# Bibliometric analysis of ongoing projects: Innovative Medicines Initiative Joint Undertaking (IMI)

## IMI EXECUTIVE OFFICE

## Fourth report:  April 2014

Prepared by Thomson Reuters on behalf of IMI JU Executive Office under a public procurement procedure

# 1 EXECUTIVE SUMMARY

This report presents a bibliometric analysis of IMI research associated with funding Calls 1 to 6, using citations as an index of research quality and co-authorship as an index of collaboration. The analyses use two sets of research publications, publications from IMI projects (IMI project research) and publications from IMI-supported researchers (IMI researchers).

The overall volume of IMI project research has increased rapidly since 2009 and the initiative continues to show rapid growth. This is partly to be expected as the number of funded projects rises and those projects funded earliest in the program begin to publish. To date, IMI projects have produced 657 publications, of which 629 have been matched to the Thomson Reuters *Web of Science*. Around three quarters of IMI project research has been published in high impact journals, i.e. those journals in the highest quartile by Journal Impact Factor.

The volume of IMI research has also increased at the level of individual projects. BTCure (Call 2) is the most prolific project in any of the funding Calls with output surpassing NEWMEDS and EUROPAIN (the most prolific projects funded in Call 1) with a substantial expansion in 2013 output accompanied by rising citation impact. Among more recent projects, EU-AIMS (Call 3) has shown substantial growth in output, and its research is now cited nearly four times (3.96) world average.

IMI project research is wide-ranging – the research portfolio from IMI projects covers diverse research fields. IMI project research has been published most frequently in Rheumatology, Pharmacology & Pharmacy, Neurosciences and Endocrinology & Metabolism journals.

IMI project research is well-regarded. The quality of IMI project research (as indexed by citation impact) has been maintained while output has grown. The citation impact of this research is, at twice world average (2.05), internationally influential. Around one quarter of papers from IMI projects are 'highly-cited', that is, they belong to the world's top ten percent of papers in that journal category and year of publication, when ranked by number of citations received. Relative to established funding organisations IMI project research performs well. Trends in the quality of this research are generally comparable to the Wellcome Trust, a leading UK funder of biomedical research.

Researchers funded by IMI are well-regarded by their peers. The total research published by IMI researchers funded in Calls 1 to 5 (as opposed to that directly associated with IMI funding) is well cited with an average citation impact over twice world average and about one-fifth of papers being 'highly-cited'. This is similar performance to IMI project research indicating that IMI funds both research and individuals performing at a high overall level.

Researchers and projects funded by IMI are highly collaborative. About two-fifths of all publications by IMI researchers were cross-sector (for example, between academic institutions and the pharmaceutical industry or small- or medium-sized enterprises - SMEs) and over half of all papers from IMI projects were cross-sector.

A more detailed summary of key findings from this report cross-referenced to associated analyses is presented overleaf.

## Summary of key findings – IMI project research

As of April 2014, there are 46 IMI projects from Calls 1 to 8, of which 23 were launched since 1 January 2012, and four since 1 January 2014. It may take several years for a project to progress from inception to the point where it has generated sufficient data for a publication. It may take further years until it has produced its most valuable results. The IMI projects that are analysed here are still relatively young, and early bibliometric indicators may not fully reflect their eventual impact.

- Of 657 publications from IMI-supported projects, a total of 609 publications were identified for inclusion in this report on IMI research activity up to end-December 2013. These publications were linked to Thomson Reuters citation data and 599 (98.4%) of these documents were substantive articles and reviews (Section 4.1, Section 4.2).
- IMI project research continues to be published in highly-regarded journals, and much of its research is published in journals ranked in the top quartile of journals by journal impact factor (74.4%). The average journal impact factor of all the journals IMI project research is published in is 6.05. The core set of journals used by IMI projects continues to highlight the diversity of IMI-supported research with titles focused on bioinformatics, genetics and psychiatry as well as disease areas such as arthritis and diabetes (Section 4.4, Annex 1).
- IMI project research is most frequently published in Rheumatology and Pharmacology & Pharmacy journals. Output in Rheumatology journals is predominantly associated with BTCure. The most frequently used journal categories continue to reflect breadth and depth in the IMI research portfolio which contains cross-cutting and more specialised research.
- The average citation impact for IMI project research is 2.05 for the 4-year period, 2010-2013, (where world average is 1.0) and one quarter (24.4%) of papers are highly-cited. This is exceptionally highly-cited research. For comparison, the EU's average citation impact relative to world baselines for the same 4-year period in similar research fields was 1.15. The percentage of highly-cited research has increased substantially since the third report, up from 19.3%. This highlights the rapid uptake of research published by IMI projects.
- IMI project research has similar citation impact to research acknowledging Wellcome Trust funding. Research from both funders is internationally influential with citation impact around twice world average and around a quarter of papers are highly-cited (Section 4.7).
- IMI project research is collaborative at sector, institution and country level. Well over half (64.3%) of all IMI project papers have been published by researchers affiliated with different sectors and half (50.6%) of all IMI project papers have an international co-author (Section 4.8).
- Output has increased for each of the first three IMI funding Calls, though Call 1 continues to account for the highest share of IMI research. NEWMEDS and EUROPAIN continue to be the most prolific projects funded in Call 1, however, output from the Call 2 project, BTCure, is by far the largest IMI project with 107 publications (Section 5).
- Research published for the projects EU-AIMS, PROTECT and U-BIOPRED have exceptionally high citation impact (around or over three times world average). EUROPAIN, NEWMEDS and Pharma-Cog also have citation impact greater than that of IMI projects overall (2.05 or more).

## Summary of key findings – IMI researchers

The productivity, research performance and collaboration of researchers funded by IMI through Calls 1 to 5 were assessed by analysing their total publication output (not limited to publications acknowledging IMI funding). Some 4 861 researchers were included in the analysis and 29 064 of their publications were identified for the period January 2007-December 2013 (Section 6.1).

- Publication output, as previously, is higher for IMI-supported researchers based in academic institutions and other research environments compared to industry and SMEs (Section 6.3).
- Since the third report, researchers associated with patient organisations have continued to publish well-regarded research and continue to show the strongest research performance. Six (28.6%) have published at least one 'hot' paper, 5 (23.8%) have an h-index of at least 10 and 19 (90.5%) have published exclusively in top-quartile journals (Section 6.4).
- Collaboration analysis was performed on the basis of co-authorship between IMI-supported researchers. Three-quarters (75.7%) of IMI researchers co-authored with at least one other IMI researcher during the period of analysis (Section 6.5).
- Again, co-authorship is more common between researchers in the same sector than among researchers in different sectors. Cross-sector co-authorship accounts for just over one-third (36.6%) of all co-authorship activities during the analysis period (Section 6.6).
- The same is true of co-authorship activities by project. The majority of collaborative relationships are among researchers associated with the same project with two-fifths (40.5%) being cross-project. The share of cross-project activity between researchers has, however, increased relative to the third report (Section 6.6).
- Since receiving IMI funding, most researchers have become more collaborative. Between 2007 and when they first received funding, the average IMI-funded researcher co-authored with 2.55 other researchers. Since receiving IMI funding, this figure has increased to 5.89 (Section 6.7).
- This increase applies for most disease areas (Section 6.7), within and cross-sector (Section 6.8) and within and between countries (Section 6.9). Iceland, Finland and Sweden are the most collaborative countries with IMI-funded researchers.

# 2 INTRODUCTION

## 2.1 OVERVIEW

The Innovative Medicines Initiative Joint Undertaking (IMI) has commissioned Thomson Reuters to undertake a periodic evaluation of its research portfolio using bibliometric and intellectual property indicators.

The commissioned evaluation comprises a series of bi-annual reports focussing on research publications and patents produced by IMI funded researchers. This report is the fourth evaluation in the series. Since the number of applications and awards specifically generated by IMI projects to date is small, IMI personnel have advised that patent analyses are not required for this fourth evaluation.

## 2.2 INNOVATIVE MEDICINES INITIATIVE JOINT UNDERTAKING (IMI)

The Innovative Medicines Initiative Joint Undertaking (IMI) is a public private partnership between the European Union and the European Federations of Pharmaceutical Industries and Associations (EFPIA). The purpose of the IMI is to increase the efficiency and effectiveness of the drug development process, thereby increasing production of safer and more effective medicines. IMI pools resources from the public and private sectors and is funded jointly through Framework Programme Seven, EFPIA and EFPIA member companies. IMI supports pre-competitive pharmaceutical research and development to deliver new approaches, methodologies, and technologies.

With a €2 billion budget, IMI supports collaborative research projects and builds networks of industrial and academic experts in Europe that will boost innovation in healthcare. By acting as a neutral third party to support the creation of innovative partnerships, IMI aims to build a more collaborative ecosystem for pharmaceutical research and development (R&D).

IMI supports research projects in the areas of safety and efficacy, knowledge management and education and training. Projects are selected through open Calls for proposals. Project participants are recruited through these open and competitive Calls based on independent peer review and concluded by a Grant Agreement and Project Agreement.

The research consortia participating in IMI projects consist of:

- large biopharmaceutical companies that are members of EFPIA

and a variety of other partners, such as:

- small- and medium-sized enterprises,
- patients' organisations,
- universities and other research organisations,
- hospitals,
- regulatory agencies,
- any other industrial partners.

To date, IMI have announced eleven Calls for proposals to be funded under the initiative. The first funding call was announced in 2008 and the latest, 11$^{th}$, funding call was launched on 11$^{th}$ December 2013.

This report will cover the research outputs (publications and papers) from Calls 1 to 6 (though Call 5 currently has no publication output associated with it) which have resulted in 40 projects.

## 2.3 THOMSON REUTERS

Thomson Reuters is the world’s leading source of intelligent information for business and professionals. We combine industry expertise with innovative technology to deliver critical information to leading decision makers in the financial, legal, tax and accounting, healthcare, science and media markets, powered by the world's most trusted news organisation. Visit our WEBPAGE for more information.

## 2.4 THOMSON REUTERS RESEARCH ANALYTICS

Thomson Reuters Research Analytics is a suite of products, services and tools that provide comprehensive research analysis, evaluation and management. 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. Visit our WEBPAGE for more information.

## 2.5 THOMSON REUTERS CUSTOM ANALYTICS & ENGINEERED SOLUTIONS

Thomson Reuters Custom Analytics & Engineered Solutions provide reporting and consultancy services within Research Analytics using customised analyses to bring together several indicators of research performance in such a way as to enable customers to rapidly make sense and interpret of a wide-range of data points to facilitate research strategy decision-making.

Our consultants have up to 15 years' experience in research performance analysis and interpretation. We have extensive experience with databases on research inputs, activity and outputs and have developed innovative analytical approaches for benchmarking, interpreting and visualisation of international, national and institutional research impact.

## 2.6 SCOPE OF THIS REPORT

One of IMI's principal objectives is to support collaborative research projects and build networks of industrial and academic experts in Europe. This will deliver socio-economic benefits to European citizens, increase Europe's competitiveness globally and establish Europe as the most attractive place for pharmaceutical R&D.

The analyses and indicators presented in this report have been specified to provide an analysis of IMI research output for research management purposes:

- To provide bibliometric indicators to identify excellence in IMI-supported research and to benchmark this research, where possible, overall and at individual project level.
- To provide bibliometric indicators at individual researcher level.
- To show that collaboration, at all levels (researcher, institutional and country), is being encouraged through the projects funded by IMI.

Outline of report

- Section 3 describes the data sources and methodology used in this report along with definitions of the indicators and guidelines to interpretation.
- Sections 4 and 5 present citation analyses of IMI project publications overall (Annex 1 provides summary analyses of IMI project publications identified since the last report to IMI).
- Section 6 presents bibliometric indicators for IMI-supported researchers and analyses of collaboration between these individuals.
- Previous reports (October 2012 and February 2013) contained patent data and analyses for IMI research. This component is not required by IMI at present but could be taken up again in future reports if IMI wish to revisit this option.

# 3 DATA SOURCES, INDICATORS AND INTERPRETATION

## 3.1 BIBLIOMETRIC DATA AND CITATION ANALYSIS

### 3.1.1 BACKGROUND

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 is 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 recognised as having a greater impact and Thomson Reuters has shown that high citation rates are correlated with other qualitative evaluations of research performance, such as peer review.[1] 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 normalised to account for such variations by field. Because citation counts naturally grow over time it is essential to account for growth by year. Normalisation is usually done by reference to the relevant global average for the field and for the year of publication.

Bibliometric indicators have been found to be more informative for core natural sciences, especially for basic science, than they are for applied and professional areas and for social sciences. In professional areas the range of publication modes used by leading researchers is likely to be diverse as they target a diverse, non-academic audience. In social sciences there is also a diversity of publication modes and citation rates are typically much lower than in natural sciences.

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 analysing 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.

### 3.1.2 PUBLICATION AND CITATION DATA SOURCES

For this project, the Thomson Reuters data platform *ScienceWire*® has been used to identify publications associated with IMI funding and individual researchers. This platform has been developed to support program evaluation and research analytics using up-to-date multi-source data on research publications, funded research projects, patents and other research-related activities. It includes publications data from MEDLINE as well as the Thomson Reuters *Web of Science*® as well as data on other entities in publicly available and proprietary databases.

Citation data have been sourced from Thomson Reuters databases underlying the *Web of Knowledge*℠, which gives access to conference proceedings, patents, websites, and chemical structures, compounds and reactions in addition to journals. It has a unified structure that integrates

[1] *Evidence* Ltd. (2002) Maintaining Research Excellence and Volume: A report by Evidence Ltd to the Higher Education Funding Councils for England, Scotland and Wales and to Universities UK. (*Adams J, et al.)* 48pp .

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* is part of the *Web of Knowledge*, and focuses on research published in journals and conferences in science, medicine, arts, humanities and social sciences. The authoritative, multidisciplinary content covers over 12 000 of the highest impact journals worldwide, including Open Access journals and over 150 000 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'. Thomson Reuters has extensive experience with databases on research inputs, activity and outputs and has developed innovative analytical approaches for benchmarking and interpreting international, national and institutional research impact.

Granularity of analysis is an important issue. Unduly fine analysis at the level of research groups provides little comparability or connectedness, while coarse analysis may miss spikes of excellence in key areas.

Journals are mapped to one or more subject categories, and every article within that journal is subsequently assigned to that category. Thomson Reuters uses these categories as the basis for bibliometric analysis because they are well-established and informed by extensive work with the research community since inception. Papers from prestigious, 'multidisciplinary' and general 'biomedical' journals such as Nature, Science, BMJ, The Lancet, 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 citing and cited references in each article. Further information about the journals included in the citation databases and how they are selected is available here: http://scientific.thomsonreuters.com/mjl/.

The bibliometric evaluation of research covered in this report has been based principally on citation analysis of research published between January 2010 and December 2013 with citation counts as at April 2014 for all 'current' indicators and citation counts as at end-2013 for all indicators calculated with reference to world citation baselines (e.g. normalised citation impact).

Annex 4 provides the standard methodology and data definitions used in bibliometric and citation analyses. A summary of bibliometric and citation data definitions is given in Section 3.1.3.

### 3.1.3 BIBLIOMETRIC AND CITATION DATA DEFINITIONS AND INDICATORS

**Citations**: 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 Thomson Reuters is estimated to attract about 95% of global citations.

**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). Impact can be calculated for papers within a specific research field such as Clinical Neurology, or for a specific institution or group of institutions, or a specific country. Citation count declines in the most recent years of any time-period as papers have had less time to accumulate citations (papers published in 2009 will typically have more citations than papers published in 2013).

**Citation velocity/hot papers**: Citation velocity is the rate at which a paper accumulates citations. Most papers reach their citation peak some time after publication. A small number of papers, however, accumulate citations rapidly (high citation velocity) and may represent breakthroughs in the field(s) to which they relate.

**Field-normalised citation impact ($NCI_F$)**: Citation rates vary between research fields and with time, consequently, analyses must take both field and 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 normalisation factor is the world average citations per paper for the year and journal category in which the paper was published. This normalisation is also referred to as 'rebasing' the citation count.

**H-index**: The h-index was developed by JE Hirsch as in indicator of both productivity and impact.[2] The value of the index h is equal to the number of papers (N) in the list that have N or more citations, while the remaining papers have fewer than N citations. Therefore, a researcher who has published 30 papers, of which 17 have received 17 or more citations while the remaining 13 have received fewer than 17 citations, has an h-index of 17. Irrespective of research impact, older researchers in more prolific fields tend to have a higher h-index.

Thomson Reuters **Hot Papers** database tracks and identifies papers with high citation velocities relative to their field and age. To identify hot papers, papers published in the last two years are selected and frequency distributions compiled for citations received in the most recent two-month period. To correct for variation in citation rates between different research fields, separate distributions are made for each field. The 22 *Essential Science Indicators*® fields used in this classification are documented here: http://archive.sciencewatch.com/about/met/fielddef/. Thresholds are set to find the top fraction of papers in each field – typically 0.1% of papers meet this threshold and are classified as **hot papers**.

**Interdisciplinarity/diffusion score**: This is indicated by the number and disparateness of the fields from which publications citing an IMI publication originate, summarised in a diffusion score developed by Carley and Porter.[3] The diffusion score is a measure of the applicability of new knowledge across subject areas and represents a measure of the robustness of the findings in the published article. The diffusion score incorporates features of traditional measures of diversity in assessing the balance and distribution of citations arising from different subject categories that in substance very different from one another. For example, while an article A receiving 5 citations from Physics, Applied and 5 citations from Chemistry, Physical and an article B receiving 5 citations from Physics, Applied and 5 citations from Physiology would have the same diversity, the diffusion score would be greater for article B since the two fields from which the citations originate are very different from one another.

**Journal-normalised citation impact ($NCI_J$)**: Another bibliometric indicator which can be very useful in small datasets is the journal-normalised citation impact, $NCI_J$. This indicator is calculated from the citation impact relative to the specific journal in which the publication appears.

For the publication in Annex 4 which has been cited 115 times to end-December 2012, the expected citation rate for a publication in Acta Biomaterialia published in 2005 would be 28.7 and the $NCI_J$ would be 4.01. Therefore, this publication has been cited more than expected for the journal.

For a set of publications, we calculate the quality index as the percentage of publications which are cited more than expected for the relevant journals.

This indicator should be considered alongside that of field-normalised citation impact as they are complementary. For example, a given set of publications may have a high quality index and relatively low average field-normalised citation impact. This would imply that these publications were well cited in relation to other papers in that journal and that year but when considered in relation to other publications in the same research field did not perform as well. The interpretation would be that the publications are in relatively low impact journals.

**Journal Impact Factor (JIF)**: In the same way that citation impact can be used as an index of research quality, the average number of citations per paper can be used to indicate the impact and/or importance of a journal. The Impact Factor for a journal (JIF) is calculated using data for a three-year period. For example, the 2012 Impact Factor for a given journal is calculated is calculated by Thomson Reuters as the average number of times which articles from the journal published in the past two years (2010 and 2011) were cited in 2012. Thus, a JIF of 2.0 means that, on average, the articles published in 2010 or 2011 have been cited twice. Citing articles may be from the same journal; however, most citing articles are from other journals.

For the journal, *Fertility and Sterility*, the 2012 journal Impact Factor would be calculated as follows:

---

[2] Hirsch, J.E. (2005) *Proceedings of the National Academy of Sciences of the United States of America* **102** (46): 16569-16572.

[3] Carley S, Porter A (2012). A forward diversity index. Scientometrics, 90:407-427.

| | | | |
|---|---|---|---|
| Cites in 2012 to items published in 2010 = | 3 259 | Number of items published in 2010 = | 744 |
| Cites in 2012 to items published in 2011 = | 2 556 | Number of items published in 2011 = | 649 |
| **Total** | **5 815** | | **1 393** |

$$\frac{Number\ of\ citations}{Number\ of\ \mathrm{items}} = \frac{5\,815}{1\,393} = \mathbf{4.174}$$

The calculation of the journal Impact Factor is fully described on the Thomson Reuters website at: http://thomsonreuters.com/products_services/science/free/essays/impact_factor/.

When looking at journal Impact Factor data it is important to remember that, as citation rates vary between research fields and publication type, these will affect the JIF. That is a JIF of 4.174 ranks the journal *Fertility and Sterility* 4th out of 77 journals in the Obstetrics & Gynaecology journal category and therefore in the top quartile. However, the journal *Journal of Alzheimer's Disease* with the same JIF of 4.174 is ranked in the second quartile (64th out of 251 journals) in the journal category Neurosciences.

**Journal top quartile**: This indicator is defined as the quartile in which the journal appears when ranked by Journal Impact Factor among all journals in that category.

**Mean normalised citation impact (mNCI)**: The mean NCI indicator for any specific dataset is calculated as the mean of the field-normalised citation impact ($NCI_F$) of all papers within that dataset.

**Papers/publications**: Thomson Reuters abstracts publications including editorials, meeting abstracts and book reviews as well as research journal articles. The terms 'paper' and 'publication' are often used interchangeably to refer to printed and electronic outputs of many types.

For clarity, in this report:

- **Publication** is used inclusively to refer to all IMI publications whether linked to Thomson Reuters citation data or not.
- ***Web of Science* publication** is used exclusively to refer to those IMI publications which have been linked to Thomson Reuters citation data.
- **Paper** is used exclusively to refer only to substantive *Web of Science* publications (journal articles, reviews and some proceedings papers) that have been linked to Thomson Reuters citation data. This definition excludes editorials, meeting abstracts or other types of publication. Papers are the subset of publications for which citation data are most informative and which are used in calculations of citation impact.

**Percentage of highly-cited papers**: For the purpose of this report, highly-cited papers have been defined as those articles and reviews which belong to the world's top decile of papers in that journal category and year of publication, when ranked by number of citations received. A percentage that is above 10 indicates above-average performance.

**Research field**: Standard bibliometric methodology uses journal category as a proxy for research field. Journals are assigned to one or more categories, and every article within that journal is subsequently assigned to that category. Publications from prestigious, 'multidisciplinary' and general medical journals such as Nature, Science, The Lancet, 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 here http://scientific.thomsonreuters.com/mjl/. For this evaluation, the standard classification of *Web of Science* journal categories has been used.

### 3.1.4 INTERPRETATION OF BIBLIOMETRIC INDICATORS AND CITATION ANALYSES

The following points should be borne in mind when considering the results of these analyses.

- IMI JU only started to fund projects in May 2009. Of the 46 active projects, 23 were launched since 1 January 2012. It may take several years for a project to progress from inception to the point where it has generated sufficient data for a publication. It may take further years until it has produced its most valuable results. The IMI JU projects that will be analysed are therefore relatively young, and early bibliometric indicators may not fully reflect their eventual impact.

- 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 analysing 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.
- As noted above many of the publications associated with IMI JU-funded projects are relatively recent. Publications accumulate citations over time and it may take years until a given publication is cited. While citation counts in early years have been shown to reflect long-term citation performance,[4] indicators based on citation counts may be relatively more volatile in the years immediately following publication.
- Citation rates vary between disciplines and fields. For example, for the UK science base as a whole, ten years produces a general plateau beyond which few additional citations would be expected. On the whole, citations accumulate more rapidly and plateau at a higher level in biomedical sciences than physical sciences, and natural sciences generally cite at a higher rate than social sciences.

INDICATOR THRESHOLDS

- **Papers**: The minimum number of papers suitable as a sample for quantitative research evaluation is a subject of widespread discussion. Larger samples are always more reliable, but a very high minimum may defeat the scope and specificity of analysis. Experience has indicated that a threshold between 20 and 50 papers can generally be deemed appropriate. For work that is likely to be published with little contextual information, the upper boundary (≥ 50) is a desirable starting point. For work that will be used primarily by an expert, in-house group then the lower boundary (≥ 20) may be approached. Because comparisons for in-house evaluation often involve smaller, more specific research groups (compared to broad institutional comparisons) a high volume threshold is self-defeating. Smaller samples may be used but outcomes must be interpreted with caution and expert review should draw on multiple information sources before reaching any conclusions.
- **Field normalised citation impact**: such values for individual papers vary widely and it is more useful to consider the average for a set of papers. This average can be at several granularities: field (either journal category or field), annual and overall (total output under consideration). When considering such average data points, care must be taken to understand that these data are highly skewed and the average can be driven by a single, highly-cited paper (this would be highlighted in accompanying text though not apparent from Tables & Figures). The world average is 1.0, so any value higher than this indicates a paper, or set of papers, which are cited more than average for similar research worldwide. For research management purposes, experience suggests that values between 1.0 and 2.0 should be considered to be indicative of research which is influential at a national level whilst that cited more than twice the world average has international recognition.

**Research field**: A problem frequently encountered in the analysis of data about the research process is that of ‘mapping’. For example, a funding body allocates money for chemistry but this goes to researchers in biology and engineering as well as to chemistry departments. Clinicians publish in mathematics and education journals. Publications in environmental journals come from a diversity of

[4] Adams, J. *et al.* (2002) Maintaining Research Excellence and Volume: A report by Evidence Ltd to the Higher Education Funding Councils for England, Scotland and Wales and to Universities UK, http://www.hefce.ac.uk/pubs/rdreports/2002/rd08_02/rd08_02.pdf

disciplines. This creates a problem when we try to define, for example, 'Parasitology research'. Is this the work funded under Parasitology programmes, the work of researchers in Parasitology units or the work published in Parasitology journals? For the first two options we need to track individual grants and researchers to their outputs, which is feasible but not within the scope of this study nor for every comparator institution. Therefore, to create a simple and transparent dataset of equal validity across time and geography, we rely on the set of journals associated with Parasitology as a proxy for the body of research reflecting the field.

### 3.1.5 DATASET DEFINITIONS USED IN THE BIBLIOMETRIC INDICATORS AND CITATION ANALYSES

**IMI project publications/papers**: This dataset comprises publications from IMI-supported projects as described in Section 4.1 and outlined in Figure 4.1.1.

**IMI researcher publications/papers**: This dataset comprises publications by IMI-supported researchers as described in Section 6 and outlined in Figure 6.2.1.

**Similar European research**: this benchmark dataset has been created using the EU-27 grouping of countries: Thomson Reuters *National Science Indicators* 2013 database and only research falling into the same journal categories as in the IMI project dataset.

**Wellcome Trust publications/papers**: this benchmark dataset has been created using specific keyword searches on funding acknowledgment data in Thomson Reuters *Web of Science* to define those publications where the Wellcome Trust has been acknowledged as a funder. This is the same process by which IMI project publications have been identified.

# 4 CITATION ANALYSIS – IMI-SUPPORTED PUBLICATIONS OVERALL

This Section of the report presents analyses of the output and citation impact of IMI projects considered overall and compared to the IMI-researcher dataset collated for all researchers supported by IMI (Section 6.1). IMI project research is also benchmarked against similar European research (see footnote on page 26) and research associated with Wellcome Trust funding (Section 4.7).

Publications for analyses include all IMI-supported publications identified in Thomson Reuters *Web of Science℠* to date – that is, publications new to this report (Annex 1) as well as publications identified in the previous reports. The census point for inclusion of publications into the third report was mid-August 2013. The census point for inclusion of publications into this fourth report was December 2013. Therefore, this report reflects changes in IMI activity between these points. Citation counts for all publications included previously have been updated from the mid-August census point used in the third report. Furthermore, the third report used citations benchmarked against end-of-2012 citations globally, whereas this report takes into account benchmarks up to the end of 2013.

When considering the analyses in this Section, earlier caveats regarding paper numbers should be borne in mind (Section 3.1.4).

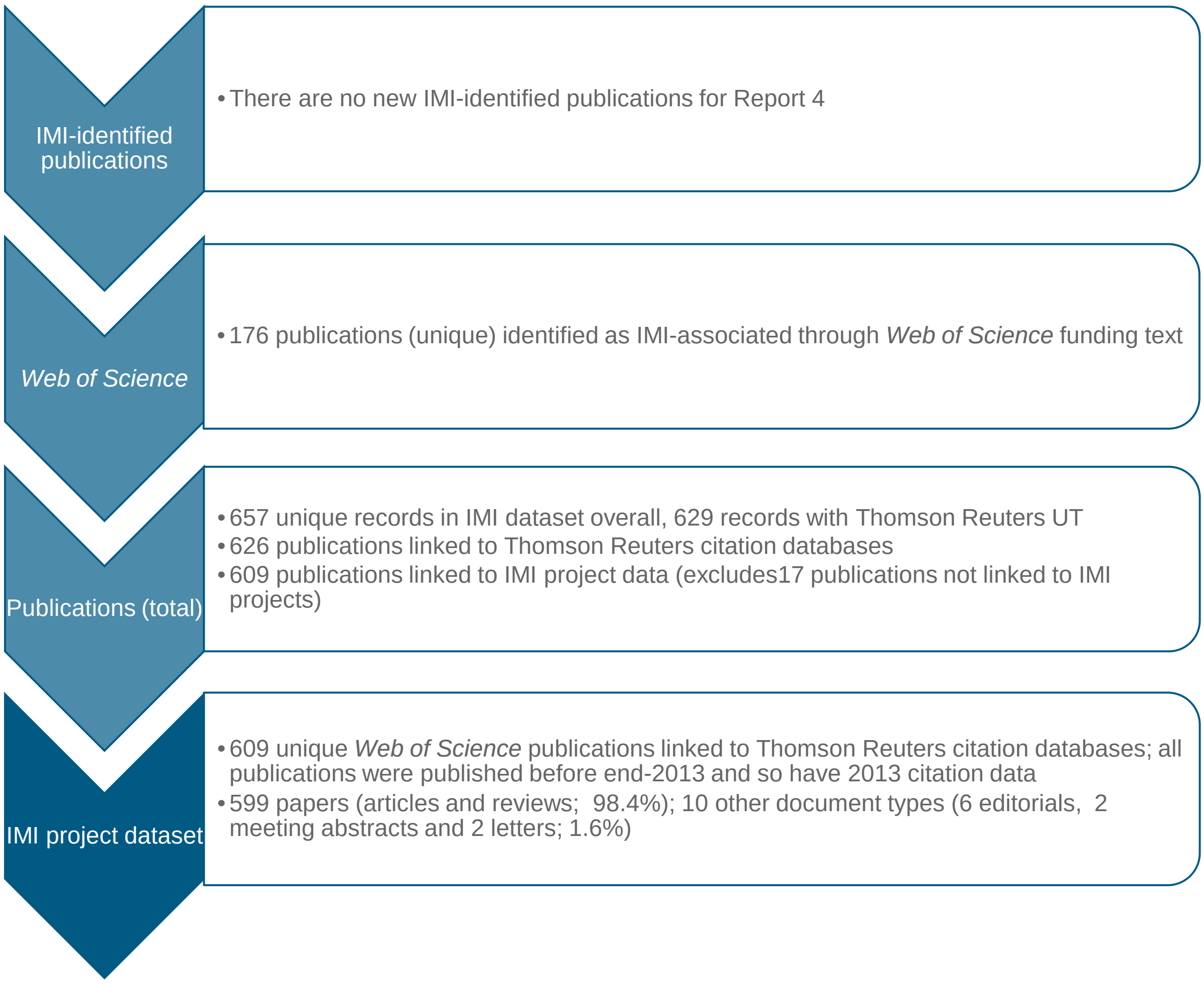
## 4.1 PUBLICATIONS FROM IMI-SUPPORTED PROJECTS

Publications from IMI-supported projects were identified using bibliographic data supplied by IMI, or through specific keyword searches using funding acknowledgment data in *Web of Science*. As in the third report, the analyses in this report include all projects from Calls 1 to 6.

The aggregated list of publications was reviewed by Thomson Reuters and supplied to IMI for further verification prior to inclusion in the analyses. Seventeen publications have not been assigned to specific projects despite review by IMI personnel. Eleven publications have been identified for Call 4 compared to 4 publications for Call 6, but none for Call 5. As these projects were only awarded funding recently (late 2012 or early 2013) and it may take years for a project to progress towards published research this should not be taken as evidence of low productivity. Three publications have been assigned to more than one project.

The process of identifying publications from IMI-supported projects which have Thomson Reuters citation data is outlined in Figure 4.1.1. The final dataset (the IMI project dataset) has changed relative to Dataset 6 in the third report as the normalised citation impact of papers at end-2013 can now be calculated. To benchmark IMI project research performance, these datasets are also compared to publications collated for IMI-supported researchers (Section 6.1), similar European research (see footnote on page 26) and research acknowledging Wellcome Trust funding (Section 4.7).

FIGURE 4.1.1 IDENTIFYING PUBLICATIONS FROM IMI-SUPPORTED PROJECTS WITH THOMSON REUTERS CITATION DATA




### 4.1.1 CITATION DATA FOR PUBLICATIONS FROM IMI-SUPPORTED PROJECTS

A total of 657 unique publication records from IMI-supported projects were identified. From these, 629 unique publications were matched to Thomson Reuters *Web of Science* of which 3 were not extracted within the database parameters used in this project and 17 were excluded on the basis that they were not linked to a specific IMI project by IMI personnel. Therefore, 609 of these unique publications were linked to records in *Web of Science*.

These citation counts have been sourced from the citation databases which underlie Thomson Reuters *Web of Science* and have been extracted at two distinct census points: current (April 2014) and end-2013. The 'current' census point allows assessment of the performance of IMI research from as up-to-date a viewpoint as possible through calculation of 'raw' citation impact (see Section 3.1.3). This, however, does not allow benchmarking of IMI research performance against the world and European average. The end-2013 census point is used to evaluate the citation impact of IMI-supported research relative to world average (normalised citation impact) and has the same end-2013 census point used in the calculation of global citation baselines (see Section 3.1.3). Normalised bibliometric indicators have been calculated using standard methodology and the Thomson Reuters *National Science Indicators* (NSI) database for 2013.

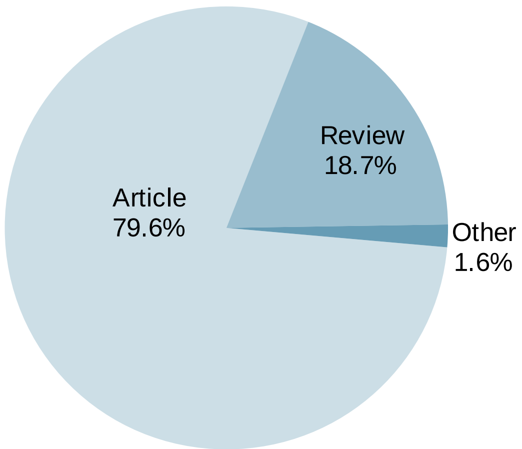
## 4.2 SHARE OF PAPERS RELATIVE TO OTHER PUBLICATION TYPES

FIGURE 4.2.1 CATEGORISATION OF IMI PROJECT RESEARCH BY DOCUMENT TYPE

Figure 4.2.1 shows the share of papers (articles and reviews) relative to other document types, for all *Web of Science* publications from IMI-associated projects. Papers are the subset of publications for which citation data are most informative and which are used in calculations of normalised citation impact.

IMI project research comprises 609 unique *Web of Science* publications (as outlined in Figure 4.1.1). Some 98.4% of these documents (599) are substantive articles and reviews with only ten documents not falling into this grouping. These documents (classified as 'Other') comprise six editorials, two meeting abstracts and two letters.

The distribution of document types is similar to that observed in the third report to IMI.



## 4.3 TRENDS IN PUBLICATION OUTPUT

Figure 4.3.1 presents the annual numbers of *Web of Science* publications in this fourth report to IMI.

- IMI project research continues to show substantial growth with just over one-third (34.1%) of publications new to this report.
- IMI projects have generated more than 300 publications in 2013 (307 in total) and growth looks set to continue.

FIGURE 4.3.1 NUMBER OF *WEB OF SCIENCE* PUBLICATIONS BY YEAR

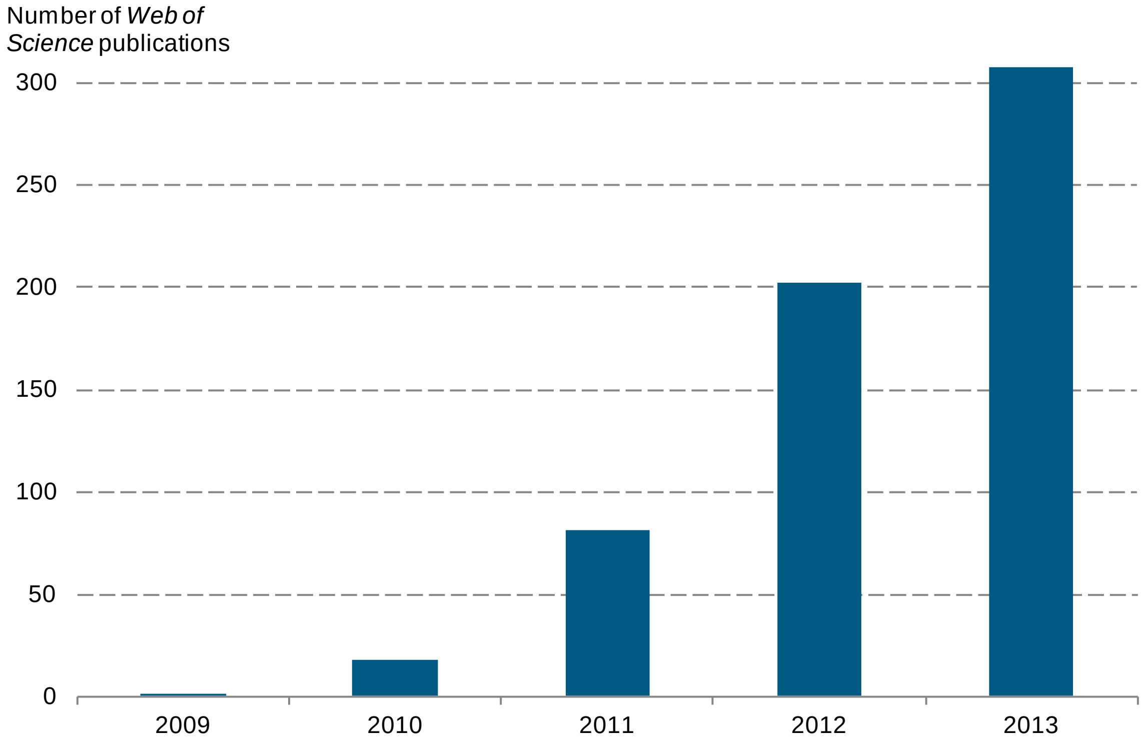
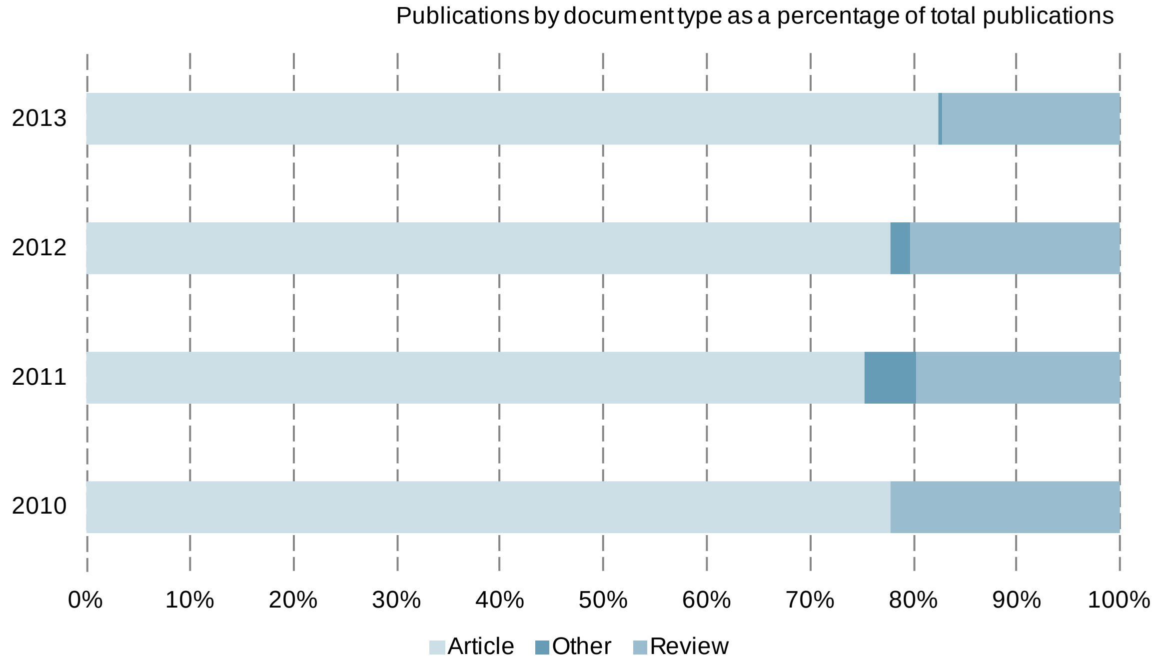


Figure 4.3.2 shows the proportion of papers (articles and reviews) relative to other document types for IMI project research from 2010 to date.[5]

IMI projects continue to generate a high proportion of papers relative to other document types with reviews typically accounting for around one-fifth of output.

FIGURE 4.3.2 CATEGORISATION OF *WEB OF SCIENCE* PUBLICATIONS BY YEAR AND DOCUMENT TYPE




[5] 2009 publications comprise a single meeting abstract – this has been omitted from Figure 4.3.2 for clarity.

## 4.4 IN WHICH JOURNALS DO IMI PROJECT PUBLICATIONS APPEAR MOST FREQUENTLY?

The 16 journals appearing most frequently[6] in the IMI project publications dataset, 2009-2013, are listed in Table 4.4.1. A total of 108 journal titles are used more than once.

Together, the 16 most frequently used journals cover 184 *Web of Science* publications, around one third (30.2%) of the total number of items in the dataset.

PLoS One is the journal in which IMI project publications appear most frequently (28 publications) followed by Annals of the Rheumatic Diseases (26 publications), exclusively from the Call 2 project BTCure and associated with the *Web of Science* journal category of Rheumatology.

Though there is a strong focus on Rheumatology, the core set of journals for IMI projects continues to highlight the diversity of IMI-supported research. There are multidisciplinary titles (such as PLoS One), as well as specialised titles in other disease areas such as diabetes (such as Diabetologia).

All but one of the journals (Molecular Informatics) in Table 4.4.1 are in the top quartile when ranked by Journal Impact Factor among all journals in that category. Molecular Informatics is ranked in the second quartile of journals within its journal category (Mathematical & Computational Biology).

IMI project publications have been published in a total of 301 journals, of which 194 are ranked in the top quartile (by Journal Impact Factor) of journals in their specific research fields. A total of 453 publications (74.4% of IMI project publications) have been published in these well regarded journals. The average journal impact factor of all the journals in which IMI project publications have been published is 6.05.

[6] Table 4.4.1 uses a frequency threshold of at least five publications. This is a change from the third report where this threshold was at least four publications. Below this threshold a further 11 journal titles would be included.

TABLE 4.4.1 JOURNALS IN WHICH IMI PROJECT PUBLICATIONS HAVE BEEN PUBLISHED MOST FREQUENTLY (2009-2013), RANKED BY NUMBER OF *WEB OF SCIENCE* PUBLICATIONS

| Journal | Number of *Web of Science* publications | Number of papers | Journal Impact Factor (2012) | *Web of Science* journal categories |
|---|---|---|---|---|
| PLoS One | 28 | 28 | 3.730 | Multidisciplinary Sciences |
| Annals of the Rheumatic Diseases | 26 | 25 | 9.111 | Rheumatology |
| Arthritis and Rheumatism | 20 | 19 | 7.477 | Rheumatology |
| Drug Safety | 18 | 18 | 3.408 | Public, Environmental & Occupational Health<br>Pharmacology & Pharmacy<br>Toxicology |
| Pain | 16 | 16 | 5.644 | Anesthesiology<br>Clinical Neurology<br>Neurosciences |
| Molecular Informatics | 13 | 13 | 2.338 | Chemistry, Medicinal<br>Mathematical & Computational Biology |
| Proceedings of the National Academy of Sciences of the United States of America | 8 | 8 | 9.737 | Multidisciplinary Sciences |
| Arthritis Research & Therapy | 7 | 7 | 4.302 | Rheumatology |
| Diabetes | 7 | 7 | 7.895 | Endocrinology & Metabolism |
| Diabetologia | 7 | 7 | 6.487 | Endocrinology & Metabolism |
| Journal of Clinical Investigation | 7 | 7 | 12.812 | Medicine, Research & Experimental |
| European Journal of Cancer | 6 | 6 | 5.061 | Oncology |
| Pharmacogenomics | 6 | 6 | 3.857 | Pharmacology & Pharmacy |
| Drug Discovery Today | 5 | 5 | 6.551 | Pharmacology & Pharmacy |
| Molecular Psychiatry | 5 | 5 | 14.897 | Biochemistry & Molecular Biology<br>Neurosciences<br>Psychiatry |
| Neuroimage | 5 | 5 | 6.252 | Neurosciences<br>Neuroimaging<br>Radiology, Nuclear Medicine & Medical Imaging |

Table 4.4.2 lists the 20 journals with the highest Journal Impact Factor (JIF) in the IMI-supported publications dataset.

Overall, there are 64 publications in journals with an impact factor of 10 or above, an increase of 8 publications over the third report. Of these, 17 publications appear in journals with an impact factor of 20 or above.

Together the top 20 journals by Journal Impact Factor account for around one twentieth (5.4%) of all IMI-supported publications.

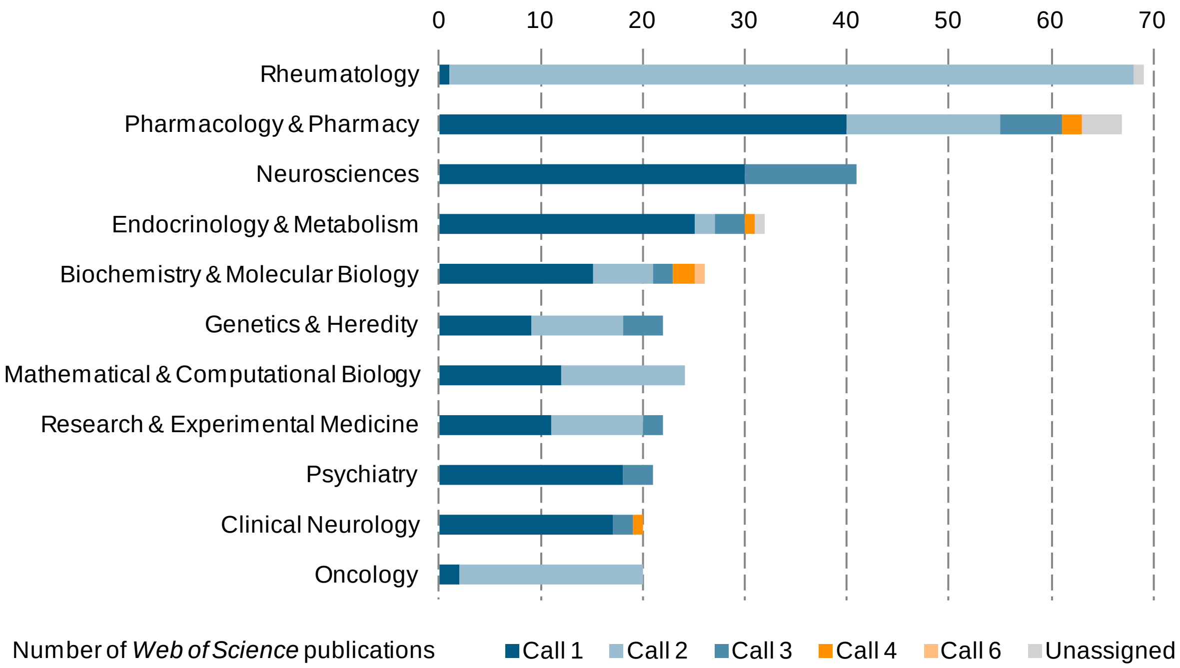
TABLE 4.4.2 JOURNALS IN WHICH IMI PROJECT PUBLICATIONS HAVE BEEN PUBLISHED MOST FREQUENTLY (2009-2013), TOP TWENTY RANKED BY JOURNAL IMPACT FACTOR

| Journal | Number of *Web of Science* publications | Number of papers | Journal Impact Factor (2012) | Journal categories |
|---|---|---|---|---|
| Nature | 2 | 2 | 38.597 | Multidisciplinary Sciences |
| Nature Genetics | 3 | 1 | 35.209 | Genetics & Heredity |
| Nature Reviews Drug Discovery | 1 | | 33.078 | Biotechnology & Applied Microbiology<br>Pharmacology & Pharmacy |
| Nature Biotechnology | 1 | | 32.438 | Biotechnology & Applied Microbiology |
| Science | 1 | 1 | 31.027 | Multidisciplinary Sciences |
| JAMA - Journal of the American Medical Association | 1 | 1 | 29.978 | Medicine, General & Internal |
| Nature Immunology | 1 | 1 | 26.199 | Immunology |
| Nature Medicine | 2 | 2 | 24.302 | Biochemistry & Molecular Biology<br>Cell Biology<br>Medicine, Research & Experimental |
| Lancet Neurology | 3 | 3 | 23.917 | Clinical Neurology |
| Nature Methods | 1 | 1 | 23.565 | Biochemical Research Methods |
| Pharmacological Reviews | 1 | 1 | 22.345 | Pharmacology & Pharmacy |
| British Medical Journal | 2 | 2 | 17.215 | Medicine, General & Internal |
| Neuron | 1 | 1 | 15.766 | Neurosciences |
| Nature Reviews Neurology | 1 | 1 | 15.518 | Clinical Neurology |
| PLoS Medicine | 2 | 2 | 15.253 | Medicine, General & Internal |
| Nature Neuroscience | 1 | 1 | 15.251 | Neurosciences |
| Nature Reviews Clinical Oncology | 1 | 1 | 15.031 | Oncology |
| Molecular Psychiatry | 5 | 5 | 14.897 | Biochemistry & Molecular Biology<br>Neurosciences<br>Psychiatry |
| American Journal of Psychiatry | 1 | 1 | 14.721 | Psychiatry |
| Alzheimers & Dementia | 2 | 2 | 14.483 | Clinical Neurology |

## 4.5 WHICH RESEARCH FIELDS ACCOUNT FOR THE HIGHEST VOLUME OF IMI PROJECT PUBLICATIONS?

Figure 4.5.1 shows the top eleven[7] *Web of Science* journal categories[8] by rank associated with IMI project research.

FIGURE 4.5.1 TOP ELEVEN *WEB OF SCIENCE* JOURNAL CATEGORIES IN WHICH IMI PROJECT RESEARCH IS PUBLISHED




- Just over one-tenth (11.3%) of IMI project research is assigned to the journal category of Rheumatology, and a similar percentage to Pharmacology & Pharmacy (11.0%). These are the top two most frequently used *Web of Science* journal categories in the IMI project dataset.
- Output in Rheumatology journals is predominantly associated with BTCure (Call 2) with one publication from PROTECT (Call 1), highlighting the productivity of BTCure. By contrast, output in Pharmacology & Pharmacy is amongst a diverse group of projects in Calls 1 and 3.
- Neurosciences is the third most frequent *Web of Science* journal category, related to a diverse group of projects in Calls 1 (EUROPAIN, NEWMEDS and PharmaCog) and 3 (EU-AIMS).
- The most frequently used journal categories in Figure 4.5.1 continue to reflect breadth and depth in the IMI research portfolio which contains cross-cutting and more specialised research.

The analysis presented in Figure 4.5.1 includes all publication types and spans the full time period of IMI-supported publications (2009-2013).

Standard descriptions of the scope of these journal categories are given in Annex 2.

[7] Eleven *Web of Science* journal categories are included due to Clinical Neurology and Oncology both having 20 publications.

[8] Journals can be associated with more than one *Web of Science* category. This analysis is based on the best-performing category (i.e. that in which it ranks highest in terms of overall citations relative to journal category and year).

## 4.6 IS IMI PROJECT RESEARCH WELL-CITED?

Citation impact of research, an indicator linked to the accumulation of citations, is subject specific. Typically, papers published in areas such as biomedical research receive more citations than papers published in subjects such as engineering even if the papers are published in the same year. All citation impact data presented in this report are therefore normalised, or rebased, to the relevant world average to allow comparison between years and fields.

Tables 4.6.1 and 4.6.2 present a summary of the citation analyses of research from IMI-supported projects compared with the IMI researcher dataset (Section 6.1). Table 4.6.1 presents a viewpoint of IMI-supported papers at the end of 2013 using indicators where citation impact has been normalised against world average values. Table 4.6.2 presents a more recent (but also more descriptive) viewpoint using indicators based on current (April 2014) citation counts (see Section 4.1.1).

### SUMMARY OF KEY FINDINGS

The citation impact of papers associated with IMI projects or papers by IMI-supported researchers is internationally influential with citation impact at or above twice world average and around three-quarters of output in top quartile journals (Table 4.6.1).

- The citation impact for IMI project papers is 2.05 (where world average is 1.0) for the 4-year period, 2010-2013. This is similar to the findings of the third report (average citation impact = 2.04), a key finding of which was that citation impact had increased substantially from the second report (average citation impact = 1.55). This indicates that the quality of IMI-associated research (as indexed by citation impact) has been maintained while output has continued to grow.
- The citation impact for IMI project papers is lower, on average, than the citation impact for the IMI researchers dataset (2.26).
- The citation impact for IMI project papers remains well above the EU's average citation impact[9,10] relative to the world baseline for the same 3-year period in similar research fields (average citation impact = 1.15).

More than one-tenth of publications from IMI projects or IMI-supported researchers (11.9% and 11.3%, respectively – Table 4.6.2) are published in open access journals that are listed in the Directory of Open Access Journals (DOAJ). DOAJ does not include all open access journals, so these figures are likely to be underestimates. These values are above the global average reported by a 2011 study reviewing accessibility in the journal literature between 2003 and 2009[11] though it should be noted that the majority of IMI research has been published since 2009 and the global average share of publications that are openly accessible is likely to have grown.

Overall, though the IMI projects dataset is small in comparison with the IMI researchers dataset, these data show that the performance of the two groups is similar.

[9] EU-27 grouping of countries: Thomson Reuters *National Science Indicators* 2013 database; similar research has been defined as including the same journal categories as in the IMI project dataset.

[10] For this analysis, only papers are considered since only these publication types have normalised citation impact data (see Section 3.1.3).

[11] Laakso *et al*. (2011) The Development of Open Access Journal Publishing from 2003 to 2009, PLOS ONE, 6(6), e20961.

TABLE 4.6.1 SUMMARY CITATION ANALYSIS FOR IMI RESEARCH – CITATIONS TO END-2013

| | Citation impact | | | | |
|---|---|---|---|---|---|
| 2010-2013 | **Number of papers** | **Normalised at field level** | **Normalised at journal level** | **Average percentile** | **% Highly-cited papers**[12] |
| IMI projects | 599 | 2.05 | 1.21 | 46.8 | 24.4% |
| IMI researchers | 16 813 | 2.26 | 1.37 | 42.3 | 20.6% |

TABLE 4.6.2 SUMMARY CITATION ANALYSIS FOR IMI RESEARCH – CITATIONS TO CURRENT

| | **IMI publications** | | ***Web of Science* publications** | | | |
|---|---|---|---|---|---|---|
| 2010 to current | **Total** | **% Open access journals**[13] | **Total** | **Citations** | **Raw citation impact** | **% Top quartile journals**[14] |
| IMI projects | 657 | 11.9% | 609 | 3 742 | 6.14 | 74.4% |
| IMI researchers | n/a | 11.3% | 21 158 | 211 055 | 9.98 | 71.9% |

Disaggregation by journal category shows strengths in the IMI project publications dataset.

Figure 4.6.1 shows that the citation impact of IMI project research in all the top eleven journal categories is, on average, well above the citation impact of similar European research and, as in the third report, in four categories is well above the citation impact for the IMI researchers dataset.

IMI project research in Pharmacology & Pharmacy, Neurosciences, Psychiatry and Research & Experimental Medicine has substantially higher citation impact, on average, than similar research by IMI-supported researchers.

IMI project research in Endocrinology & Metabolism, Mathematical & Computational Biology and Oncology has substantially lower citation impact, on average, than similar research by IMI-supported researchers.

[12] 'Highly-cited' refers to those articles and reviews belonging to the world's top decile of papers for journal category and year of publication. A percentage that is above 10 indicates above-average performance.

[13] For this report, we have considered a journal as open access if listed in the Directory of Open Access journals (http://www.doaj.org/).

[14] This indicator is based upon the quartile in which the journal appears when ranked by Journal Impact Factor among all journals in that category. Journal ranking data have been sourced from Thomson Reuters *Journal Citation Reports* database.

FIGURE 4.6.1 CITATION IMPACT OF IMI-SUPPORTED PAPERS, BY RESEARCH FIELD (JOURNAL CATEGORY) BENCHMARKED AGAINST PAPERS BY IMI-SUPPORTED RESEARCHERS AND SIMILAR PAPERS FROM THE EUROPEAN RESEARCH BASE

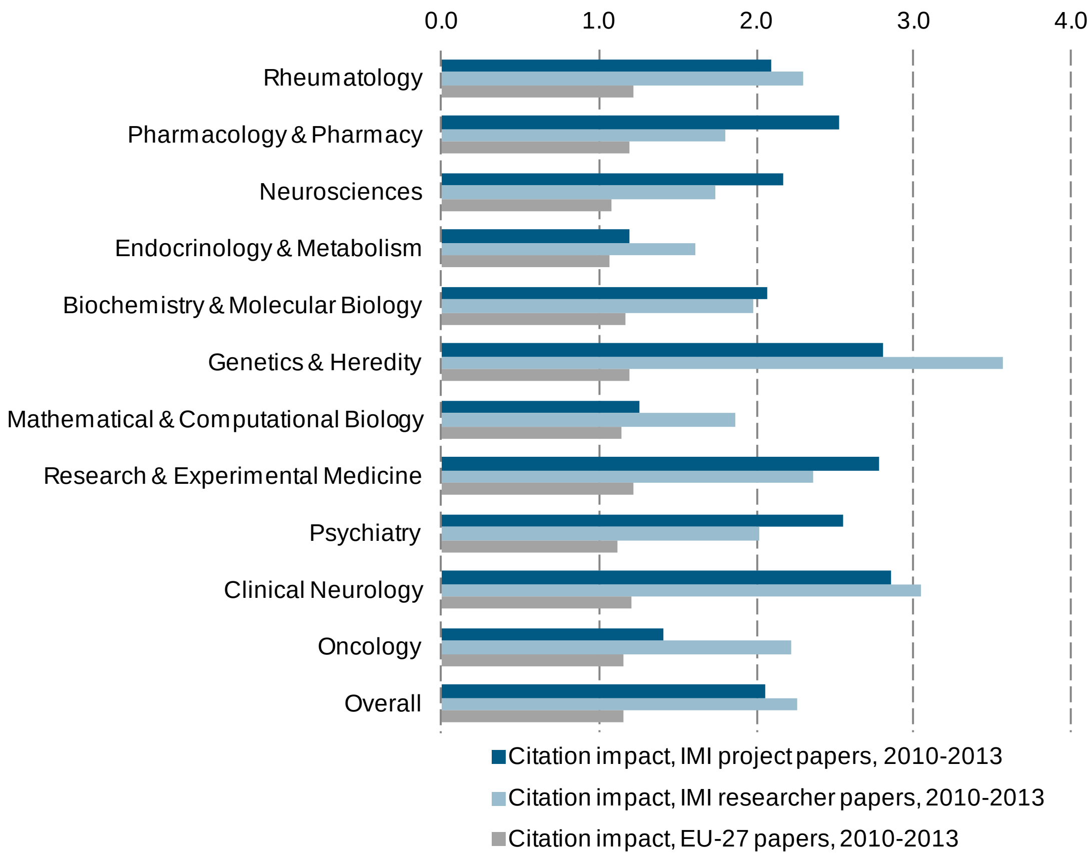



TABLE 4.6.3 SUMMARY OF PUBLICATION OUTPUT AND 3-YEAR AVERAGE CITATION IMPACT FOR IMI-SUPPORTED RESEARCH BY TOP *WEB OF SCIENCE* JOURNAL CATEGORIES, 2010-2013 BENCHMARKED AGAINST THE IMI RESEARCHERS DATASET AND SIMILAR PUBLICATIONS FROM THE EUROPEAN RESEARCH BASE

| | IMI projects | | IMI researchers | | EU-27 | |
|---|---|---|---|---|---|---|
| ***Web of Science* journal category** | **Number of papers**[15] | **Citation impact** | **Number of papers** | **Citation impact** | **Number of papers** | **Citation impact** |
| Rheumatology | 68 | 2.10 | 1 115 | 2.30 | 9 772 | 1.21 |
| Pharmacology & Pharmacy | 96 | 2.53 | 1 840 | 1.80 | 47 449 | 1.19 |
| Neurosciences | 97 | 2.17 | 1 832 | 1.74 | 59 495 | 1.08 |
| Endocrinology & Metabolism | 36 | 1.20 | 956 | 1.62 | 29 010 | 1.07 |
| Biochemistry & Molecular Biology | 49 | 2.07 | 1 547 | 1.98 | 78 214 | 1.17 |
| Genetics & Heredity | 31 | 2.81 | 1 203 | 3.56 | 30 911 | 1.19 |
| Mathematical & Computational Biology | 29 | 1.26 | 220 | 1.86 | 9 082 | 1.14 |

[15] Papers can be assigned to more than one journal category and so may be counted towards the number of papers in more than one category.

| | IMI projects | | IMI researchers | | EU-27 | |
|---|---|---|---|---|---|---|
| ***Web of Science* journal category** | **Number of papers**[15] | **Citation impact** | **Number of papers** | **Citation impact** | **Number of papers** | **Citation impact** |
| Research & Experimental Medicine | 24 | 2.78 | 522 | 2.36 | 21 130 | 1.22 |
| Psychiatry | 42 | 2.55 | 829 | 2.02 | 28 171 | 1.12 |
| Clinical Neurology | 48 | 2.86 | 1 148 | 3.05 | 40 961 | 1.21 |
| Oncology | 28 | 1.42 | 799 | 2.22 | 47 644 | 1.15 |
| **Overall** | **599** | **2.05** | **16 813** | **2.26** | **1 879 908** | **1.15** |

It is important to note that IMI projects have far fewer papers in each of these categories than either benchmark, and that low paper numbers can mean that citation impact values will be more susceptible to skew by especially well-cited papers or large numbers of uncited papers.

These analyses therefore give a useful indication of IMI project research performance relative to comparators but it should be borne in mind that this performance may change as IMI paper numbers increase further.

Standard definitions of the scope of the journal categories in Figure 4.6.1 and Table 4.6.3 are given in Annex 2.

## 4.7 HOW DOES IMI PROJECT RESEARCH COMPARE WITH RESEARCH FROM OTHER BIOMEDICAL FUNDING ORGANISATIONS?

This Section of the report evaluates trends in the performance of IMI project research and benchmarks this performance against the Wellcome Trust, a leading UK funder of biomedical research.

Wellcome Trust publications were identified by specific keyword searches using funding acknowledgment data in *Web of Science*. This is the same process by which IMI project publications have been identified (Section 4.1). This dataset is referred to as 'Wellcome Trust research' in the analyses below. Papers resulting from Wellcome Trust funding where the authors have not specifically acknowledged this funding are not covered by this dataset.

Data are presented for a 4-year time period (2010-2013) for papers where normalised citation impact can be calculated. This provides a reference for both datasets relative to world citation baselines ensuring an even basis for comparison.

### 4.7.1 TRENDS IN THE OUTPUT AND CITATION IMPACT OF IMI PROJECT RESEARCH COMPARED WITH WELLCOME TRUST RESEARCH

Figures 4.7.1 to 4.7.3 show trends in the output and performance of IMI project research compared with Wellcome Trust research. When considering these analyses, earlier caveats regarding paper numbers should be borne in mind (Section 3.1.4).

In summary:

- IMI project research shows year-on-year growth. This is to be expected as the number of projects funded by IMI increases and this growing body of projects yields results for publication (Figure 4.7.1).
- The citation impact of IMI project research is over twice world average overall (2.05). This suggests very highly-cited internationally significant research.
- The citation impact of IMI project research is similar to the citation impact of Wellcome Trust research which is just below twice world average (1.95). However, the citation impact of IMI project research has been more variable on trend, compared to Wellcome Trust research, which has remained stable at around twice world average over the time period (Figure 4.7.1). This is to be expected given the relatively small size of the IMI project dataset, and its growth relative to the output of the Wellcome Trust as a more established research institution.
- Though paper numbers are low (n=18) the citation impact of IMI project research published in 2010 is exceptional at over four times world average. This performance is driven by several highly-cited papers rather than a single paper with exceptional citation impact (Figure 4.7.1).
- IMI project research and Wellcome Trust research show a similar profile of uncited research over the time period. No IMI project papers published in 2010 are uncited (Figure 4.7.2).
- A quarter (24.4%) of IMI project research is highly-cited (ranked in the world's top decile relative to journal category and year). This is comparable to the Wellcome Trust (23.4%) indicating a body of internationally significant, highly-cited research.
- Though IMI is a 'young' funding initiative in comparison to the Wellcome Trust, the performance of IMI project research and Wellcome Trust is comparable across the majority of indicators assessed. An exception is the percentage of uncited papers at overall level, which is substantially higher for IMI than for the Wellcome Trust, though the underlying trends are similar. This is attributable to the differences in output trends observed above – as around half (51.1%) of IMI project research was published in 2013. As more recent research is less likely to be cited than older research this should not be taken as evidence that IMI project research is more likely to remain uncited (Table 4.7.1, Figure 4.7.1).

FIGURE 4.7.1 TRENDS IN OUTPUT AND CITATION IMPACT – IMI PROJECT RESEARCH COMPARED WITH WELLCOME TRUST RESEARCH

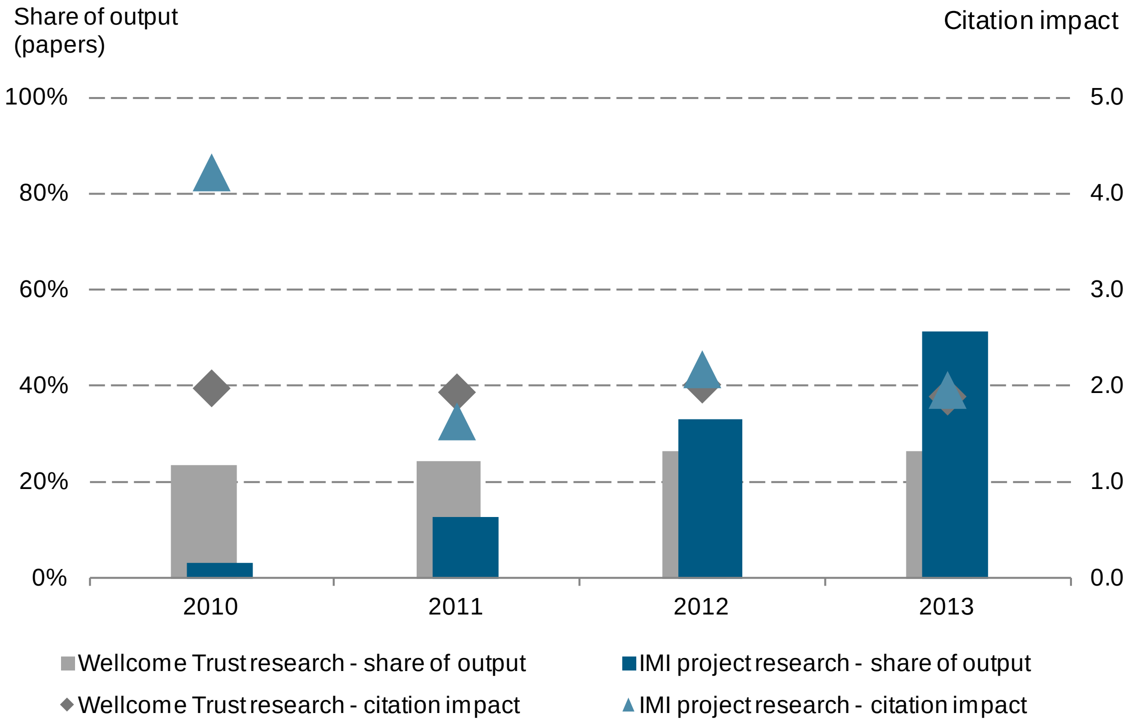



FIGURE 4.7.2 TRENDS IN UNCITED RESEARCH – IMI PROJECT RESEARCH COMPARED WITH WELLCOME TRUST RESEARCH

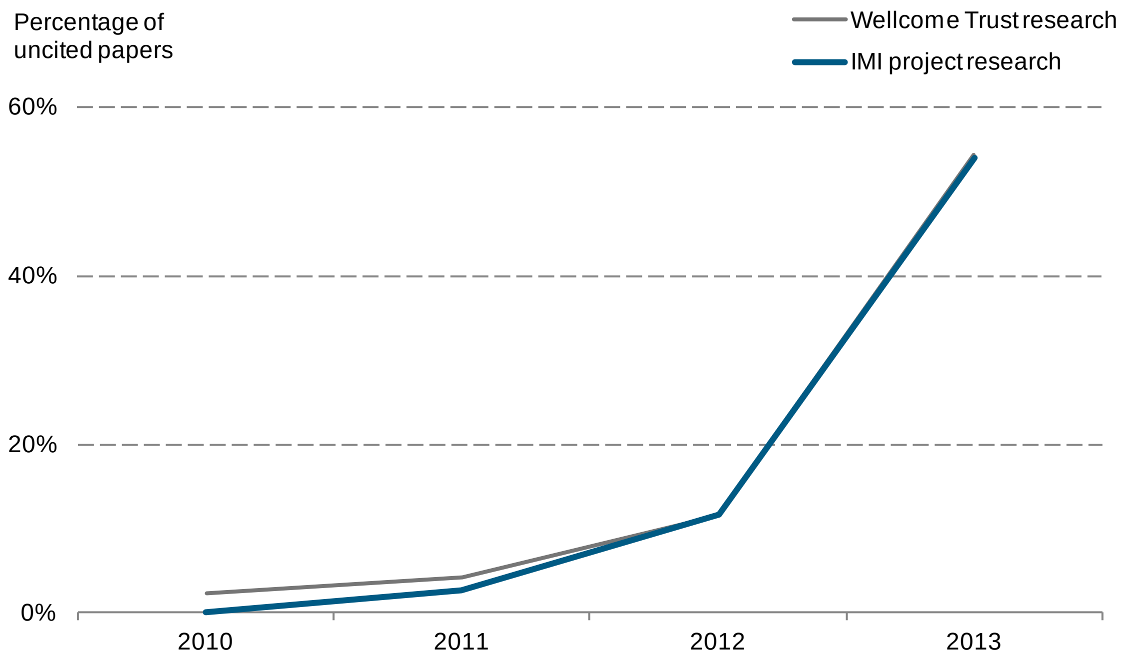


FIGURE 4.7.3 TRENDS IN HIGHLY-CITED RESEARCH – IMI PROJECT RESEARCH COMPARED WITH WELLCOME TRUST RESEARCH

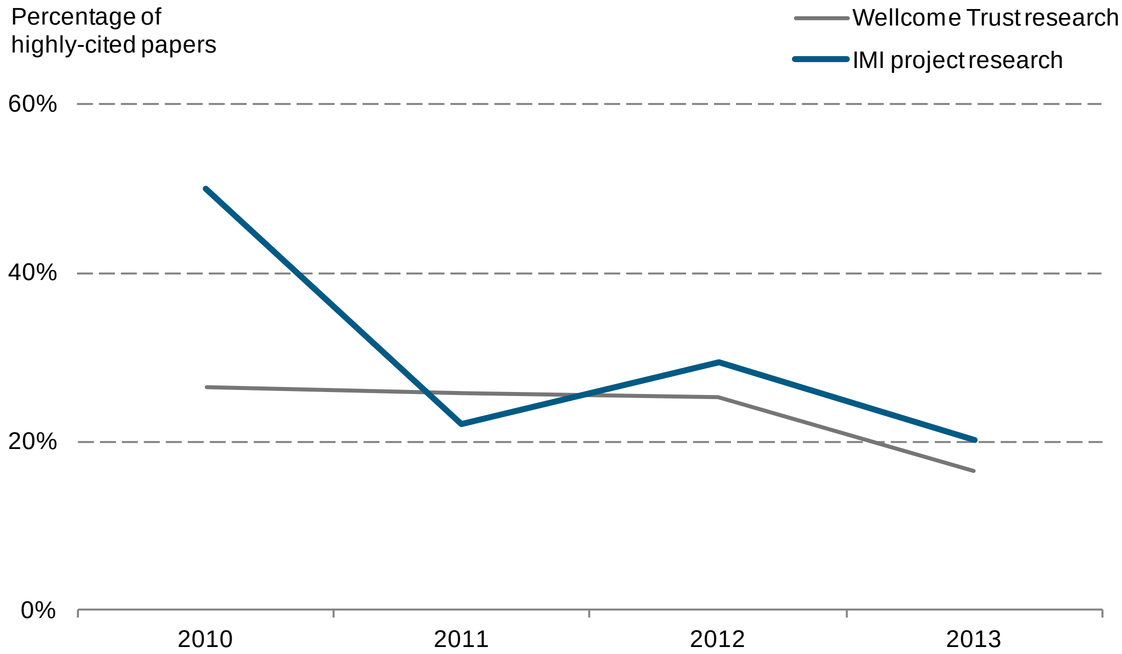



TABLE 4.7.1 SUMMARY BIBLIOMETRIC INDICATORS – IMI PROJECT RESEARCH COMPARED WITH WELLCOME TRUST RESEARCH, 2010-13

| | Number of papers | Citation impact (normalised at field level) | Percentage of uncited papers | Percentage of highly-cited papers |
|---|---|---|---|---|
| IMI project research | 599 | 2.05 | 31.7% | 24.4% |
| Wellcome Trust research | 21 421 | 1.95 | 18.9% | 23.4% |

## 4.8 HOW COLLABORATIVE IS IMI PROJECT RESEARCH?

International research collaboration is a rapidly growing element of research activity.[16] The reasons for this have not been fully clarified but include increasing access to facilities and resources, increasing access to knowledge and increasing access to people and expertise. In addition, international collaboration has been shown to be associated with an increase in the number of citations received by research papers, although this does depends on the partner countries involved.[17] Co-authorship is likely to be a good indicator of collaboration, although there will be collaborations that do not result in co-authored papers, and co-authored papers which may have required limited collaboration. Alternative data-based approaches, for example using information about co-funding or international exchanges, have limitations in terms of both comprehensiveness and validity.

In this report, co-authorship is used as a measure of collaboration. Table 4.8.1 compares the output and citation impact of IMI project papers that are co-authored between different sectors, institutions and countries. Sectors are those used in the IMI researchers dataset (Section 6.1).

The data in Table 4.8.1 show that IMI project research is collaborative at sector, institution and country level.

- Nearly two-thirds (64.3%) of all IMI project papers have been published by researchers affiliated with different sectors, including between academia and other research organisations. This is an increase from 61.3% in the third report. Nearly one-third (32.6%) of these-two-thirds are collaborations between the public sector and industry.
- Three-quarters (74.7%) of IMI project papers are collaborative between institutions.
- Half (50.6%) of all IMI project papers have are internationally collaborative.
- Collaborative IMI project research is internationally influential with citation impact well over twice world average (1.0) and a clear margin over non-collaborative IMI project research.

TABLE 4.8.1 CROSS-SECTOR, CROSS-INSTITUTION AND INTERNATIONAL OUTPUT– IMI PROJECT RESEARCH

| | **Number of papers** | **Percentage of papers** | **Citation impact (normalised at field level)** |
|---|---|---|---|
| Cross-sector | 396 | 64.3% | 2.26 |
| Single-sector | 220 | 35.7% | 1.71 |
| Cross-institution | 460 | 74.7% | 2.22 |
| Single-institution | 156 | 25.3% | 1.59 |
| International | 312 | 50.6% | 2.44 |
| Domestic | 304 | 49.4% | 1.67 |

A paper is defined as cross-sector if the listed addresses are from more than one sector. For example, if a paper has two addresses – University of Copenhagen and Novartis – it would be classified as cross-sector. If a paper has only two addresses –University of Cambridge and Utrecht University – it would be classified as single-sector.

A paper is defined as cross-institution if more than one institution is listed in the addresses.

A paper is defined as international if more than one country is listed in the addresses or domestic if a single country is listed.

[16] Adams J (2013). Collaborations: the fourth age of research. Nature, 497, 557-560.

[17] Adams, J., Gurney, K., & Marshall, S. (2007). Patterns of international collaboration for the UK and leading partners. A report by *Evidence* Ltd to the UK Office of Science and Innovation. 27pp.

# 5 CITATION ANALYSIS – AT IMI PROJECT LEVEL

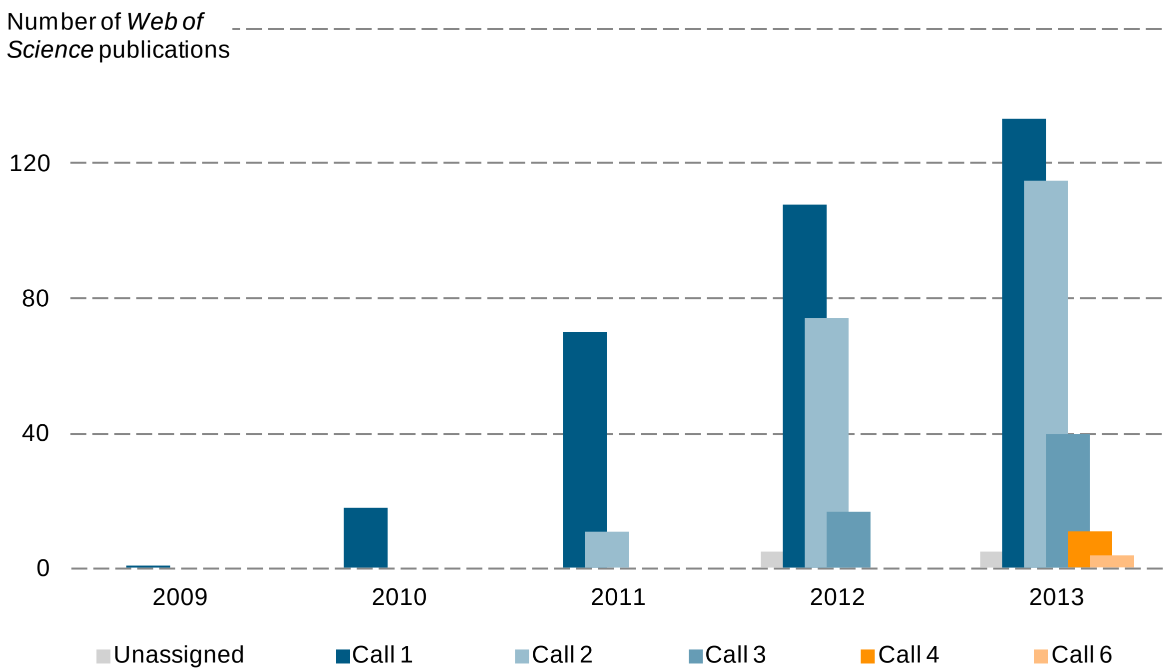
This Section of the report presents project level analyses of the publication output and citation impact of IMI research. Data are presented for projects in Calls 1 to 3 as there are insufficient publications associated with projects from Calls 4 and 6 (and no publications associated with projects from Call 5) for more detailed citation analyses.

When considering the analyses in this Section, earlier caveats regarding paper numbers should be borne in mind (Section 3.1.4).

## 5.1 TRENDS IN PUBLICATION OUTPUT BY IMI FUNDING CALL

The data in Figure 5.1.1 show that the majority of IMI-supported publications and papers are associated with Calls 1 and 2, however, output has also increased for Call 3 from 17 publications (2012) to 40 publications in 2013. There has been notable growth in output in Call 2 between 2011 and 2012, with 2013 output now approaching Call 1 totals.

FIGURE 5.1.1 NUMBER OF *WEB OF SCIENCE* PUBLICATIONS BY YEAR AND FUNDING CALL




There are 11 publications for Call 4 projects and 4 publications for Call 6 projects, and as yet no publications for Call 5 projects. This is not unexpected as projects associated with these Calls are very recent and it may take time for projects to generate results for publication.

## 5.2 SUMMARY BIBLIOMETRIC ANALYSES FOR IMI PROJECTS – CALL 1

Figure 5.2.1 presents a 'bubble-chart' visualisation of IMI-supported research for those projects with at least 10 papers – at least one of which is highly-cited – over the time period (2010-2013). The number of papers, 4-year average citation impact and share of highly-cited papers are compared. The area of the 'bubble' is proportional to the percentage of highly-cited papers. The solid horizontal line indicates the average citation impact for all IMI project papers.

FIGURE 5.2.1 PAPER NUMBERS, 4-YEAR AVERAGE CITATION IMPACT AND SHARE OF HIGHLY-CITED RESEARCH FOR SELECTED IMI PROJECTS – CALL 1

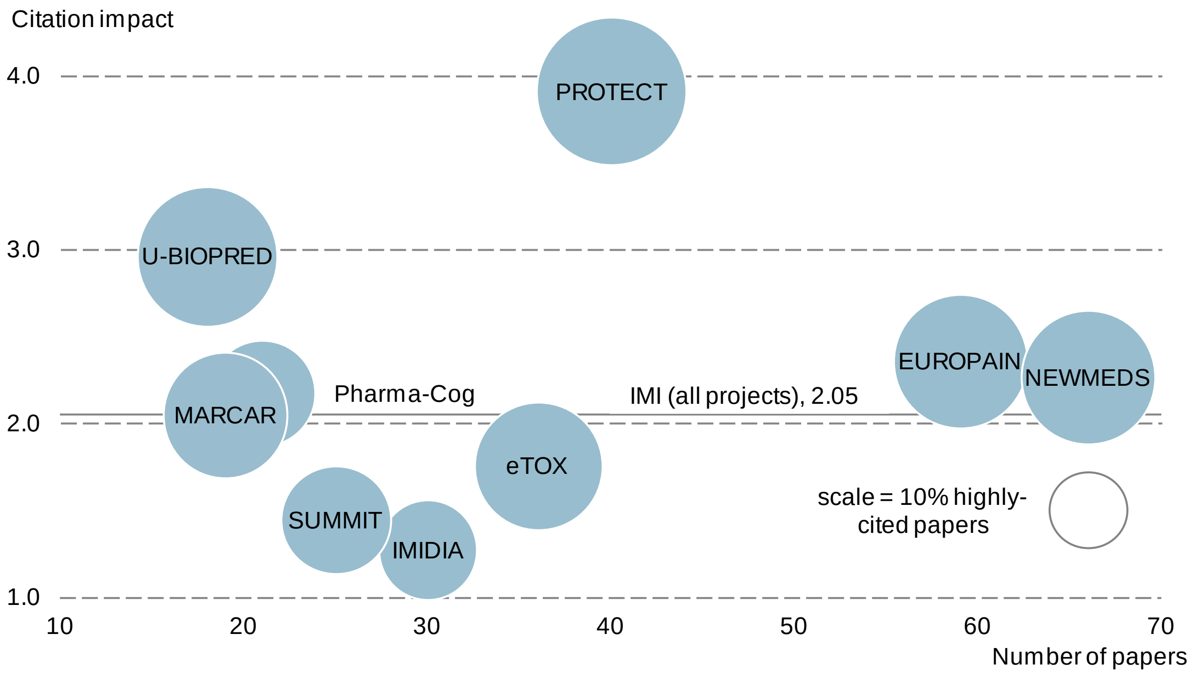



Figure 5.2.1 has been updated from the third report to cover a 4-year rather than 3-year time period; world citation baselines for 2013 became available in April 2014, enabling the calculation of normalised citation impact for 2013 papers (Section 4.1.1).

The data in Figure 5.2.1 show that:

- The average citation impact of all of these projects is above world average (1.0) and the percentage of highly-cited research is above world average (10%). This shows excellent research performance of IMI-associated research.
- Research associated with six of these projects is cited over twice world average (PROTECT, U-BIOPRED, EUROPAIN, NEWMEDS, Pharma-Cog and MARCAR).
- Research associated with the PROTECT project is cited at a level approaching four times world average (3.91) and U-BIOPRED cited nearly three times world average (2.96) though based on 18 papers. This is exceptionally highly-cited research.

Tables 5.2.1 and 5.2.2 compare bibliometric indicators for all projects in Call 1. Table 5.2.1 presents indicators where citation impact has been normalised against world average values and is an expansion of the data used in Figure 5.2.1. Table 5.2.2 presents a more recent (but also more descriptive) viewpoint using indicators based on current citation counts (see Section 4.1.1).

Three Call 1 projects (EMTRAIN, EU2P and Pharmatrain) have no *Web of Science* publications at the current time (shown in grey). Each of these projects has one publication associated with them but the journals in which the publications appear are not currently abstracted in *Web of Science*.

TABLE 5.2.1 SUMMARY BIBLIOMETRIC INDICATORS FOR IMI PROJECTS IN CALL 1 – CITATIONS TO END-2013

| Project | Number of papers | Citation impact: Normalised at field level | Citation impact: Normalised at journal level | Average percentile | % Highly-cited papers[18] |
|---|---|---|---|---|---|
| EMTRAIN | 0 | 0.00 | 0.00 | 0.0 | 0.0% |
| eTOX | 36 | 1.76 | 1.61 | 31.9 | 27.8% |
| Eu2P | 0 | 0.00 | 0.00 | 0.0 | 0.0% |
| EUROPAIN | 59 | 2.36 | 1.86 | 38.1 | 30.5% |
| IMIDIA | 30 | 1.27 | 1.00 | 50.5 | 16.7% |
| MARCAR | 19 | 2.05 | 1.23 | 51.3 | 26.3% |
| NEWMEDS | 66 | 2.27 | 1.23 | 44.3 | 30.3% |
| Pharma-Cog | 21 | 2.17 | 0.96 | 45.1 | 19.0% |
| Pharmatrain | 0 | 0.00 | 0.00 | 0.0 | 0.0% |
| PRO-active | 9 | 1.29 | 1.09 | 49.6 | 11.1% |
| PROTECT | 40 | 3.91 | 1.93 | 30.9 | 37.5% |
| SafeSciMET | 2 | 0.17 | 0.19 | 85.4 | 0.0% |
| SAFE-T | 2 | 0.96 | 0.55 | 57.2 | 0.0% |
| SUMMIT | 25 | 1.44 | 0.72 | 58.4 | 20.0% |
| U-BIOPRED | 18 | 2.96 | 1.72 | 33.1 | 33.3% |
| **Overall (IMI projects)** | **599** | **2.05** | **1.21** | **46.8** | **24.4%** |

TABLE 5.2.2 SUMMARY BIBLIOMETRIC INDICATORS FOR IMI PROJECTS IN CALL 1 – CITATIONS TO CURRENT

| Project | IMI publications: Total | IMI publications: % Open access journals | *Web of Science* publications: Total | Citations | Raw citation impact | % Top quartile journals |
|---|---|---|---|---|---|---|
| EMTRAIN | 1 | 0.0% | 0 | 0 | 0.00 | 0.0% |
| eTOX | 38 | 28.9% | 37 | 259 | 7.00 | 50.0% |
| Eu2P | 1 | 0.0% | 0 | 0 | 0.00 | 0.0% |
| EUROPAIN | 62 | 8.1% | 59 | 631 | 10.69 | 56.5% |
| IMIDIA | 34 | 2.9% | 30 | 196 | 6.53 | 76.5% |
| MARCAR | 20 | 25.0% | 19 | 60 | 3.16 | 85.0% |
| NEWMEDS | 73 | 8.2% | 66 | 498 | 7.55 | 72.6% |
| Pharma-Cog | 23 | 4.3% | 21 | 309 | 14.71 | 60.9% |
| Pharmatrain | 1 | 0.0% | 0 | 0 | 0.00 | 0.0% |
| PRO-active | 9 | 55.6% | 9 | 35 | 3.89 | 77.8% |
| PROTECT | 41 | 4.9% | 40 | 177 | 4.43 | 82.9% |
| SafeSciMET | 5 | 0.0% | 3 | 2 | 0.67 | 20.0% |

[18] 'Highly-cited' refers those articles and reviews belonging to the world's top decile of papers for journal category and year of publication. A percentage that is above 10 indicates above-average performance.

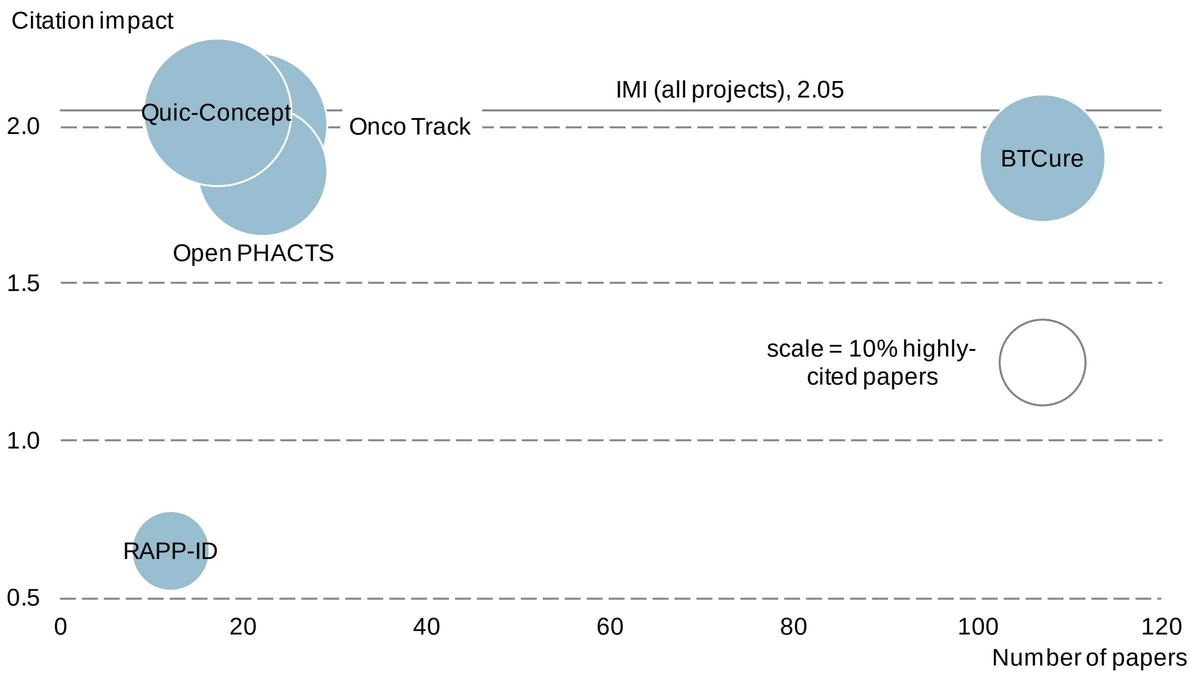
| | IMI publications | | | *Web of Science* publications | | |
|---|---|---|---|---|---|---|
| Project | Total | % Open access journals | Total | Citations | Raw citation impact | % Top quartile journals |
| SAFE-T | 4 | 25.0% | 3 | 15 | 5.00 | 50.0% |
| SUMMIT | 28 | 28.6% | 25 | 147 | 5.88 | 78.6% |
| U-BIOPRED | 18 | 0.0% | 18 | 181 | 10.06 | 83.3% |
| **Overall (IMI projects)** | **657** | **11.9%** | **609** | **3 742** | **6.14** | **74.4%** |

Bibliographic references for all highly-cited papers from IMI projects and the five papers with the highest citation velocity or interdisciplinarity (see Section 3.1.3) are listed in Annex 3.

## 5.3 SUMMARY BIBLIOMETRIC ANALYSES FOR IMI PROJECTS – CALL 2

Figure 5.3.1 presents an analysis of IMI-supported research for those projects with at least 10 papers – one of which is highly-cited – over the time period (2010-2013). The number of papers, 4-year average citation impact and share of highly-cited papers are compared. The area of the 'bubble' is proportional to the share of highly-cited papers. The solid horizontal line indicates the average citation impact for all IMI project papers.

FIGURE 5.3.1 PAPER NUMBERS, 4-YEAR AVERAGE CITATION IMPACT AND SHARE OF HIGHLY-CITED RESEARCH FOR SELECTED IMI PROJECTS – CALL 2



The data in Figure 5.3.1 show that:

- The average citation impact of most of these projects is above world average except RAPP-ID where citation impact is below world average (0.65) though based on 12 papers.
- BTCure is by far the most prolific IMI and Call 2 project with 107 publications at end-2013. The citation impact of this research is nearly twice world average (1.90).
- Research associated with Onco Track and Quic-Concept is very well-cited with citation impact around twice world average. Open PHACTS research is also well-cited (1.86).

Tables 5.3.1 and 5.3.2 compare bibliometric indicators for all projects in Call 2. Table 5.3.1 presents indicators where citation impact has been normalised against world average values and is an expansion of the data used in Figure 5.3.1. Table 5.3.2 presents a more recent (but also more descriptive) viewpoint using indicators based on current citation counts (see Section 4.1.1).

TABLE 5.3.1 SUMMARY BIBLIOMETRIC INDICATORS FOR IMI PROJECTS IN CALL 2 – CITATIONS TO END-2013

| | | Citation impact | | | |
|---|---|---|---|---|---|
| **Project** | **Number of papers** | **Normalised at field level** | **Normalised at journal level** | **Average percentile** | **% Highly-cited papers** |
| BTCure | 107 | 1.90 | 0.86 | 48.4 | 21.5% |
| DDMoRe | 13 | 0.58 | 0.65 | 74.5 | 0.0% |
| EHR4CR | 0 | 0.00 | 0.00 | 0.0 | 0.0% |
| Onco Track | 21 | 2.00 | 1.15 | 39.5 | 28.6% |
| Open PHACTS | 22 | 1.86 | 1.26 | 46.9 | 22.7% |
| Predect | 3 | 0.92 | 0.59 | 70.5 | 0.0% |
| Quic-Concept | 17 | 2.04 | 1.49 | 50.9 | 29.4% |
| RAPP-ID | 12 | 0.65 | 0.77 | 69.3 | 8.3% |
| **Overall (IMI projects)** | **599** | **2.05** | **1.21** | **46.8** | **24.4%** |

TABLE 5.3.2 SUMMARY BIBLIOMETRIC INDICATORS FOR IMI PROJECTS IN CALL 2 – CITATIONS TO CURRENT

| | IMI publications | | *Web of Science* publications | | | |
|---|---|---|---|---|---|---|
| **Project** | **Total** | **% Open access journals** | **Total** | **Citations** | **Raw citation impact** | **% Top quartile journals** |
| BTCure | 110 | 8.2% | 110 | 404 | 3.67 | 83.6% |
| DDMoRe | 14 | 21.4% | 13 | 9 | 0.69 | 64.3% |
| EHR4CR | 1 | 0.0% | 0 | 0 | 0.00 | 0.0% |
| Onco Track | 22 | 22.7% | 21 | 169 | 8.05 | 68.2% |
| Open PHACTS | 26 | 7.7% | 24 | 138 | 5.75 | 50.0% |
| Predect | 3 | 0.0% | 3 | 3 | 1.00 | 100.0% |
| Quic-Concept | 17 | 11.8% | 17 | 100 | 5.88 | 76.5% |
| RAPP-ID | 12 | 16.7% | 12 | 19 | 1.58 | 83.3% |
| **Overall (IMI projects)** | **657** | **11.9%** | **609** | **3 742** | **6.14** | **74.4%** |

There are no *Web of Science* publications associated with EHR4CR. This project had one IMI-associated publication, but the journal in which the publication appears is not currently abstracted in *Web of Science*.

Bibliographic references for all highly-cited papers from IMI projects and the five papers with the highest citation velocity or interdisciplinarity (see Section 3.1.3) are listed in Annex 3.

## 5.4 SUMMARY BIBLIOMETRIC ANALYSES FOR IMI PROJECTS – CALL 3

The numbers of papers from Call 3 projects at end-2013 are generally too few to allow a 'bubble-chart' visualisation of IMI-supported research at project level. An exception to this would be for EU-AIMS with 26 papers,11 of which are highly-cited (42.3%).

Tables 5.4.1 and 5.4.2 compare bibliometric indicators for all projects in Call 3. Table 5.4.1 presents indicators where citation impact has been normalised against world average. Table 5.4.2 presents a more recent (but also more descriptive) viewpoint using indicators based on current citation counts (see Section 4.1.1).

TABLE 5.4.1 SUMMARY BIBLIOMETRIC INDICATORS FOR IMI PROJECTS IN CALL 3 – CITATIONS TO END-2013

| | | Citation impact | | | |
|---|---|---|---|---|---|
| **Project** | **Number of papers** | **Normalised at field level** | **Normalised at journal level** | **Average percentile** | **% Highly-cited papers** |
| ABIRISK | 7 | 1.39 | 0.56 | 70.6 | 28.6% |
| BioVacSafe | 9 | 0.82 | 0.40 | 70.7 | 22.2% |
| DIRECT | 5 | 0.58 | 0.44 | 67.5 | 0.0% |
| EU-AIMS | 26 | 3.96 | 1.43 | 36.2 | 42.3% |
| EUPATI | 0 | 0.00 | 0.00 | 0.0 | 0.0% |
| MIP-DILI | 4 | 1.37 | 0.68 | 41.5 | 28.6% |
| ABIRISK | 7 | 1.39 | 0.56 | 70.6 | 22.2% |
| **Overall (IMI projects)** | **599** | **2.05** | **1.21** | **46.8** | **24.4%** |

TABLE 5.4.2 SUMMARY BIBLIOMETRIC INDICATORS FOR IMI PROJECTS IN CALL 3 – CITATIONS TO CURRENT

| | IMI publications | | *Web of Science* publications | | | |
|---|---|---|---|---|---|---|
| **Project** | **Total** | **% Open access journals** | **Total** | **Citations** | **Raw citation impact** | **% Top quartile journals** |
| ABIRISK | 7 | 28.6% | 7 | 9 | 1.29 | 42.9% |
| BioVacSafe | 11 | 9.1% | 10 | 30 | 3.00 | 45.5% |
| DIRECT | 5 | 20.0% | 5 | 4 | 0.80 | 40.0% |
| EU-AIMS | 27 | 7.4% | 27 | 290 | 10.74 | 70.4% |
| EUPATI | 0 | 0.0% | 0 | 0 | 0.00 | 0.0% |
| MIP-DILI | 5 | 20.0% | 4 | 11 | 2.75 | 80.0% |
| PreDiCT-TB | 4 | 25.0% | 4 | 1 | 0.25 | 75.0% |
| **Overall (IMI projects)** | **657** | **11.9%** | **609** | **3 742** | **6.14** | **74.4%** |

There are no IMI publications associated with EUPATI.

Bibliographic references for all highly-cited papers from IMI projects and the five papers with the highest citation velocity or interdisciplinarity (see Section 3.1.3) are listed in Annex 3.

## 5.5 TRENDS IN PUBLICATION OUTPUT AND CITATION IMPACT FOR IMI PROJECTS

Figure 5.5.1 and Figure 5.5.2 show the publication output and normalised citation impact of *Web of Science* publications associated with IMI projects.

Data for normalised citation impact are not shown where based on fewer than 5 publications. When considering the analyses in this Section, earlier caveats regarding paper numbers should be borne in mind (Section 3.1.4).

Call 1 Projects:

- NEWMEDS is the most prolific Call 1 project, followed by EUROPAIN, both cited more than twice world average. NEWMEDS research published in 2012 was cited more than three times world average (3.02) owing to one highly-cited paper.[19]
- The citation impact of the PROTECT project is markedly high (3.91). This is owing to several highly-cited papers published in the journal of Drug Safety. SUMMIT research (based on 19 publications) has one 'hot paper' associated with it.[20] The citation impact of U-BIOPRED research is associated with one highly-cited paper.[21]

Call 2 Projects:

- BTCure is the most prolific Call 2 and IMI project (107 publications) and its papers are cited nearly twice world average (1.90). Its output in 2013 expanded significantly with the addition of 72 publications, and citation impact also increased with this expansion in output.
- The projects of Open PHACTS, Onco Track and Quic-Concept are cited around twice world average, though based on around 20 publications each. By contrast, the citation impact of DDMoRe and RAPP-ID are over half world average (0.58 and 0.65 respectively) though based on smaller publication numbers.

Call 3 Projects:

- The research volumes for Call 3 projects (with the exception of EU-AIMS) are, as yet, still small. It is important not to over-interpret data for projects based on small paper numbers (such as ABIRISK, BioVacSafe and DIRECT).
- Equally publications numbers for EU-AIMS are small (26 publications) but currently cited at a level approaching four times world average (3.96). This is owing to the influence of a 'hot paper' published in Nature in 2012.[22]

19 NEWMEDS: KIROV, G et al. (2012) De novo CNV analysis implicates specific abnormalities of postsynaptic signalling complexes in the pathogenesis of schizophrenia, Molecular Psychiatry, 17: 142-153, doi: 038/mp.2011.154

20 SUMMIT: BOEKHOLDT, SM et al. (2012) Association of LDL Cholesterol, Non-HDL Cholesterol, and Apolipoprotein B Levels With Risk of Cardiovascular Events Among Patients Treated With Statins A Meta-analysis, JAMA - Journal of the American Medical Association, 307: 1302-1309

21 U-BIOPRED: MONTUSCHI, P et al. (2013) The Electronic Nose in Respiratory Medicine, Respiration, 85: 72-84, doi: 159/000340044

22 EU-AIMS: KONG, A et al. (2012) Rate of de novo mutations and the importance of fathers age to disease risk, Nature, 488: 471-475, doi: 038/nature11396

FIGURE 5.5.1 TRENDS IN OUTPUT VOLUME FOR IMI-SUPPORTED PROJECTS

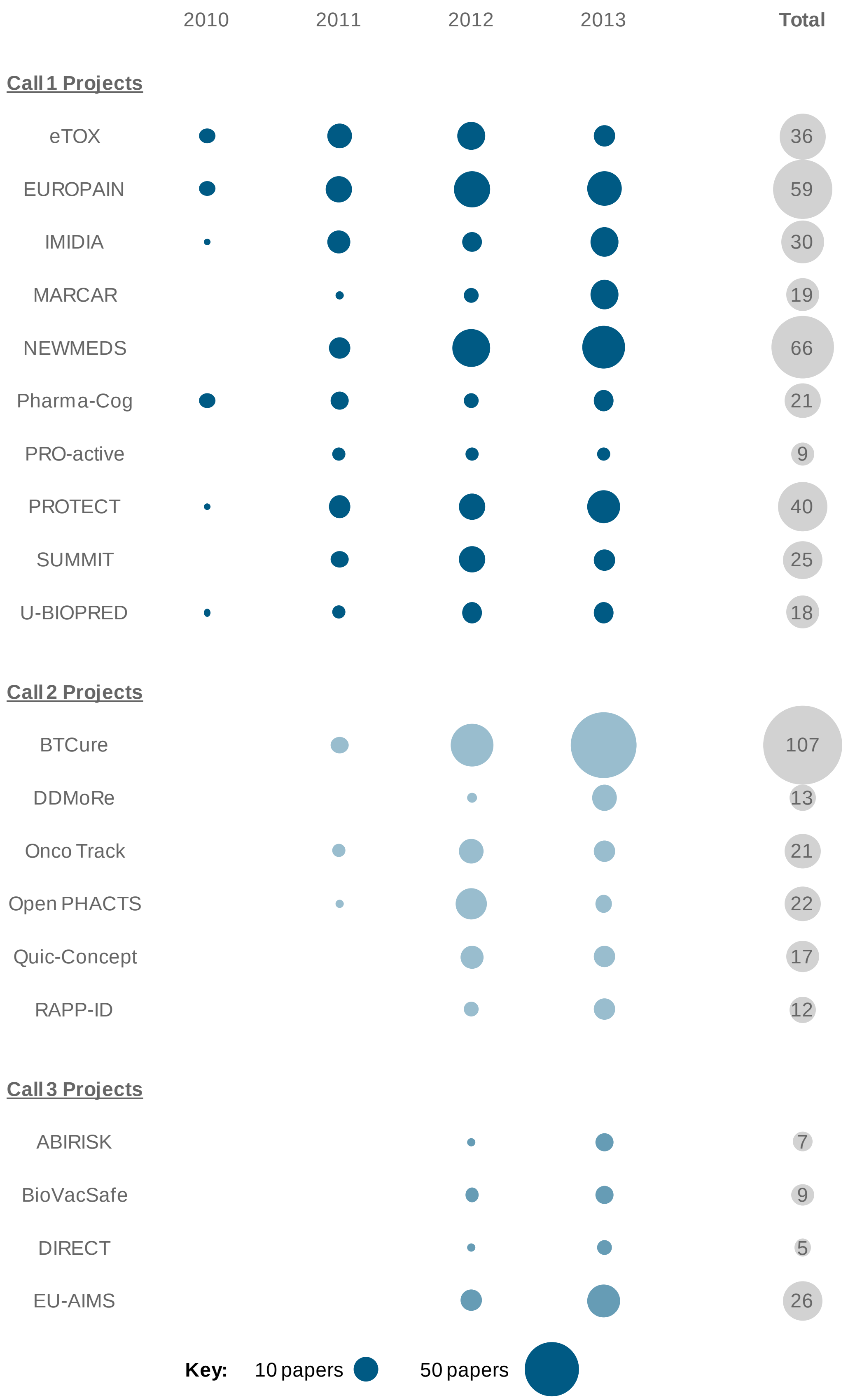
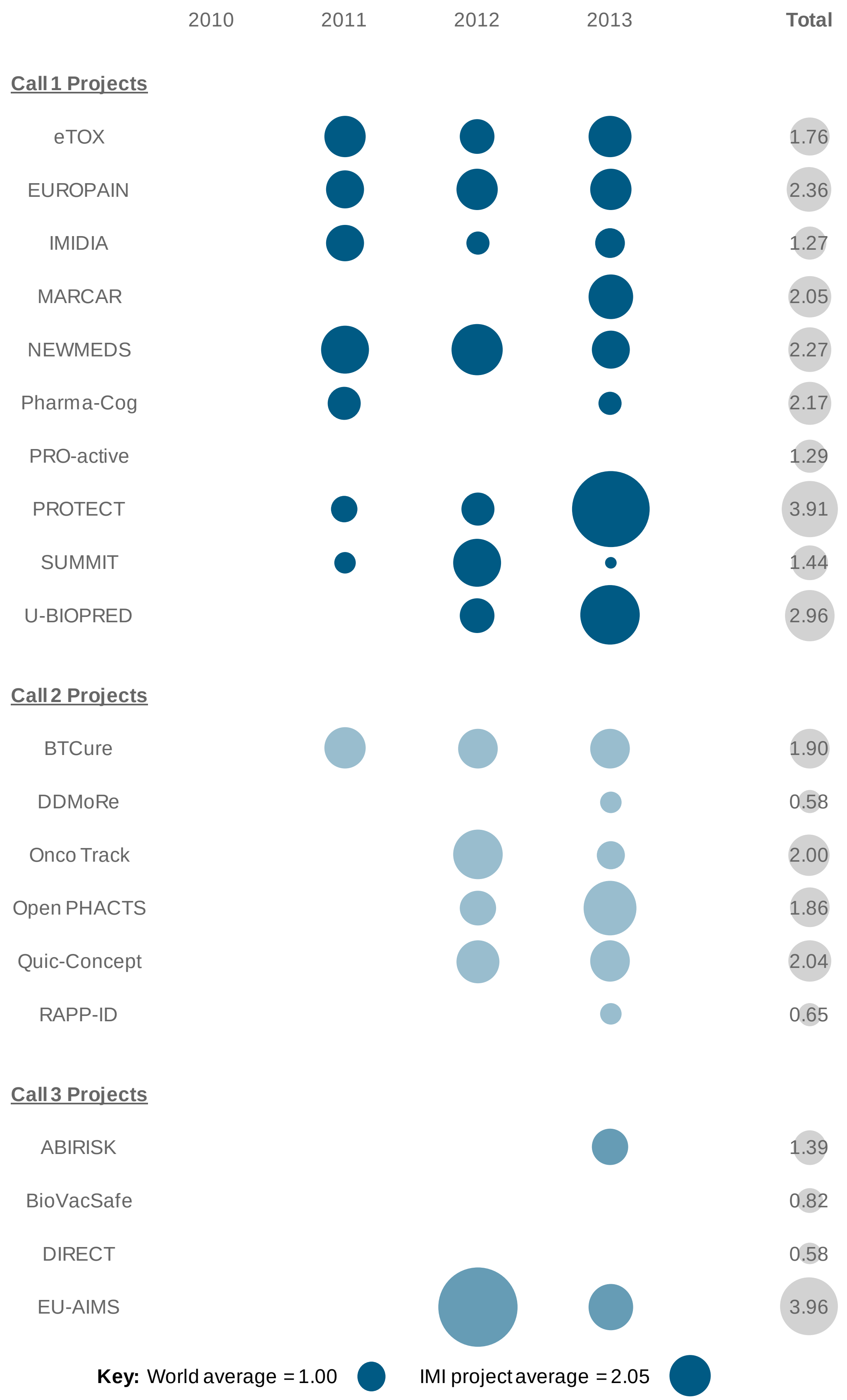


FIGURE 5.5.2 TRENDS IN CITATION IMPACT FOR IMI-SUPPORTED PROJECTS



# 6 BIBLIOMETRIC INDICATORS FOR IMI RESEARCHERS: PRODUCTIVITY, RESEARCH PERFORMANCE AND COLLABORATION

This Section of the report presents analyses of the publication output and citation impact of IMI researcher publications as well as collaborative activities between IMI researchers.

## 6.1 PUBLICATIONS BY IMI-SUPPORTED RESEARCHERS

Publications by IMI-supported researchers were identified using researcher names, projects and affiliations supplied by IMI. For this report, data and analyses are limited to those 4 861 researchers associated with 36 of the projects funded by the first five IMI funding calls (Calls 1 to 5).[23]

Names of researchers associated with funded projects were provided by IMI personnel along with organisational affiliation and sector. Combining these two data elements with the assumption that researchers from the same project are likely to co-author with one another, candidate publications authored by these individuals were identified using an automated process in *Web of Science* for the period January 2007 to December 2013. These matches were further reviewed and edited by IMI personnel.

It is important to note that this dataset includes all identified output from IMI-supported researchers as described above, and is not restricted to that output specifically resulting from IMI funding. With the assumption that the quality of the researcher does not change depending on the source of their funding, these analyses illustrate the quality of researchers who are supported by IMI funds.

These data will also provide a basis for benchmarking how well research from IMI-supported projects (Sections 4 and 5) compares with research by researchers that IMI funds.

## 6.2 CITATION DATA FOR PUBLICATIONS BY IMI-SUPPORTED RESEARCHERS

A total of 29 064 publications by IMI-supported researchers were identified. The process of identifying publications by IMI-supported researchers with Thomson Reuters citation data is outlined in Figure 6.2.1.

Citation counts for these 29 064 publications have been sourced from the citation databases which underlie Thomson Reuters *Web of Science* and were extracted at end-December 2013.

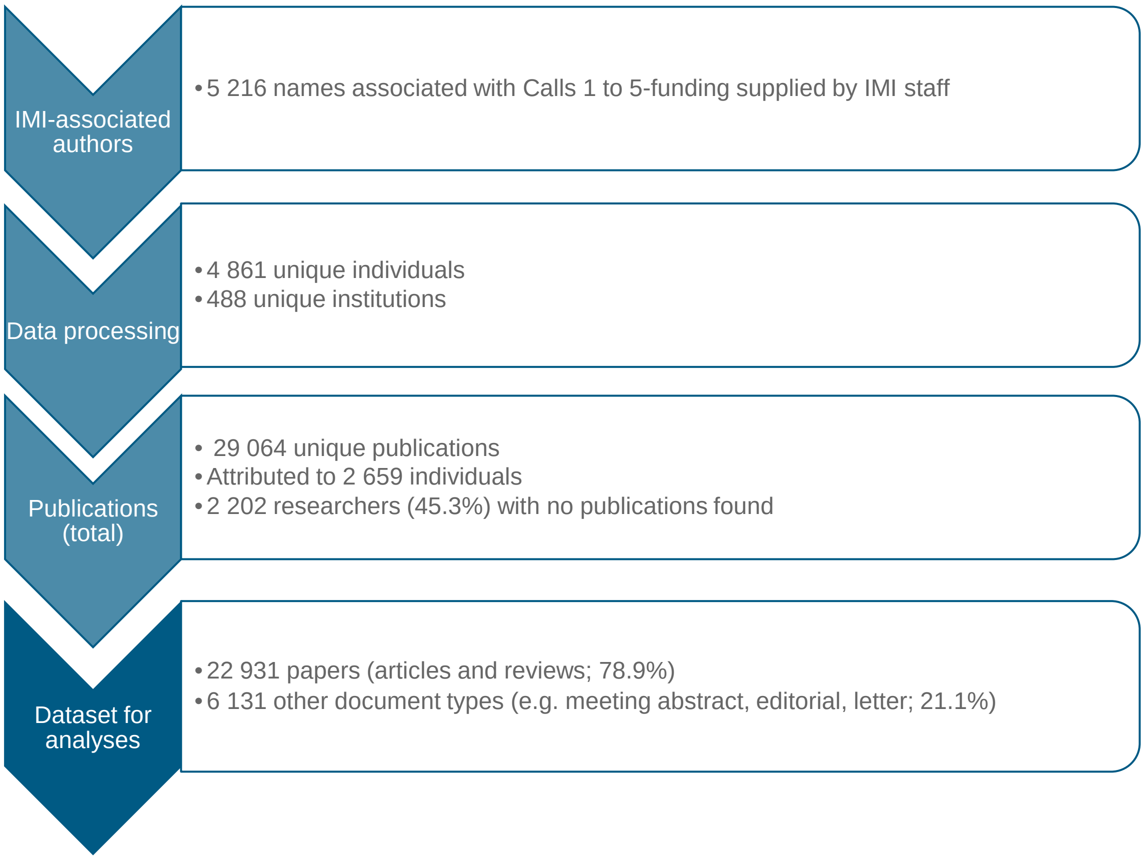
### 6.2.1 STRENGTHS AND LIMITATIONS

These data rely on publication matching from researcher productivity analysis and are restricted to the period from January 2007 to December 2013. Although this includes all document types, some publications may have been missed in the effort to only match researchers to publications for which we are fairly certain they are the author.

Researchers with multiple organisational affiliations and/or multiple sectors were assigned to a single organisational affiliation, and sector based on the judgement of IMI personnel. No similar assignment was made for investigators affiliated with multiple projects – that is, multiple project affiliations were preserved for all investigators who were involved with multiple projects.

[23] No researcher names were provided for projects in Calls 6 and 7 (Translocation, COMBACTE, ADVANCE and GETREAL), all of which started before the end of December 2013.

FIGURE 6.2.1 IDENTIFYING PUBLICATIONS BY IMI-SUPPORTED RESEARCHERS WITH THOMSON REUTERS CITATION DATA




## 6.3 BIBLIOMETRIC INDICATORS FOR IMI-SUPPORTED RESEARCHERS: PRODUCTIVITY

Table 6.3.1 summarises output by IMI-funded researchers by sector. Publication output, as previously, is higher for IMI-supported researchers based in academic institutions and other research environments.

Overall, 54.7% of researchers have at least one publication.

TABLE 6.3.1 PRODUCTIVITY: PUBLICATION OUTPUT, OVERALL AND BY SECTOR

| | Number of researchers | | |
|---|---|---|---|
| **Sector** | **Total** | **With publications** | **% researchers with publications** |
| Academia | 1 844 | 1 134 | 61.5% |
| Corporate | 1 627 | 830 | 51.0% |
| Patient organisation | 66 | 21 | 31.8% |
| Regulatory agency | 62 | 24 | 38.7% |
| Research (other) | 833 | 493 | 59.2% |
| Small Medium Enterprise | 386 | 141 | 36.5% |
| No assignment | 84 | 42 | 50.0% |
| **Total researchers** | **4 861** | **2 659** | **54.7%** |

## 6.4 BIBLIOMETRIC INDICATORS FOR IMI-SUPPORTED RESEARCHERS: RESEARCH PERFORMANCE

The bibliometric indicators presented in Table 6.4.1 have been calculated for each individual IMI-supported researcher and aggregated by sector.

As in both the second and third reports, researchers associated with patient organisations have continued to publish well-regarded research. Six (28.6%) have published at least one 'hot' paper (defined in Section 3.1.3), five (23.8%) have an h-index of at least 10 and 19 (90.5%) have published exclusively in top-quartile journals.

Of the 1 135 publishing academic-based researchers, 230 (20.3%) have published at least one 'hot paper', 263 (23.2%) have an h-index of at least 10 and 979 (86.3%) have published most frequently in top quartile journals. These percentages have all increased since the third report.

Similarly, researchers based in other research environments have also published research which has performed well. Ninety-eight of these researchers (19.8%) have published a minimum of one 'hot paper', 93 researchers (18.8%) have h-index of at least 10 and 421 (85.2%) have published in top quartile journals more frequently than in less well-regarded journals.

By contrast, as in the third report, many IMI-supported researchers working in companies and small- or medium-sized enterprises have also published relatively frequently in top quartile journals but these publications appear to be less well-cited as their 'hot papers' indicator and h-indices are generally lower.

TABLE 6.4.1 RESEARCH PERFORMANCE: BIBLIOMETRIC INDICATORS, OVERALL AND BY SECTOR

| **Sector** | **Researchers** | | **With 'hot papers'** | | **h-index ≥ 10** | | **Publishes most often in top quartile journals** | |
|---|---|---|---|---|---|---|---|---|
| | **Total** | **Publishing** | **N** | **%** | **N** | **%** | **N** | **%** |
| Academia | 1 844 | 1 135 | 230 | 20.3% | 263 | 23.2% | 979 | 86.3% |
| Corporate | 1 627 | 831 | 61 | 7.3% | 29 | 3.5% | 606 | 72.9% |
| Patient organisation | 66 | 21 | 6 | 28.6% | 5 | 23.8% | 19 | 90.5% |
| Regulatory agency | 62 | 24 | 2 | 8.3% | 0 | 0.0% | 12 | 50.0% |
| Research (other) | 833 | 494 | 98 | 19.8% | 93 | 18.8% | 421 | 85.2% |
| Small Medium Enterprise | 386 | 141 | 21 | 14.9% | 12 | 8.5% | 112 | 79.4% |
| No assignment | 84 | 42 | 5 | 11.9% | 6 | 14.3% | 41 | 97.6% |
| **Total researchers** | **4 861** | **2 659** | **414** | **15.6%** | **398** | **15.0%** | **2 164** | **81.4%** |

Table 6.4.2 presents statistics based on the diffusion index and citation velocity for researchers in each sector. For both metrics, the maximum and mean were identified for each researcher and averages then calculated by sector.

TABLE 6.4.2 RESEARCH PERFORMANCE: BIBLIOMETRIC INDICATORS, CONTINUOUS OUTCOMES

| Sector | Researchers | | Mean Diffusion Index | | Maximum Diffusion Index | | Mean Citation Velocity | | Maximum Citation Velocity | |
|---|---|---|---|---|---|---|---|---|---|---|
| | Total | Publishing | Mean | SD | Mean | SD | Mean | SD | Mean | SD |
| Academia | 1 844 | 1 135 | 0.547 | 0.121 | 0.703 | 0.124 | 0.011 | 0.016 | 0.056 | 0.104 |
| Corporate | 1 627 | 831 | 0.536 | 0.157 | 0.630 | 0.161 | 0.009 | 0.015 | 0.022 | 0.042 |
| Patient organisation | 66 | 21 | 0.611 | 0.074 | 0.697 | 0.076 | 0.052 | 0.117 | 0.112 | 0.225 |
| Regulatory agency | 62 | 24 | 0.499 | 0.182 | 0.610 | 0.174 | 0.007 | 0.009 | 0.016 | 0.024 |
| Research (other) | 833 | 494 | 0.553 | 0.132 | 0.691 | 0.134 | 0.014 | 0.029 | 0.066 | 0.160 |
| Small Medium Enterprise | 386 | 141 | 0.566 | 0.137 | 0.658 | 0.135 | 0.016 | 0.028 | 0.056 | 0.127 |
| No assignment | 84 | 42 | 0.553 | 0.152 | 0.668 | 0.143 | 0.012 | 0.013 | 0.032 | 0.049 |
| **Total researchers** | **4 861** | **2 659** | **0.546** | **0.137** | **0.674** | **0.143** | **0.011** | **0.022** | **0.046** | **0.106** |

## 6.5 COLLABORATION BETWEEN IMI-SUPPORTED RESEARCHERS AT INDIVIDUAL LEVEL

The projects funded by IMI are collaborative in nature. However, collaboration between researchers can manifest in many different ways, only one of which is in co-authorship in published materials. Using this definition of collaboration, social network analysis was used to assess the extent to which collaboration occurs, the nature of collaborations between researchers, and to identify opportunities to foster collaboration.

Overall, 2 659 researchers (54.7% of 4 861 IMI researchers in total) published any documents that were indexed in *Web of Science*. Over two-thirds of these researchers (N=2 012, 75.7% of 2 659) collaborated (co-authored) with at least one other IMI researcher during the period January 2007-December 2013.

The frequency of collaborative activities are shown over the period of January 2007 to December 2013 by year in Table 6.5.1 and in Figures 6.5.1 and 6.5.2.

The number of collaborative publications published in 2013 has fallen compared with 2012, the first such fall in the dataset. This is consistent with the interim picture presented in the third report. In particular, the number of publications published in 2013 (5 368) also fell compared with 2012 (5 897). Despite this, the number of collaborations overall has remained relatively constant at around 3 000 each year since 2011. The percentage of collaborations that are cross-sector has also remained relatively constant at around 40-45% since 2009.

TABLE 6.5.1 COLLABORATIVE ACTIVITY BY YEAR – PUBLICATIONS FROM IMI-SUPPORTED RESEARCHERS

| | 2007 | 2008 | 2009 | 2010 | 2011 | 2012 | 2013 |
|---|---|---|---|---|---|---|---|
| Publications | 69 | 814 | 939 | 1 097 | 1 404 | 1 679 | 1 468 |
| Within-Sector Collaborations | 73 | 597 | 855 | 1 224 | 1 597 | 1 673 | 1 728 |
| Cross-Sector Collaborations | 28 | 282 | 649 | 963 | 1 253 | 1 186 | 1 240 |
| % Cross-Sector | 26.9% | 31.8% | 42.5% | 43.5% | 43.0% | 40.8% | 41.0% |

FIGURE 6.5.1 NUMBER OF PUBLICATIONS CO-AUTHORED BY TWO OR MORE IMI RESEARCHERS BY YEAR

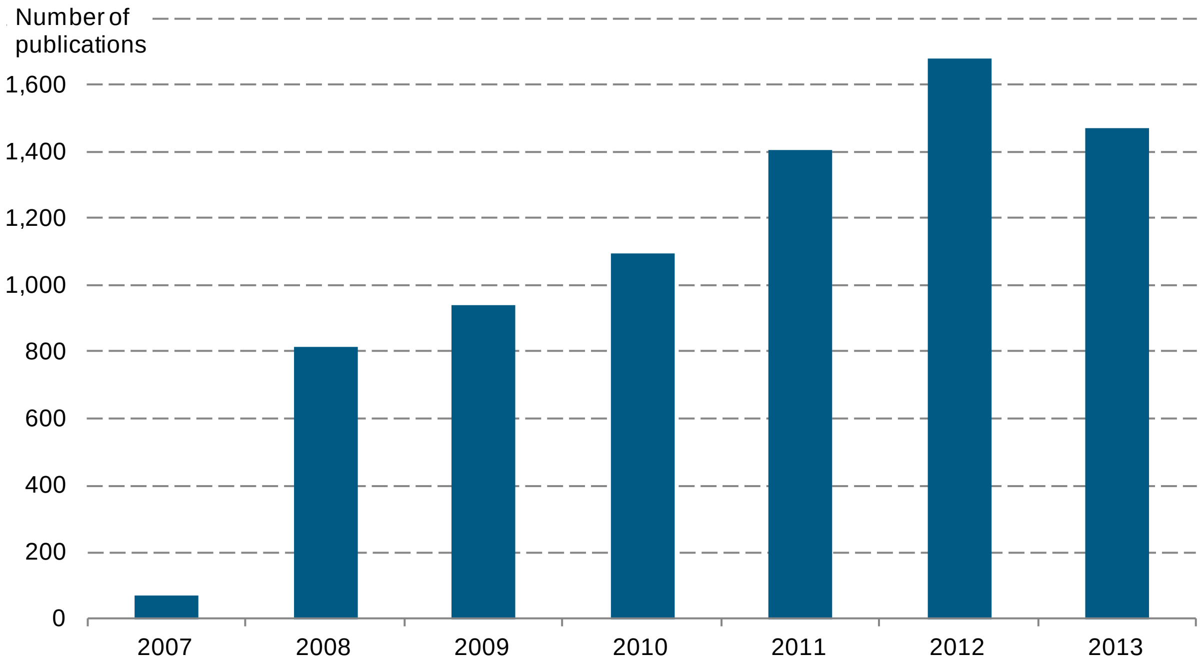



FIGURE 6.5.2 CO-AUTHORSHIP PAIRS WITHIN AND ACROSS SECTORS APPEARING ON ONE OR MORE PUBLICATIONS DURING THE GIVEN PUBLICATION YEAR

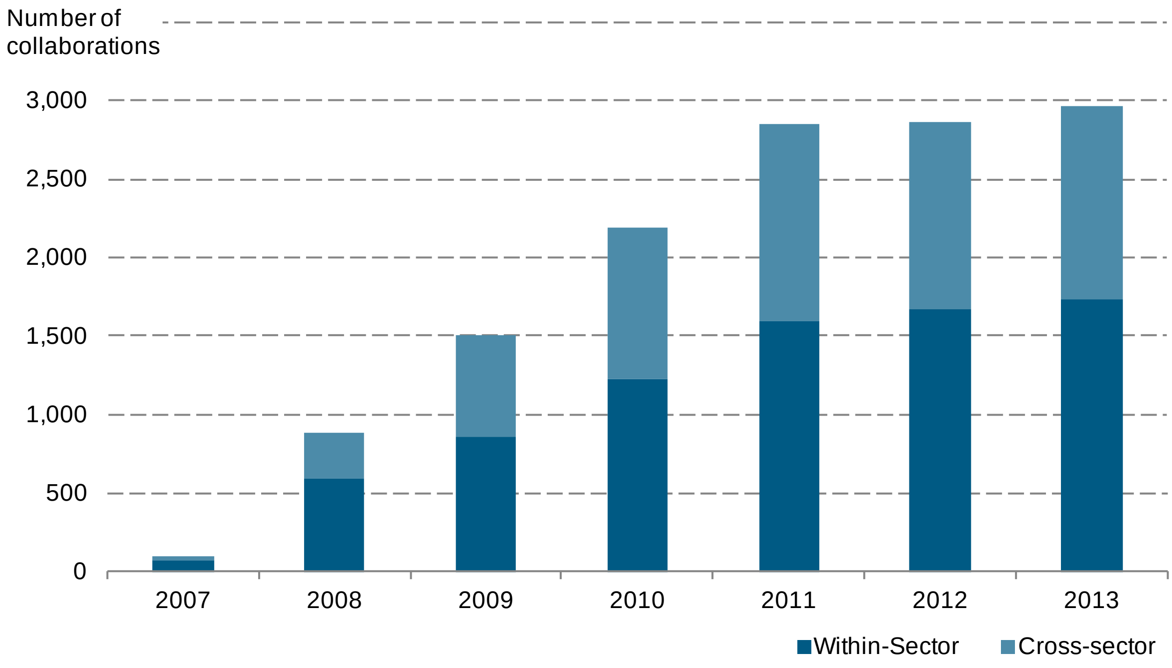
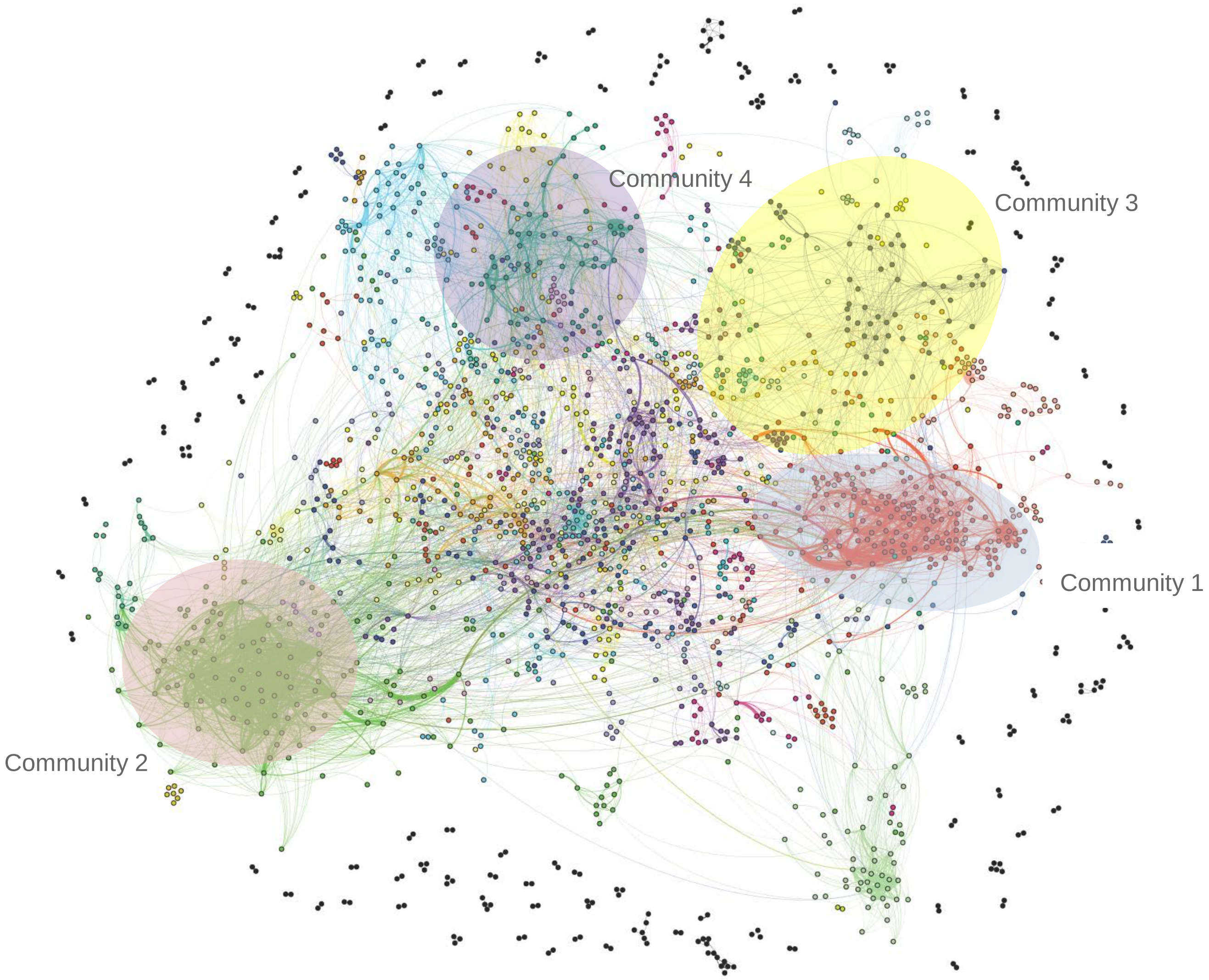



Figure 6.5.3 presents a network diagram showing all co-authorship pairs between IMI-funded researchers. Each individual is represented as a single node coloured with respect to the community of researchers they are based in. Lines between researchers are instances where co-authorship has occurred in a published work. The largest group of inter-connected researchers is composed of 31 communities of which the 4 largest are shown in shaded ovals. The diagram was produced using

Gephi, applying the Force Atlas 2 layout.[24] Communities were identified using a resolution of 1.[25] Isolated communities not connected to the main group of researchers are shown in black.

FIGURE 6.5.3 MAP OF 2 012 IMI PROJECT RESEARCHERS WHO HAVE CO-AUTHORED WITH AT LEAST ONE OTHER RESEARCHER WITHIN THE NETWORK BASED ON CO-AUTHORSHIP ACTIVITIES FROM JANUARY 2007 TO DECEMBER 2013




[24] Jacomy, M. (2009). Force-Atlas Graph Layout Algorithm. *URL: http://gephi.org/2011/forceatlas2-the-new-version-of-our-home-brew-layout/*

[25] Vincent D Blondel, Jean-Loup Guillaume, Renaud Lambiotte, Etienne Lefebvre, Fast unfolding of communities in large networks, in Journal of Statistical Mechanics: Theory and Experiment 2008 (10), P1000

## 6.6 COLLABORATION BETWEEN IMI-SUPPORTED RESEARCHERS AT SECTOR LEVEL

The largest component, shown at the centre of Figure 6.5.3 and defined as groups of researchers where all individuals are connected with one another directly or indirectly via other IMI researchers, consists of 1 758 researchers representing all six sectors. Table 6.6.1 presents data about the distribution of researchers in this main component and within the other ninety-seven smaller, isolated components shown in black around the edge of Figure 6.5.3. The largest of these smaller components consists of 12 researchers. Figures 6.6.1 and 6.6.2 show the same layout of researchers coloured by sector and disease area respectively.

Within the largest component, 31 communities were identified within which there are more frequent and closely inter-related co-authorship activities. The largest four of these 31 communities are shown enclosed by coloured ovals.

The largest community, shown in the blue oval in Figure 6.5.3, is composed of a fairly tight group of individuals suggesting that as well as there being highly collaborative activity between these researchers, collaborative activity between co-authors of any given researcher is also strong. This group is largely composed of researchers from academic and other research institutions, and also by researchers working on inflammatory disorders as part of the BTCure project.

Community 2 (shown red in Figure 6.5.3) is also a fairly tight community. Many of the co-authorship activities in this community are associated with research into metabolic disorders with many researchers involved in the DIRECT, SUMMIT and EMIF projects. While the majority of researchers are from academia, around one fifth (19.2%) are from a corporate or SME background.

Community 3 (shown in yellow in Figure 6.5.3) is far more diffuse than either of the largest two communities. Nearly half (45.3%) of the researchers in this community work on data management as part of the OPENPHACTS project. Community 4 (highlighted in purple in the same figure) is even more diffuse, and extends through the middle of the figure, with just the main portion highlighted. Researchers in this community are primarily involved in research into drug safety in the SAFE-T, MARCAR and MIPDILI projects. A much higher proportion of its researchers also come from the corporate (29.9%) or SME (6.8%) sectors.

While the majority of publishing researchers are connected to one another and are in the main connected component, 12.6% of collaborating researchers (N=254 of 2 012) collaborate within isolated communities composed of between 2 and 12 researchers. This is a similar percentage to the third report (13.5%). Ninety-seven isolated groups exist (shown on the periphery of Figure 6.5.3), of which 75 (77.3%) are composed of researchers from only one sector.

The main component includes researchers from 353 distinct organisations, 35.1% (N=124) of which span across communities. Within this set there are 145 academic organisations, 32 corporate organisations, 7 patient organisations, 4 regulatory agencies, 97 research/other entities, and 39 small- or medium-sized enterprises. The four entities which span the most communities in this main component are Astra Zeneca (corporate), Karolinska Institutet (academia), Novartis (corporate) and Roche (corporate). Overall, these four organisational affiliations include 11.3% (N=198 of 1 758) of researchers in the main component. Astra Zeneca covers 17 communities, while the other three institutions cover 12 communities each.

Co-authorship is more common among researchers in the same sector than among researchers in different sectors (Figure 6.6.1). This is expected given the principle of homophily which suggests that individuals are more likely to interact with individuals who are like them.[26] However, there are substantial co-authorship activities among researchers from different sectors. Of a total of 7 850 distinct co-authorship relationships, 3 598 are cross-sector and involve 1 193 total researchers from all 6 sectors. This accounts for 36.6% of all co-authorship activities during the analysis period.

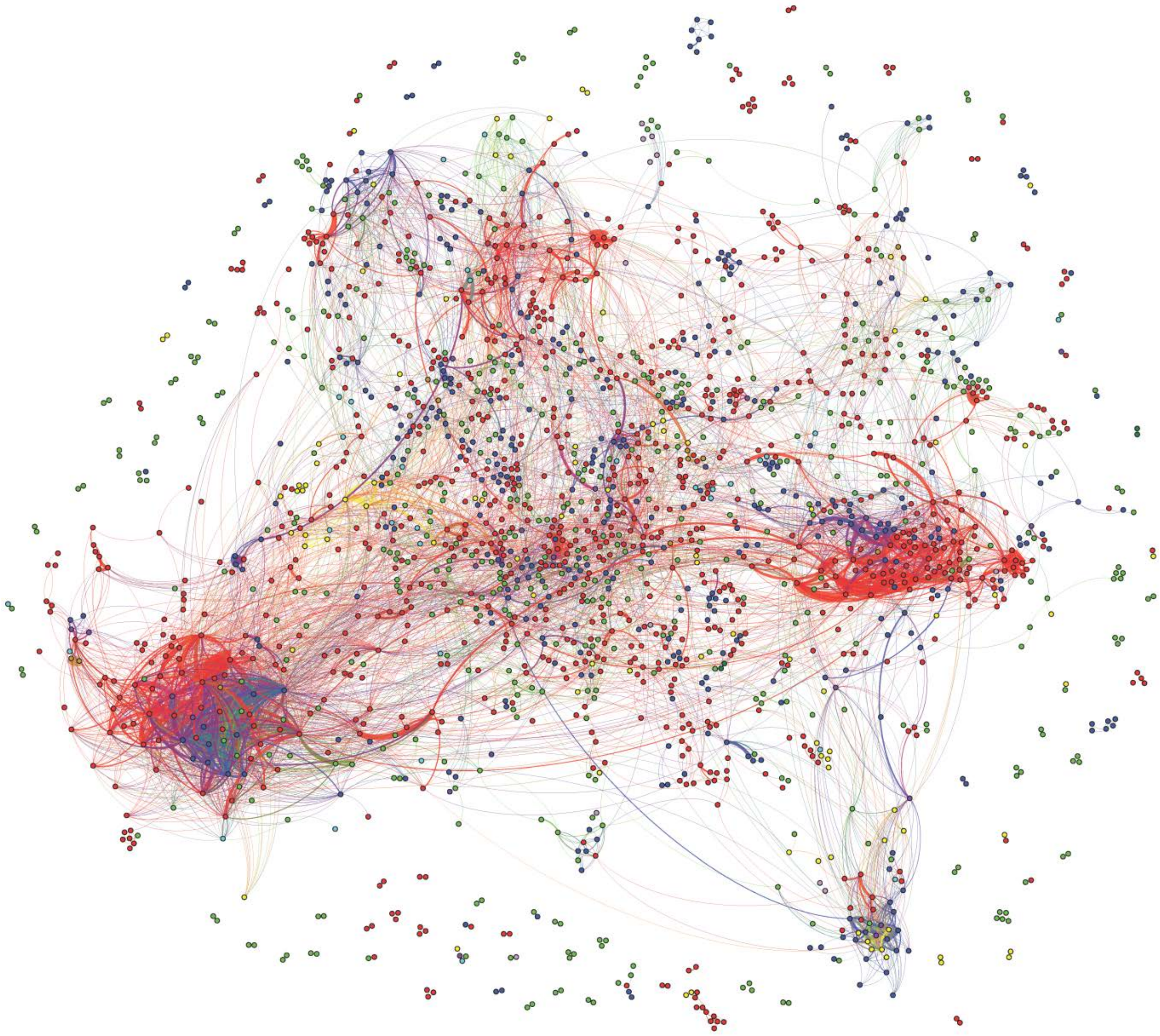
The same is true of co-authorship activities by project. The majority of collaborative relationships are among researchers associated with the same project with 3 178 of 7 850 of co-authorship relationships (40.5 %) being cross-project, an increase from 36.9% in the third report.

[26] McPherson et al. "Birds of a Feather: Homophily in Social Networks". Annu Rev Sociol, 2001, 27: 415-44.

TABLE 6.6.1 DISTRIBUTION OF SECTORS BY NUMBER OF RESEARCHERS WITHIN SELECT COMMUNITIES BASED ON CO-AUTHORSHIP ACTIVITIES FROM JANUARY 2007 - DECEMBER 2013.

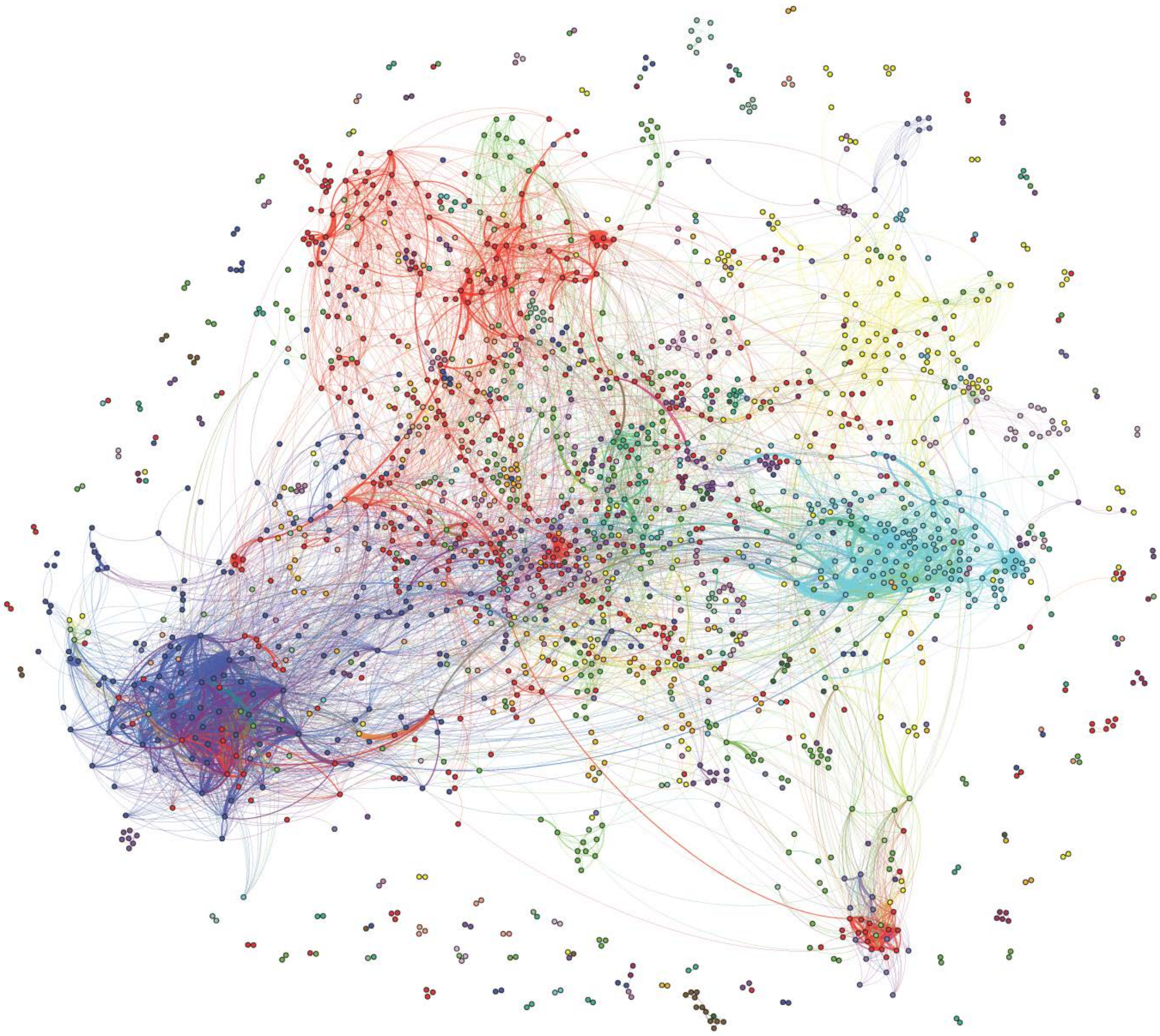
| | Isolated Communities | | Connected Communities | | Connected Community | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | 1 | | 2 | | 3 | | 4 | |
| Sector | N | % | N | % | N | % | N | % | N | % | N | % |
| Academia | 80 | 31.5% | 845 | 48.1% | 110 | 60.4% | 86 | 66.2% | 62 | 48.4% | 44 | 37.6% |
| Corporate | 114 | 44.9% | 389 | 22.1% | 18 | 9.9% | 22 | 16.9% | 29 | 22.7% | 35 | 29.9% |
| Patient organisation | -- | -- | 14 | 0.8% | 1 | 0.5% | 1 | 0.8% | 1 | 0.8% | 1 | 0.9% |
| Regulatory agency | 1 | 0.4% | 11 | 0.6% | -- | -- | -- | -- | -- | -- | -- | -- |
| Research (other) | 36 | 14.2% | 375 | 21.3% | 46 | 25.3% | 12 | 9.2% | 29 | 22.7% | 27 | 23.1% |
| Small Medium Enterprise | 18 | 7.1% | 77 | 4.4% | 6 | 3.3% | 3 | 2.3% | 4 | 3.1% | 8 | 6.8% |
| Multiple sectors | 2 | 0.8% | 19 | 1.1% | 1 | 0.5% | 2 | 1.5% | -- | -- | 1 | 0.9% |
| No assignment | 3 | 1.2% | 28 | 1.6% | -- | -- | 4 | 3.1% | 3 | 2.3% | 1 | 0.9% |
| **Total** | **254** | | **1 758** | | **182** | | **130** | | **128** | | **117** | |

FIGURE 6.6.1 RESEARCHERS WITH ANY COLLABORATIVE ACTIVITY, COLOURED BY SECTOR.



| | Sector | Proportion |
|---|---|---|
| (red) | Academia | 45.07% |
| (green) | Corporate | 25.00% |
| (blue) | Research (other) | 20.43% |
| (yellow) | Small or Medium Enterprise | 4.72% |
| (cyan) | Not assigned | 1.54% |
| (purple) | Multiple | 1.04% |
| (brown) | Patient organisation | 0.70% |
| (pink) | Regulatory | 0.60% |

FIGURE 6.6.2 RESEARCHERS WITH ANY COLLABORATIVE ACTIVITY, COLOURED BY DISEASE AREA



| | Disease area | | Disease area |
|---|---|---|---|
| | Brain disorders | | Multiple |
| | Drug safety | | Stem cells |
| | Metabolic disorders | | Education and training |
| | Data management | | Vaccines |
| | Inflammatory disorders | | Drug delivery |
| | Cancer | | Drug kinetics |
| | Lung diseases | | Sustainable chemistry |
| | Biologicals | | Drug discovery |
| | Infectious diseases | | |

In all, 570 organisations with collaborative co-authorship activity were identified, of which 421 (73.9%) span two or more communities and 327 (57.4%) span three or more communities. Table 6.6.2 lists the top 25 organisations by the number of communities they span.

TABLE 6.6.2 ORGANISATIONS AND THE NUMBER OF ASSOCIATED COMMUNITIES AND RESEARCHERS WITHIN THE MAIN INTER-CONNECTED COMPONENT

| Organisation | Sector | Number of communities | Number of researchers |
|---|---|---|---|
| Sanofi-Aventis | Corporate | 144 | 160 |
| GSK | Corporate | 140 | 157 |
| Pfizer | Corporate | 114 | 153 |
| Astra Zeneca | Corporate | 109 | 152 |
| Johnson | Corporate | 100 | 138 |
| Roche | Corporate | 91 | 129 |
| Bayer | Corporate | 81 | 85 |
| Novartis | Corporate | 67 | 88 |
| Eli Lilly | Corporate | 64 | 83 |
| Merck | Corporate | 53 | 63 |
| University of Oxford | Academia | 48 | 79 |
| INSERM | Research (other) | 48 | 65 |
| Karolinska Institutet | Academia | 44 | 94 |
| Boehringer Ingelheim | Corporate | 43 | 53 |
| Novo Nordisk | Corporate | 37 | 38 |
| Universiteit Utrecht | Academia | 36 | 53 |
| Imperial College London | Academia | 36 | 61 |
| Lundbeck | Corporate | 33 | 40 |
| Orion Pharma | Corporate | 32 | 35 |
| University of Manchester | Academia | 32 | 58 |
| Amgen | Corporate | 31 | 31 |
| Max-Planck-Gesellschaft | Research (other) | 31 | 46 |
| Vrije University Amsterdam | Academia | 30 | 36 |
| Servier | Corporate | 29 | 35 |
| UCB Pharma | Corporate | 27 | 29 |

## 6.7 COLLABORATION ANALYSIS BEFORE AND AFTER IMI FUNDING BY DISEASE AREA

This section presents an analysis of how levels of collaboration have changed as researchers have started to receive IMI funding. Each of the publications described in Figure 6.2.1 has been identified as 'pre-IMI' (before funding award) or 'post-IMI' (after funding award) based on when the researchers involved were awarded funding by IMI:

- 'Pre-IMI' publications have been defined as all publications from the associated researchers, published between January 2007 and the date of IMI funding award.
- 'Post-IMI' publications have been defined as all publications from the associated researchers, published between the date of IMI funding award and December 2013.
- Where a particular IMI-supported researcher is associated with more than one IMI project, or two researchers who co-authored a publication are associated with more than one IMI project, 'pre-IMI' and 'post-IMI' have been defined by the start date of the earliest project.

Table 6.7.1 presents the number of investigators and co-authorships by disease area, and also the average degree and average weighted degree of each author both pre and post the earliest IMI funding awarded. Figure 6.7.1 shows the average degree for the same two areas. The degree of an author is the number of distinct co-authors that author has; so a researcher who has written publications with, at most, three other researchers has a degree of three.

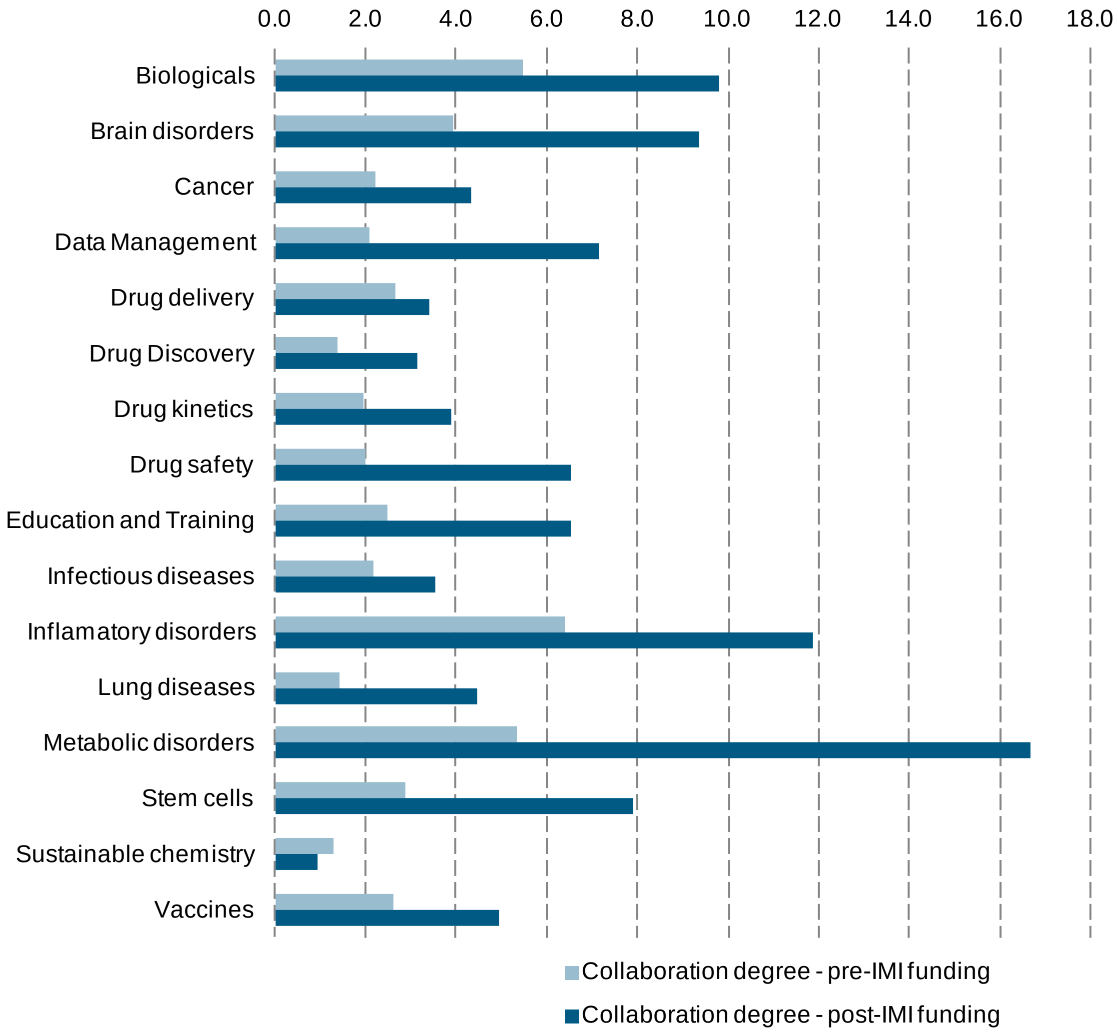
However, this statistic doesn't take into account how many publications each pair of co-authors has published together. For the average weighted degree, each connection is weighted by the number of publications that pair of researchers has co-authored. For example, for our researcher with three co-authors: if one research pair produced two papers, another pair produced four papers and the final pair published six papers, then this researcher would then have a weighted degree of four papers.

The number of co-authored papers listed in Table 6.7.1 counts each paper according to the number of co-authorship pairs associated with it. A paper with four authors would have six co-authorship pairs between each pair of authors.

For most disease areas, the average number of co-authors for each researcher has increased since the researcher first received IMI funding; sustainable chemistry (the CHEM21 project) is the only exception. Four "disease" areas (data management, drug safety, lung diseases and metabolic disorders) have all experience a more than three-fold increase in mean degree since IMI funding began.

TABLE 6.7.1 LEVEL OF COLLABORATION PRE- AND POST-IMI FUNDING BY DISEASE AREA

| **Disease area** | | **Co-authorship pairs** | **Co-authored papers** | **Number of authors** | **Mean degree** | **Mean weighted degree** |
|---|---|---|---|---|---|---|
| Biologicals | pre | 368 | 1 197 | 134 | 5.49 | 17.87 |
| | post | 569 | 1 543 | 116 | 9.81 | 26.60 |
| Brain disorders | pre | 841 | 2 116 | 425 | 3.96 | 9.96 |
| | post | 2 586 | 6 563 | 553 | 9.35 | 23.74 |
| Cancer | pre | 174 | 398 | 157 | 2.22 | 5.07 |
| | post | 358 | 707 | 165 | 4.34 | 8.57 |
| Data management | pre | 194 | 365 | 186 | 2.09 | 3.92 |
| | post | 803 | 1 354 | 225 | 7.14 | 12.04 |
| Drug delivery | pre | 84 | 206 | 63 | 2.67 | 6.54 |
| | post | 90 | 200 | 53 | 3.40 | 7.55 |
| Drug discovery | pre | 27 | 82 | 39 | 1.38 | 4.21 |
| | post | 47 | 56 | 30 | 3.13 | 3.73 |
| Drug kinetics | pre | 50 | 106 | 51 | 1.96 | 4.16 |
| | post | 88 | 128 | 45 | 3.91 | 5.69 |
| Drug safety | pre | 216 | 392 | 214 | 2.02 | 3.66 |
| | post | 1 047 | 1 999 | 320 | 6.54 | 12.49 |
| Education and training | pre | 104 | 173 | 84 | 2.48 | 4.12 |
| | post | 337 | 910 | 103 | 6.54 | 17.67 |
| Infectious diseases | pre | 99 | 232 | 91 | 2.18 | 5.10 |
| | post | 178 | 362 | 100 | 3.56 | 7.24 |
| Inflammatory disorders | pre | 562 | 1 908 | 175 | 6.42 | 21.81 |
| | post | 1 045 | 3 027 | 176 | 11.88 | 34.40 |
| Lung diseases | pre | 60 | 70 | 85 | 1.41 | 1.65 |
| | post | 303 | 452 | 136 | 4.46 | 6.65 |
| Metabolic disorders | pre | 555 | 1 232 | 207 | 5.36 | 11.90 |
| | post | 1 989 | 6 149 | 239 | 16.64 | 51.46 |
| Stem cells | pre | 133 | 215 | 92 | 2.89 | 4.67 |
| | post | 289 | 472 | 73 | 7.92 | 12.93 |
| Sustainable chemistry | pre | 29 | 86 | 45 | 1.29 | 3.82 |
| | post | 16 | 32 | 34 | 0.94 | 1.88 |
| Vaccines | pre | 76 | 154 | 58 | 2.62 | 5.31 |
| | post | 129 | 208 | 52 | 4.96 | 8.00 |
| **Overall** | **pre** | **2 567** | **6 882** | **2 016** | **2.55** | **6.83** |
| | **post** | **6 835** | **17 247** | **2 319** | **5.89** | **14.87** |




## 6.8 COLLABORATION ANALYSIS BEFORE AND AFTER IMI FUNDING BY SECTOR

Tables 6.8.1 and 6.8.2 shows the number of collaboration pairs between the different sectors, both ‘pre-IMI’ and ‘post-IMI’ funding, as defined in Section 6.7. As the number of collaboration pairs does depend on the number of researchers in each sector, this number has been included for comparison.

Figure 6.8.1 presents the mean degree for each pair of sectors in a bubble chart. The area of each bubble is proportional to the mean degree in each case. For cross-sector analyses, the number of researchers involved is the sum of the researchers for the two sectors. As a result, unless the two sectors are very collaborative, it is likely the mean degree for such pairs will be much smaller than the within-sector analyses.

While the number of researchers associated with each sector has seen a modest increase, the number of co-authorships, both within sector and cross-sector, has increased far more substantially. As a result, the mean collaborative degree has increased in all cases. In particular, the mean degree for academic-academic collaboration has more than doubled (2.91 to 6.12); for corporate-corporate collaboration has nearly trebled (0.49 to 1.41); and for academic-corporate collaboration has nearly quadrupled (0.25 to 0.96).

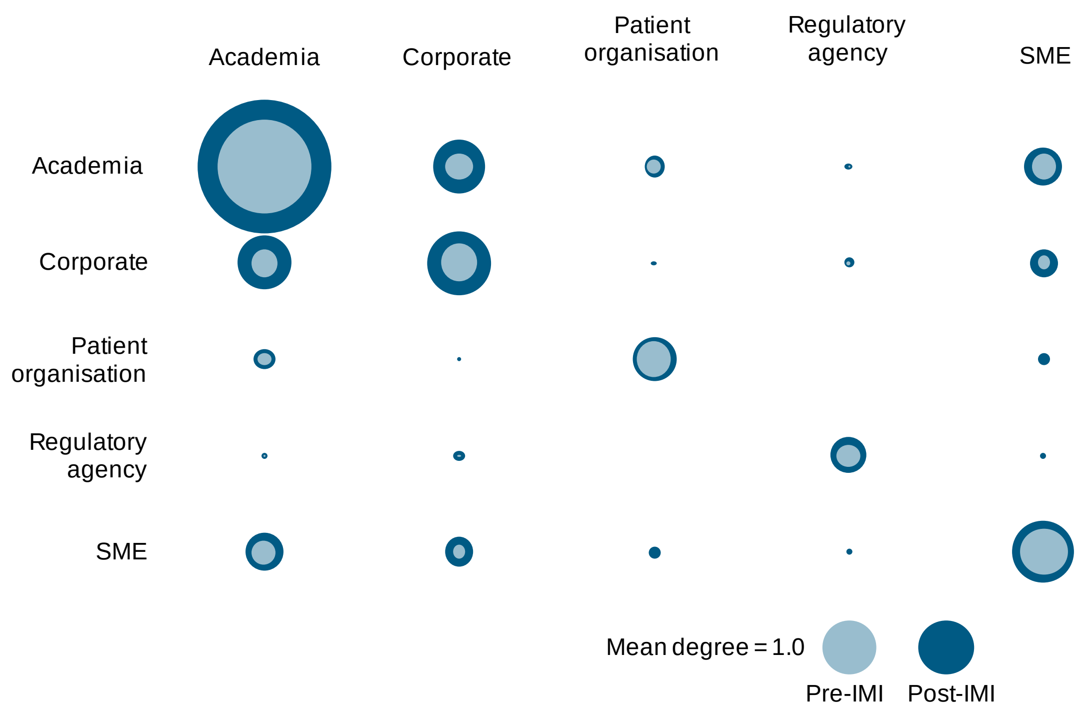
TABLE 6.8.1 NUMBER OF CO-AUTHORSHIPS BETWEEN SECTORS: PRE-IMI FUNDING

| | Academia and research institutes | Corporate | Patient org | Regulatory | SMEs | Total number of researchers |
|---|---|---|---|---|---|---|
| Academia | 1 899 | 236 | 40 | 1 | 141 | 1 304 |
| Corporate | 236 | 140 | | 1 | 20 | 573 |
| Patient organisation | 40 | | 3 | | | 14 |
| Regulatory agency | 1 | 1 | | 1 | | 11 |
| SME | 141 | 20 | | | 39 | 102 |

TABLE 6.8.2 NUMBER OF CO-AUTHORSHIPS BETWEEN SECTORS: POST-IMI FUNDING

| | Academia and research institutes | Corporate | Patient org | Regulatory | SMEs | Total number of researchers |
|---|---|---|---|---|---|---|
| Academia | 4 473 | 1 030 | 104 | 13 | 390 | 1 462 |
| Corporate | 1 030 | 478 | 3 | 13 | 114 | 677 |
| Patient organisation | 104 | 3 | 5 | | 3 | 15 |
| Regulatory agency | 13 | 13 | | 5 | 1 | 22 |
| SME | 390 | 114 | 3 | 1 | 76 | 116 |

FIGURE 6.8.1 MEAN COLLABORATION DEGREE BETWEEN RESEARCHERS FROM THE SAME AND DIFFERENT SECTORS BOTH PRE- AND POST-IMI FUNDING



## 6.9 MAPPING COLLABORATION AMONG IMI-SUPPORTED RESEARCHERS

The analyses in this section are based on the same ‘pre-IMI’ and ‘post-IMI’ subsets of publications as defined in Section 6.7.

Figures 6.9.1 and 6.9.2 provide geographical maps of collaboration among IMI researchers. These maps show:

- The mean degree of collaboration – the average number of other researchers each researcher is co-authoring with – for researchers internally within each country (shaded from white to blue). Countries with no contributing output are shaded in grey.
- The mean degree of collaboration for researchers externally between pairs of countries (shaded from white to orange). For each pair of countries, the degree was calculated based only for researchers in one country who co-authored with researchers in the other country.
- The red dots just indicate the approximate middle of each country.

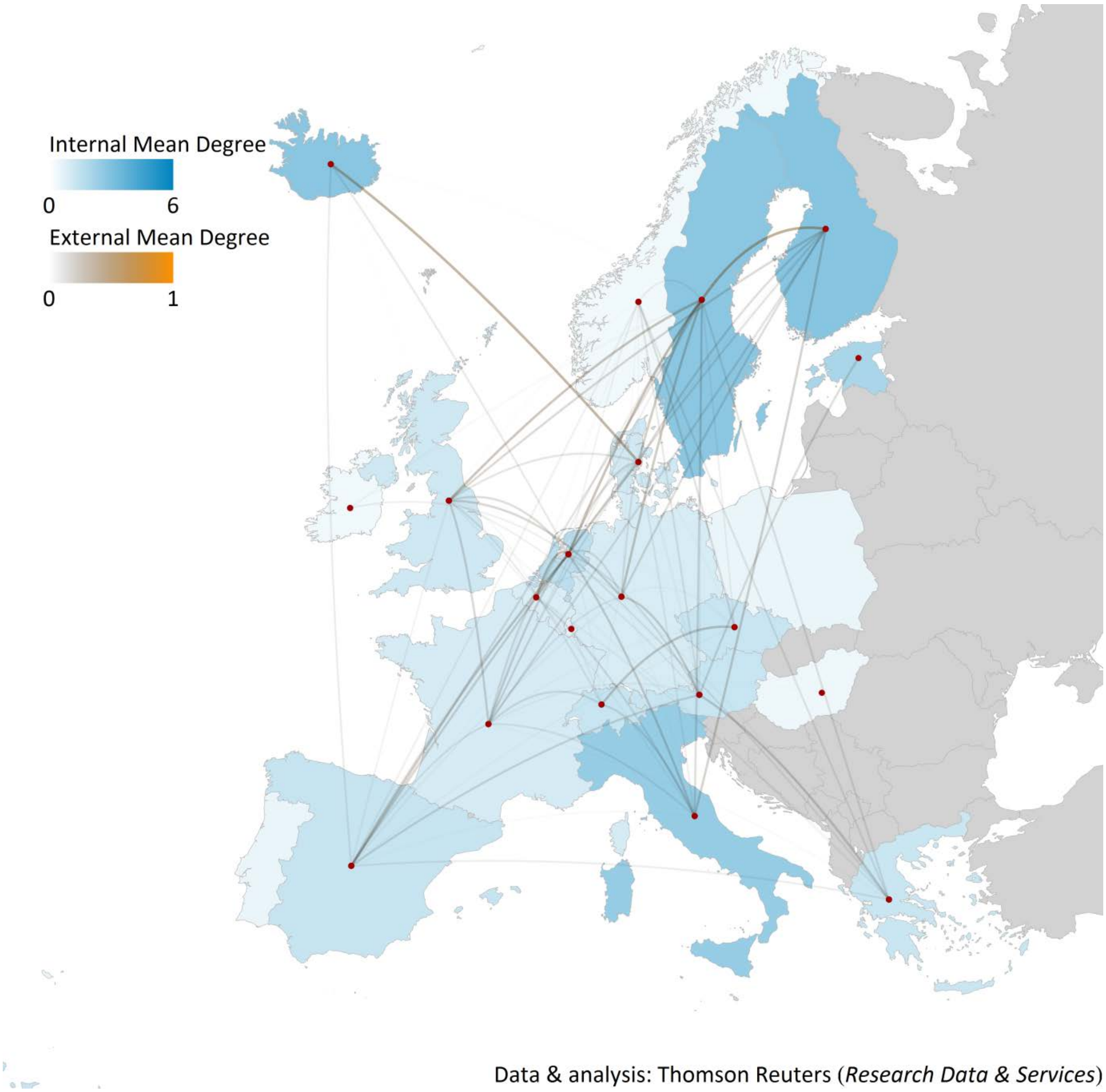
These maps show that IMI research has led to an increase in the level of co-authorship between researchers both within individual countries and between countries. While it is the case that internal collaboration within countries is higher than that between countries – researchers tend to work in local groups, after all – there are some pairs of countries showing significant levels of collaboration. Tables 6.9.1 and 6.9.2 present the ten most prolific countries and pairs of countries by mean degree post-IMI funding, together with their pre-IMI funding mean degrees for comparison.

Iceland has the highest level of internal collaboration post-IMI funding, with 5.43 co-authorship partners in Iceland per researcher, followed by Finland (4.60) and Sweden (4.24). Finland and Sweden also have the highest degree of collaboration of any pair of countries, with 116 international co-authorship pairs between a total of 211 researchers.

While these analyses point towards increased collaboration among IMI-supported researchers, however, it should be noted that researcher mobility during the time period or the recently observed global trend towards increased collaboration are not accounted for.[27]

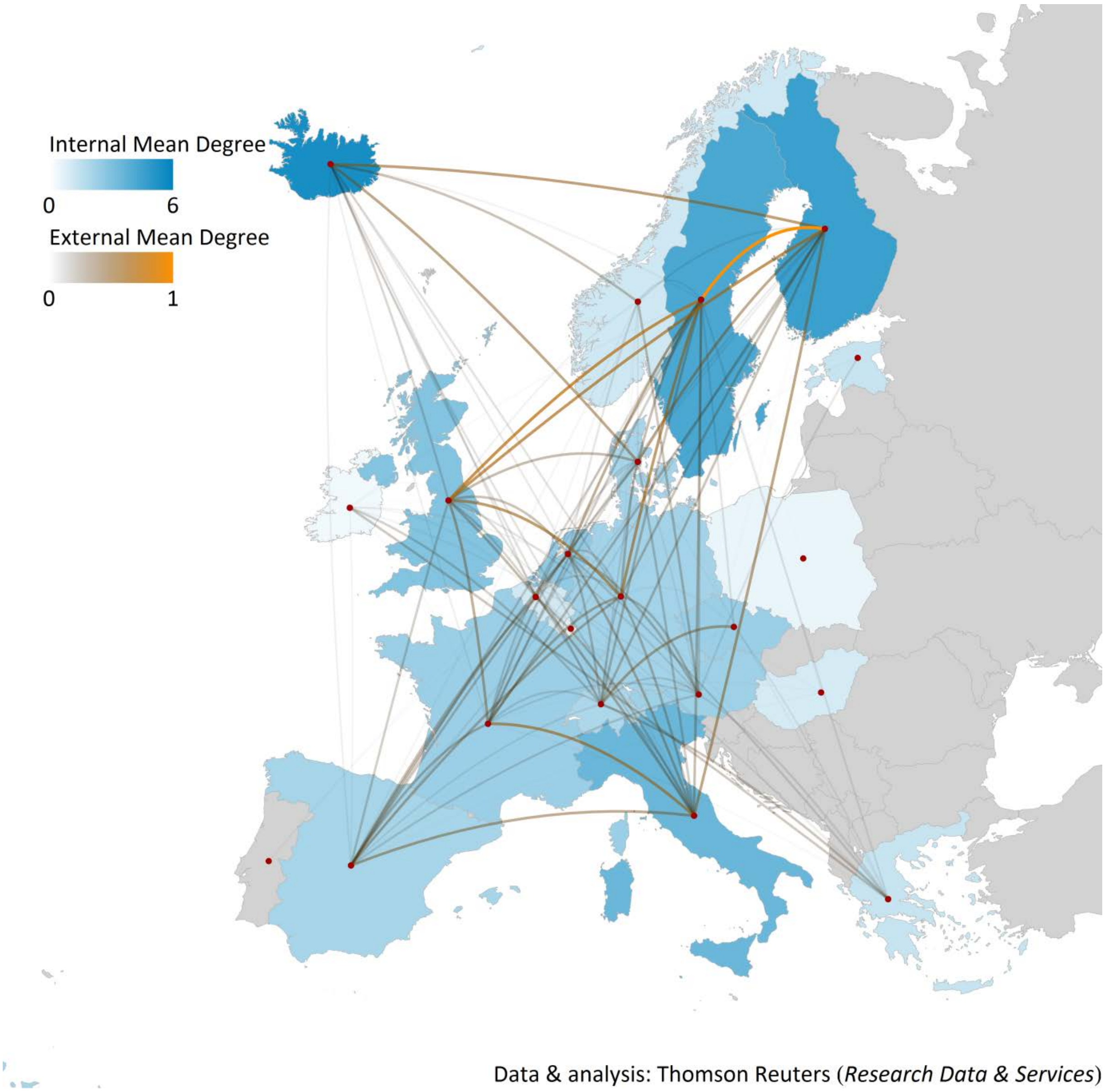
[27] Adams J (2013). Collaborations: the fourth age of research. Nature, 497, 557-560.

FIGURE 6.9.1 COLLABORATIVE RESEARCH LINKS WITHIN AND BETWEEN COUNTRIES FOR IMI-SUPPORTED RESEARCHERS, PRE-IMI FUNDING AWARD




Data & analysis: Thomson Reuters (*Research Data & Services*)

FIGURE 6.9.2 COLLABORATIVE RESEARCH LINKS WITHIN AND BETWEEN COUNTRIES FOR, IMI-SUPPORTED RESEARCHERS, POST-IMI FUNDING AWARD




Data & analysis: Thomson Reuters (*Research Data & Services*)

TABLE 6.9.1 TEN MOST COLLABORATIVE COUNTRIES INTERNALLY, BY POST-IMI MEAN DEGREE

| Country | Mean degree Pre-IMI | Mean degree Post-IMI | Percentage change |
|---|---|---|---|
| Iceland | 2.75 | 5.43 | 97% |
| Finland | 2.81 | 4.60 | 64% |
| Sweden | 2.72 | 4.24 | 56% |
| Italy | 2.44 | 3.49 | 43% |
| UK | 1.15 | 2.89 | 152% |
| Netherlands | 1.57 | 2.56 | 63% |
| Czech Republic | 1.27 | 2.33 | 83% |
| France | 0.98 | 2.33 | 139% |
| Austria | 1.30 | 2.26 | 74% |
| Germany | 0.96 | 2.14 | 123% |

TABLE 6.9.2 TEN MOST COLLABORATIVE COUNTRY PAIRS, BY POST-IMI MEAN DEGREE

| Countries | | Mean degree Pre-IMI | Mean degree Post-IMI | Percentage change |
|---|---|---|---|---|
| Finland | Sweden | 0.39 | 1.10 | 179% |
| Sweden | UK | 0.26 | 0.72 | 173% |
| Finland | UK | 0.16 | 0.64 | 290% |
| France | Italy | 0.08 | 0.60 | 649% |
| Germany | UK | 0.17 | 0.58 | 245% |
| Germany | Sweden | 0.23 | 0.57 | 151% |
| Finland | Iceland | 0.00 | 0.53 | n/a |
| Finland | Italy | 0.19 | 0.47 | 146% |
| Denmark | Iceland | 0.35 | 0.46 | 30% |
| Denmark | Finland | 0.14 | 0.44 | 219% |

# ANNEX 1: SUMMARY OF NEW IMI-SUPPORTED PUBLICATIONS

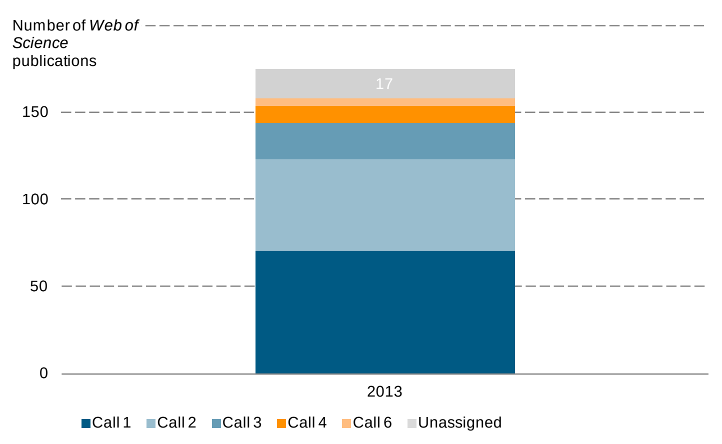
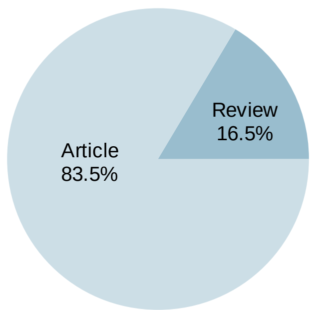
This Annex presents summary analyses of IMI publications identified since the third report to IMI. These summary analyses should be borne in mind when considering IMI project research analyses presented in Sections 4 and 5 of the report.

## A1.1 SUMMARY OF NEW IMI-SUPPORTED PUBLICATIONS – OUTPUT

A total of 176 new IMI-supported publications were identified for this report (as outlined in Section 4.1), 175 (99.4%) of which were abstracted in *Web of Science*. In total, 17 of these publications could not be assigned to IMI projects and have been excluded from the analyses in this report. The remaining 158 publications which have been assigned to projects are analysed in this Annex.

- All publications were published in 2013 (Figure A1.1.1A).
  - o 70 publications were assigned to Call 1;
  - o 53 publications were assigned to Call 2;
  - o 21 publications were assigned to Call 3;
  - o 10 publications were assigned to Call 4;
  - o 4 publications were assigned to Call 6.
- All publications are classed as papers, i.e. articles (132 publications) and reviews (26 publications), higher than the overall dataset (Section 4.2), see (Figure A1.1.1B).

FIGURE A1.1.1 SUMMARY OF NEW IMI-SUPPORTED PUBLICATION OUTPUT

| (A) Number of new *Web of Science* publications by year and call | (B) Categorisation of new *Web of Science* publications by document type |
|---|---|
|   |   |

The additional publications have extended the range of journals in which IMI project research is published. Table A1.1.1 shows all new journals used more than once while Table A1.1.2 presents the top ten new journals with highest Journal Impact Factor.

TABLE A1.1.1 NEW JOURNALS IN WHICH NEW IMI PROJECT PUBLICATIONS HAVE BEEN PUBLISHED MOST FREQUENTLY

| **Journal** | **Number of *Web of Science* publications** | **Number of papers** | **Journal Impact Factor (2012)** | ***Web of Science* journal categories** |
|---|---|---|---|---|
| Drug Safety | 16 | 16 | 3.408 | Pharmacology & Pharmacy<br>Public, Environmental & Occupational Health<br>Toxicology |
| PLoS One | 8 | 8 | 3.73 | Multidisciplinary Sciences |
| Arthritis and Rheumatism | 7 | 7 | 7.477 | Rheumatology |
| Annals of the Rheumatic Diseases | 5 | 5 | 9.111 | Rheumatology |
| Pain | 4 | 4 | 5.644 | Anesthesiology<br>Clinical Neurology<br>Neurosciences |
| Behavioural Brain Research | 3 | 3 | 3.327 | Behavioral Sciences<br>Neurosciences |
| Chemistry-A European Journal | 3 | 3 | 5.831 | Chemistry, Multidisciplinary |
| European Neuropsychopharmacology | 3 | 3 | 4.595 | Clinical Neurology<br>Neurosciences<br>Pharmacology & Pharmacy<br>Psychiatry |
| Genome Biology | 3 | 3 | 10.288 | Biotechnology & Applied Microbiology<br>Genetics & Heredity |

TABLE A1.1.2 TOP TEN NEW JOURNALS IN WHICH NEW IMI PROJECT PUBLICATIONS HAVE BEEN PUBLISHED, RANKED BY JOURNAL IMPACT FACTOR

| **Journal** | **Number of *Web of Science* publications** | **Number of papers** | **Journal Impact Factor (2012)** | ***Web of Science* journal categories** |
|---|---|---|---|---|
| Lancet Neurology | 1 | 1 | 23.917 | Clinical Neurology |
| Nature Methods | 1 | 1 | 23.565 | Biochemical Research Methods |
| Neuron | 1 | 1 | 15.766 | Neurosciences |
| Genome Research | 1 | 1 | 14.397 | Biochemistry & Molecular Biology<br>Biotechnology & Applied Microbiology<br>Genetics & Heredity |
| European Heart Journal | 1 | 1 | 14.097 | Cardiac & Cardiovascular Systems |
| Journal of Clinical Investigation | 1 | 1 | 12.812 | Medicine, Research & Experimental |
| ACS Nano | 1 | 1 | 12.062 | Chemistry, Multidisciplinary<br>Chemistry, Physical<br>Materials Science, Multidisciplinary<br>Nanoscience & Nanotechnology |
| Journal of the American Chemical Society | 1 | 1 | 10.677 | Chemistry, Multidisciplinary |

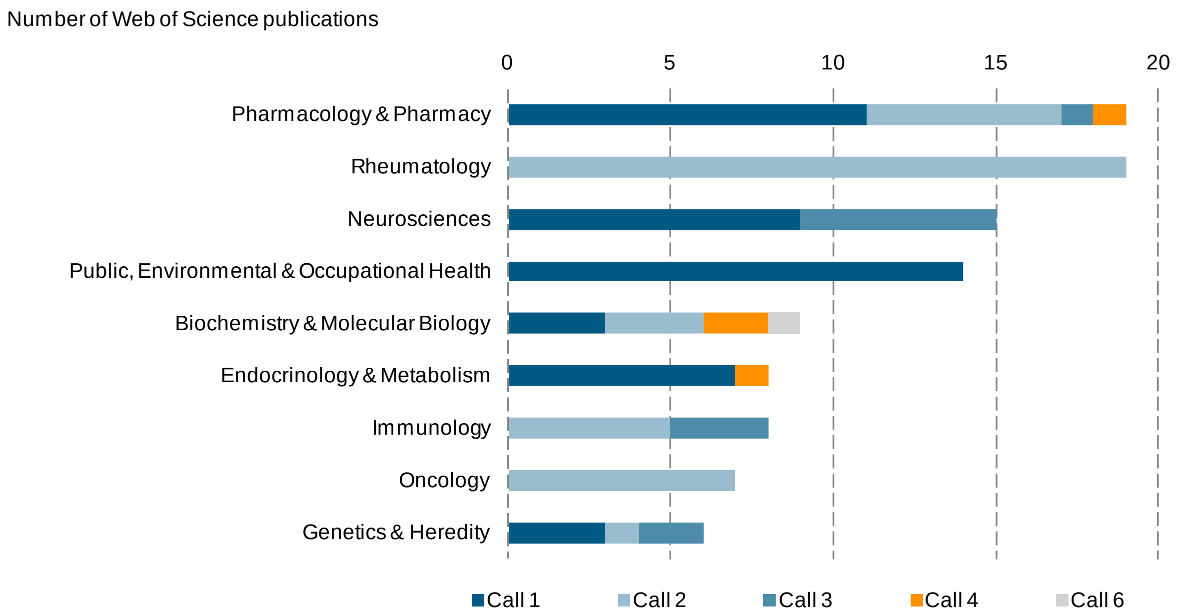
| Journal | Number of *Web of Science* publications | Number of papers | Journal Impact Factor (2012) | *Web of Science* journal categories |
|---|---|---|---|---|
| Genome Biology | 3 | 3 | 10.288 | Biotechnology & Applied Microbiology<br>Genetics & Heredity |

These journal titles encompass a broad range of research disciplines and confirm the cross-disciplinary nature of IMI project research.

Standard bibliometric methodology uses journal category as a proxy for research field. Journals are assigned to one or more categories and every publication within that journal is subsequently assigned to that category.

Figure A1.2.1 shows the top *Web of Science* journal categories associated with new IMI project publications. Only categories with at least 6 publications are included.

FIGURE A1.2.1 TOP *WEB OF SCIENCE* JOURNAL CATEGORIES IN WHICH NEW IMI PROJECT PUBLICATIONS ARE PUBLISHED




Pharmacology & Pharmacy and Rheumatology are the most frequent *Web of Science* journal categories for new IMI publications overall (both 19 publications). However, whereas Pharmacology & Pharmacy is related to multiple projects and Calls, Rheumatology research arises from the Call 2 project BTCure. Neurosciences is the third most frequent *Web of Science* journal category. Public, Environmental & Occupational Health is the fourth most frequent journal category mainly related to the Protect project, with research published in Drug Safety.

## A1.2 SUMMARY OF NEW IMI-SUPPORTED PUBLICATIONS – CITATIONS

A summary of new IMI-supported publications is shown in Table A1.2.1. Although these publications are relatively recent, nearly two-fifths (39.2%) have already been cited.

TABLE A1.2.1 SUMMARY INDICATORS FOR NEW IMI PROJECT PUBLICATIONS

| Call | Project | Number of *Web of Science* publications | Number of cited *Web of Science* publications | Total citations (current) |
|---|---|---|---|---|
| | [Unassigned] | [17] | [10] | [27] |
| 1 | eTOX | 2 | 1 | 1 |
| 1 | EUROPAIN | 7 | 3 | 3 |
| 1 | IMIDIA | 10 | 4 | 4 |
| 1 | MARCAR | 6 | 3 | 7 |
| 1 | NEWMEDS | 14 | 3 | 4 |
| 1 | PharmaCog | 3 | 0 | 0 |
| 1 | PROactive | 3 | 1 | 4 |
| 1 | PROTECT | 16 | 14 | 64 |
| 1 | SafeSciMET | 1 | 0 | 0 |
| 1 | SUMMIT | 4 | 0 | 0 |
| 1 | U-BIOPRED | 4 | 1 | 3 |
| 2 | BTCure | 34 | 14 | 20 |
| 2 | DDMoRe | 5 | 2 | 2 |
| 2 | OncoTrack | 4 | 2 | 3 |
| 2 | Open PHACTS | 1 | 0 | 0 |
| 2 | QuIC-ConCePT | 6 | 0 | 0 |
| 2 | RAPP-ID | 3 | 0 | 0 |
| 3 | ABIRISK | 1 | 0 | 0 |
| 3 | BioVacSafe | 4 | 0 | 0 |
| 3 | DIRECT | 2 | 2 | 3 |
| 3 | EU-AIMS | 12 | 7 | 15 |
| 3 | MIP-DILI | 1 | 0 | 0 |
| 3 | PreDiCT-TB | 1 | 0 | 0 |
| 4 | CHEM21 | 4 | 1 | 1 |
| 4 | EMIF | 2 | 2 | 2 |
| 4 | K4DD | 1 | | 0 |
| 4 | ORBITO | 1 | 1 | 1 |
| 4 | STEMBANCC | 2 | | 0 |
| 6 | COMBACTE | 1 | | 0 |
| 6 | TRANSLOCATION | 3 | 1 | 2 |
| | **Overall** | **158** | **62** | **139** |

Of the remaining IMI projects with no new *Web of Science* publications (not listed in Table A1.2.1):

- Call 1: EMTRAIN, Eu2P, PharmaTrain and SAFE-T
- Call 2: EHR4CR and Predect
- Call 3: EUPATI
- Call 4: COMPACT and eTRIKS
- Call 5: ELF

## ANNEX 2: DEFINITIONS OF *WEB OF SCIENCE* JOURNAL CATEGORIES

**Biochemistry & Molecular Biology** is concerned with journals that deal with general biochemistry and molecular biology topics such as carbohydrates, lipids, proteins, nucleic acids, genes, drugs, toxic substances, and other chemical or molecular constituents of cells, microbes, and higher plants and animals, including humans. Journals that focus on biochemistry in cells, tissues or organs and those whose primary focus is the organism of study, (such as plants, microbes, and so forth) are excluded as are journals that focus on methods in biochemistry or molecular biology.

**Clinical Neurology** covers journals on all areas of clinical research and medical practice in neurology. The focus is on traditional neurological illnesses and diseases such as dementia, stroke, epilepsy, headache, multiple sclerosis, and movement disorders that have clinical and socio-economic importance. This category also includes journals on medical specialties such as paediatric neurology, neurosurgery, neuroradiology, pain management, and neuropsychiatry that affect neurological diagnosis and treatment.

**Endocrinology & Metabolism** includes journals focused on endocrine glands; the regulation of cell, organ, and system function by the action of secreted hormones; the generation and chemical/biological properties of these substances; and the pathogenesis and treatment of disorders associated with either source or target organs. Specific areas covered include neuroendocrinology, reproductive endocrinology, pancreatic hormones and diabetes, regulation of bone formation and loss, and control of growth.

**Genetics & Heredity** includes journals that deal with the structure, functions, and properties of genes, and the characteristics of inheritance. This category also considers heritable traits, population genetics, frequency and distribution of polymorphism, as well as inherited diseases and disorders of the replicative process. The category is distinguishable from Biochemistry & Molecular Biology by its specific emphasis on the gene as a single functional unit, and on the gene's effect on the organism as a whole.

**Mathematical & Computational Biology** includes journals concerning the use of mathematical, statistical and computational methods to address data analysis, modelling, and information management in biological problems, processes and systems. Among the areas covered are biostatistics, bioinformatics, biometrics, modelling of biological systems, and computational biology.

**Neurosciences** covers journals on all areas of basic research on the brain, neural physiology, and function in health and disease. The areas of focus include neurotransmitters, neuropeptides, neurochemistry, neural development, and neural behaviour. Coverage also includes journals in neuro-endocrine and neuro-immune systems, somatosensory system, motor system and sensory motor integration, autonomic system as well as diseases of the nervous system.

**Oncology** covers journals on the mechanisms, causes, and treatments of cancer including environmental and genetic risk factors, and cellular and molecular carcinogenesis. Aspects of clinical oncology covered include surgical, radiological, chemical, and palliative care. This category is also concerned with journals on cancers of specific systems and organs.

**Pharmacology & Pharmacy** contains journals on the discovery and testing of bioactive substances, including animal research, clinical experience, delivery systems, and dispensing of drugs. This category also includes journals on the biochemistry, metabolism, and toxic or adverse effects of drugs.

**Psychiatry** covers journals that focus on the origins, diagnosis, and treatment of mental, emotional, or behavioural disorders. Areas covered in this category include adolescent and child psychiatry, forensic psychiatry, geriatric psychiatry, hypnosis, psychiatric nursing, psychiatric rehabilitation, psychosomatic research, and stress medicine.

**Research & Experimental Medicine** includes journals describing general medical research with a particular emphasis on extremely novel techniques and clinical interventions in a broad range of medical specialisations and applications, including vaccine development, tissue replacement, immunotherapies, and other experimental therapeutic strategies. Journals in this category reflect clinical interventions that are in early stages of development, using in vitro or animal models, and small-scale clinical trials.

**Rheumatology** covers journals on clinical, therapeutic, and laboratory research about arthritis and rheumatism, the chronic degenerative autoimmune inflammatory diseases that primarily affect joints and connective tissue.

## ANNEX 3: BIBLIOGRAPHY OF HIGHLY-CITED PAPERS, 'HOT PAPERS' AND THOSE PAPERS WITH HIGHEST INTERDISCIPLINARITY

This Annex considers the cumulative dataset of IMI project publications that have been linked to records in Thomson Reuters citation databases.

For the purpose of this report, highly-cited papers have been defined as those articles and reviews which belong to the world's top decile of papers in that journal category and year of publication, when ranked by number of citations received. A percentage that is above 10 indicates above-average performance.

Section A3.1 lists the 146 papers (with one paper associated with two projects)[28] in the IMI project publications dataset that have been identified as highly-cited. This is a large increase over the third report to IMI where 52 papers were classified as highly-cited.

Papers are listed in ascending alphabetical order (project, first author) and those papers also identified as 'hot papers' (Section A3.2) are highlighted (bold text).

### A3.1 HIGHLY-CITED PAPERS ASSOCIATED WITH IMI PROJECTS


- ABIRISK: KIESEIER, BC et al. (2013) Disease Amelioration With Tocilizumab in a Treatment-Resistant Patient With Neuromyelitis Optica Implication for Cellular Immune Responses, JAMA Neurology, 70: 390-393, doi: 001/jamaneurol.2013.668
- ABIRISK: WENNIGER, LJMD et al. (2013) Immunoglobulin G4+clones identified by next-generation sequencing dominate the B cell receptor repertoire in immunoglobulin G4 associated cholangitis, Hepatology, 57: 2390-2398, doi: 002/hep.26232
- BioVacSafe: KAUFMANN, SHE et al. (2012) Tuberculosis vaccine development: strength lies in tenacity, Trends in Immunology, 33: 373-379, doi: 016/j.it.2012.03.004
- BioVacSafe: MAERTZDORF, J et al. (2012) Enabling biomarkers for tuberculosis control, International Journal of Tuberculosis and Lung Disease, 16: 1140-1148, doi: 588/ijtld.12.0246
- BTCure: ALBRECHT, A et al. (2013) The structural basis of MRI bone erosions: an assessment by microCT, Annals of the Rheumatic Diseases, 72: 1351-1357, doi: 136/annrheumdis-2012-201982
- BTCure: AMARA, K et al. (2013) Monoclonal IgG antibodies generated from joint-derived B cells of RA patients have a strong bias toward citrullinated autoantigen recognition, Journal of Experimental Medicine, 210: 445-455, doi: 084/jem.20121486
- BTCure: BILIAVSKA, I et al. (2013) Application of the 2010 ACR/EULAR classification criteria in patients with very early inflammatory arthritis: analysis of sensitivity, specificity and predictive values in the SAVE study cohort, Annals of the Rheumatic Diseases, 72: 1335-1341, doi: 136/annrheumdis-2012-201909
- BTCure: BONELLI, M et al. (2013) Abatacept (CTLA-4IG) treatment reduces the migratory capacity of monocytes in patients with rheumatoid arthritis, Arthritis and Rheumatism, 65: 599-607, doi: 002/art.37787
- BTCure: BRINK, M et al. (2013) Multiplex Analyses of Antibodies Against Citrullinated Peptides in Individuals Prior to Development of Rheumatoid Arthritis, Arthritis and Rheumatism, 65: 899-910, doi: 002/art.37835
- BTCure: COPE, A et al. (2011) The Th1 life cycle: molecular control of IFN-gamma to IL-10 switching, Trends in Immunology, 32: 278-286, doi: 016/j.it.2011.03.010
- BTCure: CUI, J et al. (2013) Genome-Wide Association Study and Gene Expression Analysis Identifies CD84 as a Predictor of Response to Etanercept Therapy in Rheumatoid Arthritis, PLoS Genetics, 9: , doi: 371/journal.pgen.1003394
- BTCure: FINZEL, S et al. (2013) Interleukin-6 receptor blockade induces limited repair of bone erosions in rheumatoid arthritis: a micro CT study, Annals of the Rheumatic Diseases, 72: 396-400, doi: 136/annrheumdis-2011-201075


[28] WENNIGER, LJMD et al. (2013) Immunoglobulin G4+clones identified by next-generation sequencing dominate the B cell receptor repertoire in immunoglobulin G4 associated cholangitis, Hepatology, 57: 2390-2398, doi: 002/hep.26232

- BTCure: GERLAG, DM et al. (2012) EULAR recommendations for terminology and research in individuals at risk of rheumatoid arthritis: report from the Study Group for Risk Factors for Rheumatoid Arthritis, Annals of the Rheumatic Diseases, 71: 638-641, doi: 136/annrheumdis-2011-200990
- **BTCure: HARRE, U et al. (2012) Induction of osteoclastogenesis and bone loss by human autoantibodies against citrullinated vimentin, Journal of Clinical Investigation, 122: 1791-1802, doi: 172/JCI60975**
- BTCure: KIECHL, S et al. (2013) Blockade of receptor activator of nuclear factor-kappa B (RANKL) signaling improves hepatic insulin resistance and prevents development of diabetes mellitus, Nature Medicine, 19: 358-363, doi: 038/nm.3084
- BTCure: LIN, NY et al. (2013) Autophagy regulates TNF alpha-mediated joint destruction in experimental arthritis, Annals of the Rheumatic Diseases, 72: 761-768, doi: 136/annrheumdis-2012-201671
- BTCure: LUNDBERG, K et al. (2013) Genetic and environmental determinants for disease risk in subsets of rheumatoid arthritis defined by the anticitrullinated protein/peptide antibody fine specificity profile, Annals of the Rheumatic Diseases, 72: 652-658, doi: 136/annrheumdis-2012-201484
- BTCure: NIKITOPOULOU, I et al. (2012) Autotaxin expression from synovial fibroblasts is essential for the pathogenesis of modeled arthritis, Journal of Experimental Medicine, 209: 923-931, doi: 084/jem.20112012
- BTCure: PANDIS, I et al. (2012) Identification of microRNA-221/222 and microRNA-323-3p association with rheumatoid arthritis via predictions using the human tumour necrosis factor transgenic mouse model, Annals of the Rheumatic Diseases, 71: 1716-1723, doi: 136/annrheumdis-2011-200803
- BTCure: SCHETT, G et al. (2012) Bone erosion in rheumatoid arthritis: mechanisms, diagnosis and treatment, Nature Reviews Rheumatology, 8: 656-664, doi: 038/nrrheum.2012.153
- BTCure: SHI, J et al. (2011) Autoantibodies recognizing carbamylated proteins are present in sera of patients with rheumatoid arthritis and predict joint damage, Proceedings of the National Academy of Sciences of the United States of America, 108: 17372-17377, doi: 073/pnas.1114465108
- BTCure: SUWANNALAI, P et al. (2012) Avidity maturation of anti-citrullinated protein antibodies in rheumatoid arthritis, Arthritis and Rheumatism, 64: 1323-1328, doi: 002/art.33489
- BTCure: TRENKMANN, M et al. (2013) Tumor Necrosis Factor alpha-Induced MicroRNA-18a Activates Rheumatoid Arthritis Synovial Fibroblasts Through a Feedback Loop in NF-kappa B Signaling, Arthritis and Rheumatism, 65: 916-927, doi: 002/art.37834
- BTCure: TROUW, LA et al. (2013) Autoimmunity in rheumatoid arthritis: different antigens-common principles, Annals of the Rheumatic Diseases, 72: 132-136, doi: 136/annrheumdis-2012-202349
- BTCure: WALTER, GJ et al. (2013) Interaction with activated monocytes enhances cytokine expression and suppressive activity of human CD4+CD45ro+CD25+CD127low regulatory T cells, Arthritis and Rheumatism, 65: 627-638, doi: 002/art.37832
- BTCure: WENNIGER, LJMD et al. (2013) Immunoglobulin G4+clones identified by next-generation sequencing dominate the B cell receptor repertoire in immunoglobulin G4 associated cholangitis, Hepatology, 57: 2390-2398, doi: 002/hep.26232
- BTCure: WILLEMZE, A et al. (2012) The ACPA recognition profile and subgrouping of ACPA-positive RA patients, Annals of the Rheumatic Diseases, 71: 268-274, doi: 136/annrheumdis-2011-200421
- eTOX: BAUER-MEHREN, A et al. (2012) Automatic Filtering and Substantiation of Drug Safety Signals, PLoS Computational Biology, 8: , doi: 371/journal.pcbi.1002457
- eTOX: ENOCH, SJ et al. (2011) A review of the electrophilic reaction chemistry involved in covalent protein binding relevant to toxicity, Critical Reviews in Toxicology, 41: 783-802, doi: 109/10408444.2011.598141
- eTOX: KLEPSCH, F et al. (2011) Exhaustive Sampling of Docking Poses Reveals Binding Hypotheses for Propafenone Type Inhibitors of P-Glycoprotein, PLoS Computational Biology, 7: , doi: 371/journal.pcbi.1002036
- eTOX: MUTHAS, D et al. (2013) Exploiting Pharmacological Similarity to Identify Safety Concerns - Listen to What the Data Tells You, Molecular Informatics, 32: 37-45, doi: 002/minf.201200088
- eTOX: MUTHAS, D et al. (2013) A critical assessment of modeling safety-related drug attrition, MedChemComm, 4: 1058-1065, doi: 039/c3md00072a

- eTOX: OBIOL-PARDO, C et al. (2011) A Multiscale Simulation System for the Prediction of Drug-Induced Cardiotoxicity, Journal of Chemical Information and Modeling, 51: 483-492, doi: 021/ci100423z
- eTOX: PIECHOTA, P et al. (2013) Pragmatic Approaches to Using Computational Methods To Predict Xenobiotic Metabolism, Journal of Chemical Information and Modeling, 53: 1282-1293, doi: 021/ci400050v
- eTOX: PRZYBYLAK, KR et al. (2012) In silico models for drug-induced liver injury - current status, Expert Opinion on Drug Metabolism & Toxicology, 8: 201-217, doi: 517/17425255.2012.648613
- eTOX: TABOUREAU, O et al. (2011) ChemProt: a disease chemical biology database, Nucleic Acids Research, 39: D367-D372, doi: 093/nar/gkq906
- eTOX: VIDAL, D et al. (2010) In Silico Receptorome Screening of Antipsychotic Drugs, Molecular Informatics, 29: 543-551, doi: 002/minf.201000055
- EU-AIMS: BAUDOUIN, SJ et al. (2012) Shared Synaptic Pathophysiology in Syndromic and Nonsyndromic Rodent Models of Autism, Science, 338: 128-132, doi: 126/science.1224159
- EU-AIMS: BUDRECK, EC et al. (2013) Neuroligin-1 controls synaptic abundance of NMDA-type glutamate receptors through extracellular coupling, Proceedings of the National Academy of Sciences of the United States of America, 110: 725-730, doi: 073/pnas.1214718110
- EU-AIMS: DELORME, R et al. (2013) Progress toward treatments for synaptic defects in autism, Nature Medicine, 19: 685-694, doi: 038/nm.3193
- EU-AIMS: ECKER, C et al. (2013) Translational approaches to the biology of Autism: false dawn or a new era?, Molecular Psychiatry, 18: 435-442, doi: 038/mp.2012.102
- EU-AIMS: EL-KORDI, A et al. (2013) Development of an autism severity score for mice using Nlgn4 null mutants as a construct-valid model of heritable monogenic autism, Behavioural Brain Research, 251: 41-49, doi: 016/j.bbr.2012.11.016
- **EU-AIMS: KONG, A et al. (2012) Rate of de novo mutations and the importance of fathers age to disease risk, Nature, 488: 471-475, doi: 038/nature11396**
- EU-AIMS: LAI, MC et al. (2013) Biological sex affects the neurobiology of autism, Brain, 136: 2799-2815, doi: 093/brain/awt216
- EU-AIMS: PERSICO, AM et al. (2013) Autism genetics, Behavioural Brain Research, 251: 95-112, doi: 016/j.bbr.2013.06.012
- EU-AIMS: SPOOREN, W et al. (2012) Synapse dysfunction in autism: a molecular medicine approach to drug discovery in neurodevelopmental disorders, Trends in Pharmacological Sciences, 33: 669-684, doi: 016/j.tips.2012.09.004
- EU-AIMS: STEIN, JL et al. (2012) Identification of common variants associated with human hippocampal and intracranial volumes, Nature Genetics, 44: 552-+, doi: 038/ng.2250
- EU-AIMS: WHELAN, R et al. (2012) Adolescent impulsivity phenotypes characterized by distinct brain networks, Nature Neuroscience, 15: 920-U153, doi: 038/nn.3092
- EUROPAIN: AASVANG, EK et al. (2010) Predictive Risk Factors for Persistent Postherniotomy Pain, Anesthesiology, 112: 957-969, doi: 097/ALN.0b013e3181d31ff8
- EUROPAIN: ANDERSEN, KG et al. (2011) Persistent Pain After Breast Cancer Treatment: A Critical Review of Risk Factors and Strategies for Prevention, Journal of Pain, 12: 725-746, doi: 016/j.jpain.2010.12.005
- EUROPAIN: ANDREWS, N et al. (2012) Spontaneous burrowing behaviour in the rat is reduced by peripheral nerve injury or inflammation associated pain, European Journal of Pain, 16: 485-495, doi: 016/j.ejpain.2011.07.012
- EUROPAIN: BAASTRUP, C et al. (2010) Spinal-, brainstem- and cerebrally mediated responses at- and below-level of a spinal cord contusion in rats: Evaluation of pain-like behavior, Pain, 151: 670-679, doi: 016/j.pain.2010.08.024
- EUROPAIN: BARON, R et al. (2012) Subgrouping of patients with neuropathic pain according to pain-related sensory abnormalities: a first step to a stratified treatment approach, Lancet Neurology, 11: 999-1005

- EUROPAIN: CALVO, M et al. (2012) The role of the immune system in the generation of neuropathic pain, Lancet Neurology, 11: 629-642
- EUROPAIN: DERRY, S et al. (2013) Topical capsaicin (high concentration) for chronic neuropathic pain in adults, Cochrane Database of Systematic Reviews, : , doi: 002/14651858.CD007393.pub3
- EUROPAIN: EIJKELKAMP, N et al. (2013) A role for Piezo2 in EPAC1-dependent mechanical allodynia, Nature Communications, 4: , doi: 038/ncomms2673
- **EUROPAIN: FINNERUP, NB et al. (2010) The evidence for pharmacological treatment of neuropathic pain, Pain, 150: 573-581, doi: 016/j.pain.2010.06.019**
- EUROPAIN: HAROUTIUNIAN, S et al. (2013) The neuropathic component in persistent postsurgical pain: A systematic literature review, Pain, 154: 95-102, doi: 016/j.pain.2012.09.010
- EUROPAIN: HAUSER, W et al. (2012) The Role of Antidepressants in the Management of Fibromyalgia Syndrome A Systematic Review and Meta-Analysis, CNS Drugs, 26: 297-307
- EUROPAIN: HUANG, WL et al. (2013) A clinically relevant rodent model of the HIV antiretroviral drug stavudine induced painful peripheral neuropathy, Pain, 154: 560-575, doi: 016/j.pain.2012.12.023
- EUROPAIN: LENZ, M et al. (2011) Bilateral somatosensory cortex disinhibition in complex regional pain syndrome type I, Neurology, 77: 1096-1101, doi: 212/WNL.0b013e31822e1436
- EUROPAIN: PHILLIPS, TJC et al. (2010) Pharmacological Treatment of Painful HIV-Associated Sensory Neuropathy: A Systematic Review and Meta-Analysis of Randomised Controlled Trials, PLoS One, 5: , doi: 371/journal.pone.0014433
- EUROPAIN: QUICK, K et al. (2012) TRPC3 and TRPC6 are essential for normal mechanotransduction in subsets of sensory neurons and cochlear hair cells, Open Biology, 2: , doi: 098/rsob.120068
- EUROPAIN: SERRA, J et al. (2012) Microneurographic identification of spontaneous activity in C-nociceptors in neuropathic pain states in humans and rats, Pain, 153: 42-55, doi: 016/j.pain.2011.08.015
- EUROPAIN: TREEDE, RD et al. (2013) Mechanism- and experience-based strategies to optimize treatment response to the capsaicin 8% cutaneous patch in patients with localized neuropathic pain, Current Medical Research and Opinion, 29: 527-538, doi: 185/03007995.2013.781019
- EUROPAIN: WILDGAARD, K et al. (2011) Consequences of persistent pain after lung cancer surgery: a nationwide questionnaire study, Acta Anaesthesiologica Scandinavica, 55: 60-68, doi: 111/j.1399-6576.2010.02357.x
- IMIDIA: BUGLIANI, M et al. (2013) Microarray analysis of isolated human islet transcriptome in type 2 diabetes and the role of the ubiquitin-proteasome system in pancreatic beta cell dysfunction, Molecular and Cellular Endocrinology, 367: 1-10, doi: 016/j.mce.2012.12.001
- IMIDIA: KLEE, P et al. (2011) Connexins protect mouse pancreatic beta cells against apoptosis, Journal of Clinical Investigation, 121: 4870-4879, doi: 172/JCI40509
- IMIDIA: LORTZ, S et al. (2013) Overexpression of the antioxidant enzyme catalase does not interfere with the glucose responsiveness of insulin-secreting INS-1E cells and rat islets, Diabetologia, 56: 774-782, doi: 007/s00125-012-2823-7
- IMIDIA: RAVASSARD, P et al. (2011) A genetically engineered human pancreatic beta cell line exhibiting glucose-inducible insulin secretion, Journal of Clinical Investigation, 121: 3589-3597, doi: 172/JCI58447
- IMIDIA: ROGGLI, E et al. (2010) Involvement of MicroRNAs in the Cytotoxic Effects Exerted by Proinflammatory Cytokines on Pancreatic beta-Cells, Diabetes, 59: 978-986, doi: 337/db09-0881
- MARCAR: LEMPIAINEN, H et al. (2013) Identification of Dlk1-Dio3 Imprinted Gene Cluster Noncoding RNAs as Novel Candidate Biomarkers for Liver Tumor Promotion, Toxicological Sciences, 131: 375-386, doi: 093/toxsci/kfs303
- MARCAR: REDDINGTON, JP et al. (2013) Non-canonical functions of the DNA methylome in gene regulation, Biochemical Journal, 451: 13-23, doi: 042/BJ20121585
- MARCAR: REDDINGTON, JP et al. (2013) Redistribution of H3K27me3 upon DNA hypomethylation results in de-repression of Polycomb target genes, Genome Biology, 14: , doi: 186/gb-2013-14-3-r25
- **MARCAR: THOMSON, JP et al. (2012) Non-genotoxic carcinogen exposure induces defined changes in the 5-hydroxymethylome, Genome Biology, 13: , doi: 186/gb-2012-13-10-R93**

- MARCAR: THOMSON, JP et al. (2013) Dynamic changes in 5-hydroxymethylation signatures underpin early and late events in drug exposed liver, Nucleic Acids Research, 41: 5639-5654, doi: 093/nar/gkt232
- NEWMEDS: ANACKER, C et al. (2013) Role for the kinase SGK1 in stress, depression, and glucocorticoid effects on hippocampal neurogenesis, Proceedings of the National Academy of Sciences of the United States of America, 110: 8708-8713, doi: 073/pnas.1300886110
- NEWMEDS: ARTIGAS, F et al. (2013) Serotonin receptors involved in antidepressant effects, Pharmacology & Therapeutics, 137: 119-131, doi: 016/j.pharmthera.2012.09.006
- NEWMEDS: BORTOLOZZI, A et al. (2012) Selective siRNA-mediated suppression of 5-HT1A autoreceptors evokes strong anti-depressant-like effects, Molecular Psychiatry, 17: 612-623, doi: 038/mp.2011.92
- NEWMEDS: BUSSEY, TJ et al. (2012) New translational assays for preclinical modelling of cognition in schizophrenia: The touchscreen testing method for mice and rats, Neuropharmacology, 62: 1191-1203, doi: 016/j.neuropharm.2011.04.011
- NEWMEDS: DIX, S et al. (2010) A within-subject cognitive battery in the rat: differential effects of NMDA receptor antagonists, Psychopharmacology, 212: 227-242, doi: 007/s00213-010-1945-1
- NEWMEDS: GASTAMBIDE, F et al. (2012) Selective Remediation of Reversal Learning Deficits in the Neurodevelopmental MAM Model of Schizophrenia by a Novel mGlu5 Positive Allosteric Modulator, Neuropsychopharmacology, 37: 1057-1066, doi: 038/npp.2011.298
- NEWMEDS: GILMOUR, G et al. (2012) NMDA receptors, cognition and schizophrenia - Testing the validity of the NMDA receptor hypofunction hypothesis, Neuropharmacology, 62: 1401-1412, doi: 016/j.neuropharm.2011.03.015
- NEWMEDS: HRADETZKY, E et al. (2012) The Methylazoxymethanol Acetate (MAM-E17) Rat Model: Molecular and Functional Effects in the Hippocampus, Neuropsychopharmacology, 37: 364-377, doi: 038/npp.2011.219
- NEWMEDS: INGASON, A et al. (2011) Maternally Derived Microduplications at 15q11-q13: Implication of Imprinted Genes in Psychotic Illness, American Journal of Psychiatry, 168: 408-417, doi: 176/appi.ajp.2010.09111660
- **NEWMEDS: JACQUEMONT, S et al. (2011) Mirror extreme BMI phenotypes associated with gene dosage at the chromosome 16p11.2 locus, Nature, 478: 97-U111, doi: 038/nature10406**
- **NEWMEDS: KAPUR, S et al. (2012) Why has it taken so long for biological psychiatry to develop clinical tests and what to do about it?, Molecular Psychiatry, 17: 1174-1179, doi: 038/mp.2012.105**
- NEWMEDS: KEELER, JF et al. (2011) Translating cognition from animals to humans, Biochemical Pharmacology, 81: 1356-1366, doi: 016/j.bcp.2010.12.028
- **NEWMEDS: KIROV, G et al. (2012) De novo CNV analysis implicates specific abnormalities of postsynaptic signalling complexes in the pathogenesis of schizophrenia, Molecular Psychiatry, 17: 142-153, doi: 038/mp.2011.154**
- NEWMEDS: LLADO-PELFORT, L et al. (2012) 5-HT1A Receptor Agonists Enhance Pyramidal Cell Firing in Prefrontal Cortex Through a Preferential Action on GABA Interneurons, Cerebral Cortex, 22: 1487-1497, doi: 093/cercor/bhr220
- NEWMEDS: MCALLISTER, KAL et al. (2013) Dissociation between memory retention across a delay and pattern separation following medial prefrontal cortex lesions in the touchscreen TUNL task, Neurobiology of Learning and Memory, 101: 120-126, doi: 016/j.nlm.2013.01.010
- NEWMEDS: PLICHTA, MM et al. (2012) Test-retest reliability of evoked BOLD signals from a cognitive-emotive fMRI test battery, Neuroimage, 60: 1746-1758, doi: 016/j.neuroimage.2012.01.129
- NEWMEDS: SMITH, JW et al. (2011) A comparison of the effects of ketamine and phencyclidine with other antagonists of the NMDA receptor in rodent assays of attention and working memory, Psychopharmacology, 217: 255-269, doi: 007/s00213-011-2277-5
- **NEWMEDS: SULLIVAN, PF et al. (2013) A mega-analysis of genome-wide association studies for major depressive disorder, Molecular Psychiatry, 18: 497-511, doi: 038/mp.2012.21**
- NEWMEDS: TANSEY, KE et al. (2013) Contribution of Common Genetic Variants to Antidepressant Response, Biological Psychiatry, 73: 679-682, doi: 016/j.biopsych.2012.10.030

- NEWMEDS: UHER, R et al. (2012) Depression symptom dimensions as predictors of antidepressant treatment outcome: replicable evidence for interest-activity symptoms, Psychological Medicine, 42: 967-980, doi: 017/S0033291711001905
- Onco Track: ALGAR, WR et al. (2012) Quantum Dots as Simultaneous Acceptors and Donors in Time-Gated Forster Resonance Energy Transfer Relays: Characterization and Biosensing, Journal of the American Chemical Society, 134: 1876-1891, doi: 021/ja210162f
- Onco Track: HILDEBRANDT, N et al. (2011) Biofunctional Quantum Dots: Controlled Conjugation for Multiplexed Biosensors, ACS Nano, 5: 5286-5290, doi: 021/nn2023123
- Onco Track: HOTZER, B et al. (2012) Fluorescence in Nanobiotechnology: Sophisticated Fluorophores for Novel Applications, Small, 8: 2297-2326, doi: 002/smll.201200109
- Onco Track: JIN, ZW et al. (2012) Semiconductor quantum dots for in vitro diagnostics and cellular imaging, Trends in Biotechnology, 30: 394-403, doi: 016/j.tibtech.2012.04.005
- Onco Track: LANDEGREN, U et al. (2012) Opportunities for Sensitive Plasma Proteome Analysis, Analytical Chemistry, 84: 1824-1830, doi: 021/ac2032222
- Onco Track: TAIWO, O et al. (2012) Methylome analysis using MeDIP-seq with low DNA concentrations, Nature Protocols, 7: 617-636, doi: 038/nprot.2012.012
- Open PHACTS: CLARK, AM et al. (2012) Redefining Cheminformatics with Intuitive Collaborative Mobile Apps, Molecular Informatics, 31: 569-584, doi: 002/minf.201200010
- Open PHACTS: FURLONG, LI et al. (2013) Human diseases through the lens of network biology, Trends in Genetics, 29: 150-159, doi: 016/j.tig.2012.11.004
- Open PHACTS: KJAERULFF, SK et al. (2013) ChemProt-2.0: visual navigation in a disease chemical biology database, Nucleic Acids Research, 41: D464-D469, doi: 093/nar/gks1166
- Open PHACTS: WILD, DJ et al. (2012) Systems chemical biology and the Semantic Web: what they mean for the future of drug discovery research, Drug Discovery Today, 17: 469-474, doi: 016/j.drudis.2011.12.019
- **Open PHACTS: WILLIAMS, AJ et al. (2012) Towards a gold standard: regarding quality in public domain chemistry databases and approaches to improving the situation, Drug Discovery Today, 17: 685-701**
- Pharma-Cog: BABILONI, C et al. (2013) Resting state cortical electroencephalographic rhythms are related to gray matter volume in subjects with mild cognitive impairment and Alzheimers disease, Human Brain Mapping, 34: 1427-1446, doi: 002/hbm.22005
- Pharma-Cog: CARRILLO, MC et al. (2012) Worldwide Alzheimers Disease Neuroimaging Initiative, Alzheimers & Dementia, 8: 337-342, doi: 016/j.jalz.2012.04.007
- **Pharma-Cog: FRISONI, GB et al. (2010) The clinical use of structural MRI in Alzheimer disease, Nature Reviews Neurology, 6: 67-77, doi: 038/nrneurol.2009.215**
- Pharma-Cog: LANGUILLE, S et al. (2012) The grey mouse lemur: A non-human primate model for ageing studies, Ageing Research Reviews, 11: 150-162, doi: 016/j.arr.2011.07.001
- PRO-active: VAN REMOORTEL, H et al. (2012) Validity of Six Activity Monitors in Chronic Obstructive Pulmonary Disease: A Comparison with Indirect Calorimetry, PLoS One, 7: , doi: 371/journal.pone.0039198
- PROTECT: HANSEN, RA et al. (2013) How Well Do Various Health Outcome Definitions Identify Appropriate Cases in Observational Studies?, Drug Safety, 36: S27-S32, doi: 007/s40264-013-0104-0
- PROTECT: MADIGAN, D et al. (2013) Empirical Performance of the Case-Control Method: Lessons for Developing a Risk Identification and Analysis System, Drug Safety, 36: S73-S82, doi: 007/s40264-013-0105-z
- PROTECT: NOREN, GN et al. (2013) Empirical Performance of the Calibrated Self-Controlled Cohort Analysis Within Temporal Pattern Discovery: Lessons for Developing a Risk Identification and Analysis System, Drug Safety, 36: S107-S121, doi: 007/s40264-013-0095-x
- PROTECT: REICH, CG et al. (2013) Alternative Outcome Definitions and Their Effect on the Performance of Methods for Observational Outcome Studies, Drug Safety, 36: S181-S193, doi: 007/s40264-013-0111-1

- PROTECT: REICH, CG et al. (2013) The Impact of Drug and Outcome Prevalence on the Feasibility and Performance of Analytical Methods for a Risk Identification and Analysis System, Drug Safety, 36: S195-S204, doi: 007/s40264-013-0112-0
- PROTECT: RYAN, PB et al. (2013) Defining a Reference Set to Support Methodological Research in Drug Safety, Drug Safety, 36: S33-S47, doi: 007/s40264-013-0097-8
- PROTECT: RYAN, PB et al. (2013) Empirical Performance of a New User Cohort Method: Lessons for Developing a Risk Identification and Analysis System, Drug Safety, 36: S59-S72, doi: 007/s40264-013-0099-6
- PROTECT: RYAN, PB et al. (2013) Empirical Performance of a Self-Controlled Cohort Method: Lessons for Developing a Risk Identification and Analysis System, Drug Safety, 36: S95-S106, doi: 007/s40264-013-0101-3
- PROTECT: RYAN, PB et al. (2013) A Comparison of the Empirical Performance of Methods for a Risk Identification System, Drug Safety, 36: S143-S158, doi: 007/s40264-013-0108-9
- PROTECT: RYAN, PB et al. (2013) Evaluating Performance of Risk Identification Methods Through a Large-Scale Simulation of Observational Data, Drug Safety, 36: S171-S180, doi: 007/s40264-013-0110-2
- PROTECT: SCHUEMIE, MJ et al. (2013) Empirical Performance of LGPS and LEOPARD: Lessons for Developing a Risk Identification and Analysis System, Drug Safety, 36: S133-S142, doi: 007/s40264-013-0107-x
- PROTECT: SCHUEMIE, MJ et al. (2013) Replication of the OMOP Experiment in Europe: Evaluating Methods for Risk Identification in Electronic Health Record Databases, Drug Safety, 36: S159-S169, doi: 007/s40264-013-0109-8
- PROTECT: STANG, PE et al. (2013) Variation in Choice of Study Design: Findings from the Epidemiology Design Decision Inventory and Evaluation (EDDIE) Survey, Drug Safety, 36: S15-S25, doi: 007/s40264-013-0103-1
- PROTECT: SUCHARD, MA et al. (2013) Empirical Performance of the Self-Controlled Case Series Design: Lessons for Developing a Risk Identification and Analysis System, Drug Safety, 36: S83-S93, doi: 007/s40264-013-0100-4
- PROTECT: VAN STAA, TP et al. (2012) Glucose-lowering agents and the patterns of risk for cancer: a study with the General Practice Research Database and secondary care data, Diabetologia, 55: 654-665, doi: 007/s00125-011-2390-3
- Quic-Concept : ASSELIN, MC et al. (2012) Quantifying heterogeneity in human tumours using MRI and PET, European Journal of Cancer, 48: 447-455, doi: 016/j.ejca.2011.12.025
- Quic-Concept : LAMBIN, P et al. (2012) Radiomics: Extracting more information from medical images using advanced feature analysis, European Journal of Cancer, 48: 441-446, doi: 016/j.ejca.2011.11.036
- Quic-Concept : LAMBIN, P et al. (2013) Predicting outcomes in radiation oncology-multifactorial decision support systems, Nature Reviews Clinical Oncology, 10: 27-40, doi: 038/nrclinonc.2012.196
- Quic-Concept : NGUYEN, QD et al. (2012) Imaging apoptosis with positron emission tomography: Bench to bedside development of the caspase-3/7 specific radiotracer [F-18]ICMT-11, European Journal of Cancer, 48: 432-440, doi: 016/j.ejca.2011.11.033
- Quic-Concept : SOLOVIEV, D et al. (2012) [F-18]FLT: An imaging biomarker of tumour proliferation for assessment of tumour response to treatment, European Journal of Cancer, 48: 416-424, doi: 016/j.ejca.2011.11.035
- RAPP-ID: BARNKOB, R et al. (2012) Measuring acoustic energy density in microchannel acoustophoresis using a simple and rapid light-intensity method, Lab on a Chip, 12: 2337-2344, doi: 039/c2lc40120g
- **SUMMIT: BOEKHOLDT, SM et al. (2012) Association of LDL Cholesterol, Non-HDL Cholesterol, and Apolipoprotein B Levels With Risk of Cardiovascular Events Among Patients Treated With Statins A Meta-analysis, JAMA - Journal of the American Medical Association, 307: 1302-1309**
- SUMMIT: BONI, E et al. (2012) A Reconfigurable and Programmable FPGA-Based System for Nonstandard Ultrasound Methods, IEEE Transactions on Ultrasonics Ferroelectrics and Frequency Control, 59: 1378-1385, doi: 109/TUFFC.2012.2338

- SUMMIT: ROCCA, B et al. (2012) The recovery of platelet cyclooxygenase activity explains interindividual variability in responsiveness to low-dose aspirin in patients with and without diabetes, Journal of Thrombosis and Haemostasis, 10: 1220-1230, doi: 111/j.1538-7836.2012.04723.x
- SUMMIT: SANDHOLM, N et al. (2012) New Susceptibility Loci Associated with Kidney Disease in Type 1 Diabetes, PLoS Genetics, 8: , doi: 371/journal.pgen.1002921
- SUMMIT: WILLIAMS, WW et al. (2012) Association Testing of Previously Reported Variants in a Large Case-Control Meta-analysis of Diabetic Nephropathy, Diabetes, 61: 2187-2194, doi: 337/db11-0751
- U-BIOPRED: ANTO, JM et al. (2012) Understanding the complexity of IgE-related phenotypes from childhood to young adulthood: A Mechanisms of the Development of Allergy (MeDALL) Seminar, Journal of Allergy and Clinical Immunology, 129: 943-U421, doi: 016/j.jaci.2012.01.047
- U-BIOPRED: AUFFRAY, C et al. (2010) An Integrative Systems Biology Approach to Understanding Pulmonary Diseases, Chest, 137: 1410-1416, doi: 378/chest.09-1850
- U-BIOPRED: BEL, EH et al. (2011) Diagnosis and definition of severe refractory asthma: an international consensus statement from the Innovative Medicine Initiative (IMI), Thorax, 66: 910-917, doi: 136/thx.2010.153643
- U-BIOPRED: CARRARO, S et al. (2013) Asthma severity in childhood and metabolomic profiling of breath condensate, Allergy, 68: 110-117, doi: 111/all.12063
- U-BIOPRED: MONTUSCHI, P et al. (2013) The Electronic Nose in Respiratory Medicine, Respiration, 85: 72-84, doi: 159/000340044
- U-BIOPRED: SANTONICO, M et al. (2012) Electronic noses calibration procedure in the context of a multicentre medical study, Sensors and Actuators B-Chemical, 173: 555-561, doi: 016/j.snb.2012.07.042
- Unassigned: AMELINK, M et al. (2013) Three phenotypes of adult-onset asthma, Allergy, 68: 674-680, doi: 111/all.12136
- Unassigned: BOUSQUET, J et al. (2012) Allergic Rhinitis and its Impact on Asthma (ARIA): Achievements in 10 years and future needs, Journal of Allergy and Clinical Immunology, 130: 1049-1062, doi: 016/j.jaci.2012.07.053
- Unassigned: MAHMOUDPOUR, SH et al. (2013) Pharmacogenetics of ACE inhibitor-induced angioedema and cough: a systematic review and meta-analysis, Pharmacogenomics, 14: 249-260, doi: 217/PGS.12.206

‘Hot papers’ have been defined as papers which are cited quickly compared with their research field (Section 3.1.3).

Section A3.2 lists the 12 papers from IMI projects that have been identified as ‘hot papers’. Papers are listed in ascending alphabetical order (project, first author) and those papers also identified as highly cited (Section A3.1) are highlighted (bold text). 11 out of 12 of these papers are classified as highly-cited.

## A3.2 ‘HOT PAPERS’ ASSOCIATED WITH IMI PROJECTS

- **BTCure: HARRE, U et al. (2012) Induction of osteoclastogenesis and bone loss by human autoantibodies against citrullinated vimentin, Journal of Clinical Investigation, 122: 1791-1802, doi: 172/JCI60975**
- eTOX: ARIGHI, CN et al. (2011) Overview of the BioCreative III Workshop, BMC Informatics, 12: , doi: 186/1471-2105-12-S8-S1
- **EU-AIMS: KONG, A et al. (2012) Rate of de novo mutations and the importance of fathers age to disease risk, Nature, 488: 471-475, doi: 038/nature11396**
- **EUROPAIN: FINNERUP, NB et al. (2010) The evidence for pharmacological treatment of neuropathic pain, Pain, 150: 573-581, doi: 016/j.pain.2010.06.019**
- **MARCAR: THOMSON, JP et al. (2012) Non-genotoxic carcinogen exposure induces defined changes in the 5-hydroxymethylome, Genome Biology, 13: , doi: 186/gb-2012-13-10-R93**
- **NEWMEDS: JACQUEMONT, S et al. (2011) Mirror extreme BMI phenotypes associated with gene dosage at the chromosome 16p11.2 locus, Nature, 478: 97-U111, doi: 038/nature10406**
- **NEWMEDS: KAPUR, S et al. (2012) Why has it taken so long for biological psychiatry to develop clinical tests and what to do about it?, Molecular Psychiatry, 17: 1174-1179, doi: 038/mp.2012.105**
- **NEWMEDS: KIROV, G et al. (2012) De novo CNV analysis implicates specific abnormalities of postsynaptic signalling complexes in the pathogenesis of schizophrenia, Molecular Psychiatry, 17: 142-153, doi: 038/mp.2011.154**
- **NEWMEDS: SULLIVAN, PF et al. (2013) A mega-analysis of genome-wide association studies for major depressive disorder, Molecular Psychiatry, 18: 497-511, doi: 038/mp.2012.21**
- **Open PHACTS: WILLIAMS, AJ et al. (2012) Towards a gold standard: regarding quality in public domain chemistry databases and approaches to improving the situation, Drug Discovery Today, 17: 685-701**
- **Pharma-Cog: FRISONI, GB et al. (2010) The clinical use of structural MRI in Alzheimer disease, Nature Reviews Neurology, 6: 67-77, doi: 038/nrneurol.2009.215**
- **SUMMIT: BOEKHOLDT, SM et al. (2012) Association of LDL Cholesterol, Non-HDL Cholesterol, and Apolipoprotein B Levels With Risk of Cardiovascular Events Among Patients Treated With Statins A Meta-analysis, JAMA - Journal of the American Medical Association, 307: 1302-1309**

Papers with the highest interdisciplinarity have been defined as those with highest diffusion score as defined Carley and Porter (Section 3.1.3).

Section A3.3 lists the five papers from IMI projects scoring highest on interdisciplinarity. These papers were selected using similar criteria as in the third report to IMI. It should be noted that as the total volume of papers from IMI projects increases, a threshold in diffusion score is less easy to define.

Papers are listed in ascending alphabetical order (project, first author). Four of these papers were also identified as highly-cited (Section A3.1) and are highlighted (bold text). None of the five papers were also identified as 'hot papers'.

## A3.3 TOP FIVE PAPERS WITH HIGHEST DIFFUSION SCORE THAT ARE ASSOCIATED WITH IMI PROJECTS

- **U-BIOPRED: MONTUSCHI, P et al. (2013) The Electronic Nose in Respiratory Medicine, Respiration, 85: 72-8410.1159/000340044**
- **Quic-Concept : ASSELIN, MC et al. (2012) Quantifying heterogeneity in human tumours using MRI and PET, European Journal of Cancer, 48: 447-45510.1016/j.ejca.2011.12.025**
- U-BIOPRED: VAN DER SCHEE, MP et al. (2013) Effect of transportation and storage using sorbent tubes of exhaled breath samples on diagnostic accuracy of electronic nose analysis, Journal of Breath Research, 7: 10.1088/1752-7155/7/1/016002
- **EU-AIMS: PERSICO, AM et al. (2013) Autism genetics, Behavioural Brain Research, 251: 95-11210.1016/j.bbr.2013.06.012**
- **BTCure: TRENKMANN, M et al. (2013) Tumor Necrosis Factor alpha-Induced MicroRNA-18a Activates Rheumatoid Arthritis Synovial Fibroblasts Through a Feedback Loop in NF-kappa B Signaling, Arthritis and Rheumatism, 65: 916-92710.1002/art.37834**

# ANNEX 4: BIBLIOMETRICS AND CITATION ANALYSIS

Bibliometrics are about publications and their citations. The academic field emerged from ‘information science’ and now usually refers to the methods used to study and index texts and information.

Publications cite other publications. These citation links grow into networks, and their numbers are likely to be related to the significance or impact of the publication. The meaning of the publication is determined from keywords and content. Citation analysis and content analysis have therefore become a common part of bibliometric methodology. Historically, bibliometric methods were used to trace relationships amongst academic journal citations. Now, bibliometrics are important in indexing research performance.

Bibliometric data have particular characteristics of which the user should be aware, and these are considered here.

Journal papers (publications, sources) report research work. Papers refer to or ‘cite’ earlier work relevant to the material being reported. New papers are cited in their turn. Papers that accumulate more citations are thought of as having greater ‘impact’, which is interpreted as significance or influence on their field. Citation counts are therefore recognised as a measure of impact, which can be used to index the excellence of the research from a particular group, institution or country.

The origins of citation analysis as a tool that could be applied to research performance can be traced to the mid-1950s, when Eugene Garfield proposed the concept of citation indexing and introduced the Science Citation Index, the Social Sciences Citation Index and the Arts & Humanities Citation Index, produced by the Institute of Scientific Information (currently the IP & Science business of Thomson Reuters).[29]

We can count citations, but they are only ‘indicators’ of impact or quality – not metrics. Most impact indicators use average citation counts from groups of papers, because some individual papers may have unusual or misleading citation profiles. These outliers are diluted in larger samples.

## DATA SOURCE

The data we use come from the Thomson Reuters *Web of Science*™ databases which give access not only to journals but also to conference proceedings, books, patents, websites, and chemical structures, compounds and reactions. 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* focuses on research published in journals, conferences and books in science, medicine, arts, humanities and social sciences.

The *Web of Science* was originally created as an awareness and information retrieval tool but it has acquired an important primary use as a tool for research evaluation, using citation analysis and bibliometrics. Data coverage is both current and retrospective in the sciences, social sciences, arts and humanities, in some cases back to 1900. Within the research community this data source was previously referred to by the acronym ‘ISI’.

Unlike other databases, the *Web of Science* and underlying databases are selective, that is: the journals abstracted are selected using rigorous editorial and quality criteria. The authoritative, multidisciplinary content covers over 12,000 of the highest impact journals worldwide, including Open Access journals, and over 150,000 conference proceedings. The abstracted journals encompass the majority of significant, frequently cited scientific reports and, more importantly, an even greater proportion of the scientific research output which is cited. This selective process ensures that the citation counts remain relatively stable in given research fields and do not fluctuate unduly from year to year, which increases the usability of such data for performance evaluation.

[29] Garfield, E (1955) Citation Indexes for Science – New dimension in documentation through association of ideas. *Science*: **122**, 108-111.

Thomson Reuters has extensive experience with databases on research inputs, activity and outputs and has developed innovative analytical approaches for benchmarking and interpreting international, national and institutional research impact.

## DATABASE CATEGORIES

The source data can be grouped in various classification systems. Most of these are based on groups of journals that have a relatively high cross-citation linkage and naturally cluster together. Custom classifications use subject maps in third-party data such as the OECD categories set out in the Frascati manual.

Thomson Reuters frequently uses the broader field categories in the *Essential Science Indicators* system and the finer journal categories in the *Web of Science*. There are 22 fields in *Essential Science Indicators* and 254 fields in *Web of Science*. In either case, our bibliometric analyses draw on the full range of data available in the underlying database, so analyses in our reports will differ slightly from anything created 'on the fly' from data in the web interface.

The lists of journal categories in these systems are attached at the end of this document.

Most analyses start with an overall view across the data, then move to a view across broad categories and only then focus in at a finer level in the areas of greatest interest to policy, programme or organisational purpose.

## ASSIGNING PAPERS TO ADDRESSES

A paper is assigned to each country and each organisation whose address appears at least once for any author on that paper. One paper counts once and only once for each assignment, however many address variants occur for the country or organisation. No weighting is applied.

For example, a paper has five authors, thus:

| Author | Organisation | Country | | |
|---|---|---|---|---|
| Gurney, KA | Univ Leeds | UK | Counts for Univ Leeds | Counts for UK |
| Adams, J | Univ Leeds | UK | No gain for Univ Leeds | No gain for UK |
| Kochalko, D | Univ C San Diego | USA | Counts for UCSD | Counts for USA |
| Munshi, S | Gujarat Univ | India | Counts for Gujarat Univ | Counts for India |
| Pendlebury, D | Univ Oregon | USA | Counts for Univ Oregon | No gain for USA |

So this one paper with five authors would be included once in the tallies for each of four universities and once in the tallies for each of three countries.

Work carried out within Thomson Reuters, and research published elsewhere, indicates that fractional weighting based on the balance of authors by organisation and country makes little difference to the conclusions of an analysis at an aggregate level. Such fractional analysis can introduce unforeseen errors in the attempt to create a detailed but uncertain assignment. Partitioning credit would make a greater difference at a detailed, group level but the analysis can then be manually validated.

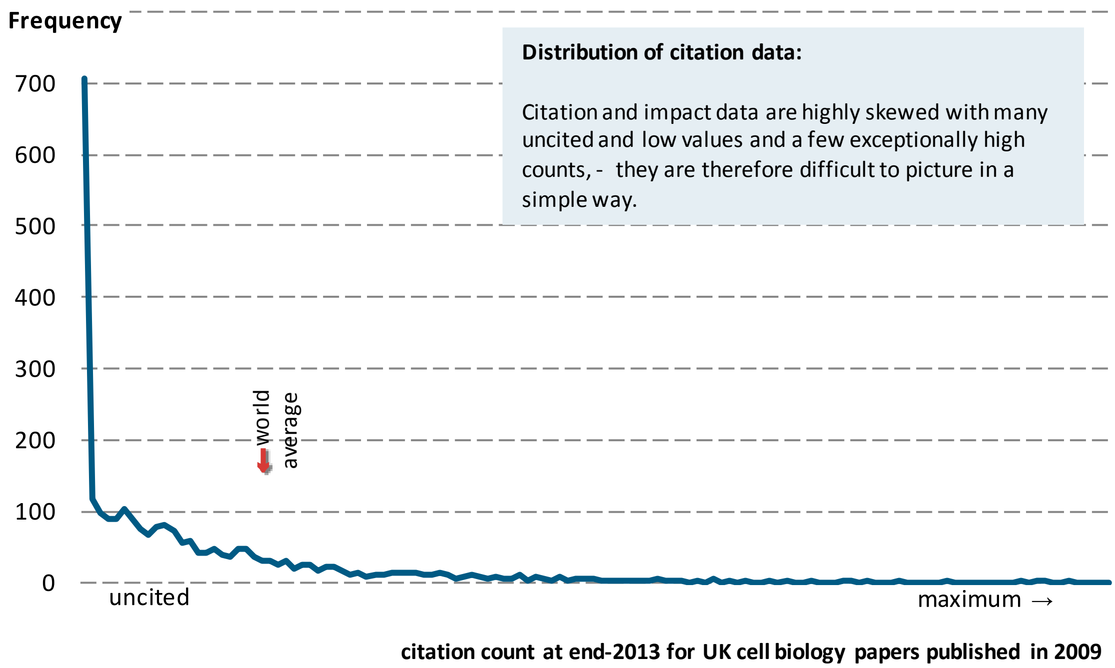
## CITATION COUNTS

A publication accumulates citation counts when it is referred to by more recent publications. Some papers get cited frequently and many get cited rarely or never, so the distribution of citations is highly skewed.

Why are many papers never cited? Certainly some papers remain uncited because their content is of little or no impact, but that is not the only reason. It might be because they have been published in a journal not read by researchers to whom the paper might be interesting. It might be that they represent important but 'negative' work reporting a blind alley to be avoided by others. The publication may be a commentary in an editorial, rather than a normal journal article and thus of general rather than research interest. Or it might be that the work is a 'sleeping beauty' that has yet to be recognised for its significance.

Other papers can be very highly cited: hundreds, even thousands of times. Again, there are multiple reasons for this. Most frequently cited work is being recognised for its innovative significance and impact on the research field of which it speaks. Impact here is a good reflection of quality: it is an indicator of excellence. But there are other papers which are frequently cited because their significance is slightly different: they describe key methodology; they are a thoughtful and wide-ranging review of a field; or they represent contentious views which others seek to refute.

Citation analysis cannot make value judgments about why an article is uncited nor about why it is highly cited. The analysis can only report the citation impact that the publication has achieved. We normally assume, based on many other studies linking bibliometric and peer judgments, that high citation counts correlate on average with the quality of the research.




The figure shows the skewed distribution of more or less frequently cited papers from a sample of UK authored publications in cell biology. The skew in the distribution varies from field to field. It is to compensate for such factors that actual citation counts must be normalised, or rebased, against a world baseline.

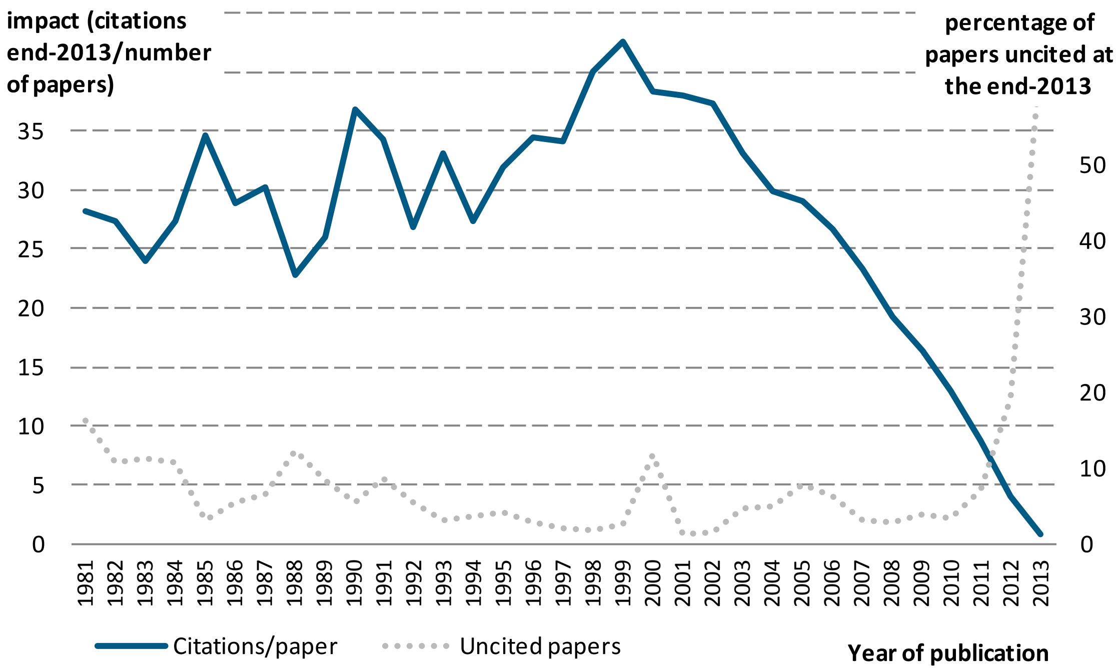
We do not seek to account separately for the effect of self-citation. If the citation count is significantly affected by self-citation then the paper is likely to have been infrequently cited. This is therefore only of consequence for low impact activity. Studies show that for large samples at national and organisational level the effect of self-citation has little or no effect on the analytical outcomes and would not alter interpretation of the results.

## TIME FACTORS

Citations accumulate over time. Older papers therefore have, on average, more citations than more recent work. The graph below shows the pattern of citation accumulation for a set of 33 journals in the journal category ***Materials Science, Biomaterials***. Papers less than eight years old are, on average, still accumulating additional citations. The citation count goes on to reach a plateau for older sources.

The graph shows that the percentage of papers that have never been cited drops over about five years. Beyond five years, between 5% and 10% or more of papers remain uncited.

Account must be taken of these time factors in comparing current research with historical patterns. For these reasons, it is sometimes more appropriate to use a fixed five-year window of papers and citations to compare two periods than to look at the longer term profile of citations and of uncitedness for a recent year and an historical year.




## DISCIPLINE FACTORS

Citation rates vary between disciplines and fields. For the UK science base as a whole, ten years produces a general plateau beyond which few additional citations would be expected. On the whole, citations accumulate more rapidly and plateau at a higher level in biological sciences than physical sciences, and natural sciences generally cite at a higher rate than social sciences.

Papers are assigned to disciplines (journal categories or research fields) by Thomson Reuters, bringing cognate research areas together. The journal category classification scheme has been recently revised and updated. Before 2007, journals were assigned to the older, well established Current Contents categories which were informed by extensive work by Thomson and with the research community since the early 1960s. This scheme has been superseded by the 252 *Web of Science*℠ journal categories which allow for greater disaggregation for the growing volume of research which is published and abstracted.

Papers are allocated according to the journal in which the paper is published. Some journals may be considered to be part of the publication record for more than one research field. As the example below illustrates, the journal *Acta Biomaterialia* is assigned to two journal categories: ***Materials Science, Biomaterials*** and ***Engineering, Biomedical***.

Very few papers are not assigned to any research field and as such will not be included in specific analyses using normalised citation impact data. The journals included in the Thomson Reuters databases and how they are selected are detailed here http://scientific.thomsonreuters.com/mjl/.

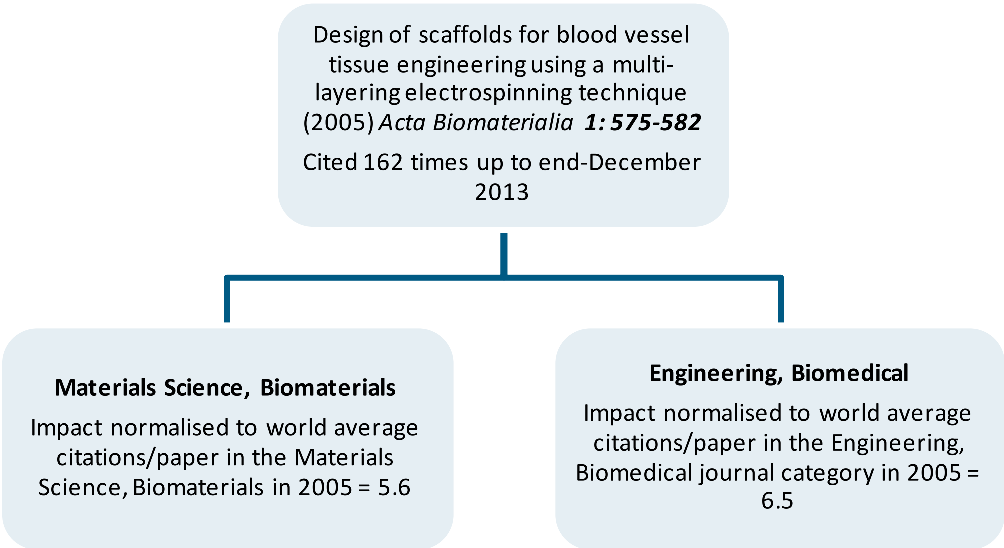
Some journals with a very diverse content, including the prestigious journals *Nature* and *Science* were classified as ***Multidisciplinary*** in databases created prior to 2007. The papers from these ***Multidisciplinary*** journals are now re-assigned to more specific research fields using an algorithm based on the research area(s) of the references cited by the article.

## NORMALISED CITATION IMPACT

Because citations accumulate over time at a rate that is dependent upon the field of research, all analyses must take both field and year into account. In other words, because the absolute citation count for a specific article is influenced by its field and by the year it was published, we can only make comparisons of indexed data after normalising with reference to these two variables.

We only use citation counts for reviews and articles in calculations of impact, because document type influences the citation count. For example, a review will often be cited more frequently than an article in the same field, but editorials and meeting abstracts are rarely cited and citation rates for conference proceedings are extremely variable. The most common normalisation factors are the average citations per paper for (1) the year and (2) either the field or the journal in which the paper was published. This normalisation is also referred to as 'rebasing' the citation count.

Impact is therefore most commonly analysed in terms of 'normalised impact', or NCI. The following schematic illustrates how the normalised citation impact is calculated at paper level and journal category level.




This article in the journal *Acta Biomaterialia* is assigned to two journal categories: ***Materials Science, Biomaterials*** and ***Engineering, Biomedical***. The world average baselines for, as an example, ***Materials science, Biomaterials*** are calculated by summing the citations to all the articles and reviews published worldwide in the journal *Acta Biomaterialia* and the other 32 journals assigned to this category for each year, and dividing this by the total number of articles and reviews published in the journal category. This gives the category-specific normalised citation impact (in the above example the category-specific $NCI_F$ for ***Materials Science, Biomaterials*** is 5.6 and the category-specific $NCI_F$ for ***Engineering, Biomedical*** is higher at 6.5). Most papers (nearly two-thirds) are assigned to a single journal category whilst a minority are assigned to more than 5.

Citation data provided by Thomson Reuters are assigned on an annual census date referred to as the Article Time Period. For the majority of publications the Article Time Period is the same as the year of publication, but for a few publications (especially those published at the end of the calendar year in less main-stream journals) the Article Time Period may vary from the actual year of publication.

World average impact data are sourced from the Thomson Reuters National Science Indicators baseline data for 2013.

## MEAN NORMALISED CITATION IMPACT

Research performance has historically been indexed by using average citation impact, usually compared to a world average that accounts for time and discipline. As noted, however, the distribution of citations amongst papers is highly skewed because many papers are never cited while a few papers accumulate very large citation counts. That means that an average may be misleading if assumptions are made about the distribution of the underlying data.

In fact, almost all research activity metrics are skewed: for research income, PhD numbers and publications there are many low activity values and a few exceptionally high values. In reality, therefore, the skewed distribution means that average impact tends to be greater than and often significantly different from either the median or mode in the distribution. This should be borne in mind when reviewing analytical outcomes.

The average (normalised) citation impact can be calculated at an individual paper level where it can be associated with more than one journal category. It can also be calculated for a set of papers at any level from a single country to an individual researcher's output. In the example above, the average citation impact of the *Acta Biomaterialia* paper can be expressed as ((5.6 + 6.5)/2) = 6.1.

## IMPACT PROFILES®

We have developed a bibliometric methodology[30] that shows the proportion of papers that are uncited and the proportion that lie in each of eight categories of relative citation rates, normalised (rebased) to world average. An Impact Profile® enables an examination and analysis of the strengths and weaknesses of published outputs relative to world average and relative to a reference profile. This provides much more information about the basis and structure of research performance than conventionally reported averages in citation indices.

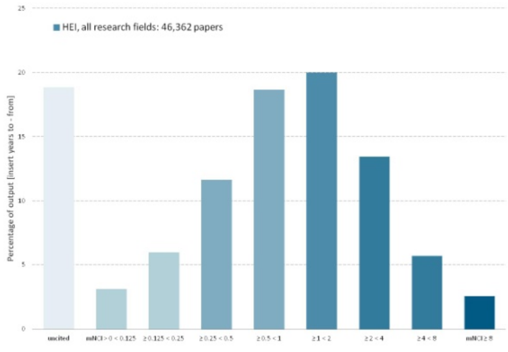
Papers which are "highly-cited" are often defined in our reports as those with an average citation impact ($NCI_F$) greater than or equal to 4.0, i.e. those papers which have received greater than or equal to four times the world average number of citations for papers in that subject published in that year. This differs from Thomson Reuters database of global highly-cited papers, which are the top 1% most frequently cited for their field and year. The top percentile is a powerful indicator of leading performance but is too stringent a threshold for most management analyses.

The proportion of uncited papers in a dataset can be compared to the benchmark for the UK, the USA or any other country. Overall, in a typical ten-year sample, around one-quarter of papers have not been cited within the 10-year period; the majority of these are, of course, those that are most recently published.

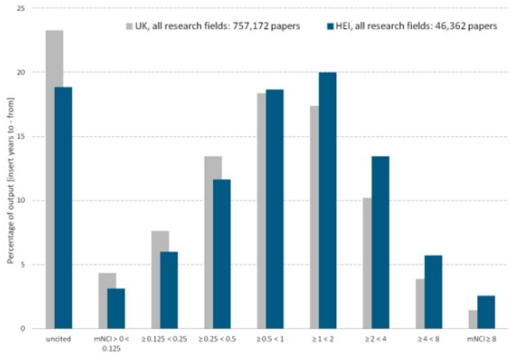
[30] Adams J, Gurney K & Marshall S (2007) Profiling citation impact: A new methodology. *Scientometrics* **72**: 325-344.

The Impact Profile® histogram can be presented in a number of ways which are illustrated below.

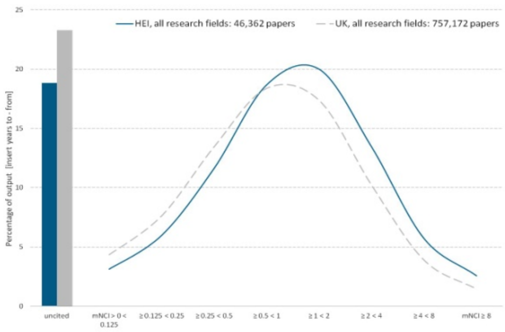
**A**



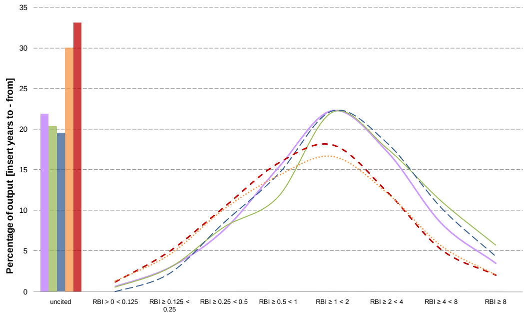

**B**




**C**




**D**




**A**: is used to represent the total output of an individual country, institution or researcher with no benchmark data. Visually it highlights the numbers of uncited papers (weaknesses) and highly cited papers (strengths).

**B** & **C**: are used to represent the total output of an individual country, institution or researcher (**client**) against an appropriate benchmark dataset (**benchmark**). The data are displayed as either histograms (B) or a combination of histogram and profile (C). Version C prevents the ‘travel’ which occurs in histograms where the eye is drawn to the data most offset to the right, but can be less easy to interpret as categorical data.

**D**: illustrates the complexity of data which can be displayed using an Impact Profile®. These data show research output in defined journal categories against appropriate benchmarks: **client, research field X**; **client, research field Y**; **client, research field Z**; **benchmark, research field X+Y**; **benchmark, research field, Z**.

Impact Profiles® enable an examination and analysis of the balance of published outputs relative to world average and relative to a reference profile. This provides much more information about the basis and structure of research performance than conventionally reported averages in citation indices.

An Impact Profile® shows what proportion of papers are uncited and what proportion are in each of eight categories of relative citation rates, normalised to world average (which becomes 1.0 in this graph). Normalised citation rates above 1.0 indicate papers cited more often than world average for the field in which that journal is categorised and in their year of publication.

Attention should be paid to:

- The proportion of uncited papers on the left of the chart
- The proportion of cited papers either side of world average (1.0)

- The location of the most common (modal) group near the centre
- The proportion of papers in the most highly-cited categories to the right, (≥4 x world, ≥8 x world).

## WHAT ARE UNCITED PAPERS?

It may be a surprise that some journal papers are never subsequently cited after publication, even by their authors. This accounts for about half the total global output for a typical, recent 10-year period. We cannot tell why papers are not cited. It is likely that a significant proportion of papers remain uncited because they are reporting negative results which are an essential matter of record in their field but make the content less likely to be referenced in other papers. Inevitably, other papers are uncited because their content is trivial or marginal to the mainstream. However, it should not be assumed that this is the case for all such papers.

There is variation in non-citation between countries and between fields. For example, relatively more engineering papers tend to remain uncited than papers in other sciences, indicative of a disciplinary factor but not a quality factor. While there is also an obvious increase in the likelihood of citation over time, most papers that are going to be cited will be cited within a few years of publication.

## WHAT IS THE THRESHOLD FOR 'HIGHLY CITED'?

Thomson Reuters has traditionally used the term 'Highly Cited Paper' to refer to the world's 1% of most frequently cited papers, taking into account year of publication and field. In rough terms, UK papers cited more than eight times as often as relevant world average would fall into the Thomson Highly Cited category. About 1-2% of papers (all papers, cited or uncited) typically pass this hurdle. Such a threshold certainly delimits exceptional papers for international comparisons but, in practice, is an onerous marker for more general management purposes.

After reviewing the outcomes of a number of analyses, we have chosen a more relaxed definition for our descriptive and analytical work. We deem papers that are in the world's top 10% of most frequently cited papers, taking into account year of publication and field, to be relatively highly-cited for national comparisons.

## A4.11 *WEB OF SCIENCE* JOURNAL CATEGORIES

Acoustics
Agricultural economics & policy
Agricultural engineering
Agriculture, dairy & animal science
Agriculture, multidisciplinary
Agriculture, soil science
Agronomy
Allergy
Anatomy & morphology
Andrology
Anaesthesiology
Anthropology
Applied linguistics
Archaeology
Architecture
Area studies
Art
Asian studies
Astronomy & astrophysics
Automation & control systems
Behavioural sciences
Biochemical research methods
Biochemistry & molecular biology
Biodiversity conservation
Biology
Biology, miscellaneous
Biophysics
Biotechnology & applied microbiology
Business
Business, finance
Cardiac & cardiovascular systems
Cell biology
Chemistry, analytical
Chemistry, applied
Chemistry, inorganic & nuclear
Chemistry, medicinal
Chemistry, multidisciplinary
Chemistry, organic
Chemistry, physical
International relations
Language & linguistics
Language & linguistics theory
Law
Limnology

Classics
Clinical neurology
Communication
Computer science, artificial intelligence
Computer science, cybernetics
Computer science, hardware & architecture
Computer science, information systems
Computer science, interdisciplinary applications
Computer science, software engineering
Computer science, theory & methods
Construction & building technology
Criminology & penology
Critical care medicine
Crystallography
Dance
Demography
Dentistry, oral surgery & medicine
Dermatology
Developmental biology
Ecology
Economics
Education & educational research
Education, scientific disciplines
Education, special
Electrochemistry
Emergency medicine
Endocrinology & metabolism
Energy & fuels
Engineering, aerospace
Engineering, biomedical
Engineering, chemical
Engineering, civil
Engineering, electrical & electronic
Engineering, environmental
Engineering, geological
Engineering, industrial
Engineering, manufacturing
Engineering, marine
Engineering, mechanical
Mining & mineral processing
Multidisciplinary sciences
Music
Mycology
Nanoscience & nanotechnology

Engineering, multidisciplinary
Engineering, ocean
Engineering, petroleum
Entomology
Environmental sciences
Environmental studies
Ergonomics
Ethics
Ethnic studies
Evolutionary biology
Family studies
Film, radio, television
Fisheries
Folklore
Food science & technology
Forestry
Gastroenterology & hepatology
Genetics & heredity
Geochemistry & geophysics
Geography
Geography, physical
Geology
Geosciences, multidisciplinary
Geriatrics & gerontology
Health care sciences & services
Health policy & services
Haematology
History
History & philosophy of science
History of social sciences
Horticulture
Humanities, multidisciplinary
Imaging science & photographic technology
Immunology
Industrial relations & labour
Infectious diseases
Information & library science
Instruments & instrumentation
Integrative & complementary medicine
Psychology
Psychology, applied
Psychology, biological
Psychology, clinical
Psychology, developmental

Linguistics
Literary reviews
Literary theory & criticism
Literature
Literature, African, Australian, Canadian
Literature, American
Literature, British Isles
Literature, German, Dutch, Scandinavian
Literature, romance
Literature, Slavic
Management
Marine & freshwater biology
Materials science, biomaterials
Materials science, ceramics
Materials science, characterization & testing
Materials science, coatings & films
Materials science, composites
Materials science, multidisciplinary
Materials science, paper & wood
Materials science, textiles
Math & computational biology
Mathematics
Mathematics, applied
Mathematics, interdisciplinary applications
Mechanics
Medical ethics
Medical informatics
Medical laboratory technology
Medicine, general & internal
Medicine, legal
Medicine, research & experimental
Medieval & renaissance studies
Metallurgy & metallurgical engineering
Meteorology & atmospheric sci
Microbiology
Microscopy
Mineralogy
Urban studies
Urology & nephrology
Veterinary
Veterinary sciences
Virology
Water resources
Women's studies
Zoology

Neuroimaging
Neurosciences

Nuclear science & technology
Nursing
Nutrition & dietetics
Obstetrics & gynaecology
Oceanography
Oncology
Operations research & management science
Ophthalmology
Optics
Ornithology
Orthopaedics
Otorhinolaryngology
Palaeontology
Parasitology
Pathology
Paediatrics
Peripheral vascular disease
Pharmacology & pharmacy
Philosophy
Physics, applied
Physics, atomic, molecular & chemical
Physics, condensed matter
Physics, fluids & plasmas
Physics, mathematical
Physics, multidisciplinary
Physics, nuclear
Physics, particles & fields
Physiology
Planning & development
Plant sciences
Poetry
Political science
Polymer science
Psychiatry

Psychology, educational
Psychology, experimental
Psychology, mathematical
Psychology, multidisciplinary
Psychology, psychoanalysis
Psychology, social
Public administration
Public, environmental & occupational health
Radiology, nuclear medicine & medical imaging
Rehabilitation
Religion
Remote sensing
Reproductive biology
Respiratory system
Rheumatology
Robotics
Social issues
Social sciences, biomedical
Social sciences, interdisciplinary
Social sciences, mathematical methods
Social work
Sociology
Soil science
Spectroscopy
Sport sciences
Statistics & probability
Substance abuse
Surgery
Telecommunications
Theatre
Thermodynamics
Toxicology
Transplantation
Transportation
Transportation science & technology
Tropical medicine