

OCTOBER 2008

Likely Impact of Methodological Decisions on the Research Excellence Framework

A report to HEFCE by RAND Europe

By Linda Butler
RAND Europe

Contents

Background	4
1. How to populate publication databases	6
2. Level of aggregation	7
3. What a university can claim	10
4. What publications a university submits	11
5. How field of research is determined	13
6. Complexity of bibliometric measures	14
7. Choice of citation database	16
Final words.....	17

Likely Impact of Methodological Choices

Background

In response to a request from HEFCE, this discussion paper expands on issues I raised in my presentation to the 'Beyond the RAE 2008: Bibliometrics, League Tables and the REF' at King's College London on 30 April 2008. Specifically, it identifies some of the key technical and methodological choices that need to be made in producing bibliometric quality indicators; outlines the perceived advantages and/or disadvantages of each main approach (in terms of, for example, robustness and acceptability, cost and practical implications, transparency, susceptibility to 'games playing', and how much difference they make to the outcomes); and advises on what the current evidence base is to support these choices.

It is by no means an exhaustive list of the possible ramifications of all methodological choices. It draws heavily on the experiences gained in my work as chair of the Metrics Working Group for Australia's Research Quality Framework (RQF), and some initial testing that had been undertaken for that exercise¹. It also draws on my experience in using bibliometric techniques to assess research performance for Australian institutions and government agencies over the last 20 years.

No matter what methodology is finally chosen for a metrics-based REF, the short-term cost implications for institutions will be significant. After sampling a number of university websites, it does not appear that many UK universities have publication databases that extend beyond items submitted to the RAE. It is hard to argue that a metrics REF will rest on such a limited number of publications, so institutions have a significant up-front cost in populating their information databases.

¹ Unfortunately when the RQF was replaced by Excellence in Research for Australia (ERA), oversight of the process moved to a new agency and the documentation relating to RQF is no longer available on-line.

In contrast, many Australian universities have established, or are establishing, integrated management systems which link their publication, staff, student and grants databases. Much of this information is presented as standard pages for each staff member on their public websites. Examples of this type of information system in operation can be seen at:

University of Queensland: <http://www.uq.edu.au/uqresearchers/>

University of Melbourne: <http://www.findanexpert.unimelb.edu.au/>

From discussions that occurred at the King's College conference in London on 30 April, this is a similar end-point for which most UK universities are aiming, but their current state is significantly behind that of their Australian counterparts.

The RQF was the catalyst for Australian universities moving more quickly down this path. Its unit of analysis was the university-defined discipline 'group' of 5 or more researchers, which in most cases did not align to an organisational unit. Universities were readily able to identify the publications belonging to the group, but found it impossible to identify grants income in the same way. It was particularly challenging where grants had been awarded to medium or large research teams whose members were distributed across a number of groups. Reporting student data presented a similar problem. Most existing management systems did not enable integrated data on individual group members to be extracted, and this in turn led to extensive revamping, or complete redesigning, of institutional management systems. Most universities sought to populate the new integrated database from existing systems, rather than undertaking a completely fresh data collection exercise. It is likely UK universities will follow the same path, and the complexity of the task and the time it will take should not be underestimated.

My comments are based on the following *assumptions*:

1. HEFCE will want to use an efficient and streamlined method of calculating the metrics;
2. institutions will want the ability to check that HEFCE 'gets it right';
3. most institutions will almost certainly want to use the same metrics for internal management purposes (whether they are valid at these lower levels of aggregation or not), and so will need access to fine-grained data; and
4. bibliometric analyses will rest on a much broader list of publications than the current 'four best' used in the RAE.

1. How to populate publication databases

In recent rounds of the RAE, UK universities have only been required to submit details of the 'four best' publications over a six year period for staff submitted in the assessment. This is only a small proportion of the total output². For a metrics based REF, the assessment will have to encompass a much broader collection of publications, which will require a significant data gathering exercise on the part of universities. There are three main choices for constructing more comprehensive lists:

a. obtained locally from individual academics at each institution

At the KCL conference, universities expressed concerns about the feasibility of approaching individual academics to obtain their full list of publications for the past x years. Many Australian universities would agree that this is fraught with problems – academics are often absent from the institution, and even if easily contactable, they do not regard the submission of their publication details to central management for the purposes of a national research assessment as something they want/need to give a high priority to. Providing publication lists to funding agencies is a different matter! The information exists – the problem is extracting it efficiently.

Most Australian universities rely on departmental administrative staff to obtain the information and enter it via web-based forms – this makes training on the functions of the databases more efficient, and such staff are much more readily contactable at key points in the timeline for submission of publication summaries to the government. Such systems have become very efficient, even for large-scale exercises, and have great potential for the REF given the pioneering work done by Australian universities – there is much to learn from their experiences. However, it may not be possible to *fully* populate such databases in time for the first REF. It is not likely to be achievable for any pilot tests conducted in 2008, unless only institutions well-advanced in this process are chosen.

b. centrally provided by HEFCE

An alternative solution, particularly for the first metrics exercise, is for HEFCE to obtain details of UK publications from existing commercial suppliers (Scopus and/or Thomson Reuters) and provide this information to each university. As this will generally provide universities with around two-thirds of their journal output³, it would be a welcome starting point for

² In a yet to be published study with Ian McAllister, we calculated that less than 20% of the total output of UK chemistry departments was submitted for assessment.

³ Based on data from Australian universities submitted annually to the Department of Education

them in the construction of their databases. HEFCE would need to work out contractual and funding arrangements for this approach, and the data costs may be significant.

c. centrally provided by Universities UK (joint negotiation to access with data supplier)

This option is a variation on b above, though with Universities UK (or a similar umbrella body) negotiating access to the data on behalf of individual universities (collective purchases are usually less expensive to individual institutions). Using this approach, the cost would be borne by individual institutions.

In summary: In the long-term, I'd suggest UK universities should be aiming to develop information management systems and procedures that allow them to collect publication data directly from their staff, via departmental administrators, but that constructing detailed publication lists for the first metrics exercise in this way may be difficult to achieve. A possible solution is for all UK publications to be obtained centrally (by HEFCE or Universities UK) to assist universities in the initial population of their publication databases. Whatever the choice, the set-up costs for the sector will be substantial, though considerably reduced in subsequent years as efficient internal systems are developed.

2. Level of aggregation

The level at which the REF assessment is conducted will have direct implications for the validity and robustness of many of the metrics that are being proposed⁴. I will discuss the issues faced by Australian universities in attempting to obtain the information required for the now-defunct RQF, and also by New Zealand universities in relation to their Performance Based Research Fund (PBRF).

a. Individual academic

New Zealand's PBRF was based on an assessment of individuals, with public reports being produced at an aggregated institution level by field. While only aggregated data was published, it is understood that many academics knew their individual assessment (either because universities made this known, or because it was possible to derive it from public

⁴ For an excellent text on the use of bibliometrics in research assessment that covers most issues, see *Citation Analysis in Research Evaluation*, Henk F. Moed, Springer Verlag, 2005.

information). This badly affected the morale of some university staff members – both those who were given a relatively poor assessment, and those who were not submitted in the first place. HEFCE will be well aware of the issues surrounding this latter group for UK staff who are not submitted to the RAE.

The one big advantage of undertaking assessments at the level of the individual academic is that it provides flexibility in how outcomes are aggregated and reported. Individual level data can be aggregated either by field of research (which may be cross-institutional), or by organisational unit. The former are of most interest to government, the latter to university vice-chancellors.⁵

In Australia, the main focus of assessment is to be at the ‘whole of university’ field level. However, the recently released Excellence in Research for Australia (ERA) discussion paper addresses the issue of flexibility in aggregation by raising the possibility of universities flagging publications and other data with the relevant organisational unit(s)⁶. Most universities will, in any event, retain this information internally, even if it is not publicly reported.

The biggest disadvantage of focussing on the individual is the problems it raises for the validity and robustness of the metrics used. Researchers are quick to point out the ‘problems’ that exist with bibliometric analysis and the ways in which it is possible to manipulate the data (citation clubs, excessive self-citations, citations to errors, the skewed nature of the distribution of citations, etc, etc). At high levels of aggregation, with a large number of publications underpinning the analysis, the effects of most of these concerns are significantly reduced. However, in any analysis based on individuals, these concerns are valid and extremely serious. I also believe it is sending the wrong signal – to assess at the individual level is to imply that the metrics are robust at this level, and that it is valid for institutions to apply them to individuals for promotion and appointments decisions.

There is another reason for concern: while publication data can readily be linked to individual academics, data underpinning other likely metrics are far more difficult to link, as discussed in more detail in the next section

⁵ Of course, university leaders are also interested in an assessment of their disciplinary strengths and weaknesses, for strategic management, and for PR purposes. However they are likely to have some reservations about external assessments, particularly when linked to funding, that do not allow them the ability to determine internal strengths and weaknesses on the same basis.

⁶ The consultation document can be found on-line at:
<http://www.arc.gov.au/era/consultation.htm>

(2b). Considerable effort (and therefore cost) would be required to artificially disaggregate such data down to the level of the individual.

b. Academic units or groups

During preparations for the Australian RQF, the problems of undertaking assessments on small groups became a major cause for concern. The difficulties did not concern publications data, which can be reported at various levels of aggregation, but rather related to external income data. Universities found it impossible to disaggregate other institutional data to the small group level, and one of the last decisions made before the change of government and abandonment of the RQF was to remove external income as one of the metrics for this reason.

In the new ERA there is also concern that data will not be available at a sufficiently fine-grained level i.e. 4-digit Field of Research code ⁷. While income from Australian competitive grant schemes (administered by the two research councils) can be reported at this level, it is far more difficult to report other external income at the 2-digit level.

c. Whole of institution

Many of the problems raised in relation to undertaking analyses at lower levels of aggregation are overcome if the focus of the assessment is at the 'whole of institution' discipline level. Concerns about the use of bibliometric indicators are to a large extent nullified because of the large number of publications analysed. It also overcomes the need to artificially divide up income and student data among the various individuals and/or departments that contribute to them, though the difficulty of disaggregating interdisciplinary research remains. This latter issue is to be a central focus of testing during the development of indicators for ERA.

In summary: There are valid concerns regarding the use of some of the proposed REF metrics at levels of aggregation that are too low. Collecting and analysing the metrics at the institutional discipline level overcomes most of these. It could also lessen the cost of gathering and manipulating the data, though senior university management probably prefer to have the information and assessments available at lower levels of aggregation.

⁷ The new Field of Research classification scheme released in March 2008 has a hierarchical structure. 2-digit codes cover broad fields of research (e.g. Mathematical Sciences, Studies in Human Society), while 4-digit codes are at the discipline level (e.g. statistics, demography). The classification system also has 6-digit codes, but it is considered impossible to obtain and assess indicators at this fine-grained level for ERA.

3. What a university can claim

There have been widespread claims that basing the RAE on the outputs of staff who are employed at a university on a given census date, rather than the outputs that staff produce while they are employed at the university (whether or not they are on staff at the census date) has led to game-playing, in the form of recruiting high profile staff just before an RAE round. The choice had little cost impact for the RAE, where only the four best publications were submitted for assessment. However for the REF, at least the first iteration, there is potential for significant cost implications in the choice between these two alternatives.

a. all data relating to researchers on the university's staff at a given census date

Obtaining the details of current staff's 'best four' publications for the RAE was relatively straightforward. But bibliometric assessment will rest on the full (or at the very least an extensive) set of publications for submitted staff (or for the whole university) for the selected assessment period. As universities work towards establishing databases of research publications for the first iteration, this will involve a massive information gathering exercise.

It might be possible to extract much of the data from Scopus by searching on a combination of author name and institution, though this would be fairly time-consuming, and laborious checking would be required for common names. Extracting the data from Web of Science would be even more problematic, as author names are not linked to institutional addresses.

In addition, data for other metrics now proposed for formal use in the assessment process (as opposed to being submitted as supporting evidence) cannot be collected on the same basis. It is likely to be impossible to obtain data on grant income and students for staff who have moved to the university during the assessment period – will their previous institution readily supply the information? and how will it be apportioned when some of the researchers on the grant, or some of the students' supervisors, remain at the original institution? The consequence of this choice is that elements of the suite of metrics will have been constructed on different bases.

b. all data relating to research undertaken at the university during the census period

There are two main advantages in making this choice. Firstly, all metrics can be constructed on the same basis (i.e. they are outputs that can be directly associated with research undertaken at the institution), and they do not rely on an artificial apportioning of outputs, particularly income and student data, for staff who have moved from another institution. Secondly,

the initial populating of a university's publications database could be significantly assisted by directly importing all publications from one or both of the commercial databases that carry the institution's address. For most universities, Web of Science can readily identify around 95%⁸ of publications in their databases from address data, and it is likely Scopus will have a similar capture rate. Some particular institutions (such as Oxford, KCL, Cambridge) present particular problems, but it is still more straightforward than attempting to gather data for specific people.

In the longer term, it will be unnecessary to retrospectively populate the publications databases as information will be entered by current staff each year. There would be no need to identify the publications of new staff, nor expunge the details of those who have left.

In summary: Focusing the analysis on the outputs of research undertaken at the university throughout the period, rather than on the outputs of researchers currently on staff, means that data for all metrics (not just bibliometrics) can be gathered on a consistent basis. Data in existing commercial databases can be used more readily to aid in the initial population of institutional publication databases and thus reduce costs. Looking to the longer-term, it is likely that maintaining management information systems in the universities will also be more efficient with this option.

However, even if this method is used in the REF, university research managers may still wish to obtain full data on recently arrived staff for internal management purposes, particularly as they become more familiar with the new system.

The choice will have little effect on overall performance assessments for the majority of universities. It will only affect outcomes for those universities where performance has rapidly improved or declined in recent years. However, it may affect outcomes at lower levels of aggregation based on smaller publication numbers – such as fine-grained discipline assessments.

4. What publications a university submits

The decision on the coverage of publications has far reaching implications for cost, 'game playing', and the validity of indicators.

a. specified upper limit on number of publications submitted per staff

One suggestion discussed during HEFCE's consultations on the REF was that the number of publications submitted for assessment be limited to a maximum number, such as 10 or 20. The rationale appears to be that it

⁸ based on an analysis of Web of Science data for Australian universities.

prevents highly productive staff members dominating the assessment. It is also a concept familiar to a sector where until now the assessment has been based on the 'best four' publications of each submitted staff member.

In a metrics based assessment system, it is difficult to support any move to limit the number of publications per staff member. Why should the work of highly productive staff be downplayed? If they are all of high quality, and the majority are attracting significant levels of citations, then this information has a legitimate place in the process. If the publications represent a push for quantity, rather than a focus on quality, then it is also important to include all publications as one driver of a full assessment could well be to focus people's attentions on quality rather than quantity.

A second, more important, reason not to use this approach is that it seriously undermines the robustness of the indicators. All world benchmarks are predicated on an assessment of ALL publications in the Web of Science/Scopus world. If universities are able to limit the number of publications put forward for the institution (or for the staff they submit), comparisons to these benchmarks will no longer be valid. Using a selective submission system also means that a metric based on the proportion of output that is among the top 1%/10%/x% most highly cited in the world (an increasingly popular measure given the skewed nature of citation research) can no longer be used because the calculation requires details of the full output of an institution.

A final argument against this approach is a concern over the extra cost involved in identifying which publications are to be submitted, and the intensive effort universities will, justifiably, get to in order to maximise their performance.

b. publications for a selected sub-set of staff

Another method being canvassed in the consultation process is that, like recent RAEs, universities will be given the option to choose which staff will be submitted, and all publications for these will be assessed. The validity of this option in the context of the REF depends on the ultimate aims of the exercise. If the REF is only used for the distribution of funds, and HEFCE is only interested in funding research that is identified as excellent, there is an argument to support this approach. But if at the same time HEFCE and other government agencies seek to use it to gain an accurate picture of the performance of institutions in different disciplines, and indeed the higher education sector as a whole, then this second option will be misleading.

To illustrate my concerns about this option, I describe (as in the previous section) the most extreme scenario that could occur: all chemistry departments in the country follow a logical path – they choose to submit

only those researchers with a good to excellent citation performance; when the bibliometric assessments are undertaken we find that virtually EVERY university in the country active in chemistry has internationally excellent research. In the RAE process, panels had additional information available to assist them in making an overall assessment of the unit, but in an automatic metrics based system, there is much less scope for adjudication except to handle obvious and serious anomalies in the data.

Experience from the UK and New Zealand has shown that allowing universities to choose which researchers to submit for assessment can have a negative impact on the morale of the staff who are not put forward. In addition, the selection process itself can be time-consuming and costly for institutions.

c. all publications

Requiring universities to submit all their publications in a discipline is unlikely to create additional compliance costs; if anything, it could result in significant savings. Even if only a sub-set is submitted, universities will still want to collect all data so that they are able to make strategic choices; and a significant amount of time will be saved in not having to determine which staff and/or which publications should be submitted for assessment. It limits the amount of 'game playing' that institutions and/or researchers can engage in. This method also ensures the robustness of the indicators used. Comparisons to international benchmarks are valid, and a more complete picture of the research performance of each university in each discipline can be gained.

In summary: While limiting the assessment to selected publications and/or selected staff has some appeal, particularly in relation to its similarity to the RAE procedures, it renders comparisons to standard bibliometric benchmarks invalid. Assessing all publications of a university results in valid comparisons, both nationally and internationally, and produces a more complete and accurate assessment of research performance.

Allowing universities to select who or what to submit will provide significant scope for universities to game the system.

5. How field of research is determined

Should universities be permitted to nominate the field of research to which a publication or staff member belongs, or should this be determined automatically by the characteristics of the data? The answer to this question is important when institutional performance is being assessed against discipline-specific benchmarks. It is absolutely essential to use such discipline specific normalisation, but allowing institutions to self-select the field to which a

publication belongs raise the possibility of a considerable amount of ‘game-playing’.

If a publication is co-authored by a mathematician and an environmental scientist, and is published in an environmental science journal, how should it be classified – as a mathematics publication? Or as an environmental science publication? Or as both? If universities are given the freedom to choose they will, if they understand the difference in citation rates between mathematics and environmental sciences, almost invariably choose mathematics. That field has a much lower citation rate than the alternative, so their relative performance will look stronger. The ERA view is that by choosing an environmental science journal, the authors have designated researchers in that discipline as their primary audience, and it is against the benchmarks of that discipline that it should be judged. In other words, the journal that carries the publication determines its field.

That said, there remain deep-seated concerns about the treatment of interdisciplinary research, or output from interdisciplinary research groups, and this will be a major focus of indicator development work for the Australian ERA. In addition, what may work well for journals may not be so straightforward for other publication types, such as books.

It is likely that data for other metrics used in the REF will already have been classified for other reporting needs (e.g. grant income, student data, etc) and at least in the first iteration is unlikely to be open to manipulation. However, a watching brief on trends in the data in future iterations would be wise.

6. Complexity of bibliometric measures

With the computing power that exists today, the choice on whether to use simple or complex bibliometric measures has few cost implications for HEFCE or the sector. However the choice can affect the transparency of the exercise, and the number of analysts who can potentially undertake the assessments. I discuss two issues below that have already been raised as areas of concern.

a. Self-citations

CWTS recommends that self-citations be excluded from any metrics used in research assessment⁹, but a number of other bibliometricians disagree

⁹ Scoping study on the use of bibliometric analysis to measure the quality of research in UK higher education institutions, CWTS 2008, available on-line at: http://www.hefce.ac.uk/pubs/rdreports/2007/rd18_07/

with their stance¹⁰. There is a serious debate about the theoretical basis for excluding self-citations; it is often valid to cite your own work and you should not be punished for doing so. Journal editors act as gatekeepers to the practice of citation clubs and an over-reliance on self-citations. They won't obviate the need of a watching brief on trends in this regard, but there is a limit to how much authors can manipulate citation data. As many have pointed out – citation data is highly skewed, and the highly cited publications that most influence citation rates have relatively few self-citations.

There is also a serious concern about transparency. Few analysts have access to the data that enables them to exclude self-citations in their calculations, and universities are likely to find it difficult to validate the calculations used in the REF, which they will certainly be keen to do.

The KCL conference reported that there was already anecdotal evidence that the possibility of self-citations being excluded from REF metrics was starting to drive perverse behaviours. They reported instances where group leaders were omitting researchers from the list of authors on a paper so that these researchers could subsequently quote the work and have their citation counted.

b. Benchmarks

The simplest method of constructing benchmarks for assessment is to use averages calculated over the *whole assessment period* (e.g. 2002-2007). This does not raise any concerns if all units/institutions being assessed have a relatively even distribution of output across the whole period. However, if the unit is expanding and the bulk of their output is in the latter part of this period, most of their publications will have had little time to attract citations, and the unit will be disadvantaged in any comparison to benchmarks based on five year averages (or indeed to other units). The reverse is that a unit that is in decline, with relatively little activity in the most recent part of the period, will be advantaged in any comparison to world or national benchmarks.

The solution is to assess publications on a year by year basis, against benchmarks specific to each year, then aggregate the data to give an overall performance assessment. While this might sound a complicated process, the calculations are straightforward and it relieves the legitimate concerns

¹⁰ See 'Seven Myths in Bibliometrics: About Facts and Fiction in Quantitative Science Studies', *Collnet Journal of Scientometrics and Information Management*, 2(1), June 2008, in press.

of researchers who are quick to raise this point¹¹. It means new activities are not disadvantaged in favour of more established activities. It also means that the output of early career researchers can be incorporated into the analysis without fear that their recency will unduly deflate the overall performance of the unit.

7. Choice of citation database

Until recently there has been only one widely accepted database for citation analysis – Thomson Reuter’s Web of Science (WoS). The arrival of Scopus, from the Elsevier stable, now means an alternative is available for use in the REF. Until now, few empirical studies have been carried out using Scopus data, so many of the comments I make below relating to the choice of citation database are as yet not based on strong empirical evidence¹². They are therefore couched in cautious terms.

a. coverage

CWTS has already provided HEFCE with an assessment of the coverage of the two databases, and a similar assessment is currently being conducted in Australia. Results from this and other studies carried out recently suggest that there is little difference in the coverage of UK or Australian output in most fields of the natural sciences, but that Scopus has superior coverage of some social science fields such as economics and management, and some applied sciences such as engineering, computing and health services. However, Thomson Reuters have announced that they will be fully integrating their proceedings database with the WoS, and hence its coverage of applied science disciplines, particularly computing and engineering, is likely to improve significantly in the near future.

b. accuracy

WoS data have been interrogated for decades, and the accuracy of their citation attributions are reasonably well determined. In its feasibility study, CWTS claim to be able to increase the number of references that can be

¹¹ This methodology is currently being used to evaluate the performance of publications resulting from the National Health and Medical Research Council’s funding schemes for the period 2002-2006. Their Program Grants scheme is new, with additional grants awarded progressively from its inception in 2001. Only 200 publications were linked to the scheme in 2002, but by 2006 the output had grown to 1300 publications. Using a five year average benchmark would seriously underestimate the performance of this scheme.

¹² However, a number of comparative studies were presented at the 2008 Science and Technology Indicators Conference held in Vienna in September, and it is likely many more studies will be published in the coming months.

attributed to specific publications by 7% over those achieved by WoS itself. At a meeting with Scopus in June 2008, their database administrators claimed a 98.5% success rate in matching the reference data. This has yet to be tested, but does appear to be achievable given the success CWTS is able to achieve with WoS data.

In discussions relating to ERA, Thomson Reuters claimed a 4-5 week turnaround in fixing errors that are identified in their database. Scopus says that it may take them around 2 months to fix errors in their data, but if they have the claimed higher initial success rate, this slow response may not be a cause for concern.

c. identifying authors and institutions

One significant advantage that Scopus has over WoS is that it links authors to their institutional addresses. The WoS failure to make this link is a shortcoming that bibliometricians have long lamented. The ability to link authors with institutions makes identifying the publications of staff with common names significantly easier. As universities start to build up their publications databases, the ability to search for a sub-set of publications that clearly links their authors with their institution could make populating the databases much simpler and more efficient.

d. either of both?

HEFCE does not necessarily have to choose a single data supplier for all REF bibliometric analysis. It may be that the preferred supplier will vary from discipline to discipline.

There may be some within the sector who would argue for measures to be calculated from both data sources. However, while HEFCE might be aiming for a basket of indicators to be used for the REF, replicating every citation measure using two (or more) different data sources is likely to over-complicate the process. I would recommend using only one data source for any given discipline, but suggest that the same supplier need not necessarily be used for all disciplines (unless the pilot studies suggest that one supplier is preferred in all cases).

Final words

I've no doubt that a number of decisions taken will be 'political', informed by intensive sector lobbying or the preferences of government stakeholders, and not necessarily nested in pragmatic rationale. Nevertheless, I felt it important to point out the implications of some of the decisions that will have to be taken.

Choosing a more costly or complex option is not necessarily wrong, provided the implications of the choice are known in advance and taken into account.

However, choosing an alternative that reduces the validity and transparency of the process is less supportable.

Whatever process is developed, universities will want it to be as open and transparent as possible. They WILL check to make sure the data relating to their institution is accurate – and they will probably also cast a close eye over the data of their closest ‘rivals’. They will want access to all the data necessary for this scrutiny.

Choices are inter-related. For example, whether data is provided centrally, or universities have to do it themselves from scratch, will heavily affect the cost implications of choosing between ‘census date’ and ‘address on publication’ (issue 3 above). If a significant amount of data is provided centrally (and this will almost certainly be on the basis of addresses listed on publications), then ‘address on publication’ becomes a significantly cheaper, more efficient option. If universities are required to build their databases from scratch without any external assistance, then the cost differential between these two choices is likely to diminish considerably.

A metrics based assessment system is completely different to a peer review system. That might seem like stating the obvious. But I reiterate it here in closing to point out that HEFCE has the opportunity to approach this as a complete break from the past, and not be tempted to apply some of the past methods to the new system to make them more ‘palatable’ to the sector. What is defensible, or indeed desirable, for a peer review system can be totally out of place in a metrics based system.