# The nature, scale and beneficiaries of research impact

An initial analysis of Research Excellence Framework (REF) 2014 impact case studies

King's College London and Digital Science

**March 2015**

DIGITAL science

KING'S College LONDON

Digital Humanities

**The Policy Institute at King's**

The word cloud on the front cover visualizes the 'Details of the impact' (Section 4) of the 2014 REF impact case studies. The clouds give greater prominence to words that appear more frequently in the source text.

In total 6,247,292 words were harvested from Section 4 of the 6,679 non-redacted case studies. After 'stop' words (eg 'and' 'the', 'but', etc.) were removed, this gave 3,730,769 words and 108,717 unique words (or 'tokens'). Removing low frequency items (10 or fewer occurrences) resulted in 19,266 unique words, with the illustrated word cloud showing the most frequent 684 words. We used a word cloud generator (www.jasondavies.com/wordcloud/) for this visualization.

# Preface

This report is based on an analysis of the 6,679 non-redacted impact case studies that were submitted to the 2014 Research Excellence Framework (REF). The Higher Education Funding Council for England (HEFCE) and its partners (Scottish Funding Council, Higher Education Funding Council for Wales, Department for Employment and Learning Northern Ireland, Research Councils UK and the Wellcome Trust), commissioned a synthetic analysis of the impact case studies to provide evidence on the impact of research in UK Higher Education Institutions (HEIs). The case studies outline changes and benefits to the economy, society, culture, public policy and services, health, the environment and quality of life. Using a mix of text-mining approaches and qualitative analysis, the nature, scale and beneficiaries of the non-academic impact of research is described. The report will be of interest to anyone interested in the assessment of research impact, REF and the impact of universities on society.

  This report is part of a wider project led by Digital Science (a division of Macmillan Science and Education) working in conjunction with its sister organization, the Nature Publishing Group, the Policy Institute and Department of Digital Humanities at King's College London. The synthetic analysis was led by the team from King's College London. Digital Science is an innovative technology company developing software and apps that change the way science is done. The Policy Institute at King's links insightful research with rapid, relevant policy analysis to stimulate edgy, forward-looking debate and shape future policy agendas. The Department of Digital Humanities at King's is an international leader in the application of technology in the arts and humanities and social sciences.

For more information about this report, please contact:

**Professor Jonathan Grant**
**Director, Policy Institute at King's**
**& Professor of Public Policy**
**King's College London**
**VB1.17, 1st Floor**
**Virginia Woolf Building**
**22 Kingsway**
**London**
**WC2B 6LE**

**Tel:** + 44 (0) 20 7848 1742
**Email:** jonathan.grant@kcl.ac.uk

# Contents

# Chapter 4: Six 'deep mines' on impact     48

# Chapter 5: Concluding reflections     70

# Figures, Tables and Boxes

## Figures

## Tables

## Boxes

# Headline findings

This report provides an initial assessment of the nature, scale and beneficiaries of the impact of UK universities' research. It is based on an analysis of the 6,679 non-redacted impact case studies that were submitted to the 2014 Research Excellence Framework (REF). Each case study aimed to showcase how research undertaken in UK Higher Education Institutions (HEIs) over the past 20 years has benefited society beyond academia – whether in the UK or globally. The case studies make an inspiring read that outline changes and benefits to the economy, society, culture, public policy and services, health, the environment and quality of life that have arisen from research.

A mix of text-mining techniques and qualitative analysis was used to synthesize the corpus of case studies. Text-mining itself can be dangerous and dirty: dangerous, as it is possible to misinterpret information in the text; and dirty, as it involves a lot of experimentation and trying by doing. Given the size of the dataset (more than 6 million words in the 'Details of the impact' section of the case studies), text-mining was useful in producing an analysis of the general patterns and themes that could be described in the case studies.

In order to probe the data further, we supplemented the text-mining method with six focused 'deep mine' questions of more than 1,000 case studies, which were read and from which qualitative inferences were made. We consider the analysis presented in this report as 'Version 1.0', and we emphasize the caveats and limitations to our analysis that are set out in Chapter 1.

In order to help digest the large amount of data generated, a series of data visualizations are presented in the main body of the report. The sections below summarize the headline findings from the analysis, and cross-reference to the key visualizations.

## The societal impact of research from UK Higher Education Institutions is considerable, diverse and fascinating

One of the most striking observations from the analysis of the REF case studies was the diverse range of contributions that UK HEIs have made to society. This is illustrated in a heat map of 60 impact topics by the 36 Units of Assessment (UOAs) (Figure 8; page 33), and the six deep mine analyses in Chapter 4, demonstrating that such diverse impacts occur from a diverse range of study disciplines.

## The research underpinning societal impacts is multidisciplinary, and the social benefit arising from research is multi-impactful

The relationship between 149 fields of research, 36 UOAs and 60 impact topics is visualized in the alluvial diagram in Figure 12 (page 39). What is evident from this visualization is that multiple fields of research underpin the case studies, leading to multiple types of impact. Overall we identified 3,709 unique pathways to impact.

## Different types of Higher Education Institutions specialize in different types of impact

Analysis of the impact topics by the size of HEI submission to REF 2014 illustrates that within this sample of case studies, small institutions are more likely to make

a disproportionate contribution to an impact topic. As shown in Figure 11 (page 37), although some larger institutions make large contributions to topics such as 'Clinical guidance' and 'Dentistry', small institutions make a greater than anticipated contribution to topics such as 'Sports', 'Regional innovation and enterprise' and 'Arts and culture'.

## UK Higher Education Institutions have a global impact

The map in Figure 13 (page 41) illustrates the geographical reach of the impacts of research undertaken over the past 20 years in UK HEIs. UK academics have made contributions to the wealth and well-being of all nations globally. The deep mine on the BRIC countries (Brazil, Russia, India and China; page 66) illustrates the richness of impact in these economies.

## The quantitative evidence supporting claims for impact was diverse and inconsistent, suggesting that the development of robust impact metrics is unlikely

While there was a large amount of numerical data included in the case studies, the way in which different numerical values were presented meant that the data could not be easily synthesized. Numerical values often differed in the units used, in the time periods used and in the assumptions underlying benefit calculations. For impact metrics to be developed such information would need to be measured, estimated and expressed in consistent ways using standard units and valuation approaches. However, there is a trade-off: one strength of the impact case studies is that they allow authors to select the appropriate and highly specific data to evidence the specific types of impact that they are claiming.

## The impact case studies provide a rich resource for analysis, but the information is collected for assessment purposes and may need to be aligned for analysis purposes

REF is an **assessment** system, and our interests in this report are on the effective **analysis** of the case studies: at times these two purposes may play against each other. Future analyses could be improved by collecting more structured and standardized information (eg drop-down boxes to indicate funding sources, standardized currencies, methods and metrics to show financial impact). There are trade-offs to this suggestion to consider – such as the potential increased burden on case study authors, and the risk that diverse and heterogeneous impacts might be discouraged by a more prescriptive framework.

# Authorship

The synthetic analysis underpinning this report was led by Jonathan Grant, Saba Hinrichs and Alastair Gill, supported by colleagues from King's College London and Digital Science.

The report was commissioned by the Higher Education Funding Council for England (HEFCE) and its partners (Scottish Funding Council, Higher Education Funding Council for Wales, Department for Employment and Learning Northern Ireland, Research Councils UK and the Wellcome Trust), and is part of a wider project led by Digital Science (a division of Macmillan Science and Education) working in conjunction with its sister organization, the Nature Publishing Group, and the Policy Institute and Department of Digital Humanities at King's College London.

In addition to this report, Digital Science developed and produced an online searchable database (http://impact.ref.ac.uk/casestudies) of the impact case studies submitted to REF 2014, that is freely available in a form and format that will enable any researcher to carry out future analysis.

The team at Digital Science was led by Jonathan Adams and included Tamar Loach and Martin Szomszor. The team at King's College London was led by Jonathan Grant and included Saba Hinrichs, Alastair Gill, Tobias Blanke, Mark Hedges and Simon Tanner. Daniel Hook at Digital Science was project director, and Joanna Wood was project manager.

Given that a large team of analysts was involved in developing the database and the synthetic analysis we suggest that this report be cited as follows:

> King's College London and Digital Science (2015). *The nature, scale and beneficiaries of research impact: An initial analysis of Research Excellence Framework (REF) 2014 impact case studies.* Bristol, United Kingdom: HEFCE.

# Acknowledgements

# Glossary

ABRC     Advisory Board for the Research Councils
ACARD   Advisory Committee for Applied Research and Development
BRIC     Brazil, Russia, India and China
DSIR     Department for Scientific and Industrial Research
EU        European Union
FOR      Field of research
FTE       Full-time equivalent
HEFCE   Higher Education Funding Council for England
HEI       Higher Education Institution
KWIC    Keyword-in-context
LDA      Latent Dirichlet Allocation
NICE     National Institute for Health and Care Excellence
QALY    Quality Adjusted Life Year
RAE      Research Assessment Exercise
REF      Research Excellence Framework
R&D      Research and development
RfP       Request for proposal
RQF      Research Quality Framework
TDM     Text and data-mining
UGC     University Grants Committee
UOA     Unit of Assessment

# 1 | Introduction

# 1 | Introduction

The purpose of this report is to provide an initial assessment of the nature, scale and beneficiaries of the impact of UK universities' research. It is based on a synthetic analysis of 6,679 non-redacted[1] impact case studies that were submitted to the 2014 Research Excellence Framework (REF). Each case study aims to showcase how research undertaken in UK universities over the past 20 years has benefited society beyond academia – whether in the UK or globally. The case studies outline changes and benefits to the economy, society, culture, public policy and services, health, the environment and quality of life. Specific examples include the following:

- Research showing the importance of same-day diagnostic tests for tuberculosis led to improvements in access to care and reductions in costs incurred by patients in Malawi, Nigeria, Yemen, Ethiopia, Nepal and elsewhere.

- The development of a super-repellent surface, created by plasmachemical techniques and invented by UK researchers, is used in millions of products worldwide, including mobile phones and hearing aids.

- Paralympic athletes' performance was improved by investigating wheelchair propulsion and optimizing configurations for competitive sport.

- Editorial and biographic analysis of the work of Virginia Woolf directly fed into the composition of *Vanessa and Virginia* (2008), a novel by Susan Sellars about Woolf's relationship with her sister, Vanessa Bell.

The remainder of this chapter provides a brief overview of the REF, describes how we used text-mining approaches to carry out a synthetic analysis of the case studies, and stresses the inevitable limitations and caveats associated with this approach. Chapter 2 provides an overview of the submitted cases studies by institution, Unit of Assessment (UOA) and other meta-characteristics associated with the case studies. Chapters 3 and 4 present an analysis of the nature, scale and beneficiaries of impact. Finally, Chapter 5 sets out some concluding thoughts and recommendations for future iterations of the REF that arose during the project.

## An overview of the Research Excellence Framework and impact case studies

The REF is a new system for assessing the quality of research in UK Higher Education Institutions (HEIs).[2] It replaced the Research Assessment Exercise (RAE), which has occurred on a (near-)quinquennial basis since 1986 (see Annex A for a brief history). The outcomes of REF 2014 were published in December 2014.[3] The REF was undertaken by the four UK higher education funding bodies, but managed by the REF team based at the Higher Education Funding Council for England (HEFCE) and overseen by the REF Steering Group consisting of representatives of the four funding bodies.

---

1  6,975 case studies were submitted to REF, of which 296 (c4%) were redacted and not part of this analysis.
2  See www.ref.ac.uk/ for further information [last accessed 27 February 2015].
3  See http://results.ref.ac.uk/ for results [last accessed 27 February 2015].

The outcomes of the REF are used in three ways:

1. Higher education funding bodies use the assessment outcomes to inform the selective allocation of their research funding to HEIs, with effect from the 2015–6 academic year.

2. The assessment provides accountability for public investment in research and produces evidence of the benefits of this investment.

3. The assessment outcomes provide benchmarking information and establish reputational yardsticks.

The REF is a process of expert review. HEIs made submissions to 36 UOAs with submissions being assessed by an expert sub-panel within each, working under the guidance of four main panels, A to D.[4] Sub-panels apply a set of generic assessment criteria to produce an overall quality profile for each submission (REF, 2011).

## Impact is defined as 'any effect on, change or benefit to the economy, society, culture, public policy or services, health, the environment or quality of life, beyond academia'

The allocation of research funding based on non-academic impact is relatively new, with the REF being the first example of its application across a research system (Morgan Jones and Grant, 2013). In 2006 a pilot exercise was carried out during the development of the national Research Quality Framework (RQF) in Australia, which would have introduced impact assessment into its national research assessment exercise (Roberts et al, 2005; Peacock et al, 2006; Donovan, 2008). This was dropped with a change of government in 2007 (Grant et al, 2010), although recently research impact assessment is back on the agenda (Australian Technology Network of Universities and Group of Eight, 2012; Morgan Jones et al, 2013).

In the UK, following a pilot exercise (Technopolis, 2010), the higher education funding bodies concluded that peer review of research impact case studies was a workable approach. It was decided that REF will assess universities on the basis of the quality of research outputs, the vitality of the research environment and the wider impact of research. The weighting for the impact assessment part of the REF was 20% of the total assessment in 2014, and there is a call from some for this to be increased to 25% in the future (Witty, 2013), as originally proposed by the funding bodies.

Impact is defined as 'any effect on, change or benefit to the economy, society, culture, public policy or services, health, the environment or quality of life, beyond academia' (REF, 2011). An impact case study is a short four-page document which has five sections:

1. Summary of the impact
2. A description of the underpinning research
3. References to the research
4. Details of the impact, and
5. Sources to corroborate the impact.

Each case study is assessed by using two criteria:

1. Reach – 'the spread or breadth of influence or effect on the relevant constituencies', and
2. Significance – 'the intensity or the influence or effect'.

---

4  Throughout this document we refer to the main panels as 'panels', unless we are referring to sub-panels, in which case the term 'sub-panel' is used.

These assessments are used to generate a sub-profile of impact case studies for each of the submissions on a four-point scale. In total, approximately £1.6 billion worth of public funding over the next five years will be determined by impact case studies.

The deadline for REF 2014 was 29 November 2013, when 154 HEIs submitted 52,077 Category A full-time equivalent (FTE) staff, 191,232 research outputs and 6,975 impact case studies.[5] The results were published on 18 December 2014.[6] The case studies contain a unique snapshot of the contributions that result from research and knowledge mobilization activities that have been undertaken in UK HEIs over the past 20 years.

## Purpose of this report

This report was commissioned by HEFCE and its partners,[7] and is part of a wider project led by Digital Science (a division of Macmillan Science and Education) working in conjunction with its sister organization, the Nature Publishing Group, and The Policy Institute at King's and Department of Digital Humanities at King's College London. The overall project had two aims:

1. Make the impact case studies freely available in a form and format that will enable any researcher to carry out analysis, using a range of techniques and methods including automated text-mining.

2. Carry out a qualitative and quantitative analysis of the impact case studies, to extract common themes and messages that will form evidence of the broad impact of higher education research on wider society.

The first objective is delivered through an online searchable database which has been developed by Digital Science.[8] The second objective is delivered through this report, led by the team at King's College London. In the request for proposal (RfP), HEFCE and its partners identified a set of key questions that it wished to address. These are summarized in Figure 1 and listed in full in Annex B although, as discussed below, given the timelines and analytical challenges, not all of the questions are addressed in this report. Indeed, it is anticipated that others will develop responses to these and other questions, and refine and improve on the analysis presented here.

## Overview of methodological approach

With a large corpus of case studies to synthesize and time constraints on the project, we had to use a variety of text-mining techniques. Text-mining denotes a set of automated methods for interrogating a large volume of documents. As with any form of mining it can be dangerous and dirty: dangerous, as it is possible to misinterpret information in the text; and dirty, as it involves a lot of experimentation and 'trying by doing'. For this reason we see the analysis presented in this report as 'Version 1.0', and hope that others in time will use these foundations to build on and improve. With this in mind we will describe the process in detail in a technical paper.[9]

We adopted three broad approaches to our analysis using text-mining: topic modelling, keyword searching and information extraction. Topic models aim to uncover hidden thematic structures or 'topics' that occur in a collection of documents (Blei, 2012). A topic consists of a cluster of words or phrases that frequently occur in similar contexts. Keyword searches allowed us to look for specific instances of impact. Information extraction occurred when we matched third party information with the case studies, typically around proper nouns such as countries, cities and institutions.

---

5   See www.ref.ac.uk/ for further information [last accessed 27 February 2015].
6   See http://results.ref.ac.uk/ for results [last accessed 27 February 2015].
7   Scottish Funding Council, Higher Education Funding Council for Wales, Department for Employment and Learning Northern Ireland, Research Councils UK and the Wellcome Trust.
8   http://impact.ref.ac.uk/casestudies
9   Please contact the authors if you are interested in receiving the paper when published. The topic modelling was performed using the Mallet open source package: we can make available details of the topics trained on the dataset of case studies, so others can reconstruct our analysis or apply our topics to other data by using the 'infer-topics' function.

**Figure 1:** Project schema relating key analytical questions with methods and report structure

## Questions (from Request for Proposal)

1. **What types of impact outcomes have been submitted?**
2. **What are the pathways by which different types of impact outcome have been realised?**
3. To what extent are "negative" findings and "learning" type impacts included?
4. **What are the time lags between underpinning research and impact outcome exampled in the impact case studies?**
5. **What quantitative data (for inputs and/or outcomes) can be extracted and synthesised from the impact case studies e.g. commercialisation revenue, visitor numbers? Is it possible to estimate overall return on investment figures?**
6. Do the case studies contain quantitative information on impact outcome that is sufficiently common and general to be developed into metrics of research impact?
7. **According to the impact case studies, what types of research users benefit from HE research, and to what extent?**
8. What are the characteristics of the underpinning research outputs on which impact case studies are based (in terms of methodologies, approaches or research topics and inter-disciplinarity?)
9. Can we learn anything about how research collaboration affects impact?
10. How does the impact described in the REF case studies relate to the Government's Industrial Strategy and other devolved administration economic/industry strategies?
11. Can we learn anything about connections between social and economic impact and citation impact?

## Report Structure

**Chapter 1**
**Introduction**

**Chapter 2**
**Overview**

**Chapter 3**
**Nature & beneficiaries of impact**
Type of impacts
Impact pathways
Beneficiaries
Time lags

**Chapter 4**
**Scale of impact**
6 'deep mines'

**Chapter 5**
**Concluding reflections**

## Methods

**Topic modelling**
**Information extraction**
**Keyword-in-context**
**Qualitative analysis**

**Box A:** Summary of text-mining approaches

Text and data-mining (TDM) is the process of deriving information from machine-read material. It works by copying large quantities of material, extracting the data and recombining it to identify patterns.

Text-mining has become an accepted method in many areas of scientific analysis to deal with the fast-growing number of texts and documents with which researchers have to work. Text-mining techniques have been successfully applied to analyse specific kinds of diabetes from UK PubMed Central, a large archive of full-text content in biomedicine and health, as well as in the humanities, to deal with millions of books that an individual researcher cannot read without help.[10]

All these domains have in common natural language processing, statistical modelling and machine learning techniques that help turn texts into research outputs by opening up the underlying semantics of textual content. Computationally this is a challenging task, because natural language texts are inconsistent and full of ambiguities that are often easier for humans to understand than for computers to process. Such ambiguities might include homonyms, where the meaning of the word only becomes clear from its context, to irony, where words are used to express a different meaning from their literal one.

Text-mining entails a range of applications and methods that often are used together to achieve the best possible results.[11] Broadly speaking, the first step is that a relevant collection of documents is selected. The second step is to extract structured data in a machine-readable format from these documents. Part-of-speech entities such as subject and noun phrases might be tagged. In the third step, useful information is extracted. Based on the discovered part-of-speech relationships, for instance, specific entities such as location or organization names are extracted. The final step discovers new knowledge and relationships in the documents.

For this study we have used a number of mature and well-developed text-mining techniques to answer specific challenges.

- Topic modelling was used to determine common topics across the whole corpus. Sixty-five topics were found (of which 60 were used) using the Apache Mallet Toolkit Latent Dirichlet Allocation (LDA) algorithm.[12] Topics are based on the frequency with which certain related words appear. A document about cars will have, next to car, terms such as 'driving', 'motor', etc. The topics are finally weighted according to how much they describe the parts of the document collection.

- Information extraction techniques aim to extract automatically structured information from unstructured documents. These are often highly complex tasks that work best in well-defined sub-domains such as news articles or company email archives. Typical subtasks of information extraction include named entity recognition, where the names of people, organizations, etc. are extracted, co-referencing of extracted named entities and relationship extraction. This was used to identify references to geographic locations.

- Keyword-in-context (KWIC). KWIC was used originally in text-mining to represent each keyword in a collection within the context it might appear. It is used in our analysis to develop an alphabetical list of keywords displayed with their surrounding text (word) context. This helps to disambiguate the meaning of these keywords.

As summarized in Figure 1 (page 15), we used topic models to identify and describe different types of impact. We typically used keywords to select a set of case studies that we read and analysed using qualitative approaches. We used information extraction approaches and keywords to identify and describe different regional groups (or beneficiaries). Box A and Annex C provide a technical summary of the text-mining approaches and as noted above, further detail will be made available in a forthcoming technical paper.

To help us prioritize our analysis we hosted a workshop in November 2014. At this workshop we asked participants to review the key questions that were identified in the RfP, and to articulate some specific questions that we could use to illustrate the richness of the information contained in the case studies (these are reported as 'deep mines' in Chapter 4). The results of the workshop are provided in Annex B.

## Caveats and limitations of analysis

There are a number of very important challenges we encountered in our analysis, and these need to be taken into account in reading this report. These fall into two groups: the first relate to the limitations of our analysis, and the second to the limitations of the case studies as research material.

With regard to limitations in the analysis, there are two points to consider. The first is the time period in which the analysis was undertaken. The project ran for six months, of which about half was dedicated to text-mining and the remainder to quality assurance, analysis and reporting. The nature of this type of exercise means that it is inevitably iterative, with each iteration improving the specificity and accuracy of what we were doing. With more time we would have been able to review and improve the text-mining approaches. Moreover, we know and accept that in our analyses there will be examples of false positives: that is, the inclusion of case studies that are not actually associated with the context of what we were analysing.

The second challenge was associated with the way in which the case studies were structured. With the benefit of hindsight, our analysis would have been much easier if the case studies had greater structure and used standardized definitions. Given that the case studies spanned a 20-year period, organization names have changed in that time and keyword searches were not sophisticated enough to capture some key information. For example, a drop-down list of funders for case study authors would have enabled that information to be captured in a structured way. Additionally, we came across examples of people using the same quantifiable information to evidence impact, but using different (and thus non-comparable) definitions. A list of standard definitions in future iterations of REF would facilitate some of the analyses that we present in subsequent chapters.

There were three challenges associated with the limitations of the case studies as research material. The first was associated with the way that impact is articulated and described. The sentiment in the language of the case studies is universally positive, reflecting its purpose as part of an assessment process. The choice of language should have been moderated through expert review of the panels in assigning 'quality profiles' (or grades) (Manville et al, 2015a). However, this synthetic analysis does not take into account different quality profiles,[13] and it is likely that it will include impacts that have been overstated in the case studies but moderated in the assessment. A more subtle point – and one for future research – is whether the use of professional writers influenced the vocabulary of the case studies, or whether the vocabulary used in different disciplines biases the formation of topics in the text-mining.

10  See www.theatlantic.com/technology/archive/2013/10/googles-ngram-viewer-goes-wild/280601/ for further information [last accessed 27 February 2015].
11  See www.jisc.ac.uk/reports/value-and-benefits-of-text-mining for further information [last accessed 27 February 2015].
12  See http://programminghistorian.org/lessons/topic-modeling-and-mallet for further information [last accessed 27 February 2015].
13  Although it is worth noting that 84% of the case studies were rated as 4* (world leading) or 3* (internationally excellent).

The second challenge was that the corpus of case studies described a very selective set of impacts, complied with a specific set of 'rules' as defined in the REF guidance documents (REF, 2011; 2012). We know from other work, for example, that certain types of impacts were deemed to be 'too risky' for submissions, as there was perceived ambiguity in the guidance documents (Manville et al, 2015b).

The third challenge is that in reading a selection of case studies, we came across examples of identical or near-identical submissions, and thus there is a risk of double counting in our analysis.[14] We used Latent Semantic Analysis to assess the degree of similarity across the case studies, and identified around 100 pairs of documents with a high degree of text similarity. Of these, 81 had been submitted either to two different UOAs but usually by the same university, or by two different universities but usually to the same UOA. However, there were a dozen instances of these similar documents being submitted by different universities to different UOAs, suggesting a degree of co-operation and collaboration between the research teams involved.

In highlighting these important caveats we do not wish to undermine the analysis in the subsequent chapters, rather to help frame its interpretation. Although implicit throughout the report, we also should emphasize that everything we have reported is a result of analysis on the written text describing impact within the REF impact case studies, and not an analysis of impact per se. Indeed, in Chapter 5 we reflect on how lessons can be learned from this pioneering assessment to ensure that future large-scale research impact assessment – and particularly the analysis of those assessments – can be improved.

---

14  REF guidance indicates: 'Where more than one submitting unit made a distinct and material research contribution to an impact, each of those submitting units may submit a case study of the impact. However, the case studies should not be identical, because each submitting unit will need to show that its research made a distinct and material contribution to the impact. This applies whether an HEI wishes to submit the same impact in different submissions, or different HEIs wish to submit the same impact'. See http://www.ref.ac.uk/about/guidance/faq/ for further information [last accessed 27 February 2015]

# 2 | An overview of the impact case studies

# 2 | An overview of the impact case studies

This chapter summarizes the basic characteristics associated with the case studies and used in the descriptive analyses in subsequent chapters. This includes information on submitting Higher Education Institutions (HEIs), distributions by Unit of Assessment (UOA) and the four main panels, and the field of research (FOR) underpinning the impact described in the case studies.

## A total of 6,975 case studies were submitted to the Research Excellence Framework, of which 6,679 form the basis of the synthetic analysis

As summarized in Table 1, 6,975 case studies were submitted to the Research Excellence Framework (REF), although 296 of these were excluded from our analysis as they were wholly redacted.[15] A further 428 case studies were partially redacted: that is, elements of the text are removed from public versions of the case studies and the redacted text is not used in the text-mining.

The distribution of analysed case studies across the panels was broadly even, ranging from 22% (Panel B – which had the highest rate of redaction and partial redaction) to 30% (Panel C). Broadly speaking, Panel A focuses on the life sciences, Panel B on the engineering and physical sciences, Panel C on the social sciences and Panel D on the arts and humanities.

## A total of 154 Higher Education Institutions made Research Excellence Framework submissions, with the number of case studies per submission ranging from 2 to 260

The size of submissions from HEIs was heavily skewed, as illustrated in Figure 2[16]: the largest number of case studies submitted was 260, and the smallest was 2 case studies.[17] The median size of submission was 28 case studies. To aid analysis we split the HEIs into three groups, based on the size of the submission:[18]

- Group I – the 25 HEIs which accounted for half of all the case studies formed
- Group II – the 39 HEIs which accounted for the next 30% of case studies
- Group III – the 90 HEIs which accounted for the remaining 20% of submitted case studies

---

15  When making submissions, HEIs were able to identify specific case studies that either should not be published at all due to their confidential nature, or that should be redacted prior to publication. HEIs were able to provide redacted versions suitable for publication after the close of submissions.
16  In Figure 2 (and the subsequent analysis of HEIs) there are 153 institutions. This is because one HEI only submitted joint submissions with another HEI.
17  The minimum number of case studies for a submission was two.
18  A similar grouping was used in selected HEIs for the evaluation submission preparation for impact assessment (Manville et al, 2015a), and for the evaluation of how panels assessed case studies (Manville et al, 2015b).

**Table 1:** Distribution of submitted case studies by panel and redaction status

| | Panel A (Life sciences) | Panel B (Engineering and Physical sciences) | Panel C (Social sciences) | Panel D (Arts & humanities) | Total |
|---|---|---|---|---|---|
| Total number of submitted case studies | 1,621 | 1,667 | 2,040 | 1,647 | 6,975 |
| Number of redacted case studies (% of total number of submitted case studies by panel) | 27 — 2% | 182 — 11% | 67 — 3% | 20 — 1% | 296 — 4% |
| **Total number of case studies analysed (as % of all analysed case studies)** | **1,594 — 24%** | **1,485 — 22%** | **1,973 — 30%** | **1,627 — 24%** | **6,679** |
| Number of partially redacted case studies (% of total number of submitted case studies by panel) | 87 — 5% | 209 — 13% | 75 — 4% | 57 — 3% | 428 — 6% |

## The number of case studies analysed by Units of Assessment ranged from 51 to 413

There was an eight-fold difference in the number of case studies analysed across the 36 UOAs. The UOA with the largest number of case studies was Business and management studies (UOA 19) with 413 case studies, and the smallest Civil and construction engineering (UOA 14) with 51 (Table 2, page 23).

## The research underpinning the case studies was multidisciplinary and diverse

As part of the database curation, Digital Science tagged the research descriptions within the 'Underpinning research' (Section 2) of the case studies using FORs developed by the Australian and New Zealand Statistical Bureaus.[19] FORs allow research and development (R&D) activity to be categorized according to the methodology used in the R&D, rather than the activity of the unit performing the R&D or the purpose of the R&D. The FOR is a hierarchical classification with three levels: namely, 22 divisions (two digits), 157 groups (four digits) and 1,238 fields (six digits), with each level identified by a unique number.[20] Classification of the underpinning research in the impact case studies required the use of 149 out of the 157 available four-digit group codes. The assignment of codes was made in an automated way by semantic analysis of the text describing the research activity. Up to three FOR codes were assigned to each case study, in order to allow for multidisciplinary and interdisciplinary research. Manual checking was used to validate these assignments, and some reclassification was applied where inappropriate assignments had been made.

---

19  http://www.arc.gov.au/pdf/ANZSRC_FOR_codes.pdf [last accessed 27 February 2015]

20  The FOR has three hierarchical levels, namely divisions (at the broadest level), groups and fields (at the finest level). The division represents a broad subject area or research discipline, while groups and fields within represent increasingly detailed dissections of these categories. Divisions, groups and fields are assigned unique two-digit, four-digit and six-digit codes respectively. The FOR classification has 22 divisions, 157 groups and 1,238 fields. The hierarchical structure of the FOR is as illustrated as follows: Division 02 Physical Sciences; Group 0206 Quantum Physics; Field 020603 Quantum Information, Computation and Communication.

**Figure 2:** Distribution of number of case studies submitted to REF 2014 by HEIs, split into three groups for analysis
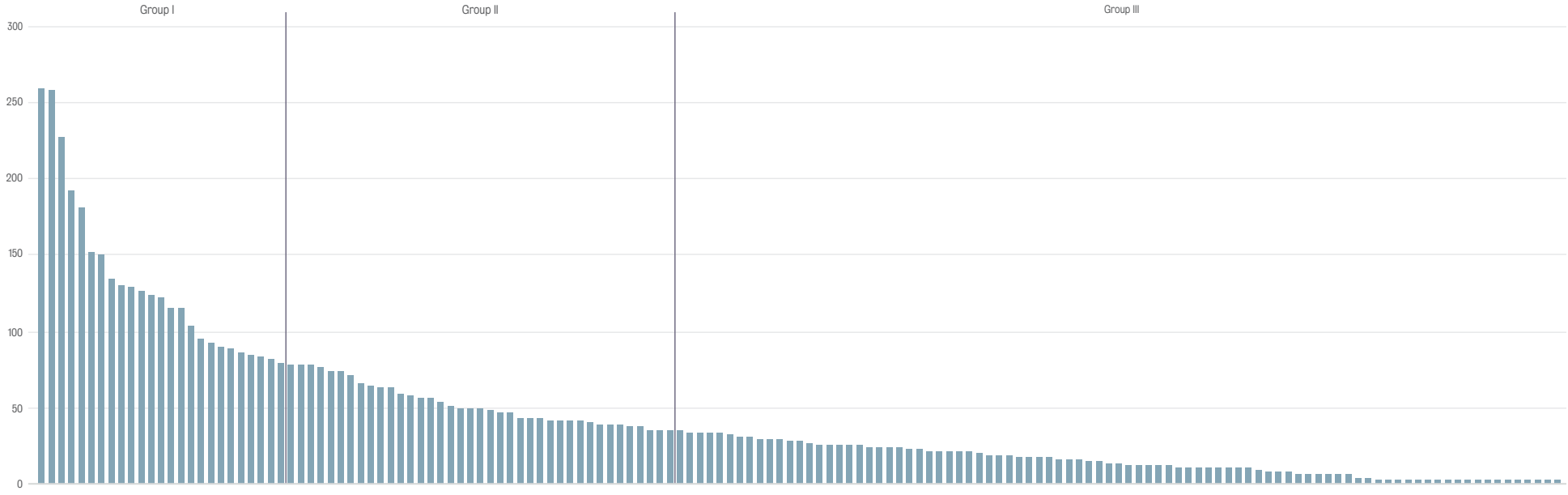
**Table 2:** Number of analysed case studies by UOA[21]

| Panel | | Unit of assessment | Number of case studies | % of total |
|---|---|---|---|---|
| Panel A | UOA 1 | Clinical medicine | 383 | 6% |
| | UOA 2 | Public health, health services and primary care | 163 | 2% |
| | UOA 3 | Allied health professions, dentistry, nursing, pharmacy | 347 | 5% |
| | UOA 4 | Psychology, psychiatry, neuroscience | 318 | 5% |
| | UOA 5 | Biological sciences | 257 | 4% |
| | UOA 6 | Agriculture, veterinary and food science | 126 | 2% |
| Panel B | UOA 7 | Earth systems and environmental sciences | 171 | 3% |
| | UOA 8 | Chemistry | 125 | 2% |
| | UOA 9 | Physics | 182 | 3% |
| | UOA 10 | Mathematical sciences | 210 | 3% |
| | UOA 11 | Computer science and informatics | 253 | 4% |
| | UOA 12 | Aeronautical, mechanical, chemical and manufacturing engineering | 124 | 2% |
| | UOA 13 | Electrical and electronic engineering, metallurgy and materials | 127 | 2% |
| | UOA 14 | Civil and construction engineering | 51 | 1% |
| | UOA 15 | General engineering | 242 | 4% |
| Panel C | UOA 16 | Architecture, built environment and planning | 141 | 2% |
| | UOA 17 | Geography, environmental studies and archeology | 235 | 4% |
| | UOA 18 | Economics and econometrics | 98 | 1% |
| | UOA 19 | Business and management studies | 413 | 6% |
| | UOA 20 | Law | 217 | 3% |
| | UOA 21 | Politics and international studies | 167 | 3% |
| | UOA 22 | Social work and social policy | 187 | 3% |
| | UOA 23 | Sociology | 97 | 1% |
| | UOA 24 | Anthropology and development studies | 80 | 1% |
| | UOA 25 | Education | 214 | 3% |
| | UOA 26 | Sport and exercise sciences, leisure and tourism | 124 | 2% |
| Panel D | UOA 27 | Area studies | 69 | 1% |
| | UOA 28 | Modern languages and linguistics | 190 | 3% |
| | UOA 29 | English language and literature | 283 | 4% |
| | UOA 30 | History | 263 | 4% |
| | UOA 31 | Classics | 59 | 1% |
| | UOA 32 | Philosophy | 98 | 1% |
| | UOA 33 | Theology and religious studies | 75 | 4% |
| | UOA 34 | Art and design: history, practice and theory | 236 | 4% |
| | UOA 35 | Music, drama, dance and performing arts | 196 | 3% |
| | UOA 36 | Communication, cultural and media studies, library and information management | 158 | 2% |
| **Total** | | | **6,679** | **100%** |

21   The colour shading in this table is to provide a key to the visualizations that are presented in subsequent chapters. Each UOA has a different tone of a colour that is associated with the main panels.

As illustrated in Figure 3, 15,963 FOR codes were assigned to the 6,679 case studies. Just under two-thirds of the case studies had two or more two-digit codes (out of a possible 22), providing an indication of the high degree of interdisciplinarity in the research underpinning the impact case studies. The vast majority (87%) of case studies had two or three four-digit codes assigned, illustrating the multidisciplinary nature of the underpinning research.

Figure 4 plots the relationship between the FORs (on the left-hand side) and UOAs (on the right-hand side). What is evident from this visualization is that multiple FORs underpin the case studies and this is weakly correlated with the UOA. In other words, the underpinning research within UOAs is multidisciplinary in its own right.

That said, 75% of all the assigned codes were for 36 fields (as illustrated in Figure 5, page 26). The most common field of research was public health (accounting for about 10% of all assigned codes), followed by historical studies, cultural studies and policy and administration.

**Figure 3:** Number of FOR codes assigned to case study



Number of Field of Research (FOR) codes assigned to a case study
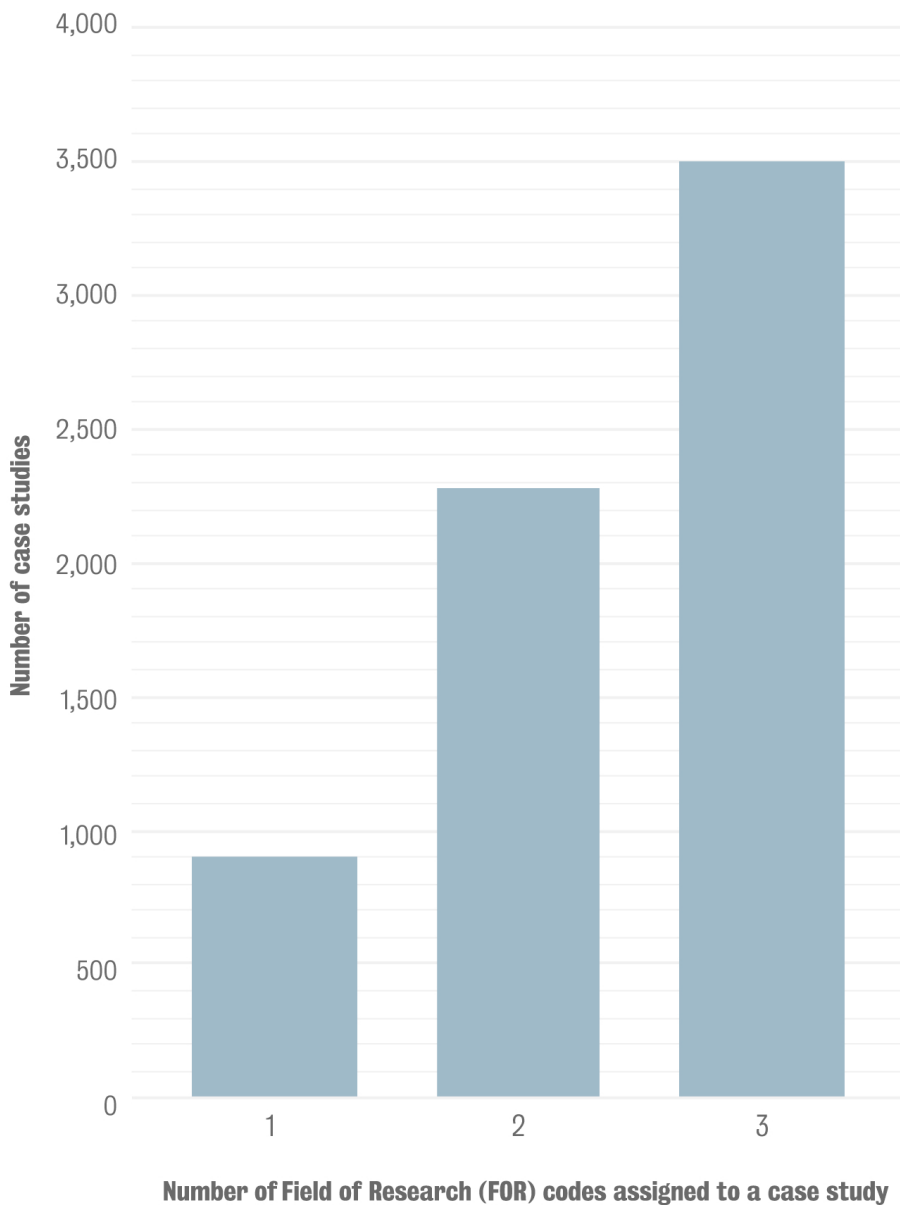
**Figure 4:** Alluvial diagram linking FORs with UOAs
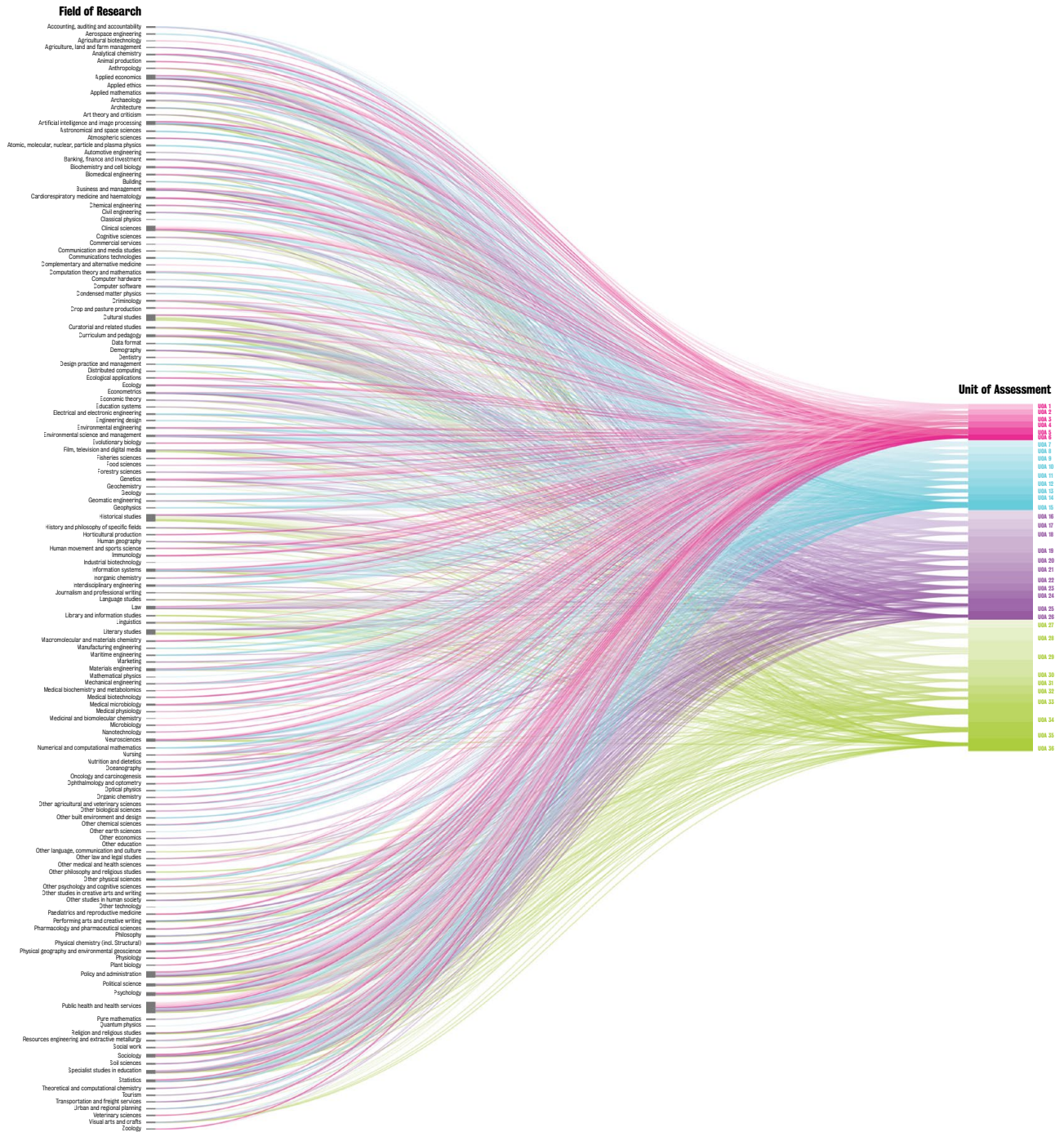
**Figure 5: Most commonly assigned codes (top 75 % of all codes assigned)**



Number of case studies coded to a Field of Research (FOR), allowing for up to three codes per case study
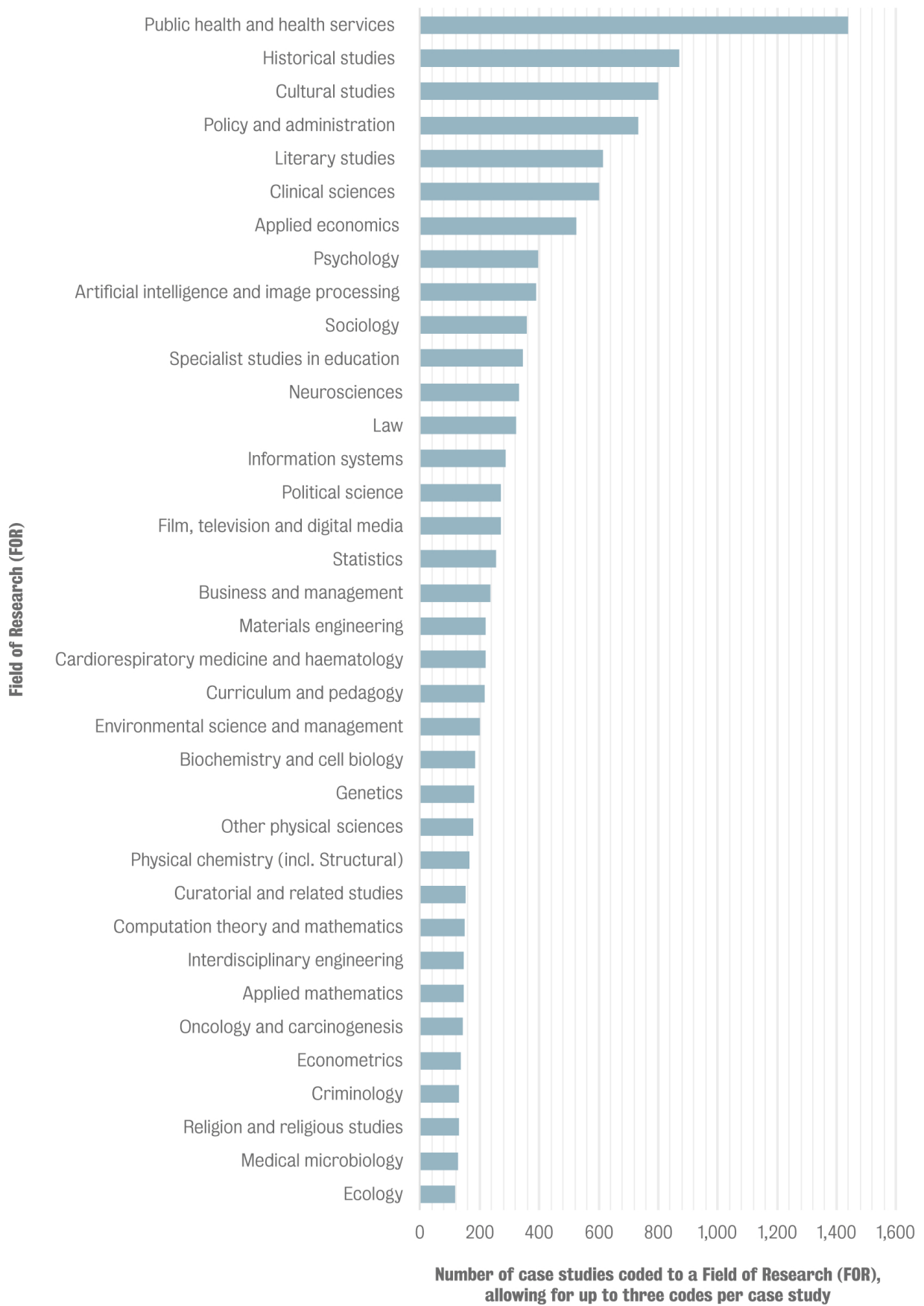
# 3 | The nature and beneficiaries of research impact

# 3 | The nature and beneficiaries of research impact

This chapter examines different types of research impact, looks at its international spread, explores who benefits and examines the time lag between the underpinning research and the realization of the impact. As noted in Chapter 1, we used topic modelling to capture different types of impact. Topic modelling is the statistical analysis of a corpus of documents, in this case the 'Details of the impact' (Section 4) of the 6,679 impact case studies. At the outset of the project we reviewed various frameworks and taxonomies of impact, but concluded that these were often context-specific and, with some exceptions, conceptual.

One of the advantages of the REF case studies is that they allow scholars to develop empirically derived taxonomies of impact. With this in mind we identified 65 potential topics for initial analysis. It became apparent that five of the topics were not relevant in identifying impacts: that is, they had words associated with the assessment process rather than substantive impact. In other words, our analysis is based on 65 topics, but we present only 60 of them: these topics are used to examine the nature of research impact (see Annex D for list of the 60 topics with the top nine words and Annex C for a methodological summary).

The topics are based on the c6.2 million words harvested from the 'Details of the impact' (Section 4) of the case studies; c3.7 million words (c60%) remained after 'stop' words (eg 'and', 'but', etc.) were removed. This resulted in c100,000 unique words. To provide an overview of the text analysis, we generated a word cloud (Figure 6 and report cover) for which we excluded words that were mentioned 10 or fewer times (leaving c20,000 unique words), with the word cloud showing the most frequent 684 words.[22]

---

22  We used the word cloud generator www.jasondavies.com/wordcloud/ for this visualization.

**Figure 6:** Word cloud of most frequent words from the 'Details of the impact' (Section 4) of the case studies

Figure 7 illustrates the relationships between the 60 topics. As can be seen in the outer circle, the largest topic is 'Informing government policy'[23], followed by 'Parliamentary scrutiny'[24] and 'Technology commercialization'[25]. Each case study could be tagged to up to three topics, and the lines between the topics illustrate these relationships. For example, the thickest line is between 'Informing government policy' and 'Parliamentary scrutiny', showing that these two topics were most frequently associated across the case studies; followed closely by the relationship between 'Informing government policy' and 'Community and local government'[26] and between 'Instrumentation'[27] and 'Technology commercialization'.

---

23  The top nine words related to this topic were: *develop polici nation plan govern inform work strategi assess*.
    Note the words have been 'stemmed' where the end of the word is truncated.
24  The top nine words related to this topic were: *polici govern report public uk committe debat evid commiss*.
    Note the words have been 'stemmed' where the end of the word is truncated.
25  The top nine words related to this topic were: *technolog compani develop product univers commerci system market*
    *industri*. Note the words have been 'stemmed' where the end of the word is truncated.
26  The top nine words related to this topic were: *local commun project citi council social peopl fund develop*.
    Note the words have been 'stemmed' where the end of the word is truncated.
27  The top nine words related to this topic were: *laser instrument materi product process imag manufactur develop industri*.
    Note the words have been 'stemmed' where the end of the word is truncated.

**Figure 7:** Chord diagram illustrating size and relationship of impact topics

## Different types of impact are more common in different disciplines

The heat map in Figure 8 illustrates the distribution of topics in the case studies by UOA and panels. To aid the reading of Figure 8 we have ordered the topics by the panel in which they frequently occur. Within each cell in the heat map is a circle that illustrates the proportion of case studies that have been allocated to a topic for the given UOA. For example, in the top left-hand corner for the UOA 1 (Clinical medicine), 41% of case studies were allocated to the 'Clinical guidance'[28] topic (fourth from the left). From this figure there are a number of important observations. First, there is a clear relationship between the type of impact and the UOA. So for example, in Panel A the topics are associated with health impacts (such as the topic 'Clinical guidance', which occurs 303 times for the 1,594 Panel A case studies, ie 19%); while in Panel D the topics are associated with a diverse range of engagement activities, with 'Media'[29] being the most common (occurring 424 times in the 1,627 Panel D case studies, ie 26%). Second, there are important examples of topics that cut across a number of UOAs. For example, 'Technology commercialization' cuts across Panels A and B, while 'Informing government policy' occurs in all four panels, albeit less so in Panel D.

---

28  The top nine words related to this topic were: *guidelin patient clinic treatment recommend stroke nice risk trial.*
   Note the words have been 'stemmed' where the end of the word is truncated.
29  The top nine words related to this topic were: *public bbc media radio programm interview time broadcast articl.*
   Note the words have been 'stemmed' where the end of the word is truncated.

**Figure 8:** Heat map illustrating distribution of topics by UOA and main panels
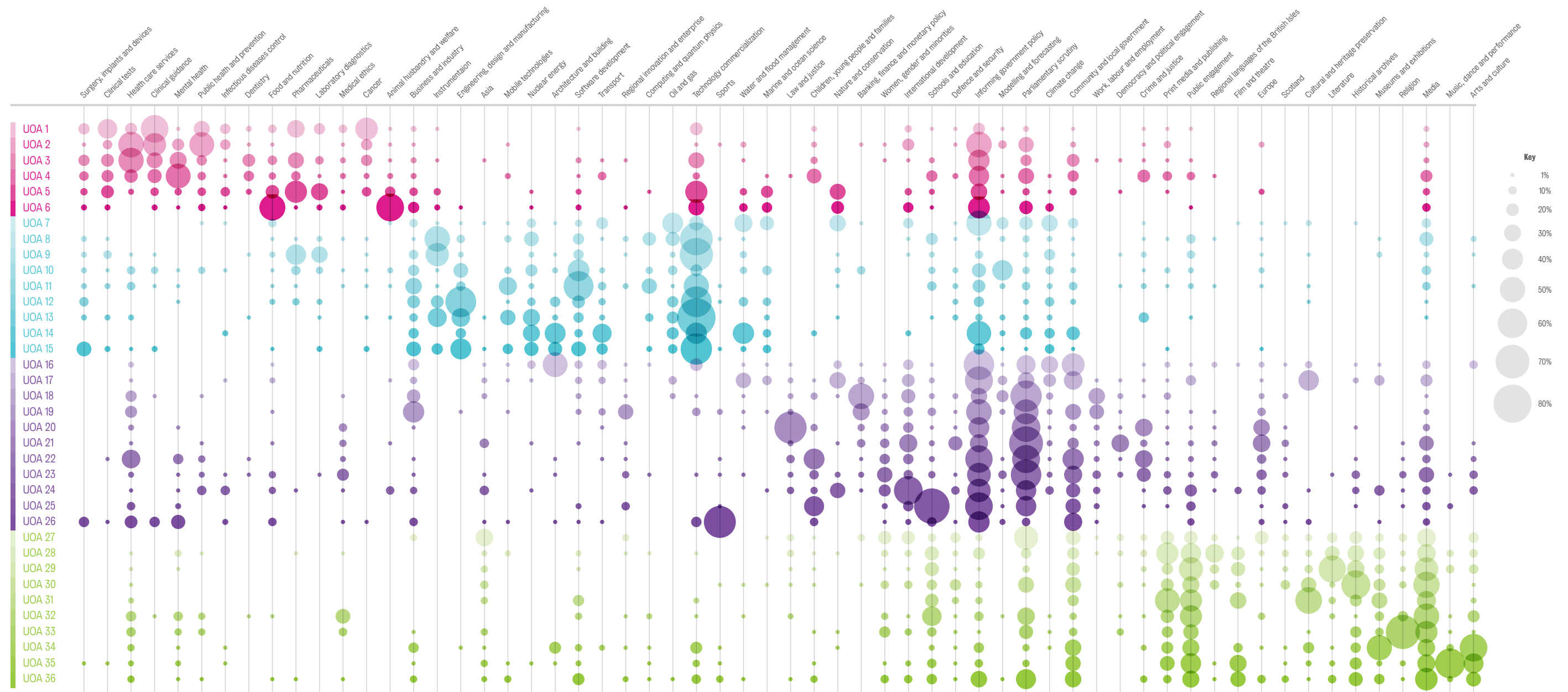
In order to assess further the degree of multiple impacts we analysed the distribution of UOAs within the 60 topics, as illustrated in the impact wheels in Figures 9 and 10. Figure 9 illustrates how the topic 'Schools and education'[30] is distributed across the four main panels and 36 UOAs. To ease interpretation we have colour-coded the panels and used different shades of the same colour to indicate associated UOAs. For example, Panel A is coloured red, with units 1–6 in different shades of red. The size of the spokes in the impact wheels indicate the frequency with which that topic occurred in that UOA: in Figure 9 the longest spoke is for UOA 25 (Education), indicating that 69% of case studies in that UOA were assigned to the topic 'Schools and education'.[31] In Figure 10 we have provided abridged impact wheels for all 60 topics (and the underlying data are provided online[32]).

**Figure 9:** Impact wheel ('Schools and education', n=381)



---
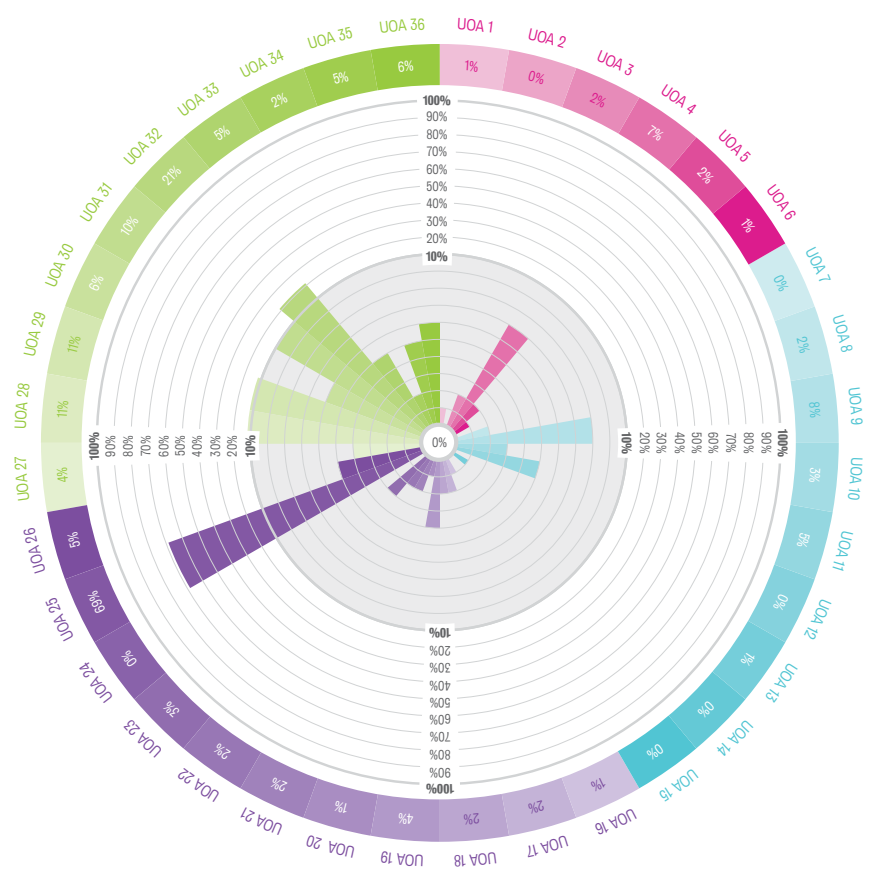
30 The top nine words related to this topic were: *educ school teacher student teach learn univers develop curriculum*. Please note the words have been 'stemmed' where the end of the word is truncated.

31 Recalling that assignment meant that the topic was in the top three highest ranking topics for each case study.

32 www.hefce.ac.uk/analysisREFimpact/

**Figure 10:** Simplified impact wheels for all 60 topics



Legend (left): UOA 1, UOA 2, UOA 3, UOA 4, UOA 5, UOA 6, UOA 7, UOA 8, UOA 9, UOA 10, UOA 11, UOA 12, UOA 13, UOA 14, UOA 15, UOA 16, UOA 17, UOA 18, UOA 19, UOA 20, UOA 21, UOA 22, UOA 23, UOA 24, UOA 25, UOA 26, UOA 27, UOA 28, UOA 29, UOA 30, UOA 31, UOA 32, UOA 33, UOA 34, UOA 35, UOA 36

Row 1: Surgery, implants and devices | Clinical tests | Health care services | Clinical guidance | Mental health | Public health and prevention | Infectious disease control | Dentistry | Food and nutrition | Pharmaceuticals | Laboratory diagnostics | Medical ethics

Row 2: Cancer | Animal husbandry and welfare | Business and industry | Instrumentation | Engineering, design and manufacturing | Asia | Mobile technologies | Nuclear energy | Architecture and building | Software development | Transport | Regional innovation and enterprise

Row 3: Computing and quantum physics | Oil and gas | Technology commercialization | Sports | Water and flood management | Marine and ocean science | Law and justice | Children, young people and families | Nature and conservation | Banking, finance and monetary policy | Women, gender and minorities | International development

Row 4: Schools and education | Defence and security | Informing government policy | Modelling and forecasting | Parliamentary scrutiny | Climate change | Community and local government | Work, labour and employment | Democracy and political engagement | Crime and justice | Media | Public engagement

Row 5: Regional language of the British Isles | Film and theatre | Europe | Scotland | Cultural and heritage preservation | Literature | Historical archives | Museums and exhibitions | Religion | Print, media and publishing | Music, dance and performance | Arts and culture

35

## Different types of Higher Education Institutions specialize in different types of impact

Figure 11 distributes the topics by the type of HEI using the three size groups described in Chapter 2. As mentioned previously, Group I is the 25 research-intensive HEIs which accounted for half of all the case studies. The 39 HEIs that accounted for the next 30% of case studies were Group II and the final 90 HEIs accounting for the remaining 20% of submitted case studies were Group III. If there were no specialization, one would expect that for a given topic 50% of the case studies would be Group I, 30% Group II and 20% Group III, as illustrated by the vertical lines in the figure.

We have ordered Figure 11 by topics that have the greatest variance from this expected distribution. As can be seen, the Group I HEIs make a disproportionate contribution to the 'Clinical guidance' and 'Dentistry'[33] topics, and the Group II HEIs to 'Marine and ocean science'[34] and 'Work, labour and employment'[35]. Perhaps the most interesting observation from Figure 11 is the degree of specialization of the Group III HEIs. As summarized in Table 3, for nine topics, 30% or more (versus an expected 20%) of the case studies came from these less research-intensive HEIs. This may be a result of the selection of case studies for submission as much as specialization itself.

**Table 3:** Topics where less research-intensive Group III HEIs make a disproportionate contribution

| Topic<br>(Top nine words associated with topic) | Proportion of case studies from Group III HEIs (expected = 20%) |
| --- | --- |
| **Sports**<br>(sport game coach footbal athlet olymp perform physic player) | 45% |
| **Regional innovation and enterprise**<br>(innov busi region sme birmingham enterpris support programm univers) | 43% |
| **Arts and culture**<br>(art artist work cultur creativ project public audienc exhibit) | 40% |
| **Music, drama and performance**<br>(music perform danc work sound audienc concert record festiv) | 37% |
| **Religion**<br>(church religi christian religion faith cathol spiritu confer bibl) | 35% |
| **Women, gender and minorities**<br>(women equal gender migrat divers ethnic commun group refuge) | 32% |
| **Schools and education**<br>(educ school teacher student teach learn univers develop curriculum) | 32% |
| **Community and local government**<br>(local commun project citi council social peopl fund develop) | 31% |
| **Asia**<br>(china chines india arab indian asian intern east foreign) | 30% |

33  The top nine words related to this topic were: *kcl dental drug oral treatment king prof scott health*.
    Please note the words have been 'stemmed' where the end of the word is truncated.
34  The top nine words related to this topic were: *marin fish fisheri sea coastal ship ocean manag environment*.
    Please note the words have been 'stemmed' where the end of the word is truncated.
35  The top nine words related to this topic were: *employ union labour trade work worker wage employe social*.
    Please note the words have been 'stemmed' where the end of the word is truncated.

**Figure 11:** Observed and expected distribution of topics by HEI group
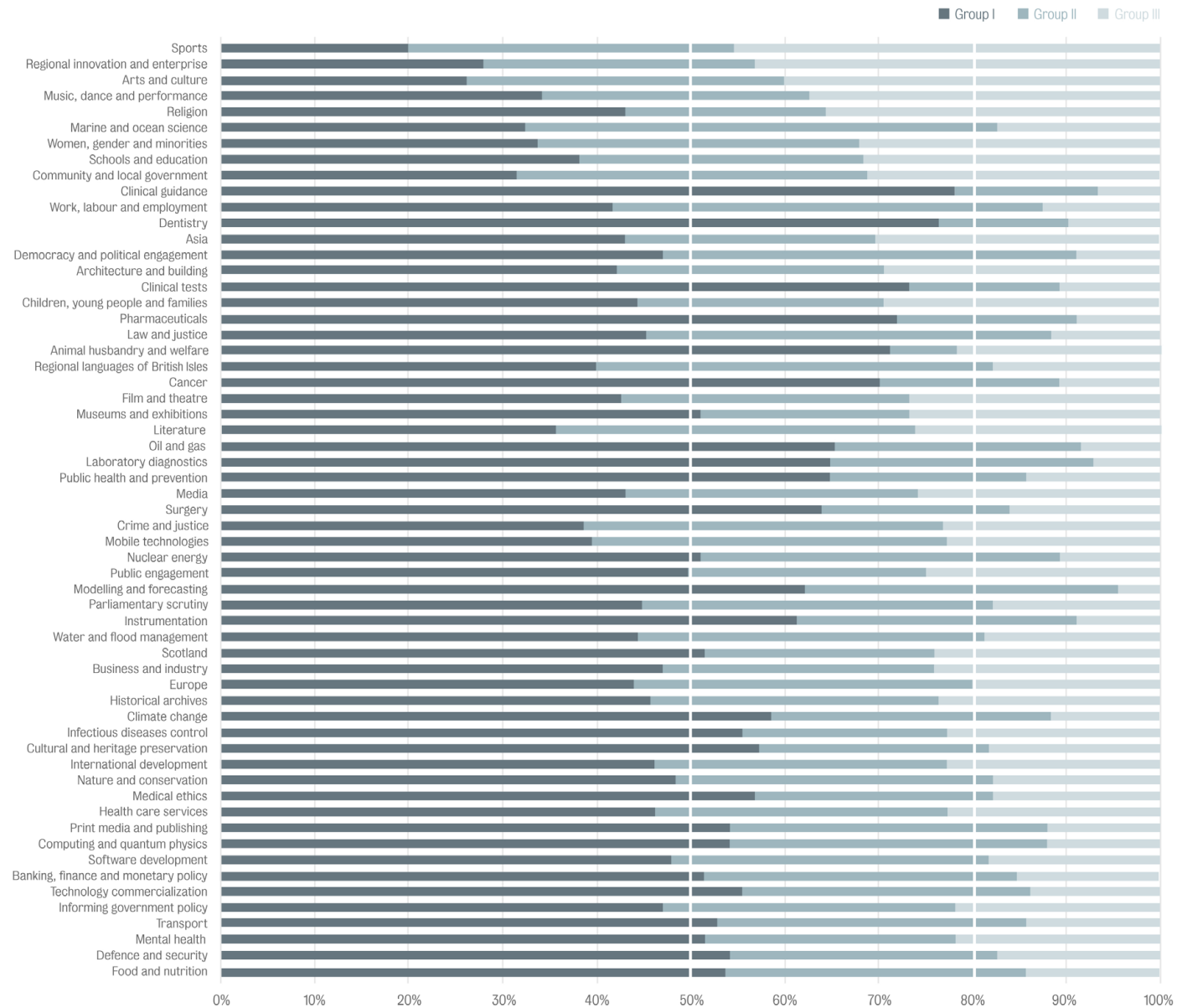
## There is a diverse range of impact pathways

The alluvial diagram in Figure 12 relates the 149 FORs within the 36 UOAs with the 60 impact topics. We have used the same colour palette as for the impact wheels – where shades of red are for units in Panel A, shades of blue for Panel B, shades of purple for Panel C and shades of green for Panel D. While Figure 12 shows a lot of crossover between the FORs and impact topics, the most striking observation to be made is the 3,709 unique different ways that the research to impact pathway takes.

**Figure 12:** Alluvial diagram linking FORs with UOAs to impact topics (impact pathways)
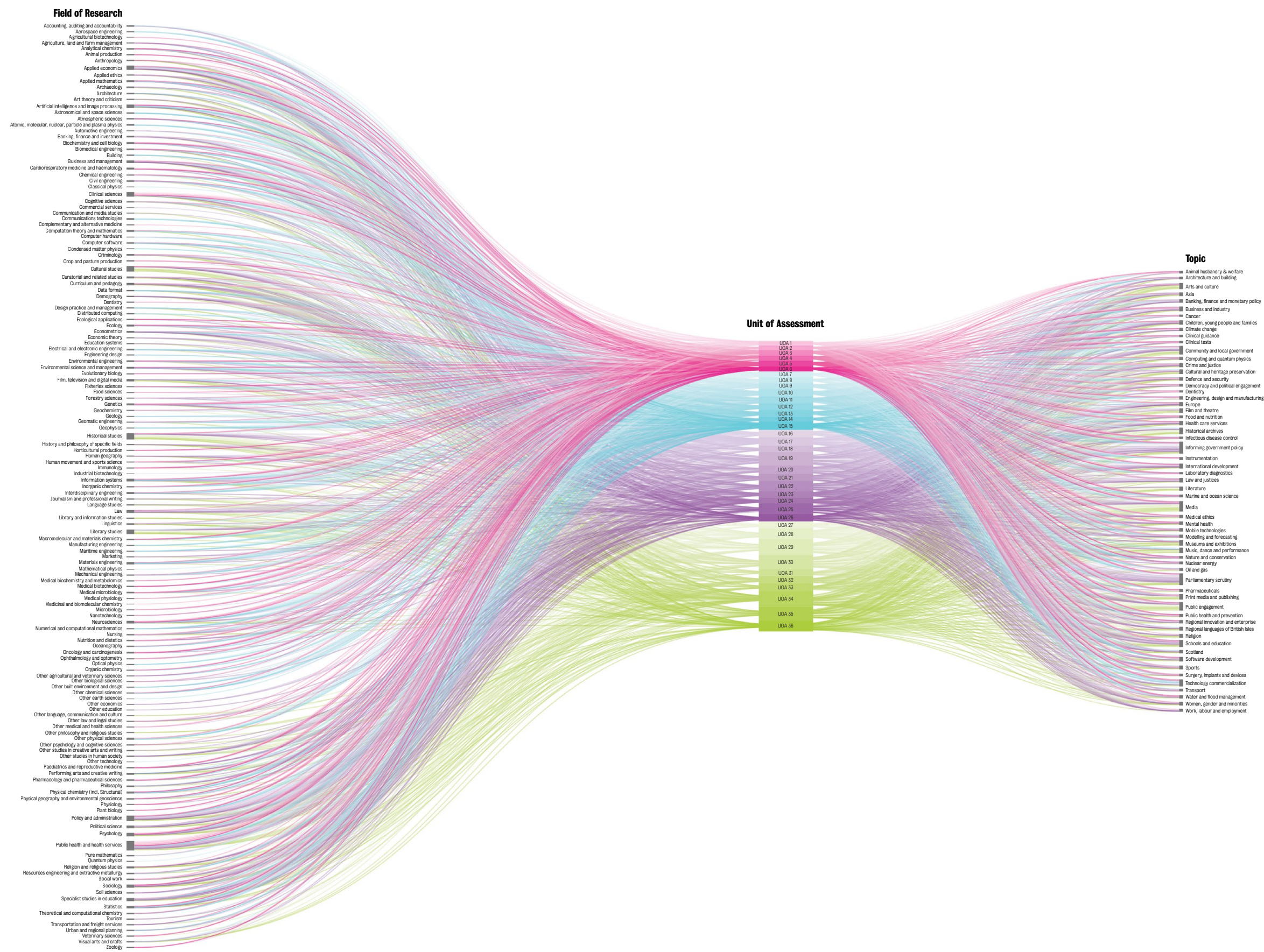
## UK Higher Education Institutions have a global impact

The map in Figure 13 illustrates the geographical reach of the impacts of research that has been undertaken in UK HEIs over the past 20 years. The information presented in this figure was generated through combining information extraction of location names with a third-party database of geolocations (known as 'geotagging').[36] Locations were tagged as either city, region or country, and form part of the searchable online database of case studies developed by Digital Science.[37] Overall there were 23,420 geotags in 'Details of the impact' (Section 4) of the case studies (an average of three per case study): of these, 17,932 were outside the UK.

The information presented in Figure 13 is at the country level, with cities and regions assigned to countries. As can be seen, UK academics have made contributions to all the countries of the world,[38] although more than 48% of those mentions were in 10 countries, as illustrated in Table 4.[39] The nature of this impact is explored further in the 'deep mine' on the BRIC (Brazil, Russia, India and China) countries, which identifies both strategic (ie collaboration with an international partner organization was created to conduct the research) and incidental (ie positive but unintended outcome in the specified country as a result of the research) types of impact.

**Table 4:** Countries where UK research had the most instances where impact has been identified, as measured by mentions in 'Details of the impact' (Section 4) of the case studies

| Country | Number of mentions | % of total (excluding UK) |
| --- | --- | --- |
| United States | 1,822 | 10% |
| Australia | 1,078 | 6% |
| Canada | 878 | 5% |
| Germany | 864 | 5% |
| France | 678 | 4% |
| Ireland | 624 | 3% |
| China | 619 | 3% |
| Netherlands | 603 | 3% |
| India | 492 | 3% |
| Italy | 484 | 3% |

---

36   Text was passed through the DBpedia Spotlight service (https://github.com/dbpedia-spotlight/dbpedia-spotlight/wiki) to search for named entities and linked to the GeoNames database (www.geonames.org/). Any countries, top-level administrative regions or cities with a population of more than 15,000 were tagged.

37   http://impact.ref.ac.uk/casestudies

38   The codes were assigned using the ISO 3166-1 standard. Currently 249 countries, territories or areas of geographical interest are assigned official codes in ISO 3166-1. A total of 205 codes were assigned from the case studies, which is higher than the official number recognized by the United Nations (as the ISO codes include unrecognized countries) – hence concluding that all countries in the world are geotagged.

39   The UK had the highest number of mentions at 5,308.

**Figure 13:** The global reach of impacts arising from research undertaken in UK HEIs

182 Poland
38 Slovakia
106 Czech Republic
216 Austria

10 Greenland

48 Iceland
5,308 United Kingdom
321 Norway
397 Sweden
281 Denmark
175 Finland
209 Russian Federation

52 Estonia
28 Latvia
44 Lithuania

10 Belarus
63 Ukraine
41
95
15 Moldova
29 Kazakhstan
18 Mongolia

624 Ireland
603 Netherlands
393 Belgium
26 Luxembourg
10 Guernsey

484
56 Romania
24 Serbia
Hungary
19 Uzbekistan
7
18 Kyrgyzstan
7 North Korea
105 South Korea

878 Canada

678 France
361 Switzerland
864 Germany
6 Slovenia
59
50
24 Macedonia (FYROM)
52 Bulgaria
165
28 Armenia Georgia Azerbaijan
5 Turkmenistan
55 Tajikistan

141 Portugal
458 Spain
Andorra
6 Monaco
136 Greece
53 Cyprus
Turkey
27 Syria Lebanon
24 Kuwait
18 Iran
72 Afghanistan

1,822 United States

38 Gibraltar
37 Morocco
21 Malta
22 Tunisia
23 Libya
40 Palestinian Territory
161 Israel
50 Iraq
59 Saudi Arabia
86 Qatar
54 UAE
18 Oman
72 Pakistan
48 Nepal
Bhutan
72 Bangladesh
22 Myanmar
619 China
443 Japan

4 Bermuda

9 Haiti
6 Dominican Republic
5 Puerto Rico
1 British Virgin Islands
2 U.S. Virgin Islands
3 Anguilla

2 Western Sahara
20 Algeria
20 Kosovo
22 Bosnia & Herzegovina
13 Montenegro
21 Egypt
26 Sudan
3 Eritrea
12 Yemen
13 Bahrain
492 India
43 Sri Lanka
86 Taiwan
128 Hong Kong
4 Macao

159 Mexico
1 Bahamas
15 Cuba
1 Caicos Islands
25 Jamaica
6 Montserrat
5 Dominica
3 Cape Verde
2 Mauritania
15 Mali
10 Niger
21 Albania
9 Djibouti
96 Thailand
14 Vietnam
56 Philippines
1 Saipan
1 Guam

1 Cayman Islands
9 Guatemala
10 El Salvador
Belize
11 Panama
4 Saint Lucia
22 Senegal
11 Gambia
4 Barbados
3 Guinea-Bissau
16 Burkina Faso
4 Guinea
81 Nigeria
4 Chad
14 South Sudan
58 Ethiopia
21 Somalia
9 Maldives
155 Singapore
126 Malaysia
11 Brunei
1 Palau

9 Guatemala
9 Honduras
15 Nicaragua
28 Ecuador
15 Venezuela
39 Peru
1 Grenada
8 Trinidad and Tobago
19 Liberia
84 Ghana
19 Cameroon
4 Central African Republic
77 Uganda
108 Kenya
100 Indonesia
3 East Timor
14 Papua New Guinea
8 Solomon Islands
1 American Samoa

56 Colombia
320 Brazil
24 Sierra Leone
8 Ivory Coast
4 Togo
9 Gabon
20 Equatorial Guinea
28 Congo – Brazzaville
24 DR Congo
28 Rwanda Burundi
8 Seychelles
British Indian Ocean Territory
10 Vanuatu

1 Guyana
1 Surinam
1 French Guyana
26 Bolivia
7 Benin
7 Angola
80 Tanzania
45 Malawi
8 Mozambique
28 Comoros
1,078 Australia
1 New Caledonia
3 Fiji

1 Cook Islands
39 Peru
35 Zambia
32 Zimbabwe
10 Madagascar
15 Mauritius
4 Tonga

81 Chile
4 Paraguay
80 Argentina
16 Uruguay
21 Namibia
20 Botswana
8 Swaziland
6 Lesotho
262 South Africa
358 New Zealand

3 Falkland Islands

0 — 1,000+

## Research within UK higher education has had an impact on a wide range of stakeholder groups

Identifying the beneficiaries of impact from the case studies was one of the original aims of the synthetic analysis (Figure 1, page 15). Rather than search the database using our own pre-selected stakeholder groups, we focused on approaches that would let the data themselves generate stakeholder groups that could be considered beneficiaries or users of the impact described in the 'Details of the impact' (Section 4) of the case studies.

In order to do this, we initially used keyword-in-context (KWIC) to generate nouns (or noun groups) that occurred near the words "stakeholders", "beneficiaries" and/or "users", and identified a number of nouns that could be considered beneficiaries. We then conducted a keyword search on a selection of the noun groups identified in this way to see how they appeared in the different panels (Figure 14). As shown in Figure 14, a wide variety of groups appear to have benefited from the research, with the caveat that further reading of these case studies would be needed to see if these nouns are used in the context of a beneficiary, or were simply referred to within the context of the research.

However, what can be seen from this text-mining exercise is that while there are groups potentially benefiting from the case studies relating to their field of research (eg writers benefiting from studies in Panel D, engineers benefiting from studies in Panel B), there are mentions of these potential beneficiaries across all the panels. Although this would have to be validated by further in-depth analysis of the case studies, it suggests that impact created can be cross-disciplinary and reaching a wide range of members of society.

We also examined the two impact topics that mapped directly onto stakeholder groups: 'Children, young people and families'[40] and 'Women, gender and minorities'[41]. The impact wheels for both of these are shown in Figures 15 and 16 (page 44). Not surprisingly, the UOAs 22 and 25 (Social work and social policy and Education, respectively) involved case studies that focused on the stakeholders identified by the topic 'Children, young people and families'. For the topic 'Women, gender and minorities', the UOAs in which this topic was prominent were UOAs 23 (Sociology) and 27 (Area studies).

These illustrative analyses suggest that there are clear mentions of beneficiaries within the corpus of data produced by the case study authors. Some of these may be potentially driving the research itself (and hence appear as a dominant topic), but other interesting stakeholder groups appear across the panels, which provides a rich data source for further analysis. We opted to use a combination of topic modelling and KWIC searches to identify the initial patterns illustrated here; further analyses would be necessary to sift through the results and find the case studies that actually use mentions of these nouns in the context of a beneficiary of the research, as opposed to participants or protagonists in the underpinning research. This method and approach is certainly worth exploring by analysts in future.

---

40  The top nine words related to this topic were: *children child young parent famili imp programm work support*.
    Please note the words have been 'stemmed' where the end of the word is truncated.
41  The top nine words related to this topic were: *women equal gender migrat divers ethnic commun group refuge*.
    Please note the words have been 'stemmed' where the end of the word is truncated.
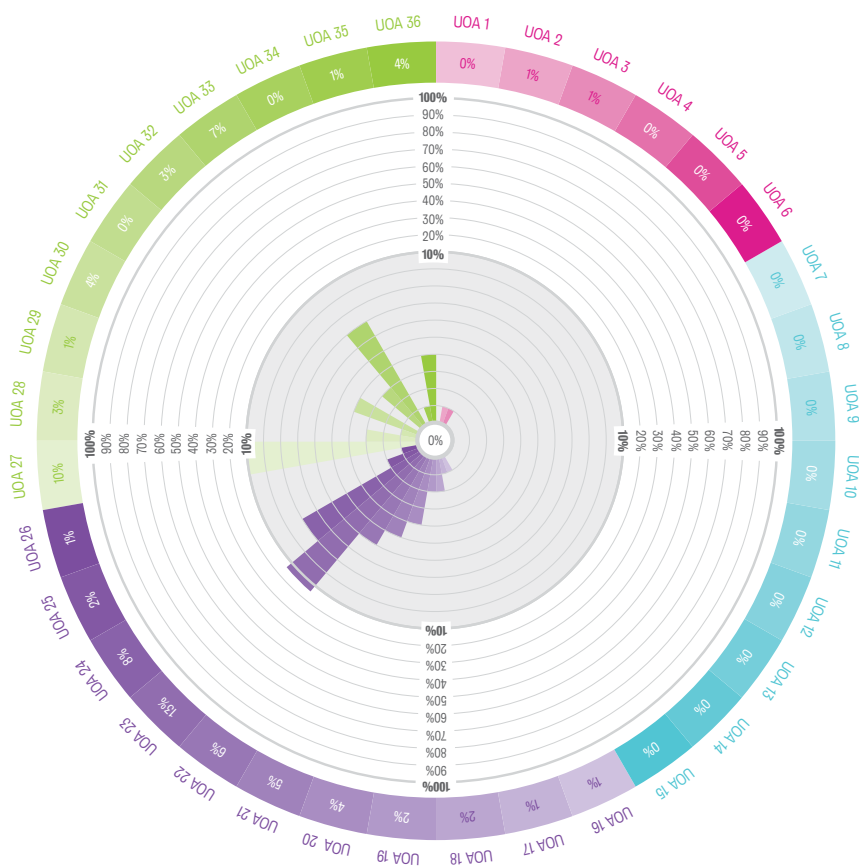
**Figure 14:** Distribution of potential beneficiaries of research found in case studies
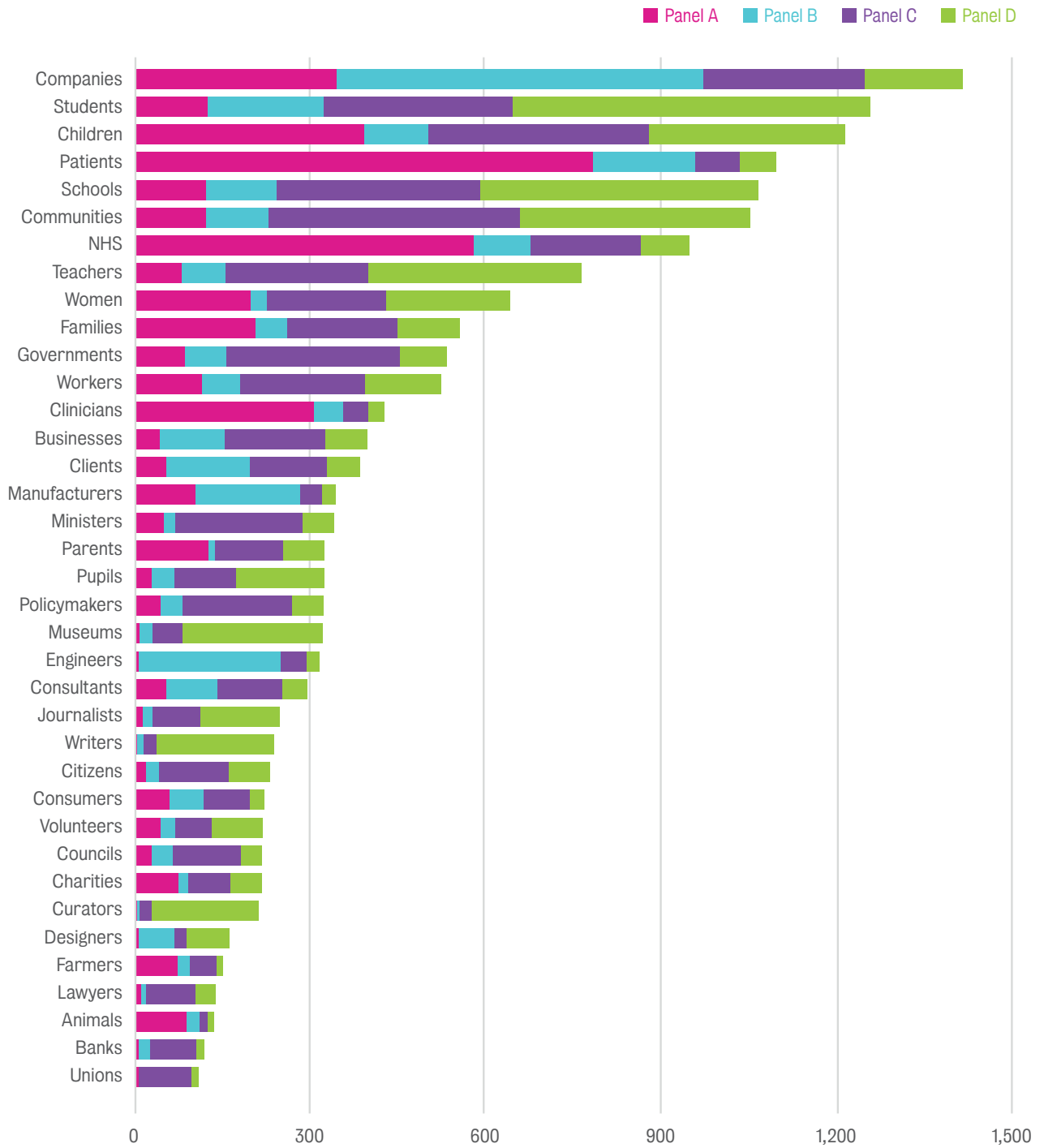
**Figure 15:** Impact wheel ('Children, young people and families', n=198)
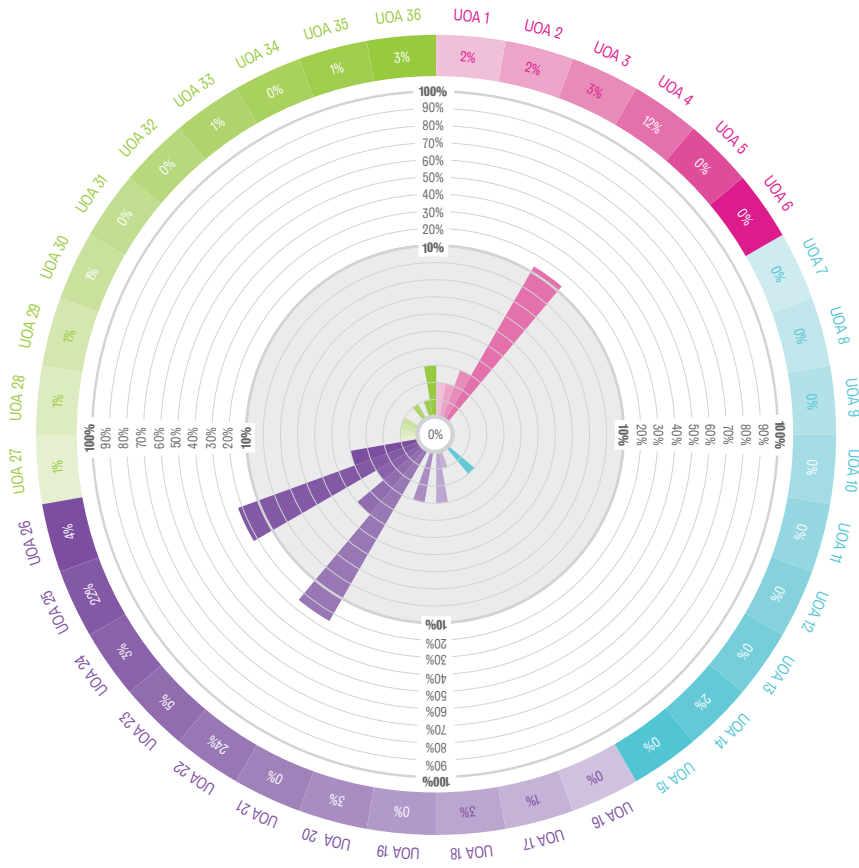


**Figure 16:** Impact wheel ('Women, gender and minorities', n=112)

## The time it takes for research to have an impact varies by discipline

Figure 17 (page 46) illustrates that it takes an estimated three to nine years for research to have an impact on society, and that the speed by which that impact occurs varies by discipline. Each line in Figure 17 shows the time lag between the underpinning research and the broader impact by UOA and the four main panels. We have estimated the time lag as being from the lower quartile to 2010 (the midpoint of when impact was to occur: ie between 1 January 2008 and 31 July 2013, according to the REF criteria).

The dates for the research were extracted from the 'References to the research' (Section 3) of the case studies. This section allowed an 'indicative maximum' of six references, with references typically being to scholarly works or research grants. Dates were identified as being four-digit numerical figures starting with either '19' or '20' (eg 1928, 2010). We then calculated the median date, the lower quartile range (ie at the 25th percentile) and dates at the fifth percentile for each UOA and for the four main panels. We used the fifth percentile as a proxy for the minimum (ie oldest) date to exclude any erroneous information or 'noise' that could have occurred given the approach for data extraction.[42] This approach meant that 97% of the dates were within the time window allowed by the REF guidance (REF, 2011) for 'Underpinning research' (Section 2) which was 1 January 1993 to 31 December 2013.[43] This distribution of dates within the allowable time window for all case studies is illustrated in Figure 18 (page 46).

It is interesting to note that the estimated time lag for Panels A and B is similar, shorter for Panel C and longer for Panel D, giving a relative indication of how fast different research disciplines may have an impact. However, the absolute numbers should be treated with considerable caution, as they are a construct of the assessment structure and not validated by other sources of information on time lags. For example, in biomedical and health research, previous estimates have suggested a significantly longer lag of 15–20 years (Morris et al, 2011; Hanney et al, 2015), although Mansfield (1990) suggests technology lags of around eight years. Indeed, Figure 18 suggests that despite the allowable period for underpinning research stretching back to 1993, the majority of research cited was published since 2008. This may be an artefact of the way that case studies were selected and drafted for submission, and, if so, the reasons for this require further investigation.

---

42  Given the information extraction approach, other numbers that met our criteria of being a 'date' could have been included. In the analysis this is unlikely to be an issue when looking at the median and interquartile ranges, but could erroneously affect the measurements of the minimum and maximum.

43  This could include, exceptionally, up to 1 January 1988 for some UOAs. Only 1.2% of dates fall within the period 1 January 1988–1 January 1993.

**Figure 17:** Estimated time lag by UOA and main panels [44]
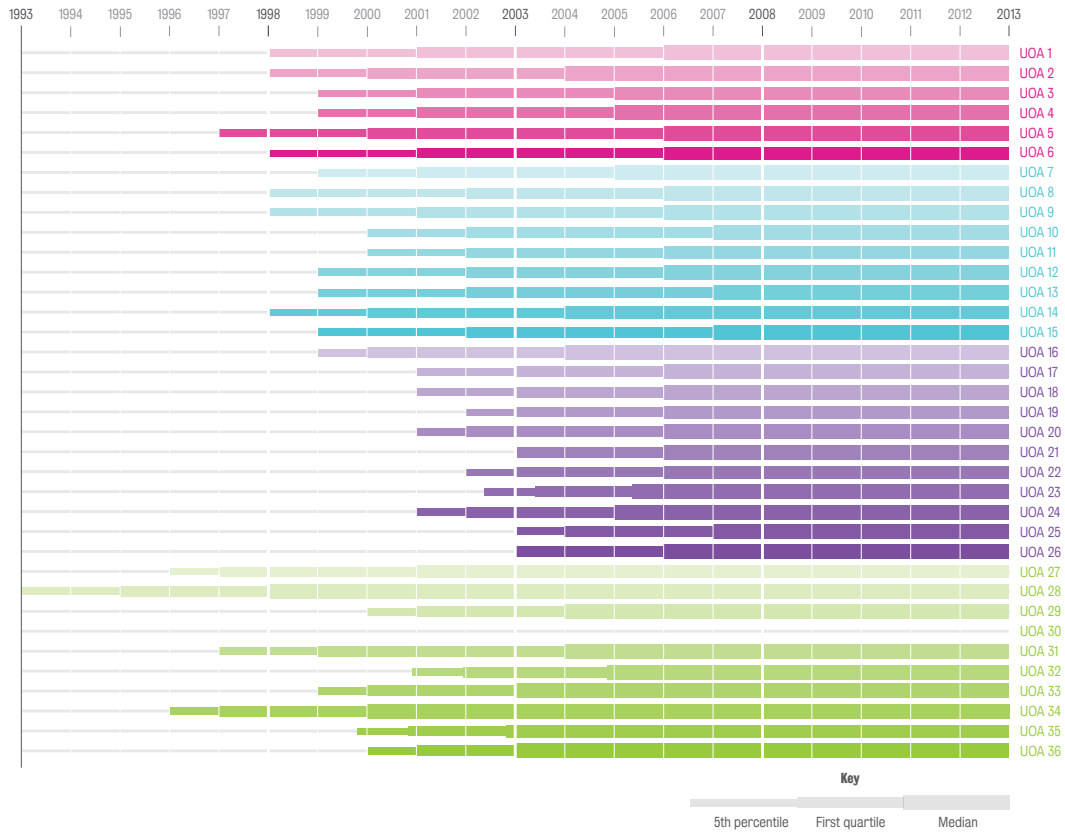


**Figure 18:** Dates of underpinning research
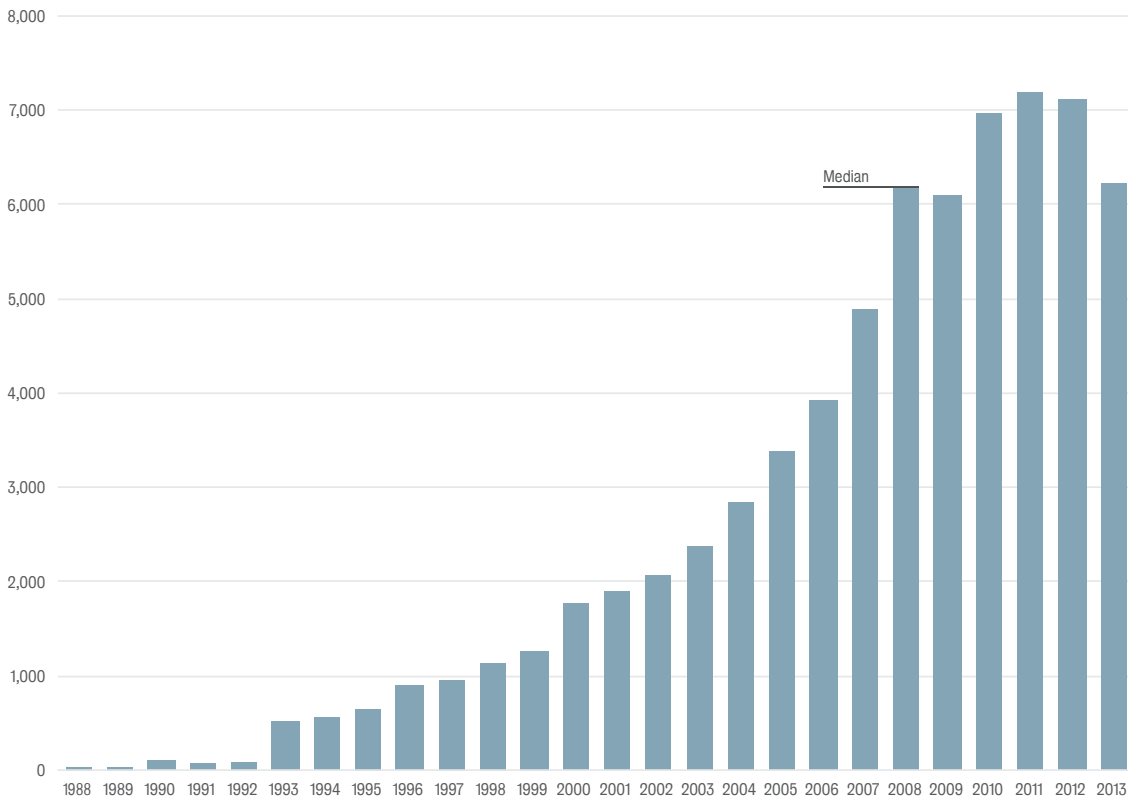


44   Dates for the UOA 30 (History) were removed as they were off the scale (eg the 5th percentile was at 1948).

# 4 | Six 'deep mines' on impact

# 4 | Six 'deep mines' on impact

This chapter presents a more focused and deeper analysis of the data in the impact case studies. This analysis serves to address a number of the key study questions (see Figure 1, page 15) through a series of focused assessments of selected subsets of case studies. The main advantage of doing this is that we could read the case studies and supplement the quantitative text-mining analysis with a more nuanced qualitative assessment. By taking this approach we illustrate both the richness of the case studies, as well as some of the challenges associated with their analysis. We termed this set of analyses 'deep mines' to capture the data-mining analogy, but also to emphasize their depth.

We have adopted this deep mine approach as initially we had envisaged being able to extract quantitative information and to group such information by various indicators, thus enabling us to develop return-on-investment type estimates and to contribute to the ongoing debate on the development of impact metrics.[45] However, early on in our analysis it became apparent that such an approach was not feasible in the time available, for two crucial reasons.

First, there was a very large amount[46] of numerical data in the case studies that were inconsistently used and that would need converting to standard units. For example, financial information was expressed in various currencies, measures and calculations of health gain were inconsistent, and researchers used varying metrics for expressing the scale of interest for cultural impacts, such as exhibition visitors or online statistics.

Second, some of this numerical data was not related to the actual impact, for example, being associated with background information (such as the size of a population) or, crucially, potential impact (eg the number of people who could benefit from a new drug).

We could have developed a number of different deep mines, but selected six based on our own preliminary analysis and the suggestions arising from the stakeholder workshop (Annex B). Four out of the six deep mines are broadly associated with one of the four main panels. The other two are relevant to all the panels. As with the other analyses presented in this report, we anticipate that others will pick up and develop these and other deep mines in the future.

The six deep mine questions were as follows:

1. What is the impact and value of research on clinical practice and health gain?
2. What has been the impact of research on industry in terms of spin-out companies, patents, or licences?
3. What has been the impact of research on public policy and parliamentary debate?
4. What has been the impact of research on film and theatre?
5. What has been the influence of the Wellcome Trust and British Academy?
6. What has been the impact of research on the BRIC (Brazil, Russia, India and China) countries?

---

45  http://hefce.ac.uk/whatwedo/rsrch/howfundr/metrics/ [last accessed 27 February 2015].
46  In 'Details of the impact' (Section 4), we identified c70,000 pieces of numerical information, excluding dates.

Over the following pages we examine these six questions. At the end of each section we provide a short commentary on some of the analytical challenges that we faced in each deep mine, along with some suggestions for further analyses that could be undertaken using similar approaches to those taken in this synthetic analysis.

## Question 1: What is the impact and value of research on clinical practice and health gain?

Unsurprisingly, an important impact mechanism for biomedical and health research is through improvements in health. We found that 416 case studies (6% of the total) were tagged with the topic 'Health care services';[47] 326 (5% of total) were tagged with the topic 'Clinical guidance'[48] and 195 (3% of total) with the topic 'Public health and prevention'.[49] In this deep mine we focused our analysis on the 'Clinical guidance' topic and on an analysis of health gains using a keyword search.

### Contributions to clinical guidance are predominantly present in health-related case studies

There are a number of different mechanisms or pathways by which biomedical and health research translates 'from bench to bedside'. One path is for research to inform clinical guidelines: this was identified as one of the topics in the text-mining and labelled as 'Clinical guidance'. Of the 326 case studies classified as addressing this topic, the vast majority (93%) were from Panel A, although as illustrated in Figure 19 (page 50), there were examples from other panels. We note, in particular, the existence of this topic within 'Sports and exercise' (UOA 26).

The citation of research in a clinical guideline provides an indicator that the research has moved into some formal recommendation for clinical practice in a health system – however, it is important to note that it does not necessarily mean that the practice has been taken up. Specific types of clinical guidelines in England and other parts of the UK are those that are produced by the National Institute for Health and Care Excellence (NICE)[50]. Since 2013 NICE has been established in primary legislation as a non-departmental public body and is responsible for developing guidance and quality standards in health and social care. A keyword search on the word "NICE" identifies 428 case studies; again, the majority of these are in Panel A, with a few exceptions (Figure 20, page 50). We note that this impact wheel is based on absolute numbers of case studies per UOA.

### Approaches to calculate and express value of health gain vary across case studies

One approach for calculating the health gain from an intervention is to use Quality Adjusted Life Years (QALYs), where one QALY is a measure of the health gain from a treatment equivalent to one year of perfect health. Health economists can use a ratio of cost-to-health gain to assess whether a new technology or intervention is cost-effective, and NICE guidance methods currently refer to a threshold range of £20,000 to £30,000 per QALY (NICE, 2008). Thus it is possible to monetize a health gain using the value of a QALY in assessing the cost-effectiveness of an intervention.

Therefore, we did a keyword search on the term "QALY" and identified 25 case studies (all but two from Panel A). We read these case studies and determined that for 14 of them, the QALY was being used to illustrate cost-effectiveness of interventions; but for the remaining 11, researchers evidenced and monetized the actual or potential health gain arising from the underpinning research in the case studies.

For these case studies, we were able to then estimate the potential value of the net total gain to be (very crudely) around £2 billion in the impact period 2008 to 2012.

---

47  The top nine words related to this topic were: *health care servic nh hospit patient nation improv practic*. Please note the words have been 'stemmed' where the end of the word is truncated.

48  The top nine words related to this topic were: *guidelin patient clinic treatment recommend stroke nice risk trial*. Please note the words have been 'stemmed' where the end of the word is truncated.

49  The top nine words related to this topic were: *health screen hiv vaccin women programm recommend prevent nation*. Please note the words have been 'stemmed' where the end of the word is truncated.

50  http:// nice.org.uk/about [last accessed 27 February 2015].

**Figure 19:** Impact wheel ('Clinical guidance', n=326)
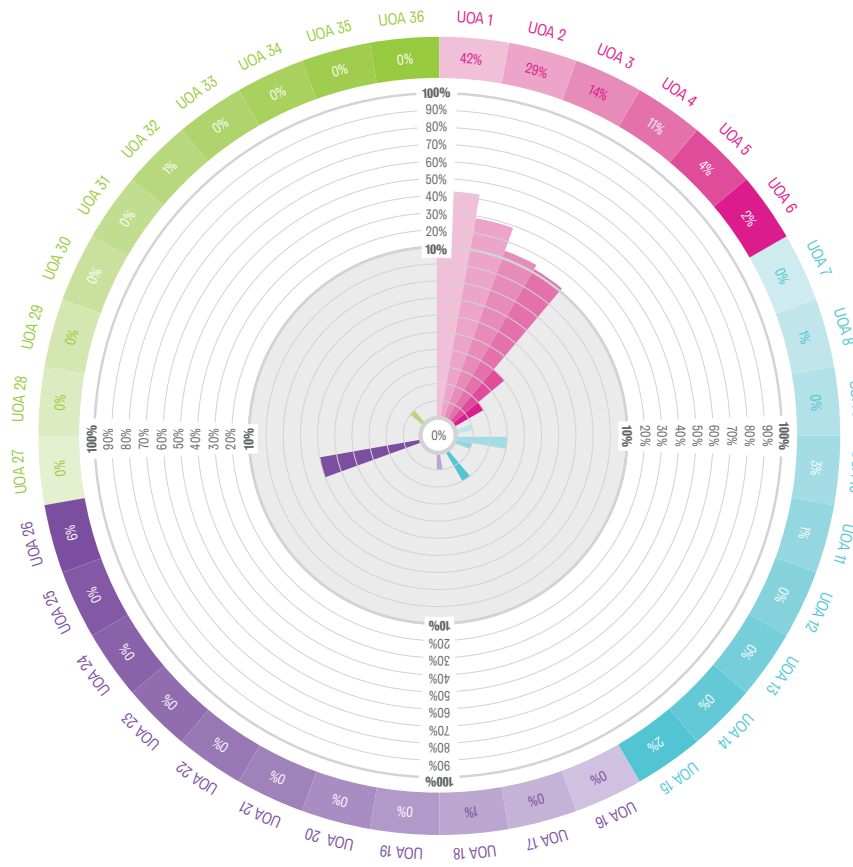


**Figure 20:** Impact wheel for case studies containing keyword "NICE" (n=428)
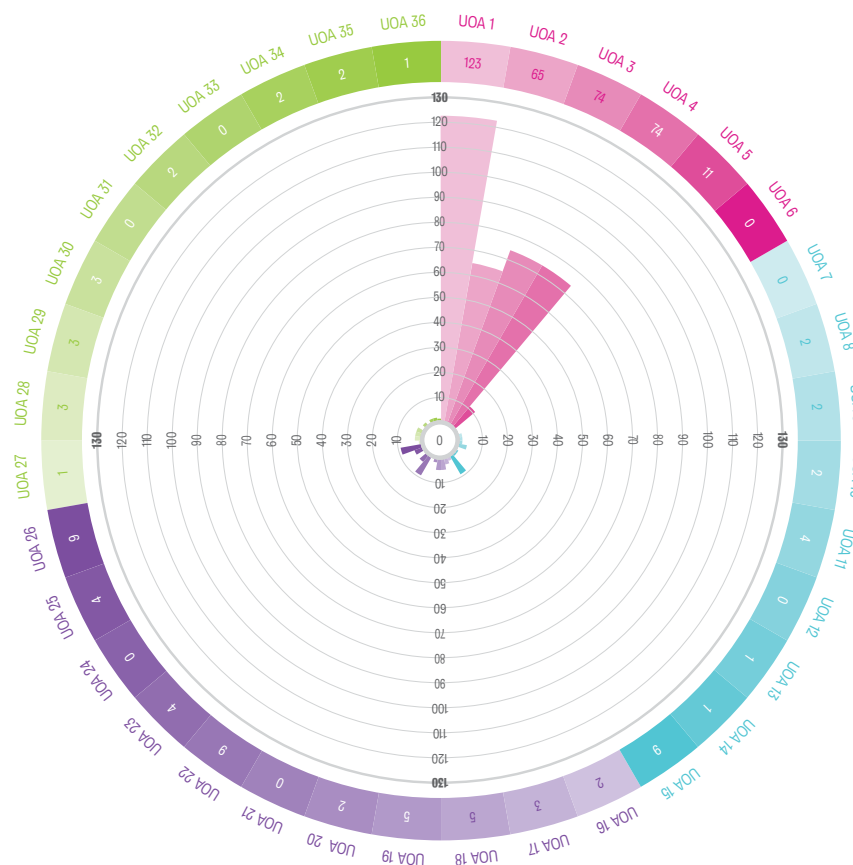
In order to obtain this figure we calculated the monetary benefit from data given in the case studies by the authors, but had to supplement and manipulate it in various ways due to a number of key issues that arose. First, the information presented in these case studies was neither consistent nor standardized. Some case studies presented the QALY gain for an individual patient, others for a patient population. Second, while some provided an estimate of the net monetary benefit, they used different figures for the value of a QALY (ranging from £25,000 to £40,000). Finally, in some of the case studies we had to go to the cited material to find additional information, and in one case we referred to an external source of information that was not cited in the case study (it was published after the case study had been submitted).

### Lessons learned

There are two key lessons arising from this deep mine on the value of health gains arising from research. First, we note that it may be possible, with additional work, to use the case studies as a source for developing return-on-investment-type figures (one of the key questions identified for this synthetic review). With more time it would be feasible to look at the other case studies which, at minimum, cite "NICE" or fit within the 'Clinical guidance' topic and, crucially, with additional work, generate estimates of health gain and monetize these. Not all of those case studies citing "NICE" or fitting within the 'Clinical guidance' topic would be relevant or amenable to such an analysis, illustrating the importance of reading the case studies to supplement the strengths of text-mining techniques in identifying them. Furthermore, in such an exercise one would want to validate the claims and normalize the assumptions and impact models.

Second, a further lesson is one we learned on reading the case studies in more detail and noting the variance in the way that the data are presented in each case study. In order to facilitate future analyses of this kind, or indeed other types of return-on-investment analyses, future iterations of REF could provide guidelines for the way that the evidence in the case studies is presented. For example, where case studies contribute to NICE guidelines, they should be encouraged to provide a total health gain estimate (ie total number of QALYs gained, for a given population, over a specified time period), and use a standard value for a QALY.

## Question 2: What has been the impact of research on industry in terms of spin-outs, patents, or licences?

Understanding the direct economic impact of research remains a common pursuit of various value-for-money and return-on-investment studies. As with the other deep mines, a consistent and systematic search on all returns was not possible because of the way that the information is collated. Our approach involved searching for case studies that would suggest actual or potential commercialization.

Initially we focused on the topics and noted that from the 60 topics used in our analysis, three of them were oriented towards the commercialization of research and a potential direct financial return: 'Business and industry'[51] (n=320, 5% of all case studies), 'Regional innovation and enterprise'[52] (n=97, 1% of all case studies) and 'Technology commercialization'[53] (n=882, 13% of all case studies).

However, on reading a selection of case studies in more depth, we noted that those tagged by the first two topics ('Business and industry' and 'Regional innovation and enterprise') focused mostly on either the process of commercialization and enterprise (eg training in entrepreneurship, manufacturing processes), or the study of such processes (ie the subjects of management and business schools). The case

---

51   The top nine words related to this topic were: *compani busi manag industri product market servic improv sector*. Please note the words have been 'stemmed' where the end of the word is truncated.

52   The top nine words related to this topic were: *innov busi region sme birmingham enterpris support programm univers.* Please note the words have been 'stemmed' where the end of the word is truncated.

53   The top nine words related to this topic were: *technolog compani develop product univers commerci system market industry*. Please note the words have been 'stemmed' where the end of the word is truncated.

studies in the third topic ('Technology commercialization') provided examples of commercial activity such as spin-outs from the research unit, agreement of licences and patent development. Therefore, we focused our analysis even further and used a keyword search to mine for specific keywords that would demonstrate types of commercialization activities: "spin-outs"[54], "patents"[55], and "licences"[56].

## Commercial activity exists and is clearly identified within the case studies

Figure 21 shows that commercial activity in terms of spin-outs, patents and licences is present within the research activities – primarily within Panel B, but scattered among other UOAs as well. The figures show the frequency of case studies within each UOA that mention these three keywords (and their variants). A case study may contain more than one of these words, and we noted that the largest overlap was between licences and patents (n=212), followed by overlaps between patents and spin-outs (n=163), and spin-outs and licences (n=125).

We also investigated how these case studies identified through the keyword searches were tagged by topic modelling. We found that most of them were tagged with the 'Technology commercialization' impact topic, as expected from our earlier brief analysis, followed closely by 'Pharmaceuticals'[57] (Figure 22, page 54). However, the other topics imply a range of applications and sectors for commercial activity within these case studies.

## The degree of commercialization reached can range from investments and partnership opportunities to actual monetary returns

Of the case studies mentioning spin-outs, patents or licences, a total of 457 case studies were tagged with the impact topic 'Technology commercialization'. We read a random sample of 200 of these case studies, and noted that the content provided a rich source of narratives describing the routes to commercialization and the varying degree to which each study reached the stage of producing actual revenue or monetary benefit. We categorized these 200 case studies and found that the claims for commercialization centred on a number of themes (not mutually exclusive):

- Revenue created from product sales (n=64)
- Secured industrial investment (n=62)
- Secured industrial partner (no financial information given) (n=38)
- Spin-out company created staff employment (n=18)

In reading the case studies we noted the potential for demonstrating some form of quantifiable information, such as total investments per year, sales and revenue figures and total number of staff in employment. However, as with the QALYs example in the first deep mine, such crude estimates could be made only with additional calculations, references to external data sources and assumptions, which would not have been possible for a sample of this size in the given time.

## Lessons learned

This deep mine particularly emphasizes the importance of combining different search and text-mining methods for analysing the case studies. In this case, the keyword search worked well in identifying specific types of commercial activity (eg spin-outs and licences). Future analyses would be enabled and better automated by requesting specific economic outputs (such as figures of revenue generated, sales or, indeed, the number of staff employed) in the guidance provided on how to present impact information in the case studies.

---

54  We included the plural and singular of this term and variants such as "spin-out", "spinout" and "spin out" in our search.
55  We included the plural and singular forms of "patent" in our search.
56  We included the plural and singular forms and variants of "licence" in our search.
57  The top nine words related to this topic were: *drug develop pharmaceut trial compani clinic phase discoveri industri.* Please note the words have been 'stemmed' where the end of the word is truncated.

**Figure 21:** Impact wheels for keywords "spin-outs" (n=344), "patents" (n=570) and "licences" (n=654)



Spin-outs



Patents



Licences

**Figure 22** Impact topics of the case studies containing keywords "spin-outs", "patents" and "licences"



Spin-outs

Patents

Licences
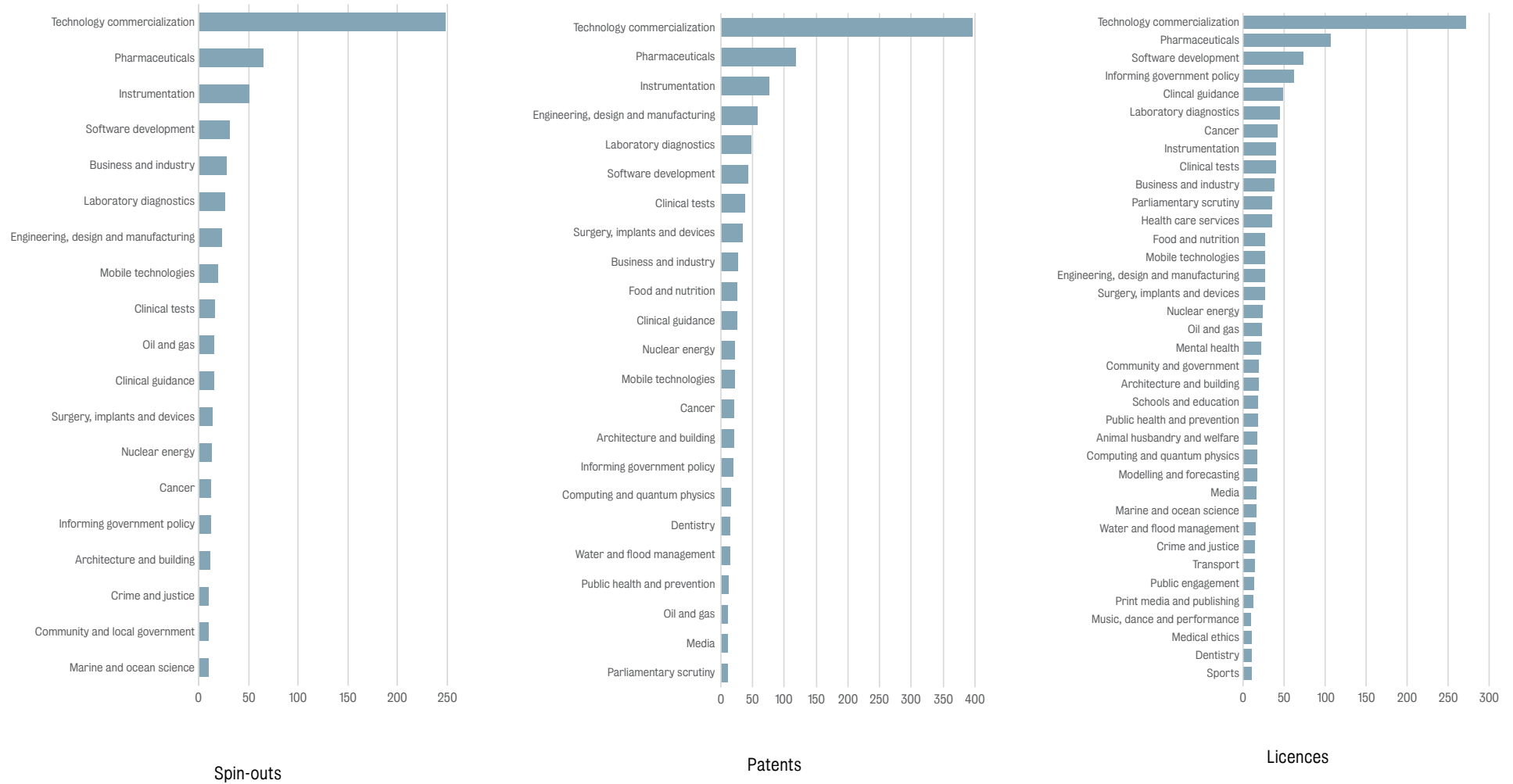
Note: values under n=10 not shown.

## Question 3: What has been the impact of research on public policy and parliamentary debate?

The challenge of using research effectively in the formulation of public policy is an enduring theme for those interested in evidence-based policymaking. Often, it is suggested that there is a gap between research and policy. A recent systematic review (Oliver et al, 2014) of the barriers to, and facilitators of, the use of evidence by policymakers suggests that little has changed over the past decade: timely access to good-quality and relevant research evidence, collaborations between researchers and policymakers, relationship and skill-building with policymakers are the factors that continue to be the most important in influencing the use of evidence.

Therefore, for our third deep mine we decided to explore the impact of academic research on public policy and political scrutiny in the REF impact case studies. We based the analysis on two of the topics that were generated in the text-mining – 'Informing government policy'[58] and 'Parliamentary scrutiny'[59] – and specifically explored academic contribution to Parliamentary Select Committees.

### The most frequent type of impacts were on public policy and parliamentary debate

The impact of research on public policy was cited often as an exemplar in preparations for REF. For example, all four of the main panels cited public policy impacts in the *Panel Criteria and Working Methods* (REF, 2012). Thus it may not be too surprising that the word "policy" commonly occurred in 'Details of the impact' (Section 4) in the case studies: it was mentioned at least once in 3,206 case studies (as detailed in Table 5), and is observed in the word cloud (Figure 6, page 29 and front cover). Furthermore, the two largest topics identified through topic modelling were 'Informing government policy' (n=1233, 20% of case studies) and 'Parliamentary scrutiny' (n=983, 17% of case studies), as illustrated in Figures 23 and 24 (page 56). Despite the two topics being those most commonly occurring together in the case studies (Figure 7, page 31), the size of the overlap was small between the two topics: only 192 case studies had both in their 'top 3' ranked topics. 'Parliamentary scrutiny' is more frequent in Panel C UOAs, and captures a greater degree of political engagement, while 'Informing government policy' is evident in Panel A and Panel B units.

**Table 5:** Selected keywords related to policy[60]

| Keyword | Number of case studies with key word |
|---|---|
| Policy | 3,206 |
| Government | 2,605 |
| Parliament | 610 |
| Lord | 268 |
| Select Committee | 265 |
| MP | 264 |

---

58  The top nine words related to this topic were: *develop polici nation plan govern inform work strategi assess.* Please note the words have been 'stemmed' where the end of the word is truncated.

59  The top nine words related to this topic were: *polici govern report public uk committe debat evid commiss.* Please note the words have been 'stemmed' where the end of the word is truncated.

60  These figures are an illustration of the distribution of these words only, and do not include plural forms (this also more closely resembles the way that these words appear in the word cloud).

**Figure 23:** Impact wheel ('Informing government policy', n=1,233)
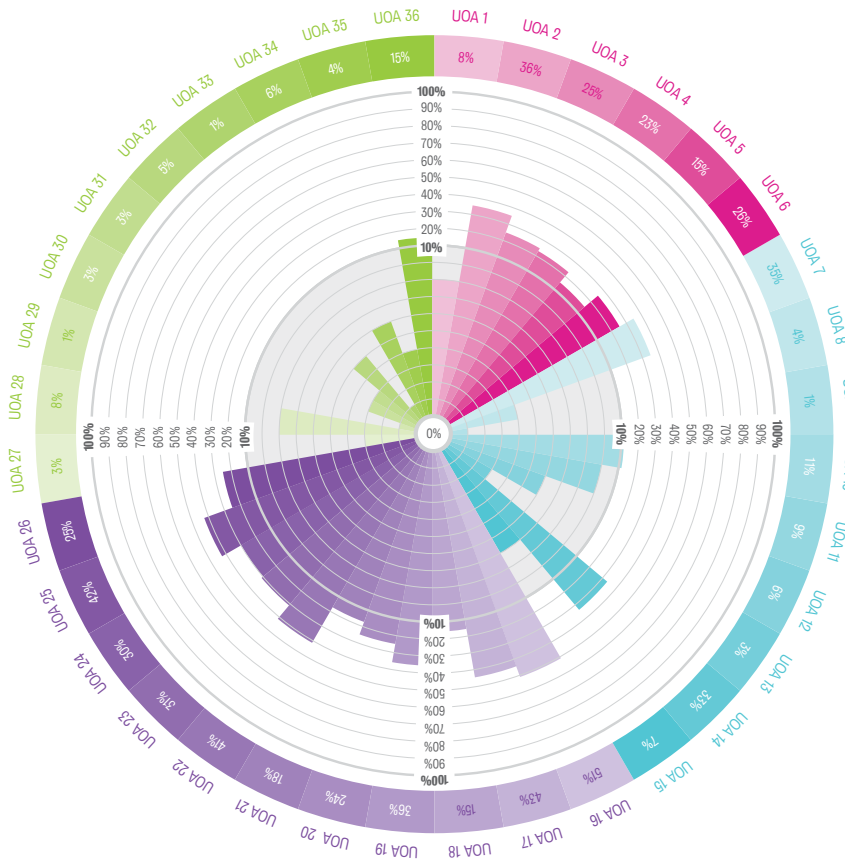


**Figure 24:** Impact wheel ('Parliamentary scrutiny', n=983)
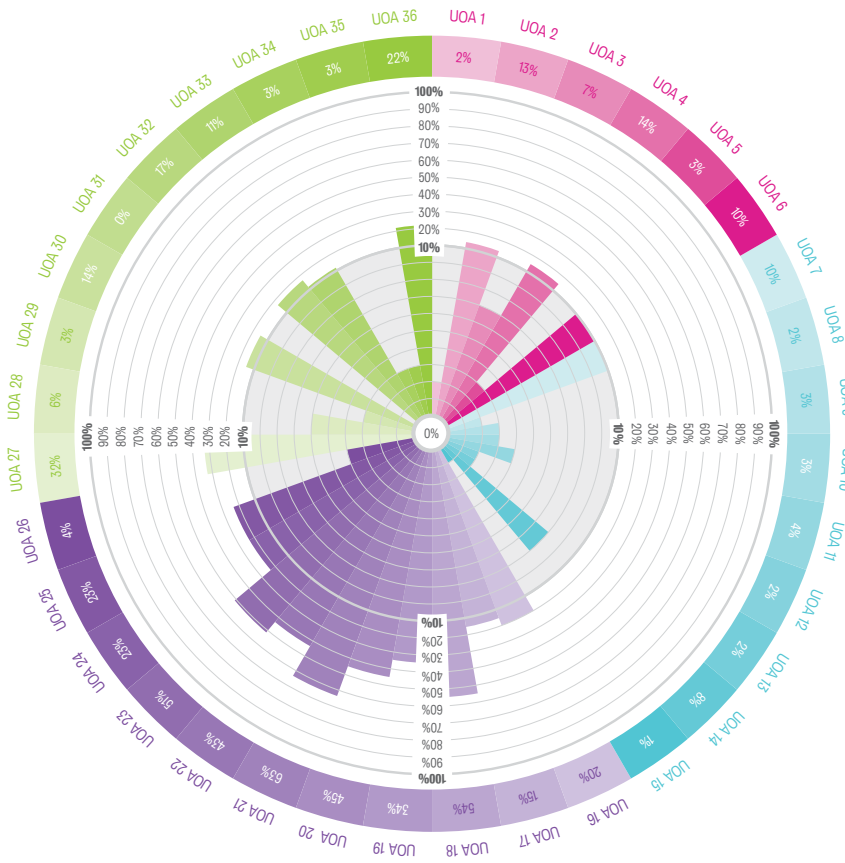
### Research supports parliamentary scrutiny in different ways

In order to explore how academic research contributes to parliamentary debate, we further analysed and read the 265 case studies that mentioned "Select Committee". We chose this because select committees often take evidence from interested parties, and this provides opportunities for academics to engage with and support parliamentary debate. We note the term "Select Committee" is used mainly for parliamentary scrutiny in Westminster, and thus excludes analysis in the devolved administrations.

   Of the 265 mentions of "Select Committee" in the case studies, 185 (70%) occurred in the 'Parliamentary scrutiny' topic; 67 (25%) in 'Informing government policy' topic; and 42 (or 16%) in neither topic. (There were 29 occurrences in both topics.) As illustrated in Figure 25, 70% of the mentions of "Select Committee" occurred in Panel C UOAs, with the remaining 30% being distributed evenly across the other three main panels.

**Figure 25:** Impact wheel for keyword "Select Committee" (n=265)
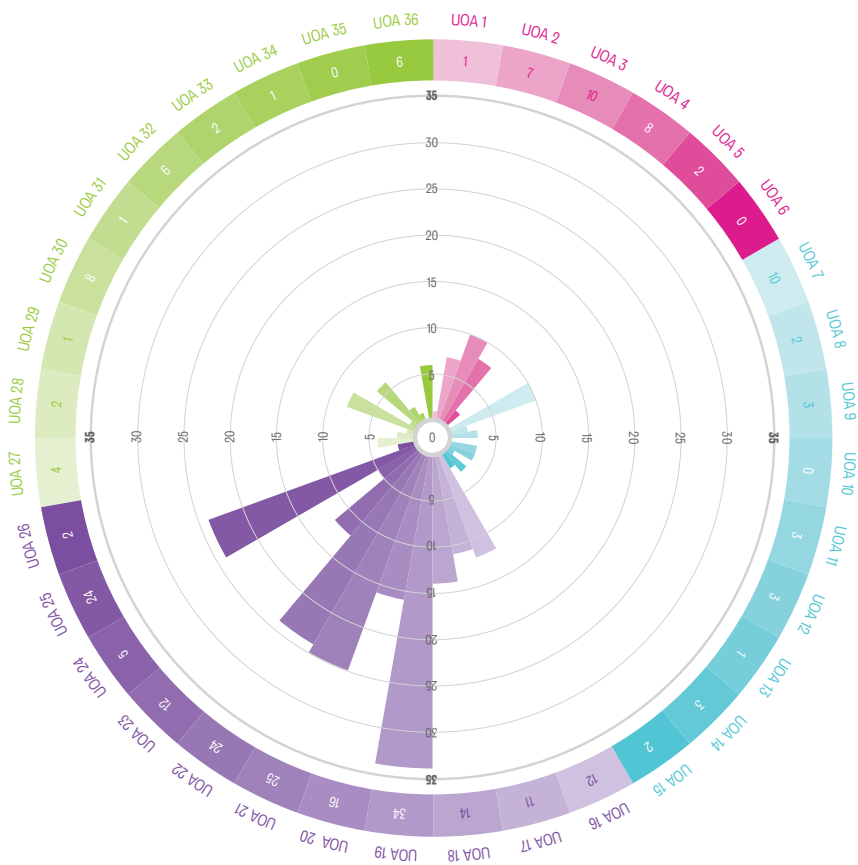
In reading the case studies we initially had hoped to see which select committees were more likely to engage academic input. However, this was not possible in the time available for the project, due to the inconsistent way in which the committees were described (sometimes incorrect or abbreviated titles were used, and each mention would need to be systematically validated against a list of historical and contemporary select committees).

Nonetheless, we did manage to examine the nature of academic input (as summarized in Figure 26). In 30 instances the academics cited in the case studies were appointed as specialist advisors to the committee, there were 101 examples of academics giving oral evidence to a hearing, and a further 23 written submissions. In addition to these direct contributions, there were 33 examples of third parties using research to support their arguments in making a case to a committee. Finally, there were 32 incidences where the select committee directly cited the research noted in the case study.

**Figure 26:** The nature of academic input to Select Committees
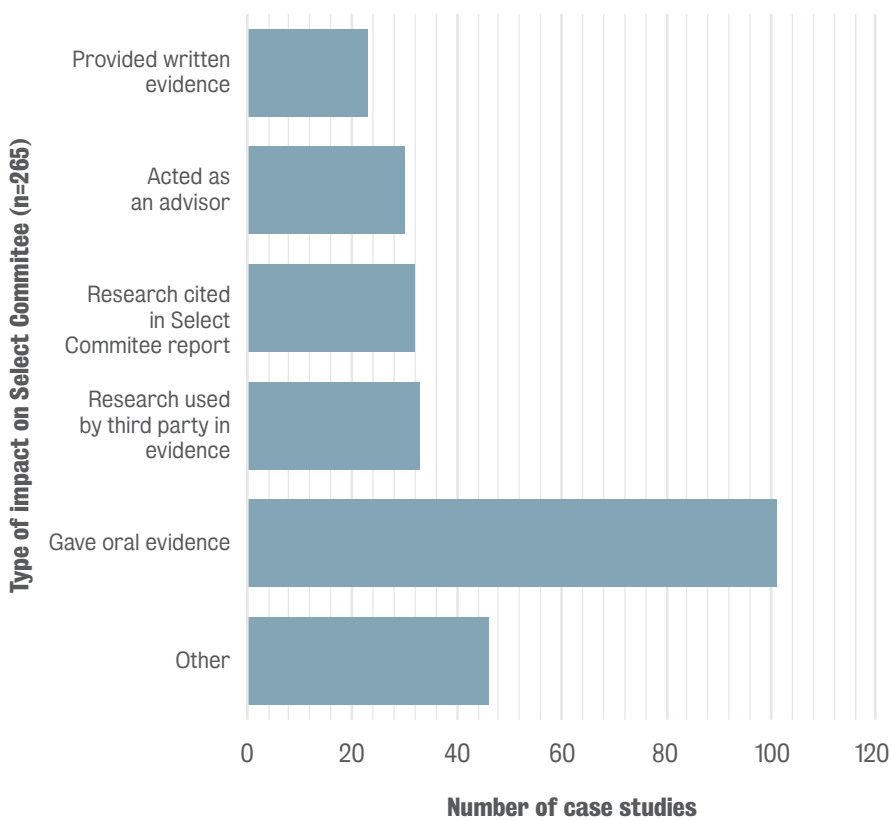
### Lessons learned

From this initial excavation of the case studies describing how research supports policymaking, it would seem there is a rich vein of information that can be further tapped in the future. This may help contribute to ongoing research to understand and improve the way that research is used in policy and decision-making. One lesson for future iterations of REF would be to provide a standardized list of select committee names that case study authors could use, as this would aid subsequent analyses.

## Question 4: What has been the impact of research on film and theatre?

Describing the value and benefit of research conducted in the arts and humanities can be challenging (Levitt et al, 2010). As suggested in the REF guidelines (REF, 2011), researchers in these fields have provided some quantitative data to support their impact case studies, such as number of visitors to exhibitions and web downloads. However, it is on reading these case studies in detail that the pathways and depth of the contributions from the arts and humanities can be more directly understood.

In this deep mine we focused on case studies tagged under the topic 'Film and theatre'[61] (see Figure 27, page 60). We selected this topic because we aimed to find one that both would be concentrated within a panel (which is more commonly the case for the case studies in Panel D), and relatively distributed across the UOAs within that panel. What we found most interesting on reading the case studies were the types of beneficiaries mentioned, as well as the process or pathways to creating impact.

### Research involving film and theatre has benefitted a wide range of people

Given the potential reach of the area of film and theatre, these case studies often highlighted more than one beneficiary beyond the general public and audiences, although the public constituted more than two-thirds of the mentions of beneficiaries in the case studies we read. Of the 139 case studies in this topic, we found impacts on those in schools and education (n=53 studies, 38%); on future film and theatre practitioners (n=84, 60%); and a smaller proportion (n=31, 22%) of impacts on 'other' stakeholders, such as policymakers, offenders and ex-offenders, local communities and hospitals.[62] As expected, many of the case studies also noted the influence of research on future practitioners and on advancing methods used in film and theatre.

Although these are very specific target groups, it can only be assumed that a wider range of people had access to the results of the work described in these case studies, given the frequency of web downloads, social media and media mentions cited throughout this selection. Often the authors of the case studies did not distinguish between direct and indirect beneficiaries. In addition, often it was unclear what effect press coverage had other than the common citation of increased public awareness.

### Impact in the arts was achieved through participatory research and media coverage

There were three distinct stages in the research pathways in these case studies that contributed to their impact story:

1. Impact as a result of the research process
2. Impact as an outcome of the research, and
3. Impact through final coverage of the research, often mainly through the media.

---

61  The top nine words related to this topic were: *film theatr perform plai audienc product festiv screen director*. Please note the words have been 'stemmed' where the end of the word is truncated

62  These categories were non-mutually exclusive and therefore the total within each is greater than the total in the sample (n=139).

Around one-fifth of the case studies involved some form of participatory process within their research, which as the authors noted, resulted in an impact on the research participants. Examples include:

- The involvement of students and members of the local public in performances

- Individuals with sickle-cell disease highlighting their own experience on film

- Connecting festival practitioners to systematically analyse their work and learn from shared experience to improve operations, and

- Creating a network for public debate on Shakespeare.

These case studies mostly report on the influence of this participatory process either to raise awareness and understanding of a specific issue, or to encourage debate. Other wider outcomes from the case studies include increasing the availability of previously unavailable historical resources in a new archive or a newly-translated text, and consultancy or advisory roles created for the academics involved in the research. The types of media coverage varied from local and regional press coverage to national and international coverage, including online and social media presence.

**Figure 27:** Impact Wheel ('Film and theatre', n = 139)
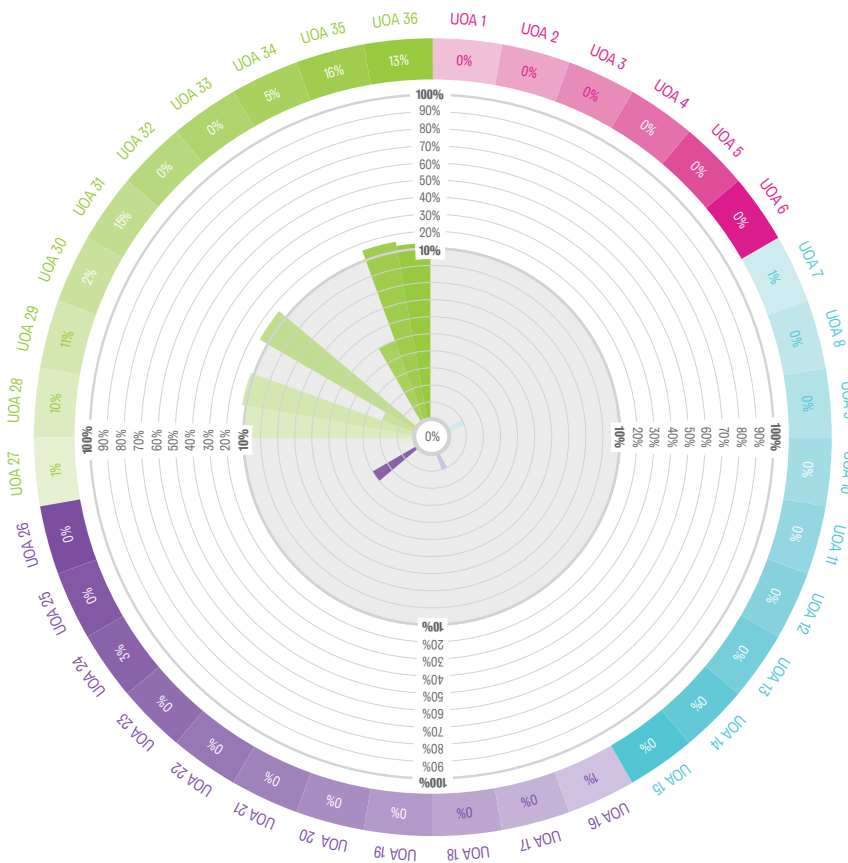
While quantitative metrics could be mined from this data by future analysts (eg the number of museum visitors, web downloads), we focused this deep mine on a very specific set of case studies to understand the ways in which impact is described in a particular topic area. The topic modelling approach was particularly helpful for selecting the case studies to read, or to select particular areas of focus. Qualitative assessment of the case studies showed the different ways in which impact was achieved (eg, through participatory research) and the range of people who benefited from the research, reinforcing the value of keeping this approach when extracting information on pathways or process of impact.

## Question 5: What has been the influence of the British Academy and Wellcome Trust?

One of the key questions raised in the request for proposal (RfP) and at the workshop was the impact of specific research funders (see Annex B). We were unable to add metadata for research funders because not all case studies provide funding information, and where this is provided, it is not in a consistent form. Therefore, we decided to devote one of the deep mines to this issue, and focused on two funders: the British Academy and Wellcome Trust. We chose these two institutions for a number of pragmatic reasons.

First, they are operating in different areas: the British Academy largely in the social sciences and humanities; the Wellcome Trust in the biomedical sciences, but with a portfolio of activity in the humanities. Second, the Wellcome Trust is a very large research funder (spending around £600 million a year), while the British Academy is relatively small in comparison (£15 million). Third, they are relatively established institutions with limited name changes, mergers or suchlike over the period for underpinning research (1993–2013).

We searched 'References to the research' (Section 3) and 'Details of the impact' (Section 4) of the case studies for mentions of each institution, using the terms "British Academy" and "Wellcome Trust"[63] for our searches (Table 6). We then read the 180 case studies with mentions in the 'Details of the impact' (Section 4) of the case studies. It is interesting to note in Table 6 that more case studies cite the British Academy than the Wellcome Trust in 'Reference to the research' (Section 3) – which is somewhat implausible, given the difference in size of the two institutions. This most probably illustrates the inconsistent nature of acknowledging the funders of underpinning research across the corpus of case studies, and perhaps potential cultural differences between disciplines.[64]

**Table 6:** Summary of case studies citing British Academy or Wellcome Trust

| Keyword | Cited in Section 3 ('References to research') | Cited in Section 4 ('Details of impact') | Cited in Section 3 or 4 |
|---|---|---|---|
| British Academy | 260 | 75 | 307 |
| Wellcome Trust | 245 | 105 | 311 |

---

63 We looked at "Wellcome Foundation" (which had no mentions) and "Wellcome" on its own (which had 11 mentions in Sections 4 and five mentions in Section 3), but this was picking up Wellcome Trust-funded centres.

64 It was notable, for example, that a number of the British Academy citations were for small grants between £5,000 and £10,000.

### The British Academy and Wellcome Trust support a diverse range of research

Unsurprisingly, and as illustrated in Figure 28, the research cited in the case studies as underpinning societal impact is focused on the social sciences for the British Academy, and the biomedical sciences for the Wellcome Trust. The 260 case studies that cited the British Academy in Section 3 of the case studies came from 23 of the 36 UOAs. Similarly, for the Wellcome Trust, the 245 case studies came from 30 units.

### The Wellcome Trust and British Academy have multiple and diverse range of impacts

Figure 29 (page 64) illustrates that that both institutions have a diverse range of impacts, often in areas that are outside their core areas of interest. These figures show the frequency with which case studies citing the two institutions in the 'Details of the impact' (Section 4) are included in the different impact topics.

As noted earlier, it is possible that a case study may occur in up to three topics, hence the total classified between the topics exceeds the number of case studies. For the British Academy the biggest impact is on 'Parliamentary scrutiny', followed by 'Media'; while for the Wellcome Trust the biggest impact is on 'Media' followed by 'Technology commercialization'. Perhaps the most impressive observation is the diversity of topics where the two funders are making a contribution: for the British Academy, 32 out of the possible 60 topics are listed, and for the Wellcome Trust, 40 topics.

### The British Academy and Wellcome Trust enable research impact

One of the key observations from reading the case studies was that both institutions not only fund research, but also actively help to enable impact. We read the 75 case studies that cited the British Academy in 'Details of the impact' (Section 4), and the 105 that cited the Wellcome Trust. A number of these case studies were excluded from further analysis as the citation was referring to research funding from either institution, which under the REF guidance is not a form of impact.

Additionally, there were a number of false results picked up by the search term that were excluded, such as the British Academy of Management or a Wellcome Trust building where the research was located. This left 46 and 59 case studies citing the British Academy and Wellcome Trust respectively. We assigned these 105 case studies to six mutually exclusive groups according to the way in which impact had been enabled, as illustrated in Figure 30 (page 66). What is clear from this analysis is that both institutions provide a venue and voice for research translation to occur. That is, they have provided physical locations for various events and helped to amplify the research through speaker opportunities, press releases or citing the research in institutional publications.

### Lessons learned

It would seem that one of the major limitations of this deep mine is that it depends on case study authors citing the funding institution in the references to research. While this was allowable in the REF guidance, it was not a requirement. Indeed, it seems implausible that had this been required, only 245 case studies would have cited the Wellcome Trust in this way, given the size of its funding portfolio. Thus, as will be noted in Chapter 5, a lesson for future iterations of REF impact assessment would be to capture formally and in a consistent manner the institution or institutions that funded the underpinning research. A related issue is that the search for the British Academy picked up seven false positives, which emphasizes the dangers of relying solely on text searching.

**Figure 28:** Impact wheel for keywords "British Academy" and "Wellcome Trust" in 'Reference to research' (Section 3) of the case studies
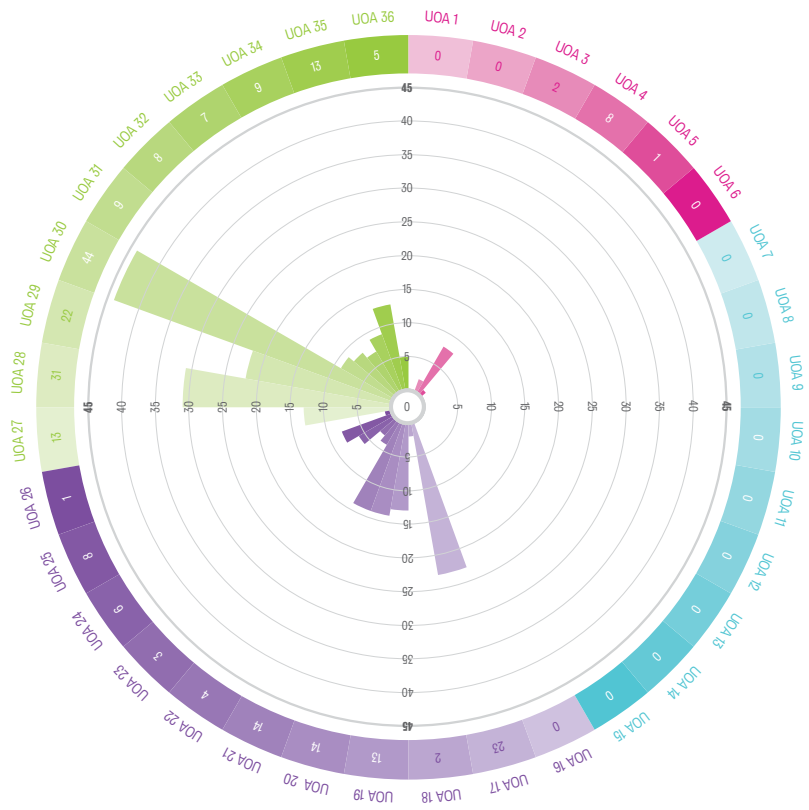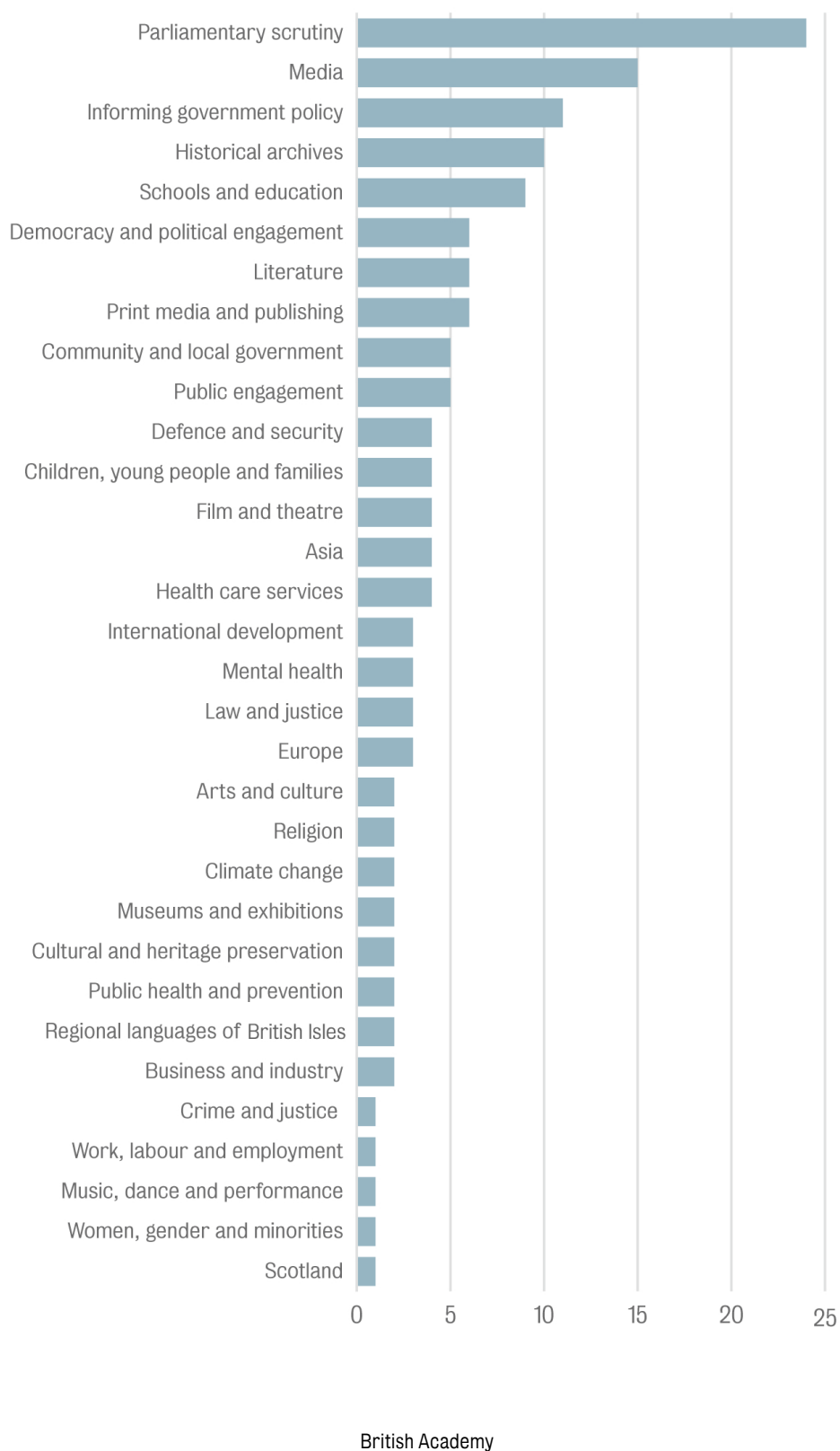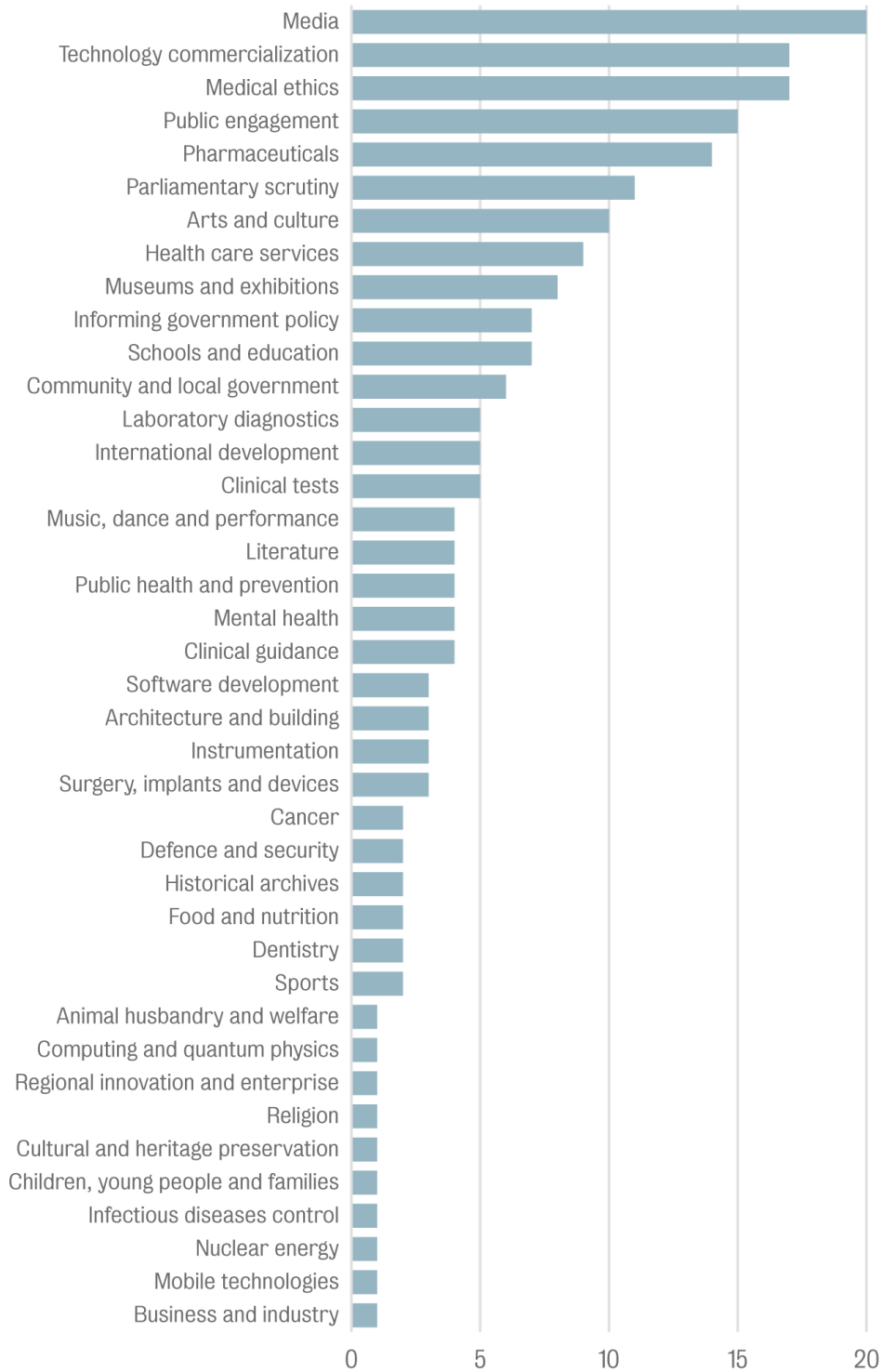


British Academy
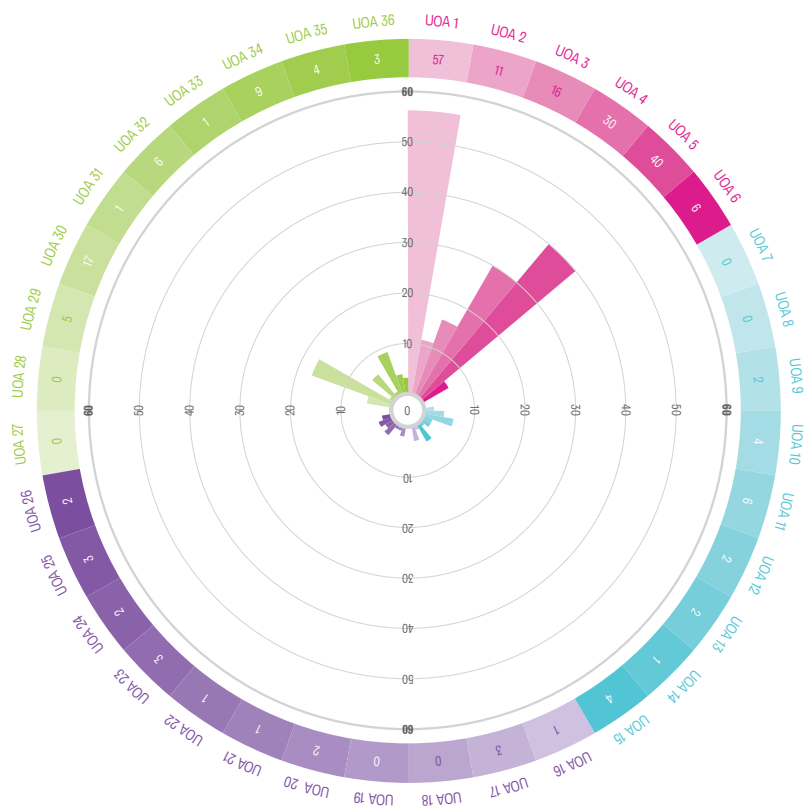


Wellcome Trust

**Figure 29a:** Comparison of the nature of impact by two research funders: the British Academy and the Wellcome Trust
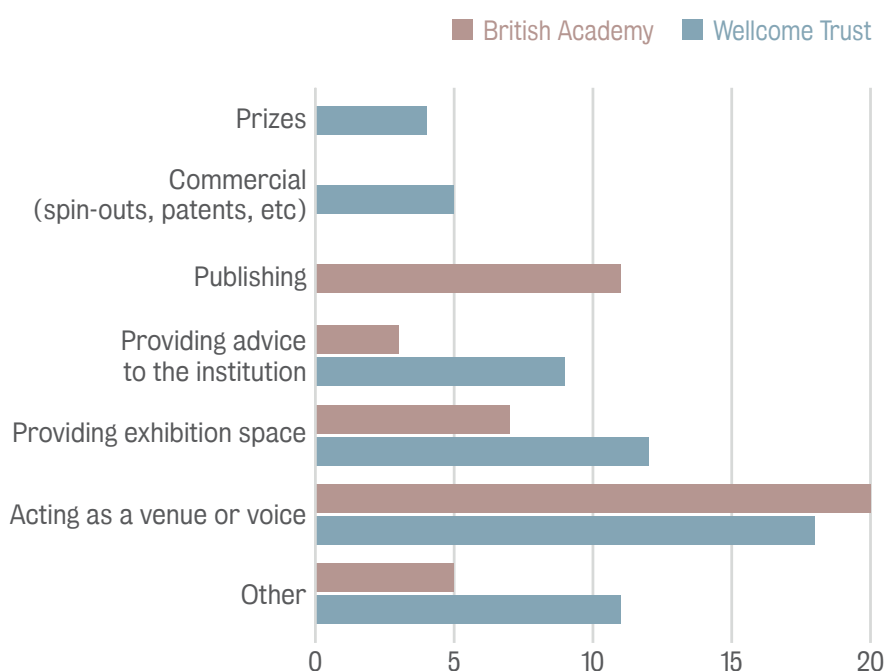
British Academy

Wellcome Trust

**Figure 30:** Different ways that the British Academy and Wellcome Trust have enabled impact



## Question 6: What has been the impact of research on the BRIC countries?

Figure 13 (page 41) in Chapter 3 shows that the research from UK HEIs undertaken in the past 20 years has had a global impact. In this last deep mine we investigated the impacts of research on the BRIC countries (grouped for their similar stage of newly-advanced economic development) (O'Neill, 2001). Of the 17,932 non-UK geotags, 1,640 (or c9%) were assigned to the BRIC countries: Brazil (n=320), Russia (n=209), India (n=492) and China (n=619).

### UK Higher Education Institutions research has had an impact on BRIC countries

As shown in Figure 31 (page 68), the impacts in BRIC countries are distributed across all UOAs, suggesting that their impacts arise from a wide range of research fields. Brazil stands out as the countries mostly mentioned by a relatively large proportion of cases across all the UOAs. However, in the case of China in particular, there is a greater relative concentration of impacts within Panels A and B, compared to Panels C and D, which could be a reflection of the clinical and scientific advances and collaborations conducted in partnership.

## Impact on BRIC countries was varied and comprehensive, and sometimes strategic

We selected a random sample of 50 case studies from each of the four BRIC countries (n=200 in total). As mentioned previously, we found both impacts that were described as strategic (ie collaboration with an international partner organization was created to conduct the research) and incidental (ie positive but not specified intended outcomes in the country as a result of the research). Some case studies (n=7) noted the creation of spin-out companies and agreement of licences. Other key types of impact included:

- Informing government policy in that country

- The creation of new technologies to develop in that country

- Creation of online resources for wide public use

- Facilitating international collaboration, often between academics, and

- Creation of resources and training for teaching.

We also focused our analysis on the types of beneficiaries mentioned in this random sample of case studies in the BRIC countries. These beneficiaries included people affiliated to schools and education institutions (n=55), policymakers (n=11), offenders and ex-offenders (n=3), hospitals or health-related beneficiaries (n=3) and the local community (n=8).

We found on closer analysis that in those studies in which the impact was strategic and focused on a specific country, a greater number of observed benefits were noted. For example, in Brazil, a consortium formed with UK HEIs led to the development of legal frameworks for widening participation and equal opportunities. Research on regulatory mechanisms to overcome economic and legal barriers to the deployment of renewable energy in Russia led to the World Bank Group taking action to enable developments in Russia through the EU Renewable Energy Directive (with the academic as legal advisor to the World Bank Group Russia Renewable Energy Program). In India, mathematical modelling of the HIV/AIDS epidemic had an influence on policy and decision-making; and in China, research into the delivery of unemployment insurance systems to individuals helped 25,000 direct beneficiaries in six Chinese cities receive more reliable services.

## Lessons learned

Conducting analyses on specific countries is possible using geotagging tools; this provides a clear selection of case studies which then can be analysed further. Given that we primarily focused on geotagging within 'Details of the impact' (Section 4), the large number of studies that were extracted only by being tagged with BRIC countries was notable. Enough material can be extracted to then conduct cross-analyses on any of the previous topics on each of these countries (eg, spin-outs within the BRIC countries, or cultural impact cited in BRIC locations) and we recommend that analysts try out some of these approaches in future.

**Figure 31:** Impact wheels for Brazil (n=320), Russia (n=209), India (n=492) and China (n=619)
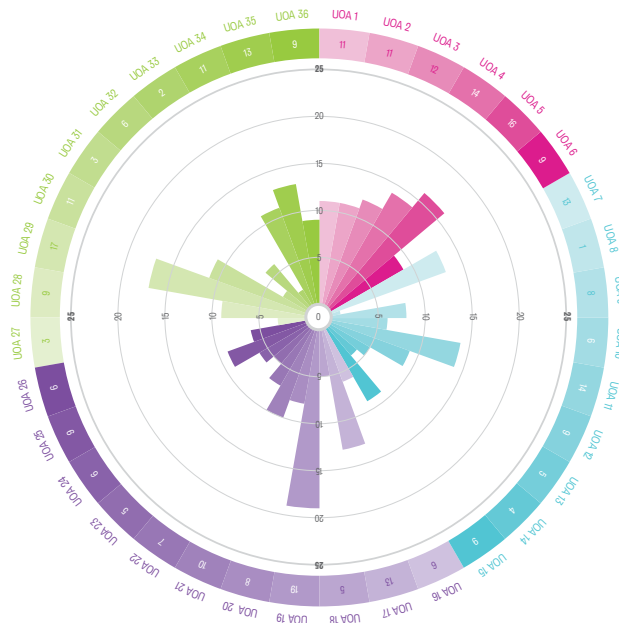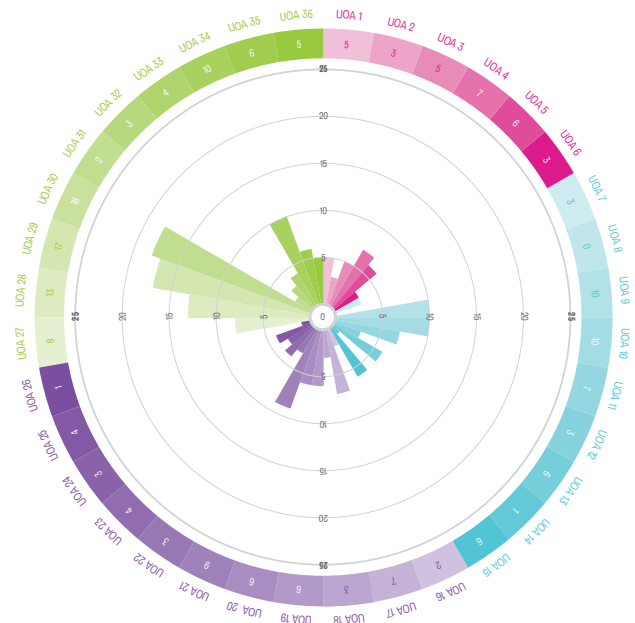
# 5 | Concluding reflections

# 5 | Concluding reflections

This report presents the first synthetic analysis of the 6,679 non-redacted impact case studies that were submitted to the Research Excellence Framework (REF) 2014. We focused on assessing the nature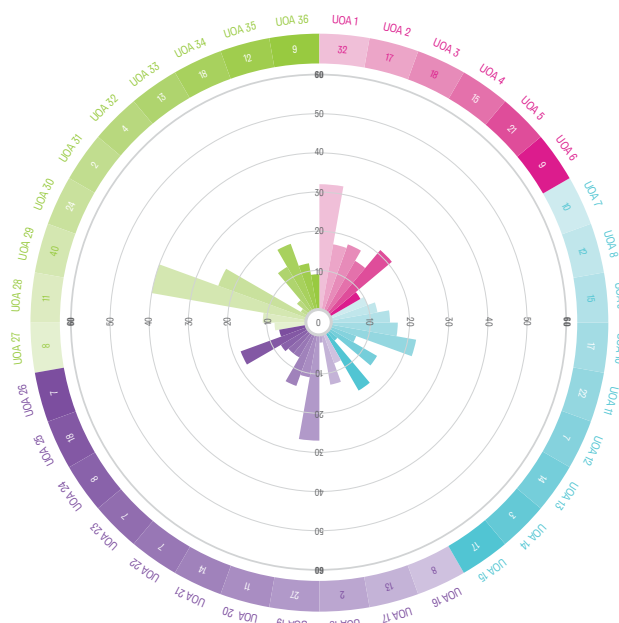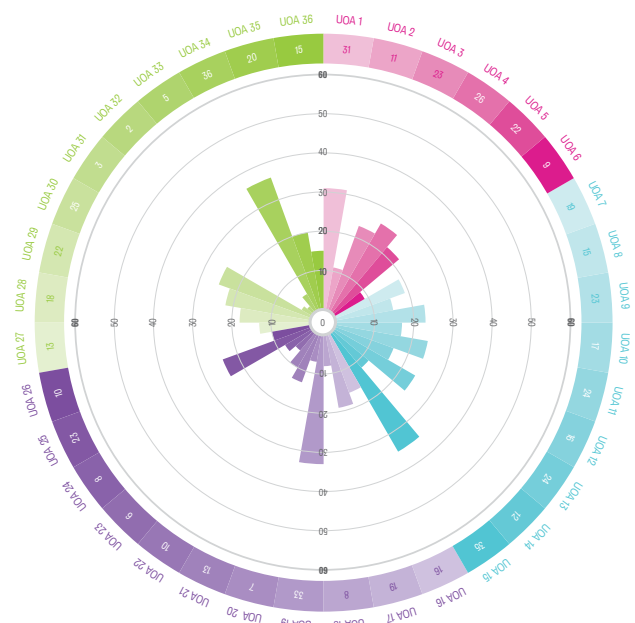, scale and beneficiaries of research impact by applying text-mining techniques and reading just over 1,000 case studies. As noted in the introduction, the analysis is more illustrative than final: with the case studies being made openly available, other analysts will develop, improve and add to this work over the coming months.[65] Indeed, with the development and application of more sophisticated text-mining techniques, a new set of tools and approaches are likely to emerge that will support the 'science of science' and provide evidence to inform future policy.

That said, it is also important to recall and acknowledge the limitations of the current analysis. As described in Chapter 1, these fall into five broad categories:

1. The limited amount of time available for the analysis, restricting what we could do – with more time we would have developed a number of the analyses further

2. The lack of consistent definitions used in the case studies – making it very difficult to compare and synthesize case studies on a common topic

3. The positive sentiment in the way that impact was articulated and described

4. The selective nature of the case studies for the purpose of REF assessment, and

5. The risk of double counting due to identical or near-identical case studies.

As noted earlier, we also emphasize that everything we have reported is a result of analysis on the written text describing impact within the REF impact case studies, and not an analysis of impact per se.

As also indicated in Chapter 1 (Box A, page 16), the topic modelling method produced an output of 65 clusters (of which we presented 60) containing words that were related to a theme or topic. Each set of topics is generated through a probabilistic approach, which means that sets of results differing in detail may arise each time the topic modelling algorithm is run. Additionally, it is worth noting that we assigned a topic label to each of these clusters, based on our assessment of the words found in that topic (see full list in Annex D). We went through a number of iterations of labelling based on internal review and further reading of the content of these case studies tagged by a particular topic. Therefore, the labelling process was subjective. We also note that for most of our analyses, we took the top three topics that were most represented within each case study to find the distribution of topics across the case studies and create the impact wheels and heat map.

We finish the report with two sets of reflections. The first consists of some concluding policy observations arising from the analysis, and the second is a set of lessons we learned in undertaking the analysis on how future iterations of REF could be implemented, if further consideration were given to metadata and structure of the case studies.

---

65  http://impact.ref.ac.uk/casestudies

## Policy observations

As illustrated in the preceding chapters, there are many policy observations to make from the synthetic analysis. Here we highlight four high-level observations.

**The societal impact of research from UK Higher Education Institutions (HEIs) is considerable, diverse and fascinating**

One of the most striking observations from the analysis of the REF case studies is the diverse range of contributions that UK HEIs have made to society. This is illustrated in the heat map of 60 impact topics by the 36 Units of Assessment (UOAs; Figure 8, page 33), impact wheels (Figure 10, page 35) and the six deep mine analyses in Chapter 4.

While this observation may be in part an artefact of the topic modelling approach that we adopted, it also should be noted that this is one of the first empirically driven attempts to catalogue research impacts on this scale. Indeed, in reading the case studies, the abiding impression is the heterogeneity in the types of impact that arise from academic research. From a policy perspective, this reinforces evidence on the broad contribution that research makes to the economy, society, culture, public policy and services, health, the environment and quality of life in the UK and globally. It also suggests that the use of case studies in REF 2014 was the appropriate way to assess research impact, as it allowed the case study authors to define and articulate a range of impacts, some of which inevitably would not been have captured through a 'top-down' taxonomy.

**The research underpinning societal impacts is multidisciplinary, and the social benefits arising from research are multi-impactful**

The relationship between 149 field of research (FORs), 36 UOAs and 60 impact topics is visualized in the alluvial diagram in Figure 12 (page 39). What is evident from this visualization is that multiple FORs underpin the case studies, leading to multiple types of impact – and this is weakly correlated with the UOA. It is interesting to note that there were 3,709 unique impact pathways (from FOR to impact topic) in the alluvial diagram. This would suggest that any attempt to define a standard route to research impact could be counterproductive, and that both incentives to encourage researchers to consider impact (such as Research Council 'pathways to impact') and reward mechanisms for achieving impact (such as REF case studies) should continue to be narrative-based.

**Different types of HEIs specialize in different types of impact**

Analysis of the impact topics by the size of HEI submission to REF 2014 illustrates that within this sample of case studies, small institutions make a disproportionate impact contribution, and to a larger number of topics, than larger institutions. As shown in Figure 11 (page 37), although some larger institutions make large contributions to topics such as 'Clinical guidance' and 'Dentistry', small institutions make a greater than anticipated contribution to topics such as 'Sports', 'Regional innovation and enterprise' and 'Arts and culture'. This raises some interesting research policy questions with regard to the nexus of selective funding, concentration and research excellence (Annex A).

**UK HEIs have a global impact**

The map in Figure 13 (page 41) illustrates the geographical reach of the impacts of research which have been undertaken in UK HEIs over the past 20 years. From a policy perspective this illustrates the global contribution of research undertaken by UK HEIs. The deep mine on the BRIC countries (Brazil, Russia, India and China; page 66) further illustrates the diversity of this impact, with benefits occurring across all UOAs for each of the four key countries. It was also interesting to note the different impact wheels for the different countries, with Brazil being preferentially cited in UOA 19 (Business and management studies), Russia UOA 30 (History) followed by 29 (English language and literature), India UOA 29 (English language and literature) and China UOA 34 (Art and design: history, practice and theory) followed by and UOA 15 (General engineering).

## Lessons for future iterations of the Research Excellence Framework

Below we draw out three areas that HEFCE and the other funding bodies may wish to reflect on for future iterations of research impact assessment as part of REF. We include a caveat to these observations by noting that REF is an **assessment** system, but that our interests in this report are in the effective **analysis** of the case studies – and that at times the two purposes may play against each other. For example, as noted below, we suggest that analyses could be improved through the collection of more structured and standardized information; however, this may add to the burden for case study authors, and inevitably would tend to privilege certain types of impact over others.

### The impact case studies provide a rich resource for future analysis

The publication of the 6,679 non-redacted case studies provides a unique resource for further in-depth analysis and the development of new methodologies and approaches for the analysis of research impact. It will be important that research funders support the development of this nascent field of research in its own right. There are numerous questions that could be addressed through further analysis of the case studies, and some of this evidence will help to create, secure, accelerate and maximize future research impacts. Some of these questions are summarized in Annex B, but others will arise as people consider this and other reports.

In undertaking future analyses, we re-emphasize the limitations of the case studies as primary research material: they articulate a select set of impacts, written to specific rules as defined in the REF guidance, and in a style and tone that aims to 'sell' the impact to the assessment panels.

### The quantitative evidence supporting claims for impact was diverse and inconsistent, suggesting that the development of robust impact metrics is unlikely

There was a large amount of numerical data (ie, c170,000 items, or c70,000 with dates removed) that was inconsistent in its use and expression and could not be synthesized. In order for impact metrics to be developed, such information would need to be expressed in a consistent way, using standard units. However, as noted above, the strength of the impact case studies is that they allow authors to select the appropriate data to evidence their impact. Given this, and based on our analysis of the impact case studies, we would reiterate the conclusion made in the original study that informed the formulation of the approach for research impact assessment adopted for REF: 'impact indicators are not sufficiently developed and tested to be used to make funding decisions' (Grant et al, 2010).

### The use of standardized lists of information and the definitions in the case studies would aid future analysis

As noted at a number of points throughout this report, the non-standardized use of information hampered our analysis. This was the case, for example, in the deep mines on select committees (page 55) and the funders (page 61). In both cases ideally we would envisage a drop-down list (or equivalent) of either select committees or research funders that case study authors can use to select the appropriate entity, thereby ensuring the standard use of language in the case studies and the tagging of that information as metadata for future analysis.

For numeric information that is used to provide evidence of impacts, a set of standardized definitions would aid future analysis. For example, as noted in the deep mine on the health benefits of research (page 49), different case study authors assigned different monetary values to Quality Adjusted Life Years (QALYs). If HEFCE and its funders were to make a statement asking case study authors to use a particular standard value for a QALY (eg, £30,000), this would facilitate future analysis. Similarly, a statement for those wishing to use financial information in demonstrating impact could ask authors to do so in pound sterling at a given exchange rate and purchasing power parity (with such information provided), which would significantly help to improve return-on-investment-type analyses.

## Closing comment

We would like to finish this report by acknowledging what a privilege and pleasure it has been to analyse and read the case studies submitted to REF 2014. The stories that are told in the impact case studies capture what is great about academic research in the UK: the range of interests, the expertise and experience, the commitment of individuals and the benefits to communities across the world. The case studies make an inspiring read that demonstrate the value of research to today's society, and to the future.

# References & Annexes

# References

Advisory Board for the Research Councils (ABRC) (1987). 'A Strategy for the Science Base'. Discussion document prepared for the Secretary of State for Education and Science by the Advisory Board for the Research Councils. London: HMSO.

Advisory Board for the Research Councils (ABRC) and University Grants Committee (UGC) (1982). *Report of a Joint Working Party on the Support of University Scientific Research* (the Merrison Report), Cmnd 8567. London: HMSO.

Al Sumait L, Barbará D, Gentle J, Domeniconi C (2009). Topic Significance Ranking of LDA Generative Models. Lecture Notes in *Computer Science: Machine Learning and Knowledge Discovery in Databases.* 5781, 67-82.

Australian Technology Network of Universities and Group of Eight (2012). 'Excellence in Innovation Research Impacting Our Nation's Future – Assessing the Benefits', Australian Technology Network of Universities. Available at: www.atn.edu.au/eia/Docs/ATN-Go8-Report-web.pdf [last accessed 27 February 2015].

Blei DM (2012). 'Probabilistic Topic Models'. *Communications of the ACM*, 55 (4), 77–84. Available at: www.cs.princeton.edu/~blei/papers/Blei2012.pdf [last accessed 27 February 2015]

Donovan C (2008). 'The Australian Research Quality Framework: A Live Experiment in Capturing the Social, Economic, Environmental and Cultural Returns of Publicly Funded Research'. In CLS Coryn and M Scriven (eds), *Reforming the Evaluation of Research.* New Directions for Evaluation, 118, 47–60.

Grant J, Brutscher P-B, Kirk S, Butler L and Wooding S (2010). 'Capturing Research Impacts: A Review of International Practice'. Report prepared for the Higher Education Funding Council for England. Cambridge: RAND Europe. Available at: www.rand.org/content/dam/rand/pubs/documented_briefings/2010/RAND_DB578.pdf [last accessed 27 February 2015]

Hanney SR, Castle-Clarke S, Grant J, Guthrie S, Henshall C, Mestre-Ferrandiz J, Pistollato M, Pollitt A, Sussex J and Wooding S (2015). 'How long does biomedical research take? Studying the Time Taken between Biomedical and Health Research and its Translation into Products, Policy, and Practice'. *Health Research Policy and Systems*, 13(1), doi: 10.1186/1478-4505-13-1. Available at: www.health-policy-systems.com/content/13/1/1 [last accessed 27 February 2015]

Levitt, R. Celia C, Diepeveen S, Chonaill S, Rabinovich L, Tiessen J (2010). *Assessing the Impact of Arts and Humanities Research at the University of Cambridge.* Cambridge: RAND Europe.

Mansfield E, (1990). 'Academic Research and Industrial Innovation'. *Research Policy*, 20 (1), 1–12.

Manville C, Guthrie S, Henham M-L, Garrod B, Sousa S, Kirtley A, Castle-Clarke S, and Ling T (2015a). *Assessing impact submissions for REF 2014: An evaluation (2015).* Cambridge: RAND Europe.

Manville C, Morgan Jones M, Frearson M, Castle-Clarke S, Henham M-L, Gunashekar S and Grant J (2015b). *Evaluation of Submission Preparation for Impact Assessment REF 2104. Findings and Observations.* Cambridge: RAND Europe.

Morgan Jones M and Grant J (2013). 'Making the Grade. Methodologies for Assessing and Evidencing Research Impact. 7 Essays on Impact'. DESCRIBE Project Report for Jisc. Dean et al. (eds.), Exeter: University of Exeter, 25-43.

Morgan Jones M, Castle-Clarke S, Manville C, Gunashekar S and Grant J (2013). *Assessing Research Impact. An International Review of the Excellence in Innovation for Australia Trial.* Cambridge: RAND Europe. Available at: www.rand.org/content/dam/rand/pubs/research_reports/RR200/RR278/RAND_RR278.pdf [last accessed 27 February 2015].

Morris Z, Wooding S and Grant J (2011). 'The Answer is 17 Years, What is the Question: Understanding Time Lags in Medical Research', *Journal of the Royal Society of Medicine*, 104(12), 510–20.

National Institute for Health and Clinical Excellence (NICE) (2008). *Guide to the Methods of Technology Appraisal*. London: NICE.

O'Neill, Jim (2001). 'Building Better Global Economic BRICs', Global Economics Paper No. 66. Goldman Sachs.

Oliver K, Innvaer S, Lorenc T, Woodman J and Thomas J (2014). 'A Systematic Review of Barriers to and Facilitators of the Use of Evidence by Policymakers'. *BMC Health Services Research*, 14(2), doi: 10.1186/1472-6963-14-2. Available at: www.biomedcentral.com/1472-6963/14/2/abstract [last accessed 27 February 2015].

Peacock J, Harvey L, Barber M, McFadden P, Marceau J, Davey I, Clark P, Fell C, King W, Enright T, Schreuder D and Smith I (2006). *Research Quality Framework. Assessing the quality and impact of research in Australia: Research Impact.* Development Advisory Group for the RQF.

REF (2011). *Assessment Framework and Guidance on Submissions. REF 02.2011* (as of 28 July 2014) Available at:: www.ref.ac.uk/pubs/2011-02/ [last accessed 27 February 2015].

REF (2012). Panel Criteria and Working Methods. REF 01.2012 (as of 28 July 2014). Available at: www.ref.ac.uk/pubs/2012-01/ [last accessed 27 February 2015].

Roberts G, Arthur E, Barber M, Batterham R, Callaghan P, Chubb I, Clark P, Edwards A, Hoj P, Milbourne R, Pettigrew A, Sheehan P and Smith I (2005). Research Quality Framework: Assessing the quality and impact of research in Australia, The Preferred Model. Expert Advisory Group for the RQF. RQF Issues paper

Shattock, M (1994). *The UGC and the Management of British Universities.* Buckingham: Society for Research into Higher Education and Open University Press.

Technopolis (2010). *REF Research Impact Pilot Exercise Lessons-Learned Project: Feedback on Pilot Submissions*: Final report, November. Available at www.ref.ac.uk/pubs/refimpactpilotlessons-learnedfeedbackonpilotsubmissions/ [last accessed 27 February 2015].

University Grants Committee (UGC) (1985). *A Strategy for Higher Education into the 1990s: The University Grants Committee's Advice.* London: HMSO.

Varcoe I (1974). *Organizing for Science in Britain.* Oxford: Oxford University Press.

Wilkie T (1991). *British Science and Politics since 1945.* Oxford: Blackwell.

Witty A (2013). *Encouraging a British Invention Revolution: Sir Andrew Witty's Review of Universities and Growth – Final Report and Recommendations.* London: Department for Business Innovation & Skills. Available at: www.gov.uk/government/uploads/system/uploads/attachment_data/file/249720/bis-13-1241-encouraging-a-british-invention-revolution-andrew-witty-review-R1.pdf [last accessed 27 February 2015].

# Annexes

## Annex A: A short history of research assessment

In 1916, the UK Government created the Department for Scientific and Industrial Research (DSIR) to support civil science and co-ordinate research (Varcoe, 1974). In this small 'dual support' system, the University Grants Committee (UGC) paid for salaries while the DSIR gave funds for specific research. Selectivity was unimportant as there were only 24 DSIR postgraduate researchers. After 1945 the UGC took a more active stance. Researcher numbers doubled during the 1950s, and research spend in 1962 was ten times that in 1945 (Wilkie, 1991). From that point the UGC operated selective funding through expert subject committees and policy (Shattock, 1994; Wilkie, 1991).

The UGC's quinquennial cycle ended in 1976–7 when, despite cuts, it selectively allocated £500,000 to enable 11 institutions in areas of high priority. The Annual Survey for 1979–80 announced that:

'[T]he current distribution of equipment grant (£72M for 1980/81 cf. recurrent grant of £987M) takes into account each university's past record of attracting outside research grants and thus provides a slightly better equipment base for those with a proven research capability.' (UGC, 1985, pp 27-28)

Selectivity was endorsed by all sides of the research support system, including the UGC, the Advisory Board for the Research Councils (ABRC) and the Advisory Committee for Applied Research and Development (ACARD). The 1982 Merrison Committee concluded 'that whatever research is done should be of high quality and properly supported', and that 'universities will need to concentrate research funds into selected areas' (ABRC and UGC, 1982). The joint report of the chairs of the ABRC and ACARD proposed that there should be a national and overt policy of selectivity among research objectives.

The UGC's *A Strategy for Higher Education into the 1990s* suggested a 'more selective allocation of research support among universities' (UGC, 1985, p 17). In 1985, the UGC told universities that the distribution of research funds would take account of work of special strength and promise, so as to maintain quality. This led, in 1986, to the first Research Selectivity Exercise, applying the ratings to the selective allocation of part of the research resource. This assessment evolved into the Research Assessment Exercise (RAE) of 1989, modified into a recognizable, long-term format in 1992.

In 1987, the ABRC's *A Strategy for the Science Base* recognized that 'selectivity and more directive management … can be seen as the inevitable response to the challenge of managing science within finite resources'. It suggested that the allocation of Research Council grants to scientists in below-average departments was 'not conducive to the concentration of effort that we believe generally to be in the national interest' (ABRC, 1987, p 5).

The ABRC concluded that greater concentration was required to maintain the international competitiveness of university research (among just 60 institutions, not the 150-odd of 2014). It proposed an 'R-T-X' (Research, Teaching and Hybrid) system differentiated according to breadth of research strength, with just 15 universities in the 'R' category of international research across most fields.

Over 30 years, UK research funding has regularly involved:

- Selective resource allocation among disciplines

- Selective allocation among researchers, mainly through peer review

- Selective distribution among departments by internal management, and

- Selective distribution across institutions, sometimes via obscure committee decisions.

The introduction of selective funding based on the peer assessment of wider research impact is in keeping with the continuing development of UK research assessment.

## Annex B: Analytical questions arising from the request for proposal and workshop

### Questions from the Request for Proposal (RfP)

1. What types of impact outcomes have been submitted to the REF? How does this vary by discipline/by user type (beneficiary)/by institution?

2. What are the pathways by which different types of impact outcome have been realized? How does this vary by discipline/by user type (beneficiary)/by institution?

3. To what extent are 'negative' findings included in the impact case studies or are only 'positive' stories submitted? Are 'learning'-type impacts reported?

4. What are the time lags between underpinning research and impact outcome exampled in the impact case studies? How does this vary between types of impact, user and disciplines?

5. What quantitative data (for inputs and/or outcomes) can be extracted and synthesized from the impact case studies: eg commercialization revenue, visitor numbers? Is it possible to estimate overall return-on-investment figures?

6. Do the case studies contain quantitative information on impact outcome that is sufficiently common and general to be developed into metrics of research impact?

7. According to the impact case studies, what types of research users benefit from higher education research, and to what extent?

8. What are the characteristics of the underpinning research outputs on which impact case studies are based (in terms of methodologies, approaches or research topics)? Do these vary by type of impact? What is the role of interdisciplinary and multidisciplinary research in leading to impact?

9. Can we learn anything about how research collaboration affects impact?

10. How does the impact described in the REF case studies relate to the Government's industrial strategy and other devolved administration economic/industry strategies?

11. Can we learn anything about connections between social and economic impact and citation impact?

| Questions from Panel A group | Number of votes |
|---|---|
| How many patents resulted from the work? | 1 |
| How many case studies only describe qualitative impacts? | 9 |
| Are negative findings that have impact mentioned? | 1 |
| What proportion of case studies involve interdisciplinary work? | 4 |
| Are the references cited in high-impact journals? | 3 |
| Are impacts related to underpinning research in previous exercise (RAE)? | 0 |
| Can 'boilerplate' words be systematically excluded? Biomedical specific terms related to impact (or other domain areas)? | 0 |
| How many case studies mention industrial collaboration? | 4 |
| How many case studies mention international collaboration? | 3 |
| How many carbon credits were saved by research in the REF? | 5 |
| Who is described as the primary audience? Policymakers/management/public? | 10 |
| How many spin-outs or start-ups were caused by the work? | 0 |
| What intermediaries of impact are mentioned (are they invisible)? | 4 |
| How many engagement events take place/stakeholder groups/ total attendance/ outreach? | 0 |
| How many policies were affected? | 3 |

| Questions from Panel B group | Number of votes |
|---|---|
| What facilities were used to provide impact? | 6 |
| How many studies exemplify cost reduction? | 3 |
| What endpoint was defined for the impact? | 0 |
| Engineering projects: what was monetary value of impact? | 0 |
| How parliament (vs. departmental/policy) was cited: beneficiary pathway to impact? | 13 |
| How many schoolchildren have been affected and reached? Search on curriculum? | 4 |
| Where does impact come from (eg, from multiple collaboration between universities, companies, public/third sector) ? How is collaboration initiated? | 7 |
| Panel D: What is the value of creative industries (examples thereof)? | 1 |
| What are contributions to standards (eg BSI)? | 5 |
| What kind of corroborative evidence is cited as proof of impact? | 7 |
| Where has research been used to fuel business innovation and growth – small and medium-sized enterprises in particular? | 4 |
| Can the company's turnover, job creation and inward investment be identified/collated? | 1 |
| How is research delivering impact in challenge areas, climate change, autonomous vehicles and crime? Where is the research coming from? | 3 |

| Questions from Panel C/D group | Number of votes |
|---|---|
| How and when do researchers engage with specific beneficiary groups (e.g. pick one of parents, kids, etc.)? | 1 |
| What types of productive interactions or engagement between academics and non-academics have led to certain types of impact (eg impact on policy)? | 12 |
| What is the proportion of research impact that is collaboratively delivered (i.e. non-academic partners)? | 0 |
| What is the role of mediators/translators in creating impact for social science and humanities research, where academics worked with external organizations to disseminate/translate in order to influence policy and practice? | 2 |
| What is the proportion of early career researchers' involvement (or PhD) in delivering research impact? | 11 |
| Is there evidence of sustainable partnership beyond the life of the project? | 0 |
| How many impacts were planned vs. came about due to serendipity? | 8 |
| What are the types of impact derived from basic vs. applied research (plus scale of impact/timeframe) ? | 2 |
| What are the educational benefits of research? | 0 |
| Was the impact realized/progressed by a deliberate (funding) programme? | 1 |
| How do researchers perceive 'impact', and what are their sources of evidence? | 1 |
| Is the attribution/contribution of individual research groups to claimed impacts convincing? | 0 |
| Has economic research been used to start/grow businesses? | 1 |
| How many/what proportion of case studies mention patent application/approval? | 0 |
| What is the scale/success outreach activiy undertaken: discipline, residence, background? | 1 |
| What is the presence of academics in the media? (eg appearances on TV/radio, blogs, news, etc.) | 0 |
| How is social media being used to communicate research and contribute to impact? | 5 |
| Have cultural places generated visitors to the local economy? | 1 |
| Is there evidence of audience response to the performance/exhibition? What form does this take? (eg questionnaire, focus group) | 1 |
| What is the economic value of university–arts collaboration? | 4 |
| What value can be given to cultural impact on local/regional/city-regional economies? | 6 |
| How far does collaborative research (in arts and humanities) inform arts funding in local government? | 0 |
| How many/what proportion refer to public understanding of 'science' (perhaps via mention of altmetrics)? | 0 |
| Has university research influenced the artistic development of arts organizations? | 0 |
| Was there an increased proportion of public engagement cited for Panels C and D? | 0 |
| What is the range of impacts on culture/cultural value captured by the impact case studies? | 1 |
| Does it offer ways of measuring cultural values beyond commercial and policy? Can these be extrapolated to apply in future case studies? | 0 |
| How have examples of cultural impact been demonstrated/evidenced? What is the value of cultural impact? | 0 |
| How does research impact on the policy process? What are the mechanisms of engagement? | 7 |
| What is the (direct) influence in policy process? (Counting things from more specific: taking part in select committee; to less specific: narratives of informal influence.) | 0 |
| What is the number of examples where research (social science) has had an impact on government policy? | 10 |
| How many case studies have affected UK and overseas national policy? | 0 |
| How does research influence government policy? | 0 |
| How many/what proportion explicitly reference 'legislation'? | 0 |
| To what extent are research findings cited/used as the basis of final government policy? | 0 |
| What is the relationship between the main source of funding and impact? | 0 |
| Does practice-based research in the performing arts affect/impact funding of these art forms? | 2 |
| What impact have British Academy funding schemes generated? | 3 |
| How many impact case studies reference QR funding? | 2 |
| What is the value of improvements to processes of services by this research? | 0 |

## Annex C: Approach to data analysis and topic modelling

The original texts from the impact case studies were supplied to us in PDF format, and these had to be cleaned and processed for subsequent analysis. In this annex we describe the text-cleaning and processing steps, and our approach to topic modelling and the text-mining methods applied in our analysis.

### Text cleaning and processing

The contents of the impact case studies were supplied in PDF format, from which we extracted plain text content for subsequent analysis. The extraction was performed using 'pdftotext' – a UNIX command line utility – with the 'layout' command option used in order to capture as much as possible of the text structure and line breaks. Given that each PDF contained all the sections of the REF submitted impact case study, the various sections of the document were extracted for separate analysis using regular expressions (ie pre-specified language patterns in the texts) implemented in a custom Python script (written by the report authors).

We then performed various iterations of processing the text to look for patterns, find nouns, entities, etc. to find out what types of processing and analysis needed to be done to answer our research questions. This exploration of the texts was performed using the Natural Language Toolkit (www.nltk.org) for entity extraction and part-of-speech tagging. Custom scripts (written by the report authors) were used to find regular expressions, word count information and perform keyword-in-context (KWIC) analysis.

The results reported in this study were based on the extracted text which have been cleaned-up to remove non-ASCII characters and punctuation. The topic modelling analysis required further processing steps as follows:

- 'Details of the impact' (Section 4) were extracted from each case study

- The various indicators used to represent redacted text were consistently replaced with the same marker – ie 'xxxx'

- 'Stop words' (eg 'and', 'but', etc.) were removed, along with custom words or phrases specific to the language of the case studies (eg numbers that indicate pages or sections)

- Apostrophes and other punctuation were removed

- Words were stemmed (using the Porter Stemming Algorithm, http://tartarus.org/martin/PorterStemmer/) – that is, different forms of the same word (eg 'runs', 'running', 'ran') were reduced to a consistent or most basic form (eg in this case, 'run'). Note that stemming may create an output of words that do not contain a real form of the original word (eg 'poetri'; see Annex D for other examples).

### Background to topic modeling

Topic modelling is a machine learning technique which relates closely to clustering algorithms. However, instead of attempting to group whole documents into clusters, topic modelling describes a set of topics to which the actual word content of the documents – as a whole – may be related. Documents may relate to more than one topic, and topic modelling calculates a weight with which each topic relates to a particular document.

We used Latent Dirichlet Allocation (LDA) to conduct the topic modelling analysis. One of the most important aspects of topic modelling as implemented in LDA is that rather than simply basing topics on word features occurring in documents together, it uses contextual information of word occurrences in documents, and so can capture words with similar meanings that are used interchangeably within similar contexts.

As a generative technique, LDA starts with a model that is then used to describe the data by adjusting the parameters to fit the model. Here, the assumption is that the whole corpus of documents contains k number of topics, and that each document talks about these k topics to a greater or lesser extent. Therefore, each document is assumed

to be a combination of these topics, each with different probabilities. One of LDA's strengths is that it performs clustering at two levels: first, for words into topics; and second, for documents into these topics. LDA works by starting with an empty topic model, then reading all the documents in parallel to calculate the probability for each topic for each word in the document. After each run, the model is updated and the process is repeated until the model better explains the documents: that is, how well the model fits the data.

LDA is the accepted state-of-the-art in topic modelling and is implemented in many standard toolboxes for machine learning. Thus we can compare results easily. However, there are some known limitations to LDA that mainly stem from the fact that it is a probabilistic model. In practice, these models are known to be sensitive to data changes and adjustments in the overall processing of the machine learning (such as the stemming of its constituent words). Furthermore, relatively small document collections such as ours can produce unstable topics if new documents are added (although this is less of an issue in our case, as we work with a fixed total of documents). Finally, topics can be too suggestive[66], and require careful review by human evaluators often only relying on ad hoc techniques to be useful.

All this means that topic modelling should be applied carefully, using iterative cycles of evaluation and testing, and only with a good understanding of the underlying data and domain. However, in our experience this applies to most successful machine learning techniques. LDA belongs to a family of algorithms that are known to work well with small shifts in meanings where similar ideas are described using different language expressions, and is therefore well suited to our data.

## Topic modelling process for Research Excellence Framework impact case studies

The LDA was performed on the data using the Mallet machine learning tool;[67] specifically the beta version was used in order to generate diagnostic information of the models for their evaluation;[68] a low alpha parameter (0.01) was specified in order to generate fewer topics which loaded more strongly onto the documents.

Since a topic modelling exercise requires human evaluation of the quality of the topics, a range of models was generated, with the number of topics ranging from 10 to 100 at coarse intervals (eg, 10, 25, 50, 75, 100). The reason for this was to identify a reasonable number of topics to extract, since problems relate to topics being 'too general' or 'too specific/narrow' typically arise. Topics were evaluated initially by visual inspection of the topic keys output (ie the top-loading words for each topic). As a result of this, we concluded that 50–100 topics were likely to be appropriate. Once the range of number of topics for the model had been narrowed by the first iteration, a second human evaluation was applied using topic keys for a narrower range (50, 60, 70, 80, 90, 100).

We sorted the topics in ascending order of their weighting, then read the top 9 keys (words loading on each topic). First, all topics in a model were checked for repetition or duplication of similar topics; second, topic keys were checked for internal consistency and plausibility (ie, did the words appear semantically related, in that they have similar and/or overlapping meanings; in particular, were they more semantically related to each other than to words in other topics).

Following visual inspection of the topics in these models, a third set of a yet more concentrated range of topic models was generated around the most promising number of topics (eg, 65, 70, 75, 80), using ad hoc evaluation. Then the main diagnostics relating to these models were plotted to check general trends: for example, do the models appear to perform better or worse with greater or fewer topics in our range (for each model, the average values for each topic were used).

There were four main standard diagnostic measures generated by LDA which we

---

66  http://sappingattention.blogspot.co.uk/2011/11/compare-and-contrast.html [last accessed 27 February 2015].
67  http://mallet.cs.umass.edu [last accessed 27 February 2015].
68  http://article.gmane.org/gmane.comp.ai.mallet.devel/1483/ [last accessed 27 February 2015].

considered as explained below, and which are regarded as important in identifying 'junk' or 'insignificant' topics (see eg Al Sumait et al. 2009 for further discussion).

1. Coherence – this is measured via the semantic similarity of words to those previously loading within that topic. When this semantic similarity is maximized, it indicates topics with the greatest internal consistency. For assessing the number of topics to specify in the analysis, the mean coherence for each of the different number of possible topic models was considered.

2. Distance from corpus – this is a measure which indicates the distinctiveness of the words in a topic relative to the rest of the corpus: that is, what the likelihood is of words occurring in a topic occurring similarly in the background corpus. We sought to maximize the distinctiveness of our topics, and selected the number of topics specified for the topic model which maximized the average value for the analysis.

3. Documents at Rank 1 – this relates to the number of documents that this topic best describes, and therefore indicates topics which are more distinctive. We selected the number of topics for the model based on higher average scores across all topics of a model.

4. Distance from uniformity – this indicates the specificity of a topic by measuring the probability of a topic relating to a smaller number of words. In this case we selected the number of topics for our model based on a lower average score across all topics in each model.

Using this process, we judged 65 topics to be the optimal number to analyse and model the data. On inspection, we noted that five of these had outputs that were descriptors of impact scale (eg 'increased', 'numbers', 'effici'), but gave no indication of context or theme (that is, they could be regarded as insignificant topics; Al Sumait et al, 2009) and therefore were excluded, leaving us with 60 topics to use in the study.

## Limitations

In common with other studies that use these methods, we note that there are limitations to the use of topic modelling and LDA. Firstly, it is a probabilistic approach and therefore the output can be different each time the analysis is run. Secondly, the quality of the model depends on the number of topics specified and how well suited this number is to the data. To obtain our 65 topics we ran over 20 iterations of the analysis using different numbers of topics between 5 and 100. We settled on 65 topics, since this gave the greatest consistency and sensible combination of topics (as determined by human judges and diagnostic tests). The 65 topics were generated to include the top 9 'key' words which load onto each topic, and which are used to interpret the topics (displayed in Annex D).

# Annex D: Topic model labels and associated words

| Topic label | Words related to this topic |
| --- | --- |
| Animal husbandry and welfare | anim welfar farm veterinari breed diseas control uk farmer |
| Architecture and building | design build construct standard industri structur project architectur engin |
| Arts and culture | art artist work cultur creativ project public audienc exhibit |
| Asia | china chines india arab indian asian intern east foreign |
| Banking, finance and monetary policy | bank financi polici econom financ credit tax risk central |
| Business and industry | compani busi manag industri product market servic improv sector |
| Cancer | cancer patient treatment clinic trial uk breast guidelin therapi |
| Children, young people and families | children child young parent famili imp programm work support |
| Climate change | climat chang energi carbon emiss uk environment adapt wast |
| Clinical guidance | guidelin patient clinic treatment recommend stroke nice risk trial |
| Clinical tests | test patient clinic genet diseas diabet diagnosi diagnost treatment |
| Community and local government | local commun project citi council social peopl fund develop |
| Computing and quantum physics | comput secur light ibm physic intel scienc particl imag |
| Crime and justice | polic crime prison justic xxxx offic violenc offend victim |
| Cultural and heritage preservation | heritag archaeolog site visitor histor museum project cultur tourism |
| Defence and security | defenc militari secur war conflict uk forc arm offic |
| Democracy and political engagement | polit elect parti democraci elector vote candid poll pd |
| Dentistry | kcl dental drug oral treatment king prof scott health |
| Engineering, design and manufacturing | engin design process manufactur fuel develop materi industri improv |
| Europe | european eu europ intern commiss polici human countri state |
| Film and theatre | film theatr perform plai audienc product festiv screen director |
| Food and nutrition | food product industri nutrit health crop agricultur uk seed |
| Health care services | health care servic nh hospit patient nation improv practic |

87

| | |
|---|---|
| **Historical archives** | histori archiv public histor project librari heritag cultur materi |
| **Infectious diseases control** | malaria control health diseas resist infect treatment drug programm |
| **Informing government policy** | Develop polici nation plan govern inform work strategi assess |
| **Instrumentation** | laser instrument materi product process imag manufactur develop industri |
| **International development** | develop countri intern world africa polici global govern african |
| **Laboratory diagnostics** | test assai diagnost dna detect protein laboratori sequenc develop |
| **Law and justice** | law legal court justic case judg act legisl lawyer |
| **Literature** | book read poetri write literari writer publish literatur translat |
| **Marine and ocean science** | marin fish fisheri sea coastal ship ocean manag environment |
| **Media** | public bbc media radio programm interview time broadcast articl |
| **Medical ethics** | ethic disabl human transplant cell donat donor uk medic |
| **Mental health** | mental health clinic servic train treatment intervent patient psycholog |
| **Mobile technologies** | mobil system technolog network servic digit app phone commun |
| **Modelling and forecasting** | model data method statist forecast predict estim risk measur |
| **Museums and exhibitions** | exhibit museum visitor art galleri collect curat displai public |
| **Music, dance and performance** | music perform danc work sound audienc concert record festiv |
| **Nature and conservation** | conserv natur manag forest land speci biodivers environment project |
| **Nuclear energy** | nuclear power energi nois electr system industri monitor oper |
| **Oil and gas** | oil ga space explor industri model field bp mission |
| **Parliamentary scrutiny** | polici govern report public uk committe debat evid commiss |
| **Pharmaceuticals** | drug develop pharmaceut trial compani clinic phase discoveri industri |
| **Print media and publishing** | univers book intern translat world publish de public uk |
| **Public engagement** | peopl particip wai experi comment engag cultur discuss life |
| **Public health and prevention** | health screen hiv vaccin women programm recommend prevent nation |

| | |
|---|---|
| **Regional innovation and enterprise** | innov busi region sme birmingham enterpris support programm univers |
| **Regional languages of British Isles** | languag ireland wale welsh northern cardiff irish english commun |
| **Religion** | church religi christian religion faith cathol spiritu confer bibl |
| **Schools and education** | educ school teacher student teach learn univers develop curriculum |
| **Scotland** | scottish scotland glasgow edinburgh govern aberdeen public dunde commun |
| **Software development** | softwar develop tool system user data model project comput |
| **Sports** | sport game coach footbal athlet olymp perform physic player |
| **Surgery, implants and devices** | patient clinic surgeri hospit medic imag implant surgic devic |
| **Technology commercialization** | technolog compani develop product univers commerci system market industri |
| **Transport** | transport safeti road rail risk fire oper train uk |
| **Water and flood management** | water flood environ risk manag environment uk qualiti pollut |
| **Women, gender, and minorities** | women equal gender migrat divers ethnic commun group refuge |
| **Work, labour and employment** | employ union labour trade work worker wage employe social |

www.kcl.ac.uk/sspp/policy-institute
@policyatkings