

The Metric Tide

Correlation analysis of REF2014 scores and metrics

Supplementary Report II to the
Independent Review of the Role of
Metrics in Research Assessment
and Management

July 2015



The Metric Tide

Correlation analysis of REF2014 scores and metrics

Supplementary Report II to the Independent Review of the Role of Metrics in Research Assessment and Management

July 2015

Report by

HEFCE Analytical Services

Cite as

HEFCE (2015). *The Metric Tide: Correlation analysis of REF2014 scores and metrics (Supplementary Report II to the Independent Review of the Role of Metrics in Research Assessment and Management)*. HEFCE. DOI: 10.13140/RG.2.1.3362.4162

© HEFCE 2015, except where indicated.

Cover image © JL-Pfeifer / Shutterstock.com.

The parts of this work that are © HEFCE are available under the Open Government Licence 2.0: www.nationalarchives.gov.uk/doc/open-government-licence/version/2

ISBN: 1902369297

DOI: 10.13140/RG.2.1.3362.4162

Twitter: #HEFCEmetrics

Contents

Executive summary	iii
Purpose	iii
Key points.....	iii
Introduction	1
Data	1
Coverage of metrics data	1
Details of metrics obtained	6
Methodology	8
Findings	8
Initial overview.....	8
Detailed analysis.....	10
Influence of metrics by unit of assessment.....	14
Modelling of additional characteristics	15
Annex A: Summaries of indicators	16
Summary of citation count as REF quality profile predictor.....	16
Summary of field weight citation impact as REF quality profile predictor.....	20
Summary of percentile of highly cited publications as REF quality profile predictor	24
Summary of source normalised impact per paper as REF quality profile predictor.....	28
Summary of SCImago journal rank as REF quality profile predictor	32
Summary of Collaboration as REF quality profile predictor.....	36
Summary of number of authors as REF quality profile predictor	37
Summary of number of countries associated with authors as REF quality profile predictor	39
Summary of academic and corporate authors as REF quality profile predictor.....	41
Summary of WIPO patent citations as REF quality profile predictor	42
Summary of Mendeley readership as REF quality profile predictor	47
Summary of ScienceDirect downloads as REF quality profile predictor	51
Summary of Scopus full text requests as REF quality profile predictor.....	55
Summary of Tweets as REF quality profile predictor	59
Summary of Google Scholar citations as REF quality profile predictor	63
Annex B: Modelling results.....	68
Annex C: Definitions and abbreviations	78

Executive summary

Purpose

This supplementary report provides the details of an analysis of the Research Excellence Framework 2014 results and metric indicators. It forms a key input of the work of the Independent Review of the Role of Metrics in Research Assessment and Management, chaired by Professor James Wilsdon, the findings of which are set out in the main report, *The Metric Tide*. The work was undertaken by HEFCE's Analytical Services Directorate with additional metrics provided by Elsevier.

For the first time, this report presents analysis which associates anonymised scores from the Research Excellence Framework (REF) with 15 bibliometric and altmetric indicators of research performance at a paper-by-author output level. Furthermore, this work provides the first sector-wide look at the coverage, correlation and predictive qualities of metrics in relation to REF 2014 output scores for all units of assessment.

Key points

Overview

This work has shown that individual metrics give significantly different outcomes from the REF peer review process, showing that metrics cannot provide a like-for-like replacement for REF peer review. Publication year was a significant factor in the calculation of correlation with REF scores, with all but two metrics showing significant decreases in correlation for more recent outputs. There is large variation in the availability of metrics data across the REF submission, with particular issues with coverage in units of assessment (UOAs) in REF Main Panel D. Finally, there is evidence to suggest issues for early career researchers (ECRs) and women in a small number of disciplines, as shown by statistically significant differences in the REF scores for these groups at the UOA level.

Initial analysis and coverage

Before metrics scores could be compared to REF output scores an assessment of the coverage of metrics in the wide range of outputs submitted to the REF was considered. As expected, those UOAs in main panel D showed the poorest coverage of linking to metrics, with only 17

per cent of Classics outputs linked, whereas all UOAs in main panels A and B had over 90 per cent of their outputs linked to the metrics database.

This analysis compared metrics scores to REF output quality scores for 78 per cent of outputs returned to the REF (149,670 out of 191,080). The aim of this work was to assess the extent to which bibliometric indicators correlated or predicted the outcome of the REF peer review assessment. Previous research has not had access to this amount of individualised data, which provides consideration of 15 indicators across all 36 UOAs.

Initial analysis of the data identified publication year as a significant factor in the calculation of correlation with REF scores. The metric most impacted by the publication year was citation count, but most metrics saw significant decreases in correlation for more recent outputs. The exception to this was number of tweets, where weak correlation increased for recent publications, but this is likely to be related to the relatively recent increase in use of Twitter by the academic community.

Metrics as a predictor

In considering the predictive capabilities of the metrics data a number of statistics were considered: correlation, precision (correct predictions) and sensitivity (proportion of REF 4* outputs identified). These measures were considered, controlling for year of publication, for all UOAs¹. This indicated that individual metrics gave significantly different outcomes from the REF peer review, as one in four predictions were false positives in the best-case scenario. Furthermore, Clinical Medicine was the Unit of Assessment with the highest correlation statistics for two-thirds of the metrics considered.

Modelling of the data allowed the influence of multiple metrics in predicting REF scores to be considered at the same time, whilst controlling for year of publication effects. This indicated that there were fewer statistically significant metrics results in main panel D, which is likely related to the reduced coverage in these subject areas. Metrics found to have statistically significant² correlations with REF scores for a wide range of UOAs included: number of

¹ This work did not consider the data at individual institutions, or at institutions by UOA (submission) level.

² Where a number of metrics were correlated with each other the model selected the metric which was best at explaining the variation in REF quality scores. This means that metrics not included in this list, but which had high correlations with other metrics (such as percentile) could also have significant relationships with REF quality scores.

tweets; number of Google Scholar cites; source normalised impact per paper; SCImago journal rank and citation count.

The model showed that outputs from UOAs in main panels A and B were more likely to have statistically significant relationships with a wider range of metrics indicators than main panels C and D. It was noted that even in UOAs where metrics coverage was highest, the outputs with missing metrics scores exhibited significant differences in the propensity to achieve REF 4* scores from those with metrics data at the UOA level. This indicated that when using metrics to help inform assessment, those outputs with missing data should not be assumed to have the same quality profile as those with known data.

Consideration of author characteristics

Using the outputs from the analysis and modelling identified a number of metrics which either correlated with each other or could be grouped together as similar. It is likely that the inclusion of all 15 metrics in the original models interfered with the ability to quantify the metrics effects accurately. Therefore, a restricted model was run, which accounted for a reduced set of metrics characteristics.

Using the restricted model, the effects of the following additional characteristics were considered: sex and ECR status of submitting author. This research found that there was evidence to indicate higher REF scores for male authors and non-ECRs after holding metric scores constant for a small number of UOAs, potentially indicating issues for women or ECRs in these disciplines.

Introduction

For many years metrics have been considered a possible alternative to peer review in the assessment of research excellence within academic departments of Higher Education Institutions (HEIs). Because of a lack of scores for individual articles, previous research into the reliability of bibliometrics to predict the outcome of research assessment exercises in the UK has used department-level average scores instead.

For the first time, this report presents analysis which associates anonymised scores from the Research Excellence Framework (REF) with 15 bibliometric and altmetric indicators of research performance at an author-by-paper output level. This detailed linking of data allowed the analysis to directly compare aggregate REF scores with metrics for the same papers whilst assessing the coverage of the metrics data. A further sector-wide assessment of correlations and the predictive qualities of metrics in relation to REF 2014 output scores was conducted which included all units of assessment.

Data

Coverage of metrics data

The REF 2014 results³ reported that 191,150 research outputs were considered as part of the REF assessment process. This underlying anonymised data linked individuals returned to the REF 2014 with their submitted publications and the REF star rating each output was awarded. The data therefore included multiple rows for individuals, and multiple rows for publications that were submitted multiple times by different authors.

The number of outputs considered for linking to metrics was only 78 per cent (149,670) of the total reported, for two reasons. First, the REF results included double weighting⁴ of outputs with extended scale and scope; and second, not all outputs could be linked to metrics data. Linking to the metrics data required a digital object identifier (DOI) to be returned on the REF database and metrics data to be available from Elsevier⁵. Table 1 presents a breakdown

³ See REF 01.2014 'Research Excellence Framework 2014: The results' for more information (<http://www.ref.ac.uk/pubs/201401/>).

⁴ See REF 02.2011 'Assessment framework and guidance on submissions' for more information (<http://www.ref.ac.uk/pubs/2011-02/>).

⁵ Elsevier is an information solutions provider that was commissioned by HEFCE to link the DOIs of REF submitted outputs to the 15 metrics used in the report. For more information see: <http://www.elsevier.com/>

of the REF outputs and those included in this comparison of the REF quality profiles and metrics data.

Table 1 Coverage of outputs included in analysis by unit of assessment

Main panel	Unit of assessment	REF outputs	Outputs with DOI	% with DOI
Total		191,150	149,670	78%
A	Clinical Medicine	13,405	13,275	99%
	Public Health, Health Services and Primary Care	4,880	4,805	98%
	Allied Health Professions, Dentistry, Nursing and Pharmacy	10,360	10,025	97%
	Psychology, Psychiatry and Neuroscience	9,125	9,020	99%
	Biological Sciences	8,610	8,550	99%
	Agriculture, Veterinary and Food Science	3,920	3,840	98%
B	Earth Systems and Environmental Sciences	5,250	5,170	98%
	Chemistry	4,700	4,685	100%
	Physics	6,445	6,380	99%
	Mathematical Sciences	6,995	6,540	93%
	Computer Science and Informatics	7,665	6,995	91%
	Aeronautical, Mechanical, Chemical and Manufacturing Engineering	4,155	4,050	98%
	Electrical and Electronic Engineering, Metallurgy and Materials	4,030	3,965	98%
	Civil and Construction Engineering	1,385	1,315	95%
	General Engineering	8,695	8,470	97%
	C	Architecture, Built Environment and Planning	3,780	2,700
Geography, Environmental Studies and Archaeology		6,020	4,800	80%
Economics and Econometrics		2,600	2,370	91%
Business and Management Studies		12,205	11,385	93%

	Law	5,525	2,275	41%
	Politics and International Studies	4,365	2,965	68%
	Social Work and Social Policy	4,785	3,485	73%
	Sociology	2,630	1,925	73%
	Anthropology and Development Studies	2,015	1,265	63%
	Education	5,525	4,040	73%
	Sport and Exercise Sciences, Leisure and Tourism	2,760	2,430	88%
D	Area Studies	1,725	855	49%
	Modern Languages and Linguistics	4,945	1,790	36%
	English Language and Literature	6,935	1,940	28%
	History	6,460	2,410	37%
	Classics	1,390	240	17%
	Philosophy	2,175	1,300	60%
	Theology and Religious Studies	1,560	450	29%
	Art and Design: History, Practice and Theory	6,355	1,305	21%
	Music, Drama, Dance and Performing Arts	4,260	1,010	24%
	Communication, Cultural and Media Studies, Library and Information Management	3,520	1,640	47%

Notes: Numbers rounded to the nearest five.

Table 1 shows that all units of assessment (UOAs) in main panels A and B had over 90 per cent of research outputs with a DOI. By contrast almost all UOAs in main panel D had less than 50% coverage, with Classics getting only 17 per cent coverage. This supports a view that UOA sub-panels might individually decide the extent to which metrics should inform the assessment of research quality. In particular, it is likely that the outputs with metrics in main panels C and D will be biased or clustered in specific institutions and subject areas.

Part of the explanation for the variability in DOI coverage relates to the types of output typical in different UOAs. Table 2 shows that those UOAs with low DOI coverage tended to also have low proportions of journal articles. However, even after the proportion of journal

articles was accounted for, UOAs in main panel D still had lower proportions of journal articles with DOIs. This indicates that even if metrics were only used for journal articles, Law and UOAs in main panel D would still have reduced coverage compared to other disciplines.

Table 2 Journal article coverage of outputs by unit of assessment

Main panel	Unit of assessment	REF outputs	% of outputs that were journal articles	% of journal articles with a DOI
Total		191,150	80%	95%
A	Clinical Medicine	13,405	100%	99%
	Public Health, Health Services and Primary Care	4,880	100%	99%
	Allied Health Professions, Dentistry, Nursing and Pharmacy	10,360	99%	98%
	Psychology, Psychiatry and Neuroscience	9,125	100%	99%
	Biological Sciences	8,610	100%	100%
	Agriculture, Veterinary and Food Science	3,920	99%	99%
B	Earth Systems and Environmental Sciences	5,250	99%	99%
	Chemistry	4,700	100%	100%
	Physics	6,445	99%	100%
	Mathematical Sciences	6,995	96%	97%
	Computer Science and Informatics	7,665	72%	97%
	Aeronautical, Mechanical, Chemical and Manufacturing Engineering	4,155	99%	98%
	Electrical and Electronic Engineering, Metallurgy and Materials	4,030	99%	99%
	Civil and Construction Engineering	1,385	97%	97%
	General Engineering	8,695	98%	98%
C	Architecture, Built Environment and Planning	3,780	77%	92%

	Geography, Environmental Studies and Archaeology	6,020	82%	96%
	Economics and Econometrics	2,600	92%	99%
	Business and Management Studies	12,205	96%	97%
	Law	5,525	61%	64%
	Politics and International Studies	4,365	70%	96%
	Social Work and Social Policy	4,785	77%	93%
	Sociology	2,630	75%	95%
	Anthropology and Development Studies	2,015	65%	95%
	Education	5,525	78%	92%
	Sport and Exercise Sciences, Leisure and Tourism	2,760	97%	91%
D	Area Studies	1,725	56%	85%
	Modern Languages and Linguistics	4,945	46%	75%
	English Language and Literature	6,935	33%	78%
	History	6,460	38%	94%
	Classics	1,390	25%	66%
	Philosophy	2,175	59%	92%
	Theology and Religious Studies	1,560	34%	81%
	Art and Design: History, Practice and Theory	6,355	26%	74%
	Music, Drama, Dance and Performing Arts	4,260	28%	80%
	Communication, Cultural and Media Studies, Library and Information Management	3,520	51%	87%

Notes: Numbers rounded to the nearest five.

Details of metrics obtained

Table 3 provides descriptive information of the metrics used in this analysis and details of the type of data returned, as this restricted the level of analysis carried out. The metrics data was linked to the anonymised REF output data where metrics were available.

Table 3 Descriptive summary of metrics data

Type of metric	Metric	Description	Type
Bibliometric	Citation_count	Absolute number of citations per publication.	Numeric, continuous
	FWCI	Field-weighted citation impact - this normalises citations in a field using the world benchmark in that field.	Numeric, continuous, bounded
	Percentile	Top first, fifth, 10th, 25th, 50th or over-50th percentile of highly cited publications.	Categorical, numeric
	SNIP	Source-normalised impact per paper – the ratio of citation count and citation potential in that subject field.	Numeric, continuous, bounded
	SJR	SCImago Journal Rank – measures the scientific influence of journals, accounting for citation count and importance of journal citation.	Numeric, continuous, bounded
	Collaboration	Single author, same institution, same country or at least one author from outside UK.	Categorical, character
	Authors	Number of distinct authors.	Numeric, continuous
	AuthorCountries	Number of distinct countries associated with authors.	Numeric, continuous, bounded
	CrossAcademicCorporate	At least one author from academia and one from the corporate sector.	Categorical, character

	WIPO_patent_citations	Number of times cited by World Intellectual Property Organisation (WIPO).	Numeric, continuous
Altmetric	MendeleyRead	Number of Mendeley article bookmarks and article sharing.	Numeric, continuous
	SciDir_Dwnld	Number of ScienceDirect publication downloads or full-text views.	Numeric, continuous
	ScopusFullTextClicks	Number of full-text requests on scopus.com (user must be subscribed to journal).	Numeric, continuous
	Tweet	Number of times tweeted (this is not restricted to the reference REF dates).	Numeric, continuous
	GS_count	Number of times cited on Google Scholar (this is not restricted to the reference REF dates).	Numeric, continuous

Note: Some of the altmetric data items are known to have coverage limitations. See Annex A for more detail.

The question of interest to this analysis was the extent to which these metric indicators were related to the REF output quality scores. The output quality scores were assessed in the REF against three criteria: originality, significance and rigour. Outputs were assessed against five definitions of quality, with each output awarded either: four star (4*); three star (3*); two star (2*); one star (1*) or unclassified. For the purposes of assessing the relationship between this rating and the metrics the star ratings were re-coded⁶ in the data as numeric values ranging from 0 to 4.

Where necessary, the percentile metric was also re-coded to a numeric field, where 1 signified the most highly cited outputs (those in the top 1 percentile) and 100 indicated outputs in the lowest 50 to 100 percentiles. This enabled the percentile metric to be included in the numeric assessments of relationship between metric and REF quality score.

⁶ It is important to note that outputs receiving an unclassified score were rated as such for one or more of a variety of reasons, some of which did not relate to the originality, significance or rigour of the research itself (for instance, the individual being deemed not to have contributed to the research). We have nonetheless decided to include unclassified outputs in this analysis, as the proportion of unclassified outputs which were rated as such for these other reasons is unknown, and the proportion of the total number of outputs that received an unclassified score overall is otherwise quite small.

Methodology

Each of the 15 metrics were considered as predictors for the REF quality profile outcomes. This analysis evaluated the distribution of the metric across the REF quality profile and, where possible, used the metric to predict whether the output would achieve various REF star ratings. This was conducted for all units of assessment, but should be considered alongside the coverage statistics in Table 1.

This initial analysis did not account for the combined effects of using metrics. So, binary dependent variables indicating those outputs achieving 4* quality were modelled using logistic regression methods. More sophisticated multinomial models were also considered for this data, but did not significantly change the conclusions drawn.

As expected, the UOAs with the least coverage, main panel D, were also the least likely to have significant results in the modelling. The findings discussed in this report have been verified using the model output, and the detailed model output is included in Annex B of this report, presenting findings by main panel and UOA.

The modelling was run for a full model, which included all metrics, and a restricted model, which removed metrics which did not significantly add to the model's ability to explain variation in REF outcomes. Additional characteristics variables were added to the restricted model to test whether there were differences in REF outcomes, by sex or early career researcher (ECR) status of the submitting author, not captured by the metrics scores.

Findings

Initial overview

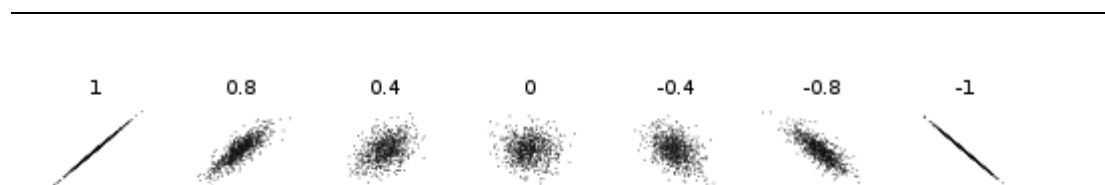
Table 4 provides overall descriptive statistics for the metrics and their relationship to the REF quality profile. This uses the Spearman's rank and Kendall's Tau-b correlation⁷ statistics as these provide non-parametric⁸ measures of dependence between two variables.

⁷ Interpretation of correlation statistics is subjective: in this report the reader is typically left to make their own conclusions about the relationship the correlation statistics indicate. However, the graphic provided below illustrates correlation examples from a sample set of data:

Table 4 Descriptive statistics for numerical metrics data

Variable name	Number non-missing	Standard		Median	Minimum	Maximum	Spearman correlation coefficient	Kendall's Tau-b correlation
		Mean	Deviation					
REF quality profile	148,755	2.9	0.8	3.0	0.0	4.0	1.000	1.000
SJR	112,450	2.5	2.8	1.6	0.0	21.6	0.340	0.269
SNIP	112,045	2.0	1.5	1.6	0.0	17.8	0.327	0.257
Percentile	135,310	39.1	35.8	25.0	1.0	100.0	-0.293	-0.250
FWCI	131,895	3.2	8.2	1.5	0.0	451.7	0.284	0.226
GS_count	148,690	34.2	95.4	13.0	0.0	9,579.0	0.273	0.217
Citation_count	148,755	15.7	46.0	4.0	0.0	2,656.0	0.246	0.201
MendeleyRead	148,685	11.3	41.5	0.0	0.0	3,251.0	0.165	0.146
Authors	148,755	10.3	110.8	3.0	0.0	3,222.0	0.151	0.123
AuthorCountries	148,755	1.6	2.2	1.0	0.0	43.0	0.142	0.125
ScopusFullTextClicks	148,755	10.6	17.0	6.0	0.0	1,166.0	0.107	0.085
Tweet	148,690	2.1	15.8	0.0	0.0	1,443.0	0.090	0.081
WIPO_patent_citations	148,755	0.0	0.2	0.0	0.0	13.0	0.069	0.065
SciDir_Dwnld	41,210	1,256.0	2,247.0	664.0	0.0	59,243.0	-0.017	-0.013

Notes: All correlations were statistically significant. Percentile has been treated as a numeric field for this analysis, so that 1 refers to outputs in the top 1% of highly cited papers and 100 refers to the bottom 50% of highly cited papers. Hence the calculated correlations are negative, because ‘good’ outputs have a low percentile value.



⁸ As stated above, the REF quality score was forced to be numeric for this analysis. However, the relationship between outputs with a 2* rating and those with a 4* rating cannot be assumed to be linear (with 2* outputs half the quality of 4*), so the underlying data is not assumed to be parametric.

Table 4 indicates that the science direct metric (SciDir_Dwnld) had a high number of missing returns, with 72 per cent returned missing. Furthermore, overall correlation between the metric indicators and REF quality profile was generally low, with all correlation statistics below 0.4. This relationship obviously varied by a number of characteristics and these are considered in more detail in the next section.

In addition to the analysis of numeric metrics, chi-squared tests of association were carried out for the two remaining categorical variables (Collaboration and CrossAcademicCorporate). This also indicated statistically significant relationships with the REF quality profile.

Detailed analysis

A key question for this data was whether the metrics obtained at output level were predictive of the final REF peer review assessment of that output. Annex A contains detailed analyses of the relationship between each metric and the REF quality profile by UOA and year of publication. The following paragraphs outline the broad findings from this detailed analysis.

The average metric scores were summarised by the REF quality profile. Many metrics exhibited high scores for REF outputs with 4* or 3* outputs as well as those with unclassified ratings (see Annex A). This demonstrates that an output with a high metrics score could relate to a poor quality or ineligible output, which could be attributed to the fact that citations can be negatively or positively worded.

Publication year effect

Each metric presented a summary of the effect of publication year on the results. This demonstrated that publications from the most recent year (2013) generally had much lower metrics scores than those in the earliest year (2008). The detailed analysis focussed on publication outputs from the earliest year available (2008) to ensure that the effect of year on the analysis was minimised. However, this means that the correlations presented are likely to be the highest available from this data and should not be assumed to be typical of all REF submissions.

Table 5 presents a summary of the metrics and the impact of publication year on the Spearman correlations (see Annex A for more detailed summaries). This shows that the metric most impacted by the publication year was citation count, but most metrics saw significant decreases in correlation for more recent outputs. The notable exception to this was

the number of tweets metric, and this is likely to relate to the ever growing popularity in using Twitter as a medium for publicising and discussing academic works. Thus, more recent works are more likely to have been referred to on Twitter and so the correlation is stronger in more recent publications, although remains relatively weak overall.

Table 5 Correlation statistics for numerical metrics data by publication year

Metric	Spearman correlation coefficient		
	Publication year = 2008	Publication year = 2013	Difference
Citation_count	0.382	0.154	-0.228
FWCI	0.376	0.169	-0.207
Percentile	-0.394	-0.172	0.222
SNIP	0.350	0.136	-0.214
SJR	0.372	0.177	-0.195
Authors	0.168	0.114	-0.054
AuthorCountries	0.164	0.100	-0.064
WIPO_patent_citations*	0.108	-	-
MendeleyRead	0.190	0.196	0.006
SciDir_Dwnld	0.029	-0.042	-0.071
ScopusFullTextClicks	0.143	0.057	-0.086
Tweet	0.065	0.153	0.088
GS_count	0.357	0.247	-0.110

Notes: WIPO patent citations were not calculated for outputs in the most recent publication year.

Relationship between UOA and metrics

The metrics considered could be grouped into two categories: those that were numeric and could be ordered into a hierarchy, called ordinal data, and those that were either numeric or characteristic but had no obvious hierarchy, called nominal data. The predictive nature of

ordinal metrics could be considered more rigorously than the nominal metrics, and Table 6 summarises some of these results. More detail can be found in Annex A. For each ordinal metric, Table 6 identifies the unit of assessment with the highest Spearman correlation coefficient and presents the measures of predictive capability considered.

Table 6 Summary of prediction statistics for most predictive unit of assessment (publication year = 2008) for ordinal metrics

Metric (ordinal)	UOA with highest Spearman correlation				
	UOA name	Spearman correlation coefficient	Precision	Sensitivity	Number of outputs
Citation_count	Clinical Medicine	0.676	48.8%	93.3%	2,070
FWCI	Clinical Medicine	0.635	60.9%	76.6%	2,025
Percentile	Clinical Medicine	-0.670	74.7%	64.7%	2,040
SNIP	Economics and Econometrics	0.665	68.6%	67.8%	265
SJR	Economics and Econometrics	0.751	71.1%	73.6%	265
WIPO_patent_citations	Clinical Medicine	0.229	55.3%	24.3%	2,070
MendeleyRead	Clinical Medicine	0.441	55.0%	60.0%	2,070
SciDir_Dwnld	Chemistry	-0.593	12.0%	7.0%	120
ScopusFullTextClicks	Clinical Medicine	0.376	68.4%	30.1%	2,070
Tweet	Art and Design: History, Practice and Theory	0.234	75.0%	15.8%	130
GS_count	Clinical Medicine	0.600	51.1%	84.8%	2,070

Table 6 shows that for some UOAs, correlation did exceed 0.5 when considering the relationship between the REF quality profile and ordinal metrics. The UOAs with consistently high correlation scores across a range of metrics were Clinical Medicine, and Economics and Econometrics.

The predictive element of this analysis took each metric separately and considered the distribution of metric values available. This was then compared to the distribution of the REF quality profile, and metric cut-off values corresponding to the REF quality proportions were identified. This allowed assessment of whether the metric values alone could predict those outputs assessed as high quality in the REF.

Precision and sensitivity

Two measures of predictive capacity were considered: precision, which is the proportion of predictions of REF 4* outcomes that were indeed assessed as REF 4* (correct predictions); and sensitivity, which is the proportion of REF 4* outputs identified by the metric prediction. Table 6 presents the overall statistics for each metric, but Annex A shows that this disguises wide variations by UOA.

These statistics should be considered together, as typically the results showed that either both proportions were low and so the predictive capabilities were poor; the precision was high but the sensitivity was low and so the predictions were accurate but not comprehensive; or the sensitivity was high but the precision was low and so the predictions were inclusive but inaccurate.

The highest precision presented in Table 6 was 75 per cent for both the percentile and number of tweets metrics. This signifies that three-quarters of the metric predictions agreed with the REF 4* assessment, but also that one-quarter of predictions did not. This shows that even in the UOAs with the highest precision, individual metrics gave significantly different outcomes from the REF peer review.

There has been much research into the effect of ‘playing the system’ when it comes to using metrics as independent measures of quality. It is likely that if a metric such as number of tweets were to be used to measure quality, a lot of work would be needed to identify genuine tweets citing the work and those produced as part of the author’s or institution’s publicity for the work, if this is even possible.

The statistics presented do not overwhelmingly support the use of metrics as a replacement for a peer-review driven model of research quality assessment. However, a number of factors could influence this decision and given the variation of the results by both UOA and metric this decision would need to be made on a case-by-case basis.

Influence of metrics by unit of assessment

Combining the results from the detailed analysis in Annex A and the modelled results in Annex B allowed assessment of the most influential metrics in explaining the variation in REF 4* scores. Stepwise regression was conducted (using the full model as the starting point) for each of the units of assessment to ascertain which metrics were significant in predicting the variation in REF 4* scores. This helped to inform the metrics required for the restricted model (see Annex B).

Metrics with the widest coverage of UOAs (identified as significant for at least half of the 36 UOAs considered) were GS_count and SJR. This suggests that these metrics were good at identifying high-scoring REF outputs for a wide range of units of assessment. Given the lack of metrics coverage in main panel D, it was not surprising that the publications in these UOAs were least likely to significantly correlate REF scores with the metrics scores.

Percentile metrics were significant for the Physics UOA outputs, while citation and index measures such as Citation count, SJR, FWCI and SNIP were significant for the Earth Systems and Environmental Sciences UOA. Altmetrics such as Mendeley readership, Google Scholar and tweets were all significant for Biological Sciences outputs, suggesting that these UOAs correlated well with a number of metrics measures.

Metrics such as number of authors and author countries were significant for some UOAs, but the direction of this significance was mixed (with more authors sometimes associated with higher quality outputs and sometimes with lower quality outputs). The reasons for this are unclear.

International collaboration was significant as an indicator of 4* REF scores within the following UOAs: Clinical Medicine; Earth Systems and Environmental Sciences; Computer Science and Informatics. This indicates that international collaboration in these UOAs distinguished high-scoring outputs better than the other metrics measures. Corporate academic collaborations identified high-performing outputs in Computer Science and Informatics and General Engineering.

The model showed that outputs from UOAs in main panels A and B were more likely to have statistically significant relationships with a wider range of metrics indicators than those from main panels C and D. It was noted that even in UOAs where metrics coverage was highest, the outputs with missing metrics scores exhibited significant differences in the propensity to

achieve REF 4* scores from those with metrics data at the UOA level. This indicated that when using metrics to help inform assessment, those outputs with missing data should not be assumed to have the same quality profile as those with known data.

Modelling of additional characteristics

After regression modelling was used to explore the relative relationships between metrics and UOA effects. A restricted model was derived to investigate the additional effects of the sex and ECR status of the submitting author. The following paragraphs summarise some of the main findings from this modelling work, more detailed summaries are provided in Annex B.

Papers submitted by early career researchers in main panel C were significantly less likely to achieve 4* than non-ECRs with the same metrics ratings. Furthermore, looking at the UOA models showed that this was driven by significant findings in Economics and Econometrics and Social Work and Social Policy.

Female authors in main panel B were significantly less likely to achieve a 4* output than male authors with the same metrics ratings. When considered in the UOA models, women were significantly less likely to have 4* outputs than men whilst controlling for metric scores in the following UOAs: Psychology, Psychiatry and Neuroscience; Computer Science and Informatics; Architecture, Built Environment and Planning; Economics and Econometrics.

On a more technical level, model interaction terms for sex and ECR status of the submitting author and metric scores were considered as part of this work. If significant, these terms would identify those metrics which were good predictors of 4* REF scores for women or ECRs in particular. However, there were no consistent findings for this across UOA or main panel groups.

The modelling of additional characteristics showed that there was evidence to indicate higher REF scores for male authors and non-ECRs after holding metric scores constant for a small number of UOAs, potentially indicating issues for women or ECRs in these disciplines.

Annex A: Summaries of metrics

Summary of citation count as REF quality profile predictor

This analysis summarises the correlation and predictive relationship between citation count and the Research Excellence Framework (REF) quality profile scores. The effect of publication year has been controlled for by focussing on the earliest year available (2008). Table 5 provides a summary of the impact of this choice of year.

Table A1 Summary of citation count by REF quality profile (publication year = 2008)

This summarises the metric values at each level of REF quality profile. Outputs with a 0 quality profile were unclassified, and either fell below the standard of nationally recognised work or did not meet the published definition of research. There were no outputs with missing citation count.

REF quality profile	Frequency	Mean	Minimum	Median	Maximum
4	5040	74.0	0.0	40.0	2,264.0
3	10885	26.8	0.0	17.0	761.0
2	4355	14.1	0.0	8.0	528.0
1	525	12.9	0.0	4.0	573.0
0	260	40.8	0.0	15.0	2,264.0
Total	21060	35.3	0.0	17.0	2,264.0

Figure A1 Scatterplot of citation count against REF quality profile (publication year = 2008)

(Spearman correlation = 0.3819)

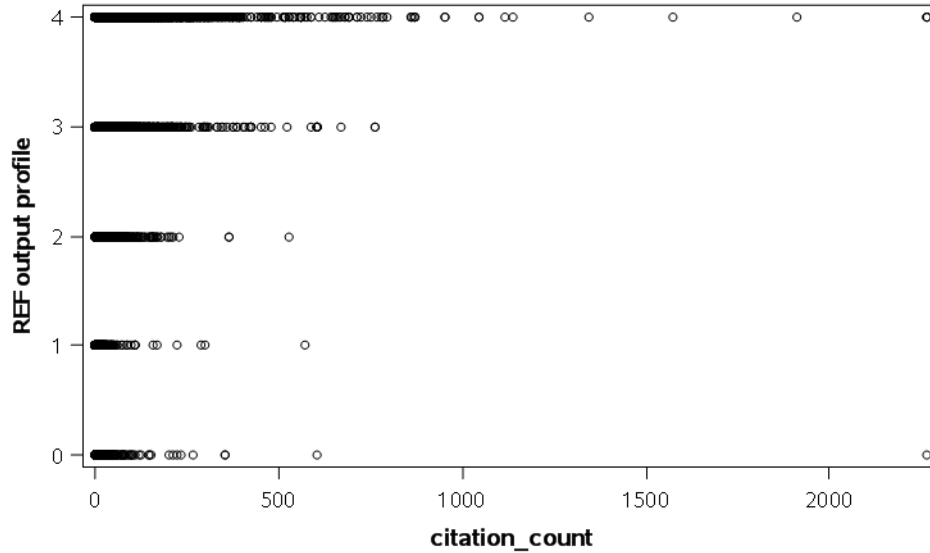


Table A2 Two-way summary of REF 4* indicator by citation count prediction (publication year = 2008)

(prediction = 1 when citation count \geq 41)

REF 4*	metric prediction		Total
	1	0	
output 1	2495	2545	5040
0	2435	13585	16020
Total	4930	16130	21060

Precision = true positives / total positive predictions = $2495 / (2495 + 2435) = 50.6\%$

Precision can be interpreted as the proportion of correct predictions.

Sensitivity = true positives / total REF 4* records = $2495 / (2495 + 2545) = 49.5\%$

Sensitivity can be interpreted as the proportion of 4* records identified by the prediction.

Table A3 Summary table of precision and sensitivity of REF 4* predictions and correlation between REF quality profile and citation count by unit of assessment (UOA) (publication year = 2008)

UOA id	Name	Total	Precision	Sensitivity	Spearman correlation
.	Overall	21060	50.6%	49.5%	0.382
1	Clinical Medicine	2070	48.8%	93.3%	0.676
2	Public Health, Health Services and Primary Care	715	41.8%	70.2%	0.394
3	Allied Health Professions, Dentistry, Nursing and Pharmacy	1455	57.8%	49.2%	0.444
4	Psychology, Psychiatry and Neuroscience	1365	47.9%	57.0%	0.407
5	Biological Sciences	1400	63.3%	76.8%	0.589
6	Agriculture, Veterinary and Food Science	550	50.9%	47.4%	0.423
7	Earth Systems and Environmental Sciences	920	48.1%	73.0%	0.491
8	Chemistry	815	52.1%	82.5%	0.609
9	Physics	1225	51.2%	86.8%	0.608
10	Mathematical Sciences	830	49.4%	21.4%	0.291
11	Computer Science and Informatics	1035	70.4%	33.1%	0.484
12	Aeronautical, Mechanical, Chemical and Manufacturing Engineering	610	39.1%	22.3%	0.226
13	Electrical and Electronic Engineering, Metallurgy and Materials	640	61.7%	45.4%	0.464
14	Civil and Construction Engineering	195	47.6%	32.3%	0.267
15	General Engineering	1330	30.5%	23.5%	0.243
16	Architecture, Built Environment and Planning	300	38.5%	7.7%	0.226
17	Geography, Environmental Studies and Archaeology	550	38.0%	32.1%	0.314
18	Economics and Econometrics	275	78.9%	16.9%	0.333
19	Business and Management Studies	1400	40.1%	18.5%	0.294
20	Law	240	25.0%	1.9%	0.114
21	Politics and International Studies	325	20.0%	1.7%	0.151
22	Social Work and Social Policy	420	50.0%	7.5%	0.265
23	Sociology	250	30.0%	8.3%	0.131
24	Anthropology and Development Studies	150	33.3%	16.7%	0.301
25	Education	455	36.8%	7.6%	0.183

UOA id	Name	Total	Precision	Sensitivity	Spearman correlation
26	Sport and Exercise Sciences, Leisure and Tourism	255	42.1%	16.3%	0.262
27	Area Studies	110	0.0%	0.0%	0.079
28	Modern Languages and Linguistics	175	0.0%	0.0%	0.242
29	English Language and Literature	170	100%	2.6%	-0.015
30	History	245	0.0%	0.0%	0.151
31	Classics	15	0.0%	0.0%	-0.109
32	Philosophy	115	0.0%	0.0%	-0.086
33	Theology and Religious Studies	45	0.0%	0.0%	-0.187
34	Art and Design: History, Practice and Theory	130	0.0%	0.0%	0.004
35	Music, Drama, Dance and Performing Arts	90	0.0%	0.0%	-0.005
36	Communication, Cultural and Media Studies, Library and Information Management	195	45.5%	10.0%	0.292

Table A4 Precision and sensitivity of predictions for a range of REF quality profile groupings (publication year = 2008)

The previous summaries considered the predictive power of citation count for identifying REF 4* work. The following table widens the population to 3* and 4* work and 2*, 3* and 4* work. This provides information on the cut-off used for the predictions and summary statistics, as defined in Table A2.

REF grouping	Prediction cut-off value	REF group indicator				Precision	Sensitivity
		1		0			
		metric	prediction				
4*	41.00	2495	2545	2435	13585	50.6%	49.5%
4* + 3*	7.00	12540	3385	2840	2295	81.5%	78.7%
4* + 3* + 2*	1.00	18465	1810	615	165	96.8%	91.1%

Table A5 Summary of prediction statistics for all publication years

In controlling for the effect of publication year, this summary has effectively ignored 85% of the data available. The following summary indicates what the prediction statistics would look

like for each year of publication. Statistical modelling of this data allows all the data available to be used in the analysis whilst controlling for year of publication effects (see ‘Publication year effect’ in report and Annex B).

Year of publication	Total	Precision	Sensitivity	Spearman correlation	Kendall's Tau-b correlation
2008	21060	53.3%	46.6%	0.382	0.305
2009	23715	49.8%	45.9%	0.360	0.289
2010	25195	46.1%	45.3%	0.347	0.280
2011	26545	41.0%	42.5%	0.300	0.244
2012	27825	35.9%	37.4%	0.249	0.209
2013	24415	36.0%	19.4%	0.154	0.140
All years	148755	40.1%	39.5%	0.246	0.201

Note: Counts and totals of outputs have been rounded to the nearest five for all tables in the analysis.

Summary of FWCI as REF quality profile predictor

This analysis summarises the correlation and predictive relationship between FWCI and the REF quality profile scores. The effect of publication year has been controlled for by focussing on the earliest year available (2008). Table A10 provides a summary of the impact of this choice of year.

Table A6 Summary of FWCI by REF quality profile (publication year = 2008)

This summarises the metric values at each level of REF quality profile. Outputs with a 0 quality profile were unclassified and either fell below the standard of nationally recognised work or did not meet the published definition of research. Outputs with missing field weight citation impact values were excluded from the analysis (missing = 1475).

REF quality profile	Frequency	Mean	Minimum	Median	Maximum	Missing
4	4785	6.8	0.0	3.7	296.4	255
3	10225	2.8	0.0	1.8	86.2	660
2	3905	1.8	0.0	1.1	58.3	450
1	425	1.9	0.0	0.9	39.7	100
0	240	4.5	0.0	1.6	296.4	15
Total	19580	3.6	0.0	1.8	296.4	1475

Figure A2 Scatterplot of FWCI against REF quality profile (publication year = 2008)

(Spearman correlation= 0.3760)

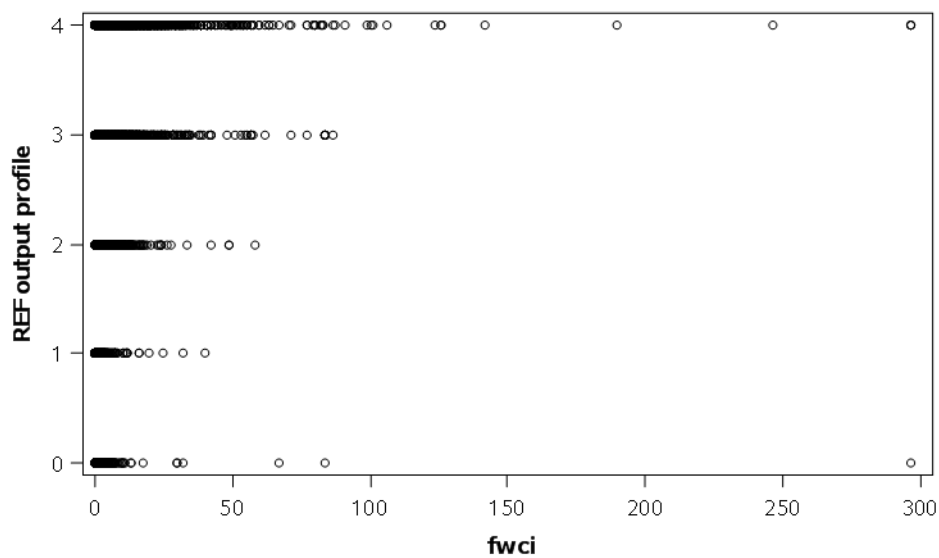


Table A7 Two-way summary of REF 4* indicator by FWCI prediction (publication year = 2008)

(prediction = 1 when FWCI \geq 3.828)

REF 4*	metric prediction		Total
	1	0	
1	2325	2460	4785
0	2360	12435	14795
Total	4685	14895	19580

Precision = true positives / total positive predictions = $2325 / (2325 + 2360) = 49.6\%$

Precision can be interpreted as the proportion of correct predictions.

Sensitivity = true positives / total REF 4* records = 2325 / (2325 + 2460) = 48.6%

Sensitivity can be interpreted as the proportion of 4* records identified by the prediction.

Table A8 Summary table of precision and sensitivity of REF 4* predictions and correlation between REF quality profile and FWCI by UOA (publication year = 2008)

UOA					Spearman
id	Name	Total	Precision	Sensitivity	correlation
.	Overall	19580	49.6%	48.6%	0.376
1	Clinical Medicine	2025	60.9%	76.6%	0.635
2	Public Health, Health Services and Primary Care	665	44.5%	75.0%	0.452
3	Allied Health Professions, Dentistry, Nursing and Pharmacy	1395	54.2%	36.5%	0.395
4	Psychology, Psychiatry and Neuroscience	1330	51.4%	47.1%	0.364
5	Biological Sciences	1375	68.6%	48.6%	0.484
6	Agriculture, Veterinary and Food Science	540	43.3%	39.5%	0.338
7	Earth Systems and Environmental Sciences	880	44.9%	59.6%	0.450
8	Chemistry	805	55.5%	67.5%	0.546
9	Physics	1195	56.9%	76.6%	0.581
10	Mathematical Sciences	780	55.2%	29.1%	0.353
11	Computer Science and Informatics	935	56.6%	54.7%	0.487
12	Aeronautical, Mechanical, Chemical and Manufacturing Engineering	590	27.1%	24.8%	0.211
13	Electrical and Electronic Engineering, Metallurgy and Materials	625	56.8%	52.2%	0.467
14	Civil and Construction Engineering	185	38.2%	43.3%	0.286
15	General Engineering	1285	29.0%	31.6%	0.206
16	Architecture, Built Environment and Planning	260	31.8%	23.7%	0.187
17	Geography, Environmental Studies and Archaeology	520	38.2%	50.9%	0.303
18	Economics and Econometrics	265	66.7%	25.3%	0.295
19	Business and Management Studies	1265	38.1%	38.8%	0.261
20	Law	155	12.5%	6.1%	0.184
21	Politics and International Studies	290	26.4%	25.9%	0.055
22	Social Work and Social Policy	375	25.5%	17.6%	0.195
23	Sociology	225	22.2%	30.3%	0.163
24	Anthropology and Development Studies	135	40.9%	37.5%	0.334

UOA					Spearman
id	Name	Total	Precision	Sensitivity	correlation
25	Education	360	25.9%	17.1%	0.149
26	Sport and Exercise Sciences, Leisure and Tourism	230	52.4%	23.4%	0.175
27	Area Studies	85	23.5%	23.5%	0.065
28	Modern Languages and Linguistics	115	55.6%	18.5%	0.368
29	English Language and Literature	95	23.5%	20.0%	-0.066
30	History	185	42.9%	16.2%	0.134
31	Classics	5	0.0%	0.0%	-0.132
32	Philosophy	80	50.0%	43.8%	0.092
33	Theology and Religious Studies	30	0.0%	0.0%	-0.100
34	Art and Design: History, Practice and Theory	95	33.3%	33.3%	0.137
35	Music, Drama, Dance and Performing Arts	50	9.1%	9.1%	-0.078
36	Communication, Cultural and Media Studies, Library and Information Management	150	46.7%	31.8%	0.273

Table A9 Precision and sensitivity of predictions for a range of REF quality profile groupings (publication year = 2008)

The previous summaries considered the predictive power of FWCI for identifying REF 4* work. The following table widens the population to 3* and 4* work and 2*, 3* and 4* work. This provides information on the cut-off used for the predictions and summary statistics, as defined in Table A7.

REF grouping	Prediction cut-off value	REF group indicator				Precision	Sensitivity
		1		0			
		metric	prediction	metric	prediction		
4*	3.83	2325	2460	2360	12435	49.6%	48.6%
4* + 3*	0.88	12180	2835	2630	1940	82.2%	81.1%
4* + 3* + 2*	0.09	18265	650	590	80	96.9%	96.6%

Table A10 Summary of prediction statistics for all publication years

In controlling for the effect of publication year this summary has effectively ignored 85% of the data available. The following summary indicates what the prediction statistics would look

like for each year of publication. Statistical modelling of this data allows all the data available to be used in the analysis whilst controlling for year of publication effects (see ‘Publication year effect’ in report and Annex B).

Year of publication	Total	Precision	Sensitivity	Spearman correlation	Kendall's Tau-b correlation
2008	19580	51.3%	45.1%	0.376	0.296
2009	22015	48.4%	45.2%	0.361	0.285
2010	23630	45.1%	44.2%	0.344	0.271
2011	24625	40.8%	42.2%	0.299	0.236
2012	25250	36.4%	38.3%	0.235	0.190
2013	16790	32.3%	31.8%	0.169	0.148
All years	131895	42.8%	41.8%	0.284	0.226

Note: Counts and totals of outputs have been rounded to the nearest five for all tables in the analysis.

Summary of percentile of highly cited publications as REF quality profile predictor

This analysis summarises the correlation and predictive relationship between percentile of highly cited publications and the REF quality profile scores. The effect of publication year has been controlled for by focussing on the earliest year available (2008). Table A15 provides a summary of the impact of this choice of year.

Table A11 Summary of percentile of highly cited publications by REF quality profile (publication year = 2008)

This summarises the metric values at each level of REF quality profile. Outputs with a 0 quality profile were unclassified and either fell below the standard of nationally recognised work or did not meet the published definition of research. Outputs with missing percentile of highly cited publications values were excluded from the analysis (missing = 1385).

REF quality profile	Frequency	Mean	Minimum	Median	Maximum	Missing
4	4795	15.2	1.0	5.0	75.0	245
3	10260	26.1	1.0	25.0	75.0	625
2	3930	38.3	1.0	50.0	75.0	420
1	440	43.9	1.0	50.0	75.0	85
0	250	29.0	1.0	25.0	75.0	10
Total	19675	26.3	1.0	25.0	75.0	1385

Figure A3 Scatterplot of percentile of highly cited publications against REF quality profile (publication year = 2008)

(Spearman correlation= -0.3939)

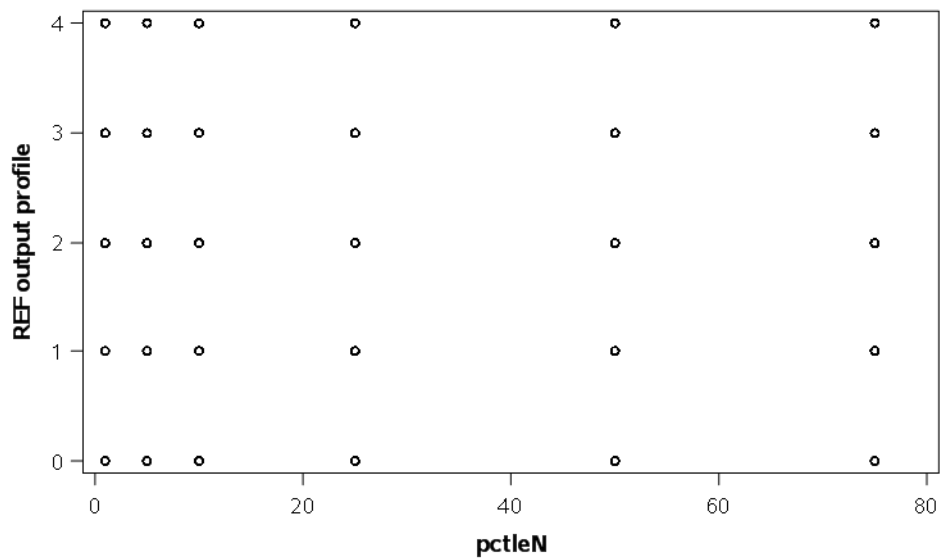


Table A12 Two-way summary of REF 4* indicator by percentile of highly cited publications prediction (publication year = 2008)

(prediction = 1 when percentile of highly cited publications \geq 1)

REF 4*	metric prediction		Total
	1	0	
output 1	1235	3560	4795
output 0	510	14365	14880
Total	1745	17930	19675

Precision = true positives / total positive predictions = $1235 / (1235 + 510) = 70.7\%$

Precision can be interpreted as the proportion of correct predictions.

Sensitivity = true positives / total REF 4* records = 1235 / (1235 + 3560) = 25.7%

Sensitivity can be interpreted as the proportion of 4* records identified by the prediction.

Table A13 Summary table of precision and sensitivity of REF 4* predictions and correlation between REF quality profile and percentile of highly cited publications by UOA (publication year = 2008)

UOA id	Name	Total	Precision	Sensitivity	Spearman correlation
.	Overall	19675	70.7%	25.7%	-0.394
1	Clinical Medicine	2040	74.7%	64.7%	-0.670
2	Public Health, Health Services and Primary Care	665	60.6%	49.4%	-0.422
3	Allied Health Professions, Dentistry, Nursing and Pharmacy	1410	77.5%	19.5%	-0.464
4	Psychology, Psychiatry and Neuroscience	1335	61.7%	24.6%	-0.401
5	Biological Sciences	1385	82.6%	36.3%	-0.574
6	Agriculture, Veterinary and Food Science	540	74.2%	20.2%	-0.423
7	Earth Systems and Environmental Sciences	890	72.8%	30.6%	-0.494
8	Chemistry	805	78.3%	46.4%	-0.584
9	Physics	1205	73.0%	51.1%	-0.597
10	Mathematical Sciences	785	45.0%	4.9%	-0.325
11	Computer Science and Informatics	940	68.4%	9.8%	-0.524
12	Aeronautical, Mechanical, Chemical and Manufacturing Engineering	590	38.9%	5.9%	-0.204
13	Electrical and Electronic Engineering, Metallurgy and Materials	625	78.6%	20.8%	-0.481
14	Civil and Construction Engineering	185	66.7%	13.3%	-0.257
15	General Engineering	1285	41.5%	9.8%	-0.227
16	Architecture, Built Environment and Planning	260	0.0%	0.0%	-0.164
17	Geography, Environmental Studies and Archaeology	525	47.6%	9.3%	-0.289
18	Economics and Econometrics	265	100%	5.7%	-0.321
19	Business and Management Studies	1270	42.1%	2.8%	-0.284
20	Law	155	0.0%	0.0%	-0.255
21	Politics and International Studies	290	0.0%	0.0%	-0.138
22	Social Work and Social Policy	375	66.7%	2.7%	-0.268

UOA					Spearman
id	Name	Total	Precision	Sensitivity	correlation
23	Sociology	225	0.0%	0.0%	-0.136
24	Anthropology and Development Studies	135	50.0%	4.2%	-0.258
25	Education	365	0.0%	0.0%	-0.149
26	Sport and Exercise Sciences, Leisure and Tourism	230	57.1%	8.5%	-0.241
27	Area Studies	85	0.0%	0.0%	-0.151
28	Modern Languages and Linguistics	115	0.0%	0.0%	-0.189
29	English Language and Literature	95	0.0%	0.0%	-0.120
30	History	185	0.0%	0.0%	-0.114
31	Classics	5	0.0%	0.0%	0.535
32	Philosophy	80	0.0%	0.0%	-0.017
33	Theology and Religious Studies	30	0.0%	0.0%	0.144
34	Art and Design: History, Practice and Theory	95	0.0%	0.0%	-0.058
35	Music, Drama, Dance and Performing Arts	50	0.0%	0.0%	-0.111
36	Communication, Cultural and Media Studies, Library and Information Management	150	0.0%	0.0%	-0.269

Table A14 Precision and sensitivity of predictions for a range of REF quality profile groupings (publication year = 2008)

The previous summaries considered the predictive power of percentile of highly cited publications for identifying REF 4* work. The following table widens the population to 3* and 4* work and 2*, 3* and 4* work. This provides information on the cut-off used for the predictions and summary statistics, as defined in Table A12.

REF grouping	Prediction cut-off value	REF group indicator				Precision	Sensitivity
		1	0	1	0		
4*	1.00	1235	3560	510	14365	70.7%	25.7%
4* + 3*	25.00	11525	3530	2295	2325	83.4%	76.5%
4* + 3* + 2*	50.00	17330	1655	540	145	97.0%	91.3%

Table A15 Summary of prediction statistics for all publication years

In controlling for the effect of publication year this summary has effectively ignored 85% of the data available. The following summary indicates what the prediction statistics would look like for each year of publication. Statistical modelling of this data allows all the data available to be used in the analysis whilst controlling for year of publication effects (see ‘Publication year effect’ in report and Annex B).

Year of publication	Total	Precision	Sensitivity	Spearman correlation	Kendall's Tau-b correlation
2008	19675	70.7%	25.7%	-0.394	-0.338
2009	22120	67.8%	23.8%	-0.370	-0.318
2010	23800	66.1%	25.8%	-0.353	-0.305
2011	24815	61.4%	23.8%	-0.309	-0.264
2012	25715	38.0%	37.2%	-0.249	-0.213
2013	19180	36.2%	24.2%	-0.172	-0.156
All years	135310	62.7%	22.2%	-0.293	-0.250

Note: Counts and totals of outputs have been rounded to the nearest five for all tables in the analysis.

Summary of source normalised impact per paper as REF quality profile predictor

This analysis summarises the correlation and predictive relationship between source normalised impact per paper and the REF quality profile scores. The effect of publication year has been controlled for by focussing on the earliest year available (2008). Table A20 provides a summary of the impact of this choice of year

Table A16 Summary of source normalised impact per paper by REF quality profile (publication year = 2008)

This summarises the metric values at each level of REF quality profile. Outputs with a 0 quality profile were unclassified and either fell below the standard of nationally recognised work or did not meet the published definition of research. Outputs with missing source normalised impact per paper values were excluded from the analysis (missing = 1930).

REF quality profile	Frequency	Mean	Minimum	Median	Maximum	Missing
4	4705	2.8	0.0	2.1	16.1	335
3	10040	1.7	0.0	1.5	15.9	845
2	3745	1.4	0.0	1.3	16.1	610
1	400	1.2	0.0	1.1	6.0	125
0	245	1.9	0.0	1.6	11.9	15
Total	19130	1.9	0.0	1.5	16.1	1930

Figure A4 Scatterplot of source normalised impact per paper against REF quality profile (publication year = 2008)

(Spearman correlation= 0.3503)

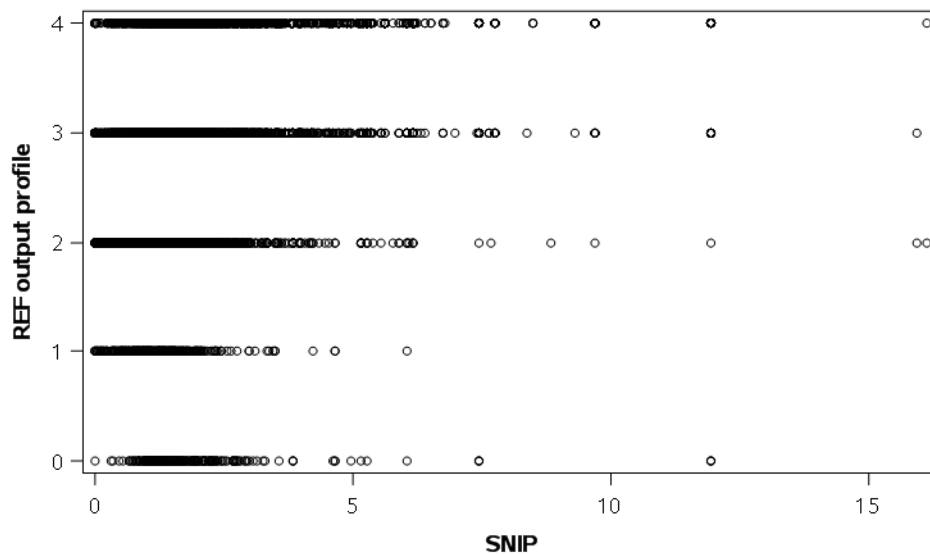


Table A17 Two-way summary of REF 4* indicator by source normalised impact per paper prediction (publication year = 2008)

(prediction = 1 when source normalised impact per paper ≥ 2.153)

REF 4*	metric prediction		Total
	1	0	
1	2200	2505	4705
0	2345	12075	14425
Total	4545	14585	19130

Precision = true positives / total positive predictions = $2200 / (2200 + 2345) = 48.4\%$

Precision can be interpreted as the proportion of correct predictions.

Sensitivity = true positives / total REF 4* records = 2200 / (2200 + 2505) = 46.7%

Sensitivity can be interpreted as the proportion of 4* records identified by the prediction.

Table A18 Summary table of precision and sensitivity of REF 4* predictions and correlation between REF quality profile and source normalised impact per paper by UOA (publication year = 2008)

UOA id	Name	Total	Precision	Sensitivity	Spearman correlation
.	Overall	19130	48.4%	46.7%	0.350
1	Clinical Medicine	2035	48.9%	82.6%	0.515
2	Public Health, Health Services and Primary Care	660	41.5%	78.6%	0.424
3	Allied Health Professions, Dentistry, Nursing and Pharmacy	1390	62.2%	35.3%	0.295
4	Psychology, Psychiatry and Neuroscience	1325	49.5%	50.7%	0.390
5	Biological Sciences	1380	70.2%	60.4%	0.551
6	Agriculture, Veterinary and Food Science	535	66.0%	28.9%	0.313
7	Earth Systems and Environmental Sciences	865	47.3%	38.5%	0.340
8	Chemistry	805	72.6%	42.3%	0.499
9	Physics	1205	45.3%	54.6%	0.317
10	Mathematical Sciences	780	49.1%	28.4%	0.309
11	Computer Science and Informatics	850	46.6%	60.9%	0.350
12	Aeronautical, Mechanical, Chemical and Manufacturing Engineering	580	27.6%	18.3%	0.165
13	Electrical and Electronic Engineering, Metallurgy and Materials	620	35.9%	49.7%	0.323
14	Civil and Construction Engineering	180	23.3%	33.3%	0.175
15	General Engineering	1270	29.1%	38.3%	0.224
16	Architecture, Built Environment and Planning	255	27.3%	5.3%	0.045
17	Geography, Environmental Studies and Archaeology	515	42.0%	27.6%	0.209
18	Economics and Econometrics	265	68.6%	67.8%	0.665
19	Business and Management Studies	1245	47.1%	43.0%	0.405
20	Law	135	25.0%	3.3%	0.130
21	Politics and International Studies	235	57.9%	21.6%	0.183
22	Social Work and Social Policy	340	36.8%	9.9%	0.291

UOA					Spearman
id	Name	Total	Precision	Sensitivity	correlation
23	Sociology	205	46.2%	20.0%	0.182
24	Anthropology and Development Studies	125	30.0%	12.5%	0.217
25	Education	350	25.0%	10.1%	0.175
26	Sport and Exercise Sciences, Leisure and Tourism	220	30.0%	6.7%	0.168
27	Area Studies	70	25.0%	6.7%	0.017
28	Modern Languages and Linguistics	100	20.0%	3.8%	0.212
29	English Language and Literature	80	0.0%	0.0%	0.187
30	History	165	16.7%	2.9%	0.147
31	Classics	5	0.0%	0.0%	-0.788
32	Philosophy	70	0.0%	0.0%	0.158
33	Theology and Religious Studies	25	0.0%	0.0%	-0.031
34	Art and Design: History, Practice and Theory	70	0.0%	0.0%	-0.070
35	Music, Drama, Dance and Performing Arts	40	0.0%	0.0%	0.185
36	Communication, Cultural and Media Studies, Library and Information Management	125	60.0%	7.7%	0.210

Table A19 Precision and sensitivity of predictions for a range of REF quality profile groupings (publication year = 2008)

The previous summaries considered the predictive power of source normalised impact per paper for identifying REF 4* work. The following table widens the population to 3* and 4* work and 2*, 3* and 4* work. This provides information on the cut-off used for the predictions and summary statistics, as defined in Table A17.

REF grouping	Prediction cut-off value	REF group indicator				Precision	Sensitivity
		1	0	1	0		
4*	2.15	2200	2505	2345	12075	48.4%	46.7%
4* + 3*	1.16	11850	2895	2615	1770	81.9%	80.4%
4* + 3* + 2*	0.55	17835	650	580	65	96.9%	96.5%

Table A20 Summary of prediction statistics for all publication years

In controlling for the effect of publication year this summary has effectively ignored 85% of the data available. The following summary indicates what the prediction statistics would look like for each year of publication. Statistical modelling of this data allows all the data available to be used in the analysis whilst controlling for year of publication effects (see ‘Publication year effect’ in report and Annex B).

Year of publication	Total	Precision	Sensitivity	Spearman correlation	Kendall's Tau-b correlation
2008	19130	51.4%	44.8%	0.350	0.275
2009	21495	48.6%	44.4%	0.340	0.268
2010	22995	47.1%	44.9%	0.338	0.267
2011	24195	42.4%	43.1%	0.315	0.248
2012	23995	42.5%	43.2%	0.305	0.240
2013	235	30.0%	33.3%	0.136	0.107
All years	112045	45.5%	44.0%	0.327	0.257

Note: Counts and totals of outputs have been rounded to the nearest five for all tables in the analysis.

Summary of SCImago journal rank as REF quality profile predictor

This analysis summarises the correlation and predictive relationship between SCImago journal rank and the REF quality profile scores. The effect of publication year has been controlled for by focussing on the earliest year available (2008). Table A25 provides a summary of the impact of this choice of year.

Table A21 Summary of SCImago journal rank by REF quality profile (publication year = 2008)

This summarises the metric values at each level of REF quality profile. Outputs with a 0 quality profile were unclassified and either fell below the standard of nationally recognised work or did not meet the published definition of research. Outputs with missing SCImago journal rank values were excluded from the analysis (missing = 1815).

REF quality profile	Frequency	Mean	Minimum	Median	Maximum	Missing
4	4710	4.2	0.0	2.8	19.1	330
3	10050	2.1	0.0	1.7	19.0	835
2	3830	1.4	0.0	1.1	16.3	525
1	410	1.0	0.0	0.7	11.5	115
0	245	2.1	0.0	1.5	16.3	10
Total	19245	2.5	0.0	1.7	19.1	1815

Figure A5 Scatterplot of SCImago journal rank against REF quality profile (publication year = 2008)

(Spearman correlation= 0.3723)

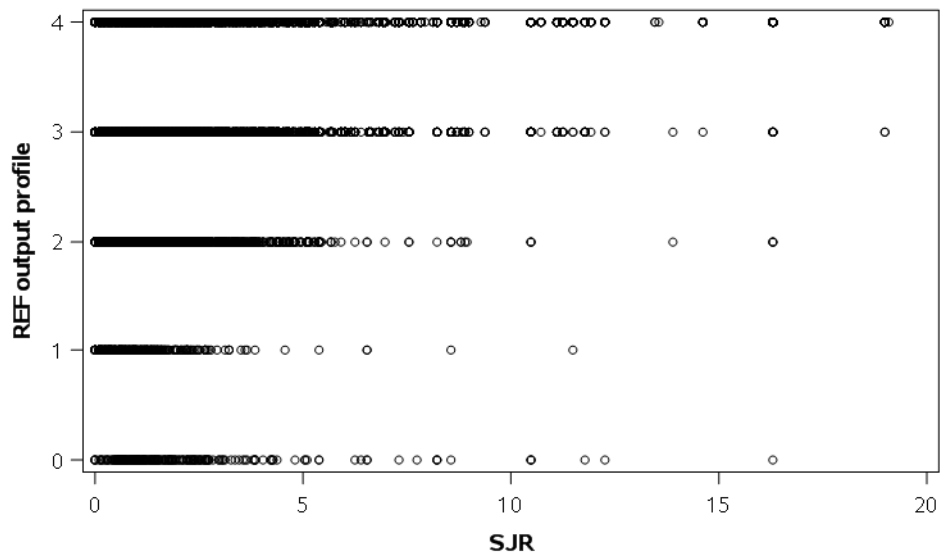


Table A22 Two-way summary of REF 4* indicator by SCImago journal rank prediction (publication year = 2008)

(prediction = 1 when SCImago journal rank \geq 3.057)

REF 4* output	metric prediction		Total
	1	0	
1	2225	2485	4710
0	2375	12160	14535
Total	4600	14645	19245

Precision = true positives / total positive predictions = 2225 / (2225 + 2375) = 48.4%

Precision can be interpreted as the proportion of correct predictions.

Sensitivity = true positives / total REF 4* records = 2225 / (2225 + 2485) = 47.2%

Sensitivity can be interpreted as the proportion of 4* records identified by the prediction.

Table A23 Summary table of precision and sensitivity of REF 4* predictions and correlation between REF quality profile and SCImago journal rank by UOA (publication year = 2008)

UOA id	Name	Total	Precision	Sensitivity	Spearman correlation
.	Overall	19245	48.4%	47.2%	0.372
1	Clinical Medicine	2030	40.0%	84.5%	0.562
2	Public Health, Health Services and Primary Care	645	50.3%	47.6%	0.386
3	Allied Health Professions, Dentistry, Nursing and Pharmacy	1395	58.6%	43.8%	0.418
4	Psychology, Psychiatry and Neuroscience	1305	50.8%	53.3%	0.437
5	Biological Sciences	1385	51.7%	89.8%	0.612
6	Agriculture, Veterinary and Food Science	540	58.3%	43.0%	0.445
7	Earth Systems and Environmental Sciences	885	47.9%	41.5%	0.389
8	Chemistry	805	47.2%	83.5%	0.511
9	Physics	1075	44.5%	64.2%	0.392
10	Mathematical Sciences	780	50.0%	27.5%	0.301
11	Computer Science and Informatics	850	61.0%	30.1%	0.421
12	Aeronautical, Mechanical, Chemical and Manufacturing Engineering	585	47.1%	13.6%	0.166
13	Electrical and Electronic Engineering, Metallurgy and Materials	620	47.0%	34.6%	0.383
14	Civil and Construction Engineering	185	60.0%	10.0%	0.333
15	General Engineering	1265	40.0%	23.5%	0.284
16	Architecture, Built Environment and Planning	260	0.0%	0.0%	0.032
17	Geography, Environmental Studies and Archaeology	520	45.2%	13.0%	0.215
18	Economics and Econometrics	265	71.1%	73.6%	0.751
19	Business and Management Studies	1265	61.9%	30.6%	0.431
20	Law	155	0.0%	0.0%	0.074
21	Politics and International Studies	285	62.5%	9.3%	0.250

UOA					Spearman
id	Name	Total	Precision	Sensitivity	correlation
22	Social Work and Social Policy	360	0.0%	0.0%	0.302
23	Sociology	220	50.0%	6.3%	0.135
24	Anthropology and Development Studies	130	0.0%	0.0%	0.184
25	Education	360	42.9%	3.7%	0.154
26	Sport and Exercise Sciences, Leisure and Tourism	230	50.0%	4.3%	0.362
27	Area Studies	85	0.0%	0.0%	0.013
28	Modern Languages and Linguistics	105	25.0%	3.7%	0.172
29	English Language and Literature	95	0.0%	0.0%	0.261
30	History	180	0.0%	0.0%	0.190
31	Classics	5	0.0%	0.0%	-0.381
32	Philosophy	80	0.0%	0.0%	-0.143
33	Theology and Religious Studies	30	0.0%	0.0%	-0.119
34	Art and Design: History, Practice and Theory	90	0.0%	0.0%	0.117
35	Music, Drama, Dance and Performing Arts	45	0.0%	0.0%	0.015
36	Communication, Cultural and Media Studies, Library and Information Management	140	0.0%	0.0%	0.199

Table A24 Precision and sensitivity of predictions for a range of REF quality profile groupings (publication year = 2008)

The previous summaries considered the predictive power of SCImago journal rank for identifying REF 4* work. The following table widens the population to 3* and 4* work and 2*, 3* and 4* work. This provides information on the cut-off used for the predictions and summary statistics, as defined in Table A22.

REF grouping	Prediction cut-off value	REF group indicator				Precision	Sensitivity
		1		0			
		metric	prediction	metric	prediction		
4*	3.06	2225	2485	2375	12160	48.4%	47.2%
4* + 3*	0.92	11955	2805	2590	1895	82.2%	81.0%
4* + 3* + 2*	0.16	17925	665	605	50	96.7%	96.4%

Table A25 Summary of prediction statistics for all publication years

By controlling for the effect of publication year this summary has effectively ignored 85% of the data available. The following summary indicates what the prediction statistics would look like for each year of publication. Statistical modelling of this data allows all the data available to be used in the analysis whilst controlling for year of publication effects (see 'Publication year effect' in report and Annex B)

Year of publication	Total	Precision	Sensitivity	Spearman correlation	Kendall's Tau-b correlation
2008	19245	51.4%	45.1%	0.372	0.294
2009	21685	48.7%	44.0%	0.355	0.281
2010	23205	47.5%	46.2%	0.345	0.274
2011	24220	44.6%	45.7%	0.323	0.256
2012	23860	42.4%	43.9%	0.311	0.246
2013	235	28.6%	31.1%	0.177	0.138
All years	112450	46.4%	45.1%	0.340	0.269

Note: Counts and totals of outputs have been rounded to the nearest five for all tables in the analysis.

Summary of collaboration as REF quality profile predictor

Table A26 Summary of collaboration by REF quality profile (publication year = 2008)

This summarises the metric values at each level of REF quality profile. Outputs with a 0 quality profile were unclassified and either fell below the standard of nationally recognised work or did not meet the published definition of research. Outputs with missing collaboration values were excluded from the analysis (missing = 1385).

	Collaboration											
	Single											
	Author		Institutional		National		International		Other		Total	
	No	%	No	%	No	%	No	%	No	%	No	%
4	390	8.1	895	18.7	1085	22.7	2395	49.9	30	0.6	4795	100.0
3	985	9.6	2335	22.7	2665	26.0	4205	41.0	75	0.7	10260	100.0
2	600	15.3	940	24.0	1165	29.7	1195	30.3	30	0.7	3930	100.0
1	85	19.8	90	20.5	145	32.7	115	26.4	5	0.7	440	100.0
0	30	11.7	55	21.8	75	30.2	90	35.9	0	0.4	250	100.0
Total	2090	10.6	4315	21.9	5135	26.1	7995	40.6	135	0.7	19675	100.0

Note: Counts and totals of outputs have been rounded to the nearest five for all tables in the analysis.

Further analysis was not completed for this metric, as there was no obvious choice for the direction of the relationship between collaboration and REF quality score (as there was no clear hierarchy to categorical responses).

Summary of number of authors as REF quality profile predictor

This analysis summarises the correlation and predictive relationship between number of authors and the REF quality profile scores. The effect of publication year has been controlled for by focussing on the earliest year available (2008). Table A28 provides a summary of the impact for this choice of year.

Table A27 Summary of number of authors by REF quality profile (publication year = 2008)

This summarises the metric values at each level of REF quality profile. Outputs with a 0 quality profile were unclassified and either fell below the standard of nationally recognised work or did not meet the published definition of research. There were no outputs with missing number of authors .

REF quality profile	Frequency	Mean	Minimum	Median	Maximum
4	5040	8.0	0.0	4.0	3,096.0
3	10885	9.7	0.0	4.0	3,096.0
2	4355	4.1	0.0	3.0	527.0
1	525	3.3	0.0	2.0	156.0
0	260	18.1	0.0	3.0	2,926.0
Total	21060	8.1	0.0	4.0	3,096.0

Figure A6 Scatterplot of number of authors against REF quality profile (publication year = 2008)

(Spearman correlation= 0.1679)

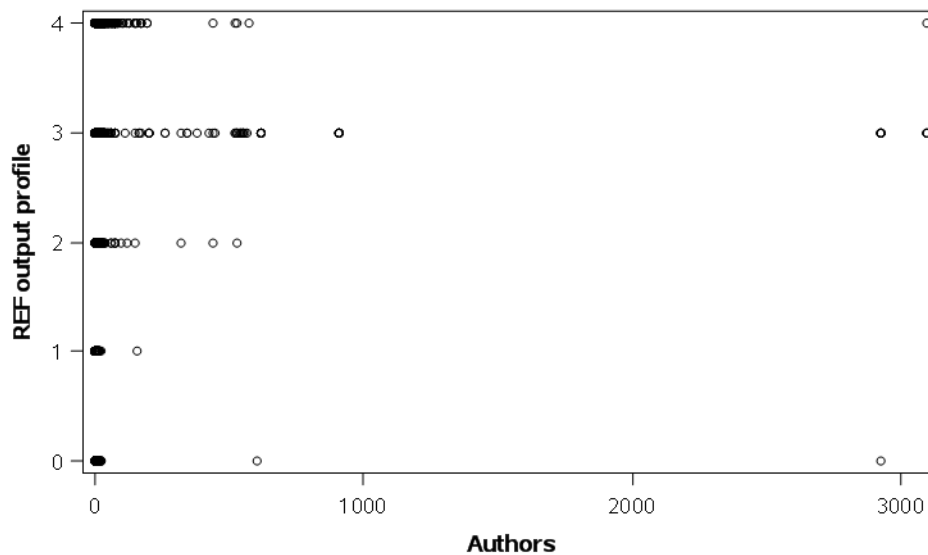


Table A28 Summary of correlations for all publication years

In controlling for the effect of publication year this summary has effectively ignored 85% of the data available. The following summary indicates what the prediction statistics would look like for each year of publication. Statistical modelling of this data allows all the data available to be used in the analysis whilst controlling for year of publication effects (see ‘Publication year effect’ in report and Annex B).

Year of publication	Total	Spearman correlation	Kendall's Tau-b correlation
2008	21060	0.168	0.136
2009	23715	0.166	0.135
2010	25195	0.165	0.135
2011	26545	0.148	0.121
2012	27825	0.143	0.117
2013	24415	0.114	0.094
All years	148755	0.151	0.123

Note: Counts and totals of outputs have been rounded to the nearest five for all tables in the analysis.

Further analysis was not completed for this metric, as there was no obvious choice for the direction of the relationship between number of authors and REF quality score (as a high metric value did not necessarily imply high quality research).

Summary of number of countries associated with authors as REF quality profile predictor

This analysis summarises the correlation and predictive relationship between number of countries associated with authors and the REF quality profile scores. The effect of publication year has been controlled for by focussing on the earliest year available (2008). Table A30 provides a summary of the impact of this choice of year.

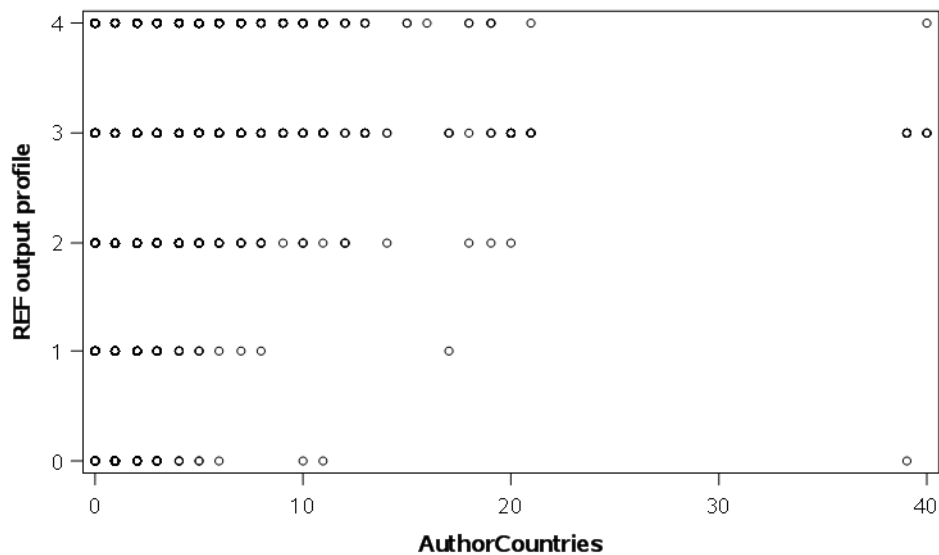
Table A29 Summary of number of countries associated with authors by REF quality profile (publication year = 2008)

This summarises the metric values at each level of REF quality profile. Outputs with a 0 quality profile were unclassified and either fell below the standard of nationally recognised work or did not meet the published definition of research. There were no outputs with missing number of countries associated with authors.

REF quality profile	Frequency	Mean	Minimum	Median	Maximum
4	5040	1.9	0.0	1.0	40.0
3	10885	1.6	0.0	1.0	40.0
2	4355	1.3	0.0	1.0	20.0
1	525	1.1	0.0	1.0	17.0
0	260	1.6	0.0	1.0	39.0
Total	21060	1.6	0.0	1.0	40.0

Figure A7 Scatterplot of number of countries associated with authors against REF quality profile (publication year = 2008)

(Spearman correlation= 0.1636)



Note: Counts and totals of outputs have been rounded to the nearest five for all tables in the analysis.

Table A30 Summary of correlations for all publication years

In controlling for the effect of publication year this summary has effectively ignored 85% of the data available. The following summary indicates what the prediction statistics would look like for each year of publication. Statistical modelling of this data allows all the data available to be used in the analysis whilst controlling for year of publication effects (see ‘Publication year effect’ in report and Annex B).

Year of publication	Total	Spearman correlation	Kendall's Tau-b correlation
2008	21060	0.164	0.144
2009	23715	0.156	0.138
2010	25195	0.148	0.131
2011	26545	0.141	0.125
2012	27825	0.145	0.128
2013	24415	0.100	0.088
All years	148755	0.142	0.125

Note: Counts and totals of outputs have been rounded to the nearest five for all tables in the analysis.

Further analysis was not completed for this metric, as there was no obvious choice for the direction of the relationship between number of countries associated with authors and REF quality score (as a high metric value did not necessarily imply high quality research).

Summary of academic and corporate authors as REF quality profile predictor

Table A31 Summary of academic and corporate authors by REF quality profile (publication year = 2008)

This summarises the metric values at each level of REF quality profile. Outputs with a 0 quality profile were unclassified and either fell below the standard of nationally recognised work or did not meet the published definition of research. Outputs with missing academic and corporate authors values were excluded from the analysis (missing = 1385).

	CrossAcademicCorporate					
	No		Yes		Total	
	No	%	No	%	No	%
4	4480	93.5	315	6.5	4795	100.0
3	9770	95.2	490	4.8	10260	100.0
2	3805	96.8	125	3.2	3930	100.0
1	425	97.0	15	3.0	440	100.0
0	235	94.8	15	5.2	250	100.0
Total	18725	95.2	950	4.8	19675	100.0

Note: Counts and totals of outputs have been rounded to the nearest five for all tables in the analysis.

Further analysis was not completed for this metric, as there was no obvious choice for the direction of the relationship between academic and corporate authors and REF quality score (as. there was no clear hierarchy to categorical responses).

Summary of WIPO patent citations as REF quality profile predictor

This analysis summarises the correlation and predictive relationship between World Intellectual Property Organisation (WIPO) patent citations and the REF quality profile scores. The effect of publication year has been controlled for by focussing on the earliest year available (2008). Table A36 provides a summary of the impact of this choice of year.

Table A32 Summary of WIPO patent citations by REF quality profile (publication year = 2008)

This summarises the metric values at each level of REF quality profile. Outputs with a 0 quality profile were unclassified and either fell below the standard of nationally recognised work or did not meet the published definition of research. There were no outputs with missing WIPO patent citations.

REF quality profile	Frequency	Mean	Minimum	Median	Maximum
4	5040	0.1	0.0	0.0	8.0
3	10885	0.0	0.0	0.0	10.0
2	4355	0.0	0.0	0.0	4.0
1	525	0.0	0.0	0.0	1.0
0	260	0.0	0.0	0.0	4.0
Total	21060	0.0	0.0	0.0	10.0

Note: While the missing statistics indicate that there was good coverage for this metric, over 90 per cent of the data returned was recorded as 0 which indicates a highly skewed distribution.

Figure A8 Scatterplot of WIPO patent citations against REF quality profile (publication year = 2008)

(Spearman correlation= 0.1077)

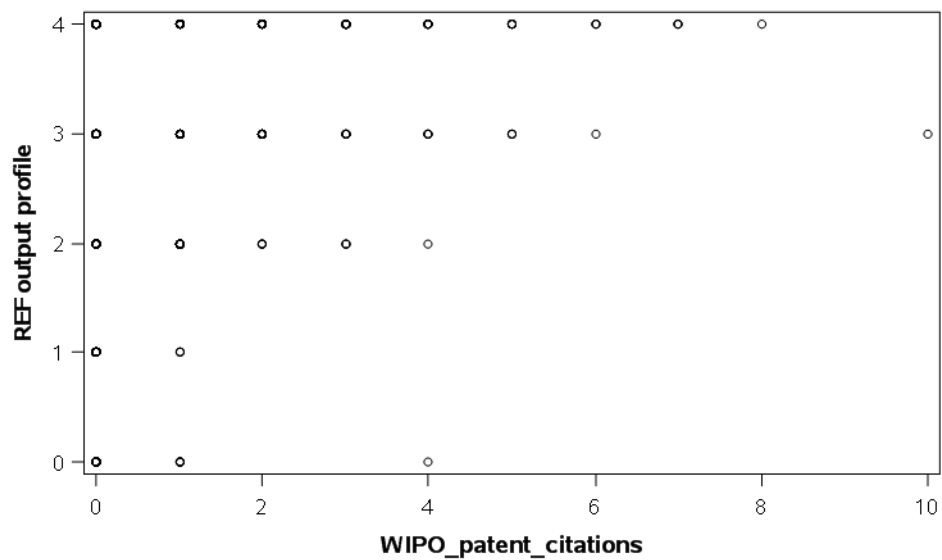


Table A33 Two-way summary of REF 4* indicator by WIPO patent citations prediction (publication year = 2008)

(prediction = 1 when WIPO patent citations \geq 1)

REF 4* output	metric prediction		Total
	1	0	
1	330	4710	5040
0	315	15700	16020
Total	645	20415	21060

Precision = true positives / total positive predictions = $330 / (330 + 315) = 50.9\%$

Precision can be interpreted as the proportion of correct predictions.

Sensitivity = true positives / total REF 4* records = $330 / (330 + 4710) = 6.5\%$

Sensitivity can be interpreted as the proportion of 4* records identified by the prediction.

Table A34 Summary table of precision and sensitivity of REF 4* predictions and correlation between REF quality profile and WIPO patent citations by UOA (publication year = 2008)

UOA id	Name	Total	Precision	Sensitivity	Spearman correlation
.	Overall	21060	50.9%	6.5%	0.108
1	Clinical Medicine	2070	55.3%	24.3%	0.229
2	Public Health, Health Services and Primary Care	715	53.3%	4.5%	0.071
3	Allied Health Professions, Dentistry, Nursing and Pharmacy	1455	40.3%	7.6%	0.084
4	Psychology, Psychiatry and Neuroscience	1365	37.0%	2.4%	0.035
5	Biological Sciences	1400	59.6%	11.1%	0.118
6	Agriculture, Veterinary and Food Science	550	33.3%	5.2%	0.071
7	Earth Systems and Environmental Sciences	920	77.8%	3.6%	0.119
8	Chemistry	815	55.6%	12.9%	0.165
9	Physics	1225	75.0%	3.8%	0.110
10	Mathematical Sciences	830	0.0%	0.0%	-0.050
11	Computer Science and Informatics	1035	56.5%	4.5%	0.113
12	Aeronautical, Mechanical, Chemical and Manufacturing Engineering	610	40.0%	5.0%	0.056

UOA id	Name	Total	Precision	Sensitivity	Spearman correlation
13	Electrical and Electronic Engineering, Metallurgy and Materials	640	50.0%	9.8%	0.132
14	Civil and Construction Engineering	195	33.3%	3.2%	0.079
15	General Engineering	1330	26.7%	5.2%	0.052
16	Architecture, Built Environment and Planning	300	0.0%	0.0%	.
17	Geography, Environmental Studies and Archaeology	550	0.0%	0.0%	.
18	Economics and Econometrics	275	0.0%	0.0%	.
19	Business and Management Studies	1400	0.0%	0.0%	0.008
20	Law	240	0.0%	0.0%	.
21	Politics and International Studies	325	0.0%	0.0%	.
22	Social Work and Social Policy	420	0.0%	0.0%	.
23	Sociology	250	0.0%	0.0%	.
24	Anthropology and Development Studies	150	0.0%	0.0%	.
25	Education	455	0.0%	0.0%	.
26	Sport and Exercise Sciences, Leisure and Tourism	255	50.0%	2.0%	0.080
27	Area Studies	110	0.0%	0.0%	.
28	Modern Languages and Linguistics	175	0.0%	0.0%	.
29	English Language and Literature	170	0.0%	0.0%	.
30	History	245	0.0%	0.0%	.
31	Classics	15	0.0%	0.0%	.
32	Philosophy	115	0.0%	0.0%	.
33	Theology and Religious Studies	45	0.0%	0.0%	.
34	Art and Design: History, Practice and Theory	130	0.0%	0.0%	.
35	Music, Drama, Dance and Performing Arts	90	0.0%	0.0%	.
36	Communication, Cultural and Media Studies, Library and Information Management	195	0.0%	0.0%	.

Table A35 Precision and sensitivity of predictions for a range of REF quality profile groupings (publication year = 2008)

The previous summaries considered the predictive power of WIPO patent citations for identifying REF 4* work. The following table widens the population to 3* and 4* work and

2*, 3* and 4* work. This provides information on the cut-off used for the predictions and summary statistics, as defined in Table A33.

REF grouping	Prediction cut-off value	REF group indicator				Precision	Sensitivity
		1		0			
		metric	prediction	1	0		
4*	1.00	330	4710	315	15700	50.9%	6.5%
4* + 3*	1.00	585	15340	60	5075	90.9%	3.7%
4* + 3* + 2*	1.00	635	19640	10	770	98.5%	3.1%

Table A36 Summary of prediction statistics for all publication years

In controlling for the effect of publication year this summary has effectively ignored 85% of the data available. The following summary indicates what the prediction statistics would look like for each year of publication. Statistical modelling of this data allows all the data available to be used in the analysis whilst controlling for year of publication effects (see ‘Publication year effect’ in report and Annex B).

Year of publication	Total	Precision	Sensitivity	Spearman correlation	Kendall's Tau-b correlation
2008	21060	50.9%	6.5%	0.108	0.101
2009	23715	46.4%	4.9%	0.085	0.080
2010	25195	47.3%	4.1%	0.079	0.075
2011	26545	42.7%	1.7%	0.051	0.048
2012	27825	51.4%	0.3%	0.025	0.024
2013	24415	0.0%	0.0%	.	.
All years	148755	47.8%	2.9%	0.069	0.065

Note: Counts and totals of outputs have been rounded to the nearest five for all tables in the analysis.

Summary of Mendeley readership as REF quality profile predictor

This analysis summarises the correlation and predictive relationship between Mendeley readership and the REF quality profile scores. The effect of publication year has been controlled for by focussing on the earliest year available (2008). Table A41 provides a summary of the impact of this choice of year.

Please note that the Mendeley Readership data is only collected for those articles where Altmetric.com has at least one other altmetric available for the article.

Table A37 Summary of Mendeley readership by REF quality profile (publication year = 2008)

This summarises the metric values at each level of REF quality profile. Outputs with a 0 quality profile were unclassified and either fell below the standard of nationally recognised work or did not meet the published definition of research. Outputs with missing Mendeley readership values were excluded from the analysis (missing = 10).

REF quality profile	Frequency	Mean	Minimum	Median	Maximum
4	5035	26.8	0.0	0.0	1,090.0
3	10880	5.1	0.0	0.0	862.0
2	4350	3.1	0.0	0.0	340.0
1	525	4.2	0.0	0.0	697.0
0	260	15.8	0.0	0.0	509.0
Total	21050	10.0	0.0	0.0	1,090.0

Note: While the missing statistics indicate that there was good coverage for this metric, around 85 per cent of the data returned was recorded as 0 which indicates a highly skewed distribution.

Figure A9 Scatterplot of Mendeley readership against REF quality profile (publication year = 2008)

(Spearman correlation= 0.1900)

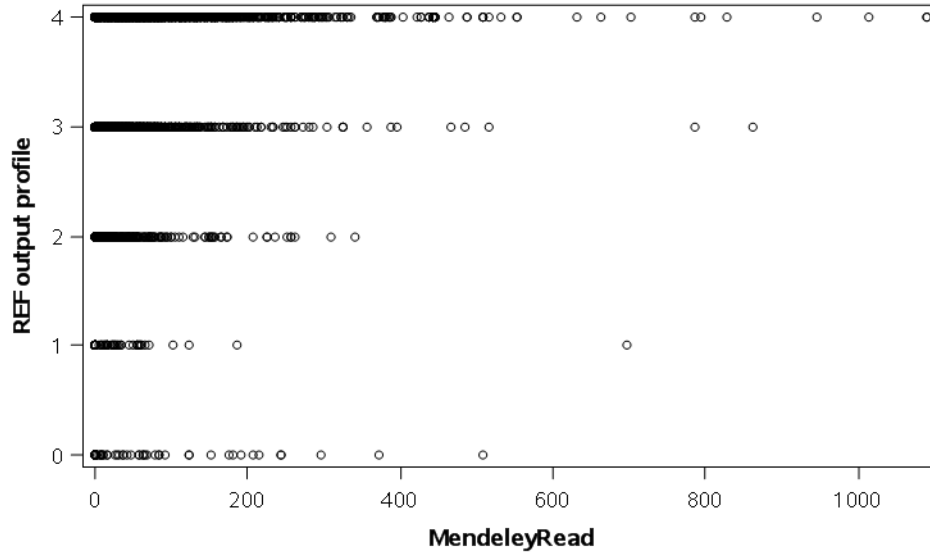


Table A38 Two-way summary of REF 4* indicator by Mendeley readership prediction (publication year = 2008)

(prediction = 1 when Mendeley readership \geq 1)

REF 4*	metric prediction		Total
	1	0	
output 1	1310	3725	5035
0	1555	14455	16010
Total	2870	18180	21050

Precision = true positives / total positive predictions = $1310 / (1310 + 1555) = 45.7\%$

Precision can be interpreted as the proportion of correct predictions.

Sensitivity = true positives / total REF 4* records = $1310 / (1310 + 3725) = 26.1\%$

Sensitivity can be interpreted as the proportion of 4* records identified by the prediction.

Table A39 Summary table of precision and sensitivity of REF 4* predictions and correlation between REF quality profile and Mendeley readership by UOA (publication year = 2008)

UOA id	Name	Total	Precision	Sensitivity	Spearman correlation
.	Overall	21050	45.7%	26.1%	0.190
1	Clinical Medicine	2070	55.0%	60.0%	0.441
2	Public Health, Health Services and Primary Care	715	39.7%	52.8%	0.233
3	Allied Health Professions, Dentistry, Nursing and Pharmacy	1450	45.5%	26.4%	0.195
4	Psychology, Psychiatry and Neuroscience	1365	47.0%	39.6%	0.207
5	Biological Sciences	1400	60.2%	51.6%	0.363
6	Agriculture, Veterinary and Food Science	550	46.5%	34.5%	0.252
7	Earth Systems and Environmental Sciences	915	44.2%	36.7%	0.269
8	Chemistry	815	52.4%	17.0%	0.190
9	Physics	1225	38.1%	10.1%	0.092
10	Mathematical Sciences	830	41.5%	8.7%	0.085
11	Computer Science and Informatics	1035	40.8%	10.1%	0.053
12	Aeronautical, Mechanical, Chemical and Manufacturing Engineering	610	22.2%	3.3%	0.006
13	Electrical and Electronic Engineering, Metallurgy and Materials	640	69.2%	5.5%	0.138
14	Civil and Construction Engineering	195	50.0%	16.1%	0.174
15	General Engineering	1330	35.7%	4.3%	0.041
16	Architecture, Built Environment and Planning	300	40.0%	6.2%	0.105
17	Geography, Environmental Studies and Archaeology	550	36.1%	20.2%	0.167
18	Economics and Econometrics	275	63.3%	21.3%	0.222
19	Business and Management Studies	1400	22.5%	6.1%	0.039
20	Law	240	15.0%	5.7%	0.034
21	Politics and International Studies	325	14.8%	6.8%	0.010
22	Social Work and Social Policy	420	13.5%	6.3%	-0.014
23	Sociology	250	28.2%	30.6%	0.128
24	Anthropology and Development Studies	150	14.3%	12.5%	0.110
25	Education	455	18.2%	11.0%	0.049

UOA id	Name	Total	Precision	Sensitivity	Spearman correlation
26	Sport and Exercise Sciences, Leisure and Tourism	255	22.5%	18.4%	0.020
27	Area Studies	110	25.0%	8.0%	0.002
28	Modern Languages and Linguistics	175	25.0%	2.9%	0.035
29	English Language and Literature	170	16.7%	2.6%	0.043
30	History	245	33.3%	10.9%	0.121
31	Classics	15	0.0%	0.0%	.
32	Philosophy	115	25.0%	3.7%	-0.029
33	Theology and Religious Studies	45	0.0%	0.0%	0.018
34	Art and Design: History, Practice and Theory	130	75.0%	15.8%	0.232
35	Music, Drama, Dance and Performing Arts	90	0.0%	0.0%	-0.073
36	Communication, Cultural and Media Studies, Library and Information Management	195	39.4%	26.0%	0.102

Table A40 Precision and sensitivity of predictions for a range of REF quality profile groupings (publication year = 2008)

The previous summaries considered the predictive power of Mendeley readership for identifying REF 4* work. The following table widens the population to 3* and 4* work and 2*, 3* and 4* work. This provides information on the cut-off used for the predictions and summary statistics, as defined in Table A38.

REF grouping	Prediction cut-off value	REF group indicator				Precision	Sensitivity
		1	0	1	0		
4*	1.00	1310	3725	1555	14455	45.7%	26.1%
4* + 3*	1.00	2455	13460	415	4715	85.5%	15.4%
4* + 3* + 2*	1.00	2790	17475	80	700	97.2%	13.8%

Table A41 Summary of prediction statistics for all publication years

In controlling for the effect of publication year this summary has effectively ignored 85% of the data available. The following summary indicates what the prediction statistics would look like for each year of publication. Statistical modelling of this data allows all the data available

to be used in the analysis whilst controlling for year of publication effects (see ‘Publication year effect’ in report and Annex B).

Year of publication	Total	Precision	Sensitivity	Spearman correlation	Kendall's Tau-b correlation
2008	21050	45.7%	26.1%	0.190	0.173
2009	23700	43.5%	28.3%	0.188	0.171
2010	25195	40.3%	32.0%	0.192	0.174
2011	26530	35.7%	37.1%	0.173	0.153
2012	27815	37.6%	39.5%	0.177	0.151
2013	24395	38.2%	38.5%	0.196	0.167
All years	148685	36.3%	35.6%	0.165	0.146

Note: Counts and totals of outputs have been rounded to the nearest five for all tables in the analysis.

Summary of ScienceDirect downloads as REF quality profile predictor

This analysis summarises the correlation and predictive relationship between ScienceDirect downloads and the REF quality profile scores. The effect of publication year has been controlled for by focussing on the earliest year available (2008). Table A54 provides a summary of the impact for this choice of year.

Table A42 Summary of ScienceDirect downloads by REF quality profile (publication year = 2008)

This summarises the metric values at each level of REF quality profile. Outputs with a 0 quality profile were unclassified and either fell below the standard of nationally recognised work or did not meet the published definition of research. Outputs with missing ScienceDirect downloads values were excluded from the analysis (missing = 14070).

REF quality profile	Frequency	Mean	Minimum	Median	Maximum	Missing
4	1825	2,163.4	0.0	44.0	29,073.0	3215
3	3330	1,385.0	0.0	948.0	40,412.0	7555
2	1540	1,167.8	0.0	797.5	13,105.0	2810
1	160	1,461.4	0.0	0.0	24,726.0	360
0	125	2,169.9	0.0	1,482.0	18,251.0	130
Total	6990	1,556.5	0.0	829.5	40,412.0	14070

Figure A10 Scatterplot of ScienceDirect downloads against REF quality profile (publication year = 2008)

(Spearman correlation= 0.0289)

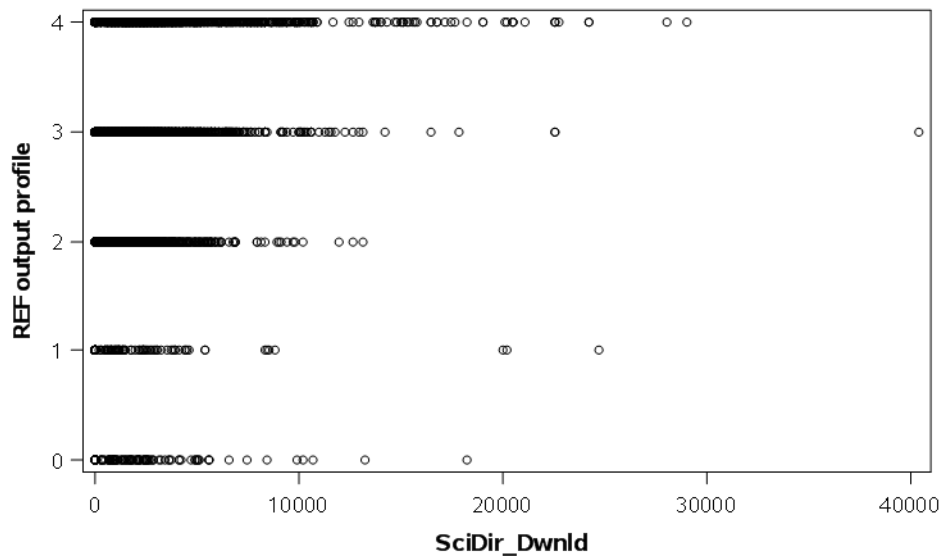


Table A43 Two-way summary of REF 4* indicator by ScienceDirect downloads prediction (publication year = 2008)

(prediction = 1 when ScienceDirect downloads \geq 2195)

REF 4* output	metric prediction		Total
	1	0	
1	605	1225	1825
0	1070	4095	5160
Total	1670	5315	6990

Precision = true positives / total positive predictions = 605 / (605 + 1070) = 36.1%

Precision can be interpreted as the proportion of correct predictions.

Sensitivity = true positives / total REF 4* records = 605 / (605 + 1225) = 33.0%

Sensitivity can be interpreted as the proportion of 4* records identified by the prediction.

Table A44 Summary table of precision and sensitivity of REF 4* predictions and correlation between REF quality profile and ScienceDirect downloads by UOA (publication year = 2008)

UOA id	Name	Total	Precision	Sensitivity	Spearman correlation
.	Overall	6990	36.1%	33.0%	0.029
1	Clinical Medicine	640	64.7%	44.2%	0.158
2	Public Health, Health Services and Primary Care	200	54.5%	77.1%	0.423
3	Allied Health Professions, Dentistry, Nursing and Pharmacy	480	28.3%	39.8%	0.043
4	Psychology, Psychiatry and Neuroscience	705	39.9%	37.8%	0.019
5	Biological Sciences	455	65.2%	39.6%	0.108
6	Agriculture, Veterinary and Food Science	210	27.7%	31.0%	-0.065
7	Earth Systems and Environmental Sciences	360	28.2%	22.5%	-0.094
8	Chemistry	120	12.0%	7.0%	-0.593
9	Physics	180	33.3%	2.9%	-0.313
10	Mathematical Sciences	285	50.0%	3.7%	-0.060
11	Computer Science and Informatics	290	41.5%	27.0%	0.139
12	Aeronautical, Mechanical, Chemical and Manufacturing Engineering	305	15.9%	18.5%	-0.027
13	Electrical and Electronic Engineering, Metallurgy and Materials	165	38.2%	35.1%	-0.008
14	Civil and Construction Engineering	85	27.3%	54.5%	0.116
15	General Engineering	550	15.2%	18.3%	-0.077
16	Architecture, Built Environment and Planning	145	25.0%	44.8%	0.016
17	Geography, Environmental Studies and Archaeology	230	27.6%	39.0%	0.089
18	Economics and Econometrics	120	32.1%	29.0%	0.110
19	Business and Management Studies	535	23.1%	50.0%	0.116
20	Law	25	0.0%	0.0%	0.049
21	Politics and International Studies	95	0.0%	0.0%	-0.010

UOA					Spearman
id	Name	Total	Precision	Sensitivity	correlation
22	Social Work and Social Policy	130	30.8%	32.0%	0.250
23	Sociology	65	20.0%	10.0%	0.039
24	Anthropology and Development Studies	65	25.0%	23.1%	-0.060
25	Education	195	25.0%	9.3%	0.050
26	Sport and Exercise Sciences, Leisure and Tourism	80	18.5%	50.0%	0.259
27	Area Studies	30	33.3%	14.3%	-0.176
28	Modern Languages and Linguistics	35	0.0%	0.0%	0.024
29	English Language and Literature	20	0.0%	0.0%	-0.045
30	History	85	0.0%	0.0%	-0.078
31	Classics	0	0.0%	0.0%	.
32	Philosophy	15	0.0%	0.0%	0.440
33	Theology and Religious Studies	5	0.0%	0.0%	-0.333
34	Art and Design: History, Practice and Theory	30	0.0%	0.0%	-0.156
35	Music, Drama, Dance and Performing Arts	15	0.0%	0.0%	-0.309
36	Communication, Cultural and Media Studies, Library and Information Management	35	0.0%	0.0%	-0.136

Table A45 Precision and sensitivity of predictions for a range of REF quality profile groupings (publication year = 2008)

The previous summaries considered the predictive power of ScienceDirect downloads for identifying REF 4* work. The following table widens the population to 3* and 4* work and 2*, 3* and 4* work. This provides information on the cut-off used for the predictions and summary statistics, as defined in Table A43.

REF grouping	Prediction cut-off value	REF group indicator				Precision	Sensitivity
		1		0			
		metric	prediction	metric	prediction		
4*	2195.00	605	1225	1070	4095	36.1%	33.0%
4* + 3*	1.00	3105	2055	1105	725	73.7%	60.2%
4* + 3* + 2*	1.00	4035	2665	175	115	95.8%	60.2%

Table A46 Summary of prediction statistics for all publication years

In controlling for the effect of publication year this summary has effectively ignored 85% of the data available. The following summary indicates what the prediction statistics would look like for each year of publication. Statistical modelling of this data allows all the data available to be used in the analysis whilst controlling for year of publication effects (see ‘Publication year effect’ in report and Annex B).

Year of publication	Total	Precision	Sensitivity	Spearman correlation	Kendall's Tau-b correlation
2008	6990	37.6%	30.9%	0.029	0.025
2009	7235	33.4%	30.1%	0.009	0.008
2010	6945	36.8%	31.1%	-0.004	-0.001
2011	7305	34.5%	29.4%	-0.025	-0.019
2012	7045	33.2%	29.5%	-0.048	-0.038
2013	5695	33.6%	29.4%	-0.042	-0.033
All years	41210	34.0%	29.4%	-0.017	-0.013

Note: Counts and totals of outputs have been rounded to the nearest five for all tables in the analysis.

Summary of Scopus full text requests as REF quality profile predictor

This analysis summarises the correlation and predictive relationship between Scopus full text requests and the REF quality profile scores. The effect of publication year has been controlled for by focussing on the earliest year available (2008). Table A51 provides a summary of the impact for this choice of year.

Table A47 Summary of Scopus full text requests by REF quality profile (publication year = 2008)

This summarises the metric values at each level of REF quality profile. Outputs with a 0 quality profile were unclassified and either fell below the standard of nationally recognised work or did not meet the published definition of research. There were no outputs with missing Scopus full text requests.

REF quality profile	Frequency	Mean	Minimum	Median	Maximum
4	5040	20.4	0.0	12.0	830.0
3	10885	13.5	0.0	8.0	252.0
2	4355	11.8	0.0	7.0	271.0
1	525	15.3	0.0	6.0	971.0
0	260	19.2	0.0	10.0	144.0
Total	21060	14.9	0.0	9.0	971.0

Figure A11 Scatterplot of Scopus full text requests against REF quality profile (publication year = 2008)

(Spearman correlation= 0.1431)

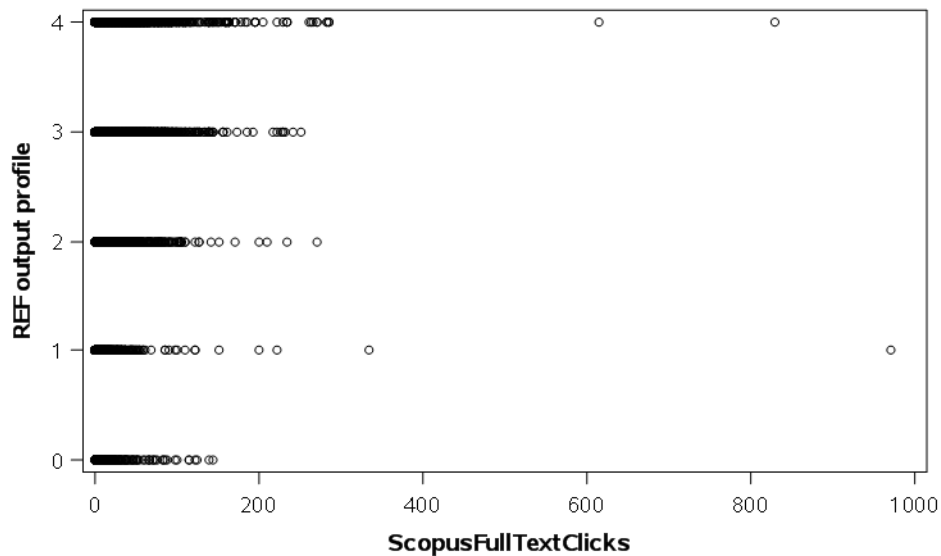


Table A48 Two-way summary of REF 4* indicator by Scopus full text requests prediction (publication year = 2008)

(prediction = 1 when Scopus full text requests ≥ 20)

REF 4* output	metric prediction		Total
	1	0	
1	1690	3350	5040
0	3320	12700	16020
Total	5005	16050	21060

Precision = true positives / total positive predictions = $1690 / (1690 + 3320) = 33.7\%$

Precision can be interpreted as the proportion of correct predictions.

Sensitivity = true positives / total REF 4* records = 1690 / (1690 + 3350) = 33.5%

Sensitivity can be interpreted as the proportion of 4* records identified by the prediction.

Table A49 Summary table of precision and sensitivity of REF 4* predictions and correlation between REF quality profile and Scopus full text requests by UOA (publication year = 2008)

UOA id	Name	Total	Precision	Sensitivity	Spearman correlation
.	Overall	21060	33.7%	33.5%	0.143
1	Clinical Medicine	2070	68.4%	30.1%	0.376
2	Public Health, Health Services and Primary Care	715	41.2%	42.1%	0.189
3	Allied Health Professions, Dentistry, Nursing and Pharmacy	1455	25.8%	28.4%	0.051
4	Psychology, Psychiatry and Neuroscience	1365	42.7%	46.4%	0.181
5	Biological Sciences	1400	52.7%	36.9%	0.213
6	Agriculture, Veterinary and Food Science	550	29.7%	50.0%	0.158
7	Earth Systems and Environmental Sciences	920	29.2%	54.6%	0.179
8	Chemistry	815	36.6%	69.6%	0.318
9	Physics	1225	46.5%	21.1%	0.129
10	Mathematical Sciences	830	39.3%	5.6%	0.072
11	Computer Science and Informatics	1035	53.2%	14.6%	0.133
12	Aeronautical, Mechanical, Chemical and Manufacturing Engineering	610	21.9%	43.0%	0.108
13	Electrical and Electronic Engineering, Metallurgy and Materials	640	36.4%	41.1%	0.169
14	Civil and Construction Engineering	195	25.0%	71.0%	0.091
15	General Engineering	1330	21.8%	40.4%	0.090
16	Architecture, Built Environment and Planning	300	23.3%	47.7%	0.130
17	Geography, Environmental Studies and Archaeology	550	22.7%	49.5%	0.140
18	Economics and Econometrics	275	54.5%	13.5%	0.121
19	Business and Management Studies	1400	28.0%	45.8%	0.182
20	Law	240	28.6%	3.8%	0.016
21	Politics and International Studies	325	12.9%	6.8%	0.102
22	Social Work and Social Policy	420	23.5%	25.0%	0.109

UOA					Spearman
id	Name	Total	Precision	Sensitivity	correlation
23	Sociology	250	15.4%	22.2%	0.058
24	Anthropology and Development Studies	150	34.6%	37.5%	0.218
25	Education	455	20.7%	18.5%	0.147
26	Sport and Exercise Sciences, Leisure and Tourism	255	15.4%	24.5%	0.107
27	Area Studies	110	33.3%	4.0%	-0.001
28	Modern Languages and Linguistics	175	25.0%	5.9%	0.111
29	English Language and Literature	170	20.0%	2.6%	0.039
30	History	245	0.0%	0.0%	0.163
31	Classics	15	0.0%	0.0%	0.132
32	Philosophy	115	0.0%	0.0%	-0.259
33	Theology and Religious Studies	45	0.0%	0.0%	-0.123
34	Art and Design: History, Practice and Theory	130	23.5%	21.1%	-0.022
35	Music, Drama, Dance and Performing Arts	90	0.0%	0.0%	-0.067
36	Communication, Cultural and Media Studies, Library and Information Management	195	40.0%	28.0%	0.196

Table A50 Precision and sensitivity of predictions for a range of REF quality profile groupings (publication year = 2008)

The previous summaries considered the predictive power of Scopus full text requests for identifying REF 4* work. The following table widens the population to 3* and 4* work and 2*, 3* and 4* work. This provides information on the cut-off used for the predictions and summary statistics, as defined in Table A48.

REF grouping	Prediction cut-off value	REF group indicator				Precision	Sensitivity
		1	0	1	0		
4*	20.00	1690	3350	3320	12700	33.7%	33.5%
4* + 3*	4.00	11615	4310	3380	1755	77.5%	72.9%
4* + 3* + 2*	1.00	17580	2695	630	150	96.5%	86.7%

Table A51 Summary of prediction statistics for all publication years

In controlling for the effect of publication year this summary has effectively ignored 85% of the data available. The following summary indicates what the prediction statistics would look like for each year of publication. Statistical modelling of this data allows all the data available to be used in the analysis whilst controlling for year of publication effects (see ‘Publication year effect’ in report and Annex B).

Year of publication	Total	Precision	Sensitivity	Spearman correlation	Kendall's Tau-b correlation
2008	21060	34.3%	30.2%	0.143	0.113
2009	23715	31.1%	29.4%	0.126	0.100
2010	25195	28.7%	27.9%	0.113	0.090
2011	26545	28.4%	28.2%	0.103	0.082
2012	27825	26.4%	25.6%	0.090	0.072
2013	24415	25.8%	26.2%	0.057	0.047
All years	148755	29.2%	28.8%	0.107	0.085

Note: Counts and totals of outputs have been rounded to the nearest five for all tables in the analysis.

Summary of Tweets as REF quality profile predictor

This analysis summarises the correlation and predictive relationship between Tweets and the REF quality profile scores. The effect of publication year has been controlled for by focussing on the earliest year available (2008). Table A56 provides a summary of the impact for this choice of year.

Please note that the Twitter data for 2008 articles were likely collected after they were published (from mid-2011 onwards).

Table A52 Summary of Tweets by REF quality profile (publication year = 2008)

This summarises the metric values at each level of REF quality profile. Outputs with a 0 quality profile were unclassified and either fell below the standard of nationally recognised work or did not meet the published definition of research. Outputs with missing Tweets values were excluded from the analysis (missing = 5).

REF quality profile	Frequency	Mean	Minimum	Median	Maximum
4	5040	0.4	0.0	0.0	225.0
3	10885	0.1	0.0	0.0	45.0
2	4350	0.2	0.0	0.0	53.0
1	525	0.2	0.0	0.0	18.0
0	260	0.5	0.0	0.0	27.0
Total	21055	0.2	0.0	0.0	225.0

Note: While the missing statistics indicate that there was good coverage for this metric, over 90 per cent of the data returned was recorded as 0 which indicates a highly skewed distribution.

Figure A12 Scatterplot of Tweets against REF quality profile (publication year = 2008)

(Spearman correlation= 0.0645)

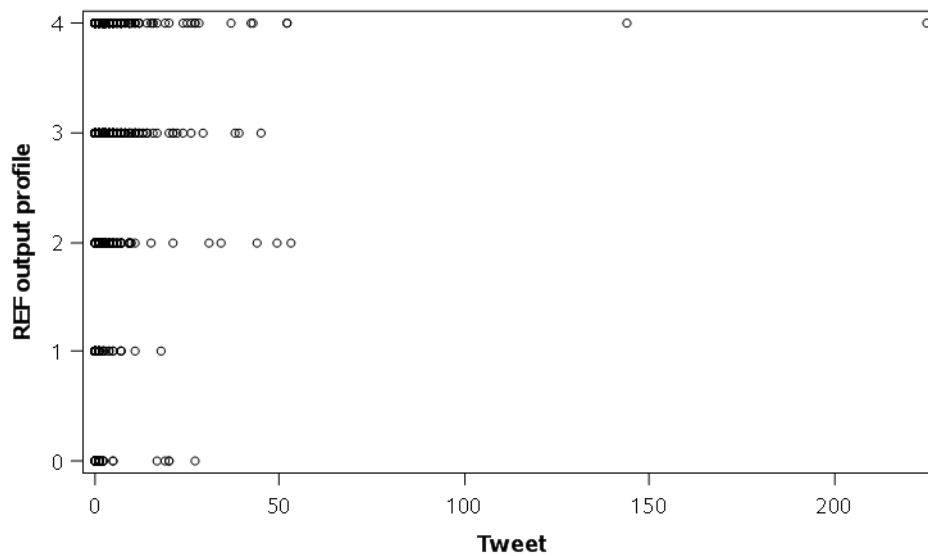


Table A53 Two-way summary of REF 4* indicator by Tweets prediction (publication year = 2008)

(prediction = 1 when Tweets \geq 1)

REF 4* output	metric prediction		Total
	1	0	
1	465	4575	5040
0	755	15260	16015
Total	1220	19835	21055

Precision = true positives / total positive predictions = $465 / (465 + 755) = 38.1\%$

Precision can be interpreted as the proportion of correct predictions.

Sensitivity = true positives / total REF 4* records = $465 / (465 + 4575) = 9.2\%$

Sensitivity can be interpreted as the proportion of 4* records identified by the prediction.

Table A54 Summary table of precision and sensitivity of REF 4* predictions and correlation between REF quality profile and Tweets by UOA (publication year = 2008)

UOA id	Name	Total	Precision	Sensitivity	Spearman correlation
.	Overall	21055	38.1%	9.2%	0.065
1	Clinical Medicine	2070	48.7%	13.6%	0.133
2	Public Health, Health Services and Primary Care	715	43.0%	32.6%	0.167
3	Allied Health Professions, Dentistry, Nursing and Pharmacy	1450	33.0%	9.1%	0.039
4	Psychology, Psychiatry and Neuroscience	1365	42.6%	16.7%	0.076
5	Biological Sciences	1400	61.1%	10.9%	0.118
6	Agriculture, Veterinary and Food Science	550	31.4%	9.5%	0.083
7	Earth Systems and Environmental Sciences	920	58.3%	17.9%	0.210
8	Chemistry	815	46.2%	3.1%	0.056
9	Physics	1225	38.5%	3.1%	0.049
10	Mathematical Sciences	830	25.0%	1.5%	0.002
11	Computer Science and Informatics	1035	31.9%	5.2%	0.012
12	Aeronautical, Mechanical, Chemical and Manufacturing Engineering	610	9.1%	0.8%	-0.054
13	Electrical and Electronic Engineering, Metallurgy and Materials	640	66.7%	2.5%	0.089

UOA id	Name	Total	Precision	Sensitivity	Spearman correlation
14	Civil and Construction Engineering	195	66.7%	6.5%	0.146
15	General Engineering	1330	30.8%	1.7%	0.035
16	Architecture, Built Environment and Planning	300	33.3%	4.6%	0.059
17	Geography, Environmental Studies and Archaeology	550	30.3%	9.2%	0.080
18	Economics and Econometrics	275	66.7%	9.0%	0.148
19	Business and Management Studies	1400	19.6%	3.7%	0.030
20	Law	240	6.7%	1.9%	0.001
21	Politics and International Studies	325	14.3%	5.1%	0.013
22	Social Work and Social Policy	420	18.2%	5.0%	-0.012
23	Sociology	250	27.0%	27.8%	0.133
24	Anthropology and Development Studies	150	14.3%	8.3%	0.070
25	Education	455	20.0%	8.8%	0.054
26	Sport and Exercise Sciences, Leisure and Tourism	255	16.1%	10.2%	-0.003
27	Area Studies	110	12.5%	4.0%	-0.031
28	Modern Languages and Linguistics	175	25.0%	2.9%	0.033
29	English Language and Literature	170	22.2%	5.1%	0.036
30	History	245	35.7%	10.9%	0.083
31	Classics	15	0.0%	0.0%	.
32	Philosophy	115	40.0%	7.4%	0.035
33	Theology and Religious Studies	45	0.0%	0.0%	0.001
34	Art and Design: History, Practice and Theory	130	75.0%	15.8%	0.234
35	Music, Drama, Dance and Performing Arts	90	0.0%	0.0%	-0.073
36	Communication, Cultural and Media Studies, Library and Information Management	195	37.0%	20.0%	0.044

Table A55 Precision and sensitivity of predictions for a range of REF quality profile groupings (publication year = 2008)

The previous summaries considered the predictive power of Tweets for identifying REF 4* work. The following table widens the population to 3* and 4* work and 2*, 3* and 4* work. This provides information on the cut-off used for the predictions and summary statistics, as defined in Table A53.

REF grouping	Prediction cut-off value	REF group indicator				Precision	Sensitivity
		1	0	1	0		
		metric prediction					
4*	1.00	465	4575	755	15260	38.1%	9.2%
4* + 3*	1.00	975	14945	245	4890	80.0%	6.1%
4* + 3* + 2*	1.00	1170	19105	50	730	95.8%	5.8%

Table A56 Summary of prediction statistics for all publication years

In controlling for the effect of publication year this summary has effectively ignored 85% of the data available. The following summary indicates what the prediction statistics would look like for each year of publication. Statistical modelling of this data allows all the data available to be used in the analysis whilst controlling for year of publication effects (see ‘Publication year effect’ in report and Annex B).

Year of publication	Total	Precision	Sensitivity	Spearman correlation	Kendall's Tau-b correlation
2008	21055	38.1%	9.2%	0.065	0.060
2009	23700	37.4%	11.3%	0.069	0.064
2010	25195	37.1%	15.8%	0.088	0.082
2011	26530	32.0%	33.1%	0.123	0.112
2012	27815	35.1%	34.7%	0.146	0.128
2013	24395	35.5%	31.9%	0.153	0.131
All years	148690	32.4%	21.6%	0.090	0.081

Note: Counts and totals of outputs have been rounded to the nearest five for all tables in the analysis.

Summary of Google Scholar citations as REF quality profile predictor

This analysis summarises the correlation and predictive relationship between Google Scholar citations and the REF quality profile scores. The effect of publication year has been controlled for by focussing on the earliest year available (2008). Table A61 provides a summary of the impact for this choice of year.

Table A57 Summary of Google Scholar citations by REF quality profile (publication year = 2008)

This summarises the metric values at each level of REF quality profile. Outputs with a 0 quality profile were unclassified and either fell below the standard of nationally recognised work or did not meet the published definition of research. Outputs with missing Google Scholar citations values were excluded from the analysis (missing = 5).

REF quality profile	Frequency	Mean	Minimum	Median	Maximum
4	5040	124.6	0.0	72.0	4,146.0
3	10885	50.9	0.0	31.0	4,516.0
2	4350	32.0	0.0	17.0	8,291.0
1	525	29.2	0.0	13.0	661.0
0	260	89.3	0.0	29.0	4,516.0
Total	21055	64.6	0.0	32.0	8,291.0

Figure A13 Scatterplot of Google Scholar citations against REF quality profile (publication year = 2008)

(Spearman correlation= 0.3575)

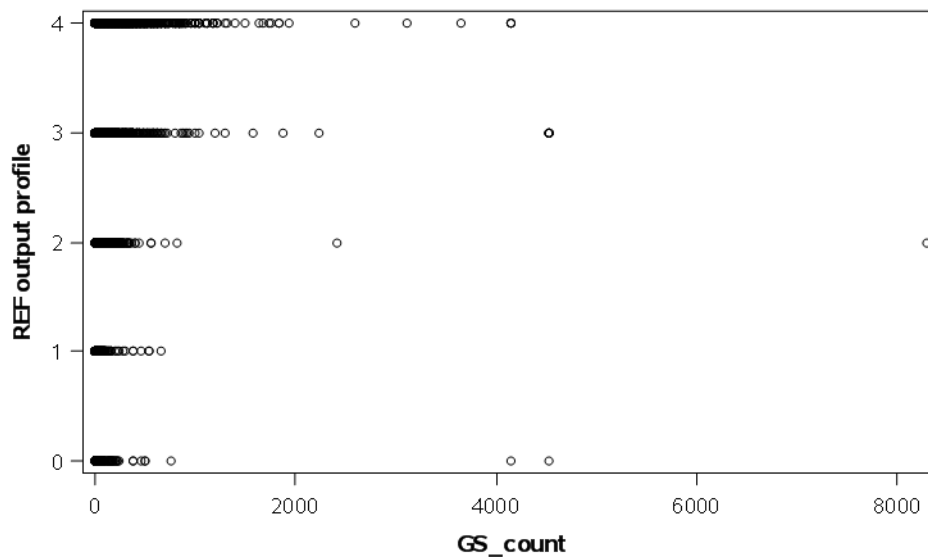


Table A58 Two-way summary of REF 4* indicator by Google Scholar citations prediction (publication year = 2008)

(prediction = 1 when Google Scholar citations \geq 73)

REF 4* output	metric prediction		Total
	1	0	
1	2495	2545	5040
0	2470	13545	16015
Total	4965	16090	21055

Precision = true positives / total positive predictions = $2495 / (2495 + 2470) = 50.3\%$

Precision can be interpreted as the proportion of correct predictions.

Sensitivity = true positives / total REF 4* records = $2495 / (2495 + 2545) = 49.5\%$

Sensitivity can be interpreted as the proportion of 4* records identified by the prediction.

Table A59 Summary table of precision and sensitivity of REF 4* predictions and correlation between REF quality profile and Google Scholar citations by UOA (publication year = 2008)

UOA id	Name	Total	Precision	Sensitivity	Spearman correlation
.	Overall	21055	50.3%	49.5%	0.357
1	Clinical Medicine	2070	51.1%	84.8%	0.600
2	Public Health, Health Services and Primary Care	715	37.4%	61.8%	0.326
3	Allied Health Professions, Dentistry, Nursing and Pharmacy	1450	54.5%	45.8%	0.403
4	Psychology, Psychiatry and Neuroscience	1365	46.7%	57.5%	0.354
5	Biological Sciences	1400	68.5%	68.7%	0.545
6	Agriculture, Veterinary and Food Science	550	48.6%	44.0%	0.367
7	Earth Systems and Environmental Sciences	920	46.4%	59.7%	0.375
8	Chemistry	815	58.8%	68.6%	0.557
9	Physics	1225	52.0%	83.0%	0.564
10	Mathematical Sciences	830	56.4%	29.1%	0.350
11	Computer Science and Informatics	1035	64.3%	47.0%	0.484
12	Aeronautical, Mechanical, Chemical and Manufacturing Engineering	610	38.6%	18.2%	0.207

UOA id	Name	Total	Precision	Sensitivity	Spearman correlation
13	Electrical and Electronic Engineering, Metallurgy and Materials	640	63.6%	38.7%	0.477
14	Civil and Construction Engineering	195	42.9%	19.4%	0.207
15	General Engineering	1330	35.2%	24.8%	0.223
16	Architecture, Built Environment and Planning	300	37.5%	13.8%	0.151
17	Geography, Environmental Studies and Archaeology	550	40.2%	37.6%	0.267
18	Economics and Econometrics	275	60.0%	47.2%	0.378
19	Business and Management Studies	1400	39.4%	42.8%	0.305
20	Law	240	28.6%	7.5%	0.159
21	Politics and International Studies	325	35.7%	16.9%	0.145
22	Social Work and Social Policy	420	36.7%	13.8%	0.238
23	Sociology	250	12.9%	11.1%	0.122
24	Anthropology and Development Studies	150	12.5%	8.3%	0.258
25	Education	455	27.7%	14.3%	0.193
26	Sport and Exercise Sciences, Leisure and Tourism	255	37.9%	22.4%	0.219
27	Area Studies	110	16.7%	4.0%	0.086
28	Modern Languages and Linguistics	175	28.6%	5.9%	0.195
29	English Language and Literature	170	40.0%	5.1%	-0.011
30	History	245	0.0%	0.0%	0.112
31	Classics	15	0.0%	0.0%	0.031
32	Philosophy	115	60.0%	11.1%	0.057
33	Theology and Religious Studies	45	0.0%	0.0%	-0.163
34	Art and Design: History, Practice and Theory	130	33.3%	5.3%	0.075
35	Music, Drama, Dance and Performing Arts	90	0.0%	0.0%	-0.072
36	Communication, Cultural and Media Studies, Library and Information Management	195	46.7%	14.0%	0.293

Table A60 Precision and sensitivity of predictions for a range of REF quality profile groupings (publication year = 2008)

The previous summaries considered the predictive power of Google Scholar citations for identifying REF 4* work. The following table widens the population to 3* and 4* work and

2*, 3* and 4* work. This provides information on the cut-off used for the predictions and summary statistics, as defined in Table A58.

REF grouping	Prediction cut-off value	REF group indicator				Precision	Sensitivity
		1		0			
		metric	prediction				
4*	73.00	2495	2545	2470	13545	50.3%	49.5%
4* + 3*	13.00	12745	3175	3145	1990	80.2%	80.1%
4* + 3* + 2*	1.00	19450	825	735	45	96.4%	95.9%

Table A61 Summary of prediction statistics for all publication years

In controlling for the effect of publication year this summary has effectively ignored 85% of the data available. The following summary indicates what the prediction statistics would look like for each year of publication. Statistical modelling of this data allows all the data available to be used in the analysis whilst controlling for year of publication effects (see ‘Publication year effect’ in report and Annex B).

Year of publication	Total	Precision	Sensitivity	Spearman correlation	Kendall's Tau-b correlation
2008	21055	52.5%	46.9%	0.357	0.283
2009	23700	48.5%	45.9%	0.341	0.271
2010	25195	45.8%	45.0%	0.334	0.266
2011	26530	42.1%	43.1%	0.293	0.233
2012	27815	37.7%	39.8%	0.249	0.199
2013	24395	38.1%	38.6%	0.247	0.200
All years	148690	41.8%	41.5%	0.273	0.217

Note: Counts and totals of outputs have been rounded to the nearest five for all tables in the analysis.

Annex B: Modelling results

In order to account simultaneously for the various factors identified as possible predictors of a high Research Excellence Framework (REF) outcome, the data was modelled using regression techniques.

Initial summaries of the data showed that key model factors were: the 15 metrics; year of publication; units of assessment (or main panel groups). Running this many variables as a ‘full’ model generated a number of statistically non-significant model estimates as a number of the metrics were correlated with each other. Using a combination of knowledge of the types of metric used, the Spearman correlation coefficients, the full model output and stepwise model fitting of the full model across the 36 units of assessment (UOAs), the list of model factors was reduced to a ‘restricted’ model.

Table B1 provides the correlation matrix for the 13 numeric metrics included in the model and the dependent variable (REF quality profile). Correlations over 0.5 are highlighted in dark grey and those between 0.25 and 0.5 are highlighted in light grey. This identifies some of the inter-relationships between the various metric measures.

Table B1 Spearman correlation coefficient matrix for numeric metrics and REF quality profile

	REF quality profile	Citation_count	FWCI	SNIP	SJR	Authors	AuthorCountries	WIPO_patent_citations	MendeleyRead	SciDir_Dwnld	ScopusFullTextClicks	Percentile	Tweet	GS_count
REF quality profile	1.00	0.25	0.28	0.33	0.34	0.15	0.14	0.07	0.17	-0.02	0.11	-0.29	0.09	0.27
Citation_count	0.25	1.00	0.77	0.40	0.54	0.52	0.40	0.15	0.10	0.24	0.52	-0.86	-0.09	0.79
FWCI	0.28	0.77	1.00	0.38	0.38	0.30	0.20	0.10	0.22	0.18	0.35	-0.88	0.10	0.69
SNIP	0.33	0.40	0.38	1.00	0.77	0.32	0.22	0.08	0.28	0.14	0.19	-0.46	0.17	0.40
SJR	0.34	0.54	0.38	0.77	1.00	0.52	0.34	0.12	0.31	0.08	0.15	-0.60	0.17	0.49
Authors	0.15	0.52	0.30	0.32	0.52	1.00	0.62	0.11	0.23	0.05	0.38	-0.47	0.12	0.42
AuthorCountries	0.14	0.40	0.20	0.22	0.34	0.62	1.00	0.06	0.13	0.00	0.30	-0.28	0.05	0.32
WIPO_patent_citations	0.07	0.15	0.10	0.08	0.12	0.11	0.06	1.00	0.05	-0.01	0.09	-0.13	-0.02	0.12
MendeleyRead	0.17	0.10	0.22	0.28	0.31	0.23	0.13	0.05	1.00	-0.03	0.14	-0.26	0.80	0.18
SciDir_Dwnld	-0.02	0.24	0.18	0.14	0.08	0.05	0.00	-0.01	-0.03	1.00	0.36	-0.21	-0.05	0.23
ScopusFullTextClicks	0.11	0.52	0.35	0.19	0.15	0.38	0.30	0.09	0.14	0.36	1.00	-0.38	0.03	0.44
Percentile	-0.29	-0.86	-0.88	-0.46	-0.60	-0.47	-0.28	-0.13	-0.26	-0.21	-0.38	1.00	-0.10	-0.75
Tweet	0.09	-0.09	0.10	0.17	0.17	0.12	0.05	-0.02	0.80	-0.05	0.03	-0.10	1.00	0.02
GS_count	0.27	0.79	0.69	0.40	0.49	0.42	0.32	0.12	0.18	0.23	0.44	-0.75	0.02	1.00

As discussed in the main report, metrics were not available for a number of outputs and to exclude outputs with missing data would eliminate around a quarter of the outputs (39,390). Instead, dummy variables were used in the model to ensure that the information from the non-missing fields was still used to calculate the model estimates.

A logistic regression model was fitted to the data. This required the outcome of the model to be restricted to a binary variable (in this case 4*/not 4*). This approach was taken after

consideration of more complex models (such as ordered logit and multinomial models) which would have more accurately reflected the REF quality profile distribution (ranging from 4* to unclassified). The more complex approach allowed for more detailed understanding of the relationships between the various REF scores, but did not significantly affect the overall conclusions drawn.

Table B2 outlines the list of full and restricted model factors; this shows the groupings used for the model factors and the selection made for the restricted model. The decision to include or exclude variables in the restricted model was based on an evaluation of a number of sources (as explained above), but this is a subjective choice. It is possible that the choices made could impact on the findings of the subsequent assessment of bias for female author and early career researcher (ECR) outputs. In order to minimise the risk of bias, the results included in the report were verified by running an alternative model using a second selection of variables.

The adopted logistic model assumed that the relationship between the metric distributions and the proportion achieving 4* was linear. However, initial exploration of the data showed that many of the metrics had non-linear relationships and skewed distributions. In order to avoid violation of the model assumptions, a transformation of the some of the metrics was carried out⁹. The model results identified with a # in Table B2 relate to the transformed metric values and should be interpreted as such.

Table B2 Comparison of factors in the full and restricted models

Full model factors	Restricted model factors
citation_count#	
ScopusFullTextClicks#	ScopusFullTextClicks#
Non-missing FWCI#	
Missing FWCI	
GS_count#	GS_count#

⁹ A number of transformations were considered and the final transformation adopted took the natural logarithm of the metric plus one. See Thelwall, M. & Wilson, P. (2014) *Regression for citation data: An evaluation of different methods* for more information (<http://www.scit.wlv.ac.uk/~cm1993/papers/RegressionForCitationDataPreprint.pdf>) [PDF].

Percentile (top 5%)

Percentile (top 10%)

Percentile (top 25%)

Percentile (top 50%)

Percentile (bottom 50%)

Percentile (unknown)

Non-missing SNIP#

Non-missing SNIP#

Missing SNIP

Missing SNIP

Non-missing SJR#

Non-missing SJR#

Missing SJR

Missing SJR

Authors#

Authors#

AuthorCountries

MendeleyRead#

Tweet#

Tweet#

Non-missing SciDir_Dwnld#

Non-missing SciDir_Dwnld#

Missing SciDir_Dwnld

Missing SciDir_Dwnld

WIPO_patent_citation

CrossAcademicCorporate

Collaboration (Institutional)

Collaboration (International)

Collaboration (National)

Collaboration (Other)

Year (2008)

Year (2008)

Year (2009)

Year (2009)

Year (2010)

Year (2010)

Year (2011)

Year (2011)

Year (2012)

Year (2012)

Female

ECR

Female*ECR

Note: a log transformation of the metric plus one was applied to those variables identified with #.

When modelling these variables a number were recoded to be dummy variables (that is a series of binary variables which together indicate the value held by a categorical variable). This included: Year, Collaboration, Cross Academic Corporate and Percentile. This means these variables were treated as categorical for the purposes of statistical modelling. The characteristic variables were also binary variables identifying the sex and ECR status of the submitting author.

The full model was run for model building purposes but is not presented here, as it is likely to have had inflated error estimates (as the assumption of independent variables was likely to have been violated). Instead a stepwise model was run for each UOA which selected the most significant metrics for each UOA in predicting a REF 4* outcome. Table B3 presents the odds ratios¹⁰ for the significant model factors of each UOA. These are discussed further in the main report.

¹⁰ Odds ratios are used to compare the relative odds of achieving a 4* REF output given a one unit increase in the variable of interest (typically a metric measure in this case). If the odds ratio is 1 then a one unit increase in the metric value does not significantly affect the odds of achieving a 4* output. However, if the odds ratio is greater than 1 then a one unit increase in the metric value is associated with higher odds of achieving a 4* output. Finally, if the odds ratio is less than 1, then a one unit increase is associated with lower odds of achieving 4*. Only those odds ratios that are significantly different to 1 are presented in this report (assessed at the 1% statistical significance level).

Table B3 Summary of significant odds ratios from stepwise regression of the full model by UOA

Model factor	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36		
Intercept	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.1	0.0	0.1	0.1	0.1	0.1	0.0	0.0	0.0	0.2	0.1	0.0	0.0	0.1	0.1	0.1	0.3	0.1	0.2	0.1	0.3	0.1	0.1	0.1	0.3	0.2		
Citation_count#	2.1		1.2	1.4	2.3	1.7	2.2	3.6	2.8		2.0		1.5					1.3							1.2	1.5		0.6										
Authors#	1.2	1.9	2.2	0.7		1.8				0.7	0.6			0.7				0.5																				
AuthorCountries		0.9	0.9		0.8		0.9		1.0						1.1																							
WIPO_patent_citation																																						
MendeleyRead#	1.1		1.1	1.1	1.1	1.1			1.2			1.2				1.1												1.3										
ScopusFullTextClicks#	1.1			1.1		1.2	0.9				0.7					0.8					0.8																	
Tweet#	1.3	1.3		1.1	1.1		1.4	1.8	1.3	0.5					1.3	1.3							1.2														1.2	
GS_count#			1.2		1.3			1.3		1.5	1.7	1.2	1.4		1.2	1.3	1.2	1.4	1.4	1.2	1.1	1.4		1.3	1.2		1.4			1.3		1.2					1.3	
Non-missing SciDir_Dwnld#		1.1		1.0	1.1		0.9	0.8		0.9	0.9	0.9	0.9		0.9			0.9																				
Missing SciDir_Dwnld	0.6	1.8		0.8	0.7		0.4	0.3			1.4	0.7	0.4		0.6	1.4					0.7																	
Non-missing SNIP#		3.7			0.3		0.2			3.6	1.7					2.1						4.7		2.7	0.3		3.8	3.0			3.0							
Missing SNIP	20.6	2.9	6.3		14.5			39.3	7.8	3.2	3.0	3.5		4.2	2.7	4.1	87.5					3.3		2.7														
Non-missing SJR#	10.3		4.6	6.9	10.3	6.7	13.5	11.6	2.8	1.9	2.6	2.0	2.7	5.1	2.1		15.8	6.1		2.9	2.5			9.0		2.3				3.0								
Missing SJR	9.3			20.8	30.2	9.2	13.1	12.1	3.4	2.9			5.7					3.7		3.8	2.1			2.7			2.6		2.4									
Non-missing FWCI#	1.7	2.2	1.8	1.3			1.8			2.3		1.2		1.4	1.3	1.6	1.4										1.8											
Missing FWCI											1.6																											
CrossAcademicCorporate											1.5				1.5																							
Collaboration (Institutional)																		0.7																				
Collaboration (International)	1.6						1.8				1.3																											

Collaboration (National)											1.3		
Collaboration (Other)													
Percentile (top 5%)						0.6				1.3		8.4	
Percentile (top 10%)						0.4							
Percentile (top 25%)	0.6		0.8			0.3	1.6	0.7		0.5			
Percentile (bottom 50%)						0.2	1.6	0.7		0.5			
Percentile (unknown)	4.6	2.5	1.9	1.5	5.2	3.8		2.7	2.0		1.6	0.6	
Percentile (unknown)	15.4	13.7	8.4	6.0	5.3	3.2	3.0					1.6	
Year (2008)								0.2			0.4	0.6	0.5
Year (2009)								0.3			0.4	0.7	0.5
Year (2010)								0.3			0.4	0.6	
Year (2011)	0.6	0.5				1.8		0.3			0.5		0.5
Year (2012)		0.4		1.4		1.9		0.4			0.5		

Note: Significance was defined at the 1% level for these results. A log transformation of the metric plus one was applied to those variables identified with #.
For more detailed model output please contact the HEFCE Analysis for Policy team.

The restricted model was considered alongside the additional characteristics of sex and ECR status of submitting author. This analysis identified whether there were still significant differences in REF quality scores between male and female and between ECR and non-ECR authors, after holding the metric scores of the output constant. The metrics are assumed to be a proxy for quality and this identifies any UOAs where biases could be occurring. Table B4 presents the initial differences by UOA; Table B5 presents the outcomes of the restricted model by unit of assessment and is discussed further in the main report.

Table B4 Summary of submitting authors by UOA and additional characteristics

Main panel	Unit of assessment	Outputs	Outputs with ECR author	% achieving 4*		Outputs with female author	% achieving 4*	
				ECR	Non-ECR		Female	Male
Total UOAs		149,670	13,580	23%	21%	40,660	19%	22%
A	Clinical Medicine	13,275	995	28%	23%	3,525	20%	24%
	Public Health, Health Services and Primary Care	4,805	450	19%	23%	2,060	19%	25%
	Allied Health Professions, Dentistry, Nursing and Pharmacy	10,025	800	27%	21%	4,320	20%	23%
	Psychology, Psychiatry and Neuroscience	9,020	910	31%	26%	3,330	21%	29%
	Biological Sciences	8,550	845	39%	28%	1,870	26%	30%
	Agriculture, Veterinary and Food Science	3,840	265	19%	18%	1,080	15%	20%
	B	Earth Systems and Environmental Sciences	5,170	425	19%	18%	960	17%
Chemistry		4,685	310	32%	21%	665	22%	22%
Physics		6,380	495	30%	20%	790	17%	22%
Mathematical Sciences		6,540	580	26%	21%	825	17%	23%
Computer Science and Informatics		6,995	645	25%	23%	1,045	18%	24%
Aeronautical, Mechanical, Chemical and Manufacturing Engineering		4,050	340	18%	18%	580	14%	19%
Electrical and Electronic Engineering, Metallurgy and Materials		3,965	245	18%	20%	420	21%	20%

	Civil and Construction Engineering	1,315	150	17%	18%	220	14%	19%
	General Engineering	8,470	875	22%	17%	1,010	16%	18%
C	Architecture, Built Environment and Planning	2,700	235	24%	22%	650	18%	23%
	Geography, Environmental Studies and Archaeology	4,800	505	18%	21%	1,375	18%	21%
	Economics and Econometrics	2,370	225	18%	29%	395	18%	30%
	Business and Management Studies	11,385	1,080	18%	21%	3,085	19%	22%
	Law	2,275	240	14%	21%	890	20%	21%
	Politics and International Studies	2,965	375	20%	19%	780	16%	20%
	Social Work and Social Policy	3,485	310	14%	18%	1,780	19%	17%
	Sociology	1,925	170	14%	16%	870	14%	18%
	Anthropology and Development Studies	1,265	135	16%	16%	495	14%	17%
	Education	4,040	265	19%	22%	1,885	20%	24%
	Sport and Exercise Sciences, Leisure and Tourism	2,430	285	15%	20%	690	16%	21%
	D	Area Studies	855	75	18%	21%	285	26%
Modern Languages and Linguistics		1,790	165	23%	21%	840	19%	23%
English Language and Literature		1,940	235	17%	18%	900	16%	20%
History		2,410	275	20%	21%	750	19%	21%
Classics		240	30	18%	17%	90	16%	18%
Philosophy		1,300	170	22%	22%	290	20%	23%
Theology and Religious Studies		450	30	26%	15%	125	13%	17%
Art and Design: History, Practice and Theory		1,305	130	13%	15%	560	16%	14%
Music, Drama, Dance and Performing Arts		1,010	165	19%	23%	470	21%	23%
Communication, Cultural and Media Studies, Library and Information Management		1,640	145	13%	21%	755	20%	21%

Additional interaction terms were considered for this model, which tested whether the interaction between metrics scores and an author's sex or ECR status significantly added to the model's ability to predict 4* REF output. These terms were not consistently significant and this suggests that there was no evidence of any metrics being better predictors of 4* outputs for female authors or ECRs.

Table B5 Summary of significant odds ratios from restricted model with REF 4* outputs by unit of assessment

Model factor	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36		
Intercept	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.1	0.0	0.1	0.1	0.2	0.1	0.0	0.0	0.2	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.3				
ScopusFullTextClicks#	1.6			1.3	1.1	1.2					0.8		1.2			0.8																						
GS_count#	1.5	1.1	1.6	1.3	1.8	1.5	1.8	2.7	1.9	1.9	2.3	1.3	2.2	1.5	1.4	1.4	1.4	1.6	1.4	1.4	1.2	1.5		1.3	1.2	1.4		1.5		1.3		1.3		1.4	1.4	1.4		
Non-missing SciDir_Dwnld#		1.2		1.0	1.1		0.9	0.8		0.9		0.9	0.9	0.9		0.9		0.9							0.9													
Missing SciDir_dwnld		2.1			0.8		0.5	0.3					0.5	0.6							0.7				0.7													
Non-missing SJR#	9.5		4.9	6.9	11.5	8.9	15.4	8.9	1.9		3.2	2.1	3.0	4.7	2.0	1.8	17.0	6.4		2.8							12.5											
Missing SJR	7.3			18.9	4.3				3.3									3.0		2.4								4.6										
Non-missing SNIP#	3.2	8.2				0.4	0.2		2.3	4.9	1.6					2.1							4.0		2.8	0.2		5.2	4.2					3.8				
Missing SNIP	58.4	28.8	4.4		25.2		5.1	61.6	2.3			4.5				5.3	57.3								2.7													
Tweet#	1.5	1.5	1.2	1.2	1.1		1.4	1.5	1.6	0.5	0.8		1.5	1.2		1.3																						
Authors#	1.2	1.5	1.6	0.8		1.6	1.3			0.8													1.3															
Year (2008)		3.0						0.2		0.3	0.2			0.2	0.4	0.4	0.4		0.4			0.5			0.4	0.4												
Year (2009)		3.1						0.2		0.4	0.3			0.3	0.5	0.4	0.4	0.4	0.4		0.5	0.5			0.5	0.4										0.4		
Year (2010)		3.0						0.1		0.4	0.2	0.5		0.3	0.6	0.4	0.5	0.4	0.4			0.5			0.5													
Year (2011)				0.6	0.4			0.2		0.5	0.2				0.5	0.3	0.5		0.5			0.6															0.5	
Year (2012)				0.5	0.4			0.1		0.5	0.3				0.6	0.5		0.4	0.6			0.6																
Female				0.8							0.7					0.7		0.6																				
ECR																		0.5																				
Female*ECR																							0.3															

Note: Significance was defined at the 1% level for these results. A log transformation of the metric plus one was applied to those variables identified with #. The models for UOAs 31 and 33 did not converge so no results are shown here. For more detailed model output please contact the HEFCE Analysis for Policy team.

Annex C: Definitions and abbreviations

The following table provides the summary of unit of assessment codes and names.

Table C1 Summary of REF main panel and unit of assessment codes

Main panel	Unit of assessment	
A	1	Clinical Medicine
	2	Public Health, Health Services and Primary Care
	3	Allied Health Professions, Dentistry, Nursing and Pharmacy
	4	Psychology, Psychiatry and Neuroscience
	5	Biological Sciences
	6	Agriculture, Veterinary and Food Science
B	7	Earth Systems and Environmental Sciences
	8	Chemistry
	9	Physics
	10	Mathematical Sciences
	11	Computer Science and Informatics
	12	Aeronautical, Mechanical, Chemical and Manufacturing Engineering
	13	Electrical and Electronic Engineering, Metallurgy and Materials
	14	Civil and Construction Engineering
	15	General Engineering
C	16	Architecture, Built Environment and Planning
	17	Geography, Environmental Studies and Archaeology
	18	Economics and Econometrics
	19	Business and Management Studies
	20	Law
	21	Politics and International Studies
	22	Social Work and Social Policy
	23	Sociology
	24	Anthropology and Development Studies
	25	Education
	26	Sport and Exercise Sciences, Leisure and Tourism
D	27	Area Studies
	28	Modern Languages and Linguistics
	29	English Language and Literature

30	History
31	Classics
32	Philosophy
33	Theology and Religious Studies
34	Art and Design: History, Practice and Theory
35	Music, Drama, Dance and Performing Arts
36	Communication, Cultural and Media Studies, Library and Information Management

The following list provides a summary of the abbreviations used in the paper and their meaning in this paper.

DOI	Digital object identifier
ECR	Early career researcher
FWCI	Field weighted citation impact
GS	Google Scholar
HEI	Higher education institution
REF 2014	Research Excellence Framework 2014
SNIP	Source normalised impact per paper
SJR	SCImago journal rank
UOA	Unit of assessment
WIPO	World Intellectual Property Organisation