details about staff circumstances and author contribution. In future, it would be beneficial to collect such information in a more structured way with specific fields to capture key data, and/or provide training, templates or detailed guidance to encourage all institutions to provide sufficient information in their original submissions.

7. Preparation for assessment

7.1. Recruitment of assessors and specialist advisers

Key points

- After the survey of submission intentions, a second round of recruitment for assessors was carried out. Overall this was effective in supplementing the panels with the expertise they required to assess submissions. 287 additional assessors were recruited (158 for impact and 129 for outputs).
- While effective, this more targeted approach to recruit people with specific individual expertise was more onerous and protracted than the initial recruitment of panel members.
- Users were generally willing to participate as impact assessors, provided that relevant individuals could be contacted within the organisation.
- We hosted briefing events for assessors prior to the assessment phase. This was much appreciated by them, particularly impact assessors.
- Many of the outputs submitted in languages other than English were reviewed within the sub-panels (including through cross-referral). Specialist advisers were successfully appointed to review the remaining outputs that were not assessable by the sub-panels.

Recommendations

- For the next REF, the funding bodies may wish to reconsider the timing of member and assessor recruitment, and whether a single round of recruitment nearer to the assessment phase might be preferable to the two phase process used in REF2014. (This depends on other decisions about the level of sub-panel members' involvement in the criteria phase.)
- Users requested details about the exact time commitments involved in joining as an assessor. The REF team provided this in a generic way and in future, it ought to be possible to provide more specific details at an early stage in the recruitment.
- The process of identifying the requirements for specialist advisers could be simplified and made more efficient in future by gathering information at an early stage about the language competencies of panel members.

243. The sub- and main panels met in early 2013 to prepare for the assessment. They reviewed the analysis of the survey of submission intentions to identify, for each UOA:

- Discipline areas where there were gaps in the sub-panel's expertise required to assess outputs, or where the volume of outputs in particular areas would lead to excessive workloads for existing members.
- The main types of user expertise required to assess impacts.

244. Based on the survey each sub-panel made recommendations for additional assessor requirements, comprising:

- Output assessors. These were practising researchers with relevant expertise to participate in assessing the output element of submissions.

- Impact assessors. These were people with professional experience of making use of, applying or benefiting from academic research, to participate in assessing the impact element of submissions.
- In reviewing their expertise requirements, a few sub-panels also recommended the appointment of additional members (where additional expertise was required to assess all elements).

245. The recommendations were discussed at each main panel, to ensure a broadly consistent approach to recruiting additional assessors. This involved some clarification about the new assessor role; a few sub-panels were wary about the distinction between members and output assessors. At the main panel discussion, opportunities to appoint 'joint' assessors were also identified. At this stage the recommendations focused on identifying the number of assessors required on each sub-panel and their specific areas of expertise. Significant further work was required to identify (and then recruit) specific individuals to fill each role.

246. Once the numbers and areas of expertise were agreed, the sub-panel chairs went about identifying suitable individuals. Those originally nominated in 2010 were considered, as well as an additional 750 user nominations that the REF team had sought through an open call in autumn 2012, in anticipation of the need for further user nominees. Sub-panels were able to fill many of the vacancies through these two rounds of nominations, but a number of specific areas of expertise required further nominations. These were targeted by first identifying the appropriate organisations and then inviting them to nominate individuals with specified expertise. These were sought during the course of 2013 and eventually resulted in a further 109 individuals being nominated, and the filling of all the vacancies. Many of these targeted appointments required protracted communications involving the sub-panel chair and possibly deputy chair or members, the panel secretary, the REF team, nominating body or bodies, and the potential candidates.

247. In a small number of instances, the sub-panel knew of a suitable individual and recommended they be co-opted. Co-option was undertaken as a last resort and in total 14 individuals were co-opted to the panels (four international main panel members, seven panel members or assessors and three specialist advisers).

248. Where further nominations were invited, the REF team encouraged the nomination of individuals from underrepresented groups, and asked nominating bodies to explain the relevant expertise of nominees as well as to state how they would help enhance the diversity of the panels concerned. In recommending assessors for appointment, sub-panel chairs took account of the diversity of their panels, but remained limited to selecting from amongst those who had been nominated.

249. Users were generally willing to participate as impact assessors, once we had been able to identify suitable individuals within the organisations to approach. A small number declined due to lack of time rather than interest, and a few (particularly those who were self-employed) declined due to the fee levels. A number of the users, particularly in the private sector, understandably wanted specific details about the timing and workload that would be involved. The REF team provided generic information about this, estimating the range of case studies to be reviewed by each assessor as between 20 and 40. It was difficult at this early stage for sub-panels to predict

the exact timing and volume of work, and we encouraged sub-panel chairs to discuss these issues with the assessors at the briefing sessions described below. With the benefit of the REF experience, it ought to be possible in future to provide users with more detailed practical information at an early stage in the recruitment.

250. Overall, 287 additional assessors were recruited, comprising 158 impact assessors and 129 output assessors. (17 of these had been recruited earlier, during 2012, where a clear gap in the expertise of a sub-panel had already been known about).

251. In the initial stages of this round of recruitment, panel secretaries supported the process by analysing the survey of submission intentions in each UOA; working with the panel chairs to identify suitable individuals; and carrying out initial correspondence with nominating bodies where further nominations were necessary. The secretariat had been allocated time for this until March 2013, by which stage some progress had been made, but the recruitment turned out to take longer than anticipated. Around half the appointments were outstanding and passed back to the REF team to pursue.

252. Overall the recruitment of assessors was more resource intensive and protracted than the original recruitment of members, which had been almost entirely drawn from a single round of nominations. In addition to the need for multiple communications and responses from numerous parties, other reasons for the protracted recruitment included: difficulties experienced by some sub-panels in analysing the survey; desirable candidates being employed by HEIs already represented on the sub-panel; and some desirable candidates declining after being nominated, which meant the process had to start again. The majority of assessors were appointed by June 2013, but additional appointments continued until well into 2014 (for example, where the actual submissions required some unanticipated additional expertise or involved an excessive workload for members, or to address panel member resignations mostly due to health issues or changes in personal circumstances).

253. In June 2013 we first published the names of appointed assessors alongside the panel membership. We updated the published membership lists on a monthly basis as additional appointments were completed. By the end of the exercise, the total numbers of members and assessors who participated (that is, including panellists who resigned but had participated at some stage) were as follows:

Main panels	
Chairs	4
International members	23
User members	17
Sub-panel chairs	36
Total main panel membership	80

Sub-panels	
Chairs	36
Members (including deputy chairs)	765
Output assessors	129
Impact assessors	158
Specialist advisers	25
Total sub-panel membership	1,113

254. In late October and early November 2013, we held two assessor briefing events. The REF manager, main panel chairs and main panel user members gave plenary presentations, and sub-panel chairs led discussions with their assessors. The events provided a morning session on output assessment and an afternoon on impact assessment, with assessors attending the relevant session. The briefings provided an introduction to the REF assessment, an opportunity for assessors to meet their sub-panel chairs and fellow assessors, and discuss what would be expected of them. This was particularly beneficial for the impact assessors who were often unfamiliar with the exercise. Given that the assessment of impact was new for everyone, the impact session was also appreciated by the newly recruited sub-panel members who attended; later on a number of sub-panel members commented informally that they would have appreciated the opportunity to attend such a session prior to the assessment phase.

Specialist advisers

255. The survey of submission intentions collected information on the numbers, languages and specialisms of any outputs in languages other than English that institutions intended to submit. The REF team reviewed these data to identify requirements for recruiting specialist advisers. Before recruiting specialist advisers we corresponded with the sub-panel chairs and secretariat to establish which outputs could be reviewed by existing panel members with relevant language skills and expertise. This enabled us to develop a picture of the language competencies available on each sub-panel.

256. We then had further correspondence to identify which outputs could be cross-referred between sub-panels, rather than recruiting specialist advisers. For any remaining outputs, we drew up a list of requirements for specialist advisers. Sub-panel chairs considered existing nominations, or recommended which organisations should be approached for further nominations. By autumn 2013 we had appointed the majority of specialist advisers so they would be ready to start in the early stages of the assessment.

257. We appointed 32 specialist advisers but after submissions were received (and some further cross-referral opportunities identified) we found that seven were not required to undertake any assessment duties. The names of the 25 individuals that were called on to provide specialist advice were [published](#) at the end of the assessment; between them they reviewed around 200 outputs in the following languages or forms of specialist notation: Bactrian/Sogdian, Bengali, Bulgarian, Catalan, Czech, Ethnomusicology, Finnish, French, German, Greek, Hungarian, Irish, Korean, Macedonian (Cyrillic), Mandarin, Medieval notation/Latin, Polish, Scots Gaelic, Serbian

and Welsh. Particular attention was given to the assessment of outputs in the medium of Welsh, as outlined in the ['Guidance on submissions'](#) (paragraph 129), involving a combination of panel members, output assessors and specialist advisers.

258. Overall the process succeeded in enabling the panels to review outputs in the full range of languages in which they were submitted. However, the process of identifying the precise requirements for recruiting specialist advisers was somewhat drawn out and onerous in relation to the total volume of outputs involved. It would be useful in future to gather information at an early stage about the language competencies of panel members, to simplify this process.

7.2. Assessment scheduling

Key points

- The REF team worked with the main panels in early 2013 to plan out the assessment process and associated meeting schedules, in order to start arranging meetings well ahead of the assessment.
- By mid-2013 all meeting dates and the key business for each day of each meeting was established. It was inherently difficult to anticipate the whole of the process so long in advance, but the planned schedules proved to be workable. The panels generally stuck to them, with relatively few adjustments.
- A common sequence of assessment was agreed for the sub-panels within each main panel, which proved to be essential for the main panels to maintain effective oversight.
- Given the newness of the impact element and the REF team's experience of the pilot, the REF team developed and proposed a schedule for assessing impact. It proved to be effective although was tight for some sub-panels.
- Many of the meetings were planned to take place in 'clusters' which proved to be useful, if adding some complication to the scheduling.
- Sourcing and booking meeting venues was a significant task that commenced well in advance of the start of the meetings. During 2013, the REF administration team booked suitable venues within budget for all 283 panel meetings, held over 467 meeting days. When surveyed, over 55 per cent of members considered the venues to be good or very good and over 90 per cent satisfactory.

Recommendations

- It would be preferable in future if all sub-panels within a main panel agree to follow the same assessment schedule, without any deviations.
- To improve the quality of venues in future more time could be spent selecting and testing venues; budget levels might be reconsidered; and it would be beneficial for the administration team to attend more of the early meetings.

Assessment workplan

259. During 2012, the REF team began to plan out the whole of the assessment phase, to identify how many meetings the panels would need, when, and what the main business would be at each round. These details needed to be finalised during the early part of 2013 to start booking

venues with the right capacity for the right number of days; to get meetings into members' diaries; and to invite the output and impact assessors to the relevant meetings. Early planning of the process also helped panels to prepare and enabled the main panels to provide effective oversight.

260. We used the RAE experience as a starting point, and developed proposed schedules to discuss with the panels at their round of planning meetings held in early 2013. Proposed schedules were based on estimates of the timing and procedures that would be required to assess each element of the assessment (outputs, impact and environment) including:

- initial calibration exercises
- allocating work to members and assessors
- dealing with cross-referrals
- audit queries
- members and assessors reviewing their allocated work and agreeing scores
- producing feedback.

261. As well as the assessment by sub-panels, we planned the sequence of main panel meetings and identified the key interactions between the sub- and main panels throughout the process. This was to ensure that each calibration exercise would involve an exercise at main panel level that fed into the sub-panel calibrations; and that each main panel could review emerging results at interim stages, and sign-off final results. These intended interactions required that each group of sub-panels within a particular main panel would follow the same sequence and timetable for assessment.

262. At their early 2013 meetings, each main panel discussed and agreed a broad sequence and overall timetable for their assessments. For all main panels this involved making an early start on outputs (which would take the longest to complete) with interim and completion points identified; and impact and environment being assessed at particular times between March and September.

263. After the early 2013 meetings at which each main panel agreed a broad schedule, there followed an extended period of refining the workplans in more detail for each sub- and main panel, so that specific meeting dates and suitable locations could be identified. A small number of sub-panels started requesting deviations from the agreed main panel workplan. These tended to be from some of the sub-panel chairs who had been former RAE panel chairs, and appeared to want to repeat parts of the RAE schedule rather than the process agreed by the REF main panel. In consultation with the main panel chairs, these amendments were agreed only where they entailed completing some of the work ahead of the main panel schedule. This was to ensure that the main panel would be able to maintain oversight of each aspect of the assessment with access to comparable data for all its sub-panels.

264. The common main panel assessment schedules proved to be highly effective in enabling the main panels to inform calibrations and to maintain oversight throughout the process, including thorough discussions of emerging results. The small number of deviations by sub-panels – on the basis outlined above – did not compromise this although it did involve a small number of sub-panels assessing environment or impact ahead of the main panel calibrations:

their results were therefore provisional pending the outcomes of the main panel calibrations. It would be preferable in future if all sub-panels within a main panel agree to follow the common sequence agreed by the main panel, without such deviations.

265. As part of the planning for the meetings, we agreed with the main panels that it would be beneficial to arrange most of the sub-panel meetings in 'clusters' (groups of around three sub-panels meeting at the same time at the same location). The intention was that this would enable main panel members – especially the internationals – to observe multiple sub-panels at each stage and might also help with sub-panels sharing issues or seeking advice on specific items during the course of meetings. In the event the clusters were valuable in enabling attendance of international members as well as main panel chairs and main panel user members, and were helpful to the panel secretariat. Arranging the clusters did however add some complication for the REF admin team and limited the range of available venues.

266. It was inherently challenging for both the REF team and the panels to anticipate the whole of the assessment process so long in advance, but the early decisions and plans proved to be workable. The panels generally stuck to them, with relatively few – and only minor – adjustments. Adhering to the assessment schedule involved a heavy workload and demanding deadlines for many panel members. It also depended on the sub-panel chairs (and many of their deputies) engaging in the operational details; the support of a highly proficient panel secretariat; and effective IT systems and support provided by the REF team.

Meeting arrangements

267. After the panels agreed their broad meeting schedules for the assessment phase, first the panel chairs and then members were canvassed for all 2014 meeting dates, using a tool built into the panel members' website (PMW). Schedules of potential dates and locations were then developed. Taking into account the clustering requirements and member availability, the REF team identified preferred dates for all meetings of all sub- and main panels and confirmed these with the chairs. In parallel to this there was correspondence with chairs about various proposed amendments to some of the details in the schedules, such as which groups of assessors were required to attend which particular days of meetings.

268. Sub-panel chairs were also invited to indicate preferences for the location and types of venue for meetings and these were taken into account when venues were booked. There were some differences in preference for city-centre versus out-of-town locations and we established a general agreement that one-day meetings would be held in city-centre locations (for ease of travel) while multi-day meetings would normally take place elsewhere.

269. The REF administration team undertook a large number of visits to venues around the UK to find venues with the required facilities and of sufficiently high quality. As a result of the visits a ranked list of appropriate venues was established for each location taking into account cost (a strict budget limit was in place), quality, ease of transport, capacity (with some panels requiring large rooms and many meetings to be clustered), facilities, and – to minimise the burden on panel secretariat – quality of on-site support. The admin team developed relationships with preferred suppliers and negotiated a number of benefits based on high-volume bookings:

- reduced rates
- more flexible terms and conditions (particularly cancellation policies for overnight accommodation)
- more flexible minimum number policies (enabling rooms to be booked for greater capacities for days attended by impact and output assessor without paying increased rates for the days with lower attendance)
- securing additional syndicate rooms within the standard budget, where these were requested by sub-panels.

270. The REF admin team sought to achieve a balance in the number of different venues used. On the one hand using a range of venues provides variety for panel members and would reduce the impact if any one venue failed to deliver sufficient quality. On the other hand a small number of different venues enabled volume-based discounts and reduced the 'learning curve' for venues in meeting our particular requirements (for example, in providing many more power sockets than usual). In all, 21 separate venues were used and given the variety in preference expressed by panel members in the survey this number seemed to be appropriate.

271. During the course of 2013, the admin team booked suitable venues within budget for the full range of panel meetings. This was a significant task and it was vital to commence this well in advance of the start of the meetings. In total, meeting arrangements were made for 283 separate sub-panel and main panel meetings, held over 467 days. An additional 20 meetings for the REF Steering Group, EDAP, assessor briefing sessions and secretariat training sessions were also arranged.

272. Once the assessment started, occasional problems were encountered with some of the venues (typically relating to overnight accommodation service quality), and there were repeated service problems at one particular venue. The REF admin team attended a number of meetings at specific venues to ensure that services were adequately provided. When panel members were surveyed in mid-2014, over 55 per cent considered the meeting venues to be good or very good and over 90 per cent considered they were satisfactory. To improve the quality of venues in future more time could be spent selecting and testing venues prior to making bookings (including staying overnight to check the quality of accommodation services); budget levels might be reconsidered; and it would be beneficial for the administration team to attend more of the early meetings.

7.3. Assessment systems

Key points

- New IT systems were developed to support the panels' work in the assessment phase. This development, and the subsequent user support provided to panels, was a substantial project for the REF team and HEFCE IT staff.
- We started designing the systems in late 2011, in discussion with a user group of panel members. The systems were developed in time for the start of the assessment in December 2013. Time for testing was constrained and some bugs were present in the initial stages, and some non-critical aspects were rolled out during 2014.
- The systems and associated training and support proved to be effective in enabling the panel members to work efficiently, and with flexibility to suit their own working practices. When surveyed, 85 per cent of panellists rated the systems as effective or highly effective, and a further 13 per cent as adequate.
- The systems succeeded in maintaining the security and confidentiality of the data used and generated throughout the assessment process.
- The systems provided the main panels with real-time summary data, enabling them to provide effective oversight of the assessment.

Recommendations

- Panel members with basic IT skills, and the user members/impact assessors, tended to find the systems more difficult and relied more heavily on support from the REF team and panel secretariat. In future we suggest that these panellists be identified early in the process so that targeted training/support can be provided.
- Spreadsheets were used extensively to enable members to work offline, which required regular downloading from and uploading to the panel members' website (PMW). In future, it may be possible to develop more user friendly methods of enabling offline working, while synchronising with the central database when online.
- The information provided to enable panellists to identify and resolve discrepancies in scores with their co-reviewers could be improved.

Overview

273. In preparing for the assessment phase we anticipated that suitable IT systems would be needed to enable the REF panels and the REF team to process large volumes of data throughout the assessment phase. From late 2011 through to early 2014, we specified and developed new IT systems to support the assessment, which proved to be vital to the success of the exercise. The aims in developing the assessment phase systems were to:

- Enable the REF panels and the REF team to carry out their work as efficiently as possible throughout the assessment phase.

- Provide systems that were usable by all panellists, with very different IT skills and a range of platforms.
- Provide generic systems that met the core requirements of all panels, while also providing flexibility to be used in different ways by individual panels/panel members to suit their detailed working methods.
- Ensure the security and confidentiality of the information used and generated through the assessment process.

274. In October 2011 and May 2012 the REF team held two sets of focus groups with a selection of panel members (including one member from each sub-panel) and secretaries, to identify their broad requirements and priorities for the assessment systems. The focus groups highlighted that panellists' main priorities were to: ensure speedy and efficient distribution of submissions and outputs to panels (in a range of formats); enable both offline and online working; and to develop tools to assist in allocating work to panel members and for recording members' scores.

275. In broad terms, we expected the assessment to involve the following range of processes, which we developed systems to support:

- Panellists required access to the submissions and to the outputs submitted in their UOA.
- The panel executive groups (the sub-panel chair, deputy, panel adviser and secretary) would allocate all the items within their UOA to panellists for assessment.
- Panellists would review all the items allocated to them, arrive at provisional scores to be recorded and reported to the sub-panel.
- Sub-panels would cross-refer some items to other sub-panels for advice, and seek specialist advice on non-English outputs.
- Sub-panels would raise audit queries and be provided with the audit outcomes.
- Sub-panels would agree and record a final score for each assessed item, and generate sub-profiles and overall quality profiles to recommend to the main panel.
- Main and sub-panels would frequently access reports on sub-panels' progress with their assessments and on the emerging and recommended results.
- Throughout the whole process the panellists would need to be able to communicate confidentially with each other.

276. During 2012 and 2013 the REF team developed specifications for the systems in collaboration with HEFCE IT staff. The IT systems were specified and developed in four areas:

- Administration: The PMW was developed to support panellists with the administrative aspects of their membership and meeting arrangements.
- Communications: A REF webmail system was provided to allow panellists (and the REF team) to communicate securely and confidentially on assessment matters between meetings. This system was procured from an external contractor.
- Distribution of submissions and outputs to panellists: Each panellist was supplied with an encrypted USB pen containing all submissions and electronic outputs for their UOA. These data were also available online through the PMW. Outputs collected in

physical form were despatched to panellists from the REF warehouse on request; each panellist was provided with a 'reading list' on the PMW from which they could request physical outputs. Where required, panel members were also able to request hard copies of any part of a submission (including PDF outputs) via the PMW.

d. Assessment systems: Spreadsheets were provided for sub-panel chairs and deputies to allocate work to panellists for assessment; and to individual panellists with details of the items allocated to them. They were used for recording scores for each assessed item and for managing other processes such as cross-referral and audit queries. 'Panel' spreadsheets containing data for the whole UOA were provided to the panel chairs, deputies, advisers and secretaries (who acted as an 'executive group' for each sub-panel) and a 'personal' spreadsheet was provided to each individual panellist containing data about the specific items allocated to them for assessment. Panellists generated spreadsheets and saved their work by downloading from and uploading to the PMW. They stored the spreadsheets locally on their encrypted USB pens to enable offline working. The spreadsheets included links to PDFs stored on the USB pens, and links to data available online on the PMW, to support efficient working practices. All the assessment data generated by panels was stored in a database on HEFCE servers, updated when panellists uploaded their spreadsheets to the PMW. Validation checks were in place and processes were built in to resolve any potential data clashes, given that multiple users could update the same data items. A range of reports were available via the PMW for different categories of users to monitor progress, identify discrepancies, review emerging outcomes and generate the final profiles.

Development

277. Development of these systems was a significant project within the REF team, with the bulk of the development from mid-2012 to the end of 2013. The project involved two full-time developers and also drew on significant resource from within HEFCE's IT team. Overall project management was provided by the core REF team who also led on specifying the user requirements for the systems. The REF team and IT staff worked highly collaboratively with weekly meetings to discuss system requirements and development activity.

278. A panel user group (PUG) was established to provide user input on the systems requirements and user testing of the systems as they were developed. This group consisted of two sub-panellists (including deputy chairs) from within each main panel, and three members of the panel secretariat. The group met four times within the project timeframe, initially to advise on system requirements and at later meetings to view demonstrations and test the developing systems.

279. A range of testing was also conducted by members of the ITS team, though in some areas timescales were tight and this limited the amount of testing which was carried out. As a result, a number of bugs needed to be fixed after the system went live. Mostly these were minor but some caused initial difficulties for some panel chairs or members.

280. As the timescales proved to be tight, during the course of the development we prioritised the delivery of different elements of the systems. All the vital elements were delivered in time for the start of the assessment, with the highest priorities being the rapid delivery of USB pens and the spreadsheets that enabled chairs to allocate work. The reports and some other features (for example, to manage cross-referrals) were rolled out after the assessment phase commenced.

Training and user support

281. In discussion with PUG we gave significant attention to usability, training and support for panel members, especially given the very wide range of IT skills amongst panellists. The spreadsheets to be used by the executive groups were necessarily large and complex. We sought to make the spreadsheets for individual panellists as simple as possible, and also tried to make other parts of the system user-friendly (for example, by synchronising passwords across the PMW, USB pens and webmail system). Nevertheless we anticipated that panellists would need a significant amount of support in using the systems.

282. A full day's training was provided for the panel executive groups ahead of the assessment phase, in December 2013, to prepare them for allocating work, using the spreadsheets and various reports to manage and monitor the sub-panels' work, and using the systems to manage other processes such as cross-referrals. The training sessions were timed so that the people attending could receive their loaded USB pens, and access the live system to generate spreadsheets for the first time. The training sessions proved to be vital and enabled the executive groups to make a rapid start to their first major task of allocating work, although for some individuals this was impeded by bugs that were discovered in the initial weeks.

283. We did not have the resources to provide face-to-face training sessions for (around 1,000) other panel members and assessors, and in discussion with PUG developed appropriate support mechanisms. A 'quick start' user guide was supplied to each person along with their USB pen, and a more detailed user guide made available on the PMW. User support was provided by email and by phone during office hours, initially by two full-time members of staff (reducing to one in the latter stages). There was a peak in user support queries in the early stages as a large number of users 'got up to speed' with the system. The user support team managed around 670 individual queries in January and February 2014. The majority of these were to resolve issues with the spreadsheets or with the USB pens. When surveyed in mid-2014, 89 per cent of panellists rated the support as good or very good, and a further 10 per cent as adequate.

284. In addition, the REF admin team provided initial support with members accessing their USB pens, and throughout the assessment phase the panel secretariat, who had rapidly become proficient in using the systems, provided a significant amount of support and advice to panel members on using the systems.

285. The systems and associated support proved to be effective in enabling the panels to conduct their work efficiently, and with flexibility to suit their own working practices. When surveyed, 85 per cent of panellists rated the systems overall as effective or highly effective, and a further 13 per cent as adequate. Of those who had been panel members in the 2008 RAE, 80 per cent reported that the IT systems had improved or significantly improved for the REF. By the

end of the exercise many chairs and members had provided very positive feedback; the Computer Science and Informatics sub-panel, for example, formally reported its congratulations to the REF team for the effectiveness of the systems.

286. Nevertheless, panellists who rated themselves as having basic IT skills and the user members/impact assessors (who used the systems less extensively) tended to be less satisfied with some aspects of the IT systems. Both of these groups also tended to rely more heavily on support from the REF team and panel secretariat. In future we suggest that these panellists be identified early in the process so that targeted training and/or support can be provided.

287. Throughout the assessment phase we maintained a log of all reported data security incidents. Overall the IT systems succeeded in maintaining the confidentiality and security of data as no serious incidents occurred. Some panel members lost or misplaced their USB pens (but these were encrypted), and some inadvertently wiped the data on their USB pens (which was recoverable from the point of their last upload to the PMW). The REF admin team retrieved virtually all the issued USB pens at the end of the exercise.

288. The systems proved to be vital in enabling the main panels to exercise their oversight function effectively. The main panel chairs and advisers were able to frequently generate reports summarising all the work completed within each UOA, and the emerging profiles of scores, which provided a sound basis for regular main panel discussions on scoring patterns and issues of comparability.

289. Though the assessment systems were effective, there is scope to improve them further in future. Based on the user support queries and feedback from panellists, we recommend improvements in the following particular areas:

- The requirement for panellists to be able to work offline drove the decision to use spreadsheets which could be stored locally on USB pens, and downloaded/uploaded regularly via the PMW. These aspects of the system generated the vast majority of user support queries. In future, it may be possible to develop more user friendly methods of enabling offline working, synchronising with the central database when online.
- The information provided to panellists to enable them to identify and resolve discrepancies in scores with their co-reviewers could be improved.

8. Assessment

8.1. Distribution of submissions and outputs

Key points

- Submissions data and electronic outputs were distributed to panel members on encrypted USB pens, and were also made available online. The USB pens were loaded and dispatched very rapidly, enabling panels to make an early start to the assessment, and proved to be an efficient and effective means of enabling panel members to access the material.
- Improvements were made to the system for panel members to request physical outputs (and printouts of PDF outputs) from the REF warehouse. Members placed their own orders via 'reading lists' on the PMW, which enabled the REF team to respond directly to their needs.
- A total of 3,898 orders were placed and a total of 25,828 physical outputs were dispatched from the warehouse. Panel members were positive about the distribution service and the support provided by warehouse staff.
- Members made more requests for printouts of PDFs than we had anticipated: in total 15,793 PDF outputs were printed and dispatched.
- Panel members made very few requests for us to source outputs that underpinned impact case studies (they were encouraged to try to find the outputs online first).

Recommendations

- Further improvements could be made to the reading lists in future, to display and track the location of physical outputs more clearly.
- To minimise the unnecessary depositing of large boxes of material, HEIs should be encouraged or required to submit portfolio-based outputs in electronic form (by uploading PDF files or on digital storage media).
- The amount of physical material sent to the warehouse and associated cost and environmental impact could be reduced further in future if HEIs are required to submit all book chapters electronically.

Electronic materials

290. During 2013, as part of the assessment systems development, we designed methods of providing all the relevant submissions and outputs to panel members and assessors. This involved:

- a. Supplying each panel member and assessor with an encrypted USB pen, loaded with all the submissions data and electronic outputs in their UOA.
- b. Panel members could also access this material online via the PMW, which also included updated material not available on the USB pens (such as data adjustments, and submitted outputs that were 'pending publication' in January 2014).

c. Once the panel chairs allocated work, each panel member was provided with an online 'reading list' on the PMW, from which they could place orders for physical outputs – and printouts of PDF materials – from the REF warehouse.

291. In early 2013 IT staff at HEFCE tested a range of encrypted USB pens and identified the most suitable product, which provided encryption without the need for local installation of software by panel members. Estimates were made of the storage capacity required for each UOA and in July 2013 we procured 1,200 USB pens with capacities ranging from 8 to 32 GB. Each one was manually registered and a temporary password set.

292. In the lead up to the submissions deadline, the REF admin manager drew up a detailed plan for loading and dispatching the USB pens. HEFCE IT staff developed and tested code for generating all the necessary data and associated indexes, for each UOA. Panel members' delivery addresses were checked, three temporary administrators were appointed and a bank of PCs set aside for loading. As soon as the submission deadline passed, the data was processed and organised into the correct file structures for loading, and loading commenced. The REF admin manager tested samples to ensure the data were correct.

293. The team succeeded in dispatching the USB pens, each with the correct data for the recipient and to the correct address, to all panel members and assessors by 11 December 2013, just eight working days after the submissions deadline.

294. To help panel members work efficiently with the large volumes of data contained on the USB pens, a browser-based index was provided on each USB pen, and their personal spreadsheets were designed to include direct links from each item they had been allocated, to the relevant PDF document stored on the USB pen.

295. Although a 'quick start' guide was posted with the USB pens, a high volume of queries were received initially, typically for support with logging in and updating passwords. A very small number of panel members were unable to use the USB pens due to security restrictions on their work computers and we loaned them laptops instead. After the initial login difficulties, panellists generally found the USB pens to be an effective and efficient way of accessing and working with the material. When surveyed, 80 per cent of panellists rated the USB pens as good or very good, and a further 14 per cent rated them as satisfactory.

296. During the assessment period some USB pens malfunctioned, were lost or accidentally wiped of data. In these instances a replacement was provided by the REF team and a log of all lost data was maintained. Panel members were required by the REF confidentiality arrangements to return their USB pens at the end of the assessment period. They were collected by the panel secretaries and the admin team who managed to retrieve around 96 per cent of all issued USB pens, by February 2015.

Physical outputs and printouts

297. Outputs collected in physical form were despatched to panellists from the REF warehouse on request; each panellist placed orders from their 'reading list' on the PMW. Panel members

were also able to request hard copies of any part of a submission (including PDF outputs) via the PMW. As orders were placed, they appeared on the warehouse system (ROSS) for processing.

298. Panel members began ordering items from 15 December 2013 – before many of the physical outputs had been sent by HEIs to the warehouse – and there was a backlog of around 100 orders by the time the warehouse staff began dispatching outputs, a few days ahead of schedule, on 15 January 2014. Orders were normally processed chronologically. Some panel members placed successive orders for a large number of items, and these were broken down and dispatched in stages in order to ensure that as many panel members as possible were receiving material for assessment. 3,898 orders were placed during the assessment period. The last order was placed and dispatched on 17 September 2014. A total of 25,828 physical outputs were dispatched from the warehouse.

299. Panel members were able to order printed copies of PDF outputs, but were generally discouraged from doing so (given they had access to these outputs on their USB pens and on the PMW). This service had been available during the RAE, and demand had been high. We had expected that developments in technology, increased familiarity with working online, and provision of the USB pens would have substantially reduced the demand: however, this proved not to be the case. In particular, at the beginning of the assessment period the demand for printing was much higher than anticipated. This added to the order backlog and so delayed the dispatch of physical outputs. Extra resource was needed from the REF admin team to complete the printing orders. 15,793 outputs in total were printed and dispatched to panel members during the assessment period. Ironically, the largest number of printouts were requested by the sub-panel with 'Environmental Studies' in its title.

300. In total the warehouse team dispatched 3,815 consignments during the assessment period, 2,105 of which were Royal Mail small parcels and 1,710 consignments of one or more boxes sent using the courier firm used more generally by HEFCE (which had been established according to approved government tender procedures). They proved not to be the ideal choice as they had only recently begun to expand their next-day delivery service and a number of deliveries and remote collections were outsourced to their partner companies. This caused some communication problems and errors. We suggest that a fuller procurement procedure be undertaken for any future output distribution service.

301. Most outputs were assessed by more than one panellist. Items returned to the warehouse were therefore processed promptly so that they could be sent out to a second panellist where necessary. The 'reading list' on the PMW provided members with links to information about the current availability of each item ordered, but many panel members were unaware that they had access to these details, which were not displayed as prominently as they might have been. The warehouse team therefore frequently dealt with queries from panel members as to why items they had ordered had not yet been dispatched; the answer being that the output was still with another panel member. Much clearer presentation of this information on the PMW would be beneficial in future, as would improved procedures for panel members to co-ordinate the multiple assessment of physical outputs.

302. Panel members were given several options for returning outputs to the warehouse. They could post the outputs and claim back the costs. The warehouse team also provided postage paid address labels valid for small parcels. These were sent out automatically with small parcels, and could also be requested by panel members. A third option was that remote courier collections could be booked by the warehouse team, so that one or more boxes of items could be collected from a panel member's home or office and delivered to the warehouse. This service proved very popular with panel members. It also benefited the warehouse team, as collections could be timed to reduce the likelihood of large quantities of material arriving on any one day. This enabled more efficient processing of returned outputs, particularly in June and July when the daily volume of returned material was very high.

303. Panel members could pass outputs directly to each other, either by post or in person at panel meetings. The 'reading list' allowed panel members to record that an output had been passed to another member, so that the location was automatically updated on ROSS and visible to warehouse staff. However, panel members often did not record this on the PMW, and in future this feature should be made more prominent. The warehouse team could also change the location of an output, in cases where panel members had not recorded that an item had been forwarded. There was a recurring problem where items were returned to the warehouse with no sender indicated, and warehouse staff were unable to tell whether a second order that had been placed for the output still needed to be fulfilled, or whether the output had already been passed on to the second panellist. Better procedures for panel members to co-ordinate the multiple assessment of outputs could help to reduce this problem for future exercises.

304. The physical outputs supplied in UOA 34 (Art and Design: History, Practice and Theory), posed particular challenges for output dispatch and return. Several institutions supplied sizeable portfolio boxes, often containing not only the output but a large number of supporting items in hard copy. In some cases, the output depositing guidance had not been followed, and actual artefacts and products had been sent to the warehouse, rather than the photographic or other visual record requested. Members of Sub-panel 34 therefore often had to receive and then return numerous large and heavy boxes. These consignments also took considerably more time to pack up securely, which mean that the warehouse staff spent a disproportionate amount of time on orders from one sub-panel. The environmental impact and expense were also significant, with the cost of sending UOA34 courier consignments being more than a fifth of the total courier dispatch cost during the assessment period. We suggest that improved guidance be provided on the provision of portfolios, and that HEIs should be encouraged or required to submit such outputs in electronic form (by uploading PDF files or on digital storage media), rather than supplying boxes of physical material.

305. The warehouse and audit teams liaised frequently to resolve a range of issues, for example where an institution incorrectly supplied an output in both PDF and physical form, and one or other had to be removed from the REF systems.

306. In addition to the submitted outputs, the warehouse also collected outputs that underpinned impact case studies, where these were requested by panels and were not available as PDFs. Software (the output distribution tool) was developed by HEFCE IT staff to allow the REF audit team to enter details of such outputs so they then appeared on ROSS, and could be

logged by the warehouse on receipt and ordered by panel members. However, the development work took longer than anticipated and an interim spreadsheet was used when impact case study outputs were first requested. In the end the number of such outputs requested was small and the manual records kept would have been sufficient for the warehouse to manage the process.

307. Overall, the systems for panel members to request outputs through their reading lists were significantly improved compared to the RAE (in which secretaries placed all orders on behalf of panel members), and the warehouse provided an effective distribution service. When surveyed:

- 78 per cent of members who used the reading list to order outputs rated it as good or very good (and a further 18 per cent satisfactory).
- 70 per cent rated the output distribution service as good or very good (and 24 per cent satisfactory).
- 82 per cent rated the support provided by warehouse staff as good or very good (and 17 per cent satisfactory).

Returning physical outputs to HEIs

308. After the September 2014 round of panel meetings, members were contacted by their panel secretaries and requested to return any remaining physical outputs to the warehouse by mid-October, so they could be returned to the institutions. Around 1,100 outputs were still outstanding after this date and panel secretaries assisted the warehouse manager in tracking these down. It might have been helpful to inform panel members of this deadline earlier in the year.

309. The warehouse team returned outputs to HEIs between mid-November and mid-December 2014. The couriers provided a logistics plan grouping consignments by geographical area, which reduced costs and environmental impact. Outputs were collected from the warehouse in batches, usually of around 150-200 boxes at a time and transported to regional depots across the country, for delivery to institutions. A delivery schedule was drawn up to allow warehouse staff time to pack up each consignment, and HEIs were provided with their delivery dates in advance. In total 1,346 boxes of outputs were returned and the logistics operation generally ran smoothly, with only one regional delivery experiencing delays and no boxes lost in transit.

310. Throughout the assessment, a total of 214 physical outputs went missing. Around two-thirds of these were book chapters supplied by the institutions as photocopies. Items which had gone missing during the assessment and were still required by panellists had been replaced (either by ordering a new copy, or requesting one from the HEI). The REF team undertook to replace any books which were identified as missing at the end of the assessment. HEIs were given the option of replacing the item themselves and invoicing the REF team, or requesting the warehouse manager to order and dispatch a replacement. In most cases HEIs preferred the latter option. The ISBN provided in the output metadata on the submission system was used to identify and source the replacement, so it was important that it referred to the correct edition.

311. Book chapters were deposited by HEIs either as a photocopy of the chapter or the whole book. However, there was no record on our systems of which format had been supplied – and HEIs also often held no such records – and this caused some difficulties. Panel members

frequently mistook a photocopied book chapter for a printout of a PDF file, and disposed of it accordingly. When book chapters went missing the warehouse staff did not know whether a photocopy or a book needed to be replaced. This led to extra work for the warehouse manager and panel secretaries chasing missing outputs at the end of the process. The amount of material sent to the warehouse and associated cost and environmental impact could be reduced in future if HEIs are required to submit all book chapters electronically.

8.2. Assessment process

Key points

- Sub-panel chairs' (and deputies') first major task was to allocate work to panellists. The REF systems generally enabled them to do this efficiently.
- Calibration exercises at main and sub-panel level required some additional meeting time and early workload, but were important to help ensure comparability of assessment standards.
- The main operational challenge for the panels was to review the large volume of outputs. Most panels therefore wanted to start this as early as possible, and this was achieved by allocating and distributing outputs as rapidly as possible.
- Panels succeeded in reviewing all outputs on schedule, but the workload for many panellists was very heavy and the time allowances made by panel members' institutions varied greatly.
- Panels achieved strong consensus on their assessments of impact; although the volume of material to review was comparatively small, this did require a significant amount of discussion and meeting time.

Recommendations

- Careful consideration should be given in future to how the panellists' workload issues might be addressed, including the possibility of sampling outputs.
- All main and sub-panels systematically carried out calibration exercises for outputs and impact; calibrations for environment were more variable and in future it would be beneficial to plan these more systematically from the outset.
- A larger number of outputs were cross-referred than anticipated. The assessment system was largely effective in supporting this but could be improved in future to deal with 'bulk' as well as individual requests.
- Some sub-panels found the timescale for assessing impact (usually within a two month period) to be tight and in future it would be helpful to extend this.

312. The assessments were made by the sub-panels, overseen by the main panels, as described in their [overview reports](#). The REF manager's report need not repeat such information or comment on issues that the panels were far better placed to reflect on. The observations below, therefore, focus largely on the operational aspects of the assessment process, which were of direct concern to the REF team.

Outputs assessment

313. The main workload challenge for panels was to review the large volume of outputs in the time available, and they therefore wanted to make as early a start as possible. This required several processes to be completed rapidly at the start of the assessment phase, as follows.

314. The REF team supplied the chairs (and deputies) with full details of submitted outputs during December 2013, through the assessment systems. Sub-panel chairs (often with substantial input from their deputies) used the spreadsheets to allocate outputs to members and assessors. The spreadsheets enabled them to do this efficiently and according to their particular working practices. Some sub-panels used straightforward processes (for example, the Computer Science and Informatics sub-panel had collected the subject classification for every output, which directly informed the allocations); others relied on manually reviewing the topic of each output, or involved many members in the process.

315. Efficient mechanisms to avoid major conflicts of interest during the allocation process were in place. Each panel member declared their major conflicts on the PMW; the spreadsheets used by chairs and deputies to allocate work prevented any item from being allocated to a member with a major conflict. (Minor conflicts of interest were dealt with on an ad hoc basis.)

316. As soon as the work had been allocated, panellists were instructed to generate their personal spreadsheet on the PMW, which listed the outputs they were allocated and provided direct links to the available PDFs. Their 'reading lists' on the PMW also went live, from which they could start ordering physical outputs.

317. During the process of allocation (and subsequently on the advice of individual members), sub-panels identified which outputs should be cross-referred to other sub-panels for advice. The assessment system generated spreadsheets listing the outputs identified for cross-referral to each sub-panel, which the chairs of the advising sub-panels used to specify which individuals on the sub-panel would provide the advice. After this, the REF manager approved the cross-referrals and the system generated a REF webmail message to invite the individual to provide advice on each output. If a PDF was available, it was attached to the webmail; if not the output was added to their 'reading list' so they could order it in the normal way. After reviewing the output they provided advice by replying to the webmail message.

318. In the event we found that a larger number of cross-referrals were requested than we had anticipated. In total, 5,248 outputs were cross-referred between sub-panels, with a significant amount of traffic across, as well as within, the four main panel areas. The panel [overview reports](#) provide further details at UOA level. Our systems generated a separate webmail message for each output being cross-referred. This proved to be efficient and effective in dealing with small numbers of referrals. However, there were also some large 'batches' of cross-referrals (especially from SP9 to SP18) which either generated excessive volumes of webmail messages, or were avoided by using workarounds. In future, cross-referrals could be managed more efficiently if the system could be designed to deal with 'batches' as well as individual items.

319. In order that calibration exercises could be conducted early on, each main panel developed a broad plan well in advance (at the early 2013 meetings). All main and sub-panels

undertook a systematic calibration exercise for output either in December 2013 or January 2014. These were very valuable in establishing comparable standards of assessment. The panel [overview reports](#) provide further details.

320. Each main panel established targets for the proportion of outputs that each of its sub-panels should review, at various interim points. By establishing these targets and using the data that was recorded on the assessment systems, the main panels were able to regularly review progress and the emerging patterns of outcomes at UOA level.

321. Individual panel members in turn were set targets for completing their work by the sub-panel chairs. They used their personal spreadsheets to access materials and record their scores and comments, and could generate reports to identify which outputs they needed to discuss with co-reviewers to reach a consensus. When surveyed, 73 per cent of panellists rated the spreadsheets as very good or good for this purpose, and a further 19 per cent satisfactory.

322. Panels succeeded in reviewing all outputs on schedule. However, for many the workload was heavy and for some it was excessively so. When surveyed in mid-2014 (some had completed reviewing outputs by this point but many had not), 18 per cent reported the workload was excessively heavy; 75 per cent reported it was heavy but manageable; and 7 per cent (mostly the impact assessors) that it was not particularly heavy. We also asked how many days (equivalent) they had spent since December 2013 reviewing materials (excluding meeting days); and about the extent to which their institutions had made allowances for their time spent on REF work:

Number of days spent reviewing	% of respondents
1-10	15%
11-30	35%
31-50	28%
51 or more	22%

Allowances made by employing institution	% of respondents
Majority of time required	19%
Some allowance made	37%
No allowance made	29%
Not applicable	14%

323. Given the heavy workload and very variable allowances made by institutions, we recommend that careful consideration should be given in future to how the panellists' workload issues might be addressed. It should be noted that at an early stage in the discussions we had invited sub-panels to consider whether they could sample the outputs rather than review them all. Sub-panels universally rejected this suggestion, but in the light of the workload issues the possibility of sampling ought to be reconsidered in future.

Impact assessment

324. Given that impact was new to the panels, the REF team developed a proposed workplan for the panels based on our experience of managing the pilot exercise, which they generally adopted. The process involved:

- Briefing the sub-panels to remind them of the criteria and ensure clarity and consistency in interpreting them. The REF team produced briefing materials which were tailored as appropriate and delivered by the panel advisers, to each sub-panel.

- The sub-panel chairs and deputies allocated impact case studies and templates to members and assessors, after some discussion at main panel level to ensure comparable approaches.
- Sub-panels were invited to identify audit queries, as discussed in section 6.2.
- Each main panel and its sub-panels carried out calibration exercises, through which all members of the main or sub-panel reviewed and discussed a common sample of case studies and impact templates.
- Sub-panel members and impact assessors reviewed the items allocated to them and discussed their scores before and/or during meetings to reach consensus on each item; these scores were then reviewed by the sub-panel as a whole and proposed to the main panel.

325. Given the much smaller volume of material to review for impact compared to outputs, and the intention to limit the call on users' time to the minimum necessary, two rounds of meetings (in addition to the briefing events in 2013) had been scheduled for the impact assessment. All sub-panels completed their assessments on schedule and although the volume of reviewing was modest, a significant amount of discussion and meeting time was required for familiarisation and to establish consensus on scores. By the end of the process the sub-panels appeared to reach high degrees of consensus and confidence in the outcomes.

326. Also, for some sub-panels the timetable for assessing impact (usually within a two month period) proved to be tight. This was in relation to preparing fully for the first round of meetings (for which panellists were expected to have read the briefing document, reviewed the calibration sample and 'skim read' their allocated case studies to identify potential audit queries) and for raising and completing audit queries in time for the second round of meetings. Some sub-panels anticipated this challenge and added an additional short meeting earlier on in the process for an early briefing to prepare members for these tasks. This was found to be helpful and would be worth planning in from the outset in a future exercise, although it would involve in additional meeting for users.

Environment assessment

327. As we expected the general process for assessing environment to be similar to the RAE, the REF team gave less attention to planning the environment assessment, other than identifying the timing (usually in the latter stages of the assessment) and number of meeting days required. As the environment assessment approached, the REF team encouraged the main panels to carry out a calibration exercise, especially as environment was typically assessed in a single round of meetings with limited opportunity for the main panels to review interim results and inform the sub-panels' final assessments. The main and sub-panels carried out calibration exercises but these were smaller scale than for outputs and impact and in future it would be beneficial to plan these in more systematically at an earlier stage.

8.3. Roles and responsibilities

Key points

- The main panels were highly effective in providing oversight of the sub-panels' assessments.
- The sub-panel chairs made a very substantial personal commitment to the exercise and were pivotal in its successful conduct.
- The sub-panels were effectively supported by a range of roles:
 - The role of 'impact assessor' was created to involve a wide range of users in assessing impact, with the minimum necessary time commitment. This worked well: their input was highly valued and they felt well integrated in the assessment.
 - Output assessors made a large contribution and successfully provided the panels with sufficient breadth and depth of expertise to review outputs.
 - Specialist advisers were effective in reviewing outputs in a range of languages that the sub-panels were unable to assess.

Recommendations

- Sub-panel deputy chairs often made a much more substantial contribution than originally anticipated. In future their role could be more fully specified at an earlier stage.
- International members on each main panel played a vital role in providing assurance about assessment standards and in supporting overall governance of the process. In practice their levels of involvement varied from observing meetings to participating actively in discussions. There is scope for more consistent – and more widespread – involvement of international members in future.
- Main panel users also played an important role in oversight of the impact assessment, although their levels of individual involvement also varied. Given the oversight role, it would be beneficial for the next REF to recruit main panel user members who have experience of the 2014 exercise.

328. The key roles and responsibilities of the main and sub-panels, and the roles of the members and assessors, were set out in the 2010 [publication](#) that announced the UOA structure. There were a wider range of distinct roles than in the RAE, and although the key responsibilities had been established at an early stage, the specific roles, levels of involvement and tasks associated with different categories of participant were to some extent worked out in practice in the lead up to and during the course of the assessment.

329. The basic role of each sub-panel was to assess the submissions made in its UOA and recommend the outcomes to the main panel. Each sub-panel included people with the following roles:

- a. A chair, whose role in chairing and leading the sub-panels' work involved a variety of activities. These included: ensuring there was an appropriate breadth and depth of expertise available to review submissions; working with the panel secretariat to plan all the work of the sub-panel, allocating work to members and assessors and monitoring

progress; setting agendas and chairing the meetings; guiding discussions, feeding in decisions and advice from the main panel; ensuring (with support from the panel adviser) that the published criteria and procedures were adhered to at all stages; requesting and accepting cross-referrals to and from other sub-panels; taking decisions on a range of detailed issues raised by members and assessors; ensuring consistent standards of assessment within the sub-panel (including keeping panellists' scoring under review); producing the sub-panel's section of the overview report and overseeing the production of feedback statements to institutions. In anticipation of the workload involved, many (but not all) chairs decided not to take on a full load in reviewing outputs. This was beneficial, particularly as it allowed them some capacity to act as an additional reviewer on outputs where consensus was not reached.

b. A deputy chair, who was chosen by the sub-panel from amongst the members during the criteria setting phase. In some cases two deputies were appointed. Their role was initially defined narrowly – in terms of deputising when the chair held a conflict of interest, or was unable to attend a meeting. As the assessment began, the deputy chairs' roles expanded. They were included in the 'executive group' for each sub-panel, initially so they could allocate work from the chairs' institutions, but many played a much larger role working very closely with the chairs and sharing many of the activities outlined above. The extent of their input was often not anticipated at the time when the deputies were selected, and in future the role could be more fully specified at an earlier stage.

c. Panel members, who played a full role in assessing all elements of submissions (outputs, impact and environment), and attended all rounds of meetings. Members also led on the production of the feedback statements to institutions. In some sub-panels, members were also assigned a co-ordinating role for individual submissions, which often eased pressure on the chair and deputy, and was helpful when it came to producing the feedback statements. A small number of the members on each sub-panel were research users, appointed at the outset so they help develop the panel criteria. During the assessment phase many of the research user members focused on assessing the impact element but some, where they had the necessary expertise, also assessed outputs and/or environment.

d. Output assessors, who were appointed to cover any substantive gaps in expertise and/or assist in dealing with large volumes of outputs in particular areas. They played a full role – usually with an equivalent workload to members – in assessing outputs and attended all meetings where outputs were discussed. The creation of this role for the REF was essentially an upgrade of the previous specialist adviser role, following feedback from RAE panels that reviewing of outputs should be more integrated into the work of the panels rather than 'outsourced'. The output assessor role succeeded in providing the panels with the breadth and depth of expertise they needed to review outputs, in a more integrated way. A few sub-panels were wary of the distinction made between full members and output assessors, and tried at times to include assessors in additional activities and meetings. The REF team upheld the view that members and assessors should each fulfil the roles for which they had been appointed, in accordance with the published specification of these roles. In a very small number of cases, where an output assessor's expertise was needed

to review impacts and environment as well, their appointment was changed to that of a member.

e. Impact assessors, who were appointed to assess the impact element alongside the panel members, on a full and equal basis. They were each allocated a full load of case studies and impact templates, and attended all meetings where impact was discussed. This new role was defined specifically so that most of the users involved on panels could take part in a very focused way with the minimum necessary time commitment, that is, by assessing only the impact element. This worked well, although it did require much more careful planning of the assessment schedules at an early stage, to identify exactly which days of which meetings the impact assessors would need to attend. When surveyed, the impact assessors generally fed back that their workload was reasonable and they felt well integrated into the assessment process, although they did raise concerns that the fee levels did not adequately compensate for their time. The sub-panel members valued their views and input, and there was a genuine sense that the assessment of impact was a joint enterprise between the users and academic members. For some impact assessors, their involvement could have been simplified a little further by clarifying at an earlier stage that academic members would largely deal with the threshold decisions (relating to the underpinning research), with users focusing their attention on the grading of the impact.

f. A panel advisor and a panel secretary. These roles are discussed in section 8.5.

330. The key roles of each main panel during the assessment were to guide the sub-panels in carrying out their work, ensure the criteria were adhered to and assessment standards applied on a consistent basis, and sign-off the outcomes. Each main panel included the following roles:

a. A chair, whose role was to oversee the work of the main panel. They worked closely with the panel advisers to plan the main panel's business; chaired meetings and guided discussions on a variety of issues relating to procedures and decision-making; built consensus and collective decision-making amongst sub-panel chairs; monitored emerging outcomes at UOA level and raised issues as necessary; attended a variety of sub-panel meetings; steered the activities of the main panel user and international members; and led on the production of the main panel overview reports.

b. A deputy chair, chosen from amongst the members of the main panel. Their role was to deputise where the chair held a conflict of interest or could not attend a meeting, and in some cases to assist the chair with handling particular issues. In practice, as the main panels rarely discussed individual submissions, there were few instances where the deputy was called on due to a chair's conflict of interest.

c. The chair of each sub-panel within the main panel's remit. Their role on the main panel was both to represent their sub-panel (providing reports and feedback, and presenting recommendations), and to act as a member of the main panel in making collective decisions. This included collective responsibility for signing-off the results for all submissions within the main panel. At times, there appeared to be a tendency for sub-panel chairs' input to focus more on representing their sub-panel, which is unsurprising.

Given this, it was valuable to also include a number of members on the main panel who were independent of any specific sub-panel.

d. International members on each main panel played a vital role both in providing assurance about standards of assessment (especially for outputs, though they also tended to be keenly interested in impact and environment), as well as contributing as main panel members to the general governance of the process. International members were assigned to 'clusters' of sub-panels and most of them attended numerous sub-panel meetings as well as the main panel meetings. In practice their engagement varied somewhat from observing meetings to participating actively in calibration exercises and discussions throughout the exercise. There is scope for more consistent – and more widespread – use of international members in future. However, the costs and associated with a significant number of international members who would require more than short-haul flights would be substantial.

e. Main panel user members on each main were important in helping provide governance of the impact assessment. Many were actively engaged in the main panel calibrations and participating in sub-panel discussions of impact. Some, given their seniority, found it difficult to commit much time beyond participating in the main panel meetings. There is scope to make this role more consistent in future, and it would be beneficial, given the oversight role, to recruit individuals who already have experience of assessing impact in the 2014 REF.

f. Observers from the Research Councils, who valued the experience of observing the proceedings and were called upon to comment and provide information at various points.

g. Three panel advisers. This role is discussed in section 8.5.

331. Because of the dual role held by sub-panel chairs, and the substantial responsibilities and activities involved in chairing a sub-panel, they played a pivotal role throughout the assessment. The successful conduct of the exercise relied very heavily on their substantial personal commitment as well as expertise and skills.

332. Overall, the main panels proved to be very effective in guiding and providing oversight of the sub-panels' assessments. They met in between each round of sub-panel meetings, receiving reports on progress and discussing issues at each stage, feeding into the subsequent round of sub-panel meetings. The assessment systems enabled them to have thorough discussions at each stage about the emerging profiles, reaching an understanding and justifying any substantial differences in outcomes at UOA level.

333. In addition to the roles outlined above, specialist advisers assisted specifically with outputs in languages (or in specialist forms of notation) that were not assessable by the sub-panel members and assessors. This was deliberately a much more constrained role than the specialist advisers used in the 2008 RAE, following feedback from the panels. The original intention was to limit the role to outputs in languages other than English. A few sub-panels requested that the role be extended to also advise on outputs that the panels were unable to assess for other reasons,

such as highly specialist notation. In the event, many of the non-English outputs were assessed by the sub-panels and through cross-referral, and few requests were made for advice on outputs in specialist forms of notation. Overall the role of specialist adviser was an effective way of dealing with outputs submitted in a wide variety of languages, albeit in comparatively small numbers.

334. It should be noted that the much more restricted role of specialist advisers in REF probably contributed to the increase in cross-referral since the RAE. In the RAE, sub-panels made use of many specialist advisers – and sought specialist advice on a total of 11,769 outputs – as well as cross-referring a total of 3,407 outputs. In the REF, specialist advice was sought on fewer than 200 outputs, and a total of 5,248 outputs were cross-referred.

335. The roles of the REF team in supporting the assessment were as follows (a number of which are discussed in more detail in the sections below):

- Providing administrative support
- Recruiting, training and deploying the panel secretariat
- Distributing submissions and outputs
- Providing IT systems and user support
- Auditing submissions and providing the outcomes to panels
- Approving requests for cross-referrals and specialist advice
- Provide guidance to panels on policies and procedures
- Working with the main panels and their chairs to monitor overall progress and promote consistency across the four main panels
- Preparing the results for publication.

8.4. Administrative support

Key points

- The panel members' website (PMW) provided an efficient and effective means for members to supply and update their details, access meeting information, book accommodation and claim expenses.
- The REF admin team dealt with large volumes of queries and provided high quality administrative support: 97 per cent of members rated the support as good or very good.
- Large volumes of accommodation bookings and amendments were made. Panel members were generally satisfied with the accommodation, although there were constraints due to members making late requests, and the high cost of hotels in London.

Recommendations

- Fee levels were generally felt to be low and 30 per cent of users reported they could be a potential barrier to future participation. We recommend the funding bodies review the fee levels, especially for research users.
- Improvements could be made to the fee information provided on the PMW.
- Venues with overly strict policies for cancelling block-booked accommodation should be avoided in future.

336. The REF administrative team comprised the REF administration manager supported by two panel administrators and a team administrator. Additional temporary staff were appointed at peak periods of activity (such as loading the USB pens for panel members). Administrative processes were supported by the PMW, which was developed initially in the autumn of 2010 for the criteria phase, and then enhanced for the assessment phase. The PMW provided a portal for panellists to manage all aspects of their work, some of which had been dealt with by post or email in previous exercises. The PMW enabled more streamlined administrative processes and enabled us to operate with a smaller admin team than in the 2008 RAE.

337. Panellists did the following via the PMW:

- Entered and updated their personal and membership details (including contact details, payment details, statements of their expertise and declarations of major conflicts of interest)
- Accessed all required information about meetings (dates, locations, agendas and papers) and requested accommodation
- Claimed expenses and received information about fees
- Accessed general information and guidance
- Once the assessment phase started, accessed the submissions and the assessment systems (their reading list, personal spreadsheets and reports).

338. The initial requirements for the PMW for the criteria phase were developed by the REF admin manager in 2010 and developed by IT staff, to be tested and in place for the criteria phase in 2011. The initial system was largely effective but as the exercise progressed a number of improvements were required and these were implemented during 2013, in time for the assessment phase. For example, the system needed to be more flexible in allowing individuals to be invited to specific days of particular meetings (given the assessor roles); and it needed to hold PDF confirmations of accommodation bookings.

339. Prior to the assessment phase the admin team had provided administrative support for earlier rounds of meetings, scheduled and booked meeting venues for the assessment phase, worked with IT staff to enhance the administrative systems, and prepared for distributing the USB pens. During the assessment phase, administrative support focused on processing panellists' fees and expense claims, booking overnight accommodation, liaising with meeting venues, and dealing with queries from panel members.

Fees and expenses

340. Sub-panel members and assessors (except those employed by government organisations) were entitled to a flat-rate attendance fee of £200 for each day they attended a REF panel meeting; the rate for main panel members was £300 per day. During the assessment phase sub-panel members and assessors were entitled to an additional flat-rate preparation fee of £250 per meeting. This fee was in recognition of the work undertaken to review the submissions in preparation for each meeting, and replaced the honorarium payment made at the end of the 2008 RAE. The reason for this change was to scale the payments according to the volume of preparation work, given the large differences in workload between members, output and impact assessors.

341. Of the 1,105 panel members and assessors undertaking roles involving fees, 50 (4.5 per cent) declared themselves ineligible to receive fees, of which 7 (0.6 per cent) declared themselves ineligible to receive fees or expenses. Of the remaining members and assessors, the large majority chose to receive direct payment of fees and a small minority (around 10 per cent) chose to invoice for their fees through their employing organisation (or as a self-employed company).

342. The payment of fees via direct payment was efficient for the REF admin team (which calculated fees according to meeting attendance recorded by the panel secretaries) and for the HEFCE finance team to process payments. However, a large volume of queries were received toward the end of each financial year regarding payment for tax declaration purposes. The processing of invoices by both the REF admin team and the finance team was more time consuming as often there would be queries and errors in the amounts being invoiced. The system could be further improved in future by incorporating facilities for panel members receiving direct payment to see how much they have been paid overall in each financial year, and for panel members paid through invoicing to be able to see how much they are eligible to invoice based on attendance information entered by the REF admin team.

343. Specialist advisers were paid fees on a sliding scale relating to the number of outputs assessed.

344. Panel members' responses to the survey indicated they generally felt fees to be low but not prohibitive. However, 30 per cent of user members and impact assessors considered the low fees would be a potential barrier to future participation (for academic members and output assessors, this figure was 15 per cent). We recommend that the fee levels should be reviewed for a future exercise, particularly for users.

345. Travel and expenses claims were submitted by panel members online via the PMW and processed by the panel administrators. Claims were made in line with guidelines that were comparable to those in place for HEFCE staff and HEFCE board members. They encouraged the use of public transport and limits were in place for subsistence. The guidelines had initially been drafted by the admin manager during the criteria phase and in response to the claims and queries received during the early 2013 panel meetings, they were updated to provide additional clarity and flexibility for those panel members that preferred to travel long distances on some of the meeting days and avoid overnight stays.

346. Across the criteria and submissions phase meetings, the admin team processed over 8,000 expense claims. Although we endeavoured to make payments within two weeks of claims this proved to be unrealistic, as payment runs were only made every two weeks and some time was required to check and approve claims before adding them to the payment runs. Nevertheless, panel members were satisfied with the process for claiming and receiving expenses and fee payments (85 per cent rated it as good or very good).

Accommodation bookings

347. Hotel accommodation was requested by individual panel members and the secretariat via the PMW. The REF admin team forwarded the requests to the travel agent used more generally by HEFCE, then received and uploaded confirmation PDFs back to the PMW. Over the course of the assessment period 5,718 hotel bookings were made, including 908 booking amendments and 238 cancellations. This was a considerable resource requirement on the REF admin team throughout the year, often processing high volumes of requests in short periods of time. Panel members making late requests shortly before meetings exacerbated this and often resulted in inferior accommodation. Part way through the year the admin team sought to address this by proactively communicating with members well in advance of meetings to remind them about accommodation requests.

348. When booking the meeting venues, many of which had on-site accommodation, we negotiated as far as possible on their cancellation policies. However, some venues still had strict policies that required early release of block held rooms, which were difficult for us to book accurately when members did not make requests well in advance of meetings. Where there were strict cancellation policies combined with limited on-site room capacity and a lack of nearby alternative hotels, some problems were encountered in finding suitable accommodation for all members. We recommend avoiding such venues in future.

349. In addition to the issues above the high cost of hotel accommodation in London often constrained the choices available within the budget (especially when late requests were made). Overall, panellists were generally satisfied with the quality of accommodation but reported some variability. In the survey 58 per cent rated the accommodation as good or very good, and a further 36 per cent as satisfactory.

Enquiries and user support

350. The REF admin team were the first point of query for all panellists regarding any administrative matters as well as IT and systems issues. They provided first-line user support on IT issues, referring issues as necessary to IT staff for second-line support.

351. The team consistently dealt with high volumes of queries throughout both the submission and assessment process, with particular peaks related to the release of new parts of the assessment systems. The first significant instance was in October 2013 when the REF webmail was rolled out to all panel members. As the REF IT user support team were being fully occupied by providing user support for the submission system, primary support of the webmail system was provided by the REF admin team. This generated high volumes of queries relating to synchronising log-in details for the PMW and webmail systems.

352. From December 2013 to February 2014 as panel members familiarised themselves with the assessment systems, additional support was provided, with members of the IT and admin teams covering Saturdays on a rota. This received very positive feedback from the panel members making contact.

353. Overall, the panel members' survey indicated that panellists were highly satisfied with the administrative support provided by the REF admin team, with 97 rating it as good or very good. The administrative systems also received positive feedback:

- 94 per cent rated the PMW as good or very good as a means of accessing meeting information
- 98 per cent rated the process of booking accommodation as good or very good
- 85 per cent rated the process of claiming fees and expenses as good or very good.

8.5. Panel secretariat

Key points

- The panel secretaries and advisers were seconded from HEIs to provide committee servicing support and advice to the main and sub-panels. They brought a wealth of invaluable skills and expertise, with many having managed their institutions' REF submissions.
- The secretariat provided excellent levels of support which was greatly appreciated by the panels, and was vital to the effective conduct of the assessment.
- Assigning advisers and secretaries to work in pairs, with each pair supporting two-three sub-panels was effective. The workload for the secretariat was heavy but manageable.

Recommendation

- When scheduling panel meetings in future, the workload implications for the secretariat at peak periods should be better taken into account.

Recruitment and deployment

354. For each of the criteria and assessment phases, we recruited a team of panel secretaries and panel advisers, to provide committee servicing support and advice to the panels. They were recruited on secondment from HEIs, although applications were openly advertised and welcomed from individuals based at any organisation that could support the secondment.

355. Key criteria for appointing advisers included a thorough knowledge and understanding of the REF criteria and procedures, and experience in successfully providing guidance to academic committees. The advisers we appointed typically had substantial responsibilities in managing their institutions' REF submissions. Secretaries were required to have a range of skills in servicing academic committees; many were appointed with institutional REF experience although some came from other administrative or professional service backgrounds within their HEIs. The collective skills and experience that the panel secretariat brought with them proved to be a tremendous benefit to the panels and to the overall conduct of the assessment.

356. For the criteria phase, we arranged for all the sub-panels within each main panel to meet simultaneously in a conference format with plenary sessions and break-out sessions for the sub-panels. We therefore recruited 11 secretaries, so that one could support each sub-panel (11 being the maximum number of sub-panels within any one main panel). We also recruited eight advisers (two for each main panel). As well as supporting meetings, a key task of the secretariat

in this phase was to analyse responses to the consultation on the draft panel criteria. The advisers also helped with drafting and refining the criteria documents. Overall this worked effectively although it was more challenging for the secretaries to work across the main panels without being paired with an adviser, as they were in the later stages.

357. For the round of planning meetings in early 2013, many of the secretaries returned supplemented by REF team members, to support the panels (including the significant tasks of analysing responses to the survey of submission intentions, and helping the panels to gain suitable nominations for assessors).

358. For the assessment phase a larger cohort of advisers and secretaries was required, with more substantial time commitments. We decided to appoint a team of 26 in total: 13 full-time secretaries and 13 part-time advisers. Seven of the earlier team members continued with the assessment phase, requiring 19 new appointments for the assessment phase.

359. Each main panel and its sub-panels was assigned a team of three advisers and three secretaries. Within each team, an adviser and a secretary worked as a pair to support three sub-panels within the main panel remit (two within the Main Panel A remit). This gave them a more even workload than in the RAE, when each pair supported one of the 15 main panels regardless of the number of UOAs within that main panel. The Main Panel A secretariat were also expected to provide cover support for other sub-panels where required (for example, due to sickness) but in the event this was rarely required.

360. At the time of scheduling the assessment phase panel meetings (before we had appointed the secretariat), sub-panels were grouped together in clusters which often met simultaneously. We therefore assigned each adviser and secretary pair to work across the clusters, and scheduled all the cluster meetings to avoid clashes, so that the secretariat would be available to service all meetings of the sub-panels they had been assigned to.

Training and co-ordination

361. Having recruited and deployed the secretariat, the REF team ran a programme of events to train the secretariat, with specific sessions focused on:

- Orientation (before the start of the assessment)
- IT systems (alongside sub-panel chairs and deputies)
- Meeting and assessment procedures (such as agendas, papers and minutes, conflicts of interest, audit procedures)
- Individual staff circumstances
- Impact assessment
- Feedback statements and overview reports.

362. In addition to specific training by the REF team, meetings with the secretariat devoted time to sharing ideas and information, co-ordinated work across the four main panels, and time for detailed planning within each main panel team. Some meetings were also held specifically for the four main panel chairs, the panel advisers and REF team to discuss progress and issues across the exercise (see section 8.7).

363. The secretariat fed back that they found the training and information sharing meetings extremely valuable, although would have appreciated some additional training in the early stages, especially regarding the IT systems. In addition, there were many ad hoc meetings amongst the secretariat which were facilitated by the clustering of meetings where several sub-panels were meeting on the same date in the same location.

Work of the secretariat

364. Throughout the assessment phase the secretariat supported the work of their panels, working very closely with their chairs and deputies, liaising regularly with each other and the REF team, and dealing directly with panel members and assessors at all stages.

365. Enhancements to the IT systems to support the assessment phase meant that the secretariat spent less time than in the RAE on clerical tasks (such as ordering outputs on behalf of panel members or manually updating scores). Instead, we expanded their advisory role, especially in reviewing the individual staff circumstances on behalf of the sub-panels (discussed in the next section). Key tasks and roles of the secretariat included:

- Helping to plan the work of the sub-panels and main panels
- Helping chairs and deputies to allocate work to panellists
- Monitoring progress and producing and presenting reports
- Writing and circulating meeting agendas, papers and minutes
- Raising audit queries with the REF audit team and communicating the outcomes to panellists
- Reviewing all clearly defined staff circumstances and making recommendations to the panels
- Following up cross-referral and specialist advice procedures
- Providing briefings and advice to panels on a range of issues and procedures
- Dealing with a range of individual queries and following up with the REF team where appropriate
- Supporting panels' use of the assessment systems
- Keeping accurate and complete records of all scores and decisions
- Collating feedback statements and contributing to the overview reports.

366. Throughout the assessment, the secretariat provided excellent levels of support to the panels, which was critical to the successful conduct of the exercise. When surveyed, 84 per cent of panellists rated the secretariat support as very good, and a further 14 per cent as good. At the end of the exercise the sub-panels expressed emphatic appreciation for the secretariat, especially their dedication, professionalism and good humour.

367. At the end of the assessment we held a review meeting with the secretariat. Their feedback concurred with a number of the points and recommendations already made in this report relating to the assessment. In terms of their own work, overall they were positive about the experience; the pairing of secretaries and advisers each assigned to two-three sub-panels worked well; and they found the clustering of meetings useful so they could meet regularly with the co-secretariat around the panel meetings. They considered the workload was heavy but

manageable, although the panel advisers felt the role realistically took up more than 0.5 FTE, and there were some significant ‘pinch points’. Especially when their sub-panels had multi-day meetings one after the other, the secretariat spent long periods away from home and worked long hours not just to attend the meetings but to keep up with the preparation and follow-up activities for their sub-panels. This challenge was not fully appreciated at the time of scheduling the meetings and should be better taken into account in future.

8.6. Individual staff circumstances review

Key points

- All 16,361 individual staff circumstance cases were reviewed to check that the criteria for output reductions had been applied correctly. This was a significant task for the panel secretariat (who reviewed clearly defined circumstances) and EDAP (who reviewed complex circumstances).
- Centralising the review of complex circumstances was undoubtedly more effective than dealing with them at sub-panel level.
- Although a significant number of cases required further information through audits, the submitted cases overwhelmingly met the criteria.

Recommendation

- The need for auditing in future could be reduced by developing a template to clarify what information is required for particular circumstances, and further guidance given on the appropriate level of detail to be provided.

368. During the assessment, all individual staff circumstance cases were reviewed to determine whether the criteria for output reductions had been applied correctly. This was a significant task: 15,329 staff were submitted with clearly defined circumstances, and a further 1,032 with complex circumstances. EDAP reviewed all complex circumstances; this is discussed in the [EDAP final report](#) and the information need not be repeated here. The panel secretariat reviewed all the clearly defined circumstances and made recommendations on output reductions to their sub-panels, who were responsible for these decisions.

369. The role of the secretariat was to check that, based on the information provided in the staff circumstance statements, the output reductions had been applied and calculated correctly. The REF team provided training for the secretariat as well as ongoing advice and guidance on any specific issues that arose, to ensure consistency in how the circumstances were reviewed. Given the large volume of circumstances, and the expectation that the vast majority would meet the criteria, we discussed with the panel secretariat options for sampling. The secretariat initially reviewed one third of the cases, after which we discussed the outcomes with them and collectively decided they should review all staff circumstance statements.

370. Many of the circumstances were straightforward to check, however some difficulties were encountered:

- a. Early on, panel secretaries reported that some submissions repeatedly lacked sufficient information in their statements for them to verify that the criteria had been applied

and calculated correctly. We collated this feedback from the secretaries and identified nine institutions had repeatedly provided insufficient information across their submissions. The audit team raised 'bulk audits' to request they check all their clearly defined circumstances statements and expand them wherever necessary to provide the required information. Later on in the review, six additional HEIs were identified as fairly frequently providing minimal information. We audited a sample of statements submitted by these HEIs.

b. The criteria for ECRs involved a degree of interpretation to decide the point at which an individual first undertook independent research. Where required, the secretariat sought discipline specific advice from their sub-panels on this point. However, the issue was often further complicated where little information had been provided in the statement to explain on what basis the HEI had determined when the individual began undertaking independent research. This led to a number of audit queries to request further information.

c. There was some variability amongst HEIs in interpreting the criteria when an individual had carried out research during a career break. Some HEIs were requested to resubmit career break calculations taking into account the time spent carrying out research.

371. Leaving aside the nine institutions that were bulk audited, we audited 6 per cent of staff submitted with individual staff circumstances, often to request further information relating to ECR status. Having requested any further information that was needed and checked the clearly defined circumstances, the secretariat and sub-panels found that the overwhelming majority of output reductions had met the criteria. A very small percentage (0.38 per cent) did not, resulting in one or more unclassified outputs.

372. Towards the end of the exercise we decided in discussion with EDAP to provide each institution with confidential feedback on the extent to which it had adhered to the staff circumstances guidelines, to help institutions learn from the experience. The feedback was generated by the audit team, based on the frequency of audits and unclassified outputs resulting from staff circumstances at institution level, in comparison to the sector average, although was often constrained by the small numbers involved.

373. Centralising the review of complex circumstances was undoubtedly more effective than dealing with them at sub-panel level, especially in terms of ensuring consistent treatment. We recommend that this approach be repeated in a future exercise. Given that a significant number of audits were required due to insufficient information in the submissions, we and the secretariat recommend developing a template to clarify what information is required for particular circumstances, and further guidance given on the appropriate level of detail to be provided.

8.7. Policy support and guidance

Key points

- To promote consistency, the REF team produced guidance for panels and worked with the panel secretariat to brief them on key procedures and issues.
- The REF manager and deputy attended and contributed to all main panel meetings, observing progress and helping to share issues and approaches across the four main panels.
- Regular meetings of the four main panel chairs and advisers were essential in identifying and addressing cross-main panel issues, and reviewing emerging profiles across the four main panels.
- An exercise was carried out to test the scoring of impact across the four main panels; it provided assurance that impacts had been judged comparably across the main panels.

Recommendations

- Experience of the REF shows that a calibration of impacts across a wide range of disciplines is feasible, and in future it may be beneficial to carry out a cross-main panel calibration of impact at an earlier stage.
- The REF team brokered agreement on the use of a single grading scheme across all panels. It may be beneficial in future to build on this by exploring the potential benefits and feasibility of more standardised processes for resolving differences in individual panellists' scores.

374. A key issue for the REF policy team was to promote consistency across the exercise both in terms of the application of the assessment procedures and criteria, and in terms of comparable standards of assessment. Our guiding principle was that differences in outcomes at UOA level should not be the result of differences in the practices of panels; they ought to reflect differences in the strength of submissions. Throughout the assessment, the REF team provided policy support and guidance to the panels and observed and co-ordinated developments across the main panels. We did this in a number of ways, including providing generic guidance on procedures and generic meeting papers on key issues; attending and contributing to all main panel meetings; working with the four main panel chairs and advisers on cross-exercise issues; and initiating a cross-main panel calibration exercise for impact.

375. We provided the panels with generic guidance documents on key procedures such as managing conflicts of interest, audit procedures, and cross-referral and specialist advice procedures. On some issues we drafted generic meeting papers which the secretariat adapted and used as appropriate (for example, on the use of citation and environment data, recommendations on staff circumstances, and the production of feedback statements and overview reports). To support the impact assessment, in addition to a generic meeting paper to explain the procedures and criteria, we provided a briefing presentation for the panel secretariat to adapt as appropriate and deliver at the panel meetings. These generic materials were efficient for the secretariat and ensured consistent guidance was provided to all panels.

376. During the planning stage in early 2013 we had worked with the main panels to plan their assessment schedules and put in place a range of processes at main panel level to maintain consistency. These included calibration exercises, ongoing monitoring and review of emerging outcomes (enabled by careful scheduling of meetings and design of the IT systems), and active engagement of main panel international and user members across sub-panels. The REF manager and deputy manager between them attended all main panel meetings, observing their progress, contributing to discussions and helping to share approaches across the main panels.

377. Also during the planning stage we identified that the grading scheme for recording scores required early resolution, to enable timely development of the assessment systems. We developed proposed schemes for recording scores for each type of item to be assessed (outputs, impact and environment). Initially, across the main panels there were different views on the appropriate level of granularity for scoring, but also agreement that consistency across panels would be important. We carried out some modelling on the effect of different schemes on the granularity of the resulting sub-profiles, after which all main panels agreed to use the same scheme. (This involved assigning an integer score from 0-4 to each output, and assigning a score to each impact case study, each impact template, and each section of the environment template on the scale 0; 0.5; 1; 1.5; 2; 2.5; 3; 3.5; 4. The 'half-marks' were to be assigned for borderline judgements.)

378. Some panels found it difficult to anticipate how these early decisions on the grading scheme would later affect the assessment, but they all made use of the agreed scheme. At the end of the assessment several sub-panels felt that more granularity for the environment element in particular would have been beneficial. While the same grading scheme was used by all panels, we decided through discussion with the four main panel chairs to retain some flexibility in terms of how individual panellists recorded their individual provisional scores, prior to agreeing scores. There was also flexibility in the procedures followed by the sub-panels for resolving differences in their individual provisional scores (for example, one-to-one discussions prior to meetings, discussions during meetings, use of algorithms, or referral to additional reviewers to arbitrate). It may be beneficial in future to explore the potential benefits and feasibility of more standardised processes across the panels for resolving scores.

379. During the assessment we met regularly with the four main panel chairs and advisers to discuss common issues. At appropriate points we compared the average profiles emerging within each main panel, for each of the three elements of assessment. The main difference that emerged during the course of the assessment was the significantly higher impact profile within Main Panel A. In response, the REF manager organised a cross-main panel impact calibration exercise, to test whether this difference was the result of differences in assessment standards, or was justified by the submissions; and to test more generally whether all four main panels were assessing impact in a comparable way.

380. The exercise involved a sample of case studies selected from one relatively high scoring and one relatively low scoring UOA within each of the four main panels, that the sub-panels concerned had assigned a range of scores. Cases were selected to include types of impact (such as health, environment and public understanding) that were common across the four main panels. The four main panel chairs plus two members from each main panel (mostly user

members and some internationals) reviewed and graded the sample of cases, and provide their reasons and levels of confidence in their assessments. The results of the exercise showed that:

- The reviewers were generally consistent in their gradings. Only a few cases received a wide spread of scores; and of the 12 reviewers, only one marked consistently higher, and one consistently lower than the rest.
- Overall the reviewers' scores were close to those of the sub-panels. There were no systematic differences between their and the sub-panels' scores for any of the four main panels. This provided reassurance that none of the main panels were assessing impact with generally different standards.

381. The exercise showed that it was very possible for individuals with some experience in assessing impact (as all the reviewers had by that stage in the exercise) to judge impacts reliably, beyond their immediate areas of expertise. With this experience, in a future exercise it could be helpful to carry out cross-main panel calibrations of impact at an earlier stage in the process – especially if there any significant changes to the impact requirements or guidelines.

8.8. Results and outcomes

Key points

- Results were published to schedule on 18 December 2014 without errors, and received widespread media attention. During the week of publication there were 86,888 visitors to the results website and 30,000 tweets using the REF2014 hashtag.
- Additional analysis was published alongside the results, which usefully addressed issues such as equality and diversity, interdisciplinary research, and improvement in quality since RAE2008.
- In January 2015 we published the submissions, overview reports and EDAP report, and provided confidential feedback statements to heads of HEIs.

Recommendation

- Substantial effort went into producing the feedback statements, to make them more useful to institutions than in the RAE. However, it was unclear whether the amount of effort was proportionate to the benefit gained by HEIs. We suggest the funding bodies ask institutions how useful they found the feedback and develop the future approach accordingly; and that the amount of effort involved could be reduced by developing a very tightly specified template for sub-panels to use.

Results and analysis

382. Key decisions on the nature and scope of the results, reports and feedback that would be provided at the end of the exercise had been taken in 2011, and were set out in the [guidance on submissions](#). More detailed decisions were taken by the Steering Group in early 2014, including the decision to publish sub-profiles at the same time as the overall profiles (given the expected level of interest in impact); the exact timing of release and publication of results; and the publication of results in hard-copy as well as on the web. The decision was taken to provide

institutions with slightly longer access to their results under embargo, prior to publication, than in the RAE.

383. A results website was developed by HEFCE ITS staff during early 2014. It was also to be used for publishing the submissions. The site allows users to browse results and submissions by UOA or by institution, and also provides download files.

384. In order to support HEIs in their preparations for the publication of results, we held a briefing event in London in November 2014. The event was attended by 200 delegates and included Q&As with the REF manager and HEFCE communications staff. A webinar-version of the briefing was also held to make the opportunity available to those who could not attend.

385. The four main panels signed-off all the results at their meetings in late October or early November 2014. From this point there was a tight turnaround to prepare them for publication in hardcopy (dispatched to arrive at HEIs on 18 December) as well as test the online publication and produce files for circulation, under embargo, to HEIs. A press conference was held on 17 December and the full results published on 18 December. The results were published on schedule without any known errors.

386. The publication date had been selected partly to coincide with HESA's publication of contextual REF data, which provided approximations of the number of eligible staff in each UOA at each HEI. However, as the HESA data were released to the media later than the REF results, the initial league tables were based entirely on the REF results; some subsequent league tables then incorporated the HESA data. Neither the REF team nor the funding bodies endorsed any particular form of league table.

387. The results generated substantial interest. Data from 'Google Analytics' indicates that during the week of the results publication (18 – 24 December) the main REF website (excluding the results pages) received 53,954 visitors and the results section received 86,888 visitors. On results day, 60,962 visitors to the site generated 666,302 page views.

388. This was the first exercise in which results also featured on social media. On 18 December the REF results trended on Twitter throughout the day and during that week there were over 30,000 tweets using the REF2014 hashtag. Over 160 online articles and blogs on the REF results were published by January 2015.

389. The REF team and colleagues in HEFCE's Analytical Services Group prepared additional analysis for publication alongside the results. This was facilitated by the assessment systems which allowed us to analyse results across the exercise in various ways. The [analyses](#) can be found on the REF website and were helpful to anticipate and address some key issues that were likely to be raised once the results were published, such as equality and diversity, interdisciplinary research, the degree of discrimination between HEIs, and the extent of improvement in profiles since RAE 2008.

390. In January we published the submissions on the REF website. Some difficulties were encountered with institutions requesting very late in the process that certain information should

be redacted from publication, and we discovered some PDFs had been poorly redacted. Early in the development of the website we had anticipated substantial interest in the impact case studies, but realised that publishing the submitted PDF documents alone would not be very useful for analysis. As a result HEFCE decided in early 2014 to commission work to provide the case studies, additionally, as a searchable database, which was planned for publication in March 2015.

Feedback statements and overview reports

391. During the assessment phase the sub-panels produced confidential feedback for each submission, which we provided in confidence to heads of HEIs. Our intention was to provide institutions with more helpful feedback than they had received in RAE2008 (which often consisted of a narrative description of the sub-profiles). We provided a broad template and some fictional examples at an early stage to panels, and encouraged them to highlight – for each element of the submission – the strongest aspects. Having consulted the REF Steering Group, we accepted that this might well include, if necessary, the identification of the strongest case studies, to help HEIs learn from this first assessment of impact.

392. Producing the feedback statements proved to involve a significant amount of work for the panels, the secretariat, and the REF team. Many of the sub-panels spent much time discussing the appropriate level of detail to provide, issues such as how or indeed whether to provide any negative feedback, concerns about identifying case studies, appropriate phrasing and so on. Discussions at and between main panels were also held to discourage the sub-panels from reaching different conclusions on these points. Initial drafting of each statement was carried out by panel members, who often produced initial drafts in different styles with different levels of detail, requiring a significant amount of iteration and often further detailed discussions at panel meetings. Most panels eventually produced feedback with broadly the level of detail we had intended, with some variations.

393. At the end of this process it was unclear to us whether the amount of effort involved was worth the benefit that institutions gained from the feedback. For the next exercise we suggest:

- the funding bodies ask institutions how useful they found the feedback, and develop the approach accordingly
- the amount of effort involved could be reduced by developing a very tightly specified template for sub-panels to use.

394. Each main panel produced an overview report, including sections from each sub-panel, which detailed how the assessment was carried out, provided an overview of the panels' observations about the state of research in the areas falling within its remit, and general reflections on the submissions and their assessment. The REF team provided a broad template early on, annotated with indications of the kind of information that would be helpful to include. We anticipated particular interest in the impact assessment and encouraged all sub-panels to include a reasonably full description of the range and nature of the impacts that had been submitted. The reports were mostly drafted by the main and sub-panel chairs and the advisers (who spent a significant amount of time working on them). The level of detail provided by sub-panels varied somewhat and was reviewed by the main panels. The four main panel chairs, the advisers and

REF team met during the drafting process to discuss the main panel sections of the reports and to ensure they all covered the relevant issues in appropriate levels of detail. The REF team supplied various data to each main panel to be included in their reports, on a consistent basis. The reports were [published](#) in January 2015, timed to coincide with the provision of feedback statements to HEIs and publication of the [EDAP final report](#).

9. Impact of the REF

9.1. Cost and regulatory impact

Key points

- The initial ambition in 2006 of greatly reducing the administrative burden of research assessment was no longer feasible, after a metrics alternative to peer review proved unacceptable to the sector.
- Within the context of a peer review exercise, efforts were made to streamline wherever possible, with notable success in establishing consistency across the exercise. However, the scope for aligning data requirements with other sources proved to be limited.
- It was apparent that in comparison to the 2008 RAE, significant additional work was carried out in institutions to prepare impact submissions, and to apply the more rigorous procedures to support equality and diversity. The cost to institutions cannot be estimated here, and an evaluation has been commissioned by HEFCE.
- The direct cost to the funding bodies of administering the REF was £14.4 million, a modest increase since RAE2008.

Recommendations

- A number of suggestions for refinements within the existing parameters of the exercise are made, which could in future achieve modest reductions in burden.
- To substantially reduce the burden associated with a future exercise, more radical changes would be needed. We suggest serious consideration of the feasibility of decoupling the selection of staff and outputs, although that may present challenges in identifying a robust volume measure for funding purposes.

395. As discussed in section 2, the initial ambition for reforming the funding and assessment framework had been to greatly reduce the administrative burden. After a metrics alternative to peer review proved unacceptable to the sector, the policy drivers for the reform shifted, and it was no longer feasible to greatly reduce the burden of the exercise. Within the context of a peer review exercise, we made efforts to streamline wherever possible. Following discussion with the EAGs and a workshop with research PVCs, these efforts focused on:

- Reducing the number of UOAs and ensuring greater consistency between them. Together, these changes were intended to reduce the complexity for HEIs of responding to numerous variations on the criteria, and to reduce the scope for tactical decision-making about which staff to submit in which UOAs.
- Simplifying the categories of staff that were eligible for selection.

- Removing or simplifying any items of data that had been collected in the RAE that were unnecessary for the REF. (Some minor data items were identified.)
- Aligning the data requirements with other existing sources. (In the event we found the scope for this was limited to aligning the definitions of research student and research income data with HESA data returns, and using these HESA data to validate REF returns.)
- Using structured templates for the narrative elements of submissions.
- Exploring options to reduce panels' workload such as sampling the outputs to be reviewed, and more extensive double-weighting of outputs. (In the event sub-panels rejected sampling, but some received significant numbers of double-weighted outputs.)

396. All of these options were carried through in some form, with notable success in establishing consistency across the UOAs, although the opportunities for data alignment proved to be rather limited. Despite these efforts, however, it was apparent during the exercise that significant additional work was carried out in institutions to prepare their impact submissions, and to apply the more rigorous procedures to support equality and diversity; and that overall the burden of the exercise had increased since the RAE. The costs to institutions and panel members, and hence the extent of the increase since the RAE, cannot be estimated here. An evaluation has been commissioned by the funding bodies. The total direct cost to the funding bodies of administering the REF was £14.4 million, a modest increase since the 2008 RAE. Further details about the direct costs are in section 10.3.

397. Based on our experience of the REF we suggest some potential refinements within the existing parameters of the assessment framework, that should be considered with a view to reducing burden, if a similar exercise is repeated:

- Developing the submissions guidance and assessment criteria as a single coherent document, by the REF team working with the main panels; and deferring the appointment of sub-panels until a later stage. (Savings could include approximately £1 million in the direct costs of the criteria phase sub-panel meetings; sub-panel members' time; and greater simplicity for institutions in the guidance documentation.)
- Simplifying the guidelines and tariffs regarding individual staff circumstances as far as possible, and providing a tightly specified template for submitting staff circumstances.
- Reducing the need for HEIs to 'reinvent the wheel' when developing codes of practice, for example by developing 'model' text for various sections of the code which HEIs could adapt to suit local circumstances.
- More closely aligning research student and income data with HESA data, now that HESA cost centres are more similar to REF UOAs. This could potentially involve direct use of the HESA data, without requiring submission by HEIs.
- Sub-panels reviewing samples of outputs, and further extending the use of double-weighting (some sub-panels noted that some HEIs appeared reluctant to request double-weighting when they could legitimately have done so).

398. The above suggestions do not include suggestions relating to impact, as this element of the exercise is subject to a full evaluation that is expected to produce such recommendations.

399. Given the substantial burden involved in the REF it is also worth considering some more radical changes for the future, with the potential to significantly reduce the burden of assessment. Some brief suggestions for further consideration are offered:

- a. If the selection of outputs required for assessment can be decoupled from the selection of staff, the considerable efforts and stresses involved in staff selection and associated equality and diversity procedures could be greatly reduced. Some other complexities, notably relating to multi-authored papers, might also be reduced. However, this approach would require very careful consideration as there will remain a need for a robust volume measure for funding purposes.
- b. Some might argue that the burden associated with staff selection can be avoided by having a comprehensive exercise in which all eligible staff (or possibly all research publications) are included. However, it is not clear that some form of selection can be avoided (institutions could effectively select staff through the eligibility criteria); the implications for institutions that are not research intensive would need careful consideration; and this approach could result in far greater volumes of outputs for panels to review.
- c. In some disciplines it might be possible and acceptable to place greater reliance on metrics, with the potential to substantially reduce the workload for some sub-panels. However, a binary system (metrics-driven for some disciplines and peer review for the others) is unlikely to be acceptable so a more variable mix of metrics and peer review along a spectrum would need to be explored. It is not however clear how such a system would substantially reduce the burden on institutions, unless staff selection can be avoided for the more metrics-based disciplines.

9.2. Impact on equalities, staff and privacy

Key points

- Significant efforts were made to strengthen the equality and diversity aspects of the REF. EDAP concluded that overall they were successful and that the REF had a positive impact on equality and diversity in the sector.
- Although the REF panels did not assess individuals, the process of staff selection could put significant pressure on individual staff. This may have been exacerbated by the increased focus on 3* and 4* outputs in the REF.
- Measures were successfully put in place to preserve the confidentiality and security of data collected and generated as part of the REF exercise, especially sensitive information about individuals' circumstances.

Recommendations

- The EDAP report makes a number of recommendations for further enhancing the equality and diversity measures.
- The impact on individual staff might be reduced in future if some means could be found of decoupling the selection of staff and outputs.

400. Significant efforts were made to strengthen the equality and diversity aspects of the REF. As part of its work EDAP considered the effectiveness of these measures, and concluded that overall they succeeded and that the REF has had a positive impact on equality in the sector. The [EDAP final report](#) sets out its conclusions and makes recommendations for building on these advances in future, which need not be repeated here.

401. The REF assessment is undertaken at the level of whole submissions, and focuses on the quality of outputs without making judgements about individual staff. Nevertheless, it involved the selection of individual staff and although the requirements for codes of practice were based on principles of fairness and transparency, the process of staff selection – and the importance to individuals of being submitted – could put significant pressure on individual staff. This was probably heightened in the REF, in comparison to the RAE, by the increased focus on 3* and 4* research. As suggested above, these pressures could be substantially reduced in a future exercise if the selection of outputs could be decoupled from the selection of staff.

402. As the REF involved collecting and processing large volume of personal data, privacy and confidentiality issues were carefully considered. Early in the submissions phase we published a [model staff data collection notice](#) to assist HEIs in ensuring that individuals included in their submissions were aware of how their data would be used. Institutions were also informed of the [arrangements](#) for submitting data that were confidential for other reasons, such as commercial sensitivity.

403. All those involved on the REF panels were bound by the terms of a [confidentiality agreement](#) which included detailed instructions for maintaining the confidentiality and security of both the submissions data they had access to, and the assessment data that they generated. The assessment systems were designed specifically to support the secure and confidential working practices.

404. A retention and disposal schedule for all data collected and generated during the REF was created and adhered to. It followed the principles that any confidential or sensitive information should be retained only for as long as it was required for the purpose of the assessment, and that all general information of wider interest should be published. Almost all the personal information collected as part of the submissions – and that created by panels during the assessment – was destroyed once no longer required, in line with good data protection practice. Only minimal details of staff were published with the submissions, and the listed staff were not linked to submitted outputs, to help preserve confidentiality about staff circumstances.

405. Particular attention was given to the confidentiality of the very sensitive personal information that could be involved in complex circumstances. Measures were successfully put in place at each stage, including requiring institutions to have confidential means of disclosure within the HEI; providing details about complex circumstances only to EDAP and not to the sub-panels; and anonymous review of the complex circumstances by EDAP. The only issue encountered was that although institutions had been instructed to anonymise the complex circumstance statements, a small number included identifying information. It was crucial therefore that such information was available only to a very small group of people (EDAP). At the

end of the exercise EDAP concluded that the arrangements were trusted by staff, who proved willing to disclose the information.

9.3. Economic and sustainability impact

Key points

- The inclusion of impact in the REF provides a clear positive incentive to enhance the benefits that research brings to the economy and sustainability.
- The scale of the REF assessment involved a significant amount of travel by panel members; to mitigate the environmental impact of this we successfully enabled and encouraged them to travel primarily by public transport.
- The vast majority of outputs were collected and distributed in electronic form, and the number of physical outputs was less than half that in the 2008 RAE. This greatly reduced the environmental impact of deliveries, although panel members requested a larger number of print outs of PDF outputs than anticipated.

Recommendation

- The environmental impact of the exercise could be reduced further by increasing the collection of outputs in electronic form.

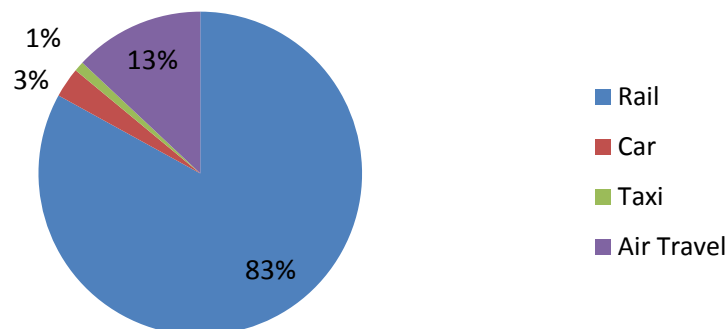
406. By explicitly assessing the non-academic impact of research, the REF provides clear rewards for academics and institutions that produced impactful research, and provides a strong incentive for them to seek to maximise impacts from their research in future. The submitted case studies already demonstrate a range of impressive impacts on the economy, sustainability and the environment. Clear incentives are now in place to enhance these benefits from research and the REF ought therefore to make a positive, if indirect, contribution to the economy and sustainability in future. However it is not possible in the present report to properly evaluate these effects.

407. The direct environmental impact of running the REF exercise primarily related to panel members' travel to meetings, and the collection and distribution of (large numbers of) outputs. With over 250 panel meetings taking place over the course of 2014, each with an average attendance of 35 individuals, we consider that travel constituted the main direct environmental impact of running the exercise. Individual members were responsible for making their own travel arrangements, and we sought to minimise and mitigate the environmental impact by:

- Holding meetings of one day's duration in central London or Birmingham, discouraging car travel and ensuring that public transport was readily available for UK based attendees.
- Holding multi-day meetings at non-city centre venues (to provide a more pleasant working environment at lower cost), but within a short distance of mainline train stations. Where possible shuttle bus arrangements were made for travel to and from the station.
- Providing explicit expenses guidance that panel members should travel to meetings by public transport wherever possible. There were strict constraints on the use of taxis.

408. After the round of planning meetings in early 2013 we analysed the available data on panel members' means of travel, which indicated that our mitigating actions were succeeding.

Primary methods of travel to REF meetings



409. The collection and distribution of outputs are described more fully in sections 5.10 and 8.1. In summary, we were able to collect and distribute the vast majority of outputs in electronic form, and the number of physical outputs was less than half the number in the 2008 RAE (21,588 compared to 53,525). This greatly reduced the volume of physical deliveries; we despatched a total of 3,815 consignments to panel members. This could be further reduced in future by collecting more of the material electronically, especially book chapters and portfolio-based outputs (which were frequently supplied in large boxes).

410. However, panel members requested a larger number of print-outs of PDF outputs than anticipated – a total of 15,793. Their willingness or otherwise to review the material electronically may constrain the savings that can be made in future.

10. Project management

10.1. Project structure and staffing

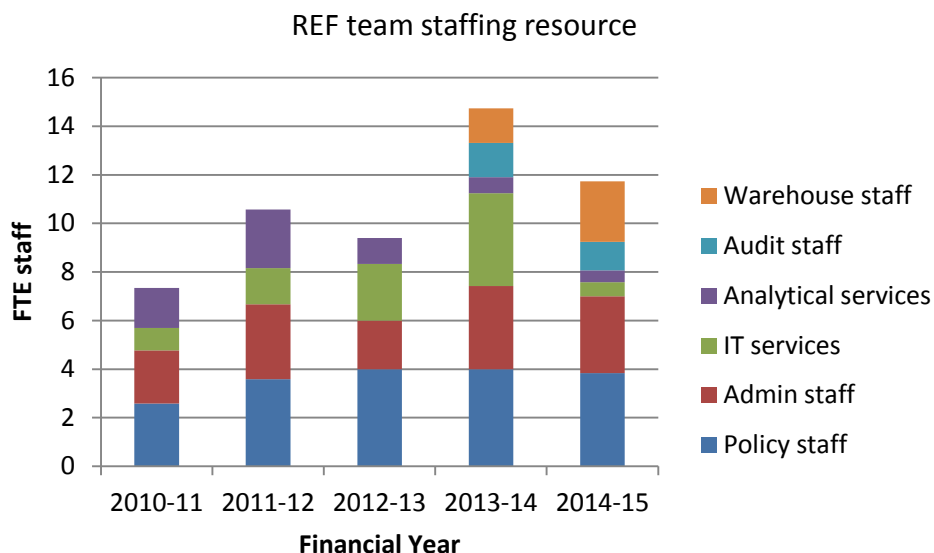
Key points

- Implementation of the REF was effectively overseen by the REF Steering Group, which represented the four UK funding bodies.
- Once the key parameters of the assessment had been decided, the REF team became established. At its peak of activity, 15 FTE staff based at HEFCE supported the REF, bringing together a range of skills and expertise.

411. Once the key parameters of the assessment had been decided by the funding bodies in 2010, the REF team became established. It was based at HEFCE and worked on behalf of the four UK funding bodies. The REF team's work was overseen by the REF Steering Group, which represented the four UK funding bodies. The Steering Group was responsible for signing-off all key guidance documents and other policy-related decisions, monitoring progress, project risks and the operating budget. Members of the Steering Group reported to and sought decisions as necessary from each of their respective Boards, chief executives or equivalent.

412. The REF team in 2010-11 was initially small, and once it developed a project plan and identified the staffing requirements, recruited staff as necessary for the implementation phases. Staff required for the whole of the exercise were recruited on open ended contracts and staff

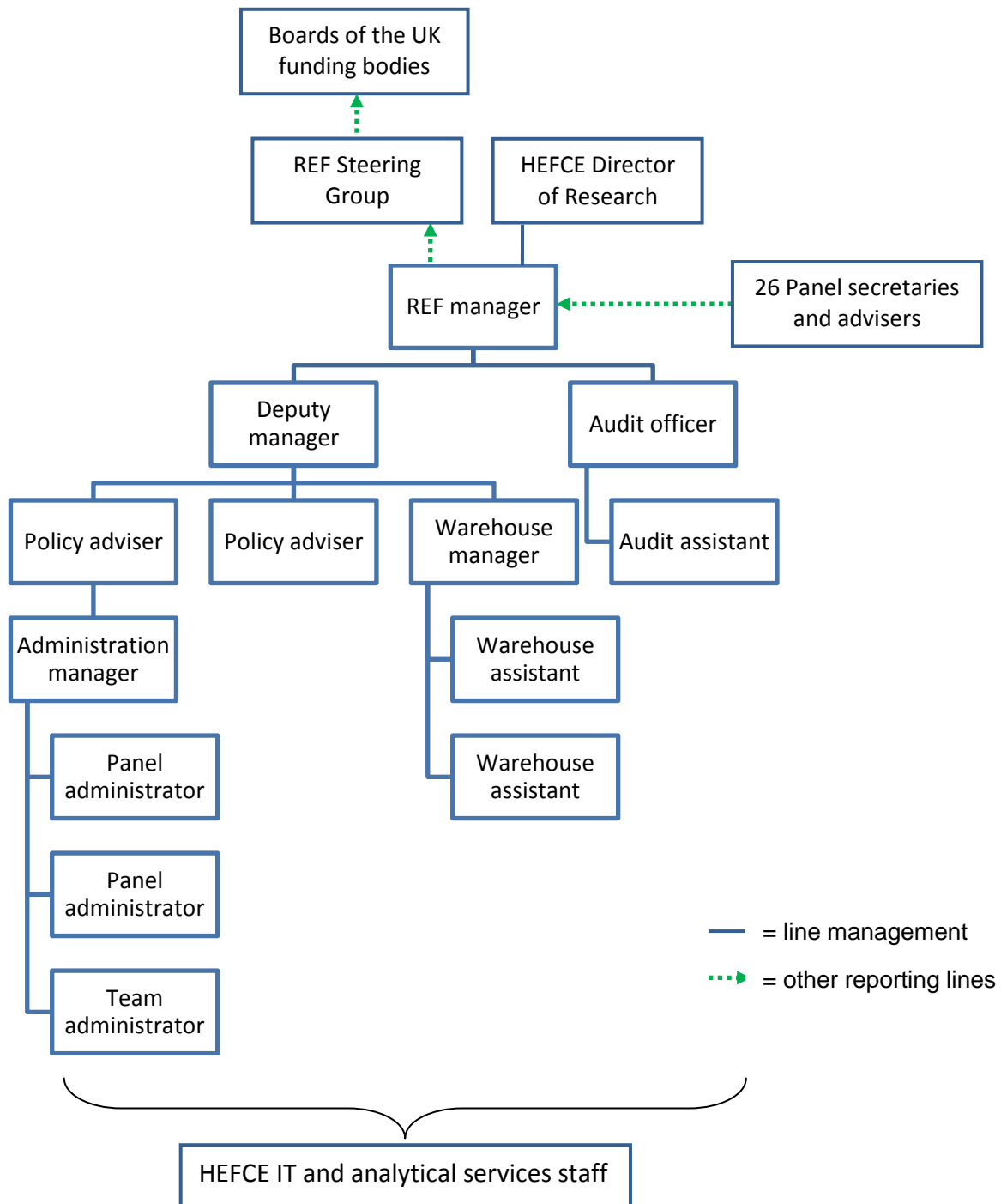
required for specific phases of work were recruited on temporary contracts. The chart below summarises the staff complement for the REF team broken down by year and main areas of work. It includes estimates of the staff resource from HEFCE's IT and Analytical Services teams that supported the REF project.



413. A range of skills were brought together to deliver the exercise, from across members of the REF team and other teams within HEFCE that supported the REF. Key areas of expertise included policy development, project management, communications, logistics and business processes, software development, database management, data processing, customer service and administrative skills. Internal project groups met very frequently to bring together the required range of expertise for key strands of work, especially for IT developments; data collection and analysis; and general project management.

414. Due to the development, testing and user support for the new assessment systems, more IT resource was required than initially anticipated.

415. The team structure, at its full complement, is below.

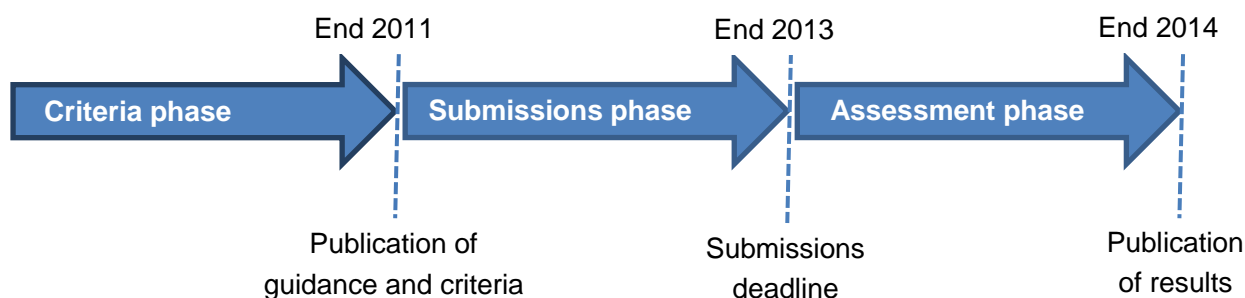


10.2. Project planning and management

Key points

- The 2008 RAE provided invaluable benchmarks for planning and budgeting for the REF, with key deadlines largely transposed from the RAE timetable.
- Long, medium and short-term planning by the REF team was highly effective and this was vital to the successful operation of the exercise.
- All major processes and deliverables were completed to schedule, with pinch-points effectively anticipated and planned for.
- Risk was effectively managed; none of the many potential risks (under the REF team's control) materialised in a way that had a significant impact on the project.

416. Following the key 'initial decisions', implementation of the REF involved three key phases, each with a major deliverable to be completed by a fixed deadline, indicated below:



417. Key points in the REF timetable were largely transposed from the 2008 RAE, including the timing for publication of guidance, the census date for staff eligibility, the submissions deadline and publication of results. The [timetable](#) for the REF, including all key dates for submitting institutions, was published early and required very few and only minor revisions during the process.

418. Effective planning was vital to the successful implementation of the exercise. The 2008 RAE provided invaluable benchmarks for planning and budgeting for the REF. The REF team developed a broad plan for the whole of the exercise, initially by adapting the RAE plan. This was reviewed and refined periodically, usually on a six-monthly basis. The broad plan was used for regular progress reporting to the Steering Group, which we updated on the status and any issues arising for each main strand of activity. The final version of this broad project plan is at Annex C.

419. The team also put in place arrangements for short- and medium-term planning and progress monitoring. Detailed plans were developed for each phase of the exercise and for specific strands of work (such as the audit exercise, IT developments, and the warehouse operation). Monthly team meetings reviewed progress and plans with a quarterly horizon, and weekly meetings were held to update on immediate activities. Specific internal project groups also reviewed plans and progress for those projects (especially data collection and the IT developments).

420. Overall the planning framework was highly effective, with all major processes and deliverables completed to schedule and to a consistently high standard. The two main pinch-points in the project were anticipated early on and effectively planned for; the most challenging being the transition from the submission to the assessment phase in late 2013; early preparation for the assessment phase proved to be essential. The only activity that could have been better planned for at an earlier stage was the assessment systems development. The resources and timescales required were initially underestimated but resolved by appointing additional staff on fixed-term contracts and prioritising developments with a phased roll-out.

421. The REF team developed a comprehensive risk register which was reviewed regularly by the team and reported on as a standing item at REF Steering Group meetings. We identified four general areas of risk to the successful implementation of the exercise:

- Stakeholder confidence in the process and outcomes
- Operational effectiveness
- Robustness of the assessment method
- Behavioural consequences.

422. Each of these areas of risk was broken down further into a number of specific risks, with assessments of their likelihood and impact, and mitigating actions. During the project the team refined its processes for risk management by evaluating the status of each risk as the project progressed and our attention shifted from some risks to others. Our reports to the Steering Group highlighted risks that were being actively managed at each stage of the project.

423. Internal HEFCE audits were carried out in most years of the REF project, and provided assurance to senior HEFCE staff and the HEFCE Audit Committee about the effective management, and especially risk management, of the REF project. The audits also led to some refinements to our project and risk management procedures.

424. Through the course of the REF they were many potential risks that were at least partly under the REF team's control, and some that were not. We successfully mitigated or avoided all risks under our control, and none materialised in a way that had a significant impact on the project. Some risks that were largely not under the REF team's control may well have transpired, particularly in relation to the behavioural consequences of the REF (such as the cost incurred by institutions). Such issues are more properly addressed as part of the funding bodies' programme of evaluating the REF.

10.3. Budget

Key points

- The direct administration cost of the REF (shared between the four UK funding bodies) was £14.4 million, a modest increase since the 2008 RAE.
- As in the previous exercise the largest component was the cost of panel meetings. These were difficult to estimate accurately at the outset (before the total number of panellists and meetings were known), and some revisions to the budget were made during the course of the exercise.

425. The direct administration cost of the REF was shared between the four UK funding bodies. The REF team made initial budget estimates using the RAE as a benchmark, estimating the staffing resources that would be required over the life of the project, and estimating the total cost of panel meetings – the largest cost element. This required a series of assumptions about the total number of panel members and assessors to be appointed; the average number of meetings and days that each category of member would attend; and the average travel, accommodation and venue hire costs. These were difficult to estimate accurately before all members and any of the assessors were appointed and before meetings were scheduled. The initial budget for the REF was set at £13.2 million.

426. During the course of the REF the budget was revised twice as the numbers of panellists and meetings became clearer, and some additional staffing requirements were identified. The funding bodies agreed increases to the budget in 2012 and 2013 to accommodate additional IT resources and increased estimates of the panel meeting costs.

427. The total expenditure was £14.4 million, a modest increase in expenditure for 2008 RAE, which was £12 million. The main reasons for the increase are the additional meeting days required for assessing impact; additional policy development and piloting in the initial stages; increased IT development costs; and more fully accounting for the indirect staffing costs (which include estimated costs of office services such as finance and human resources).

428. The breakdown of expenditure can be seen in the table below. Panel costs comprise panellists' fees, expenses, accommodation and meetings costs, and the panel secretariat secondments. Other programme costs include the costs of the impact pilot exercise, provision of citation data, warehouse costs, IT equipment and publication costs. Cost controls were established in line with those in place for HEFCE staff and Board members, and the REF admin team rigorously adhered to the limits set for overnight accommodation and venue bookings, and achieved good value for money by negotiating discounts for large-volume bookings.

Financial year	2010-11	2011-12	2012-13	2013-14	2014-15	Total
Direct staff costs (£)	387,470	471,370	406,009	642,517	492,664	2,400,030
Indirect staff costs (£)	297,641	325,317	390,121	420,736	314,956	1,748,771
Panel costs (£)	82,571	1,292,766	311,810	1,179,390	6,457,275	9,323,812
Other programme costs (£)	337,357	87,112	72,581	288,490	190,076	975,616
Total (£)	1,105,039	2,176,565	1,180,521	2,531,133	7,454,971	14,448,229

Annexes

Annex A: Publications

2015

- [REF 2014 Panel overview reports](#) (January 2015)
- [Equality and diversity in the 2014 REF: A report by the Equality and Diversity Advisory Panel \(EDAP\)](#) (January 2015)
- [Institutional codes of practice on the selection of staff](#) (January 2015)
- [Final list of panel membership, specialist advisers and conflicts of interest](#) (January 2015)
- [Publication of the REF 2014 submissions](#) (January 2015)

2014

- [REF 01.2014 Research Excellence Framework 2014: The results](#) (December 2014)
- [REF 2014 – Key facts leaflet](#) (December 2014)
- [Arrangements for the publication of results](#) (April 2014)

2013

- [Arrangements for the collection of final codes of practice and equality impact assessments](#) (December 2013)
- [Invitation to make submissions to the REF](#) (March 2013)
- [REF submission system user guide](#) (updated September 2013)

2012

- [REF Codes of Practice for the selection of staff: A report on good practice](#) (October 2012)
- [Invitation to complete the REF survey of submission intentions](#) (July 2012)
- [Summary of responses to the 'Consultation on draft panel criteria and working methods'](#) (March 2012)
- [Invitation to make requests for multiple submissions and for impact case studies requiring security clearance](#) (February 2012)
- [REF 01.2012 Panel criteria and working methods](#) (January 2012)

2011

- [Invitation to submit codes of practice on the selection of staff for the REF](#) (December 2011)
- [REF 03.2011 Consultation on draft panel criteria and working methods](#) (July 2011)
- [Analysis of panel membership](#) (July 2011)

- [REF 02.2011](#) [Assessment framework and guidance on submissions](#) (July 2011)
- [Guidance to panels](#) (July 2011)
- [Equality briefing for panels](#) (July 2011)
- [REF 01.2011](#) [Decisions on assessing research impact](#) (March 2011)

2010

- [REF 01.2010](#) [Units of assessment and recruitment of expert panels](#) (July 2010)

Other publications

- [REF Research Impact Pilot Exercise Lessons-Learned Project: Feedback on Pilot Submissions – Technopolis](#) (November 2010)
- [Research Excellence Framework impact pilot exercise: Findings of the expert panels](#) (November 2010)

Annex B: Survey of submission intentions

The table below compares, by UOA, the volume of staff that HEIs indicated in the survey in late 2012 that they intended to submit, with the actual volume of staff submitted in November 2013.

UOA no.	UOA name	Intended Cat A FTE staff to be submitted	Actual Cat A FTE staff submitted	% difference
1	Clinical Medicine	3,581	3,571	0
2	Public Health, Health Services and Primary Care	1,345	1,354	1
3	Allied Health Professions, Dentistry, Nursing and Pharmacy	2,839	2,748	-3
4	Psychology, Psychiatry and Neuroscience	2,666	2,520	-5
5	Biological Sciences	2,312	2,373	3
6	Agriculture, Veterinary and Food Science	1,061	1,042	-2
7	Earth Systems and Environmental Sciences	1,454	1,380	-5
8	Chemistry	1,242	1,229	-1
9	Physics	1,771	1,704	-4
10	Mathematical Sciences	1,924	1,930	0
11	Computer Science and Informatics	2,108	2,044	-3
12	Aeronautical, Mechanical, Chemical and Manufacturing Engineering	1,147	1,152	0
13	Electrical and Electronic Engineering, Metallurgy and Materials	1,094	1,071	-2
14	Civil and Construction Engineering	423	390	-8
15	General Engineering	2,370	2,447	3
16	Architecture, Built Environment and Planning	1,087	1,025	-6
17	Geography, Environmental Studies and Archaeology	1,815	1,686	-7
18	Economics and Econometrics	820	756	-8
19	Business and Management Studies	3,677	3,320	-10
20	Law	1,794	1,553	-13

21	Politics and International Studies	1,327	1,275	-4
22	Social Work and Social Policy	1,422	1,302	-8
23	Sociology	714	704	-1
24	Anthropology and Development Studies	589	562	-5
25	Education	1,608	1,442	-10
26	Sport and Exercise Sciences, Leisure and Tourism	843	790	-6
27	Area Studies	535	483	-10
28	Modern Languages and Linguistics	1,449	1,386	-4
29	English Language and Literature	2,012	1,971	-2
30	History	1,831	1,786	-2
31	Classics	380	383	1
32	Philosophy	566	591	4
33	Theology and Religious Studies	467	413	-12
34	Art and Design: History, Practice and Theory	1,779	1,604	-10
35	Music, Drama, Dance and Performing Arts	1,183	1,142	-3
36	Communication, Cultural and Media Studies, Library and Information Management	1,037	935	-10
	Total	54,269	52,061	-4

Annex C: Project plan

	2011												2012												2013												2014												2015		
	1	2	3	4	5	6	7	8	9	10	11	12	1	2	3	4	5	6	7	8	9	10	11	12	1	2	3	4	5	6	7	8	9	10	11	12	1	2	3	4	5	6	7	8	9	10	11	12	1	2	3
CRITERIA PHASE																																																			
Complete initial panel member recruitment	█																																																		
Develop guidance on submissions	█																																																		
Develop draft panel criteria	█																																																		
Panel meetings	█																																																		
Consultation on panel criteria	█																																																		
Publish final panel criteria													█																																						
SUBMISSIONS PHASE																																																			
Develop and test submission system	█												█												█																										
Prepare user guide and provide user support	█												█												█																										
Bulk data import/export testing	█												█												█																										
Pilot of submission system	█												█												█																										
Submission system live													█												█																										
Multiple subs and class. case study request													█												█																										
Survey of submission intentions													█												█																										
Provision of REF4 data to HEIs													█												█																										
AUDIT AND DATA VERIFICATION																																																			
Develop audit guidance and procedures													█												█																										
Recruit and train audit staff													█												█																										
Staff sample and data comparison audits													█												█																										
Panel instigated audits													█												█																										
OUTPUTS MANAGEMENT																																																			
Publisher agreements to access outputs													█												█																										
Develop outputs management systems													█												█																										
Harvest electronic outputs													█												█																										
Procure and kit-out warehouse													█												█																										
Recruit and train warehouse staff													█												█																										
Collect physical outputs													█												█																										
Distribute outputs to panel members													█												█																										
Return physical outputs to HEIs													█												█																										
EQUALITY AND DIVERSITY																																																			
ECU training events													█												█																										
EDAP review HEIs' codes of practice													█												█																										
Report on codes of practice													█												█																										
EDAP review complex circumstances													█												█																										
EDAP review EIAs													█												█																										
ASSESSMENT PHASE																																																			
Schedule panel meetings													█												█																										
Book meeting venues													█												█																										
Recruit and train panel admin team													█												█																										
Recruit panel secretariat													█												█																										
Develop assessment systems													█												█																										
Round of panel meetings (planning)													█												█																										
Appoint assessors													█												█																										
Appoint specialist advisors													█												█																										
Assessor briefing events													█												█																										
Distribute subs and PDF outputs to panels													█												█																										
Panel meetings													█												█																										
Secretariat support for panels													█												█																										
Assessment systems live / user support													█												█																										
Provide guidance to panels													█												█																										
Provide admin support to panels													█												█																										
Manage cross-referrals and spec. advice													█												█																										
Final payments and collect USB pens													█												█																										
Publish outcomes and analysis													█												█																										
Provide feedback to HEIs													█												█																										
Publish subs, EDAP and overview reports													█												█																										
Publish minutes and REF manager report													█												█																										