# EPSRC Bibliometric Study 2018

Commissioned by the Engineering and Physical Sciences Research Council (EPSRC)

April 2019

Clarivate Analytics

PAGE INTENTIONALLY LEFT BLANK

# Table of Contents



Appendices III - IX are available on request from EPSRC

Clarivate Analytics

# AUTHORS

**Ivelisse Cabrera, Ph.D.**
Scientific Analyst

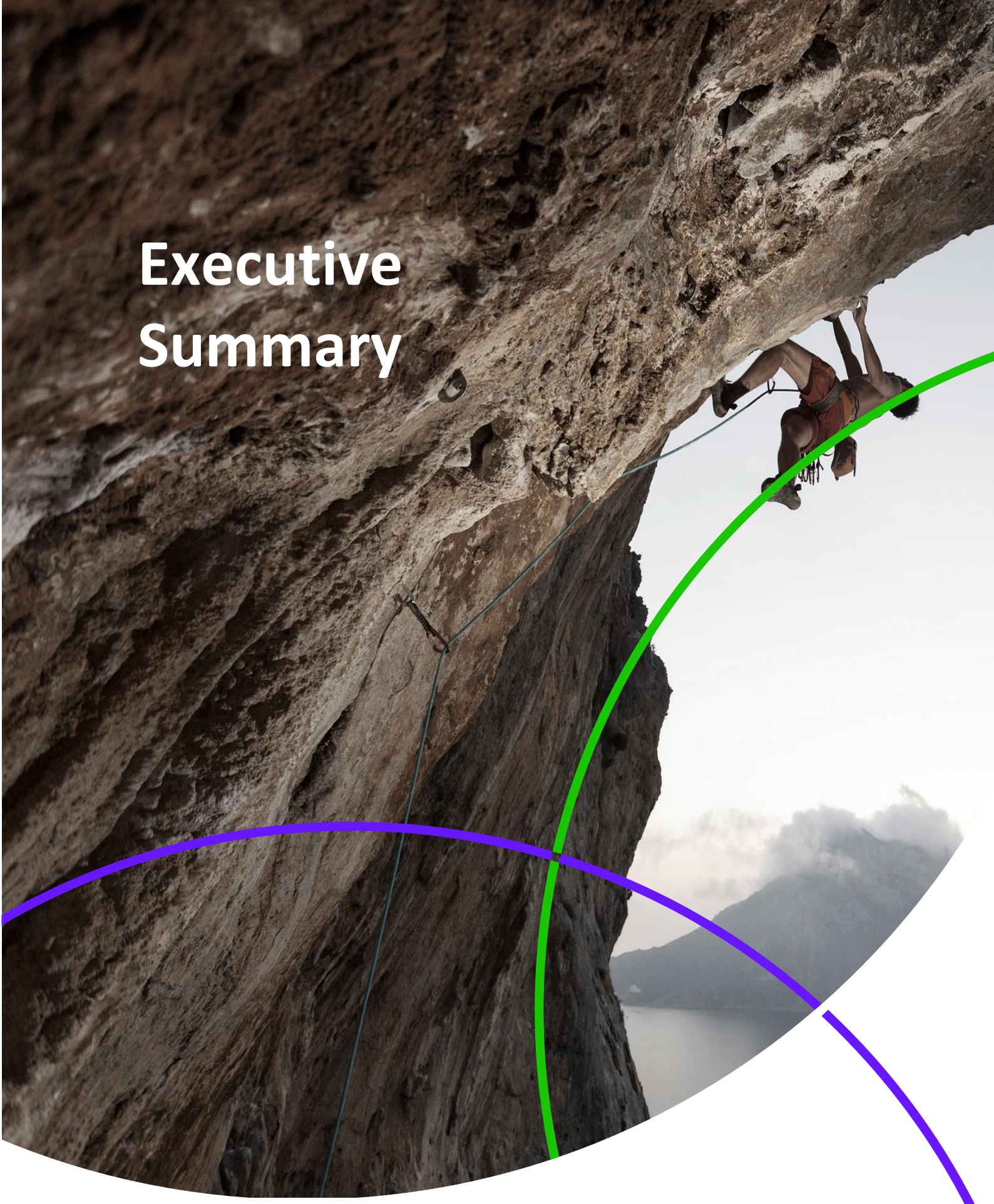
Dr. Cabrera is a data analyst at Clarivate Analytics, specializing in evaluating research activity and performance using bibliometric data. She has led and executed bibliometric and evaluation projects for government, academia, funders, and research organizations. She holds a Ph.D. in condensed matter physics from the Johns Hopkins University.

**Nan Ma, Ph.D.**
Senior Scientific Analyst

Dr. Nan Ma is a Senior Scientific Analyst at Clarivate Analytics and her primary role is as a bibliometric analysis expert. She serves in project management and project delivery capacity, supporting academic institutions, government agencies and funding organizations globally. She has over 10 years of experience in bibliometric analysis and research performance evaluation, and holds a Ph.D in Manangement Science.

Clarivate Analytics

# Executive Summary



Clarivate Analytics

# EXECUTIVE SUMMARY

*This report presents a bibliometric evaluation of the Engineering and Physical Sciences Research Council (EPSRC)-funded research output published between 2008 and 2017, which focuses on the productivity, impact, disciplinary foci, collaborations, and contribution of EPSRC-funded research to UK research.*

EPSRC research funding has been linked in the researchfish® platform to over 97,500 publications, some 92% of which were indexed in the Web of Science Core Collection. Those publications classified as papers (i.e., articles and reviews), which were published between 2008 and 2017, were selected as the basis for this study. Most of these papers were classified in disciplines that can be broadly defined as Chemistry, Computer Science, Energy & Fuels, Engineering, Materials Science, Mathematics, and Physics. However, EPSRC-funded research also generated a significant volume of papers in disciplines more commonly associated with medical, biological and health sciences, the arts, humanities and social sciences, and the environmental sciences, highlighting the breadth of research funded.

To ensure a fair benchmarking against the rest of the UK and beyond, the top 25 journal subject categories under which EPSRC-funded papers were classified were identified and used as the main criteria to produce regional and global datasets. This process, which ensured disciplinary analysis between EPSRC-funded papers and the regional and global comparators, resulted in a subset of 59,688 EPSRC papers, accounting for 72% of the selected dataset. This subset was used for the analysis presented in this report and the main findings are highlighted below.

## The role of EPSRC in the global and national Engineering and Physical Science Research

Papers from EPSRC-funded research accounted for over 1% of the global research output in relevant disciplines between 2008 and 2017, ranging from 1.1% in Mathematics and Engineering to 1.8% in Physics. These sets of papers were well-cited, with 86.5% of them receiving at least one citation during the study period. The overall normalized citation impact (NCI) of EPSRC-funded papers (NCI=1.60) was 60% higher than the global average, and the percentage of highly-cited papers (at the 10%, 5%, and 1% levels) exceeded the global averages by factors of 1.7, 1.8, and 2.0, respectively. The overall impact of EPSRC-funded papers was also stronger than that of the European countries and Brazil, Russia, India, and China (collectively known as the BRIC countries).

At the national level, EPSRC-funded papers contributed to over a quarter (27.4%) of the UK research output, ranging from 19.7% in Computer Science to 32.2% in Physics. The NCI of papers arising from EPSRC-funded research in relevant disciplines (ranging from 1.30 in Computer Science to 1.73 in Physics) was also higher than the NCI for the rest of the UK (i.e., papers not related to EPSRC-funded research), by up to 0.4 points in most cases. The largest difference in citation impact from research funded by EPSRC versus other UK research was observed in Materials Science, where EPSRC papers had an NCI=1.57 and the rest of the UK had an NCI=1.12.

The paper output funded by EPSRC was highly collaborative, with 46.8% of all papers arising from EPSRC-funded research identified as internationally collaborative. This value was higher than the global percentage (23.6%), but lower than the value for papers from the rest of the UK (63.9%). The

latter reflects the relatively high level of involvement by UK researchers in collaboration with EU-funded research. The USA, China, and European countries (particularly Germany and France) were identified as the top collaborators of EPSRC-funded papers. While the USA was the top collaborator in most disciplines, China was the top collaborator in Energy & Fuels and Engineering. Papers co-published with the top ten collaborating countries resulted in a positive impact on the overall EPSRC performance. Specifically, the NCI for these collaborative papers was higher than the overall NCI for EPSRC (NCI=1.60), ranging from NCI=1.79 for papers co-authored with researchers affiliated to a Chinese institution to NCI=2.27 for papers co-authored with researchers affiliated to a Swiss institution.

## EPSRC contributions to UK's strength areas

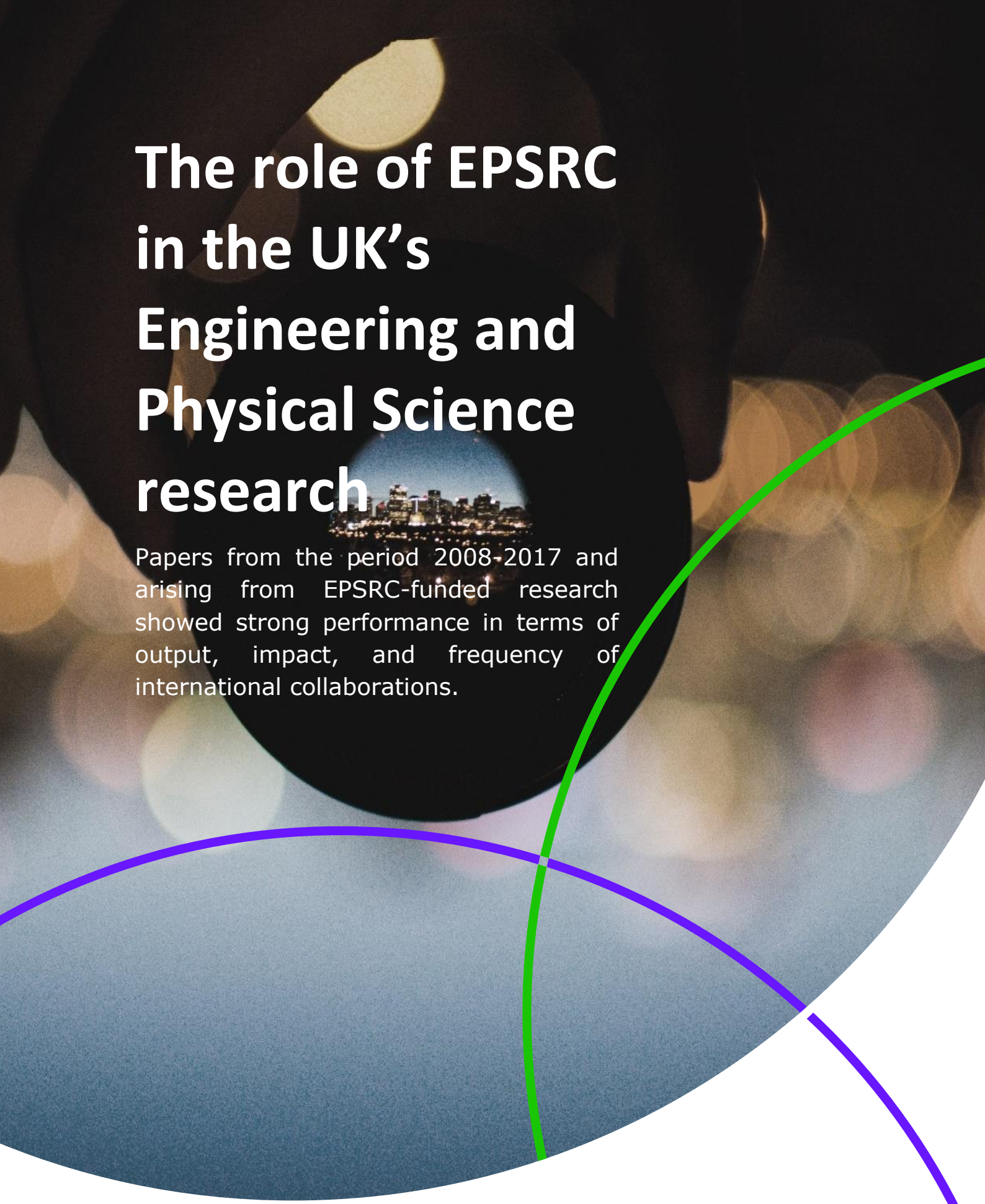
The Research Front Methodology[1] was used to identify Research Fronts where the UK played an important role in terms of co-authoring top 1% highly co-cited papers. This dataset served as the basis for producing clusters of "research topics" that would provide an overview of influential research from the UK based on frequently used keywords. Using this methodology, a total of 386 research topics were identified. Eight of these research topics, which included the largest number of papers funded by EPSRC, were identified and presented in this report.

The selected research topics included papers related to energy research (*Energy*) and domestic power (*Power*), self-assembly of systems (*Self-Assembly*), computer memory devices (*Antiferromagnetic Memory*), computer algorithms (*Monte Carlo Algorithms* and *Optimization Algorithms*), and communications (*Broadband Antennas* and *Small Cell Spectrum*). Three of the selected topics showed interdisciplinary classifications between Chemistry, Materials Sciences, and Physics (*Antiferromagnetic Memory*); Computer Science and Engineering (*Monte-Carlo Algorithms*); and Engineering and Energy & Fuels (*Power*). The EPSRC co-funded all influential papers related to *Monte-Carlo Algorithms* and *Broadband Antennas*. These eight topics are only a sample of the 140 research topics that included influential papers co-funded by EPSRC.

In summary, this report shows a strong performance by EPSRC-funded research in terms of scholarly output. It was shown that the impact of research funded by EPSRC exceeded that of the rest of the UK, European Countries, BRIC countries, and the global output, highlighting the globally influential research that EPSRC funds.

[1] See Appendix X for more details.

PAGE INTENTIONALLY LEFT BLANK



# The role of EPSRC in the UK's Engineering and Physical Science research

Papers from the period 2008-2017 and arising from EPSRC-funded research showed strong performance in terms of output, impact, and frequency of international collaborations.

Clarivate Analytics

# I. OVERALL PERFORMANCE OF EPSRC-FUNDED RESEARCH

EPSRC funds research in a broad range of areas which can be generally classified within the research areas of Physical Sciences and Engineering & Technology. However, EPSRC funding also results in a significant volume of contributions to areas outside of the "traditional" remit of EPSRC, for example in areas more commonly associated with Medical, Biological and Health Sciences, the Arts, Humanities and Social Sciences, and the Environmental Sciences. Using the researchfish® platform, EPSRC-funded researchers have confirmed references to over 97,500 uniquely identified journal publications arising from their EPSRC funding.[2] Of these, 82,589 papers (i.e., articles and reviews) published between 2008 and 2017 were identified in the Web of Science Core Collection (WoS). The dataset was further refined by selecting the subset of 59,688 papers[3] classified under the WoS journal subject categories in which EPSRC-funded research output was most frequently published. These WoS journal subject categories were mapped to the following seven disciplines[4]:

- Chemistry
- Computer Science
- Energy and Fuels
- Engineering
- Materials Sciences
- Mathematics
- Physics

Papers published in these mapped disciplines were also obtained for the following comparator groups for comparison with the EPSRC papers and to provide benchmarks against which to assess the overall performance of EPSRC-funded research:

| | |
|---|---|
| Rest of UK: | Refers to papers with UK authors, but which have not been attributed to specific instances of EPSRC funding. |
| EU-15: | Refers to Austria, Belgium, Denmark, Finland, France, Germany, Greece, Ireland, Italy, Luxembourg, Netherlands, Portugal, Spain, Sweden, and the UK |
| Rest of EU-15: | Refers to papers published by authors from affiliations within the EU-15 group, but excluding papers only attributed to UK authors. |
| EU-28: | Refers to Austria, Belgium, Bulgaria, Croatia, Cyprus, Czech Republic, Denmark, Estonia, Finland, France, Germany, Greece, Hungary, Ireland, Italy, Latvia, Lithuania, Luxembourg, Malta, Netherlands, Poland, Portugal, Romania, Slovakia, Slovenia, Spain, Sweden, and the UK |
| BRIC: | Refers to Brazil, Russia, India, and China. |
| World: | Refers to all papers in the mapped disciplines, irrespective of the authors' affiliation country. |

[2] https://www.researchfish.net/

[3] Throughout the remainder of this report, references to 'EPSRC-funded papers' are to this subset of 59,688 papers, unless otherwise specified

[4] See Appendix I for the Web of Science journal subject category-to-discipline mapping used for this comparison.

This subset accounted for 72% of the total number of papers published between 2008-2017 which researchers have attributed to their EPSRC funding using the researchfish outcomes collection platform. Papers not included in this subset correspond to those classified under Web of Science journal subject categories that contain a smaller number of publications funded by EPSRC. Although over a quarter of EPSRC papers are excluded from the benchmarking analysis, this method ensures disciplinary alignment to the regional and global comparisons.

As previously stated, the global baselines calculated for this analysis are based on the Web of Science journal subject categories schema. In this schema, papers can be assigned to one or more category, as opposed to other schemas (e.g., Essential Science Indicators fields), in which a paper is assigned to a single category. A quirk of the way these baselines are calculated (whole counting of categories for papers in more than one subject category) and the way the normalized citation impact (NCI) is calculated (fractional counting of categories for papers in more than one subject category) results in the NCI of the world not being equal to one exactly. To ease the comparison of the NCI metric from EPSRC papers against the selected comparators and the world in this report, the NCI for EPSRC papers and all comparators have been normalized against the global NCI, such that the global NCI in all cases equals 1.00. Appendix II lists the precise global NCI for all aggregates presented in this study.

Figure 1 shows a bibliometric summary of EPSRC papers, as compared with the selected comparators. The overall number of world papers published in the disciplines relevant to EPSRC was 4,286,994 and research funded by EPSRC contributed to 1% of this output (as represented by the area of the bubble in the figure). Papers funded by the EPSRC had the greatest NCI (1.60), as noted by the position of the EPSRC bubble in the vertical axis. Papers from the rest of the UK had the second highest NCI (1.26). All European comparators showed an NCI above the overall global average (NCI between 1.1 and 1.2), while the BRIC countries had an NCI below the global average (0.95). The largest percentage of cited papers (86.5%) was observed for the collection of EPSRC-funded papers, as noted by its position in the horizontal axis. All comparators had over 80% of their papers cited, which was above the global baseline (79.1%), except for the BRIC countries, which collectively had 76.2% of their papers cited.

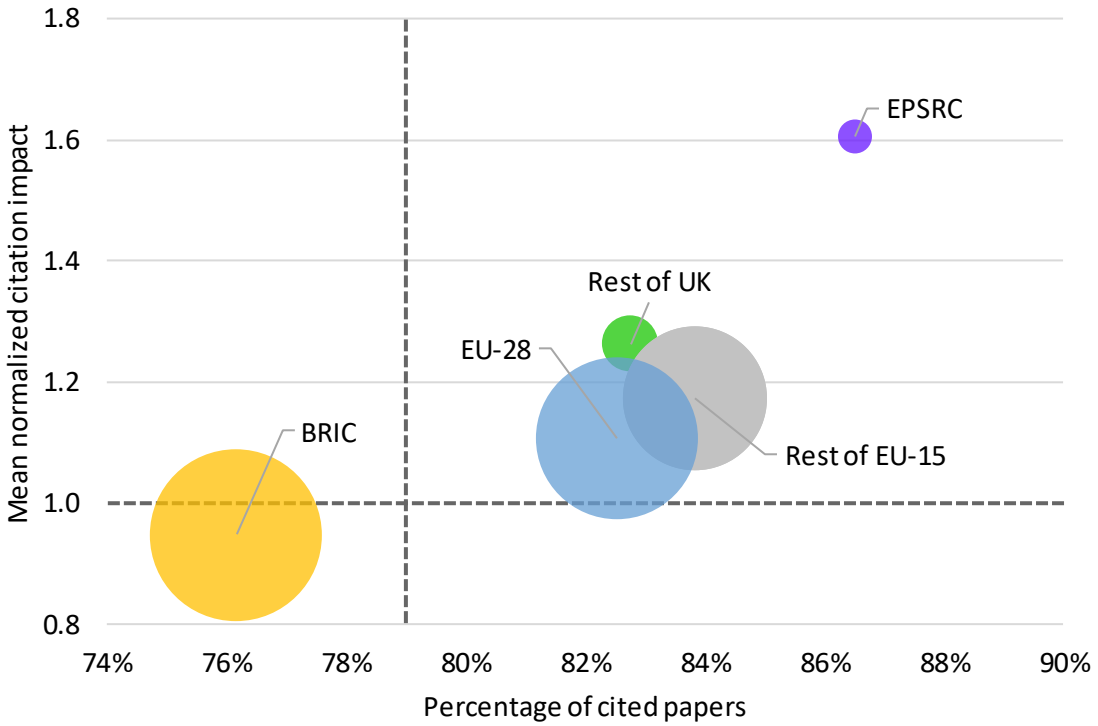



*Figure 1 Overall percentage of cited papers (x-axis), normalized citation impact (y-axis), and number of papers (area of bubbles) of EPSRC-funded papers and selected comparators, 2008-2017. The horizontal and vertical dashed lines correspond to the percentage of global cited papers and global normalized citation impact, respectively.*

Figure 2 shows the percentage of internationally collaborative papers from EPSRC and selected comparators. The world dataset showed that 23.6% of all papers were internationally collaborative. The EPSRC-funded papers had a larger percentage of international collaborations, with 46.8% of its papers listing affiliations from more than one country. The rest of the UK group showed the largest percentage of internationally collaborative papers (63.9%). It was also noted that the European Union was a top international funder of this group. All other groups had international collaborations in around half of their papers, except for the BRIC countries, which showed a percentage of international collaborations identical to the global dataset.

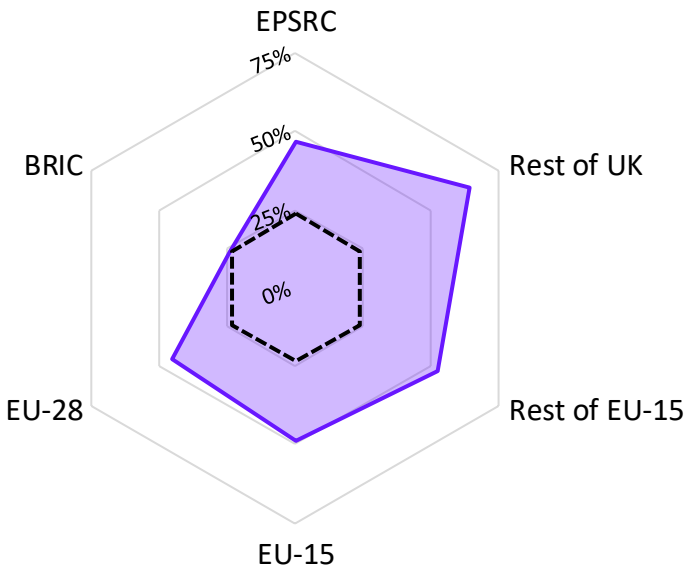



*Figure 2 Percentage of internationally collaborative papers from EPSRC and selected comparators, 2008-2017. The dashed line corresponds to the global percentage of internationally collaborative papers.*

Highly cited papers generally indicate scientific excellence, as only the most highly cited papers would make the top 10%, 5%, or 1% in their respective field, year, and document type[5]. Figure 3 - Figure 5 show the percentage of EPSRC and selected comparators' papers in the top 10%, top 5%, and top 1% highly-cited global papers. EPSRC showed the highest percentages of papers in all three classifications, with 18.6%, 9.9%, and 2.2% of its papers in each classification, respectively. The rest of the UK and European comparators had percentages of highly cited papers slightly lower than EPSRC, but higher than the global percentages (i.e., 10.7%, 5.5%, and 1.1%). Note that the global percentages are not necessarily exactly 10%, 5%, or 1%, due to the way this indicator is calculated.[6] The BRIC countries presented the lowest percentages of highly cited papers, values which were closer to the global percentages.

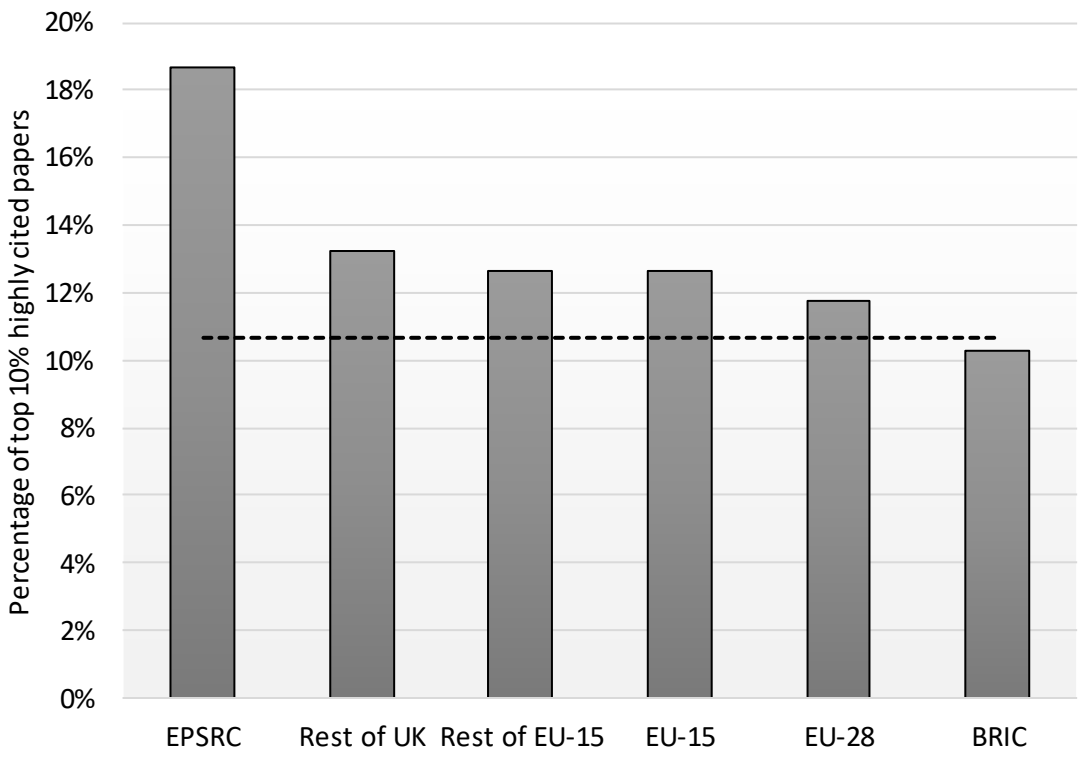



*Figure 3 Overall percentage of EPSRC and selected comparators' papers occupying the world's top 10% highly cited papers, 2008-2017. The dashed line corresponds to the global percentage of top 10% highly cited papers.*

[5] The percentage of highly cited indicator can be used in conjunction with other indicators to provide a more complete picture of performance. See the InCites Handbook for more details: http://help.prod-incites.com/inCites2Live/indicatorsGroup/aboutHandbook.html .

[6] Please refer to Appendix X for details on this metric.

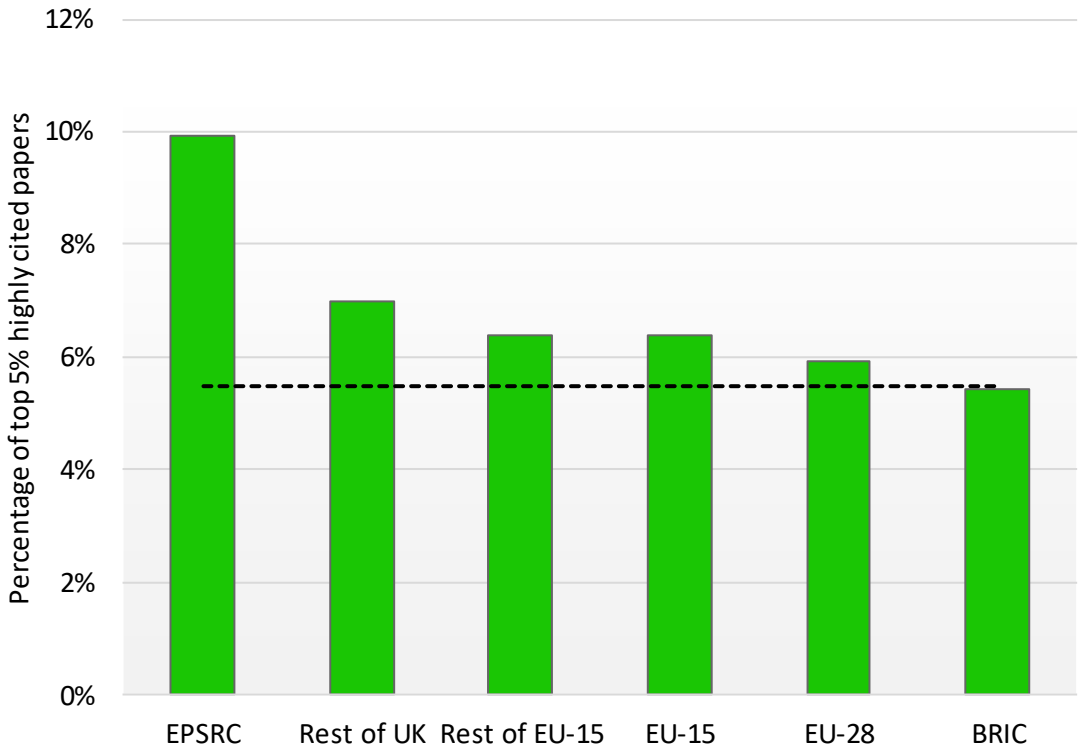



*Figure 4 Overall percentage of EPSRC and selected comparators' papers occupying the world's top 5% highly cited papers, 2008-2017. The dashed line corresponds to the global percentage of top 5% highly cited papers.*

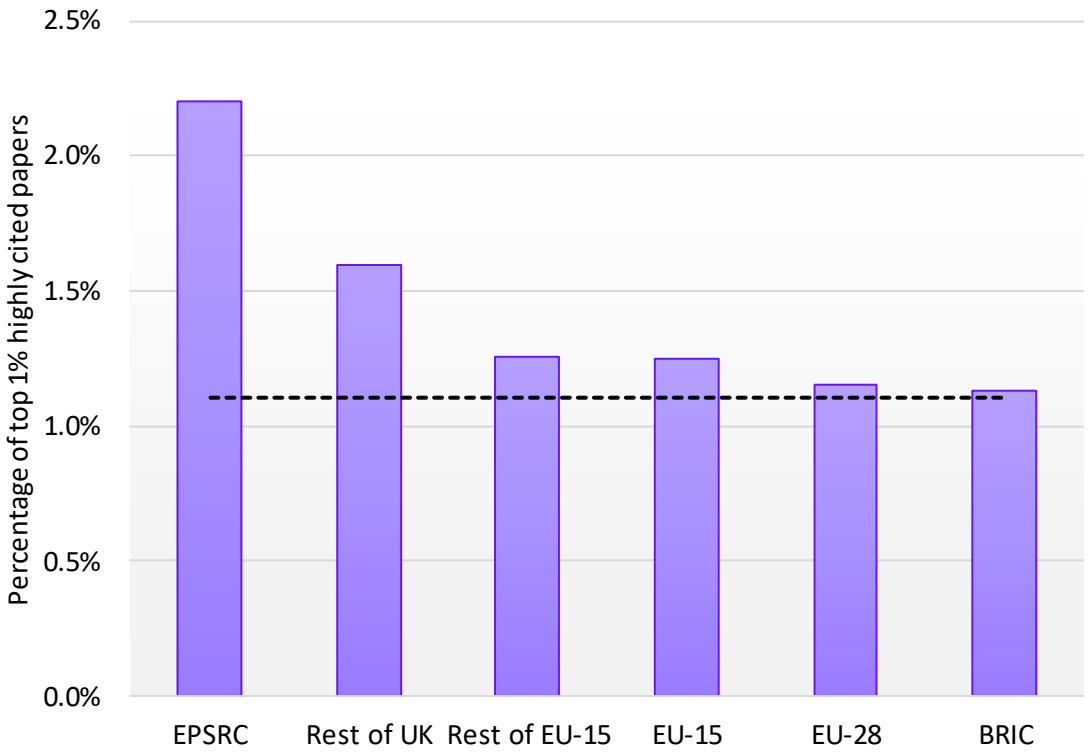



*Figure 5 Overall percentage of EPSRC and selected comparators' papers occupying the world's top 1% highly cited papers, 2008-2017. The dashed line corresponds to the global percentage of top 1% highly cited papers.*

Clarivate Analytics

## II. ANALYSIS OF EPSRC-FUNDED RESEARCH BY PUBLICATION YEAR

The number of EPSRC-funded papers by publication year within the dataset used for this study is shown in Figure 6, which should be interpreted with care. Specifically, the apparent annual increase in volume from 2008 to 2016 is an artefact reflecting improvements in the attribution of publications to EPSRC funding, and the slight decrease of publications in 2017 is also likely to be an artefact related to the processing date used for the dataset utilized in this analysis (see Appendix X for more details). However, in each year the volume is sufficient to allow robust analysis showing that the normalized citation impact (black dots connected by a black line) was consistently above the global average during the study period, varying from NCI = 1.56 in 2015 to NCI = 1.69 in 2008. Due to the limited time for papers from 2017 to accumulate citations, the NCI for this year tends to fluctuate more than expected and the value is not considered for this analysis. Consequently, all results for 2017 are presented, but in lighter text, in all figures from this section.

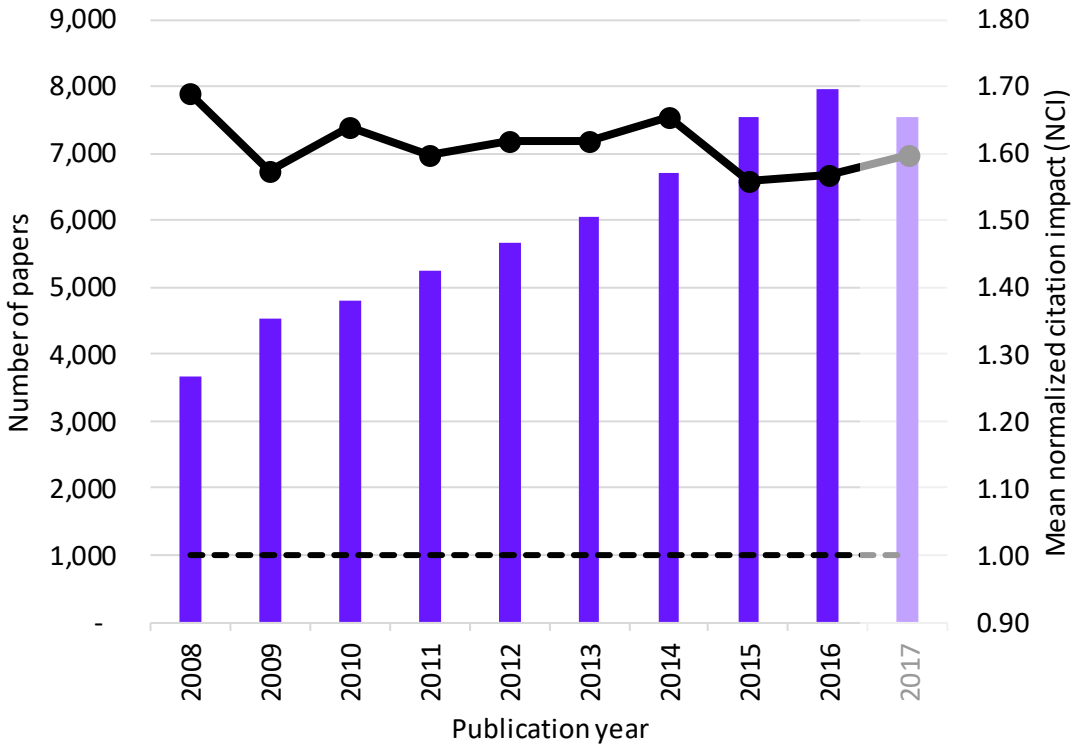



*Figure 6 Number of EPSRC-funded papers (purple bars) and mean normalized citation impact (black solid curve), 2008-2017. The dashed line corresponds to the global mean normalized citation impact.*

Most of the EPSRC-funded papers (over 90%) were cited at least once, between 2008 and 2015, as shown in Figure 7. The decrease of cited papers in more recent years is typical of what is observed for global output, and it is also seen for the EPSRC-funded papers, given that papers have had less time to be cited.

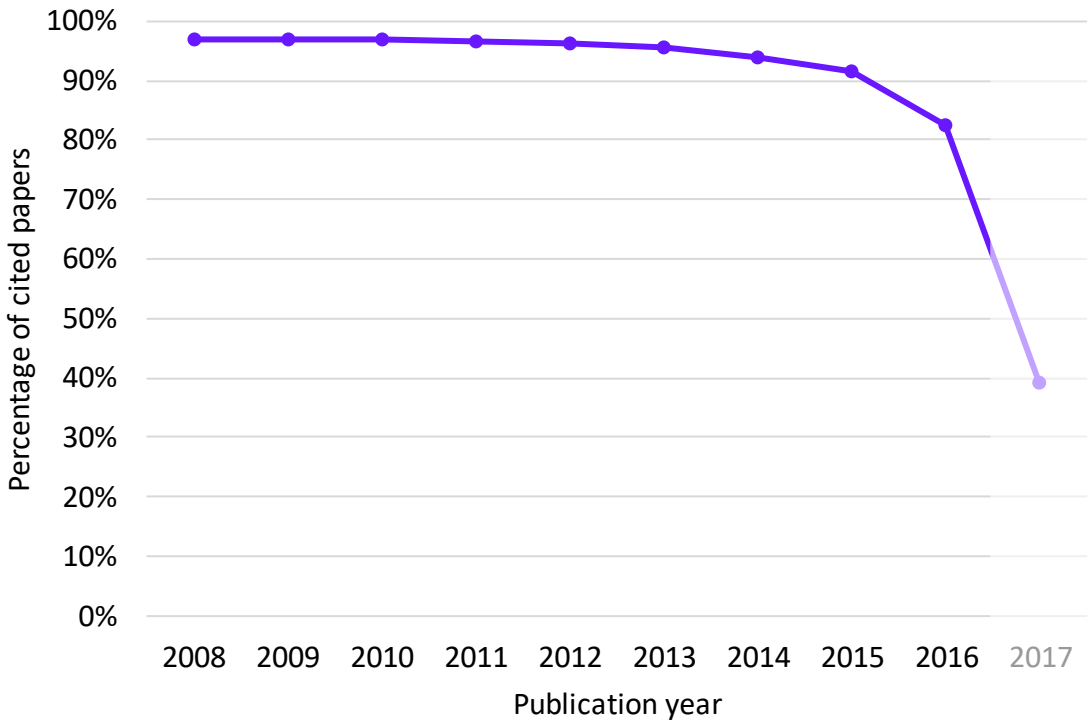



*Figure 7 Percentage of EPSRC-funded papers that have been cited at least once, 2008-2017.*

Figure 8 shows the percentage of EPSRC-funded papers that included at least one international co-author. The percentage of internationally collaborative papers steadily increased from 40.8% in 2008 to 49.2% in 2016.

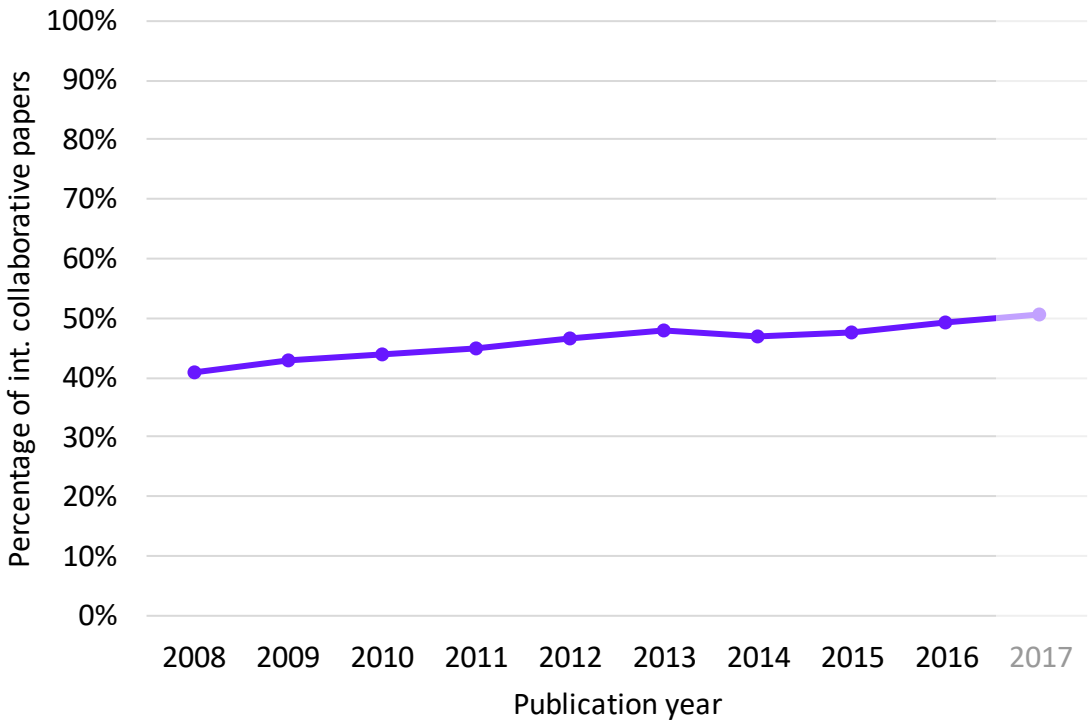



*Figure 8 Percentage of EPSRC-funded papers that are internationally collaborative, 2008-2017.*

The percentage of EPSRC highly-cited papers (at the top 10%, top 5%, and top 1% levels) was generally twice that of the global average between 2008 and 2017, as shown in Figure 9 - Figure 11. These trends suggest the high relevance of research topics chosen by EPSRC-funded researchers to the scientific community.

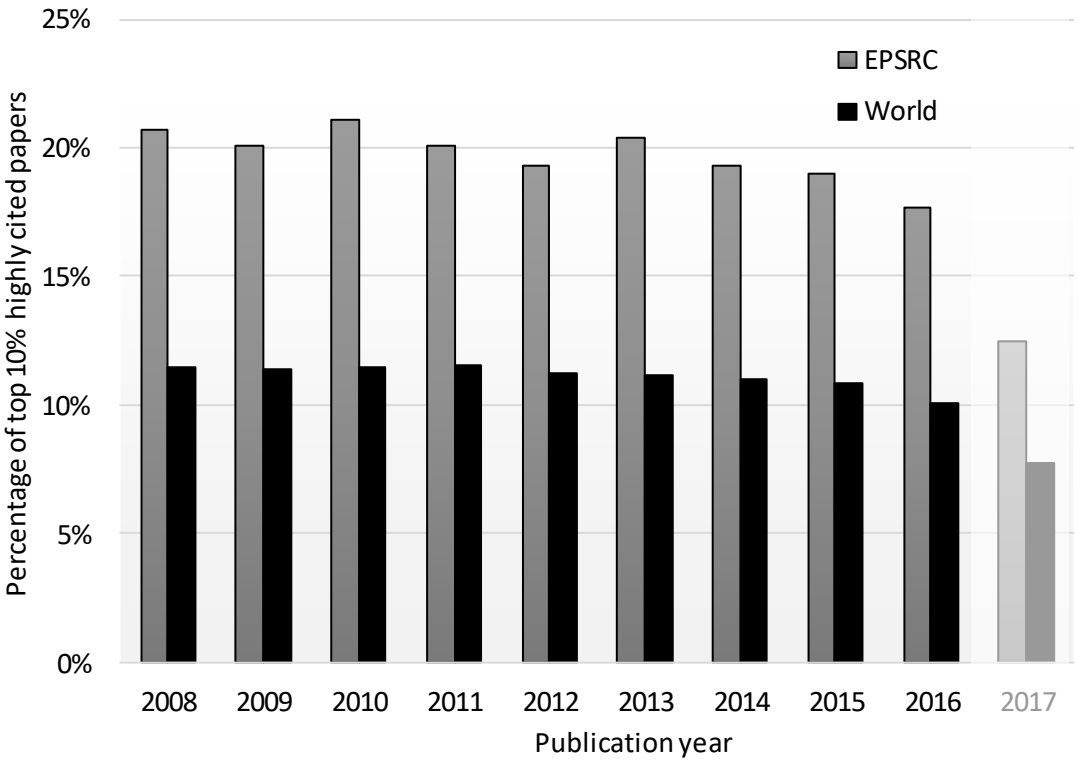



*Figure 9 Percentage of EPSRC-funded papers occupying the world's top 10% highly cited papers, 2008-2017.*

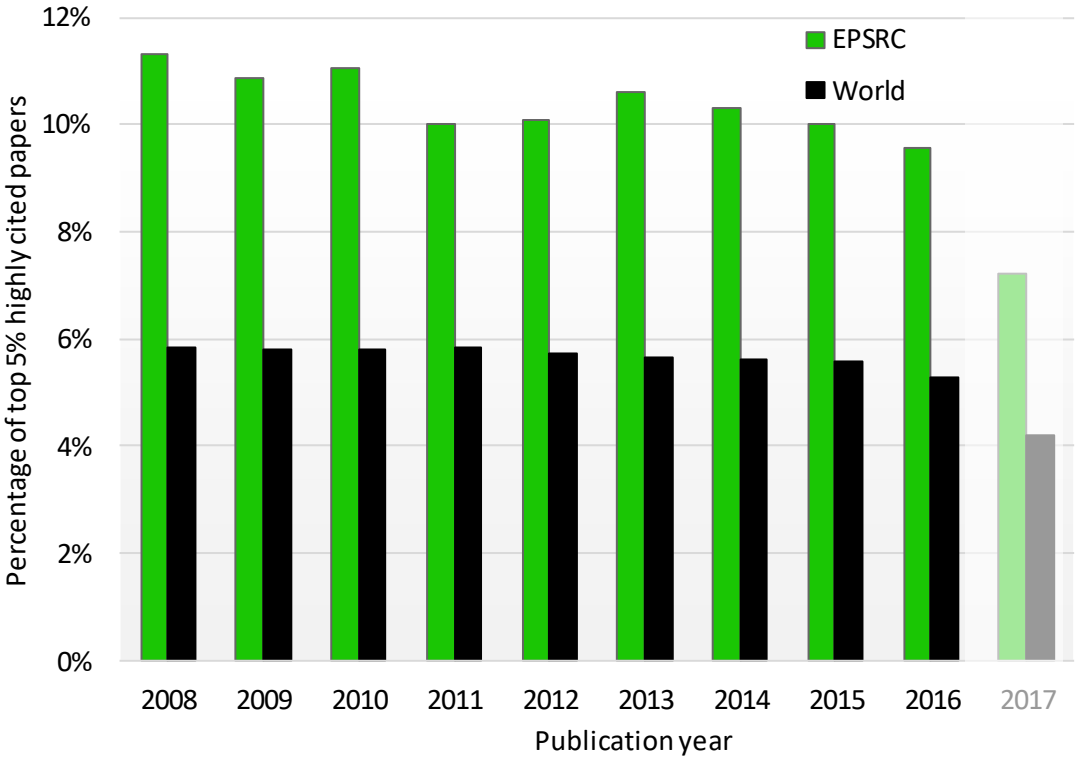



*Figure 10 Percentage of EPSRC-funded papers occupying the world's top 5% highly cited papers, 2008-2017.*

Clarivate Analytics

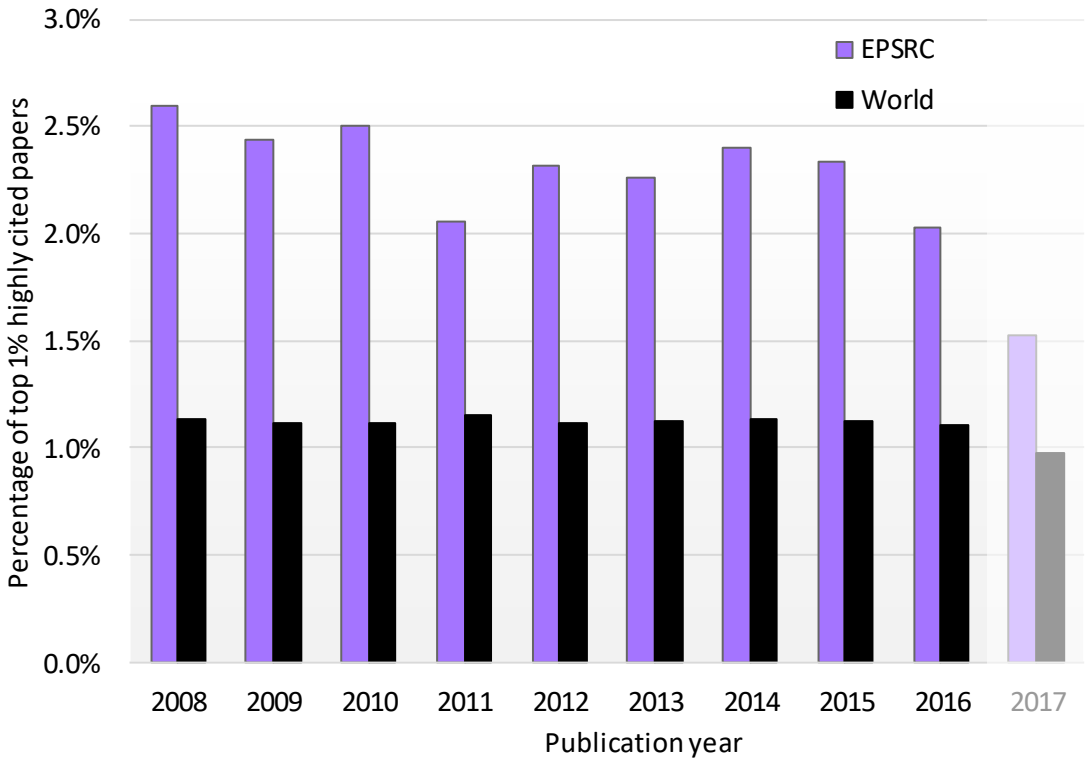



*Figure 11 Percentage of EPSRC-funded papers occupying the world's top 1% highly cited papers, 2008-2017.*

# Disciplinary Analysis of EPSRC funded research

With a sizeable contribution to UK's research output in Physics, Chemistry, and Materials Science, and the highest citation impact observed amongst all comparators, EPSRC-funded papers excelled in their core publishing disciplines.

Clarivate Analytics

## III. DISCIPLINARY ANALYSIS OF EPSRC-FUNDED RESEARCH

It is useful to understand the role EPSRC-funded researchers played in the global and regional context, in terms of output and impact of the research published in peer-reviewed journals. Figure 12 shows the percentage of global papers resulting from research funded by EPSRC. A non-negligible percentage of global papers involved researchers funded by EPSRC, from 1.1% in Mathematics and Engineering to 1.8% in Physics. The number and percentage of world papers funded by EPSRC by discipline can be found in Table 1.

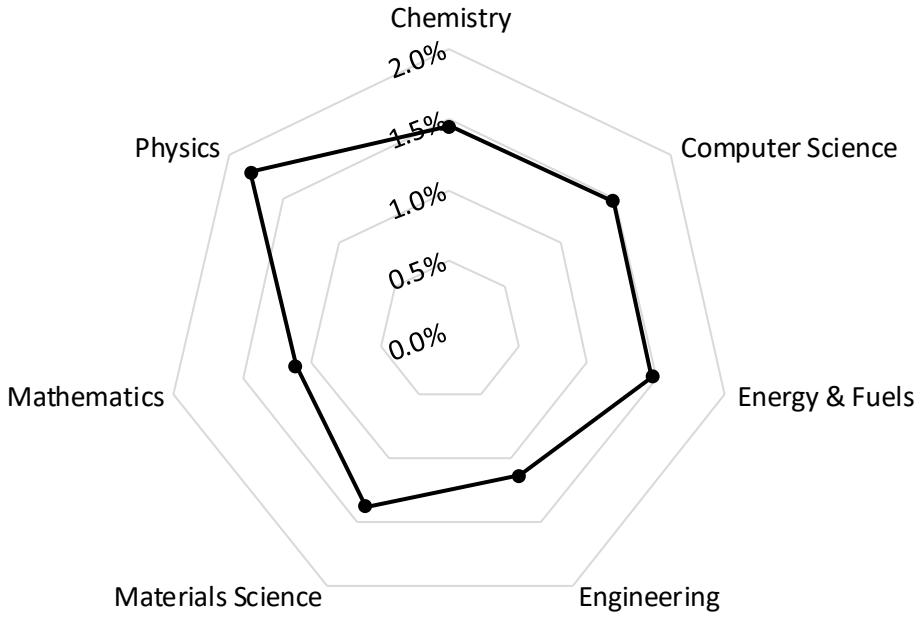



*Figure 12 Overall percentage of world papers funded by EPSRC by discipline, 2008-2017.*

*Table 1 Number and percentage of world papers funded by EPSRC by discipline, 2008-2017.*

| Discipline | Number of papers | Percentage of world papers |
|---|---|---|
| **Chemistry** | 19,251 | 1.4% |
| **Computer Science** | 4,964 | 1.5% |
| **Energy & Fuels** | 3,404 | 1.5% |
| **Engineering** | 13,260 | 1.1% |
| **Materials Science** | 10,916 | 1.4% |
| **Mathematics** | 4,328 | 1.1% |
| **Physics** | 24,158 | 1.8% |

A comparison between the share of papers across the disciplines published by EPSRC-funded researchers and selected comparators during the study period is presented in Table 2, which shows that, relative to the global distribution, EPSRC had a broadly similar distribution, but with a higher proportion in Physics, and lower proportions in Engineering and Mathematics. Compared with the rest of the UK, Physics, Chemistry, and Materials Science had higher proportions of EPSRC-funded papers, whilst most of the others, especially Engineering, had lower proportions.

*Table 2 Overall paper output of EPSRC-funded papers and selected comparators by discipline, 2008-2017. The disciplines are listed by decreasing order of percentage of EPSRC papers.*

| Discipline | EPSRC | Rest of UK | EU-15 | Rest of EU-15 | EU-28 | BRIC | World |
|---|---|---|---|---|---|---|---|
| Physics | 30.1% | 23.9% | 25.5% | 25.8% | 25.0% | 24.3% | 24.0% |
| Chemistry | 24.0% | 20.5% | 22.5% | 22.4% | 22.9% | 26.0% | 23.8% |
| Engineering | 16.5% | 23.0% | 20.0% | 19.7% | 20.1% | 19.2% | 20.9% |
| Materials Science | 13.6% | 11.6% | 12.4% | 12.5% | 12.4% | 15.7% | 14.2% |
| Computer Science | 6.2% | 9.5% | 7.5% | 7.4% | 7.2% | 4.5% | 6.0% |
| Mathematics | 5.4% | 7.1% | 8.1% | 8.3% | 8.6% | 6.0% | 6.9% |
| Energy & Fuels | 4.2% | 4.4% | 4.0% | 3.9% | 3.8% | 4.1% | 4.1% |

Figure 13 shows the normalized citation impact of papers published by EPSRC-funded researchers and selected comparators by discipline during the study period. EPSRC had the highest citation impact in all disciplines, ranging from NCI=1.30 in Computer Science to NCI=1.73 in Physics. EPSRC's NCI surpassed that of the rest of the UK's research not funded by them, by up to 0.4 points in most cases. One exception is noted in Computer Science, where NCI = 1.3 for both groups.

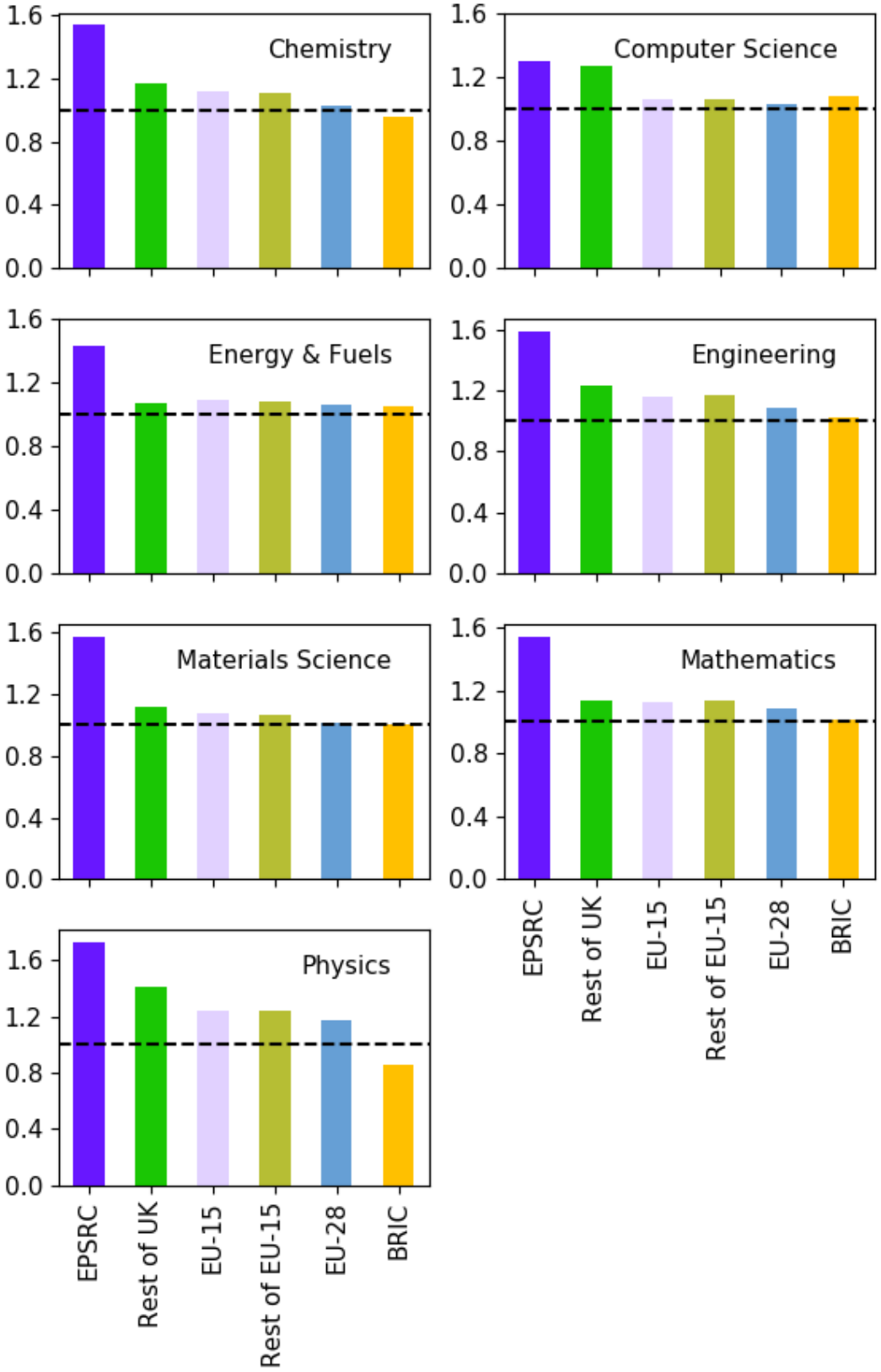



*Figure 13 Overall normalized citation impact of EPSRC-funded papers and selected comparators by discipline, 2008-2017. The dashed line corresponds to the global normalized citation impact.*

Clarivate Analytics

Figure 14 shows a summary of bibliometric indicators for papers published between 2008 and 2017, resulting from research funded by EPSRC by discipline. Specifically, the percentage of UK papers (horizontal axis), mean normalized citation impact (vertical axis), and number of papers per discipline (bubble area) are presented. At the national level, EPSRC-funded researchers collectively contributed to an average of over a quarter of UK papers (27.4%) in all disciplines presented in this section, as noted by the vertical dotted line. The most notable contributions to the UK paper outputs were observed in Physics (32.2%), as well as Chemistry and Materials Science (30.7%). The lowest contribution of UK papers funded by EPSRC was observed in Computer Science (19.7%). The collective paper output by disciplines included a group of inherently interdisciplinary papers which have been counted in each of the relevant disciplines. Focused connections between Physics and Chemistry were observed in 2,438 of these papers and broader connections between Materials Science, Chemistry, Nanoscience & Nanotechnology, and Physics were observed in 986 papers. Other interdisciplinary connections were observed in smaller groups of papers. The role of interdisciplinarity in exceptional research will be highlighted in Section IV of this report, where examples of EPSRC contributions to UK's strength areas are presented.

The normalized citation impact by discipline ranged between NCI=1.30 for Computer Science and NCI=1.73 for Physics, all exceeding the global NCI (horizontal dotted line) by 30% or more. The volume of output and impact of UK papers funded by EPSRC by discipline can be found in Table 3.

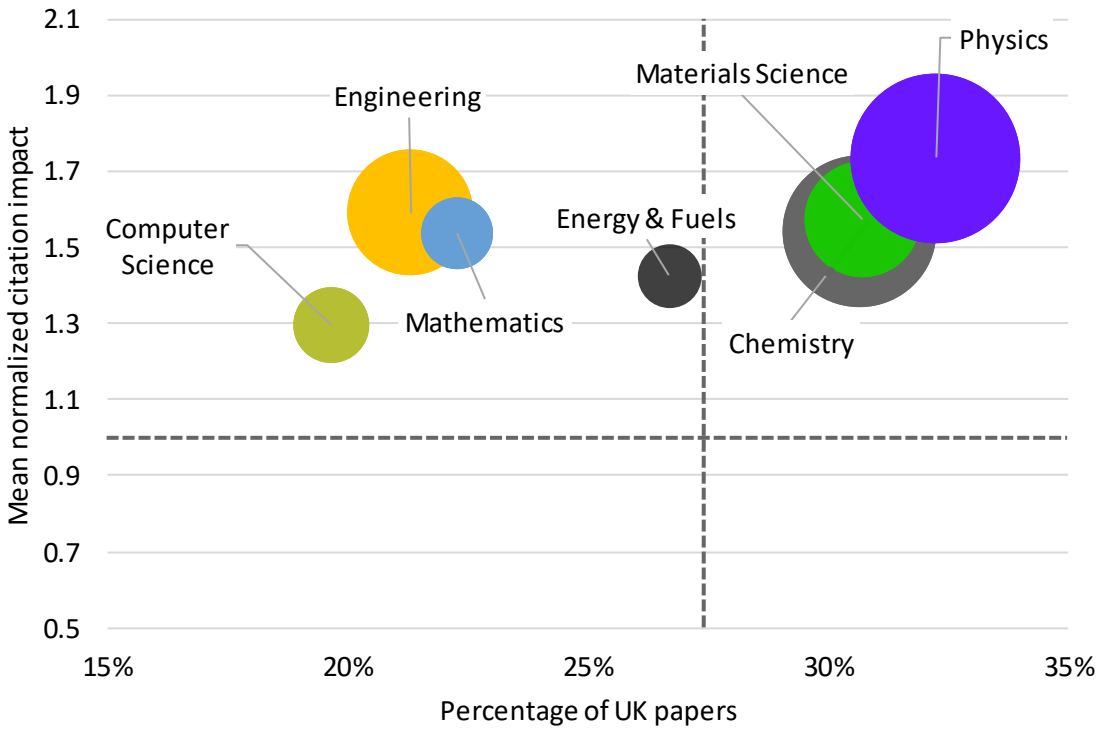



*Figure 14 Overall percentage of UK papers (horizontal axis), mean normalized citation impact (vertical axis), and number of papers (area of bubbles) funded by EPSRC by discipline, 2008-2017. The horizontal and vertical dashed lines correspond to the global mean normalized citation impact and the overall percentage of UK papers funded by EPSRC, respectively.*

*Table 3 Volume of output and impact of UK papers funded by EPSRC by discipline, 2008-2017.*

| Discipline | Number of papers | Percentage of UK papers | Mean NCI |
|---|---|---|---|
| **Chemistry** | 19,251 | 30.7% | 1.54 |
| **Computer Science** | 4,964 | 19.7% | 1.30 |
| **Energy & Fuels** | 3,404 | 26.7% | 1.43 |
| **Engineering** | 13,260 | 21.3% | 1.59 |
| **Materials Science** | 10,916 | 30.7% | 1.57 |
| **Mathematics** | 4,328 | 22.3% | 1.54 |
| **Physics** | 24,158 | 32.2% | 1.73 |

Figure 15 shows the overall percentage of EPSRC-funded papers that were cited at least once during the study period. All disciplines had a percentage of cited papers higher than the global percentages, as noted in Table 4. Chemistry had the highest percentage of cited papers (91.6%), while Mathematics had the lowest amount of cited papers with 72.9%. The largest margins between the percentage of EPSRC papers cited versus global papers cited were observed in Physics (8.5%), Engineering (8.2%), and Mathematics (7.3%). Computer Science was the only discipline showing a percentage of cited papers almost equal to the global value (76.6% versus 76.5%).

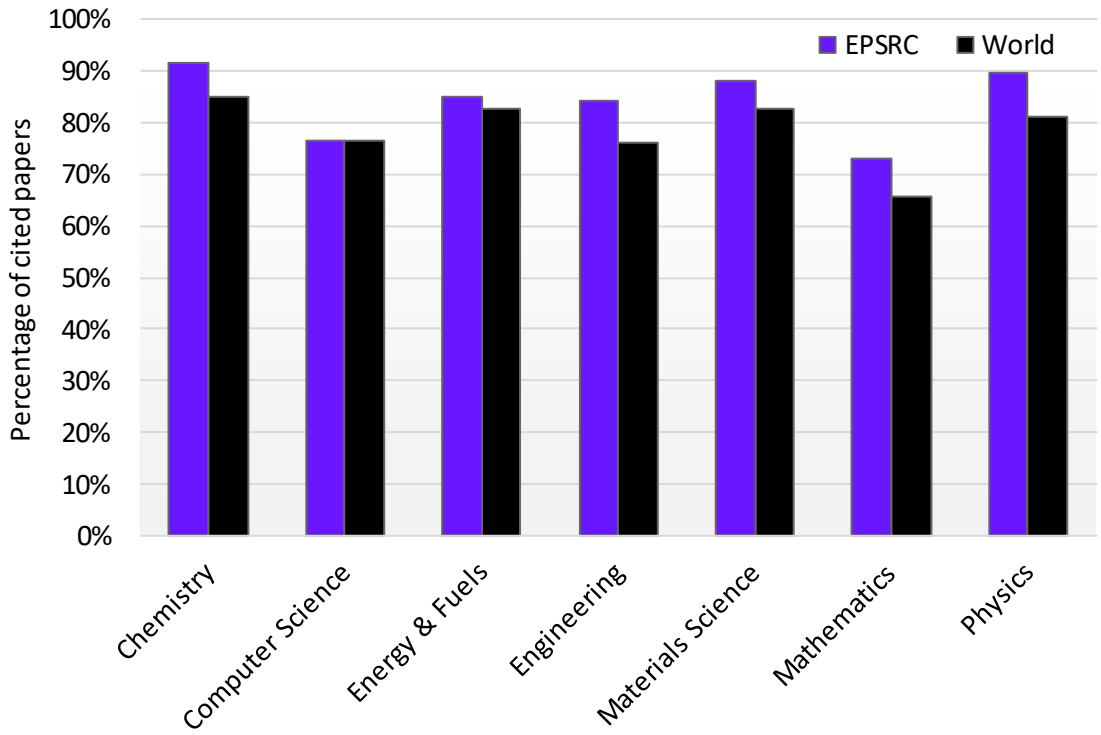



*Figure 15 Overall percentage of EPSRC-funded papers that were cited at least once by discipline, 2008-2017.*

*Table 4 Percentage of EPSRC-funded papers vs. world papers that were cited at least once by discipline, 2008-2017.*

| Discipline | EPSRC | World |
|---|---|---|
| Chemistry | 91.6% | 85.0% |
| Computer Science | 76.6% | 76.5% |
| Energy & Fuels | 84.9% | 82.8% |
| Engineering | 84.2% | 76.0% |
| Materials Science | 88.3% | 82.6% |
| Mathematics | 72.9% | 65.6% |
| Physics | 89.7% | 81.2% |

Figure 16 shows the percentage of collaborations between EPSRC-funded researchers from the UK and authors from other countries by discipline during the study period, as compared to the rest of the UK and the world. The highest percentage of collaborations were observed in Mathematics and Physics, with 54.8% and 53.5%, respectively. The least internationally collaborative discipline was Energy & Fuels, with 35.6% of its papers including one or more co-author from outside the UK The level of international collaboration on EPSRC papers in all disciplines presented, as measured by the countries of affiliation reported by authors, was consistently lower than that observed for the rest of the UK, reflecting the relatively high level of investment by UK researchers in collaboration with EU-funded research. It must be noted, however, that EPSRC papers included significantly more international collaborations than the global trends in all disciplines presented in this study, as noted in Table 5.

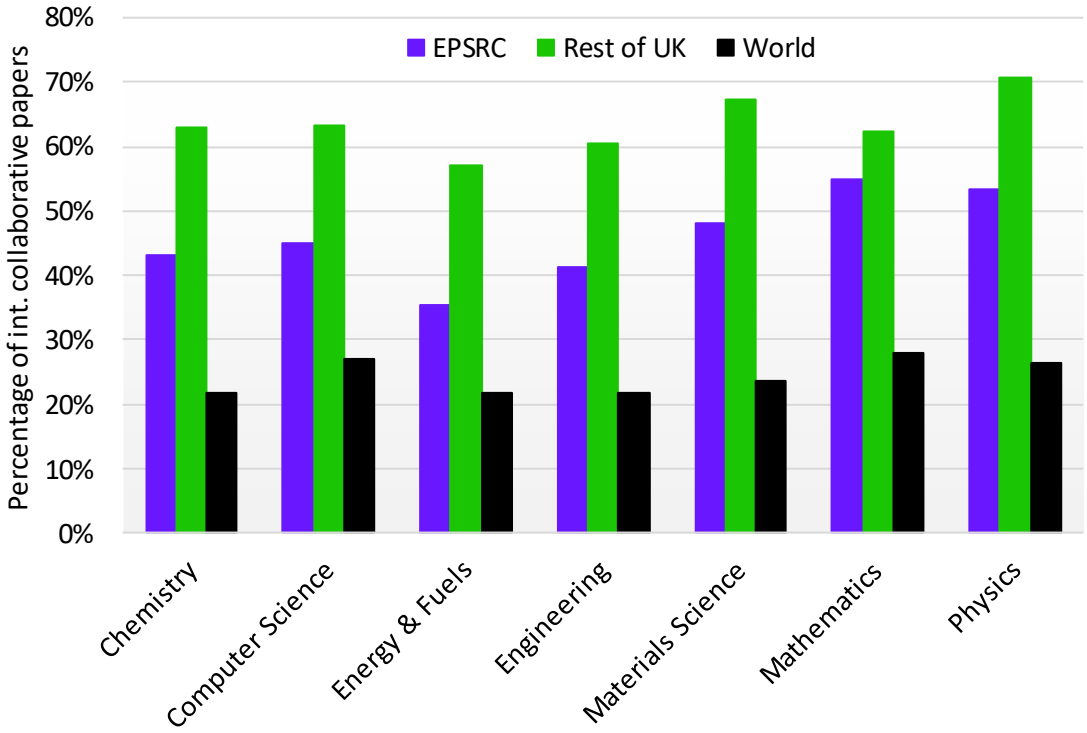



*Figure 16 Overall percentage of EPSRC-funded papers, rest of UK papers, and world papers that were internationally collaborative by discipline, 2008-2017.*

*Table 5 Percentage of EPSRC-funded papers vs. the rest of the UK and world papers that were internationally collaborative by discipline, 2008-2017.*

| Discipline | EPSRC | Rest of UK | World |
|---|---|---|---|
| Chemistry | 43.2% | 62.9% | 21.8% |
| Computer Science | 45.1% | 63.2% | 27.2% |
| Energy & Fuels | 35.5% | 57.2% | 21.7% |
| Engineering | 41.3% | 60.5% | 21.7% |
| Materials Science | 48.1% | 67.4% | 23.7% |
| Mathematics | 54.8% | 62.2% | 27.9% |
| Physics | 53.5% | 70.7% | 26.3% |

Figure 17 - Figure 19 show the overall percentage of EPSRC-funded papers occupying the world's top 10%, 5%, and 1% highly cited papers by discipline during the study period, as compared with the rest of the UK, EU-28, and the world. Overall, EPSRC-funded papers had the highest percentages of top 10% highly-cited papers, as compared with the selected comparators. One quarter (25.6%) of all papers in Energy and Fuels were within the world's top 10% highly cited papers, as compared to 17.6% for the world (see Figure 17). This discipline also showed the highest percentages of top 5% and 1% highly-cited papers within the EPSRC portfolio, i.e., 14.0% and 3.5% were part of the world's top 5% and top 1% highly cited, respectively, as observed in Figure 18 and Figure 19. Mathematics had the lowest percentage of highly cited papers in the EPSRC portfolio, albeit above the expected values, with 15.3%, 8.5%, and 1.9% of its papers contributing to the world's top 10%, top 5%, and top 1% highly cited papers, respectively. These percentages were higher than the global percentages in this discipline (i.e., 9.5%, 5.1%, and 1.1%, respectively), as well as the percentages for the rest of UK and EU-28, as noted in Table 6 through Table 8. Appendix IV lists the research output and impact of EPSRC and selected comparators by discipline during the study period. A more granular analysis is presented in Appendix V and Appendix VI, which list the research output and impact of EPSRC and selected comparators by Web of Science journal subject categories.

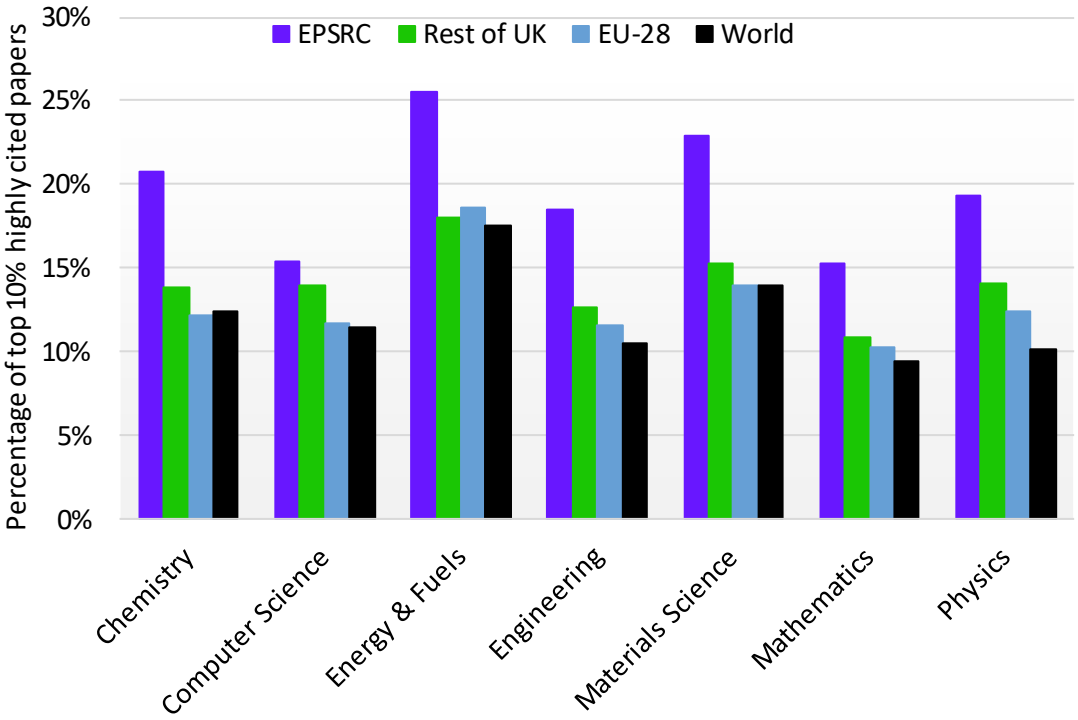



*Figure 17 Overall percentage of papers occupying the world's top 10% highly cited papers by group and discipline, 2008-2017.*

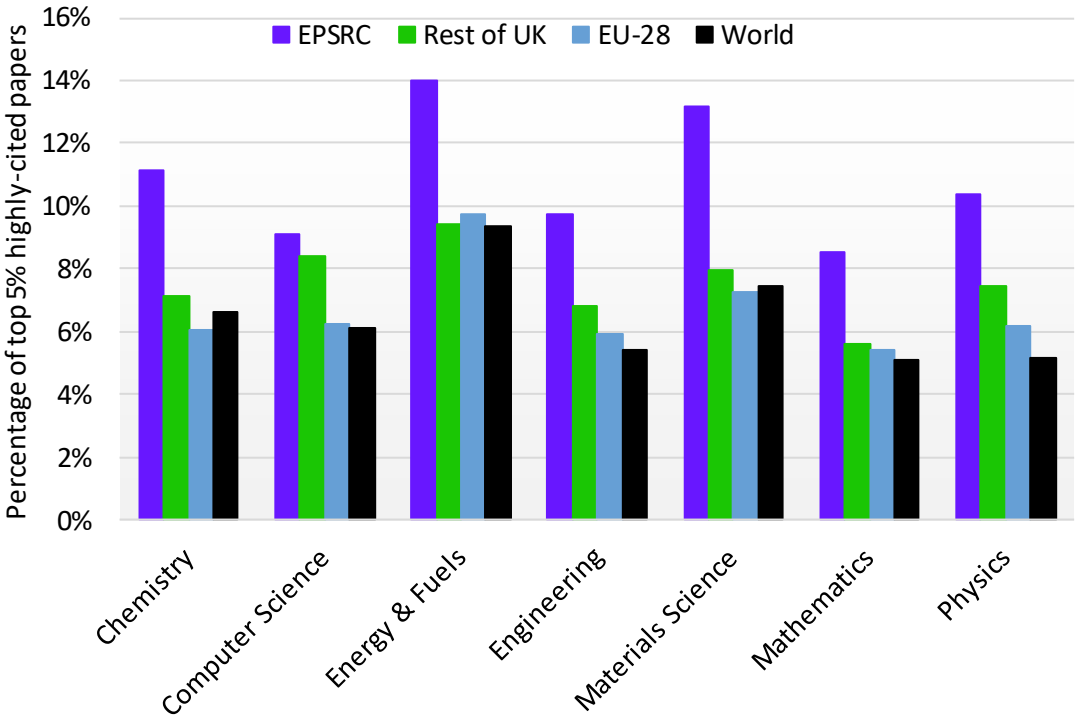



*Figure 18 Overall percentage of papers occupying the world's top 5% highly cited papers by group and discipline, 2008-2017.*

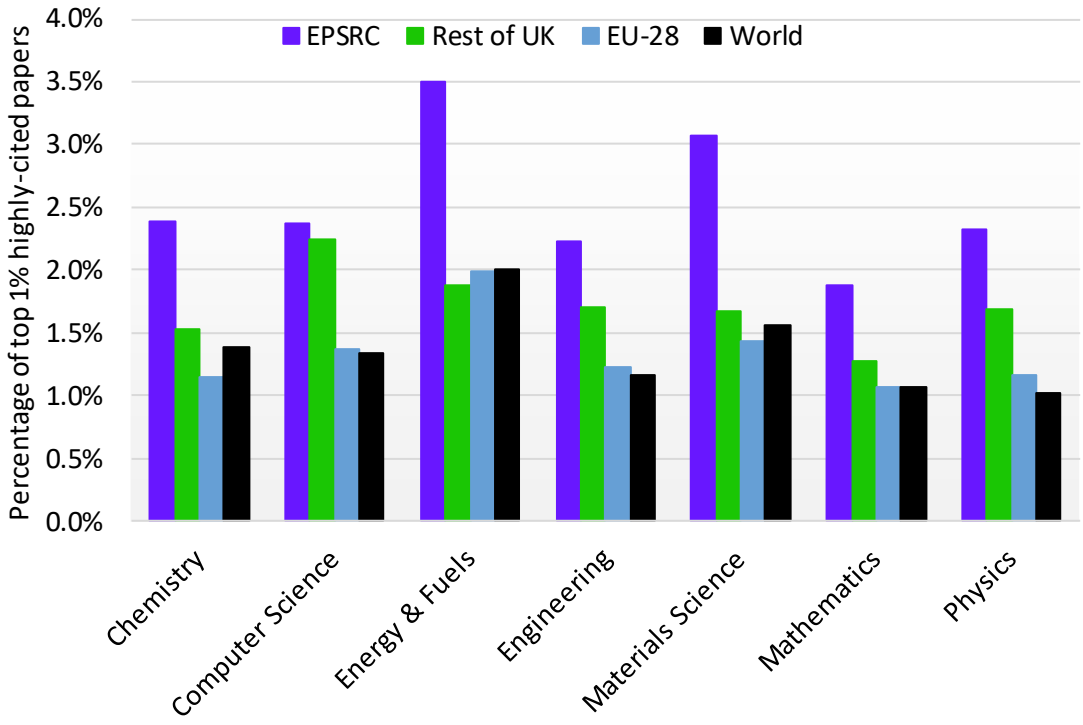



*Figure 19 Overall percentage of papers occupying the world's top 1% highly cited papers by group and discipline, 2008-2017.*

*Table 6 Number and percentage of top 10% highly-cited papers from EPSRC-funded papers and the world by discipline, 2008-2017.*

| Discipline | EPSRC | | Rest of UK | | EU-28 | | World | |
|---|---|---|---|---|---|---|---|---|
| Chemistry | 4,004 | 20.8% | 6,036 | 13.9% | 48,568 | 12.1% | 165,820 | 12.5% |
| Computer Science | 765 | 15.4% | 2,836 | 14.0% | 14,635 | 11.7% | 38,433 | 11.4% |
| Energy & Fuels | 871 | 25.6% | 1,701 | 18.0% | 12,498 | 18.6% | 41,149 | 17.6% |
| Engineering | 2,448 | 18.5% | 6,245 | 12.7% | 40,657 | 11.6% | 122,847 | 10.5% |
| Materials Science | 2,503 | 22.9% | 3,990 | 15.3% | 32,988 | 14.0% | 116,384 | 14.0% |
| Mathematics | 662 | 15.3% | 1,632 | 10.8% | 15,460 | 10.3% | 36,644 | 9.4% |
| Physics | 4,681 | 19.4% | 7,171 | 14.1% | 53,967 | 12.4% | 136,256 | 10.2% |

*Table 7 Number and percentage of top 5% highly-cited papers from EPSRC-funded papers and the world by discipline, 2008-2017.*

| Discipline | EPSRC | | Rest of UK | | EU-28 | | World | |
|---|---|---|---|---|---|---|---|---|
| Chemistry | 2,137 | 11.1% | 3,090 | 7.1% | 24,399 | 6.1% | 87,699 | 6.6% |
| Computer Science | 452 | 9.1% | 1,704 | 8.4% | 7,765 | 6.2% | 20,500 | 6.1% |
| Energy & Fuels | 477 | 14.0% | 879 | 9.4% | 6,433 | 9.7% | 21,526 | 9.3% |
| Engineering | 1,299 | 9.8% | 3,330 | 6.8% | 20,664 | 5.9% | 63,180 | 5.4% |
| Materials Science | 1,430 | 13.1% | 1,970 | 8.0% | 15,543 | 7.2% | 59,413 | 7.5% |
| Mathematics | 368 | 8.5% | 845 | 5.6% | 8,106 | 5.4% | 19,786 | 5.1% |
| Physics | 2,512 | 10.4% | 3,807 | 7.5% | 26,619 | 6.1% | 69,787 | 5.2% |

*Table 8 Number and percentage of top 1% highly-cited papers from EPSRC-funded papers and the world by discipline, 2008-2017.*

| Discipline | EPSRC | | Rest of UK | | EU-28 | | World | |
|---|---|---|---|---|---|---|---|---|
| Chemistry | 462 | 2.4% | 653 | 1.5% | 4,400 | 1.1% | 18,603 | 1.4% |
| Computer Science | 119 | 2.4% | 446 | 2.2% | 1,753 | 1.4% | 4,369 | 1.3% |
| Energy & Fuels | 119 | 3.5% | 178 | 1.9% | 1,326 | 2.0% | 4,629 | 2.0% |
| Engineering | 292 | 2.2% | 832 | 1.7% | 4,203 | 1.2% | 14,040 | 1.2% |
| Materials Science | 338 | 3.1% | 419 | 1.7% | 3,022 | 1.4% | 12,675 | 1.6% |
| Mathematics | 82 | 1.9% | 196 | 1.3% | 1,651 | 1.1% | 4,267 | 1.1% |
| Physics | 556 | 2.3% | 863 | 1.7% | 5,237 | 1.2% | 13,421 | 1.0% |

# EPSRC contributions to UK's strength areas

Research funded by EPSRC was identified within a body of exceptional UK research published between 2013 and 2018, and a selection of these research areas are highlighted in the following profiles.

# IV. SELECTED EXAMPLES OF UK STRENGTH AREAS

This section explores the contribution of EPSRC-funded researchers to areas of academic strength in the UK, as identified by using the Research Fronts Methodology.[7]

*Research Fronts dataset*

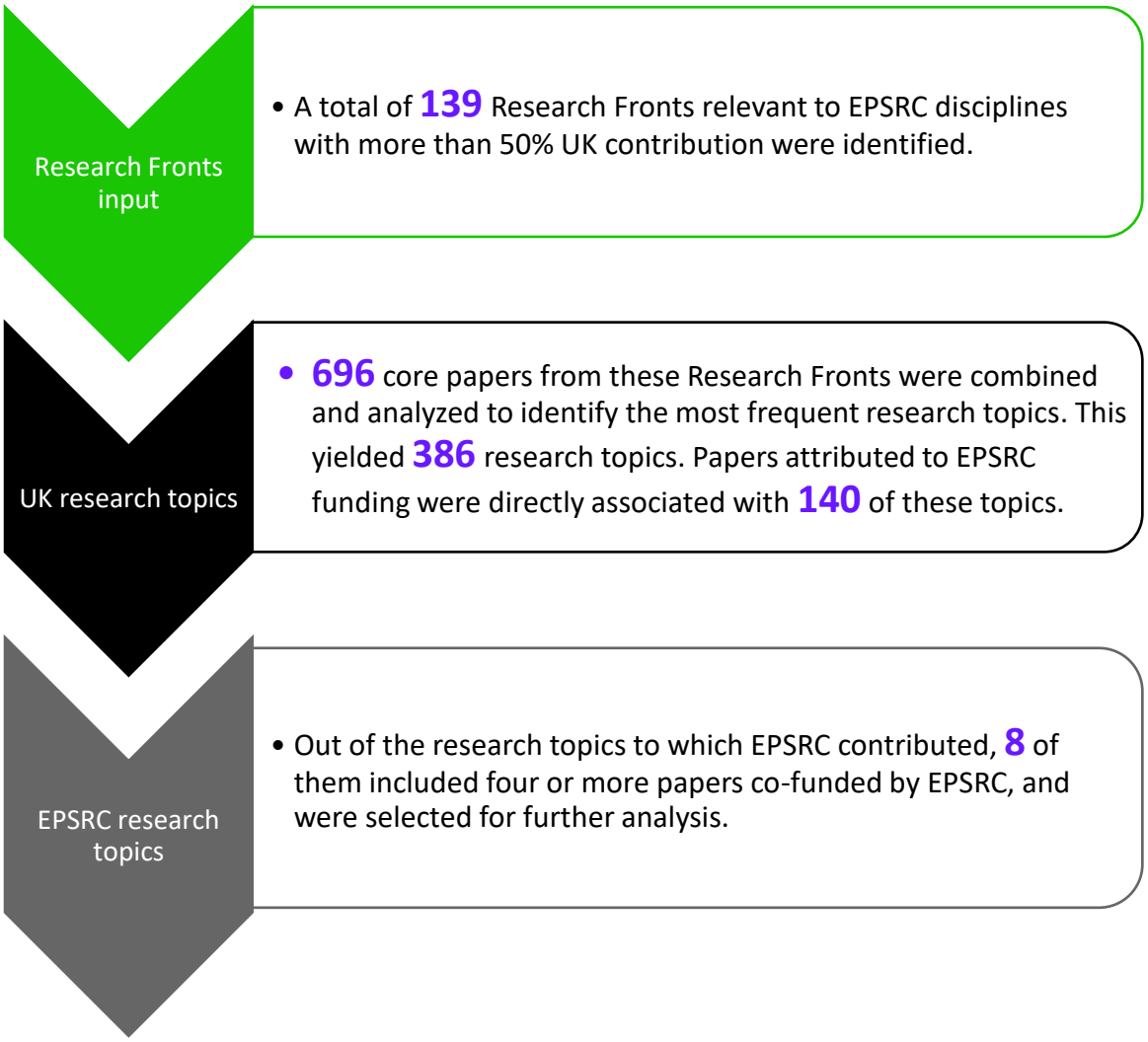
A 'Research Front' consists of a group of top 1% highly cited papers that have been frequently co-cited or cited in pairs. It has been shown that when researchers repeatedly acknowledge the same two papers in the reference lists of their publications, there is typically a close cognitive relationship between the co-cited papers. Groups of frequently co-cited papers, likewise, create a cluster of publications focusing on a problem, methodology, or solution. By analyzing this subset of all papers for co-citation, the most influential part of current research can be summarized. Thus, a Research Front consists of a number of 'core' papers, which are the highly cited papers frequently co-cited, and the citing papers that define the cohesive relationship among the core papers. The citing papers represent the most recent work and therefore represent the leading edge, or as Price said,[8] the growing epidermis, of the Research Front.

A Research Fronts dataset including highly co-cited papers published between 2013 and 2018 was obtained. These Research Fronts were filtered for relevancy (they must have included core papers classified under any of the 26 journal subject categories listed in Appendix I) and the level of UK participation (they must have had over 50% of papers with UK contributors). This process yielded 139 distinct Research Fronts derived from 696 core papers, which served as the basis for the topic analysis described next.

*Research topics dataset*

Text mining techniques were used to explore possible topics links between the Research Fronts. It was hoped that identifying the most frequent research topics, including those not necessarily linked by co-citations, would provide a basis on which to more broadly aggregate a large number of Research Fronts and hence facilitate subsequent analysis. A total of 386 widely disparate research topics were identified which could not be readily combined into clear groups; however, papers attributed to EPSRC funding were directly associated with 140 of these research topics, and eight of the topics were directly associated with four or more papers arising from EPSRC-funded research. These eight were further analysed and the results are presented here. A diagram showing the data selection and processing for the UK strengths areas analysis is shown below.

[7] See https://clarivate.com/essays/research-fronts/ for further details on the Research Fronts methodology and Appendix XI for details on the data processing.
[8] Publication reference: PubMed ID 14325149.




Overall Highlights:

Three of the topics identified were directly related to a single Research Front. However, in all other cases, there were two or more Research Fronts related to a single research topic, suggesting a topic with more than one active "pocket" of research not linked by co-citations.

Four out of the eight top research topics were classified within the discipline of Engineering;

There was a strong indication of interdisciplinarity, particularly between:

- ✓ Chemistry, Materials Sciences, and Physics;
- ✓ Computer Science and Engineering; and
- ✓ Engineering and Energy and Fuels;

EPSRC was a co-funder of all key influential papers within the topic of "Broadband Antennas" and a co-funder of all UK influential papers on "Monte Carlo Algorithms".

A qualitative and quantitative profile for each of the top eight research topics is presented in Table 9 to highlight the key metrics and contributors of these active areas of research. The bibliographies of influential papers funded by EPSRC are listed in Appendix VII.

*Table 9 Qualitative and quantitative profile of research topics where the UK exhibited a strength and to which EPSRC-funded researchers contributed.*

| Strength Area | Synopsis | Related disciplines | Average NCI | Num. of overall papers | Num. of UK influential papers | Num. of EPSRC influential papers | Other notable funders |
|---|---|---|---|---|---|---|---|
| **Energy** | *From research examining local community energy initiatives to specific challenges such as liquid air energy storage, the UK has been actively engaged in energy research. EPSRC and EDF Energy played an important role funding many of these investigations.* | Energy and Fuels | 7.06 | 33 | 26 | 9 | EDF Energy |
| **Monte Carlo Algorithms** | *Highly cited papers related to Markov chain Monte Carlo algorithms (important for the understanding of probabilities in complex systems) comprised a topic on their own. In this collection of papers spanning areas within Computer Science and Engineering, EPSRC-funded researchers played a leading role in the UK and overall.* | Computer Science, Engineering | 6.24 | 11 | 7 | 7 | European Research Council |
| **Self-Assembly** | *Researchers in the UK (many of them funded by EPSRC), contributed to research concerning the development of new approaches involving catalysis and self-assembly to synthesize functional nano-systems, controlled polymer architectures and dynamic supramolecular assemblies.* | Chemistry | 9.28 | 10 | 8 | 5 | European Research Council |
| **Antiferromagnetic Memory** | *In a quest to find novel methods to improve memory devices, researchers in the UK contributed to theoretical and experimental research on the electric control of antiferromagnets for reading and writing data, particularly in crystals such as Mn2Au.* | Chemistry, Materials Science, Physics | 13.38 | 10 | 7 | 5 | Grant Agency of the Czech Republic, Ministry of Education of the Czech Republic |

| Strength Area | Synopsis | Related disciplines | Average NCI | Num. of overall papers | Num. of UK influential papers | Num. of EPSRC influential papers | Other notable funders |
|---|---|---|---|---|---|---|---|
| **Optimization Algorithms** | *Spanning across sub-topics related to optimization, evolutionary programming, and cluster analysis, the UK conducted influential research related to these iterative, feedback-based algorithms in which convergence behavior is of prime interest.* | Computer Science | 9.18 | 8 | 7 | 4 | Wellcome Trust |
| **Power** | *Research specific to domestic UK power supply, including systems optimizations and the use of renewable sources, such as photovoltaics, was carried out by UK researchers. EPSRC and funders from China contributed to these influential papers.* | Energy and Fuels, Engineering | 9.48 | 9 | 8 | 4 | National High Tech. Res. and Dev. of China, National Natural Sci. Fdn. of China |
| **Broadband Antennas** | *From implementation techniques to the study and applications of Spatial Modulation aided Multiple-Input Multiple-Output (SM-MIMO) transmission techniques, EPSRC-funded researchers led research related to broadband antennas.* | Engineering | 22.64 | 5 | 5 | 5 | European Research Council, National Science Foundation of China |
| **Small Cell Spectrum** | *Focusing on communications and signal processing challenges, UK researchers published influential papers on small cell networks and most of these researchers were funded by the EPSRC.* | Engineering | 14.12 | 9 | 7 | 6 | N/A |

# Collaborations within EPSRC-funded research

EPSRC-funded researchers engaged in highly collaborative research output, with the USA, China, Germany, and France identified as strong collaborating countries.

Clarivate Analytics

# V. COLLABORATION LANDSCAPE

As previously indicated, EPSRC-funded researchers presented a high level of international collaborations during the study period, with almost half of the research papers funded by EPSRC classified as internationally collaborative. An assessment of these collaborations shows that the United States of America (USA), Germany, and China were frequent collaborators with EPSRC-funded researchers, as noted in Table 10. Papers co-authored with seven out of the top ten collaborating countries yielded citation impacts more than double the global average (NCI ≥ 2.00), and collaborations with Switzerland had the highest citation impact (NCI=2.27). An analysis of collaboration trends of EPSRC-funded papers (not presented here) suggests a significant increase in collaborations with Chinese institutions between 2008 and 2017 in the disciplines of Chemistry, Materials Science, and Physics.

*Table 10 Overall top collaborative countries of EPSRC-funded papers, 2008-2017.*

| Country | Number of papers | Normalized citation impact |
|---|---|---|
| USA | 6,667 | 2.18 |
| Germany | 4,273 | 2.00 |
| China | 3,923 | 1.79 |
| France | 3,375 | 2.01 |
| Italy | 2,103 | 2.11 |
| Spain | 2,001 | 2.02 |
| Japan | 1,657 | 1.81 |
| Australia | 1,401 | 1.92 |
| Switzerland | 1,265 | 2.27 |
| Netherlands | 1,099 | 2.07 |

Appendix VIII and Appendix IX list the top collaborating countries of papers published by researchers funded by EPSRC by discipline during the study period. Some highlights of each discipline are listed below:

***Chemistry***:

Top collaborating countries: USA (1,905), Germany (1,119), and China (1,018)

Highest impact: Collaborations with Italy, NCI= 2.41.

***Computer Science:***

Top collaborating countries: USA (496), China (424), and Germany (284)

Highest impact: Collaborations with Australia, NCI= 2.56.

***Energy & Fuels:***

Top collaborating countries: China (354), USA (194), and Germany (102)

Highest impact: Collaborations with Italy, NCI= 3.06.

***Engineering:***

Top collaborating countries: China (1,468), USA (994), and France (455)

Highest impact: Collaborations with the USA, NCI= 2.33.

***Materials Science:***

Top collaborating countries: USA (1,167), Germany (761), and China (736)

Highest impact: Collaborations with Switzerland, NCI= 2.96.

***Mathematics:***

Top collaborating countries: USA (752), Germany (411), and France (299)

Highest impact: Collaborations with the USA, NCI= 1.72.

***Physics:***

Top collaborating countries: USA (3,390), Germany (2,459), and France (1,780)

Highest impact: Collaborations with Australia, NCI= 2.41.

PAGE INTENTIONALLY LEFT BLANK

# Conclusions

The impact of research funded by EPRSC was exceptional, as compared with other regions and the global output, indicating that EPSRC is an important source of funding for globally relevant UK research.

Clarivate Analytics

# VI. CONCLUSIONS

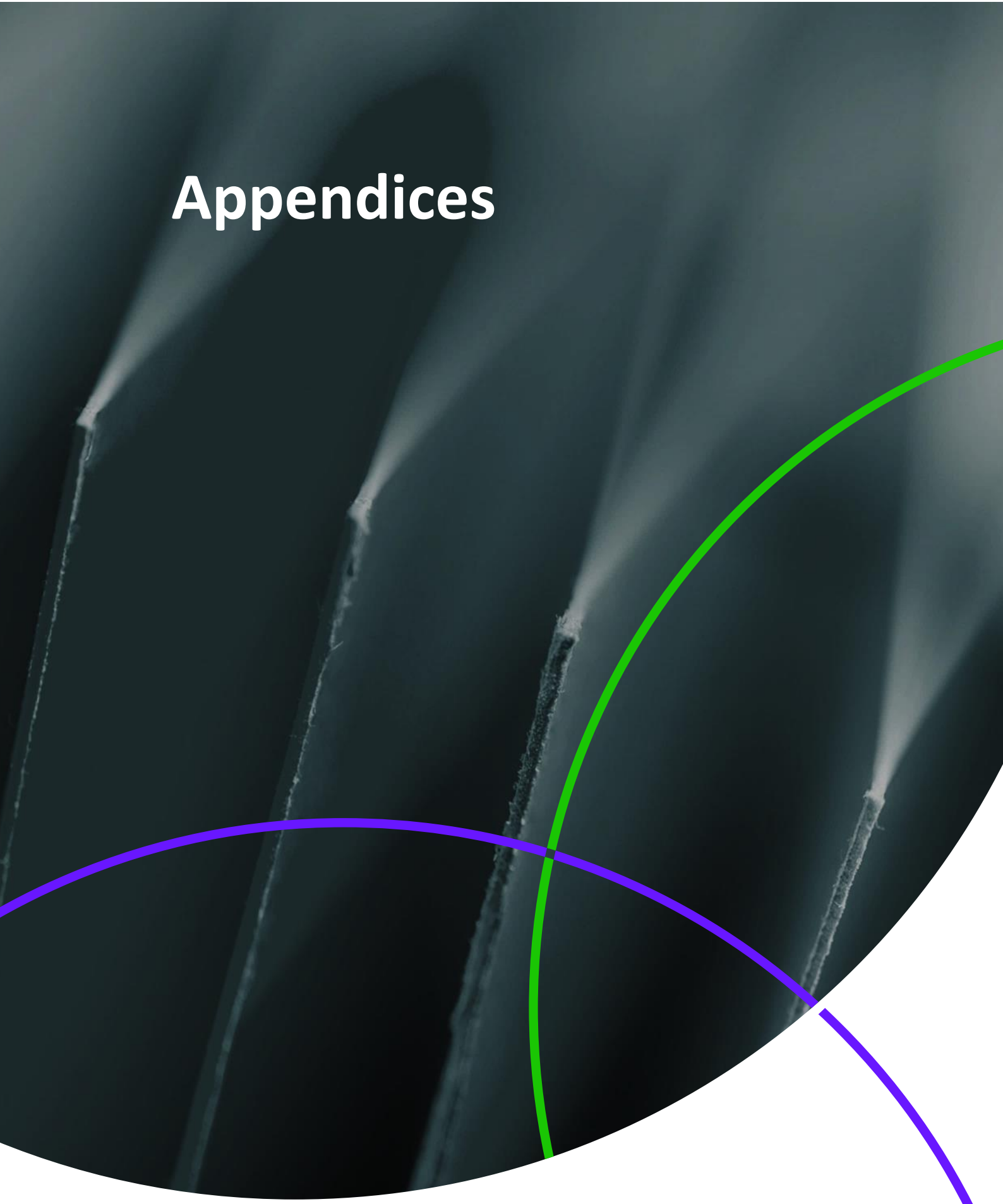
This report provided an evaluation of the research output, impact, disciplinary foci, and collaborations of EPSRC-funded papers published between 2008 and 2017, as well as their contributions to some of UK's areas of strength in Engineering and the Physical Sciences. The data indicate that EPSRC funding resulted in highly-cited scholarly output that contributed to over 1% of the global research output and 27.4% of the UK research output in categories broadly classified under Chemistry, Computer Science, Energy & Fuels, Engineering, Materials Science, Mathematics, and Physics. Roughly six out of seven EPSRC-funded papers were cited, and the citation impact of EPSRC-funded papers consistently surpassed the global average. Furthermore, it surpassed the citation impact of UK papers not funded by the EPSRC, European regions, BRIC countries and the world. This high performance in terms of citation impact was also observed in the percentage of highly cited papers, which were close to twice the global averages.

The data also shows the role EPSRC had in funding globally influential UK research, as exemplified by the research topics presented in Section IV of this report. Specifically, EPSRC funded influential papers in topics related to Energy & Fuels, Computer Science, Engineering, Chemistry, Materials Science, and Physics, among others, in some cases funding all papers related to a single research topic.

The collaborative nature of papers funded by EPSRC was also notable, with close to half of all funded papers including international co-authors, a value higher than the global average. Although the USA was the top collaborator in most disciplines, China was the top collaborator in Energy & Fuels and Engineering.

Taken together, this analysis shows a strong performance of publications arising from EPSRC-funded research and the key role EPSRC plays in supporting nationally and internationally influential research.

# Appendices



Clarivate Analytics

# APPENDIX I – MAPPING OF EPSRC-DEFINED DISCIPLINES TO WEB OF SCIENCE JOURNAL SUBJECT CATEGORIES

*Table 11 Mapping of EPSRC-defined disciplines to Web of Science Journal Subject Categories.*

| Discipline | Web of Science journal subject category |
|---|---|
| **Chemistry** | *Chemistry, Physical* |
| | *Chemistry, Multidisciplinary* |
| | *Chemistry, Organic* |
| | *Chemistry, Inorganic & Nuclear* |
| **Computer Science** | *Computer Science, Theory & Methods* |
| | *Computer Science, Artificial Intelligence* |
| | *Computer Science, Interdisciplinary Applications* |
| | *Computer Science, Software Engineering* |
| **Energy & Fuels** | *Energy & Fuels* |
| **Engineering** | *Engineering, Electrical & Electronic* |
| | *Mechanics* |
| | *Engineering, Chemical* |
| | *Engineering, Mechanical* |
| | *Telecommunications* |
| | *Instruments & Instrumentation* |
| **Materials Science** | *Materials Science, Multidisciplinary* |
| | *Nanoscience & Nanotechnology* |
| **Mathematics** | *Mathematics* |
| | *Mathematics, Applied* |
| **Physics** | *Physics, Applied* |
| | *Physics, Condensed Matter* |
| | *Optics* |
| | *Physics, Atomic, Molecular & Chemical* |
| | *Physics, Multidisciplinary* |
| | *Physics, Fluids & Plasmas* |
| | *Physics, Mathematical* |

## APPENDIX II – GLOBAL NORMALIZED CITATION IMPACT

Average normalized citation impact for the global data extracted using the Web of Science journal subject categories listed in Appendix I. These values have been normalized to NCI=1.00 in the report, and all other NCI values have been normalized against the actual global NCI value.

*Table 12 Global normalized citation impact per discipline and in aggregate, 2008-2017.*

| Aggregate | NCI |
|---|---|
| **All disciplines in aggregate** | 0.92 |
| **Chemistry** | 1.03 |
| **Computer Science** | 0.99 |
| **Energy & Fuels** | 1.20 |
| **Engineering** | 0.89 |
| **Materials Science** | 1.00 |
| **Mathematics** | 0.93 |
| **Physics** | 0.87 |

*Table 13 Global normalized citation impact per year, 2008-2017.*

| Publication year | NCI |
|---|---|
| **2008** | 0.95 |
| **2009** | 0.95 |
| **2010** | 0.94 |
| **2011** | 0.95 |
| **2012** | 0.94 |
| **2013** | 0.94 |
| **2014** | 0.93 |
| **2015** | 0.93 |
| **2016** | 0.91 |
| **2017** | 0.81 |

# APPENDIX X - AUTHORS, DATA SOURCES, AND METHODOLOGY

## A. AUTHORS

Clarivate Analytics accelerates the pace of innovation by providing trusted insights and analytics to customers around the world, enabling them to discover, protect and commercialize new ideas faster. We own and operate a collection of leading subscription-based services focused on scientific and academic research, patent analytics and regulatory standards, pharmaceutical and biotech intelligence, trademark protection, domain brand protection and intellectual property management. Clarivate Analytics is now an independent company with over 4,000 employees, operating in more than 100 countries and owns well-known brands that include Web of Science, ScholarOne, EndNote, Kopernio, Converis, Publons, Cortellis, Derwent, CompuMark, MarkMonitor and Techstreet, among others. For more information, visit https://clarivate.com .

The Web of Science Group provides reporting and consultancy services using customized analyses. By bringing together several indicators of research performance, we enable customers to rapidly make sense of and interpret a wide-range of data points to facilitate research strategy decision-making. We have extensive experience with databases on research inputs, activity and outputs and have developed innovative analytical approaches for benchmarking, interpreting and visualization of international, national and institutional research impact.

For over half a century we have pioneered the world of citation indexing and analysis, helping to connect scientific and scholarly thought around the world. Today, academic and research institutions, governments, not-for-profits, funding agencies, and all others with a stake in research, need reliable, objective methods for managing and measuring performance. For more information, visit http://webofsciencegroup.com/ .

## B. DATA SOURCES

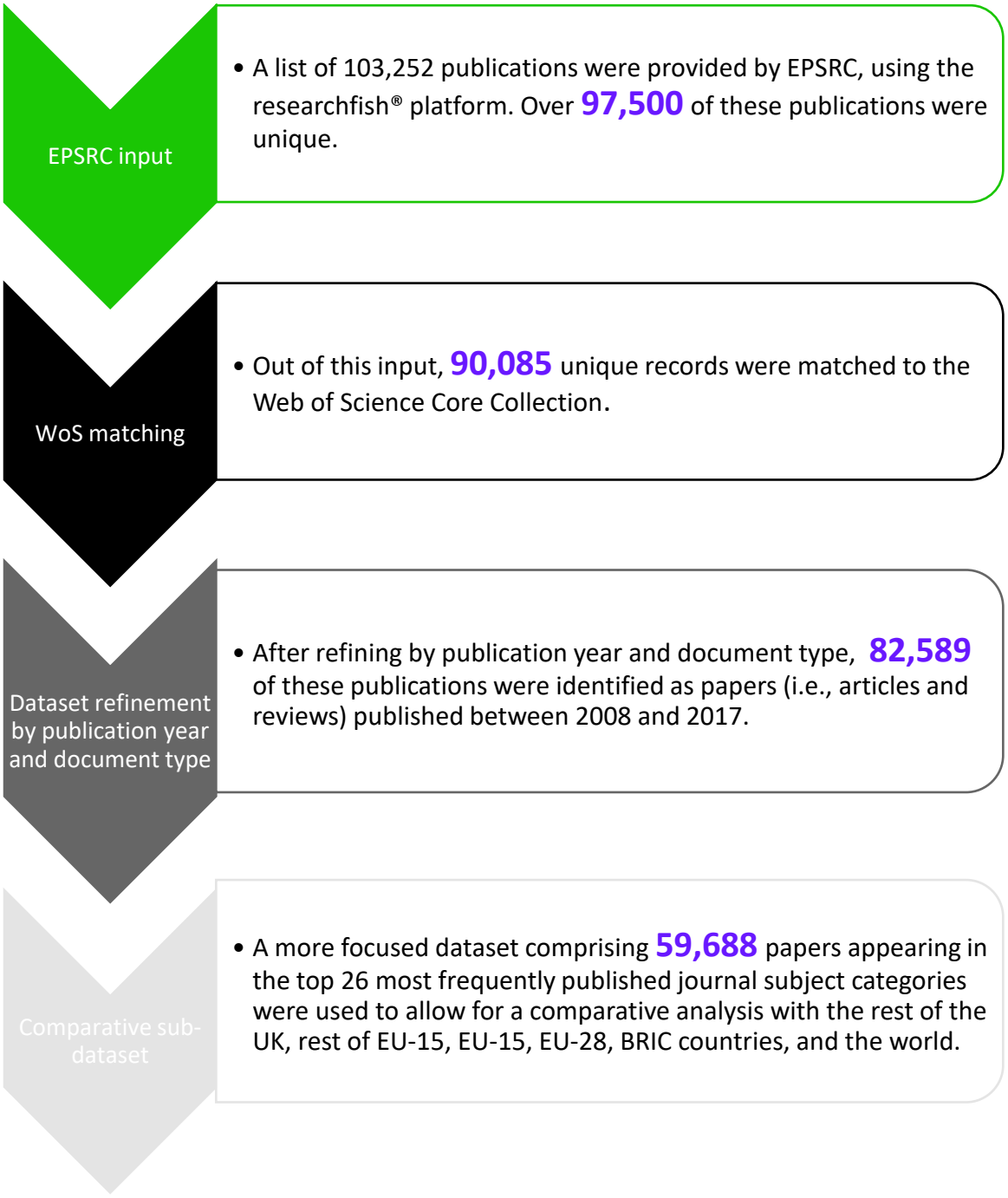
Clarivate Analytics has extensive experience with databases on research inputs, activities and outputs and has developed innovative analytical approaches for benchmarking and interpreting international, national and institutional research impact. The Web of Science gives access to conference proceedings, patents, websites, and chemical structures, compounds and reactions in addition to journals. It has a unified structure that integrates all data and search terms together and therefore provides a level of comparability not found in other databases. It is widely acknowledged to be the world's leading source of citation and bibliometric data. The Web of Science Core Collection is part of the Web of Science and focuses on research published in journals and conferences in science, medicine, arts, humanities and social sciences. The authoritative, multidisciplinary content covers over 20,000 of the highest impact journals worldwide, including Open Access journals and over 180,000 conference proceedings. Coverage is both current and retrospective in the sciences, social sciences, arts and humanities, in some cases back to 1900. Within the research community, these data are often still referred to by the acronym "ISI", referring to the Institute for Scientific Information, which was recently reinstated[10]. Clarivate Analytics has extensive experience with databases on research inputs, activity and output and has developed

[10] https://clarivate.com/blog/science-research-connect/isi-redux-the-past-inspires-the-future/ .

innovative analytical approaches for benchmarking and interpreting international, national and institutional research impact.

**EPSRC publications database:** EPSRC uses the researchfish® platform to maintain a comprehensive dataset of publications confirmed by EPSRC-funded researchers to have arisen from their EPSRC funding. References to over 100,000 uniquely identified journal publications were provided to Clarivate Analytics to define the initial dataset to be considered for the analysis presented in this report.

**Web of Science:** For the bibliometric analysis, the EPSRC dataset was matched to the databases underlying the Clarivate Analytics Web of Science, as noted in the flowchart below:



**Comparators database:** The comparators database was created by extracting all papers from the Web of Science Core Collection which were classified under the top 26 journal subject categories in which EPRSC papers were most frequently published, as listed in Appendix I. These papers where then classified into the following six groups:

| | |
|---|---|
| Rest of UK: | Refers to papers with UK authors, but which have not been attributed to specific instances of EPSRC funding. |
| EU-15: | Refers to Austria, Belgium, Denmark, Finland, France, Germany, Greece, Ireland, Italy, Luxembourg, Netherlands, Portugal, Spain, Sweden, and UK |
| Rest of EU-15: | Refers to papers published by authors from affiliations within the EU-15 group, but excluding papers only attributed to UK authors. |
| EU-28: | Refers to Austria, Belgium, Bulgaria, Croatia, Cyprus, Czech Republic, Denmark, Estonia, Finland, France, Germany, Greece, Hungary, Ireland, Italy, Latvia, Lithuania, Luxembourg, Malta, Netherlands, Poland, Portugal, Romania, Slovakia, Slovenia, Spain, Sweden, and UK |
| BRIC: | Refers to Brazil, Russia, India, and China. |
| World: | Refers to all papers in the mapped disciplines, irrespective of the authors' affiliation country. |

**UK Strengths Areas dataset:** The dataset used to identify areas of influential UK research was based on the **Research Fronts** methodology. Within the Essential Science Indicators, a Research Fronts consists of a group of top 1% highly cited papers that have been frequently co-cited or cited in pairs. Groups of frequently co-cited papers, likewise, create a cluster of publications focusing on a problem, methodology, or solution. By analyzing this subset of all papers for co-citation, the most influential part of current research can be summarized. Thus, a Research Fronts consists of a number of 'core' papers, which are the highly cited papers frequently co-cited, and the citing papers that define the cohesive relationship among the core papers.

The Research Fronts dataset used for the analysis was generated in June 2018 and included top 1% highly co-cited papers published between January 2013 and June 2018. Within this dataset, a subset of research fronts having the following characteristics were identified:

> **Relevancy**: The Research Front should include core papers that have been classified under any of the 26 journal subject categories listed in Appendix I;
>
> **UK participation**: More than 50% of the core papers within a Research Front should have at least one co-author affiliated to a UK institution.

This process yielded 139 distinct Research Fronts, derived from 696 core papers. Using all core papers from these 139 Research Fronts as a dataset, text mining techniques were used to identify the most frequent research topics, and a total of 386 research topics were obtained. Papers attributed to EPSRC funding were directly associated with 140 of these research topics; eight of them were directly

associated with four or more papers arising from EPSRC-funded research. These eight research topics were selected and analyzed in this report.

**Aggregate data:** The data underpinning the charts and figures presented in this report are available from EPSRC upon request.

## C. DEFINITIONS AND METRICS

**Papers/publications:** Clarivate Analytics abstracts publications including research journal articles, editorials, meeting abstracts and book reviews. The terms "paper" and "publication" are often used interchangeably to refer to printed and electronic outputs of many types. However, the term "paper" is used exclusively to refer to substantive journal articles, reviews and some proceedings papers and excludes editorials, meeting abstracts or other types of publication. Papers, which are the subset of publications for which citation data are most informative and which are used in calculations of citation impact, were used for the analysis presented in this report.

**Research field:** Standard bibliometric methodologies use Web of Science journal categories or **InCites: Essential Science Indicators℠** fields as a proxy for research fields.[11] Essential Science Indicators aggregate data at a higher level than the journal categories – there are only 22 Essential Science Indicators research fields compared to over 252 journal categories. Journals are assigned to one or more categories, and every article within that journal is subsequently assigned to that category. Papers from prestigious, "multidisciplinary" and general medical journals such as *Nature*, *Science*, *The Lancet*, *The BMJ*, *The New England Journal of Medicine and the Proceedings of the National Academy of Sciences* (PNAS) are assigned to specific categories based on the journal categories of the references cited in the article. The selection procedures for the journals included in the citation databases are documented at the Clarivate Analytics master journal list website. Web of Science journal categories and a custom classification of disciplines (with names resembling the Essential Science Indicators) were used for this analysis.

**Citation count:** The citation count is the number of times that a citation has been recorded for a given publication since it was published. Not all citations are necessarily recorded since not all publications are indexed. However, the material indexed by Clarivate Analytics is estimated to attract about 95% of global citations. Self-citations are included in the citation count presented in this analysis.

**Percentile in Subject Area** The percentile of a publication is determined by creating a citation frequency distribution for all the publications in the same year, subject category and of the same document type (arranging the papers in descending order of citation count) and determining the percentage of papers at each level of citation, i.e., the percentage of papers cited more often than the paper of interest. If a paper has a percentile of value of one, then 99 percent of the papers in the same subject category, year and of the same document type have a citation count that is lower. A percentile indicates how a paper has performed relative to others in its field, year, and document type and is therefore a normalized indicator. Below is a sample calculation of percentile for a set of 11 papers.

[11] Essential Science Indicators are defined by a unique grouping of journals with no journal being assigned to more than one field. These fields are focused on the science, technology, engineering and medicine subjects and arts & humanities subjects are excluded. Customized analyses, however, can be designed to include these as an additional category.

| Times cited | Percentile |
|---|---|
| **1,000** | 9.09 |
| **50** | 18.18 |
| **10** | 27.27 |
| **3** | 36.36 |
| **2** | 54.55 |
| **1** | 81.82 |
| **1** | 81.82 |
| **1** | 81.82 |
| **0** | 100.00 |
| **0** | 100.00 |

**Average Percentile** For any set of papers, an Average Percentile can be calculated as the mean of all the percentiles of all the papers in the set. In the case that a paper is assigned to more than one category, the category in which the percentile value is closest to zero is used (i.e. the best performing value). Percentile values are rounded to the second decimal place. The Average Percentile can apply to any set of papers, such as an author's body of work, all the publications in a journal or the accumulated publications of an institution, country, or region. The average percentile will represent the average performance of the papers in the set having been normalized for field, year and document type. The main advantage of the Average Percentile indicator is that it can be used to compare to peer entities regardless of size, age or subject focus. In this regard, it is quite similar to and is a complement to, the Category Normalized Citation Impact indicator. An advantage of the Average Percentile indicator is that it describes the relative position of a paper compared to similar papers. One disadvantage is that it does not necessarily indicate the actual number of citations. In the example in the table above, it can be seen that the most highly cited paper has 20 times more citations than the second most cited paper, however the percentile of the first paper has a relatively similar value to the paper in second position. In a more typical distribution, which may contain thousands of papers, these two papers may have very similar percentiles. This artifact of the methodology is advantageous as it overcomes the skewed nature of citation-based indicators, but at the same time, it is disadvantageous as it may not fully recognize the value of highly cited papers. As with other indicators, it is recommended that the percentile is used alongside and to complement other indicators.

**Highly cited papers:** Highly cited work is recognized as having a greater impact and Clarivate Analytics has shown that high citation rates are correlated with other qualitative evaluations of research performance, such as peer review. In the analysis presented here, papers that have an average percentile within the top 1%, 5%, and 10% in terms of citation frequency taking into account year of publication, type of document, and field, are considered to be highly cited. A higher value is considered to be higher performance. A value of around "1" for a set of papers represents that one percent of the papers in that set are in the top one percent of the world regardless of subject, year and document type and would therefore be considered to be performing at the same level as world average. A value above "1" represents that more than one percent of papers in the set are in the top one percent of the world and a value of less than "1" would represent that less than one percent of the papers in the set are in the top one percent of the world.

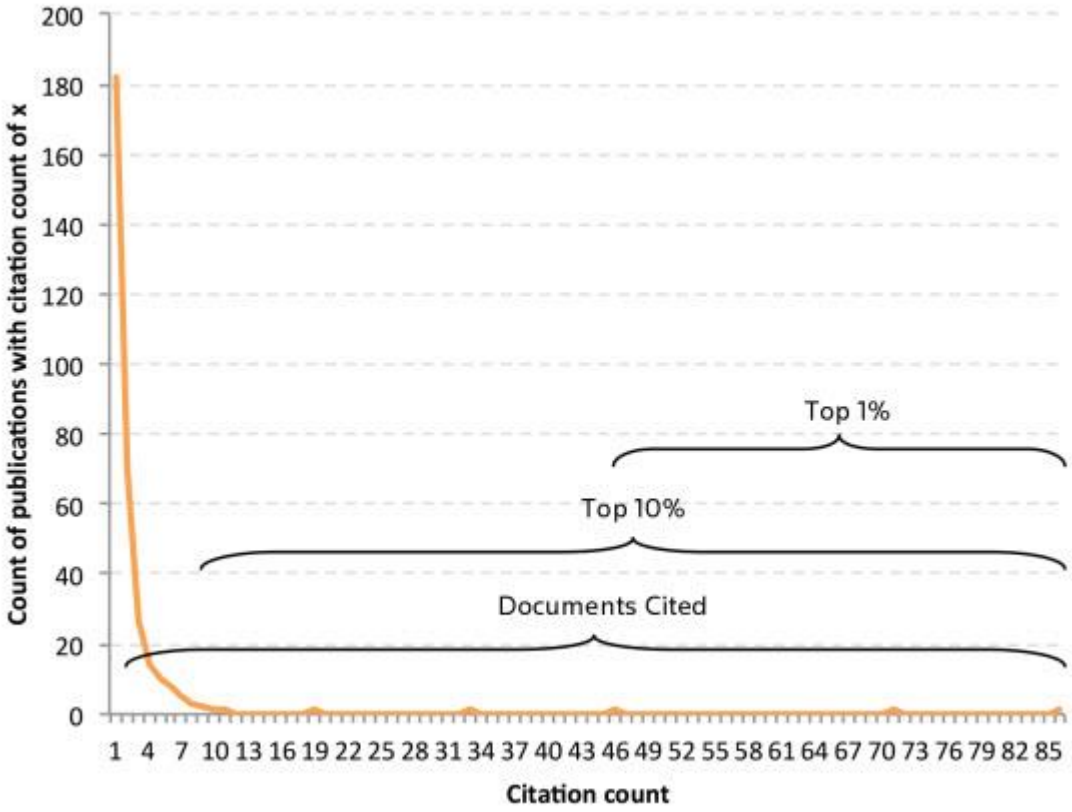
The % of papers in the top 1% indicator is considered to be an indicator of research excellence as only the most highly cited papers would make the top one percent in their respective field, year and

document type. The indicator can be used in conjunction with other indicators to provide a more complete picture of performance. The % of papers in the top 1% indicator can be applied to any level of aggregation (author, institution, national/international, field).

Although the top one percent is a relevant measure of excellence, by its nature it is typically only a small percentage of any document set and therefore the statistical relevance of small sample sizes is a significant concern.

The % of papers in the top 1% is best used with large datasets such as the accumulated publications of an institution, country or region and for a publication window of several years. The % of papers in the top 5% anad top 10% are very similar to the % of papers in the top 1% simply with a threshold of around 5 or 10 percent instead of one percent, respectively. Therefore, typical performance will be around a value of 5 or 10 and values of higher than 5 or 10 would be considered above average performance. The three indicators complement each other very well to give a broader picture of highly performing research (5 and 10 percent) and excellence (one percent).

The % of papers in the top 10% is also more appropriate than the % of papers in the top 1% when the size of the data set is smaller. However, it is still only appropriate for large to medium size data sets and should be used with a great deal of caution when looking at small datasets such as the output of an individual author. A sample citation distribution illustrating the number of publications (y axis) with citation count of x (x axis) is shown below.




**Citation impact:** "Citations per paper" is an index of academic or research impact (as compared with economic or social impact). It is calculated by dividing the sum of citations by the total number of

papers in any given dataset (so, for a single paper, raw impact is the same as its citation count). Citation count declines in the most recent years of any time-period as papers have had less time to accumulate citations (i.e., papers published in 2010 will typically have more citations than papers published in 2017).

**Category Normalized Citation Impact (NCI):** Citation rates vary between research fields and with time. Consequently, analyses must take both field and publication year into account. In addition, the type of publication will influence the citation count. For this reason, only citation counts of papers (as defined above) are used in calculations of citation impact. The standard normalization factor is the world average citations per paper for the year, journal category and document type in which the paper was published. The Category Normalized Citation Impact (NCI) of a publication is calculated by dividing the actual citations by the normalization factor. When a publication is assigned to more than one subject area or document type, an average of the ratios of the normalization factor is used.

**Average normalized citation impact (avg NCI):** The average or mean normalized citation impact (avg NCI) indicator for any specific dataset is calculated as the mean of the category normalized citation impact (NCI) of all papers within that dataset. In this report, this metric is referred to as average NCI or NCI and it includes authors' self-citations

**Co-authorship of publications:** The metadata associated with every research publication include the addresses of the authors. It is thus possible to develop an analysis of the organizations that co-author publications by extracting and examining these data. Co-authorship is generally accepted as an indicator of collaboration, although there are collaborations that do not result in co-authored publications and co-authored publications which involve limited collaboration. Conceivably other indicators of collaboration such as co-funding and international exchanges could be used but comprehensive and consistent data are not available.

**Internationally collaborative publications:** The number of internationally collaborative research publications is increasing rapidly. This is because such collaboration provides access to a wider range of resources, including intellectual resources, and accelerates the rate of discovery as well as increasing the intellectual content and therefore the impact of individual outputs. For this reason, internationally collaborative publications tend to be more highly cited than those that are solely domestic. In the analysis, publications will be considered to be international if more than one country is included in the addresses associated with a paper. Note that England, Scotland, Wales, and Northern Ireland have been aggregated and are treated as one country for the purposes of this study.

## D. BIBLIOMETRICS AND CITATION ANALYSIS

Research evaluation is increasingly making wider use of bibliometric data and analyses. Bibliometrics is the analysis of data derived from publications and their citations. Publication of research outcomes is an integral part of the research process and is a universal activity. Consequently, bibliometric data have a currency across subjects, time and location that are found in few other sources of research-relevant data. The use of bibliometric analysis, allied to informed review by experts, increases the objectivity of, and confidence in, evaluation.

Research publications accumulate citation counts when they are referred to by more recent publications. Citations to prior work are a normal part of publication and reflect the value placed on a work by later researchers. Some papers get cited frequently and many remain uncited. Highly cited

work is recognized as having a greater impact and Clarivate Analytics has shown that high citation rates are correlated with other qualitative evaluations of research performance, such as peer review. This relationship holds across most science and technology areas and, to a limited extent, in social sciences and even in some humanities subjects.

Indicators derived from publication and citation data should always be used with caution. Some fields publish at faster rates than others and citation rates also vary. Citation counts must be carefully normalized to account for such variations by field. Because citation counts naturally grow over time, it is essential to account for growth by year. Normalization is usually done by reference to the relevant global average for the field and for the year of publication.

Bibliometrics work best with large data samples. As the data are disaggregated, so the relationship weakens. Average indicator values (e.g., of citation impact) for small numbers of publications can be skewed by single outlier values. At a finer scale, when analyzing the specific outcome for individual departments, the statistical relationship is rarely a sufficient guide by itself. For this reason, bibliometrics are best used in support of, but not as a substitute for, expert decision processes. Well-founded analyses can enable conclusions to be reached more rapidly and with greater certainty and are therefore an aid to management and to increased confidence among stakeholders, but they cannot substitute for review by well-informed and experienced peers.