



Compendium of Bibliometric Science Indicators



NOTE FROM THE SECRETARIAT

This document contains the final version of the OECD *Compendium of Bibliometric Science Indicators*. The report brings together a new collection of statistics depicting recent trends and the structure of scientific production across OECD countries and other major economies that supports indicators contained in the 2015 *OECD Science, Technology and Industry Scoreboard*.

This report was prepared in partnership between the OECD Directorate for Science, Technology and Innovation ([DSTI](#)) and the [SCImago Research Group \(CSIC, Spain\)](#).

It was presented to the Committee for Scientific and Technological Policy (CSTP) and National Experts in Science and Technology Indicators (NESTI) delegates for comment and approval.

This paper was approved and declassified by written procedure by the Committee for Scientific and Technological Policy (CSTP) in May 2016 and prepared for publication by the OECD Secretariat.

Note to Delegations:

*This document is also available on OLIS under the reference code:
DSTI/EAS/STP/NESTI(2016)8/FINAL*

This document and any map included herein are without prejudice to the status of or sovereignty over any territory, to the delimitation of international frontiers and boundaries and to the name of any territory, city or area.

The statistical data for Israel are supplied by and under the responsibility of the relevant Israeli authorities or third party. The use of such data by the OECD is without prejudice to the status of the Golan Heights, East Jerusalem and Israeli settlements in the West Bank under the terms of international law.

Please cite this publication as:

OECD and SCImago Research Group (CSIC) (2016), *Compendium of Bibliometric Science Indicators*. OECD, Paris. Accessed from <http://oe.cd/scientometrics>.

Cover image: © <https://lostandtaken.com/>

FOREWORD

The *Compendium of Bibliometric Science Indicators* report brings together a new collection of statistics depicting recent trends and the structure of scientific production across OECD countries and other major economies. This report was prepared in partnership between the OECD Directorate for Science, Technology and Innovation ([DSTI](#)) and the [SCImago Research Group \(CSIC, Spain\)](#). The work on the report started in 2014 and finished in 2015, the data cover the 2003-12 period.

This report builds on the approach used by the OECD Working Party of National Experts on Science and Technology Indicators (NESTI) to work with official national experts and highly reputed research organisations to review, extend and improve the scope, relevance and international comparability of statistics on science, technology and innovation.

OECD and SCImago have collaborated in recent years to contribute to the Programme of Work of the OECD Committee for Scientific and Technological Policy and provide regular updates on statistical information on the structure and key trends applying to scientific activities in the *OECD Science, Technology and Industry Scoreboard*, in its 2011, 2013 and 2015 editions. This document is the source of a significant part of the bibliometric indicators reported in the *Scoreboard's* [2015 edition](#) and its [accompanying brief](#) on science and research today.

With this report, the OECD continues its exploration of the rich potential of bibliometric indicators and how these indicators can help advance today's science and technology policy agenda. The aim is not only to highlight the opportunities offered by the indicators that can be elaborated using bibliometric sources but to also promote a constructive discussion on the interpretation of the indicators, the biases inherent in the data and the advantages and shortcomings of alternative calculation methods. The results should provide a valuable source of information on bibliometric indicators for potential users of these data. Unlike other OECD statistics on science, technology and innovation, this compendium is not intended as a reference set of guidelines. It is also by no means a comprehensive exploration of the full potential of bibliometric data, a field of research in rapid expansion due to the increased availability of data sources, the potential to link to other databases, and improvements in computational tools to analyse the data.

This publication is informed by, but does not attempt to discuss the use that is made of bibliometric information at the level of individuals and institutions for performance management, grant allocation, personal promotion and other decisions. Its objective is to present a number of insights on the scientific performance of countries and what underpins observed aggregate patterns. As a statistical publication, it is worth noting that practical decisions made by funding organisations, administrators and several others on the basis of bibliometric information at the level of individuals and teams can have significant behavioural effects. This can in turn have an impact on how bibliometric indicators corresponding to countries, sectors or research domains can be interpreted.

It is also important to note that these indicators are based on data that are collected principally for commercial purposes by a private data provider drawing on a complex and evolving system of

scientific peer review and publishing. In this publication, the underlying data that support the production of indicators are compiled by Elsevier in the Scopus® database and curated by the SCImago Research Group. This does not represent an endorsement of a particular source of bibliometric indicators relative to other providers. The indicators presented here were calculated by the SCImago Research Group following a collaborative indicator selection and design with the OECD.

From the SCImago Research Group side, the work was led by Félix de Moya-Anegón and coordinated by Carmen López-Illescas with contributions from Zaida Chinchilla-Rodríguez and Elena Corera-Álvarez.

At the OECD, Brigitte van Beuzekom and Fernando Galindo-Rueda from DSTI's Economic Analysis and Statistics Division contributed to this report, with comments by Alessandra Colecchia. Brunella Boselli assisted in the preparation of the web version interactive charts available at <http://oe.cd/scientometrics>. The indicators in this report are available by clicking on the links immediately below each figure.

Comments were provided by delegates from the OECD Committee for Scientific and Technological Policy (CSTP) and its Working Party of National Experts on Science and Technology Indicators (NESTI). The CSTP was invited to declassify the document under the written procedure. This was completed in May 2016.

The authors have made every effort to ensure the accuracy and completeness of the information contained in this report. Any errors or omission are the authors' sole responsibility. Readers are invited to submit any comments or suggestions to brigitte.vanbeuzekom@oecd.org.

TABLE OF CONTENTS

FOREWORD	3
TABLE OF CONTENTS.....	5
READER'S GUIDE	6
Scope of the publication.....	6
Abbreviations	6
Country groupings.....	6
Scientific research field codes.....	7
Acronyms	7
EXECUTIVE SUMMARY	8
INTRODUCTION.....	10
1. SCIENTIFIC PUBLICATION OUTPUT: KEY FEATURES AND TRENDS	15
1.1. Trends in scientific production	16
1.2. Scientific publications and R&D	18
1.3. Scientific collaboration.....	20
1.4. Scientific leading authorship and collaboration.....	22
1.5. Scientific “excellence” and leading authorship	24
1.6. Trends in scientific production, top cited publications	26
1.7. The citation impact of scientific production and international collaboration	28
1.8. Scientific publishing and open access	30
Chapter notes and references	32
2. SCIENTIFIC PUBLICATION INDICATORS BY FIELDS OF SCIENCE	35
2.1. Scientific publication output and journals by fields	36
2.2. Trends in indexed scientific publication output by fields.....	38
2.3. Field specialisation by country	40
2.4. Citations by fields	42
2.5. Fields with highest citation impact by country	44
2.6. International collaboration by fields	46
2.7. Open access by fields.....	48
Chapter notes and references	50
3. SCIENTIFIC PUBLICATION INDICATORS BY SECTORS.....	51
3.1. The contribution of sectors to scientific production	52
3.2. Scientific publication output trends by sector.....	54
3.3. Scientific production by sector and field	56
3.4. Collaboration patterns by sector	58
3.5. Cross-sector collaboration	60
3.6. Scientific excellence and citation impact.....	62
3.7. High “impact” institutions	64
Chapter notes and references	66

READER'S GUIDE

Scope of the publication

In this report, the geographical coverage has been limited to the OECD countries and the BRIICS (Brazil, the Russian Federation, India, Indonesia, China and South Africa). For space availability and presentational reasons, other countries and economies have been excluded from a majority of indicators.

Abbreviations

For most of the figures, this publication uses 3-digit ISO codes for countries.

OECD

AUS	Australia	ISL	Iceland
AUT	Austria	ISR	Israel
BEL	Belgium	ITA	Italy
CAN	Canada	JPN	Japan
CHE	Switzerland	KOR	Korea
CHL	Chile	LUX	Luxembourg
CZE	Czech Republic	MEX	Mexico
DEU	Germany	NLD	Netherlands
DNK	Denmark	NOR	Norway
ESP	Spain	NZL	New Zealand
EST	Estonia	POL	Poland
FIN	Finland	PRT	Portugal
FRA	France	SVK	Slovak Republic
GBR	United Kingdom	SVN	Slovenia
GRC	Greece	SWE	Sweden
HUN	Hungary	TUR	Turkey
IRL	Ireland	USA	United States

BRIICS

BRA	Brazil	IND	India
CHN	People's Republic of China	RUS	Russian Federation
IDN	Indonesia	ZAF	South Africa

Country groupings

BRIICS	Brazil, the Russian Federation, India, Indonesia, China and South Africa
OECD	Total OECD area (34 member countries)
ROW	Rest of the world, i.e. all countries and economies, excluding OECD countries and the BRIICS
WLD	World

Scientific research field codes

The following codes were used for the fields captured in the All Science Journal Classification used in the Scopus database.

Code	Field	Code	Field
ART	Arts & Humanities	HEA	Health Professions
AGR	Agricultural & Biological Sciences	IMM	Immunology & Microbiology
BIO	Biochemistry, Genetics & Molecular Biology	MSC	Materials Science
BUS	Business, Management & Accounting	MAT	Mathematics
CHE	Chemistry	MED	Medicine
CEN	Chemical Engineering	MUL	Multidisciplinary
COM	Computer Science	NEU	Neuroscience
DEC	Decision Sciences	NUR	Nursing
DEN	Dentistry	PHA	Pharmacology, Toxicology & Pharmaceuticals
EAR	Earth & Planetary Sciences	PHY	Physics & Astronomy
ECO	Economics, Econometrics & Finance	PSY	Psychology
ENE	Energy	SOC	Social Sciences
ENG	Engineering	VET	Veterinary
ENV	Environmental Science		

Acronyms

ASJC	All Science and Journal Classification
GDP	Gross domestic product
GERD	Gross domestic expenditure on R&D
PPP	Purchasing power parity
PRO	Public research organisation
R&D	Research and development
S&T	Science and technology

EXECUTIVE SUMMARY

The *Compendium of Bibliometric Science Indicators* brings together a new collection of statistics depicting recent trends and the structure of scientific production across OECD countries and other major economies based on bibliometric data from Scopus.

The global landscape of scientific research has been shifting

The United States remains by a significant margin the largest source of scientific publications, with almost 25% of world scientific output in 2012. However, China increased its share of world publication output to 15% and has experienced an average annual growth rate of over 20% over 2003-12. Indicators of scientific excellence based on citation impact take into account differences in publication and citation norms by sector, providing the basis on which to quality-adjust simple counts of documents. Trends for the distribution of the top 10% most-cited documents among the largest producing countries are more stable. China's share of highly cited documents has grown rapidly, from 4 to 14%, compared with 6 to 17% for its share in all publications over the 2003-12 period. However, because of the low starting share of top-cited publications, under current growth rates it would take several years for China's level of top cited publications to catch up with the United States.

High impact research output is highly concentrated and reflects specialisation patterns

Four countries, the United States, the United Kingdom, Germany and China, accounted together for 50-70% of high-impact publications across all scientific disciplines. While the United States leads on all fields as expected by its size, its pre-eminent role is particularly marked in the life sciences and most health-related and social science fields. US-authored publications account for a relatively lower percentage of high-impact publications in a number of basic science domains outside the life sciences and engineering. The United Kingdom is the second largest producer of top-cited publications in the fields where the United States accounts for a large share, but also excels in the Earth, Environmental, Agricultural and Veterinary sciences. China is the second largest producer of top cited publications in Materials Science, Chemistry, Engineering, Computer Science and Chemical Engineering, Energy, Decision Sciences and Mathematics. Germany has the second largest share in Physics and Astronomy, and is the third largest producer of high-impact publications across most fields. Other top countries include Japan – a significant player in Materials Science, Chemistry and Biochemistry – Australia, Brazil, France, India and Spain.

Scientific publication output is mainly concentrated in higher education institutions, reflecting incentives

Publication growth has been mainly concentrated in higher education institutions, reflecting the increasing share of higher education R&D expenditures in total R&D but probably also incentives to publish in peer-reviewed journals featuring in major indices. Growth in government and health sectors has been significant but less marked. The government sector is relatively more important in the BRIICS than in the OECD area. Compared to BRIICS, OECD countries appear to be undertaking more scientific research within the health sector and health-related domains. After 2007, there has been virtually no growth in documents attributable to authors in the private sector (principally business). Scientific output within this sector is most likely to be under-represented in bibliometric indicators.

High-impact institutions are located in few countries

The United States and the United Kingdom host the higher education institutions with the largest number of highly-cited publications. The list of high impact government institutions is the most diverse by country location, with 12 different countries featured among the top 20 institutions and 19 among the top 50. A number of countries such as France, China, Germany, Spain and Italy feature

within this list with large central and multi-subject research institutions with presence in different locations within their countries. Some of these institutions account for more publications than the largest US-based universities. To some extent, these represent scale effects, as many of these government institutions have excellence rates that do not fare significantly above the world average.

Collaboration goes hand in hand with scientific excellence

In the global landscape of scientific research, collaboration between institutions in different countries intensified over the 2003-12 period. Increased collaboration appears to have been one of the major factors driving the surge in publication activity over the period. Data suggest a positive relationship between measures of scientific research collaboration and impact, the latter proxied by the average normalised citation index.

Collaborating with leading foreign authors abroad is a key path towards scientific excellence

For a majority of countries, approximately 20% of their domestic publications have been led by authors affiliated to institutions in other countries. Collaboration allows countries to be part of partnerships led by authors abroad, as implied by the identity of the corresponding author across documents. The United States account for a disproportionate share of leading authors and has the highest rate of domestic scientific excellence, followed by the Netherlands.

There appears to be further scope for collaboration in some sectors and research domains

International collaboration is significantly more common in Earth sciences, physics and mathematics. In contrast, collaboration is less frequent in health-related sciences and in the arts. Cross sector collaboration patterns for higher education reflect the relative importance of other sectors in domestic scientific production. Within the combined group of OECD and BRIICS countries, when authors in the higher education sector collaborate with authors in different sectors they do so mostly with authors within government sector, from 32% for the United States to 90% for the Russian Federation.

Growth in publications is principally driven by the indexing of new journals

Growth in scientific production has been principally accounted for by the indexation of new journals, explaining 60% of the rise in the number of documents. New journals often appear as means to disseminate findings in emerging or converging scientific domains, or to bring together evidence of relevance to particular application areas. It is however not possible to tell how many of these journals already existed before their new or possibly older content was indexed.

There has been a rise in multi-disciplinarity at the level of journals

Medicine continues to be the field accounting for most publications in most countries. There is a higher than average number of documents per journal in the natural sciences relative to the social sciences, where knowledge appears to be more fragmented into relatively more journals. Interestingly, the fastest growth in the number of indexed publications has been found in journals specialised in multidisciplinary research, social and health sciences, partly reflecting increased efforts to index journals in these domains. Multi and inter-disciplinarity are however best analysed at the level of the contents of individual documents rather than the journals in which they are published.

Open access journals represent a minority of journals, though less so in emerging economies

Publishing in open access (OA) journals is one approach to ensure open access to scientific research documents. In Scopus, 2 800 titles, about 12% of the 22 283 active journals, were identified as open access (OA) journals. For the OECD countries the average share of journals listed as OA is 8%. For the BRIICS countries the OA share is much higher at 25%. OA publishing appears to be more common in the life sciences and least common for engineering and economics-related domains.

INTRODUCTION

Background

Bibliometrics has been defined as “the application of statistical and mathematical methods to books and other media of communication” (Pritchard, 1969). The scientific peer review system and the body of scholarly publications it generates help provide the basis on which bibliometrics can be applied to the study of scientific research. Quantitative studies on research publications can draw on the contents and information contained in publications – that meet defined review and publication criteria – to analyse, under some assumptions, multiple dimensions of scientific production and dissemination.

For example, indexed information on scientific documents helps investigate the sources of science through the identity and affiliation of authors and the references contained in a document. Scientific collaboration can be potentially gauged by the extent of co-authorship and/or the engagement of multiple institutions. The relevance of the research to the wider scientific community may be inferred from the extent a publication is cited by other documents or the visibility of the title, e.g. the journal, in which a given document is published, based on its past citation record.

These are examples that reveal both the power as well as the limitations of science bibliometric analysis. The interpretation of bibliometric analysis is contingent on a series of norms and incentives that vary across sectors and knowledge domains. It can also evolve over time, for example, it is widely known that not all scientific discoveries and research results are published in a well-defined list of international scientific journals where they can be read and cited by other researchers and a possibly wider user community. As a matter of fact, the discipline was itself borne out of the necessity to inform decisions on which journals research libraries should physically stock with limited financial and space resources. Like other forms of administrative or commercial sources, bibliometric data do not exist to serve statistical purposes. While this does not invalidate their relevance for the statistical analysis of science, it is important to remember this key feature and apply a degree of caution when designing and interpreting bibliometric indicators.

The *Compendium of Bibliometric Science Indicators* brings together a new collection of statistics depicting recent trends and the structure of scientific production across OECD countries and other major economies. This report was prepared in partnership between the OECD Directorate for Science, Technology and Innovation ([DSTI](#)) and the [SCImago Research Group \(CSIC, Spain\)](#).

SCImago is a research group from the Consejo Superior de Investigaciones Científicas (CSIC), University of Granada, Extremadura, Carlos III (Madrid) and Alcalá de Henares, dedicated to information analysis, representation and retrieval by means of visualisation techniques. Scimago has worked closely with international partners from the publishing realm, scientific information suppliers, universities and government agencies. Through its collaboration with Elsevier, with whom they develop analysis tools, SCImago produces the SCImago Journal & Country Rank and the SIR Reports which analyses institution research performance across the world.

This report builds on the approach used by the OECD Working Party of National Experts on Science and Technology Indicators (NESTI), to work with official national experts and highly reputed research organisations to extend and improve the scope, relevance and international comparability of statistics on science and technology. Scientometric indicators complement and contribute to OECD efforts to standardise, collect, report and analyse a wide range of science, technology and innovation activities by providing evidence on a selected set of S&T outcomes. Bibliometric indicators

complement and contribute to OECD efforts to standardise, collect, report and analyse a wide range of science, technology and innovation activities by providing evidence on a selected set of S&T outcomes.

OECD and SCImago have collaborated in recent years to contribute to the Programme of Work of the OECD Committee for Scientific and Technological Policy and provide regular updates on statistical information on the structure and key trends applying to scientific activities in the *OECD Science, Technology and Industry Scoreboard*, in its 2011, 2013 and 2015 editions and the *OECD Science, Technology and Industry Outlook* (OECD, 2010, 2012, 2014).

Through this collaborative report with Scimago, the OECD continues its exploration of the rich potential of bibliometric indicators and how these indicators can help advance today's science and technology policy agenda. The aim is not only to highlight the opportunities offered by the indicators that can be elaborated using bibliometric sources but to also promote a constructive and open discussion on the interpretation of the indicators, the biases inherent in the data and the advantages and shortcomings of alternative calculation methods. The results should provide a valuable source of information on bibliometric indicators for potential users of these data. The indicators presented here were calculated by the SCImago Research Group following a collaborative indicator selection and design with OECD.

Unlike other OECD statistics on science, technology and innovation, this compendium is not intended as a reference set of guidelines. Although there is an OECD manual in the closely related area of patent statistics (OECD, 2009) the OECD has not developed standards in the area of bibliometrics. For further methodological guidance, the reader may wish to refer to established academic manuals such as the latest version of the handbook produced by the Karolinska Institutet (Rehn et al, 2014). OECD's position is very much that of a user of bibliometric data.

The OECD has also worked closely in the area of bibliometrics with the US National Science Foundation and especially Japan's National Institute of Science and Technology Policy, who helped OECD introduce custom-made analysis featured in previous editions of the *Science, Technology and Industry Scoreboard* (OECD, 2007 and OECD, 2009), the OECD monograph on *Measuring Innovation* (OECD, 2010) and also supported the development of analytical capabilities in this area within the OECD Directorate for Science, Technology and Innovation. This has also resulted in a number of working papers, starting with Okubo (1997), an initial introduction to bibliometrics for OECD, and Igami and Saka (2007) on science maps. Further information on related OECD work can be found at a dedicated webpage, <http://oe.cd/scientometrics>.

This publication does not attempt to discuss the use that is made of bibliometric information at the level of individuals and institutions for performance management, grant allocation, personal promotion and other decisions. Its objective is to present a number of insights on the scientific performance of countries and what underpins observed aggregate patterns. A rising number of bibliometricians, social scientists and research administrators have become increasingly concerned about what they perceive to be a pervasive misapplication of indicators to the evaluation of scientific performance, leading to the release of a list of principles (the so-called Leiden Manifesto) to help guide research evaluation (Hicks et al, 2015). While focused on the specifics of research evaluation, many of these principles have an almost direct translation into the statistical policy use domain of bibliometrics. The recognition of the importance of systemic effects of assessment and indicators is one of them.

The extent to which publication in a journal that is included in a major index may matter for recognition and career progression provides indeed a motivation for compiling publication indices that

is distinct from the interest in the data as a source for compiling aggregate statistical indicators. Generating statistical indicators on the basis of bibliometric data is a very distinct task from comparable exercises based on survey or other types of administrative data. As a statistical publication, it is worth noting here that practices that rely on bibliometric information at the micro level can have significant behavioural feedback effects and therefore impact on the interpretation of the aggregate indicators.

It is also important to note that these indicators are based on data that are collected principally for commercial purposes by private data providers drawing on a complex and evolving system of scientific peer review and publishing. In this publication, the underlying data that support the production of indicators are compiled by Elsevier in the Scopus® database and curated by the SCImago Research Group. The use of this source does not represent an endorsement of a particular source of bibliometric indicators relative to other offers. OECD has used different sources with different partners to serve different statistical and analytical needs.

The bibliometric data landscape has experienced some remarkable developments in recent years. The Science Citation Index from the Institute for Scientific Information (ISI, now part of Thomson Reuters) under the leadership of Eugene Garfield was for many years the sole information source for use by experts for specialist analyses. Accessibility improved in 2002, when Thomson Reuters launched an integrated web platform, making the Web of Science database widely accessible, and coverage was improved with the inclusion of indices for social sciences and the arts and humanities. Competing citation indices were created: Elsevier's Scopus was released in 2004 with publications from 1996 onwards. Google Scholar was released in beta version in 2004.

The field of bibliometrics has been morphing from the study of indices for improving information retrieval from peer-reviewed scientific publications onto a wider domain, known as scientometrics, defined as the “quantitative study of science, communication in science, and science policy” (Hess, 1997), which covers other types of documents and information sources relating to science and technology. These sources can include data sets, web pages and social media.

This report does not aim to provide a comprehensive exploration of the full potential of bibliometric data, a field of research in rapid expansion due to the increased availability of data sources, the potential to link to other databases, and improvements in computational tools to analyse the data. Standards have emerged to facilitate the disambiguation and tracking of individual authors (e.g. [ORCID](#)) and a better integration with CV and funding data. This type of work can improve the quality of experimental efforts to measure the mobility of scientific researchers (Appelt et al., 2015). As discussed in this publication, it remains a major challenge to identify the means by which statistics obtained in different domains (e.g. R&D statistics from surveys and bibliometric data) can be consistently linked at different levels of aggregation. The combination of necessary confidentiality requirements on survey data on one side and the reliance on proprietary data sources makes this task particularly complicated.

The development of suitable data infrastructures for the integrated analysis of science, technology and innovation (STI) is an area of major OECD interest and one that it expects that world-leading experts will discuss at its forthcoming Blue Sky Conference on the future of STI indicators. The OECD Blue Sky Fora take place every 10 years and help set a global agenda for -developing policy-relevant STI indicators in today's global economy (<http://oe.cd/blue-sky-indicators>).

Structure of the report

This report is structured into three main chapters:

Chapter 1 provides an overview of the bibliometric science data at the national level, indicating the key features and trends. Definitions and measurability boxes have been included in most sections. The focus is on understanding scientific output measures and citation-based quality adjustments.

Chapter 2 explores differences across countries by fields of science. Each journal/paper in Scopus is assigned to one or more fields using Elsevier's All Science and Journal Classification (ASJC). This chapter explores specialisation patterns and subject-specificities that help explain in part observed cross-country differences.

Chapter 3 investigates the contribution of different institutional sectors (higher education, government, health and private sector) as codified by SCImago, to measure scientific production output. It also explores the main features of each sector, including identifying the largest producers of highly cited publications.

Throughout the entire publication, information on definitions and measurability of key concepts supports the discussion of the main results in the figures. Data for all figures reported in this publication are available in spreadsheet format online through the links inserted at the bottom of each figure. Selected interactive online charts also allow for an exploration of some key indicators along country and field dimensions.

References

- Appelt, S., F. Galindo-Rueda, R. de Pinho and B. van Beuzekom (2015), "Which factors influence the international mobility of research scientists?", OECD Science, Technology and Industry Working Papers, 2015/02, OECD Publishing, Paris, <http://dx.doi.org/10.1787/5js1tmrr2233-en>.
- Garfield, E. (1979), "Is citation analysis a legitimate evaluation tool?", *Scientometrics*, Vol. 1, No. 4, pp. 359–375, <http://dx.doi.org/10.1007/BF02019306>.
- Hess, D. J (1997), *Science Studies: An advanced introduction*, New York: New York University Press.
- Hicks, D., Wouters, P., Waltman, L., de Rijcke, S. and Rafols I. (2015) "Bibliometrics: The Leiden Manifesto for research metrics", *Nature*, Volume: 520, 429–431, <http://dx.doi.org/10.1038/520429a>.
- Igami, M. and A. Saka (2007), "Capturing the evolving nature of science, the development of new scientific indicators and the mapping of science", OECD Science, Technology and Industry Working Papers 2007/1, OECD, Paris, www.oecd.org/sti/working-papers.
- OECD (2009), *OECD Patent Statistics Manual*, OECD Publishing, Paris, <http://dx.doi.org/10.1787/9789264056442-en>.
- OECD (2011), *Science, Technology and Industry Scoreboard, 2009*, OECD Publishing, Paris, http://dx.doi.org/10.1787/sti_scoreboard-2009-en.
- OECD (2011), *Science, Technology and Industry Scoreboard, 2011*, OECD Publishing, Paris, http://dx.doi.org/10.1787/sti_scoreboard-2011-en.

- OECD (2013), *OECD Science, Technology and Industry Scoreboard 2013: Innovation for Growth*, OECD Publishing, Paris, http://dx.doi.org/10.1787/sti_scoreboard-2013-en
- OECD (2014), *Science, Technology and Industry Outlook, 2014*, OECD Publishing, Paris, http://dx.doi.org/10.1787/sti_outlook-2014-en.
- OECD (2015), *Science, Technology and Industry Scoreboard, 2015*, OECD Publishing, Paris, http://dx.doi.org/10.1787/sti_scoreboard-2015-en.
- OECD (2014), “OECD-experts dialogue on scientometrics: Improving the use of bibliometric indicators and analysis for policy-making: Summary of discussions.” Accessed from www.oecd.org/sti/inno/scientometrics.htm .
- Okubo, Y. (1997), “Bibliometric Indicators and Analysis of Research Systems: Methods and Examples”, OECD Science, Technology and Industry Working Papers, 1997/01, OECD Publishing, Paris, <http://dx.doi.org/10.1787/208277770603>.
- Pritchard, A. (1969), “Statistical Bibliography or Bibliometrics”, *Journal of Documentation*, 25, 4.
- Rehn, C., Gortnizki, C., Larsson A. and Wadskog D. (2014), *Bibliometric handbook for Karolinska Institutet*, Karolinska Institutet University Library. http://kib.ki.se/sites/default/files/bibliometric_handbook_2014.pdf .
- SCImago (2007), SJR — SCImago Journal & Country Rank, Retrievable from www.scimagojr.com.

1. SCIENTIFIC PUBLICATION OUTPUT: KEY FEATURES AND TRENDS

1. Trends in scientific production
2. Scientific publications and R&D
3. Scientific collaboration
4. Scientific leading authorship and collaboration
5. Scientific excellence and leading authorship
6. Trends in scientific production, top cited publications
7. The citation impact of scientific production and international collaboration
8. Scientific publishing and open access

1.1. Trends in scientific production

Since 2003, the number of documents published worldwide in scholarly journals listed in Scopus, a major global index of scientific publications, has nearly doubled reaching 2.8 million in 2012. For most countries, this is well in excess of the growth of population over the same period. The number of publications provides a first, somewhat crude approximation, to a measure of the level of scientific production within countries that is subject to peer-review. The number of documents accounted for by authors affiliated to institutions located in OECD countries was 60% larger in 2012 than in 2003, while for the BRIICS group of countries (Brazil, the Russian Federation, India, Indonesia, China and South Africa), the volume of indexed documents nearly quadrupled. This was mainly driven by the fast growth in publications from China, with a five-fold increase over this period (**Figure 1.1**), supported by fast growth in Brazil and India too. On this metric, the combined total for the BRIICS first surpassed that of the United States in 2011 (see more data online).

The global distribution of scientific production has changed as a result of these growth patterns. If we compare the 2004-08 and 2008-12 periods (**Figure 1.2**), the share accounted for by the BRIICS increased from 16.4% to 22.2%, while the OECD share declined slightly from 80.6% to 79.5%. The United States remains by a significant margin the largest source of scientific publications, moving from 25.5% to 23.6%, followed by China which increased its share from 10.3% to 14.8%. Scientific production is however now more concentrated among the top 20 countries. The share of the rest of the world declined from 11.1% in 2004-08 to 7.5% in 2008-12. While the number of indexed publications continues to increase worldwide, there is some evidence of a slowdown in publication output growth among most countries (**Figure 1.3**).

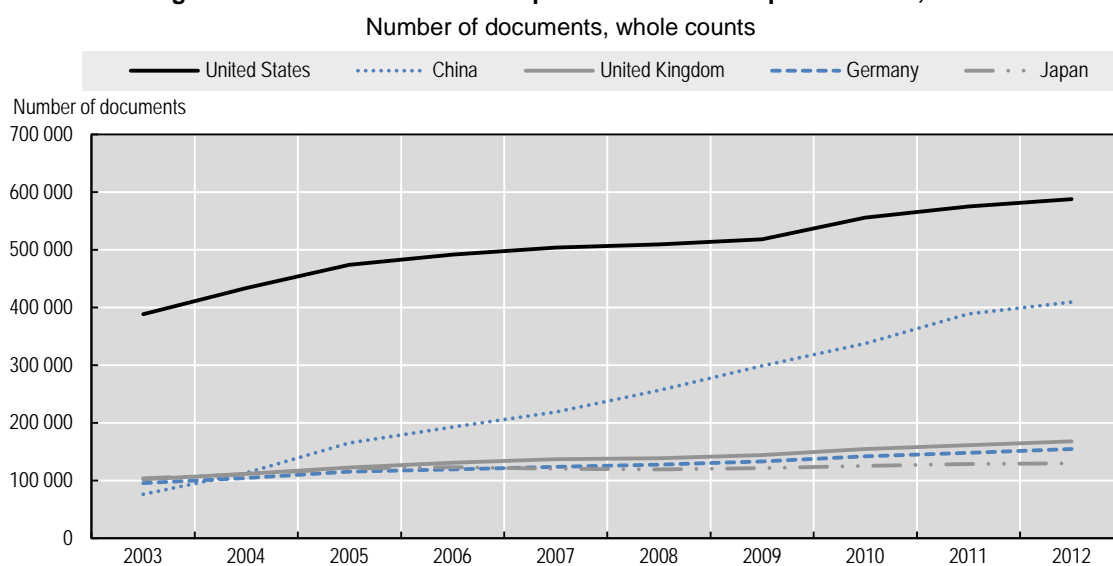
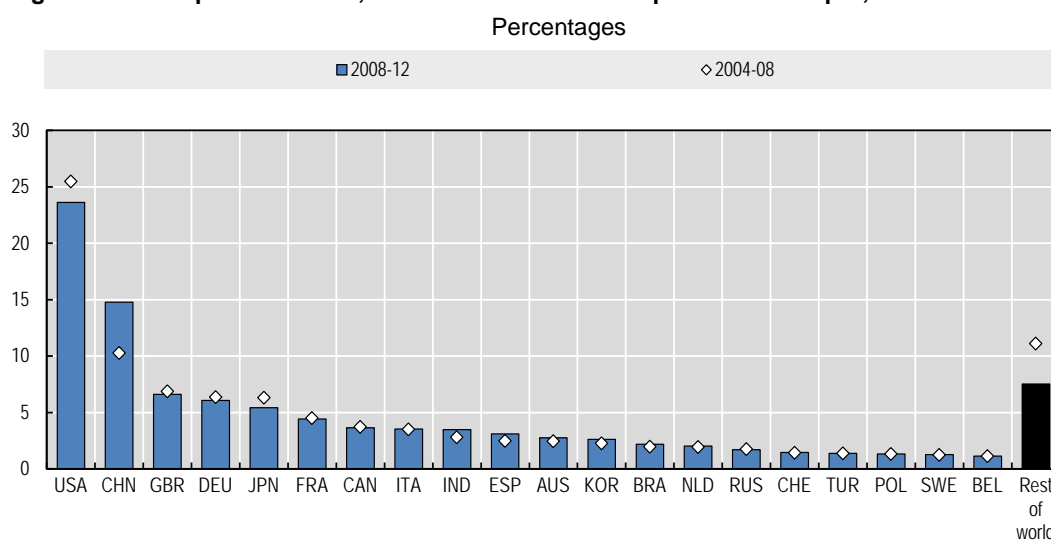
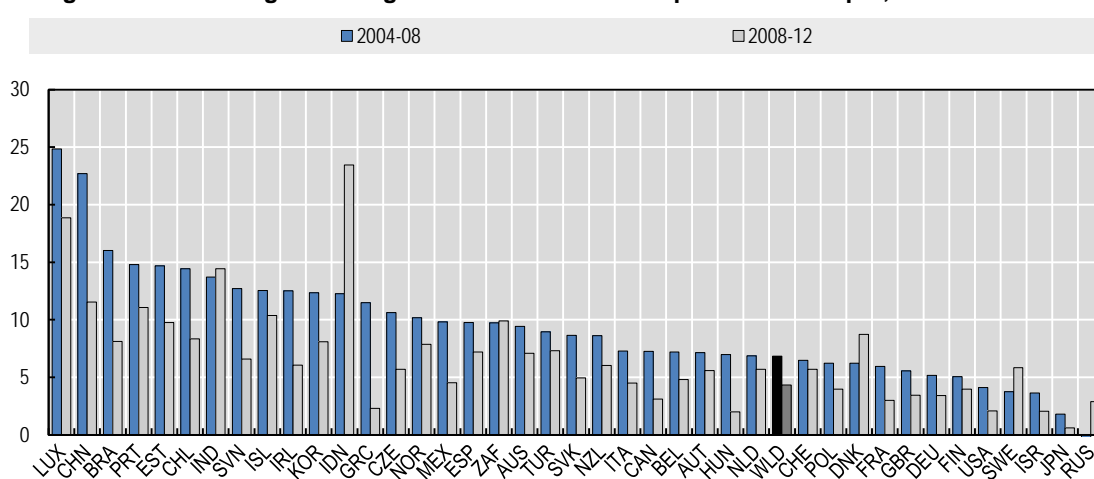
Definitions and measurability

Throughout this report, the term “**scientific production**” is used in a narrow sense to refer to the count of documents in scholarly journals in Scopus (all document types are included). These are based on whole counts of documents by authors affiliated to institutions in each country. As opposed to fractional methods, this method assigns in the final count an equal weight of one to each of the document’s authoring units, in this case distinct countries of author affiliation, regardless of how many there are (See Section 1.3 for more details).

The growth in scientific publication output in all countries, and especially in the BRIICS, is the combined result of an increasing volume of titles (e.g. journals) being added to Scopus and an increase in the number of documents within journals indexed throughout the entire period. New journals are added to Scopus through a process that results in increased coverage of previously existing journals, e.g. when they are deemed to attain the required standards for listing in the index, and newly created journals that from the outset are deemed to meet such standards. Documents can be retrospectively added in some cases. These indexing practices may as a result overstate the true increase in scientific production, and may more accurately reflect additions to the corpus of scientific work that scientists worldwide can be aware of, through its presence on a widely used index, and draw upon. Bibliometric indicators can thus not be dissociated from indexing practices, which vary from one producer to another. Overall, improvements in ICT tools allowing the use of much larger databases have led indices to attempt to capture more thoroughly the efforts of researchers and journal editors to join the global scientific community by expanding their coverage of journals all around the world and published in other languages (López-Illeras et al., 2009).

The table below shows the number of journals and documents in 2003 and 2012. There has been a net addition of more than 6 000 titles. It shows that if we calculate the number of documents using only those 2003 journals still covered in 2012, the estimate of total production for that year would be about 25% lower. New journals account for 60% of the absolute growth in the number of indexed documents over the period. This pattern differs across countries. Unfortunately, it is not possible to tell with available sources whether these new journals already existed in advance and were active in publishing peer-reviewed scientific research.

	2003		2012		
	Total	of which corresponding to journals still indexed in 2012	Total	of which corresponding to 2003 journals	of which new journals
Number of journals/titles	12 527	11 557	18 950	11 557	7 393
Number of documents	1 191 566	1 168 163	2 146 487	1 572 683	573 804
Number of journals lost from index - no longer published		970			
Number of journals added to index					7 393

Figure 1.1. Trends in scientific production for the top 5 countries, 2003-12[Download figure and data](#)**Figure 1.2. Top 20 countries, share of world scientific publication output, 2004-08 and 2008-12**[Download figure and data](#)**Figure 1.3. Average annual growth rates of scientific publication output, 2004-08 and 2008-12**[Download figure and data](#)

1.2. Scientific publications and R&D

Scientific publications disclose some of the peer-reviewed outcomes of research efforts. Despite their limitations as an indicator of research output, the comparison between estimates of research efforts, such as R&D expenditures, and scientific publications can provide additional evidence on the structure and productivity of national research systems with regards to this dimension of output.

It is possible to note (**Figures 1.4 and 1.5**) that countries with higher R&D intensity – gross domestic expenditures on R&D (GERD) as a percentage of GDP – also tend to exhibit higher rates of scientific publications per capita. Asian countries such as China, Korea and Japan tend to exhibit relatively low publication intensity relative to their R&D intensity. This may reflect a stronger focus on forms of research and development activity (by sector and research domain) that are less generally associated with scientific publications, but may also be partly due to lower coverage of scientific output in publication indices for reasons such as language biases (see box below).

The evolution of research efforts between 2003 and 2012 in OECD countries and the BRIICS exhibits some similarities but also some significant differences with respect to the recorded evolution of scientific publications. The intensity of scientific production, measured as the number of documents published in Scopus, per million inhabitants, grew for the OECD and BRIICS countries. R&D intensity, also increased for most countries although by a lesser margin.

The cross country correlation between scientific production per capita and R&D intensity declined over this period from 75% to 56%. This declining correlation may be due to a large number of factors. Publications can increase without increasing significantly R&D expenditures as a result of changing career incentives and research practices, for example in the form of increasing scientific collaboration (Moed, 2008). Publication activity may also displace other research related activities.

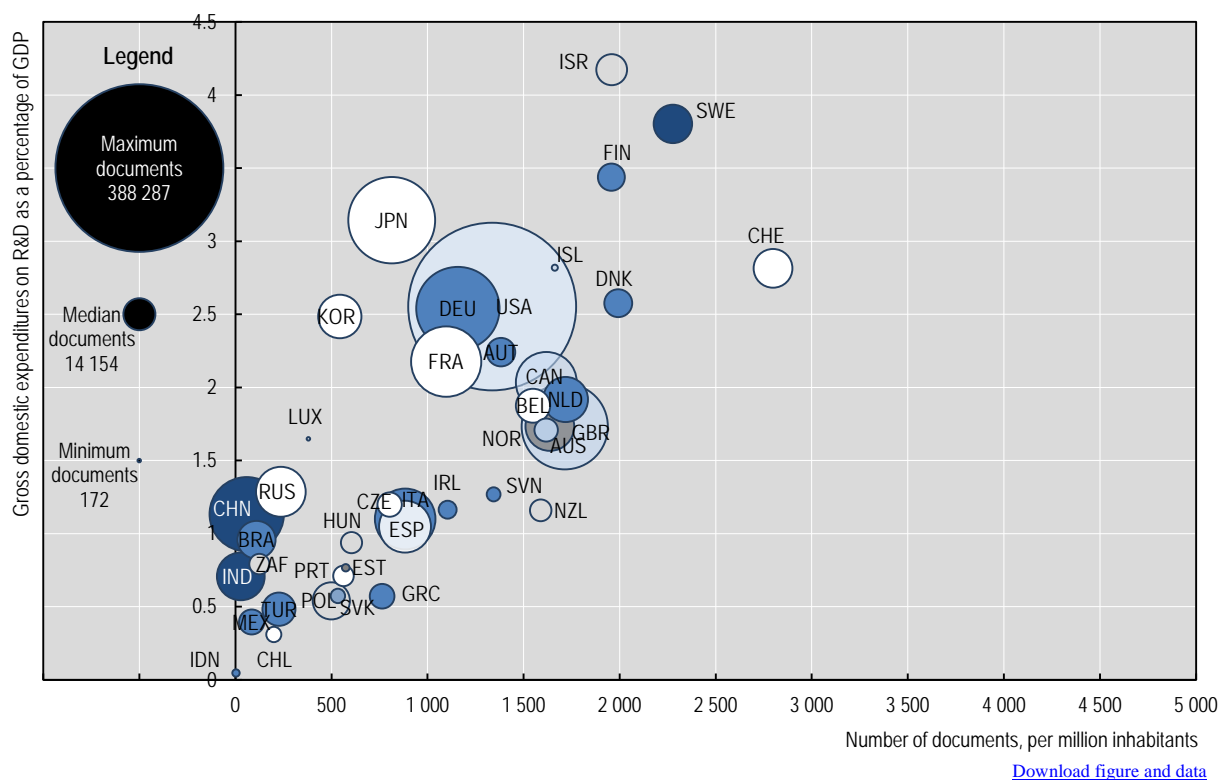
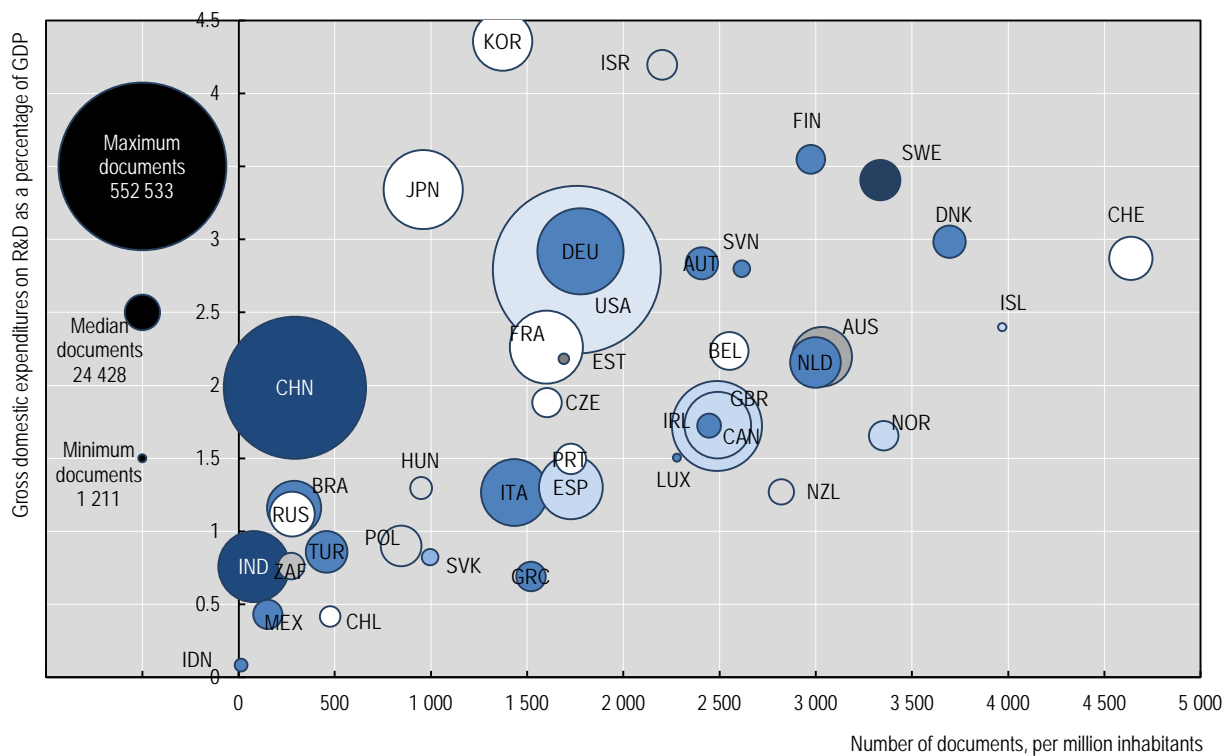
Definitions and measurability

GERD is the main aggregate used for international comparisons of expenditures incurred for the performance of R&D within an economy. The OECD *Frascati Manual 2015*, recently revised, defines R&D as “creative and systematic work undertaken in order to increase the stock of knowledge - including knowledge of humankind, culture and society - and to devise new applications” (OECD, 2015). GERD is often reported as a percentage of GDP, is used to denote the R&D intensity of an economy. Scientific publications are often presented on a per capita basis, normalised per million inhabitants, to compare scientific production among countries of different size.

The attribution of scientific production on a country basis may differ from that applied to R&D. This allocation difference can be particularly marked for some countries. One of the possible drivers for this difference is due to the fact that for R&D statistics, in line with standard statistical guidelines, the activities of international “extraterritorial” organisations (e.g. CERN, ESO) are treated as part of the rest of the world and therefore excluded. In contrast, scientific publications are typically assigned on the basis of which country the organisation is located, according to the address provided by the author. For example, over the 2003-12 period, we estimate that nearly 4 % of all documents potentially attributable to Switzerland were authored by individuals with an affiliation to CERN.

Language and related biases in indexing practices

Publications in languages other than English are not automatically excluded in Scopus. According to Elsevier, approximately 21% of titles in Scopus are published in languages other than English (Elsevier, [Scopus Coverage Guide](#), accessed December 9, 2014). However, there may be a number of implicit biases in indexing practices, for example the requirement for content to be available in Latin characters, or for the information about the journal to be available in English as demonstration of its “international” scope or outreach. These solution-oriented practices can be in some cases detrimental to reflecting scientific endeavours with a more specifically local focus in terms of subject matter and potential application. As a result, a global index may prove insufficient as a basis for representing the scientific efforts made, especially in a number of domains where the knowledge frontier is of a more local nature. These are some fundamental issues that lie at the heart of major ongoing debates concerning what should be the minimal requirements for inclusion in a scientific scholarly index and the types of uses, global or country-specific, that bibliometric databases may or may not be ready to support.

Figure 1.4. Scientific production and R&D intensity, 2003
[Download figure and data](#)
Figure 1.5. Scientific production and R&D intensity, 2012
[Download figure and data](#)

1.3. Scientific collaboration

It is widely held that scientific collaboration has become a more pervasive feature of scientific research activity in most countries. This trend is apparent when looking at the affiliations and geographic locations of the co-authors of scientific publications.

Scientific collaboration between different institutions over the 2003-12 period has been largest in Luxembourg, Iceland and France (**Figure 1.6**). In general, the incidence of international collaboration is larger than domestic inter-institutional collaboration. However, it is possible to note that the largest countries in terms of scientific production also exhibit larger domestic collaboration rates. In the case of France, its high institutional collaboration share is mainly due to its sizeable domestic collaboration share. The size of domestic collaboration in France may be explained by the widespread organisation of public research laboratories as joint ventures between at least one university and a public or private research organisation. The lowest shares of inter-institutional collaboration are found in China, Turkey and India, followed by Japan, Brazil, Korea and the United States. Further analysis is required to understand whether this is due to these countries having particularly large institutions which facilitate internal collaborations, or other factors that may prevent inter-institutional collaboration.

In 2012, Luxembourg, Iceland, Switzerland and Belgium exhibited the highest international collaboration rates (**Figure 1.7**), mainly followed by several other small (in publication terms) European countries. Although smaller countries are generally more likely to engage in international collaboration, this is not always the case. Chapter 2 will also show differences by disciplines. Several factors are known to influence international scientific collaboration, including geographical and cultural proximity (Appelt et al., 2015).

In the global landscape of scientific research, collaboration between institutions in different countries intensified over the 2003-12 period. Increased collaboration appears to have been one of the major factors driving the surge in publication activity over the period. The emergence of new players also changed the structure of global collaboration networks (OECD, 2013).

Definitions and measurability

As noted earlier, publications are attributed to countries on the basis of the authors' institutional affiliations. This requires a means of counting publications with co-authors from different units. One approach is to fractionalise publications by contributing units; so that reported figures add up to the total number of publications (each document has the same weight). An alternative is to report total counts per unit (the "whole counts" approach), which gives equal weight of one to each of the document's authoring units. These therefore have different implications for interpretation and further manipulation (fractional counts can be aggregated up while whole counts cannot). It is worth noting that the differences between results obtained by whole and fractional counts can also be a function of the database used for such measurement. Results could be affected by the balance of domestic publications and international publications.

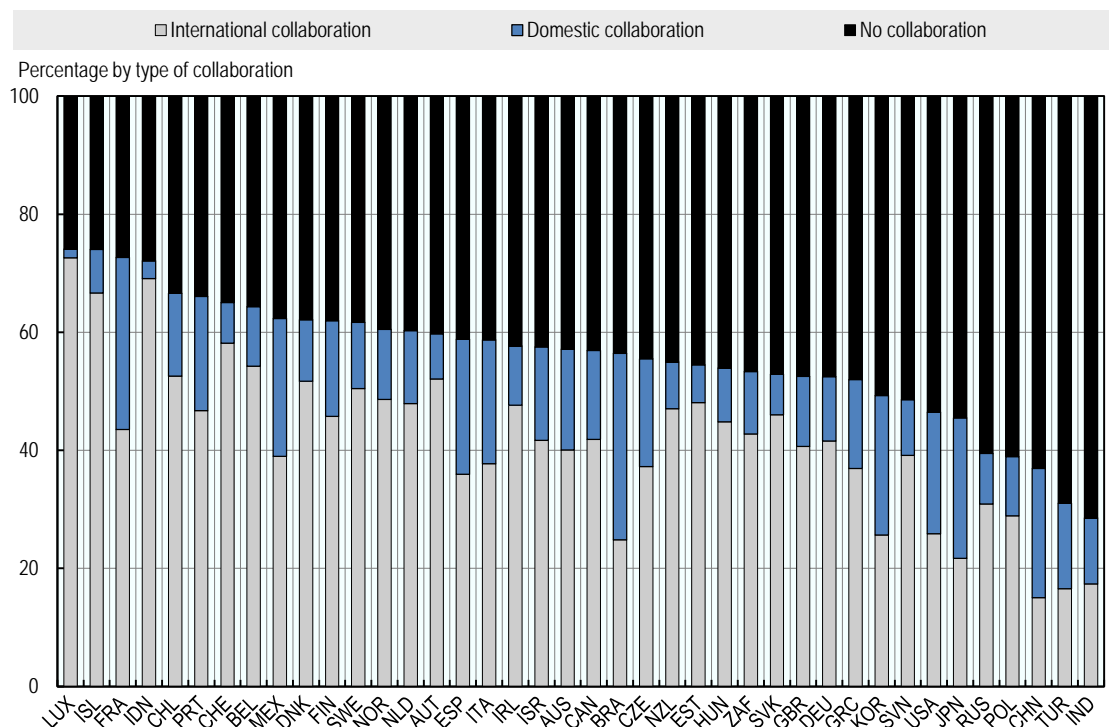
Although the choice does not affect country rankings much, care should be exercised when interpreting either type of results (Okubo, 1997; Moed et al., 2004; Javitz et al., 2010; Perianes-Rodriguez et al., 2015; Waltman et al., 2015). Estimates of scientific production and collaboration in this publication are based on whole count documents by authors affiliated to institutions in each country.

Country rankings can also be affected by how missing affiliations are dealt with. In order to achieve the highest level of precision for the different indicators SCImago carries out an extensive process of disambiguation of institutions' names contained in affiliation data.

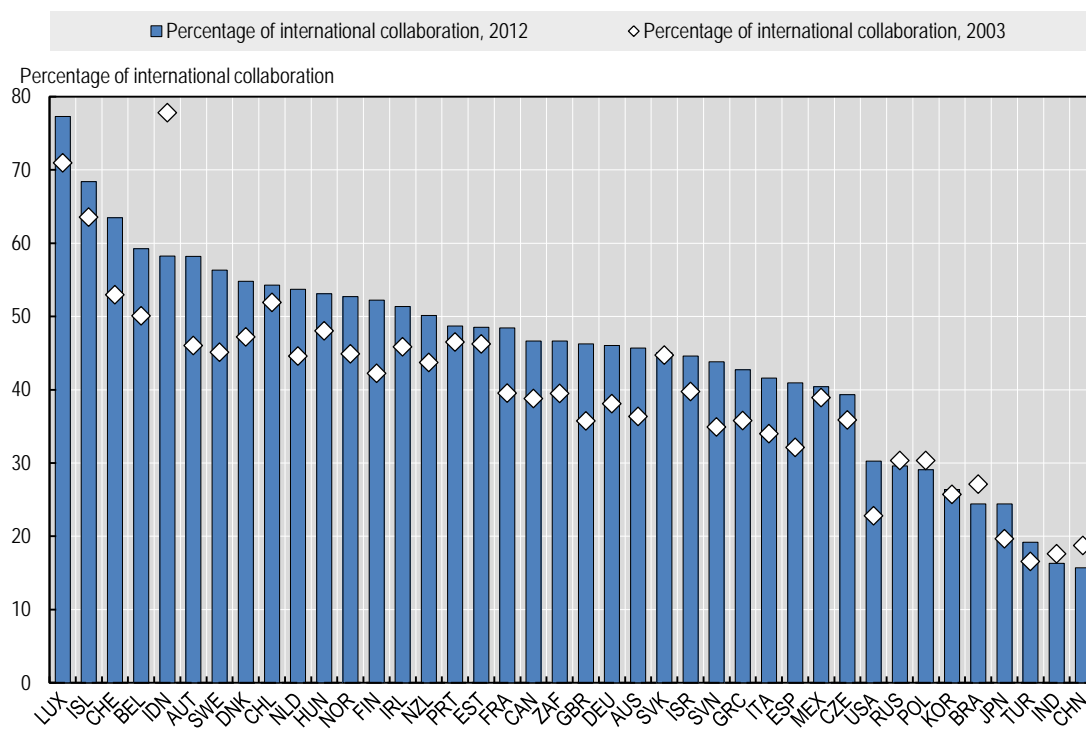
Collaboration is defined as co-authorship involving different institutions. International collaboration refers to publications co-authored among institutions in different countries. Estimates are computed for each country by counting documents for which the set of listed affiliations includes at least one address within the country and one outside. National collaboration concerns publications co-authored with different institutions within the reference country. No collaboration refers to publications not involving co-authorship across institutions. This includes singled-authored articles, as long as the individual has a single affiliation, as well as multiple-authored documents within a given institution.

Figure 1.6. Collaboration by country, 2003-12

As a percentage of all documents, whole counts

[Download figure and data](#)**Figure 1.7. Trends in International collaboration by country, 2003 and 2012**

As a percentage of all documents, whole counts

[Download figure and data](#)

1.4. Scientific leading authorship and collaboration

Individuals from different institutions and countries play different roles in the context of collaborative activities. While it is not possible to estimate accurately the share of the contribution of a given partner to any given scientific publication output, the information contained in publications allows for the identification of the institution and country of the author listed as “leading author”. This helps describe, at a global level, the likely leading role of a country in the context of international collaborations.

Over the 2008-12 period, Luxembourg, China and Portugal exhibited the highest shares of leading authorship among the documents exhibiting international collaboration (**Figure 1.8**). The picture is very similar to that over the 2003-08 period, with some exceptions. For example, in Greece, Russia and the Slovak Republic, there was an appreciable reduction in the share of international collaborations where these countries authors were listed as leading.

Combining information on leadership for international collaborations and documents involving a single country, it is possible to derive a more complete picture of a country’s integration in the global scientific landscape. Some countries with a high leadership share among collaborations may exhibit an overall low global leadership because of the small share of international collaborations in their overall output. This appears to be for example the case of China (**Figure 1.9**).

For a majority of countries, approximately 20% of their domestic publications have been led by authors affiliated to institutions in other countries. This percentage is significantly lower for a number of countries such as Brazil, China, India, Japan, the Russian Federation and the United States, which exhibit high levels of scientific production and a relatively low degree of international collaboration.

Definitions and measurability

Publications are attributed to countries on the basis of the authors’ institutional affiliations. This requires a means of counting publications where multiple authors and affiliations are listed. In addition to the fractional counting and whole counting methods, an alternative approach is to attribute the entire document to its leading author’s affiliation using information on the identity of the corresponding author (Moya-Anegón et al., 2013).

The concept of leading authorship can follow different norms within scientific groups. A main caveat is that the leadership criterion for attribution gives no weight to other contributors not listed as leading, so it should be interpreted carefully in conjunction with other indicators as it will likely give rise to a significantly more skewed distribution of output.

The scientific leadership indicator helps interpret the role of a given institution or country in collaboration activities, as reflected in publication output. The scientific leadership indicator shows the share of scientific output (in this case, documents by authors from a given country) where an author from this country is listed as leading author. At the country level, the indicator is defined only for documents involving international collaborations.

Figure 1.8. Domestic leading authorship in international collaboration, 2004-08 and 2008-12
As a percentage of documents exhibiting international collaboration

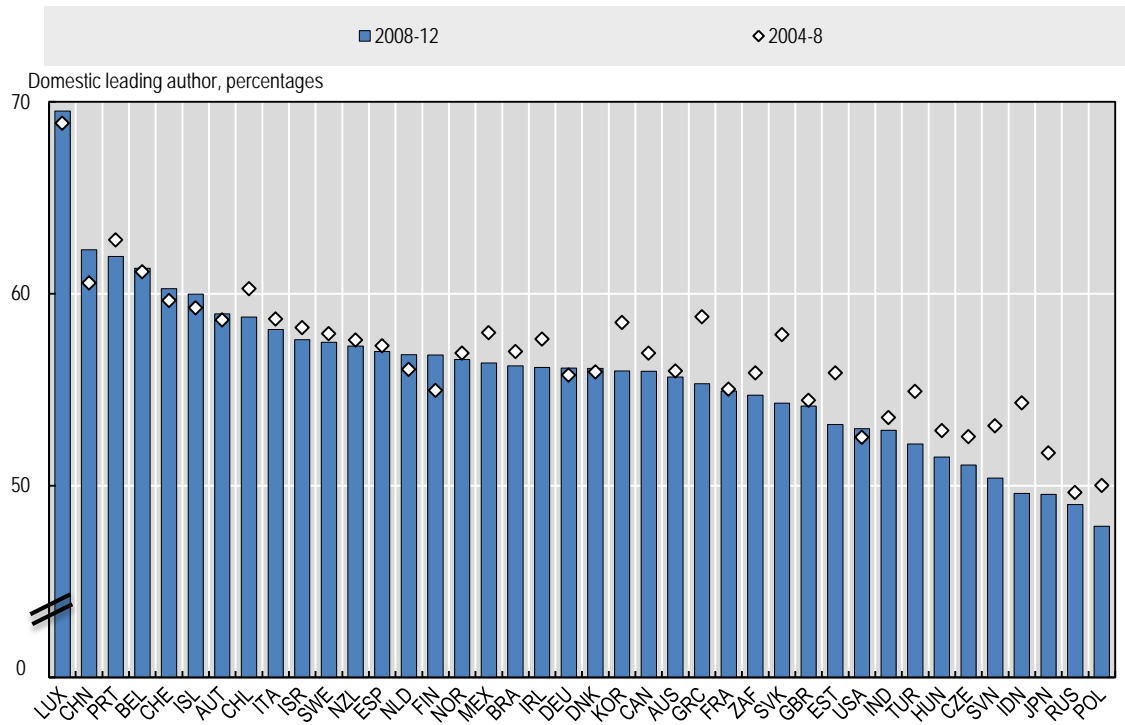
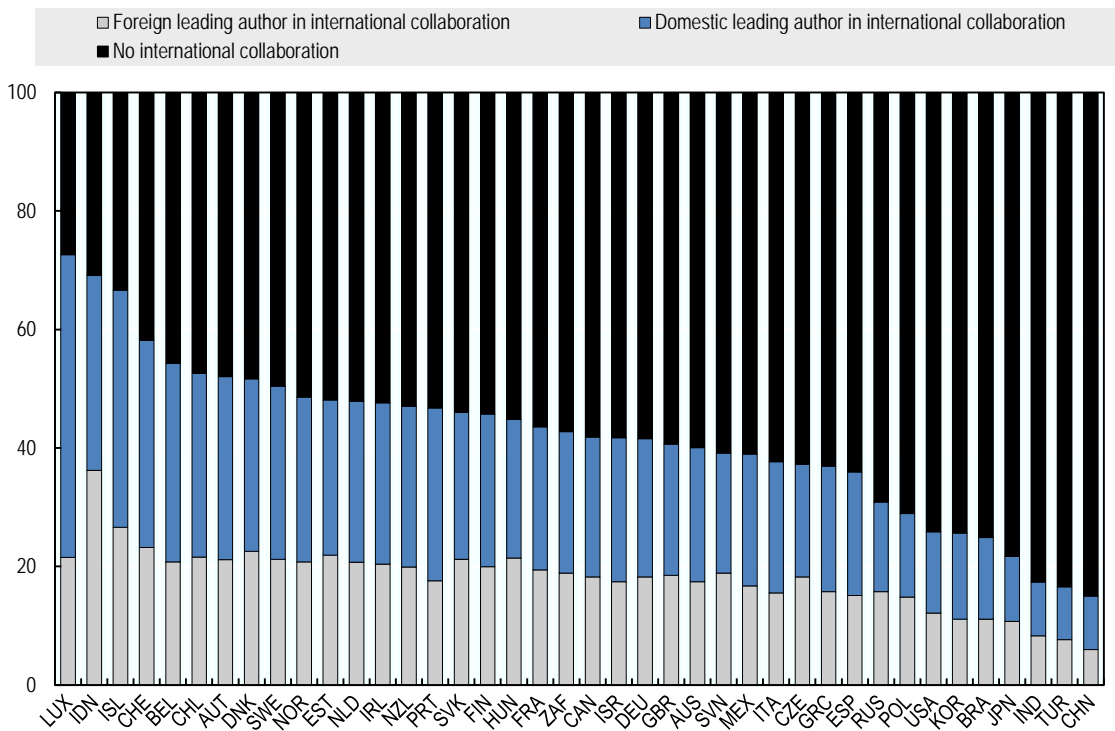

[Download figure and data](#)

Figure 1.9. Scientific leading authorship and international collaboration, 2003-12
As a percentage of all documents, whole counts


[Download figure and data](#)

1.5. Scientific “excellence” and leading authorship

Counting publications provides a very partial picture of a country’s contribution to the global scientific research effort, as simple counts give no indication as to the quality of publications. Inclusion in a scientific publication index only provides limited information on the novelty, robustness and broader relevance of a set of results. This calls for further analysis of the diversity that exists among listed documents.

Citations provide a means of inferring the relevance of a given document to the scientific community as documents are subsequently published. Counting how many of the 10% most-cited publications correspond to authors in a given country provides a quality-adjusted measure of its research output, in other words, a proxy for scientific excellence (Bornmann et al., 2012).

Switzerland has the largest share of documents (close to 20%) with a high citation impact among domestic publications, closely followed by the Netherlands and Denmark (**Figure 1.10**).

Joint analysis of excellence and leadership information can provide further insights into the source of a country’s highly cited publications. In the United States, for example, 17% of publications are among the 10% top cited, of which 14% had a US-based leading author, while only 3% are led by authors with affiliations abroad (**Figure 1.11**). Accordingly, the United States has the largest share of top cited publications led by domestic authors, followed by the Netherlands and the United Kingdom. Other countries with higher overall excellence rates display lower levels of leading excellence because of the higher importance of collaborative articles led by authors from other countries.

Definitions and measurability

Peer-reviewed scientific publications convey the research findings of scientists worldwide. Subsequent citations by other authors provide an indirect but objective source of information about the quality of research outputs, as implied by their use by the scientific community itself. Despite the limitations, for example the fact that citations do not take into account the use of the scientific information by inventors or practitioners who are less likely to publish in peer-reviewed journals, they provide one of the possible quality adjustments to raw counts of documents. Its relevance can be considered to be higher in the context of the higher education sector (see Chapter 3 on publications by sectors).

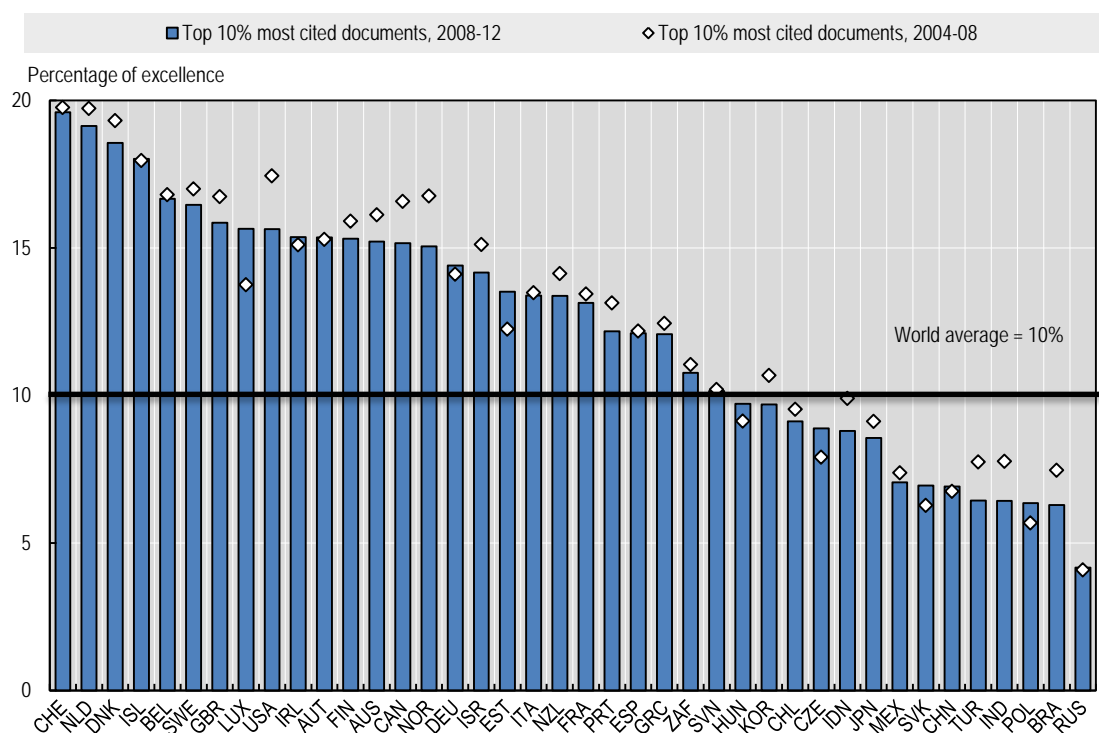
The indicator of scientific excellence indicates the amount (in %) of a unit’s scientific output that is part of the set of the 10% most-cited papers within their respective scientific fields. The choice of a given threshold is largely arbitrary. The top 10% most-cited is most frequently used, but one could also look at other thresholds such as the top 1%, top 5% or the top 25%.

The indicator of scientific excellence is calculated at the document level using whole counts. Documents – organized by ASJC field and year - are sorted in descending order based on the number of citations received. A threshold of 10% most cited documents is calculated for each category. Only documents with a fixed number of citations above the threshold are included. Documents with the same number of citations as the threshold are sorted according to the [Scimago Scientific Journal Rankings](#) (SJR) value of the journal in which they were published; those with the highest scores are selected. It is likely that this step of the procedure – used to select the citation tiebreak documents – favours countries publishing in higher prestige journals versus a more random approach. The resulting documents are included until 10% is attained. Only after this selection has been made can the identification by country, sector, and institution be made to determine their share in total output of the said unit of analysis. The citation window is based on the whole period, so top 10% most cited documents for 2003-12 uses the citations for the 2003-12 period. The world average is 10% for the period.

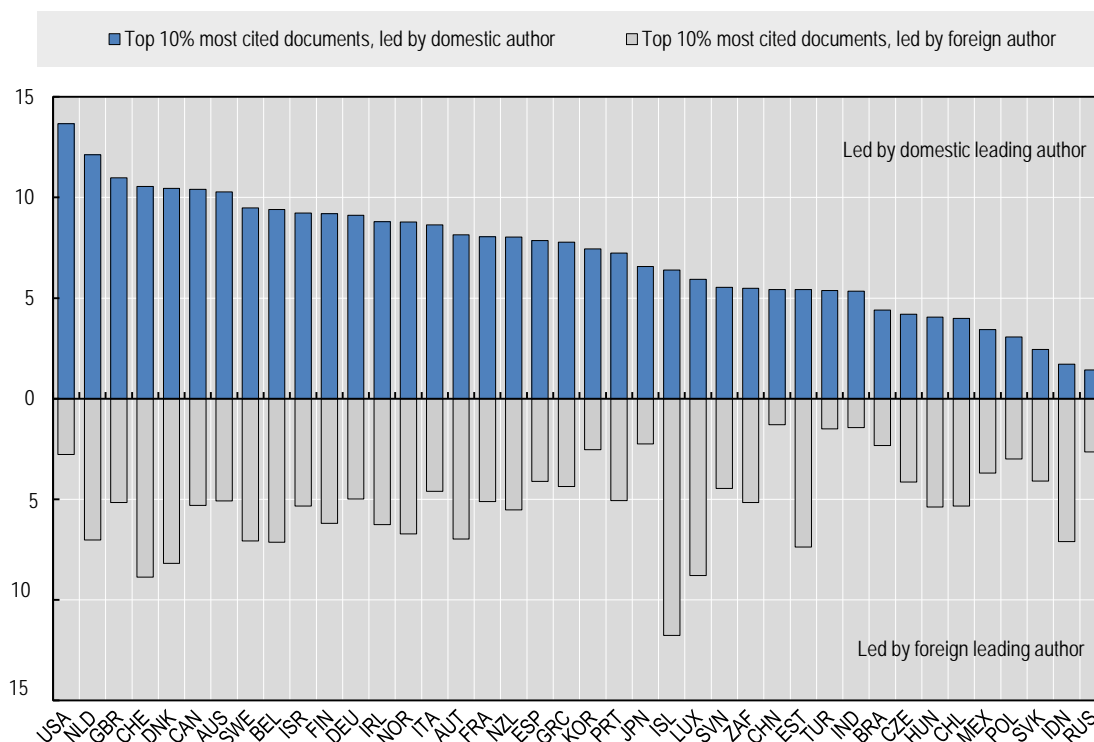
This indicator can be used in combination with the scientific leadership indicator (defined in the previous section) to better describe the role of international collaborations in driving the observed number of a country’s publications rated as “excellent”. For example, it is possible to estimate how many “top cited publications” initially attributed to a country have a leading author from another country. This type of measure provides an indication of the extent to which international collaboration is a mechanism for enhancing the quality of a country’s scientific output, by drawing on expertise elsewhere through collaboration.

Figure 1.10. Top 10% most cited documents by country, 2004-08 and 2008-12

As a percentage of all documents, whole counts

[Download figure and data](#)**Figure 1.11. Top 10% most cited documents and scientific leading authorship, 2003-12**

As a percentage of all documents, whole counts

[Download figure and data](#)

1.6. Trends in scientific production, top cited publications

It is highly relevant to establish whether the trends exhibited by scientific publications are also reflected among those that are more intensely cited – and presumably more impactful – publications. An increase in the number of publications in a country can be concentrated on higher or lower “impact” publications, reflecting the extent to which any observed expansion has been accompanied or not by a shift towards “excellence”.

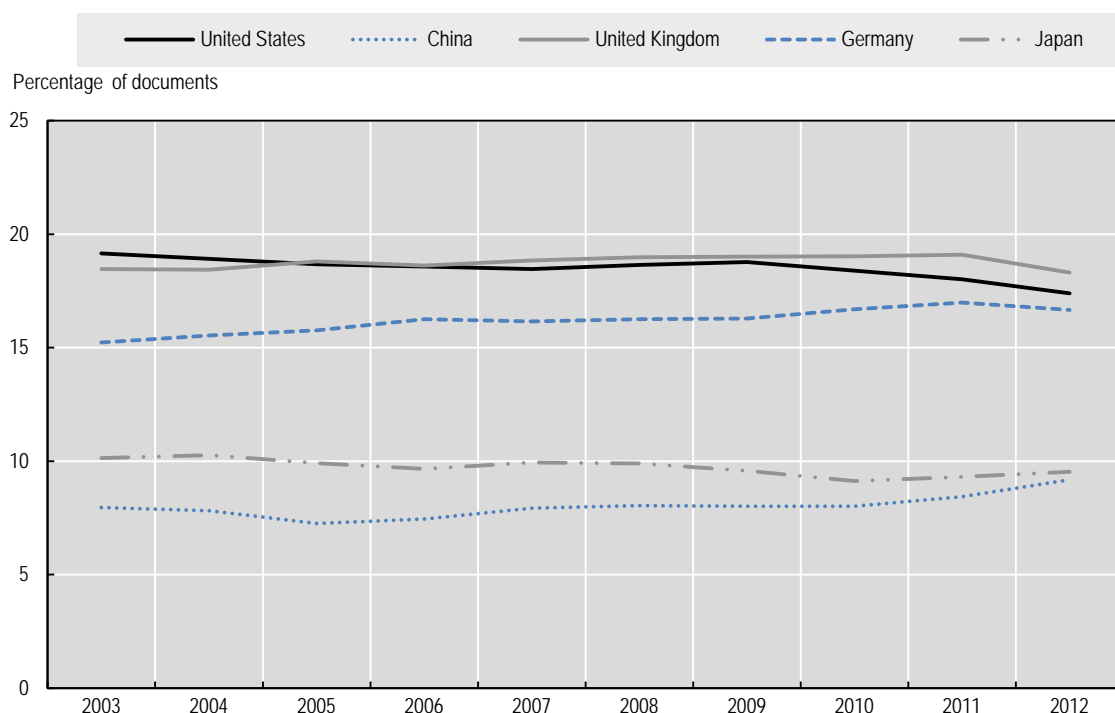
Trends for the top 10% cited documents among the largest producing countries have been fairly stable over the 2003-12 period, but some relevant patterns can be identified (**Figure 1.12**). For example, Germany’s share of top-cited publications appears to be converging with that for the United Kingdom and the United States. The latter’s rate has declined by approximately 2 percentage points over the last decade. China’s rate of (top 10%) highly cited publications has caught up over this period with Japan’s, at nearly 10%. A significant gap persists between these two countries and the United Kingdom, the United States and Germany.

These patterns translate into some significant shifts in the global distribution of highly cited documents (**Figure 1.13**). In 2003 about 92% of the top-cited documents were in the OECD countries. By 2012, this share dropped to 83%. The United States retains the largest absolute number of top-cited documents, with its share in world top cited documents slightly receding from 51% to 44% over the 2003-12 time period, compared with total publications where the share fell from 32 to 25%. China’s share of highly cited documents has grown very rapidly, rising from 4 to 16%, compared with 6 to 17% for its share in all publications. Because of the low starting point and lower rate of top cited publications, even assuming that current trends might continue into the future, it would take several years if not decades for China’s level of highly cited publication to catch up with that of the United States.

Definitions and measurability

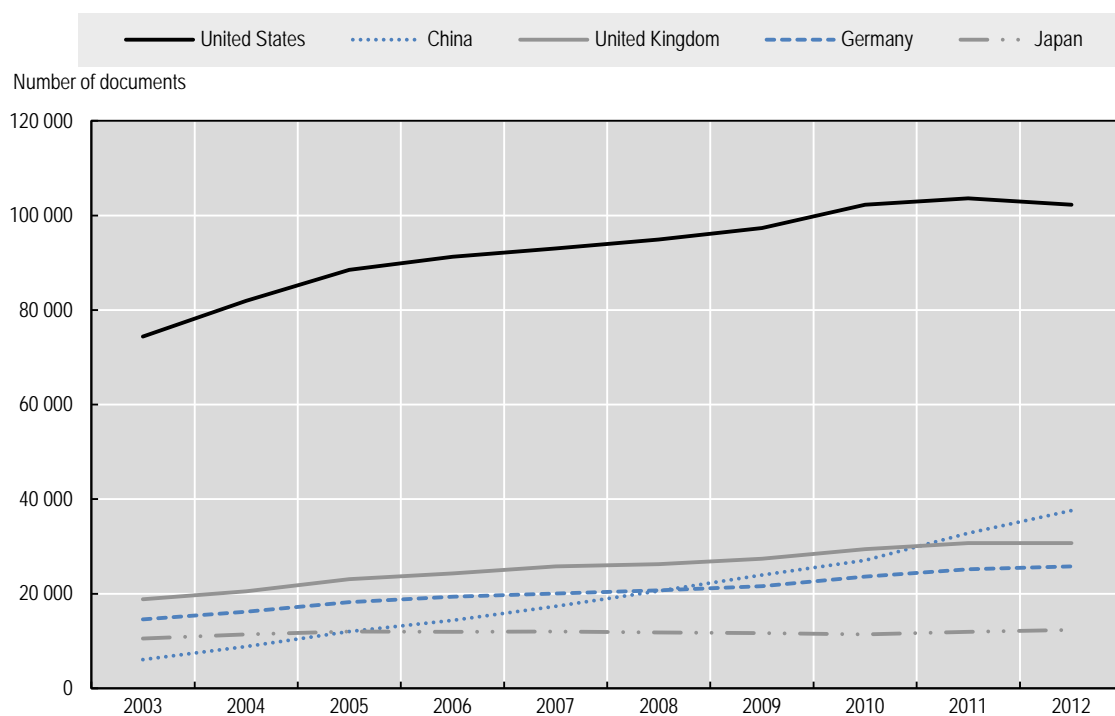
See the Definitions and measurability box in Section 1 on Trends in scientific production for Scopus indexing practices which notably affect the BRIICS countries.

Figure 1.12. Shares of top 10% most cited documents in top 5 producing countries, 2003-12
As a percentage of all documents within each country



[Download figure and data](#)

Figure 1.13. Trends in top 10% most cited documents for top 5 producing countries, 2003-12
Number of documents, whole counts



[Download figure and data](#)

1.7. The citation impact of scientific production and international collaboration

The production of scientific research is progressively shifting from individuals to groups, from single to multiple institutions, and from the national to international level. Because they draw on larger pools of expertise, international research collaborations are more likely to have a bigger impact in terms of citations in subsequent scientific publications. Differences across countries suggest a positive relationship between measures of scientific research collaboration and impact, the latter proxied in this case by the average normalised citation index (**Figure 1.14**). This relationship appears to be stronger in economies with lower levels of scientific production, suggesting the importance of scale, which smaller economies can overcome by participating in global networks. While collaboration may drive increased impact (and more efficient work), the association may be partly due to a selection effect whereby higher impact authors attract more collaborators.

In the figure, bubbles plot a country's share of documents resulting from international collaboration – as implied by the share of domestic articles co-authored with individuals affiliated with foreign institutions – against the normalised impact of its publications. The bubble size represents the volume of scientific production, with the United States and China being the largest producers of scientific output. Switzerland had both a high share of international scientific collaboration and high average impact, although its total output volume was smaller than that of countries such as France or the United Kingdom.

The relative position of countries on the basis of top quartile journals or average citation impacts is fairly similar (**Figure 1.15**), and close to the one of the top 10% most cited documents, with some exceptions, especially in the case of smaller countries, where there may be relatively few high impact articles but with high numbers of citations.

Definitions and measurability

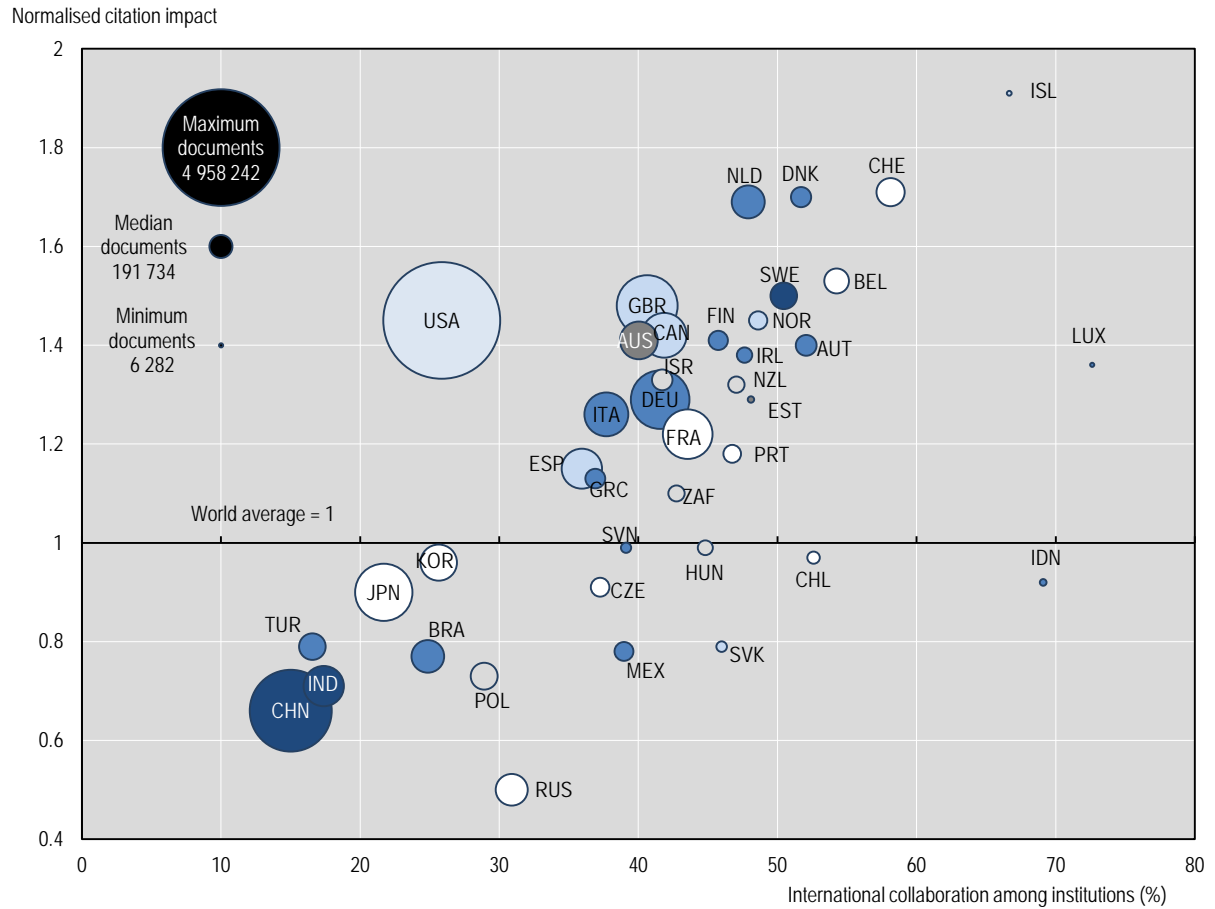
As in previous figures, the international institutional collaboration indicator is based on the proportion of documents involving institutional affiliations with other countries or economies, relative to all documents attributed to authors with an affiliation in the reference economy. Single-authored documents with multiple affiliations in different countries count as *institutional* international collaboration.

The normalised impact measure is derived as the ratio between the average number of citations received by the documents published by authors affiliated to an institution in a given economy and the world's citation average, over the same time period, by document type and subject area. The normalisation of citation values is item-oriented, i.e. carried out at the level of the individual article. If an article belongs to several subject areas, a mean value is calculated for each area. The values show the relationship of the unit's average impact to the world average, which is 1, i.e. a score of 0.8 means the unit cited is 20% less than average and 1.3 means the unit cited is 30% more than average.

Publication counts in top quartile journals are defined as publications in the reference period by authors affiliated to an institution in a given country published in the most influential 25% of the world's scholarly journals in their category, as ranked by the SCImago Journal Rank (SJR) indicator (www.scimagoir.com) on the basis of citation data. This indicator is primarily used as a predictor of actual citations, especially when the time window on which to compute citations is limited. The various types of "impact" and "excellence" measures are highly inter-related. Normalised citation impact indicators are based on citations per document relative to an area "average" benchmark. All these indicators tend to produce rather similar country rankings, but may exhibit more variability as they are applied to smaller scientific production units.

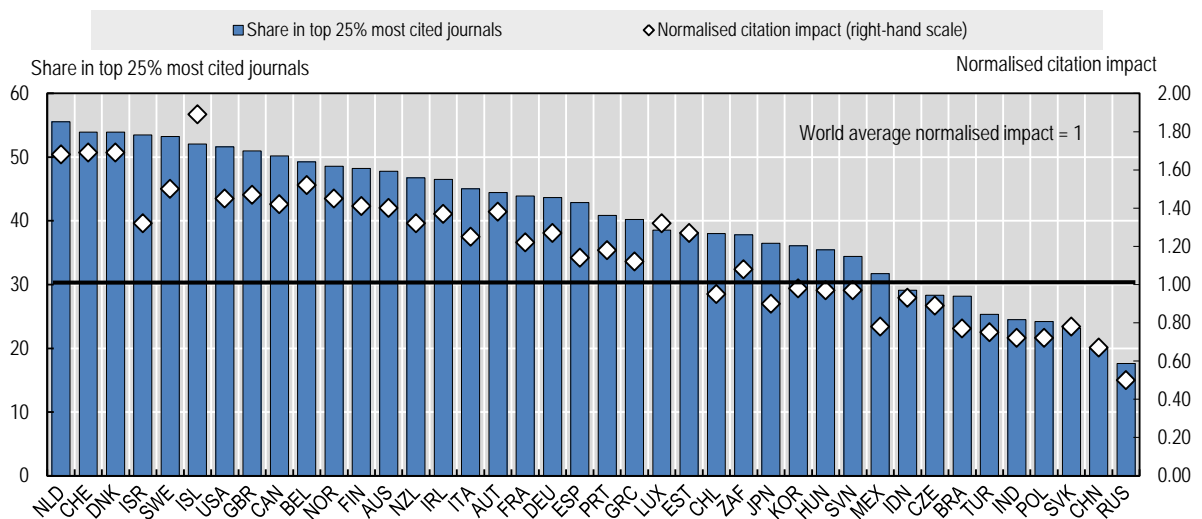
Although article citation has the advantage of focusing directly on the impact of the articles examined, it takes time for documents to be read and subsequently cited in new documents, particularly in some disciplines. A trade-off exists between the length of time over which citations are accounted for and the timeliness of the indicator: the more time allowed for measuring the impact, the less timely the indicator becomes. For this reason, in some context, an assessment is made on the basis of the average citation impact for documents previously published in the same journal, providing an indicator of the "expected" citation impact. Individual documents may subsequently exceed or fall short of this "expectation".

Figure 1.14. The impact of scientific production and international collaboration, 2003-12



[Download figure and data](#)

Figure 1.15. The quantity and quality of scientific production, share in the top 25% most cited journals and citation impact, 2003-12
As a percentage of all documents, whole counts



[Download figure and data](#)

1.8. Scientific publishing and open access

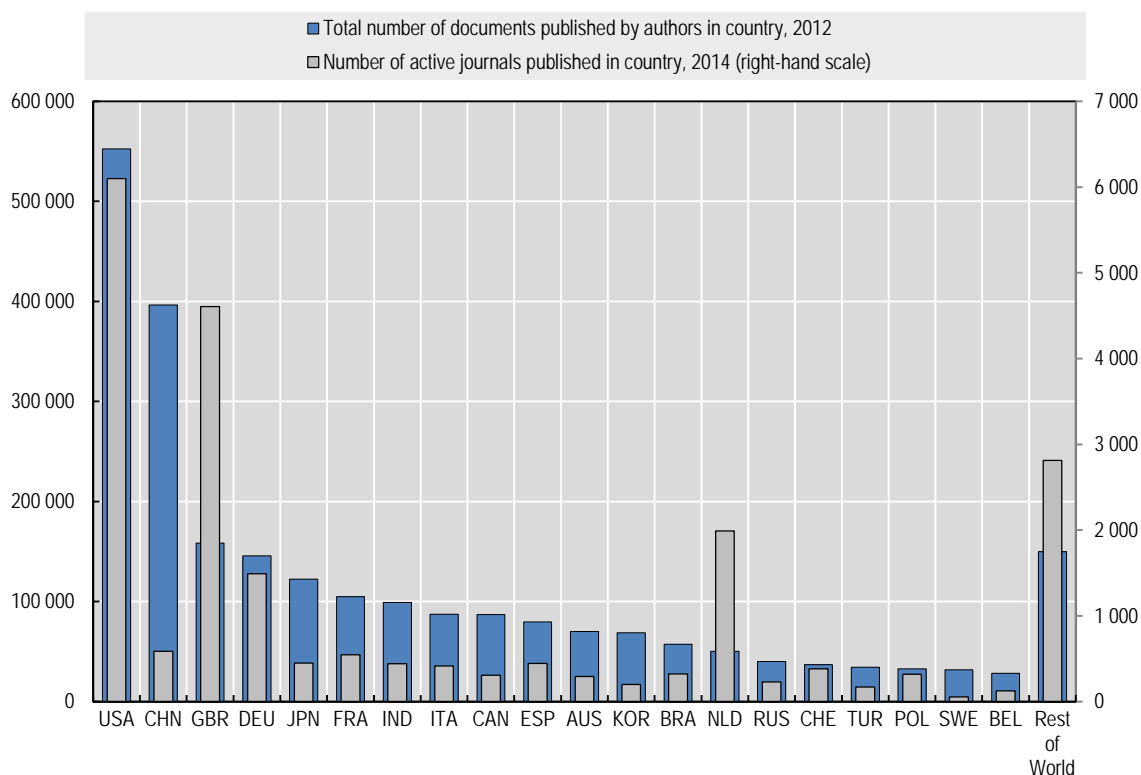
The activity of scientific publishing can play a very important role in the process of knowledge dissemination and quality assurance, contribute to shaping research careers as well as provide in its own right a significant economic activity. Many journals are international in scope but tend to have a main centre of interest depending on where its editors and publishers make major decisions. The analysis of the location of publishing activity, based on information on titles indexed by Scopus (**Figure 1.16**), reveals that it is highly concentrated among a few countries, showing the location of some major publishing houses. For example, the academic publisher Elsevier, which also owns and publishes Scopus, is based in the Netherlands and so are a large percentage of its titles. The correlation between the number of active journals in a country and the total number of publications by that country is 0.75, for OECD and BRIICS countries (0.98 if China is excluded). In 2012, China, the second largest producer of scientific documents in the world, only accounted for 2.6% of all active journals listed in Scopus. The United States, on the other side of the spectrum, with the largest amount of documents published, also “hosts” the largest share of world journals, 27.4%. With the widespread use of remote online communications, the location of publishing activity may now have a smaller impact on the extent and quality of scientific production efforts than it used to have in the past.

In Scopus, 2 800 titles, i.e. about 12% of those active journals, were identified as open access (OA) journals. For the OECD countries the average share of journals listed as OA is 8%. For the BRIICS countries the OA share is much higher at 25%. Open access data show significant differences in the use of this means of research distribution across countries. In Chile 77.6% of all active journals are open access journals, with 66 OA journals. The United States has the largest absolute number of OA journals, 268 journals. This however represents a very low share out of total journals, 4.4%. The United Kingdom has the second largest number of OA journals, 263 journals, followed by Brazil and India, with 219 and 159 OA journals respectively (**Figure 1.17**). It must be noted that some countries host very few active journals and that these figures represent the status by journal according to where it is published, not the status of a document. An analysis at the level of documents based on the status of the journals they are published in is available in OECD (2015b). The OECD is also currently doing [new survey analysis](#) of OA publication patterns by authors, also investigating the availability of documents on open registers or from the publishers themselves after an embargo period.

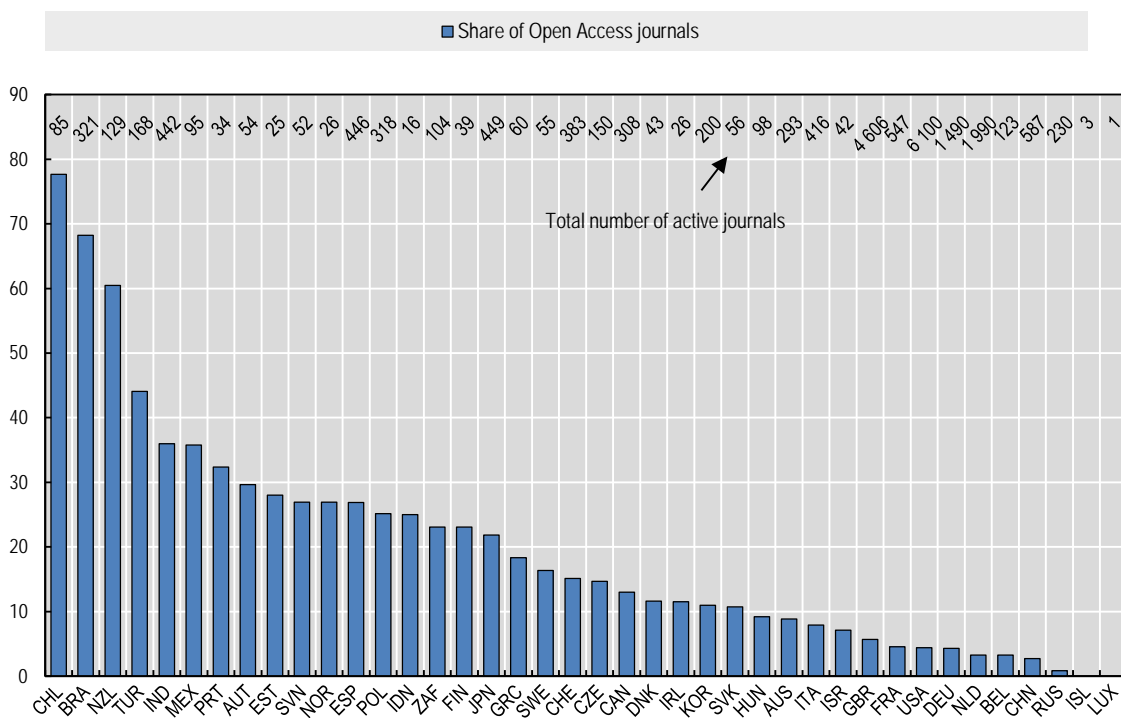
Definitions and measurability

In 2014, **Scopus** covered approximately 22 283 active journals worldwide. The OECD and BRIICS countries accounted for 92.5% of those active journals. For a journal to be included in Scopus it has to meet certain criteria and go through a board approved review process: The title should publish peer-reviewed content; The title should be published on a regular basis (i.e. have an ISSN that has been confirmed by the ISSN International Centre); The title should have English language abstracts and article titles; The title should have references in Roman script; and should have a publication ethics and publication malpractice statement. In 2014, 324 new journals were added. Once a journal has been added, the information for that journal can be added retrospectively. In 2014, 767 new journal title suggestions were received for possible inclusion. The number of suggested titles varies by subject area; the acceptance rate is about 42%. On average, once a request has been received it can take about four to twelve months to be processed.

The Open Society Foundation defines **open access** as “...an alternative publishing and distribution model that makes scholarly research literature freely available to the public online, without restrictions.” In its listing of titles, Open access is defined by Elsevier as: publishing in either an open access journal, or a journal that supports open access; giving immediate public access to the final published article; where the publishing fee is paid by the author (or on his behalf); and the use is determined by the author's choice of user license. The designation of titles as OA is largely based on inclusion in the Directory of Open Access Journals (DOAJ), and as such, focuses on the Gold-Access model of open access. DOAJ maintains a database of over 8,000 OA journals. The inclusion criteria for the DOAJ are specified on their website and include being freely accessible upon publication and implementing some form of quality control. The DOAJ and Scopus listing criteria reduce to a minimum the potential risk of basing our analysis on journals that adopt so-called “[predatory](#)” practices.

Figure 1.16. Scientific publication output and number of active journals, top 20 countries
[Download figure and data](#)
Figure 1.17. Open access publishing, 2014

As a percentage of all active journals published in each country


[Download figure and data](#)

Chapter notes and references

Country scope

In this report, the geographical coverage has been limited to the OECD countries and the BRIICS (Brazil, the Russian Federation, India, Indonesia, China and South Africa). For presentational reasons, other economies have been excluded.

Figure 1.2. Top 20 countries, share of world scientific publication output, 2004-08 and 2008-12

The reference to “top 20 countries” refers in this case to the countries within the scope of this study. In the full list of countries and economics, the Islamic Republic of Iran (IRN) would feature in 17th place in terms of the volume of indexed publications corresponding to authors affiliated to institutions in this country.

References

- Appelt, S., F. Galindo-Rueda, R. de Pinho and B. van Beuzekom (2015), “Which factors influence the international mobility of research scientists?”, OECD Science, Technology and Industry Working Papers, 2015/02, OECD Publishing, Paris, <http://dx.doi.org/10.1787/5js1tmrr2233-en>.
- Bornmann, L., F. Moya Anegón and L. Leydesdorff (2012), “The new Excellence Indicator in the World Report of the SCImago Institutions Rankings 2011”, *Journal of Informetrics*, Vol. 6, No. 2, pp. 333-335, <http://dx.doi.org/10.1016/j.joi.2011.11.006>.
- Javitz H., T. Grimes, D. Hill, A. Rapoport, R. Bell, R. Fecso and R. Lehming (2010), “U.S. Academic Scientific Publishing”, Working paper SRS 11-201. Arlington, VA: National Science Foundation.
- López-Illescas, C., F. Moya-Anegón and H.F. Moed (2009), “Comparing bibliometric rankings derived from the Web of Science and Scopus: The effect of poorly cited journals in oncology”, *Journal of Information Science*, Vol. 35, No. 2, pp. 244-256, <http://dx.doi.org/10.1177/0165551508098603>.
- Moed, H.F. (2008), “UK research assessment exercises: informed judgments on research quality or quantity?”, *Scientometrics*, Vol. 74, No. 1, pp. 141-149, <http://dx.doi.org/10.1007/s11192-008-0108-1>.
- Moed, H.F., W. Glänzel and Ulrich Schmoch (eds.) (2004), *Handbook of Quantitative Science and Technology Research: The Use of Publication and Patent Statistics in Studies of S&T Systems*, Kluwer Academic Publishers, the Netherlands.
- Moya-Anegón, F., Guerrero-Bote, L., Bornmann and H.F. Moed (2013), “The research guarantors of scientific papers and the output counting: A promising new approach”, mimeo, accessed from www.lutz-bornmann.de/icons/guarantor.pdf, June 2013.
- OECD (2015a), *Frascati Manual 2015: Guidelines for Collecting and Reporting Data on Research and Experimental Development*, The Measurement of Scientific, Technological and Innovation Activities, OECD Publishing, Paris, <http://dx.doi.org/10.1787/9789264239012-en>.
- OECD (2015b), “Making Open Science A Reality - Final Report”, Working Party on Innovation and Technology Policy, OECD, Paris, <https://www.innovationpolicyplatform.org/content/open-science>

OECD (2013), *OECD Science, Technology and Industry Scoreboard 2013: Innovation for Growth*, OECD Publishing, Paris, http://dx.doi.org/10.1787/sti_scoreboard-2013-en.

Okubo, Y. (1997), “Bibliometric Indicators and Analysis of Research Systems: Methods and Examples”, OECD Science, Technology and Industry Working Papers, 1997/01, OECD Publishing, <http://dx.doi.org/10.1787/208277770603>.

Perianes-Rodriguez, A. and J. Ruiz-Castillo (2015), “Multiplicative versus fractional counting methods for co-authored publications. The case of the 500 universities in the Leiden Ranking”, *Journal of Informetrics* Vol. 9, pp. 974–989, <http://dx.doi.org/10.1016/j.joi.2015.10.002>.

Waltman, L. and N.J. van Eck (2015), “Field-normalized citation impact indicators and the choice of an appropriate counting method”, *Journal of Informetrics* Vol. 9, pp. 872–894, <http://dx.doi.org/10.1016/j.joi.2015.08.001>.

This page is intentionally left blank

2. SCIENTIFIC PUBLICATION INDICATORS BY FIELDS OF SCIENCE

1. Scientific publication output and journals by fields
 2. Trends in indexed scientific publication output by fields
 3. Field specialisation by country
 4. Citations by fields
 5. Fields with highest citation impact by country
 6. International collaboration by fields
 7. Open access by fields
- Chapter notes and references

2.1. Scientific publication output and journals by fields

The scientific publication output captured by Scopus is distributed unevenly across 27 different fields (**Figure 2.1**) in the All Science Journal Classification (ASJC) (**Table 2.A.1.**). Over the 2003-12 period, Medicine accounted for the largest number of indexed documents, namely 4.4 million documents (24%), followed by Engineering (12%) and Biochemistry, Genetics & Molecular Biology (11%). There are some significant differences between OECD countries and the BRIICS, which underpin country-level specialisation patterns to be covered in section 2.3. For example, Engineering is the most frequent field by documents published by authors in the BRIICS group of countries.

Documents are attributed to fields in this publication on the basis of the document's journal's source classification. It is therefore relevant to compare the field distribution of journals (**Figure 2.2**) with that for documents based on their journal's classification. While Medicine accounts for both the largest number of documents and journals, Social Sciences and Arts & Humanities account for a relatively large number of journals (2nd and 3rd largest) compared to the number of documents in that field. This appears to be explained by differences in the average number of documents published per journal, which in turn may relate to differences in number of pages/words per document as well as publication frequency. This is suggestive of different publication practices by field.

Across fields, the share of journals published in OECD countries ranges from 45% to 94%, with an average of 84%. BRIICS countries have the largest of journals in Medicine and Engineering; these are also their top two fields in terms of output (**Figure 2.2**).

Definitions and measurability

Elsevier assigns each journal in Scopus to one or more subjects using its All Science and Journal Classification (ASJC). There are 27 main fields comprising 334 subjects in the classification. The 27 fields and their abbreviations are reported at the beginning of this publication.

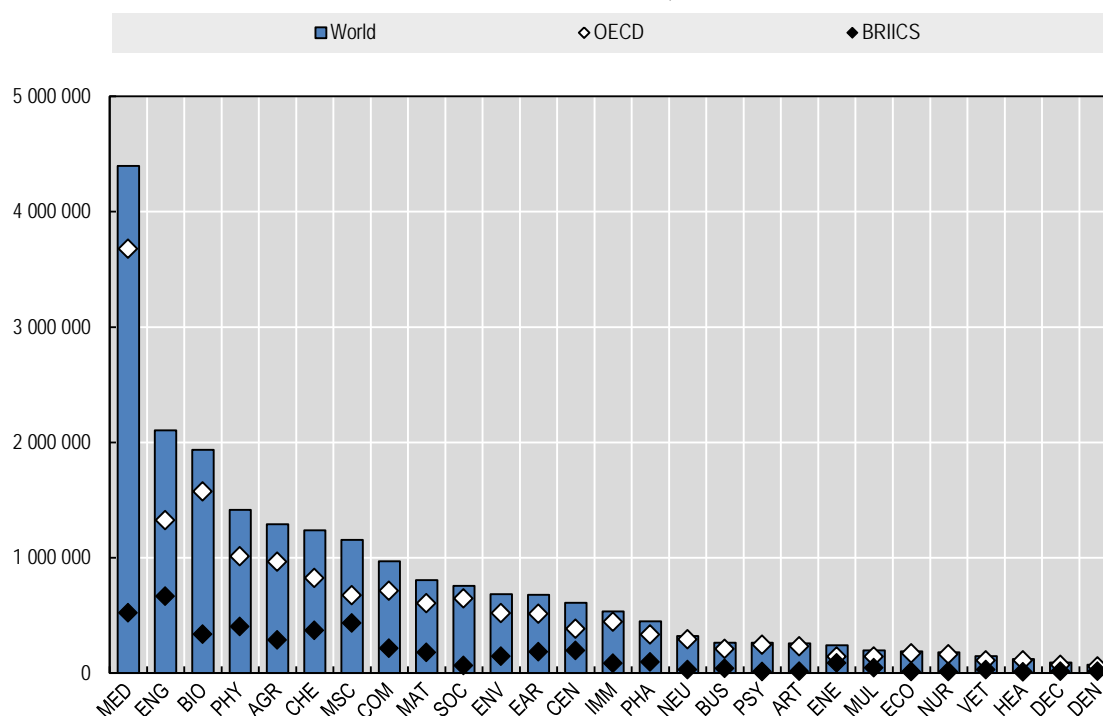
Documents are implicitly assigned to fields on the basis of the journal in which they are published. This is therefore an approximate assignment of research outputs to areas because a given journal's classification may not provide an accurate representation of a document's thematic content. In 2014, the 22 283 active journals covered in Scopus were attributed to an average of 2.06 subjects each. Only 39% of all titles were allocated to a single subject. The average for OECD countries was 1.9 subjects, for the BRIICS countries it was 1.8. Journals in the Netherlands and Spain had the highest OECD average with 2.5. For fields-based statistics, this publication adopts a whole counts approach, rather than a fractional perspective.

The ASJC classification comprises a "Multidisciplinary" field based on its editorial line. However, it is important to note that not all documents published in journals classified into this field such as *Nature* or *Science* need not represent the outcome of multidisciplinary research. This matters for the interpretation of the results in this chapter. Journal-based field assignments could lead to under/over estimating the output for certain fields. Further refinements could be envisaged such as the text mining of key words or abstracts or references at the document level. This could provide a more accurate assignment of documents and their research areas (López-Illeras et al., 2009).

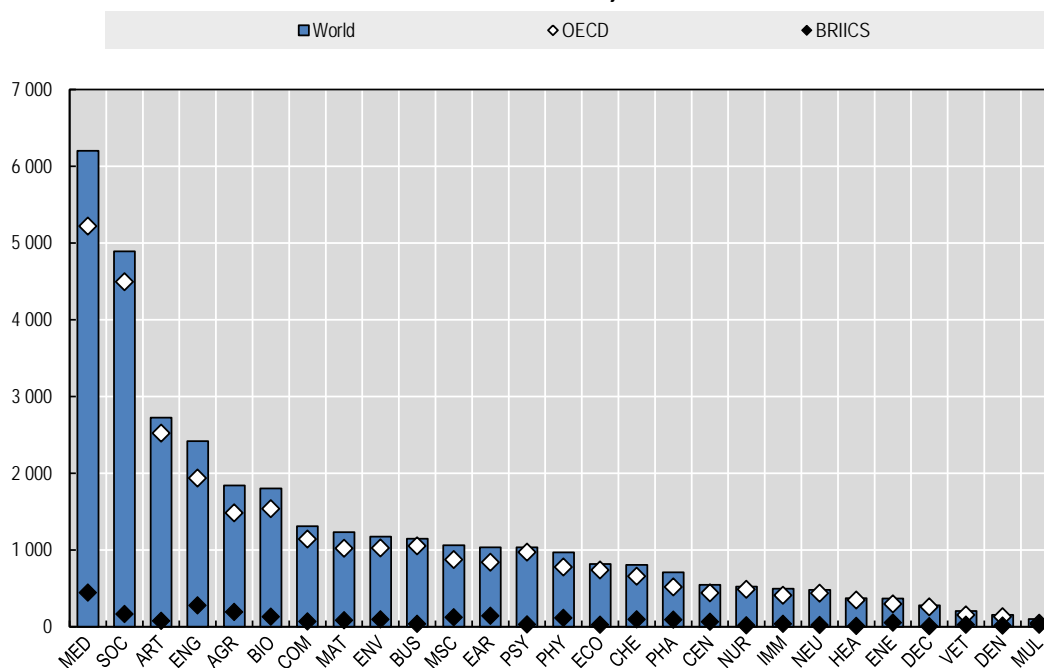
The OECD has its own classification of Fields of Science and Technology (FOS) which is applied for distributing R&D resources (OECD, 2002). With the recent revision of the *Frascati Manual*, this classification has been renamed into the OECD Fields of Research and Development (OECD, 2015) but has not been fundamentally changed. A future revision of this classification would aim at establishing closer linkages with available data-driven classifications for scientific production, technology areas used for patents, and more recent efforts to build science maps based on a range of similarity measures.

Figure 2.1. Scientific output by field, 2003-12

Number of documents, whole counts

See the *Reader's Guide* for the codes used for the fields captured in the All Science Journal Classification.[Download figure and data](#)**Figure 2.2. Scientific publishing by field, 2014**

Number of active journals

See the *Reader's Guide* for the codes used for the fields captured in the All Science Journal Classification.[Download figure and data](#)

2.2. Trends in indexed scientific publication output by fields

The total number of Scopus-indexed documents increased across all ASJC fields over the 2003-12 period. Seven fields exhibited double-digit average annual growth rates, namely: Multidisciplinary; Arts & Humanities; Social Sciences; Nursing; Decision Sciences; Energy; Economics, Econometrics & Finance (**Figure 2.3**).

Based on figures reported earlier in this chapter (**Figure 2.1**), these fast-growing fields accounted for relatively small shares of total indexed scientific output. Immunology & Microbiology represented the lowest growth rates over the period for any field with 4.5%, compared with 17.9% for Multidisciplinary.

Definitions and measurability

The analysis of the growth in the number of documents published by fields of science echoes some of the challenges arising from the use of available indices that were mentioned in the notes on definitions and measurability in **Chapter 1**. The faster relative growth in scientific production in journals associated to Multidisciplinary, Arts and Humanities and Social Sciences domains has to be carefully interpreted by taking into account a range of possible factors, including specific changes in indexing practices and not only a genuine increase in scientific production in these domains.

Scientific production in the social science and humanities domains, and to a lesser extent a number of health sciences, has been typically under-represented in bibliometric indices due to factors such as the centrality of books and their high citation rate; as well as the national orientation of social science literatures (Hicks, 1999), leading to more conscious efforts by bibliometric indices to extend indexing to these areas and thus extend the coverage and relevance of their services.

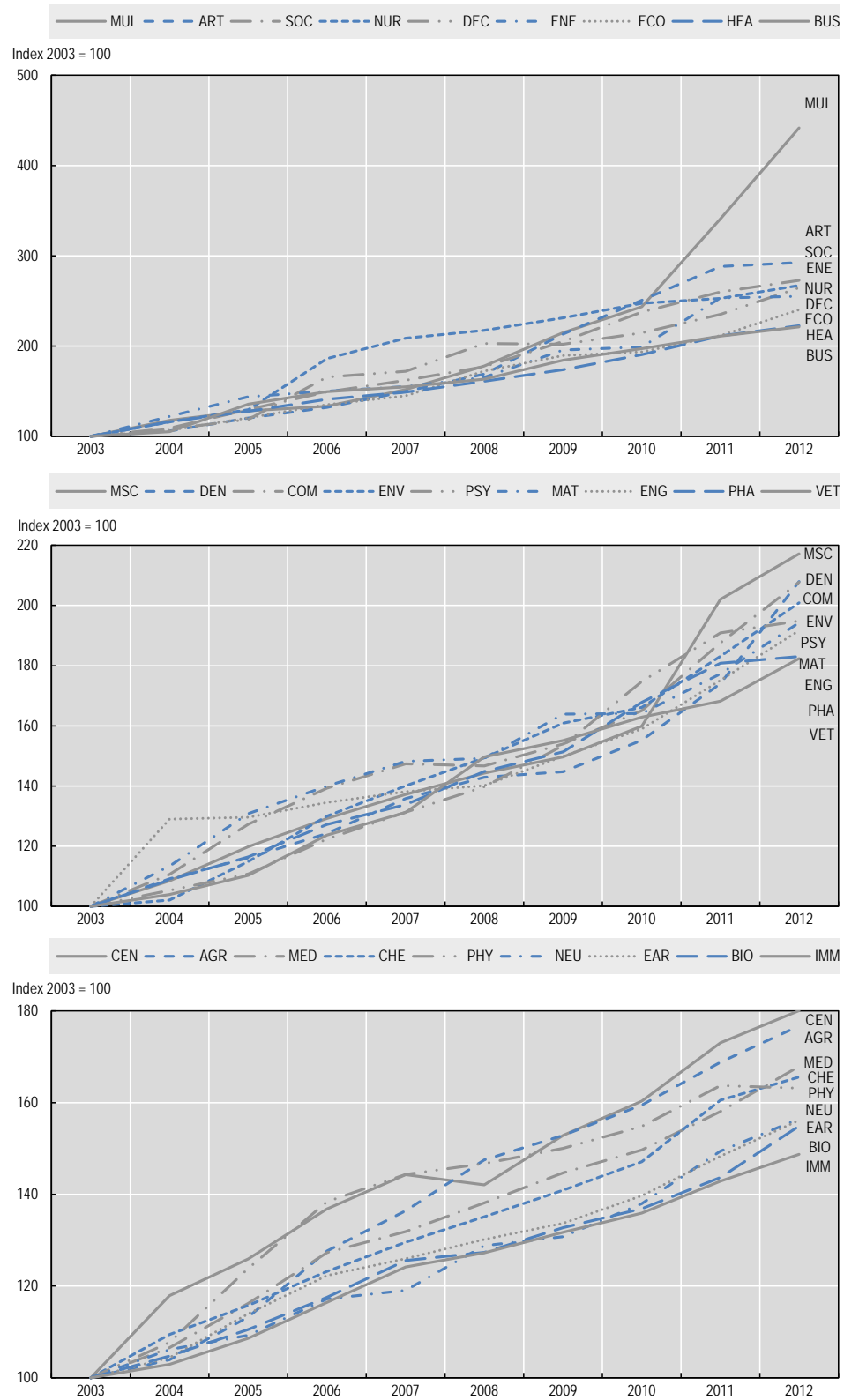
This effect may have therefore been more pronounced for countries whose scientific production in social science and humanities was not previously indexed, for example, as a result of it being published without an English abstract. There are abundant examples of journals in social sciences that have switched to publishing in English or providing selected information in that language. Although English clearly continues to be the preferred language of scientific communication, there are still plenty of disciplines within which researchers continue to publish in their native language as well.

Concerning the case of multidisciplinary research, it is important to recall that this is a classification applicable at the level of journals. The fact that scientific output in multidisciplinary journals is on the rise does not necessarily imply that science has become more multidisciplinary.

As a matter of fact, there appears to have been a surge in the number of new journals or newly indexed journals under this denomination. Inspection of the Scimago Journal Rank website (Scimago, 2015) indicates that a number of multidisciplinary journals are indexed for the first time towards the end of the reference period, joining established names with high publication volumes and citation impact such as the *Proceedings of the National Academy of Sciences (PNAS)*, *Nature* and *Science*. For example, the *World Applied Sciences Journal* is the fifth largest multidisciplinary journal in 2012 by volume, the first year it features on records, The *Australian Journal of Basic and Applied Sciences*, placed in eighth position in terms of total output, begins its records in 2010.

Figure 2.3. Trends in world output by field, 2003-12

Whole counts, index 100 = 2003, fields grouped by different scales



See the *Reader's Guide* for the codes used for the fields captured in the All Science Journal Classification.

[Download figure and data](#)

2.3. Field specialisation by country

The general pattern of scientific publishing by country is remarkably similar across most countries when examining the fields with highest levels of scientific publication output. Medicine accounts for the largest number of publications in nearly all countries, with exceptions found in China where Engineering was the dominant field, followed by Materials Science, Estonia (Agricultural & Biological Sciences), Indonesia (Agricultural & Biological Sciences), Korea (Engineering) and the Russian Federation (Physics & Astronomy, followed by Chemistry) (**Figure 2.4**). The second largest most common field is Biochemistry, Genetics & Molecular Biology (16 countries) and Agricultural & Biological Sciences (10 countries).

Given this large extent of similarity, it can be helpful to analyse relative specialisation through a relative activity index (see definitions box below). This indicator provides evidence of the fields in which a given country accounts for a relatively high share of scientific production, compared to the global norm.

For example, Brazil is highly specialised in Dentistry, as are Turkey and Sweden though to a much lesser extent. Very high levels of specialisation are also found in Iceland and Chile in the field of Earth and Planetary Sciences. Luxembourg is highly specialised in Economics and Finance, while the Russian Federation is strongly specialised in Physics (**Figure 2.5**).

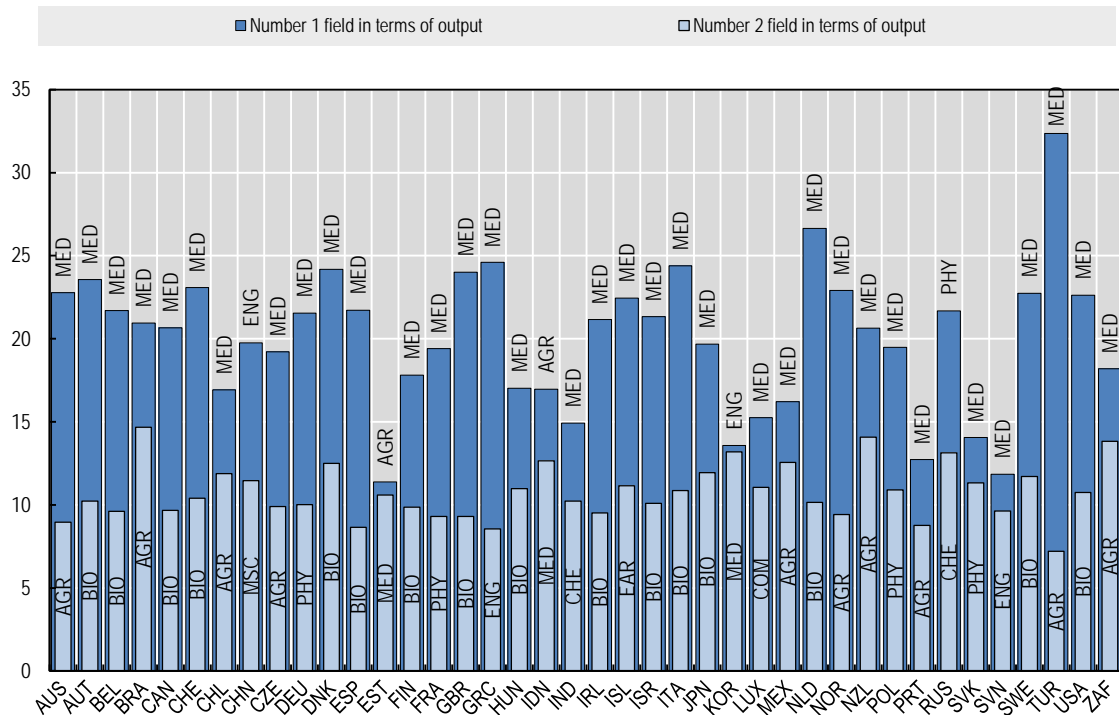
Most countries exhibit less marked specialisation levels, with their most specialised fields accounting for at most twice the field's global share. The United States is most (relatively) specialised in Psychology, while the United Kingdom appears to be specialised in Arts and Humanities as well as Social sciences. France's relative specialisation domains are found in Earth sciences and Mathematics.

Definitions and measurability

The Output indicator is a measure of scientific production. This indicator represents the total number of documents, published in scholarly journals indexed in Scopus (all document types are included).

The Relative Activity Index measures relative specialisation. The indicator is calculated by dividing a field's share of papers within a given country by the global share of that particular field (Glänzel, 2000). For a country with very similar distribution by fields to the world, specialisation values by fields should be very close to 1. A value of 2 for a given field and country indicates that the weight of that field in the country is two times as large as for the entire world. Figure 2.5 represents the two fields per country that score the highest on this measure.

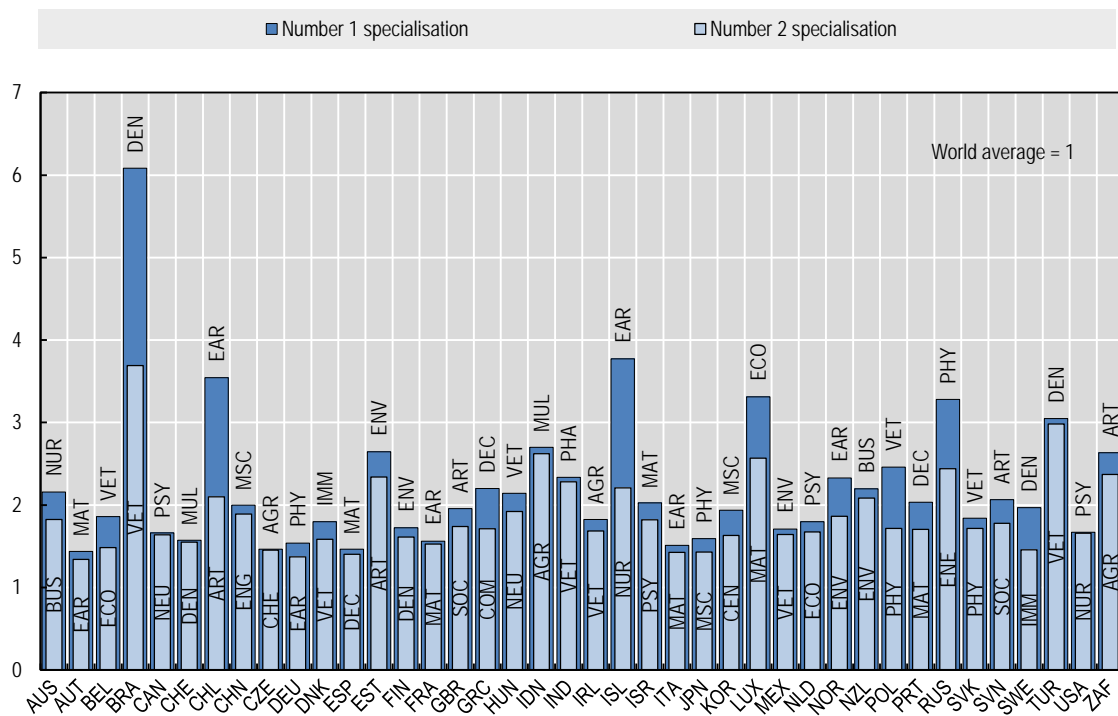
Figure 2.4. Top two largest fields of scientific output, by country, 2003-12
Share of field over total output within country, percentages



See the *Reader's Guide* for the codes used for the fields captured in the All Science Journal Classification.

[Download figure and data for all fields](#)

Figure 2.5. Top two fields of largest relative specialisation, by country, 2003-12
Relative activity indices for the top two fields per country, relative to world



See the *Reader's Guide* for the codes used for the fields captured in the All Science Journal Classification.

[Download figure and data for all fields](#)

2.4. Citations by fields

Different fields exhibit different citation patterns. Over the 2003-12 period, documents published in Multidisciplinary journals attained the largest average number of citations per document within the Scopus database (**Figure 2.6**). This may be related to the prestige and media attention that some of these journals attract, and the possibility that readers from different fields identify relevant sources which they may not have otherwise identified and cited. It is however possible to note that within this “field”, there are very marked differences concerning the country affiliation of authors. This may be related to differences in the types of multidisciplinary journals covered within Scopus, ranging from household names to newly indexed journals.

Neuroscience is the second largest field by number of citations per document, closely followed by Biochemistry, Genetics & Molecular Biology. Arts & Humanities is the field with the least number of citations per document.

It takes a very different number of citations to be characterised as highly cited in different domains. Focusing on the 10% most cited documents within each field and by OECD authors, it is possible to note that publications in this group attract a very different average number of citations, from close to 200 citations per document in Multidisciplinary journals, nearly 70 for Immunology & Microbiology and Biochemistry, Genetics & Molecular Biology to Social Sciences, with 28.8 citations per document (**Figure 2.7**).

Definitions and measurability

This indicator is defined by the number of citations recorded per document. This publication works with so called “citable documents”, i.e. focusing on articles, reviews and conference papers published within the reference period. All citations included in the Scopus database are covered. The number of citations received by those documents is considered over the same reference period and all types of documents covered in Scopus are considered. On average, only about 50% of all documents published over the 2003-12 period have been cited over that same period.

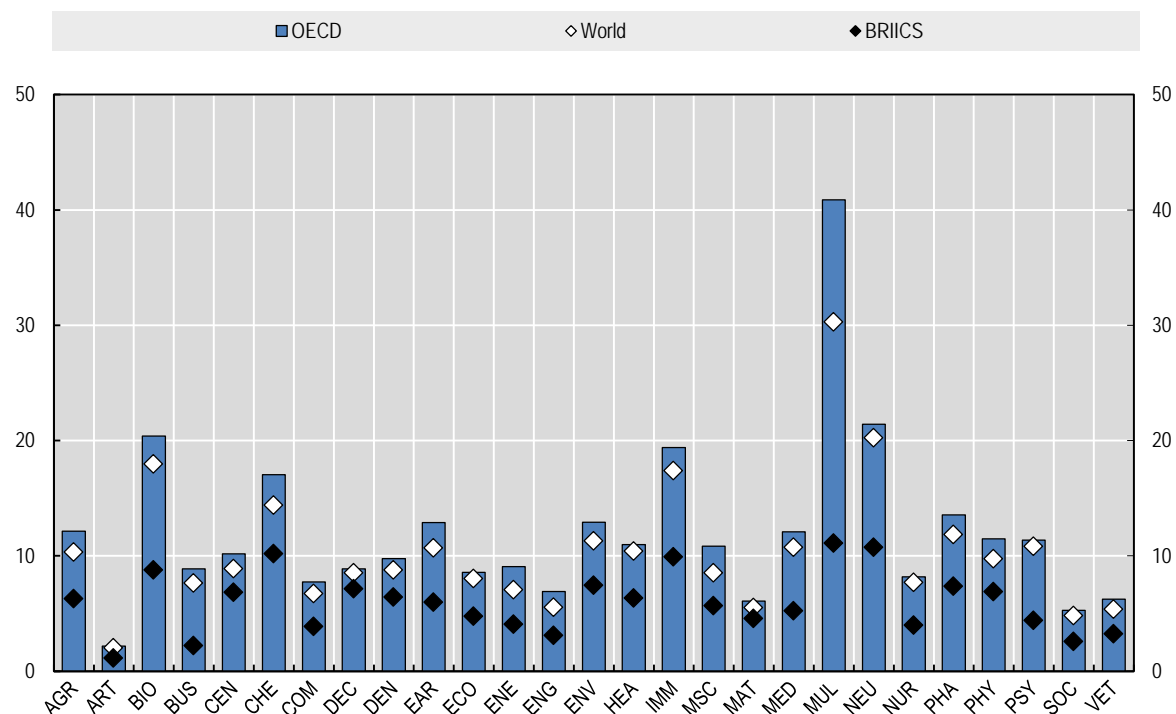
The outcome of citation count depends on a number of factors. Documents citing the reference document may or may not be included within Scopus. Citations in most types of books, working papers, etc. are not counted. To the extent that scientific discourse in different fields may rely differently on those vehicles, some field-based biases may appear when counting citations. On the other hand, by counting only citations within journals that are deemed to meet basic indexing requirements, a minimum quality standard is applied.

Field specific citation norms may be important explanatory factors, for example concerning practices such as self-citation, number of co-authors, number and length of documents per author, incidence of review articles, etc. In this publication, self-citations are included in estimates. In some fields, it may take on average longer for documents to be cited but they may be in turn cited over a longer time period (i.e. fast versus slow burners).

These systematic differences across fields call for the established practice of normalising citation counts by fields in order to derive more robust measures of citation impact.

Figure 2.6. Citations per document by field, 2003-12

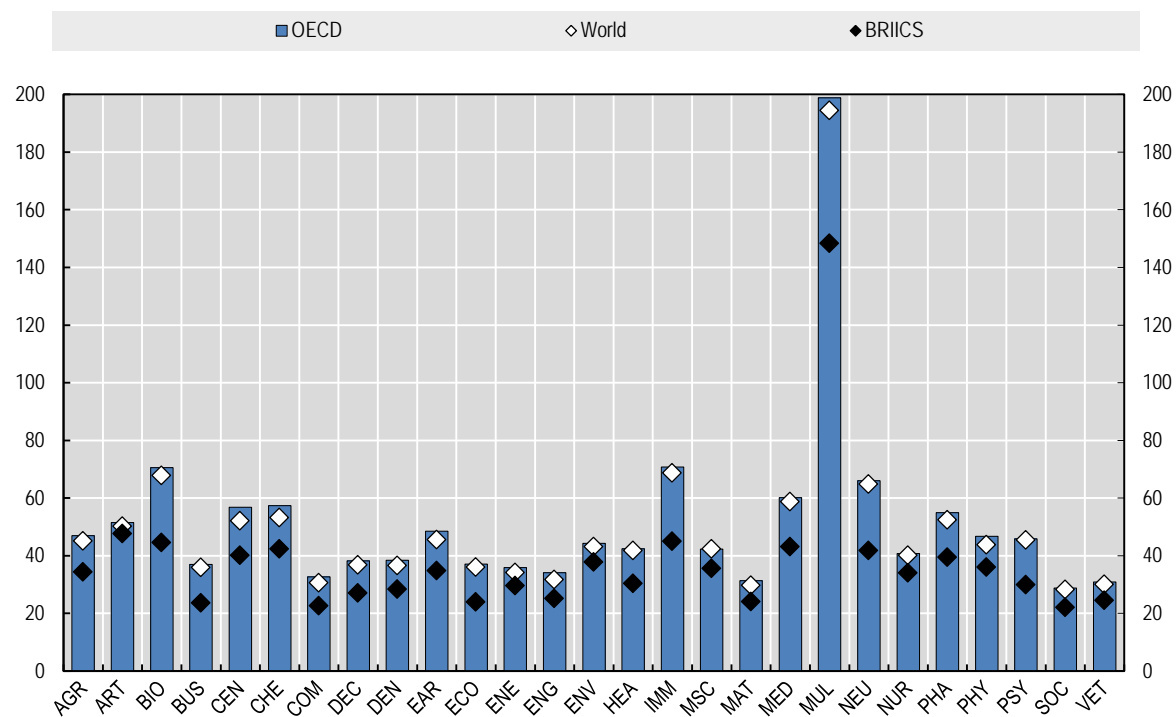
Average number of citations



See the Reader's Guide for the codes used for the fields captured in the All Science Journal Classification.

[Download figure and data](#)**Figure 2.7. Citations per document - among the top 10% most cited - by field, 2003-12**

Average number of citations



See the Reader's Guide for the codes used for the fields captured in the All Science Journal Classification.

[Download figure and data](#)

2.5. Fields with highest citation impact by country

Different countries excel in different fields, as implied by the number of citations received by their attributed documents, compared to those by authors working in the same areas and affiliated to institutions in other countries. Among the group of OECD countries, documents published in Multidisciplinary journals attain on average the highest number of citations (**Figure 2.8**). Since this is not very informative about the country's (on average) highest performing fields, relative to the field's norm, it can be useful to consider the top two fields within country. From this perspective, Medicine, Nursing and other health related fields stand out as the fields in which authors from several OECD countries attain the largest normalised citation impact compared to the world average.

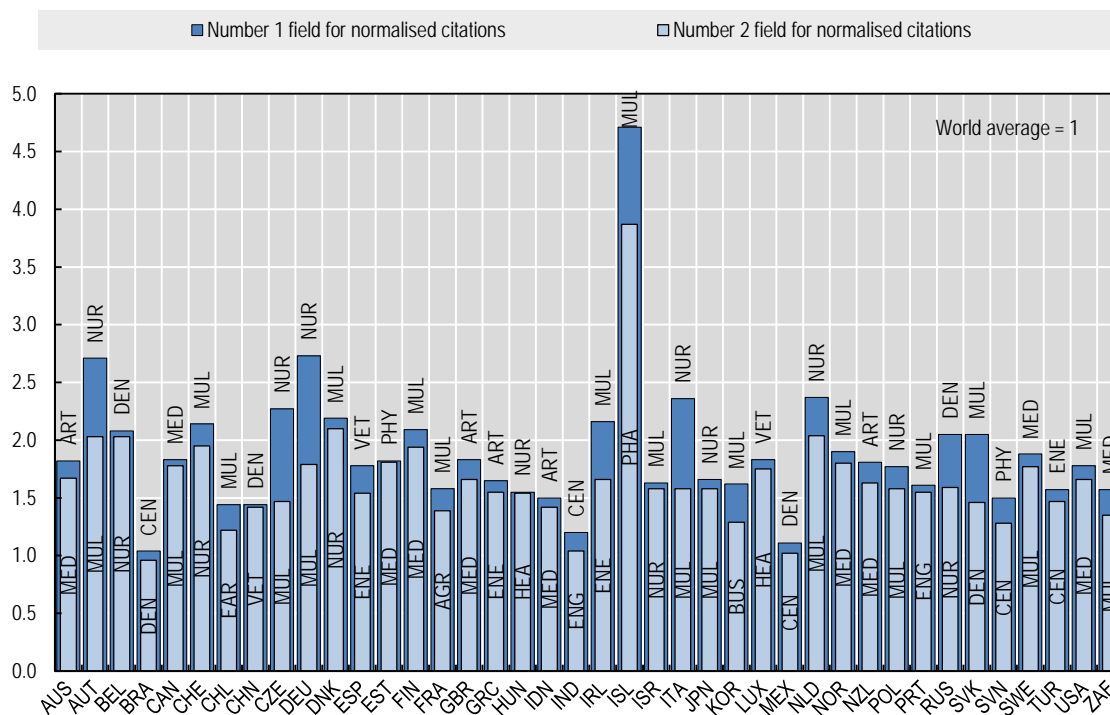
Among OECD and BRIICS countries, the distribution of top cited publications, i.e. the 10% top cited publications within each field, provides an indicator of aggregate scientific excellence across fields (**Figure 2.9**). The United States accounts for the largest number of top-cited or high impact publications across all disciplines. This pre-eminent role is particularly marked in the life sciences and most health-related and social science fields. US-authored publications account for a relatively lower percentage of high-impact publications in a number of basic science domains outside the life sciences and engineering. The United Kingdom is the second largest producer of top-cited publications, especially in the fields where the United States accounts for a large share, but also excels in the Earth, Environmental, Agricultural and Veterinary sciences. China is the second largest producer of top cited publications in Materials Science, Chemistry, Engineering, Computer Science and Chemical Engineering, Energy, Decision Sciences and Mathematics. Germany has the second largest share in Physics and Astronomy, and is the third largest producer of high-impact publications across most fields. Other countries in the top 4 include Japan – a significant player in Materials Science, Chemistry and Biochemistry – Australia, Brazil, France, India and Spain. Canada is present in several fields, often those where the United States has the most prominent position.

Definitions and measurability

Based on the differences in citation patterns identified in the previous section, some field-based normalisation is required to investigate citation impacts. The measure of normalised citation impact shows the relationship of the unit's average citation to the world's average, the unit is the country. The indicator is derived as the ratio between the average number of citations received by the documents published by authors affiliated to an institution in a given country and the world average of citations, over the same time period, by document type and subject area. Its value indicates how many times above (or below) the document has been cited above (or below) world average. For example, a value of 0.8 means the country is cited 20% below world average in that field and 1.3 means the country is cited 30% above average for that domain.

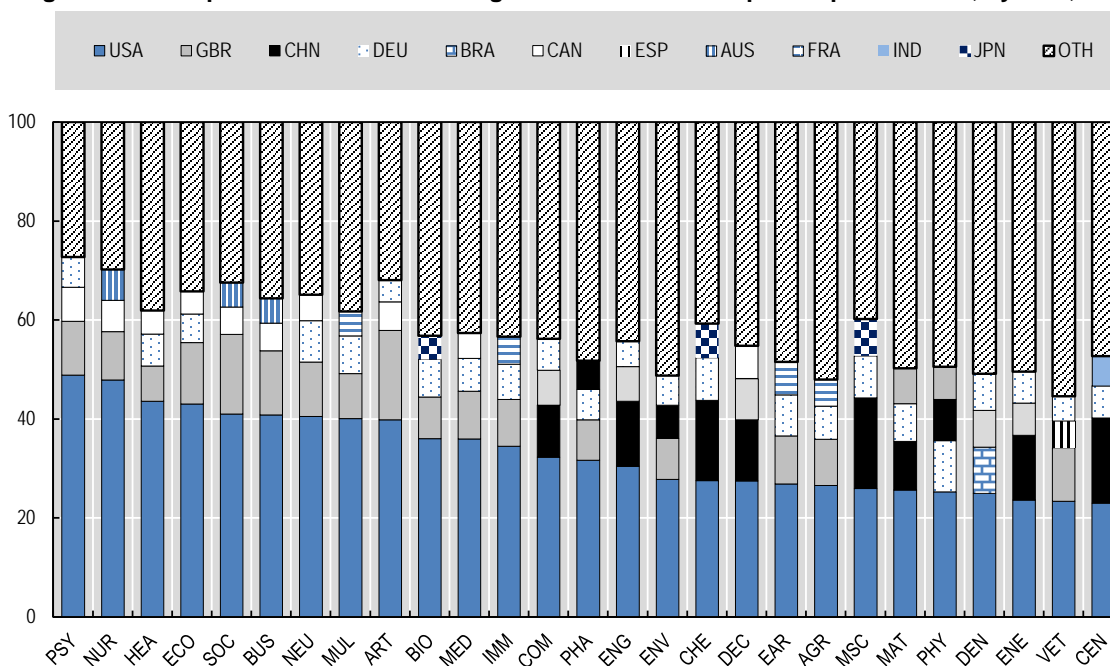
The normalisation of citation values is item-oriented, i.e. carried out at the level of the individual document. If an article has been published in a journal allocated to more than one subject area, a mean value of the areas is calculated. One potential caveat when considering this type of indicator at the country level is the potential low number of publications from which average citations are calculated. This indicator does not incorporate any information about the "quality" of the citation.

The top 10% most cited documents is an indicator of "excellence". This rate indicates the amount (in percentages) of a unit's scientific output that is included into the group of the 10% of the most cited papers in their respective scientific fields. It is a measure of high quality of research output of a unit, in this case the country.

Figure 2.8. Top normalised citation, top two fields by country, 2003-12

See the Reader's Guide for the codes used for the fields captured in the All Science Journal Classification.

[Download figure and data for all fields](#)

Figure 2.9. Top 4 countries with the largest number of 10% top cited publications, by field, 2003-12

See the Reader's Guide for the codes used for the fields captured in the All Science Journal Classification.

[Download figure and data for all fields](#)

2.6. International collaboration by fields

International collaboration by field data shows that work in some fields is characterised by higher levels of international collaboration than others. The highest incidence of international collaboration is found in documents published in journals classified into Earth & Planetary Sciences, Multidisciplinary and Physics (**Figure 2.10**).

Low levels of international collaboration are found in the Arts and Humanities, a number of health related fields and the Social Sciences.

Overall, across all disciplines, international collaboration rates are higher among OECD countries than for the world average. In most instances, international collaboration was higher in OECD countries than in BRIICS countries.

Despite differences in the average level of collaboration by country, with few exceptions, the same fields (Earth Sciences and Multidisciplinary) exhibit the largest degree of international collaboration within individual economies. The top two international collaboration fields by country frequency data confirms that the top fields in terms of collaboration are Multidisciplinary (number 1 for 26 countries), Earth & Planetary Sciences (number 1 for 9 countries). The second most collaborative fields are Physics & Astronomy (number 2 for 14 countries), Earth & Planetary Sciences (12 countries) and again Multidisciplinary (6 countries) (**Figure 2.11**).

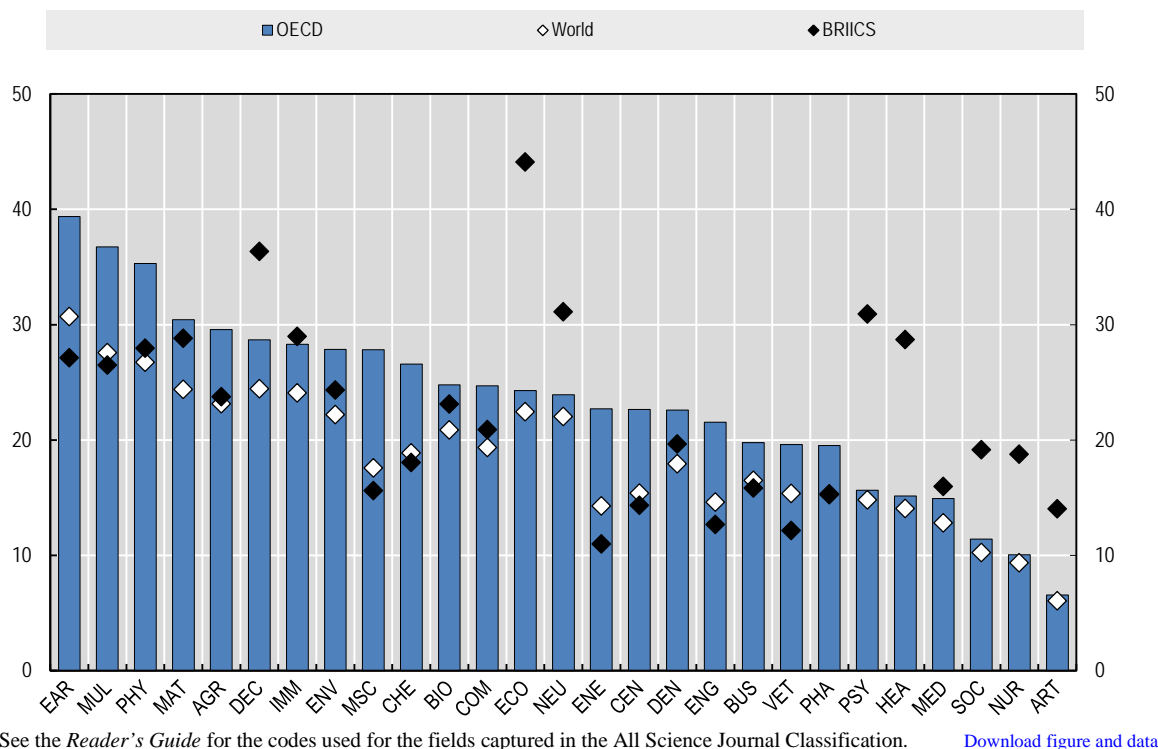
International collaboration in the top fields ranges from 35.6% to 92.2%, India for Economics, Econometrics & Finance and the Russian Federation for Dentistry.

Definitions and measurability

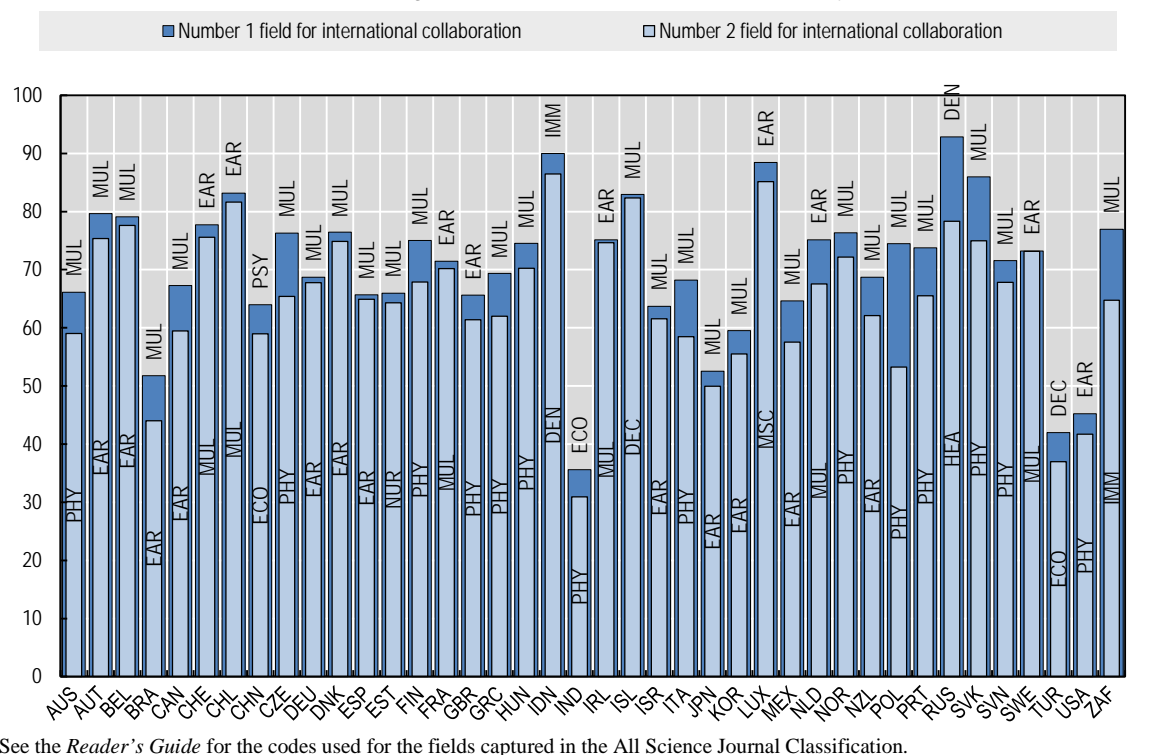
The indicator of International collaboration is based on the proportion of documents involving institutional affiliations with other countries or economies, as a proportion of documents attributed to authors with an affiliation in the reference economy. Single-authored documents with multiple affiliations in different countries count as institutional international collaboration.

Figure 2.10. International collaboration by field, 2003-12

As a percentage of publications within field and country

**Figure 2.11. Top two international collaboration fields by country, 2003-12**

As a percentage of publications within field and country



2.7. Open access by fields

About 12% of the 22 283 active journals covered by Scopus are available on an open access (OA) basis. As it is the case for the total number of documents and total number of active journals, Medicine is the field with the largest absolute number of OA journals. This field covers about 38% of all OA journals (**Figure 2.12**).

The incidence of open access publishing is more prevalent in the BRIICS countries (Miguel et al., 2011). As noted in Chapter 1, nearly a quarter of journals published in the BRIICS countries are open access, compared to 8% for journals published in the OECD area. The share of OA publishing by field is higher in the BRIICS group of countries than in the OECD area except for Earth & Planetary Sciences and Energy (**Figure 2.13**).

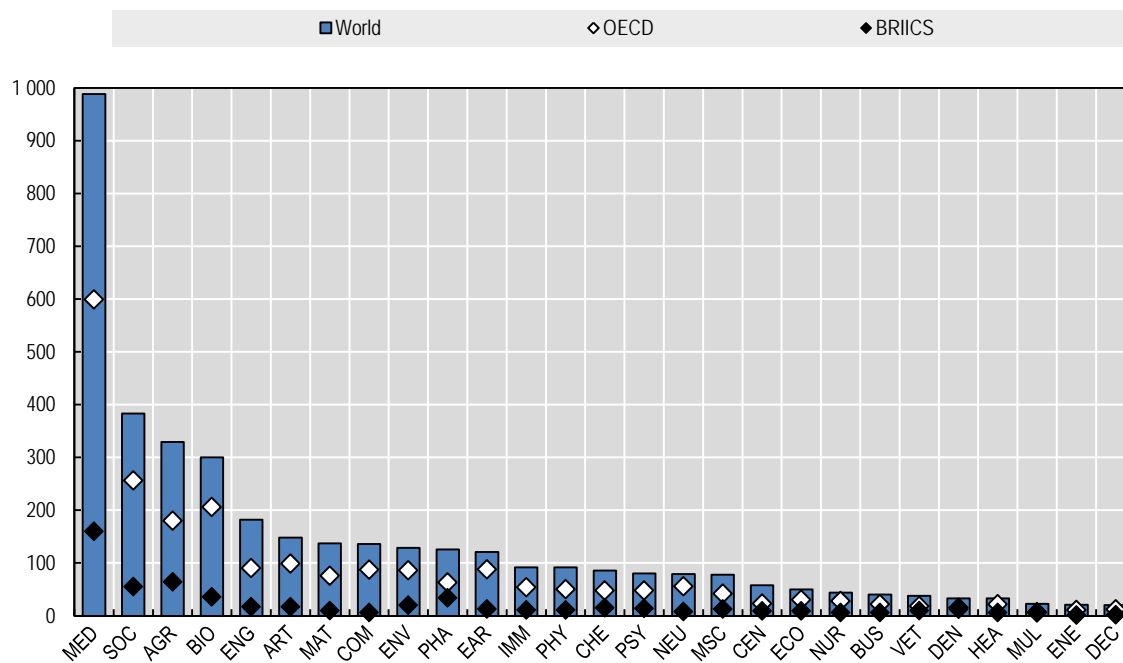
The United States and the United Kingdom are the only two countries with OA journals in all 27 fields. Brazil has OA titles in every field except Computer Science and Energy.

Definitions and measurability

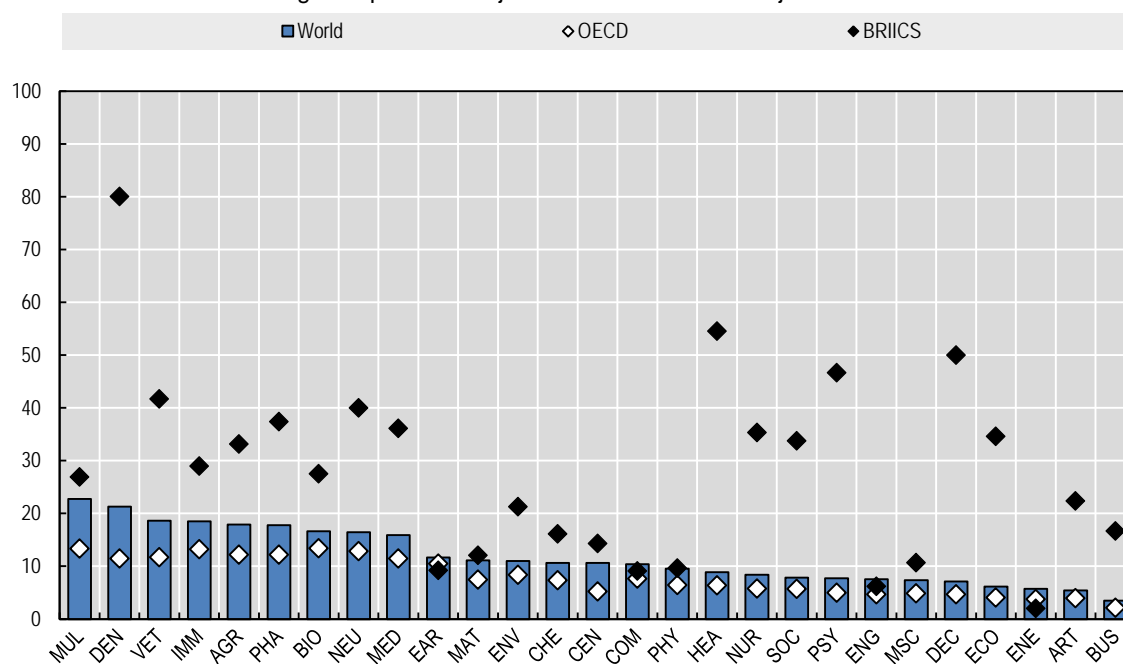
See the Definitions and measurability box in Chapter 1, section 8 on Scientific publishing and open access on open access and the Scopus database.

Figure 2.12. Open access journals by field, 2014

Number of active journals

See the *Reader's Guide* for the codes used for the fields captured in the All Science Journal Classification.[Download figure and data](#)**Figure 2.13. Share of open access journals by field, 2014**

Percentage of open access journals in zone over total journals in zone

See the *Reader's Guide* for the codes used for the fields captured in the All Science Journal Classification.[Download figure and data](#)

Chapter notes and references

- Glänzel, W. (2000), “Science in Scandinavia: a bibliometric approach”, *Scientometrics* 2000, Vol. 48, No. 2, pp. 121–150.
- Hicks, D. (1999), “The Difficulty of Achieving Full Coverage of International Social Science Literature and the Bibliometric Consequences”, *Scientometrics*, Vol. 44, No. 2 pp. 93-215.
- López-Illescas, C., E.C.M. Noyons, M.S. Visser, F. Moya-Anegón and H.F. Moed (2009), “Expansion of scientific journal categories using reference analysis: How can it be done and does it make a difference?” *Scientometrics*. Vol. 79, No. 3, pp. 473-490, <http://dx.doi.org/10.1007/s11192-007-1975-6>.
- Miguel, S., Z. Chinchilla-Rodríguez and F. Moya-Anegón (2011), “Open Access and Scopus: A New Approach to Scientific From the Standpoint of Access”, *Journal of the American Society for Information Science and Technology*, Vol. 62, No. 6, pp.1130-1145, <http://dx.doi.org/10.1002/asi.21532>.
- OECD (2015), *Frascati Manual 2015: Guidelines for Collecting and Reporting Data on Research and Experimental Development, The Measurement of Scientific, Technological and Innovation Activities*, OECD Publishing, Paris, <http://dx.doi.org/10.1787/9789264239012-en>.
- OECD (2002), *Frascati Manual 2002: Proposed Standard Practice for Surveys on Research and Experimental Development, The Measurement of Scientific and Technological Activities*, OECD Publishing, Paris, <http://dx.doi.org/10.1787/9789264199040-en>.
- SCImago (2007), SJR — SCImago Journal & Country Rank. Retrieved June 01, 2015, from: http://www.scimagojr.com/journalrank.php?area=1000&category=0&country=all&year=2012&order=item&min=0&min_type=cd

3. SCIENTIFIC PUBLICATION INDICATORS BY SECTORS

1. The contribution of sectors to scientific production
 2. Scientific publication output trends by sector
 3. Scientific production by sector and field
 4. Collaboration patterns by sector
 5. Cross-sector collaboration
 6. Scientific excellence and citation impact
 7. High “impact” institutions
- Chapter notes and references

3.1. The contribution of sectors to scientific production

The analysis of the affiliations of authors featured in scholarly journals reveals that most documents published over the 2003-12 period can be attributed to authors in “higher education” institutions. In the OECD area, this sector’s share is close to 70%, while in the BRIICS, the share is close to 80%. “Government” and “health” sectors account for most of the remaining documents, with some notable differences between OECD and the BRIICS. For the former, government and health each account for 13% of all documents. The government sector is relatively more important in the BRIICS (17%) while health accounts for only 3%. The “private” sector accounts for a very small fraction of published documents, 3% for OECD and 1% for the BRIICS (**Figure 3.1**). Comparing these results with OECD data on how R&D is distributed by sectors in the OECD area (noting different sectoral definitions explained in the box below), this confirms that scholarly publishing is a relatively marginal activity for business enterprises (which account for nearly 70% of R&D) and therefore a less meaningful measure of research output (Moya-Anegón et al., 2014).

The sectoral distribution of scientific scholarly publication output by country reveals some significant differences. Turkey, Estonia and Chile have the largest share of documents in the higher education sector, contrasting with the case of the Russian Federation, France and Spain (**Figure 3.2**). The Russian Federation has the largest government sector share followed by France and the Slovak Republic, while the lowest shares are found in Turkey, Israel and Sweden. The largest shares for health are found in Belgium, Spain and Switzerland, while the lowest shares were in Hungary, China and the Slovak Republic. Brazil exhibits the largest share of documents accounted for the private sector (close to 40%), followed by Switzerland.

Overall, it may be argued that these figures represent not only how (principally public) research resources are organised and allocated but also combined country and sector-specific factors and incentives that influence the propensity to publish in journals listed in Scopus. In the case of smaller countries, the presence of individual institutions with particular characteristics can greatly influence the overall distribution.

Definitions and measurability

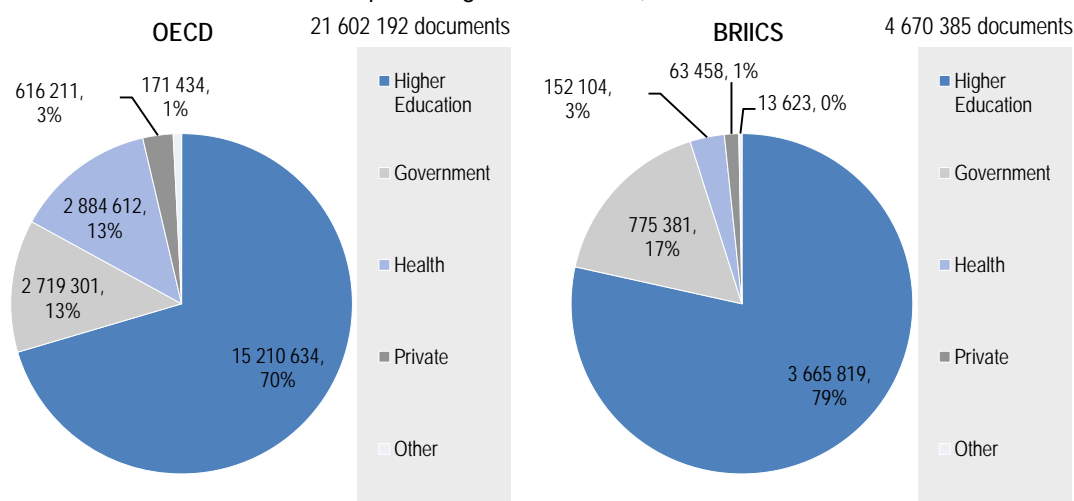
The standardisation methods used by [SCImago](#) include extensive disambiguation of institution names through the institutional affiliation of documents included in Scopus. The identification task includes the consolidation or separation of institutions, as well as dealing with name changes. The objective of the standardisation is twofold, to define and identify the institutions by drawing up a list of research institutions where every institution is correctly identified and defined, and to attribute publications and citations to each institution. To accomplish this, SCImago takes into account the institutional affiliation of each author in the field ‘affiliation’ of the Scopus database. A hybrid system comprising manual and automatic elements is applied to assign affiliations to one or more institutions, as appropriate.

Once the institutions have been correctly identified they are grouped into five “major” sectors: higher education, government, health, private and other. This classification system has notional similarities to that used in the OECD *Frascati Manual* (FM) (OECD, 2015) institutional classification (Higher education, Business enterprises, Government and Private non-profit), but differs in that the latter does not identify separately a “health” sector. Furthermore, the allocation criteria applied by Scimago differ from those applied by national statisticians in charge of producing R&D data. The classification used in the FM is a derivation of the institutional classification system proposed in the System of National Accounts (SNA) (EC et al., 2009), with the addition of the Higher Education sector in order to meet policy user interest in securing disaggregated R&D figures for this subset of R&D performers. The use of a health sector in bibliometrics is quite common and explained by the large volume of indexed medical science publication. This can however lead to some comparability problems because health-focused organisations can be formally part of higher education, government or private sectors. Source: OECD, based on OECD(2015) and <http://www.scimagoir.com/methodology.php>

In order to get around this comparability issue, developing an organisation name dictionary could allow to categorise institutions not only by sector but also by function. If such a dictionary were broadly and openly available for all countries it would be easier to classify organisations to serve different purposes.

Figure 3.1. Distribution of documents by sector, 2003-12

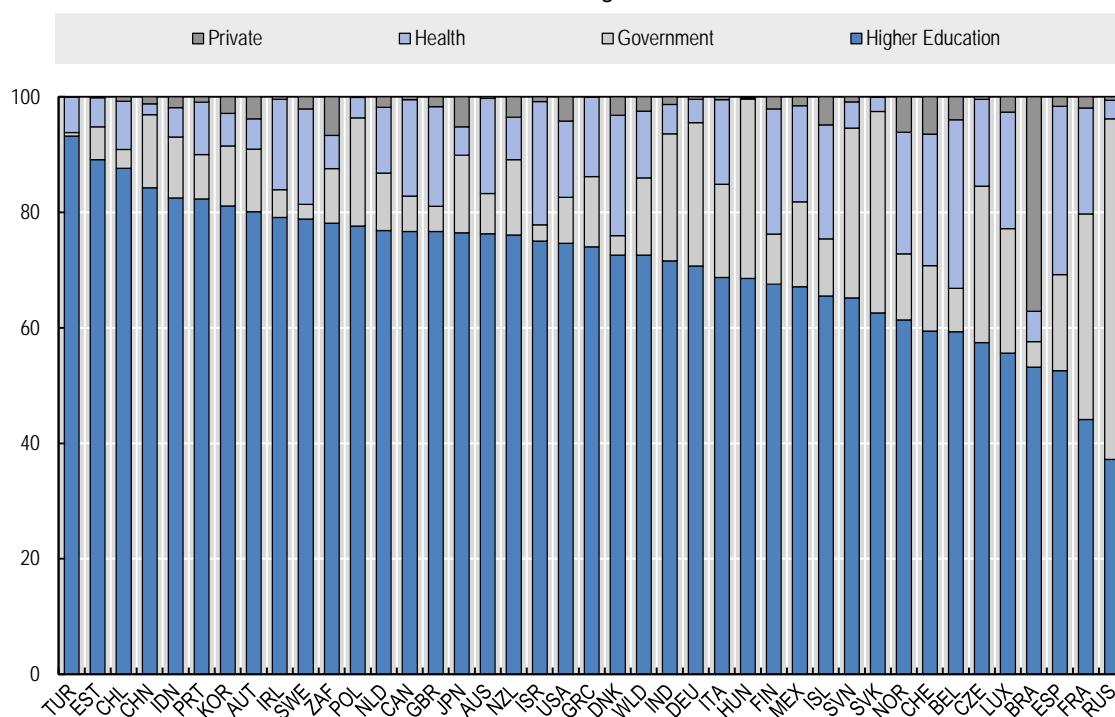
Number and percentage of documents, whole counts



[Download figure and data](#)

Figure 3.2. Distribution of scientific output by sector, by country, 2003-12

Percentages



[Download figure and data](#)

3.2. Scientific publication output trends by sector

Different sectors have contributed to a varying extent to the recorded increase in numbers of scientific publications over the past decade. Looking at the combined experience of OECD and BRIICS countries, publication growth over the 2003-12 period is positive for all sectors (**Figure 3.3**), with the higher education sector growing on average at 9.1% per year, the largest growth over the period. This compares to 6.6% and 6.1% in the government and health sectors respectively, and 3.6% in the private sector. As a result, and in contrast with a relatively more stable sectoral distribution of R&D expenditures (GERD), the higher education sector progressively increased its share of publications over the period from 69% to 74%. Over this period, the share of GERD in the OECD area accounted for by the Higher Education sector, using the definition in the OECD Frascati Manual, has also increased but by a more reduced margin, reaching 18%. A complete picture on the distribution of GERD by performing sector among all BRIICS countries is not available, but in the case of the largest R&D performer in this group, i.e. China, the share accounted for by the higher education sector has a slightly lower importance and has been slightly declining due to the fast growth of the business enterprise sector.

Despite the different sources and sectoral classifications used, it is possible to note that there is a strong relationship between the growth in R&D performance (9.3% annual average) and publication output (11.1%) over the 2003-12 period (**Figure 3.4**). For most countries the average annual growth rates for scientific production in the higher education sector exceeded those for Higher education R&D (HERD), and the overall cross-country correlation between the two variables is close to 85%.

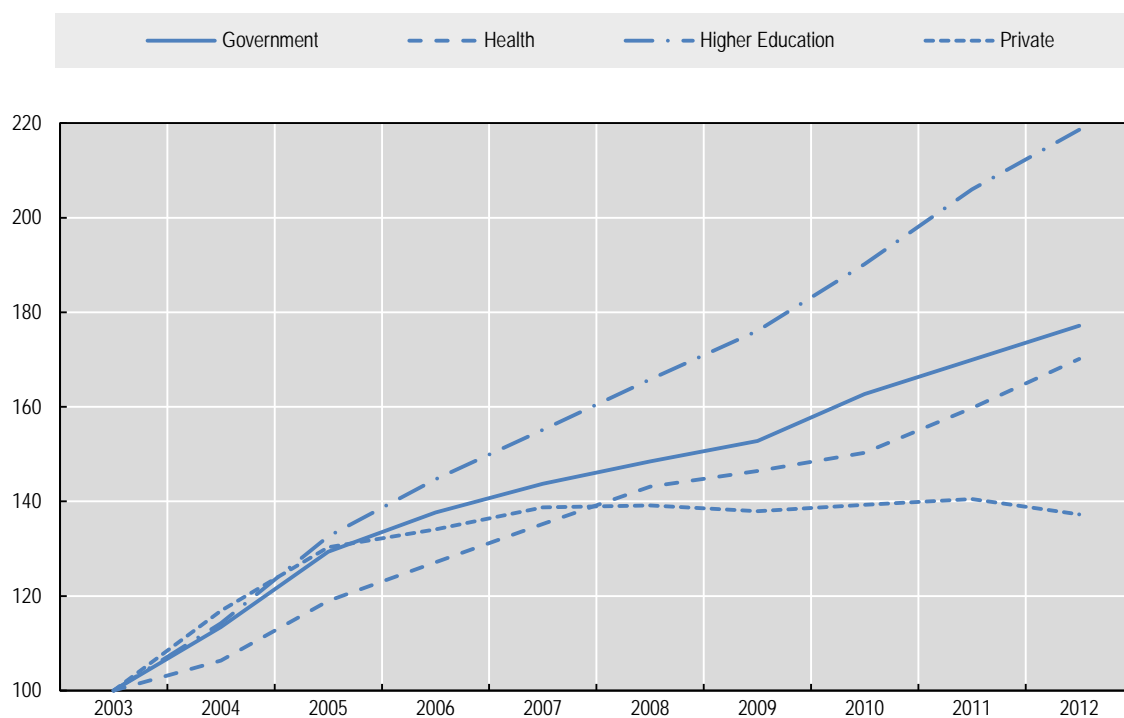
The reasons for the more rapid increase in publications in the higher education sector may be related to the role of incentives, since publishing in journals featured in major indices such as Scopus has become an increasingly factor in promotion and grant award decisions within this sector. Within higher education, it appears that funding levels have had some form of impact on publication output.

Within the so-called private sector, trends relate to a relatively small fraction of scientific publication output. Within this sector, it is however important to note that output appears to have stalled since 2007, coinciding with the onset of the economic and financial crisis.

Definitions and measurability

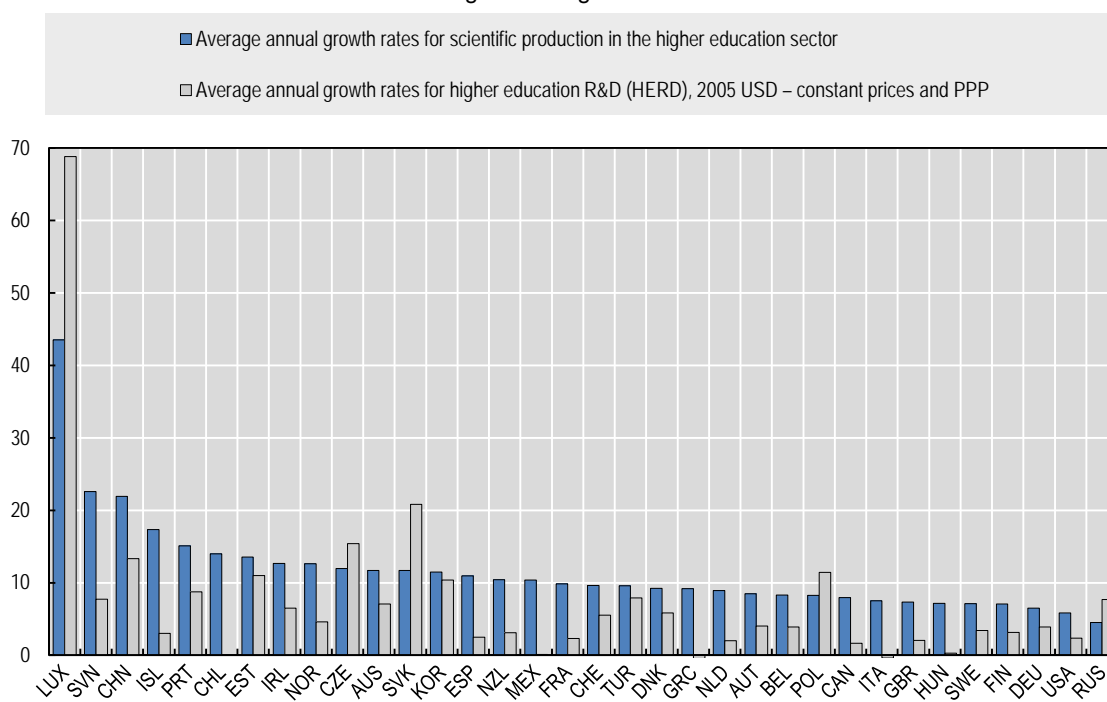
See the Definitions and measurability box in Section 1 on *Scientific production by sector and field* for information about the sector breakdown and institutional classification.

Figure 3.3. Trends in scientific publication output by sector, OECD and BRICS, 2003-12
Whole counts, index 100 = 2003



[Download figure and data](#)

Figure 3.4. Changes in the scientific publication output and R&D in higher education, 2003-12
Average annual growth rates



[Download figure and data](#)

3.3. Scientific production by sector and field

The importance of higher education as contributor to scientific production is not only widespread across countries but also across different scientific domains. The higher education sector accounts for at least 60% of all publications in each field (**Figure 3.5**).

Institutions in “government” are found to have a significant role in the fields of Physics & Astronomy, Agricultural & Biological Sciences, Earth & Planetary Sciences, Environmental Science, Materials Science and Energy, with shares over or close to 20%.

As expected, “health” institutions play significant roles within Medicine, Biochemistry, Genetics & Molecular Biology, Immunology & Microbiology, Neuroscience, Pharmacology, Toxicology & Pharmaceutics, Nursing and Health Professions, as well as Dentistry. These are domains in which institutions classified into government play a very small role, although some of the health institutions include some which are formally part of government, as noted in the examples provided at the end of this chapter.

The presence of so-called “private” institutions is globally small but also highly concentrated within Energy, Pharmacy, Engineering and Computer Science.

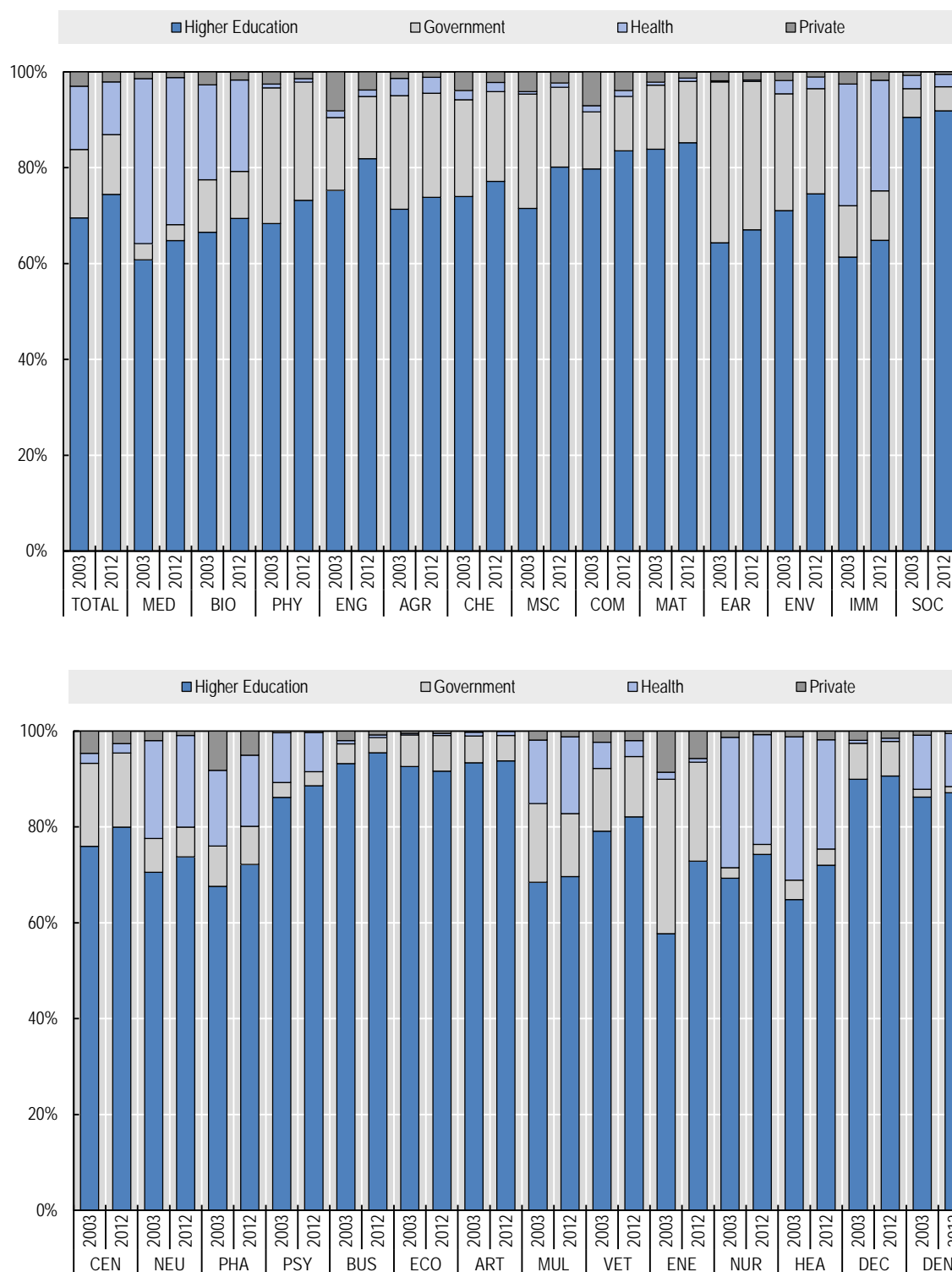
Over the 2003-2012 period, the share of scientific output accounted for by higher education is observed to increase across the entire spectrum of disciplines. This increase is particularly marked in the fields of Engineering and Energy.

Table 3.1. below provides the codes that have been used in the figures of this section.

Table 3.1. All Science and Journal Classification (ASJC) codes

Code	Field	Code	Field
ART	Arts & Humanities	HEA	Health Professions
AGR	Agricultural & Biological Sciences	IMM	Immunology & Microbiology
BIO	Biochemistry, Genetics & Molecular Biology	MSC	Materials Science
BUS	Business, Management & Accounting	MAT	Mathematics
CHE	Chemistry	MED	Medicine
CEN	Chemical Engineering	MUL	Multidisciplinary
COM	Computer Science	NEU	Neuroscience
DEC	Decision Sciences	NUR	Nursing
DEN	Dentistry	PHA	Pharmacology, Toxicology & Pharmaceutics
EAR	Earth & Planetary Sciences	PHY	Physics & Astronomy
ECO	Economics, Econometrics & Finance	PSY	Psychology
ENE	Energy	SOC	Social Sciences
ENG	Engineering	VET	Veterinary
ENV	Environmental Science		

Figure 3.5. Distribution of scientific publication output by field, OECD and BRIICS, 2003 and 2012
Percentage share of documents by sector



[Download figure and data](#)

3.4. Collaboration patterns by sector

This section investigates the patterns of scientific collaboration within sectors. For presentational reasons, the focus of the figures presented is on collaborations involving higher education, but comparable results for other sectors are available in the online edition of this publication. With the exception of four countries, institutional collaboration is more common than no collaboration in publications attributed to higher education (**Figure 3.6**). On average, 59% of publications involve some form of collaboration.

Within the ensemble of higher education in the OECD area and the BRIICS countries, 31% of documents exhibit some form of international collaboration – including those documents where there is also collaboration with other domestic institutions.

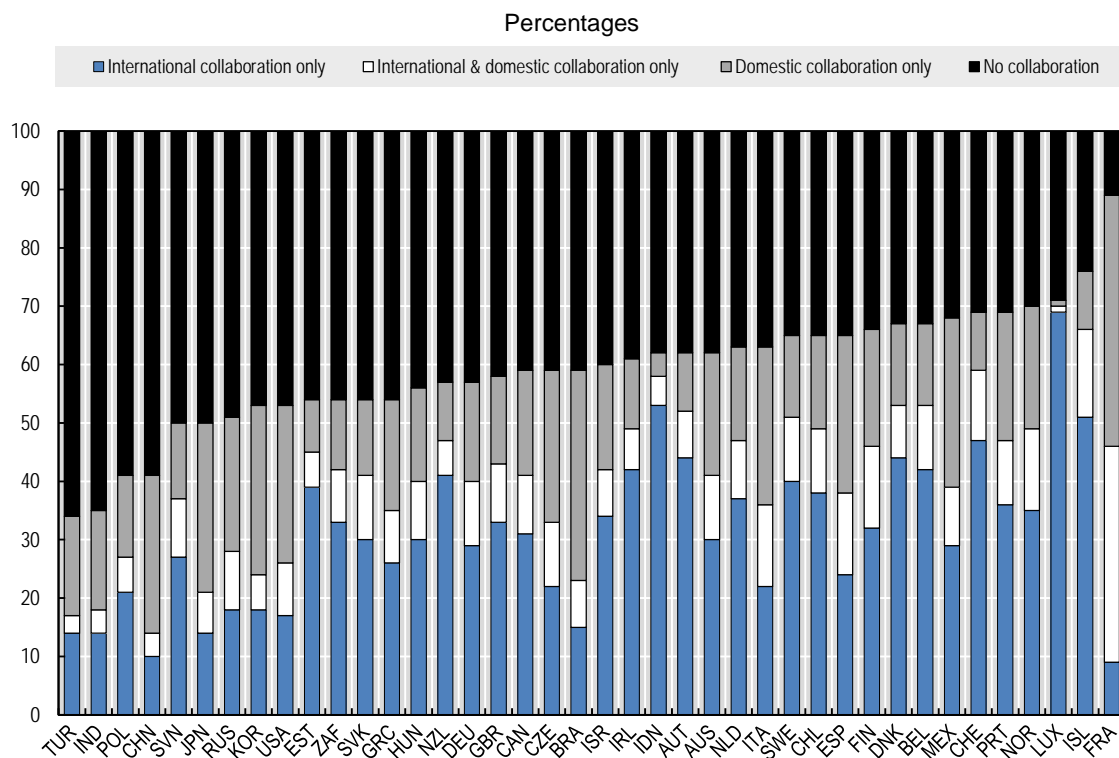
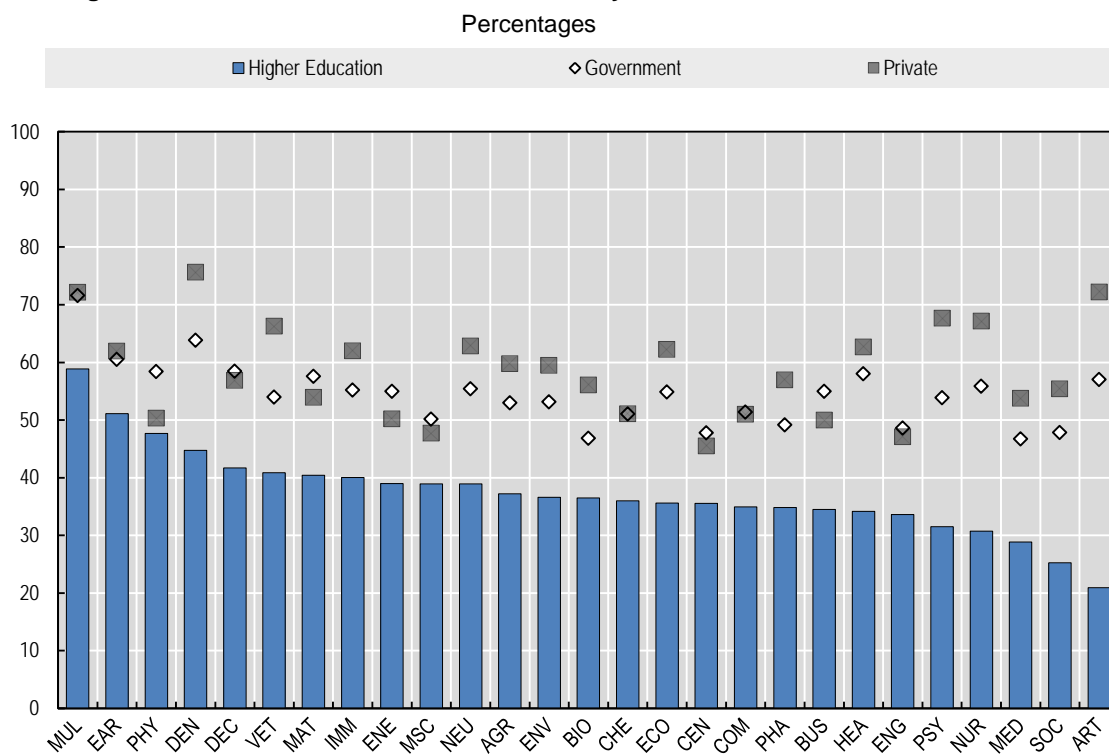
Figure 3.7 reports on overall international collaboration rates by sector and field of science within OECD countries over the 2003-12 period. International collaboration rates, as reflected by scientific publications, appear to be systematically lower within higher education, compared with government and private sectors. Does this imply that there is less international collaboration in research within this sector, compared to others? Given that incentives for producing indexable output are higher in the higher education sector, the answer to the question depends on how large are international collaboration rates in other sectors for projects that lead to research that does not result in scientific publications in journals indexed in Scopus.

Within higher education, the highest international collaboration rates are found within Multidisciplinary, Earth and Planetary Sciences and Physics and Astronomy domains. The lowest international collaboration rates are found in social sciences and humanities, medical and engineering sciences. High observed private international collaboration rates may reflect collaboration within the same organisation's affiliates in different countries.

Definitions and measurability

Collaboration is defined in this report as co-authorship involving different institutions. International collaboration refers to publications co-authored with institutions in another country. Estimates are computed for each country and sector by analysing documents for which the set of listed affiliations includes other country addresses in addition to the reference one. National collaboration concerns publications co-authored with different institutions within the reference country. No collaboration refers to publications not involving co-authorship across institutions. This includes singled-authored articles, as long as the individual has a single affiliation, as well as multiple-authored documents within a given institution.

The institutional measure of collaboration used in this publication may overstate actual inter-personal collaboration in the case of countries where it is common practice to have a double affiliation, for example in a hospital or government institute on the one hand, and a higher education institution on the other. In this case, a significant fraction of single-authored publications may be represented as involving a form of institutional collaborations.

Figure 3.6. Publication output by collaboration type in the higher education sector, 2003-12**Figure 3.7. International collaboration rates by field and sector, OECD, 2003-12**

3.5. Cross-sector collaboration

Co-authorship across different sectors is a potential indicator of collaboration across organisations with different objectives and a potential conduit for knowledge transfer. There is in particular significant interest in the extent to which authors in higher education institutions collaborate with authors in other sectors. This is the focus of this section.

The cross sector collaboration patterns for higher education reflect the relative importance of other sectors in domestic scientific production. Within the combined group of OECD and BRIICS countries, when authors in the higher education sector collaborate with authors in different sectors they do so mostly with authors within government sector (**Figure 3.8**), from 32% for the United States to 90% for the Russian Federation. Among OECD countries, the average government share is at 53%, whereas for the BRIICS this share stands at 64%. The health share is highest for the Netherlands (57%), followed by Israel (55%).

The average share of collaboration with the private sector is at about 8% for OECD (6% for BRIICS). Japan and Belgium exhibit the highest shares (close to 14%).

Cross-sector collaboration patterns do vary across scientific fields. Collaboration between the higher education and the private sector is largest in the domains of Computer Science and Pharmacology. For Computer Science, Israel, Sweden and India, have the highest rates of cross collaboration with private institutions. In the case of Pharmacology, the largest private shares for cross-sector collaboration by higher education institutions are found in Iceland, Switzerland and Denmark (**Figure 3.9**).

Definitions and measurability

Estimates of cross sector collaboration are based on differences in the sectors of institutional affiliation of the group of authors in a given paper. In line with definitions provided earlier in this publication, this is an institutional-based measure, using a whole counts approach. It is therefore possible to count cross sector collaborations in the case of single authors with multiple affiliations that span different sectors.

This indicator has been computed as a percentage of all cross-sector collaborations involving higher education. Internal collaborations within higher education are excluded from the calculation.

Co-authorship based measures of collaborations may in some cases understate the true extent of collaboration by higher education with other sectors, in particular, the so-called private sector, if units in such sectors do not systematically engage in scientific publishing. Different statistical survey approaches for capturing higher education – business collaboration have been developed both from the perspective of business (as in innovation surveys following the OECD/Eurostat Oslo Manual) and from the perspective of higher education institutions (as in the case of countries that run dedicated higher education-business interaction surveys).

Figure 3.8. Profile of cross-sector collaborations for authors in higher education, 2003-12
Percentage of collaborations, based on whole counts, all fields combined

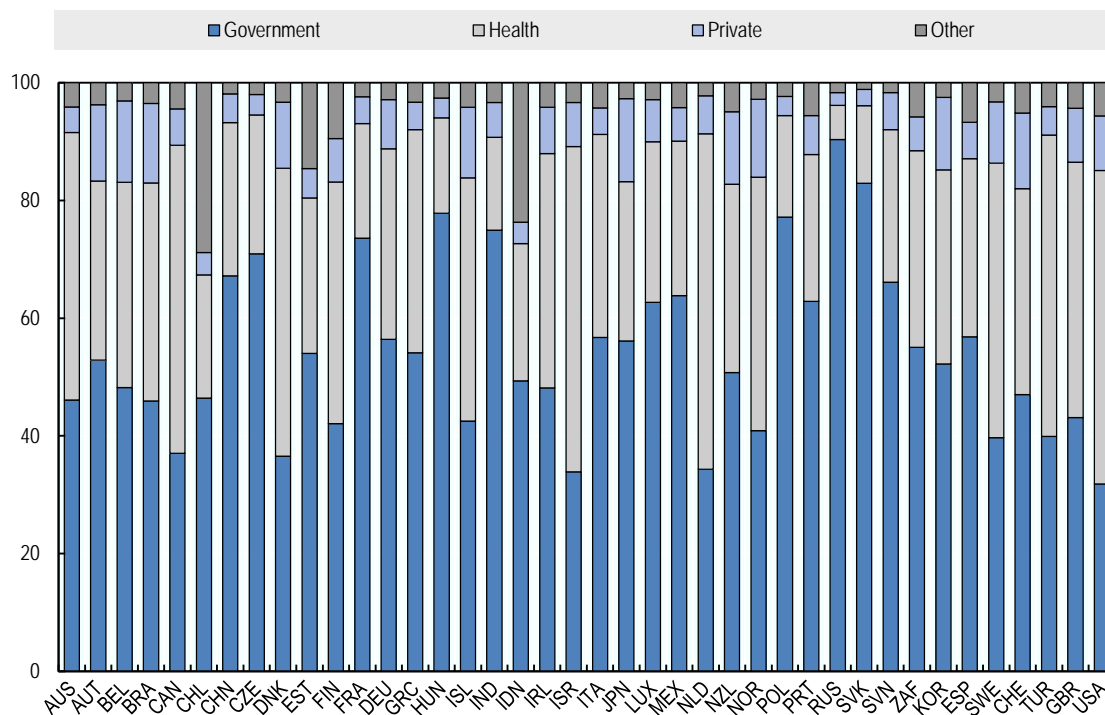
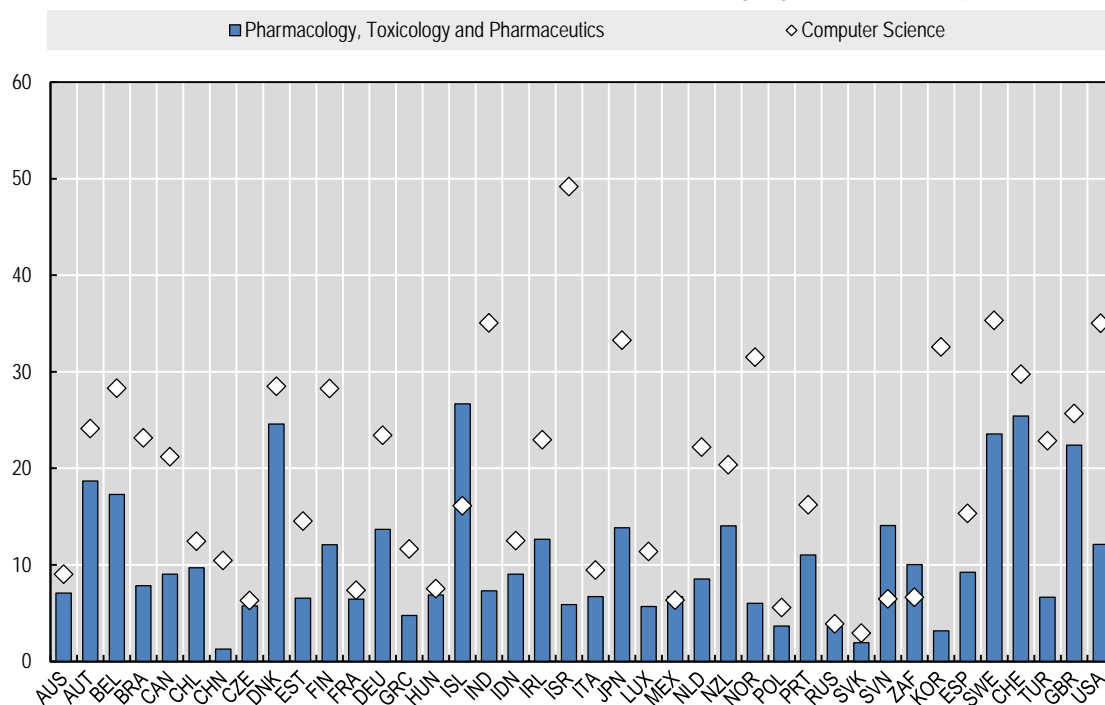


Figure 3.9. Higher education-private sector collaboration, selected fields, 2003-12
Share of private in cross-sector collaborations involving higher education, by field



3.6. Scientific excellence and citation impact

In the OECD and BRIICs areas, the higher education sector accounts for the largest sector share among the subset of the top 10% most cited documents published in 2003-12 (**Figure 3.10**). The Russian Federation is the only country where the share of top 10% most cited documents were higher in the government sector than that of the higher education sector. On a country-by-country basis, the sectoral distribution for high impact publications appears to resemble that for overall scientific publication output, as described in **Figure 3.4** in Section 1 on *Scientific production by sector*.

However, a careful comparison between these two sets of results reveals that the higher education sector is somewhat under-represented among high impact publications (70%) relative to its share in total publications (73%). This suggests that average “excellence” is somewhat lower in this sector. This may be related to the faster growth in publications from this sector relative to others, but no conclusive evidence can be provided for this type of effect without looking in detail at citation impact trends.

This perspective is confirmed by comparing the field-normalised citation impact of documents published in different sectors within each country. **Figure 3.11** allows a comparison of different sectors within countries, as well as similar sectors across different countries. This reveals that citation impact is typically lower in higher education institutions than within their counterparts in other sectors, with exceptions such as Luxembourg and Switzerland. Private institutions display the highest sectoral citation impacts in several countries, although these results are often based on few observations and may indicate some sort of selection effect. Within countries, the average performance of the health sector is relatively high in Australia, Canada, Estonia, Hungary, Indonesia, Poland, the Slovak Republic and the United States, where it is on par with government. Government institutions score relatively high in Austria, Chile, Ireland, Portugal and Turkey.

Definitions and measurability

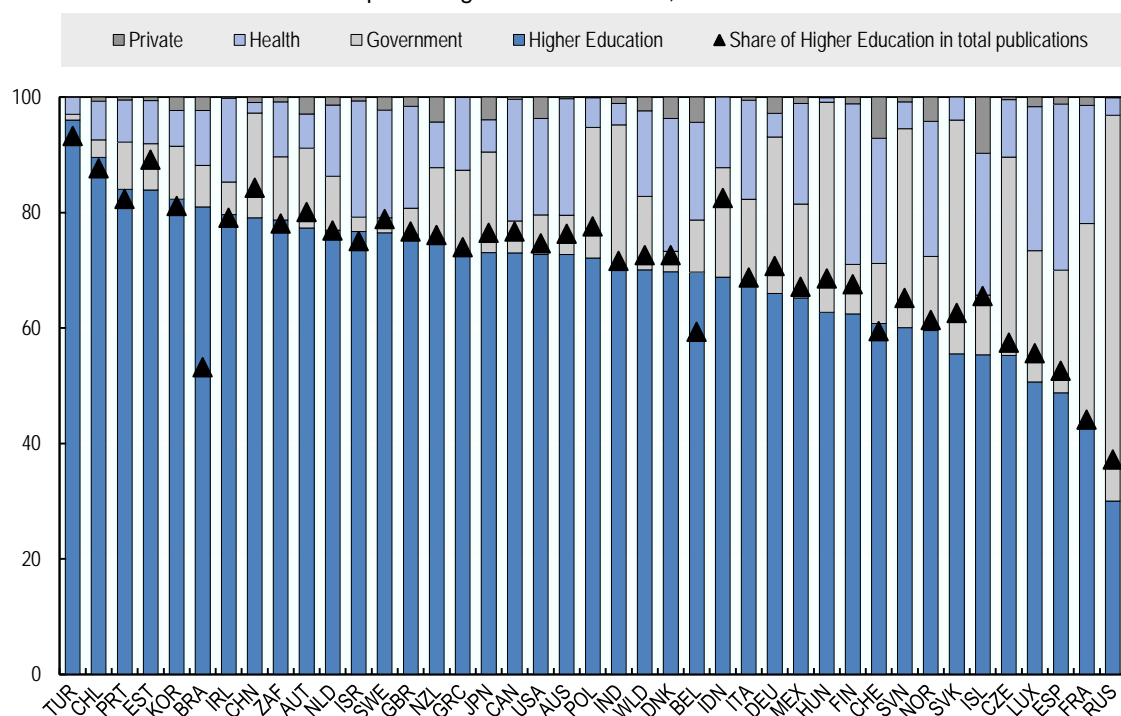
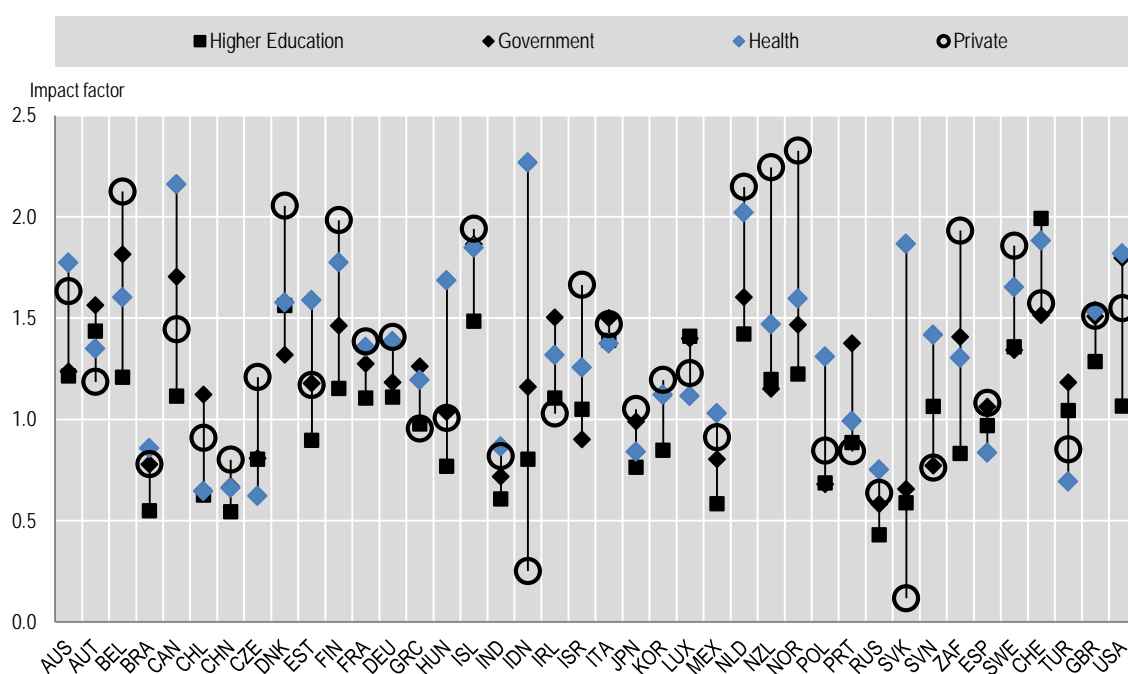
The definitions used here are consistent with those introduced in **Chapter 1** and applied to sectors within countries. The figure representing the 10% most cited documents is an indicator of scientific “excellence”. This rate indicates the amount (in %) of a unit’s scientific output that is included into the set of the 10% of the most cited papers in their respective scientific fields. It is therefore defined with reference to the citation norm within a subject area. It is a measure of high quality of research output of a group of units, in this case the publications attributed to institutions affiliated to a given sector in a given country.

The measure of normalised citation impact shows the relationship of the unit’s average impact to the world average, the unit is the sector in a given country. The indicator is derived as the ratio between the average number of citations received by the documents published by authors affiliated to an institution in a given country and the world average of citations, over the same time period, by document type and subject area. The normalisation of citation values is item-oriented, i.e. carried out at the level of the individual article. If an article belongs to several subject areas, a mean value of the areas is calculated. The value indicates how many times above (or below) the document has been cited above (or below) world average. A value of 0.8 means the country is cited 20% below world average and 1.3 means the country is cited 30% above average.

For some countries and sectors with few observations, the results may be less reliable as indicators and they may exhibit a higher degree of volatility over time.

Figure 3.10. Distribution of top 10% most cited documents – excellence – by sector, 2003-12

As a percentage of all documents, whole counts

[Download figure and data](#)**Figure 3.11. Citation impact of different sectors, all fields, 2003-12**[Download figure and data](#)

3.7. High “impact” institutions

It is widely recognised that individual institutions that attain a critical scale and level of excellence may have a substantial impact on the overall performance of sectors and countries and may also attract a number of related innovation activities in their vicinity. This section examines which institutions are responsible for the largest number of highly cited publications within each sector. **Figure 3.12** provides the abbreviated names and locations of the 20 (50 in the online version) institutions that account for the largest number of the 10% most-cited documents, published over the 2003-12 period, within each of the four sectors covered.

The results point to widely known, large institutions. They confirm the pre-eminent role of US-based institutions in these lists of “high impact” institutions in most sectors, especially in “higher education”. The presence of US-based institutions in the higher education list is above this country’s share of highly cited documents for this sector. This reveals a significant diversity across HE institutions within this country in terms of size and publication outcomes. Institutions based in the United Kingdom and Canada also play a prominent role within this group.

The list of high impact government institutions is the most diverse by country location, with 12 different countries featured among the top 20 institutions and 19 among the top 50. A number of countries such as France, China, Germany, Spain and Italy feature within this list with large central and multi-subject research institutions with presence in different locations within their countries. Some of these institutions account for more publications than the largest US-based universities. To some extent, these represent scale effects, as many of these government institutions have excellence rates that do not fare significantly above the world average.

It is possible to find high impact institutions within the BRIICS, especially in government, where 6 economies hosting these institutions are not in the OECD, namely China, the Russian Federation, India, Singapore, Argentina and Chinese Taipei. The health sector features only OECD countries.

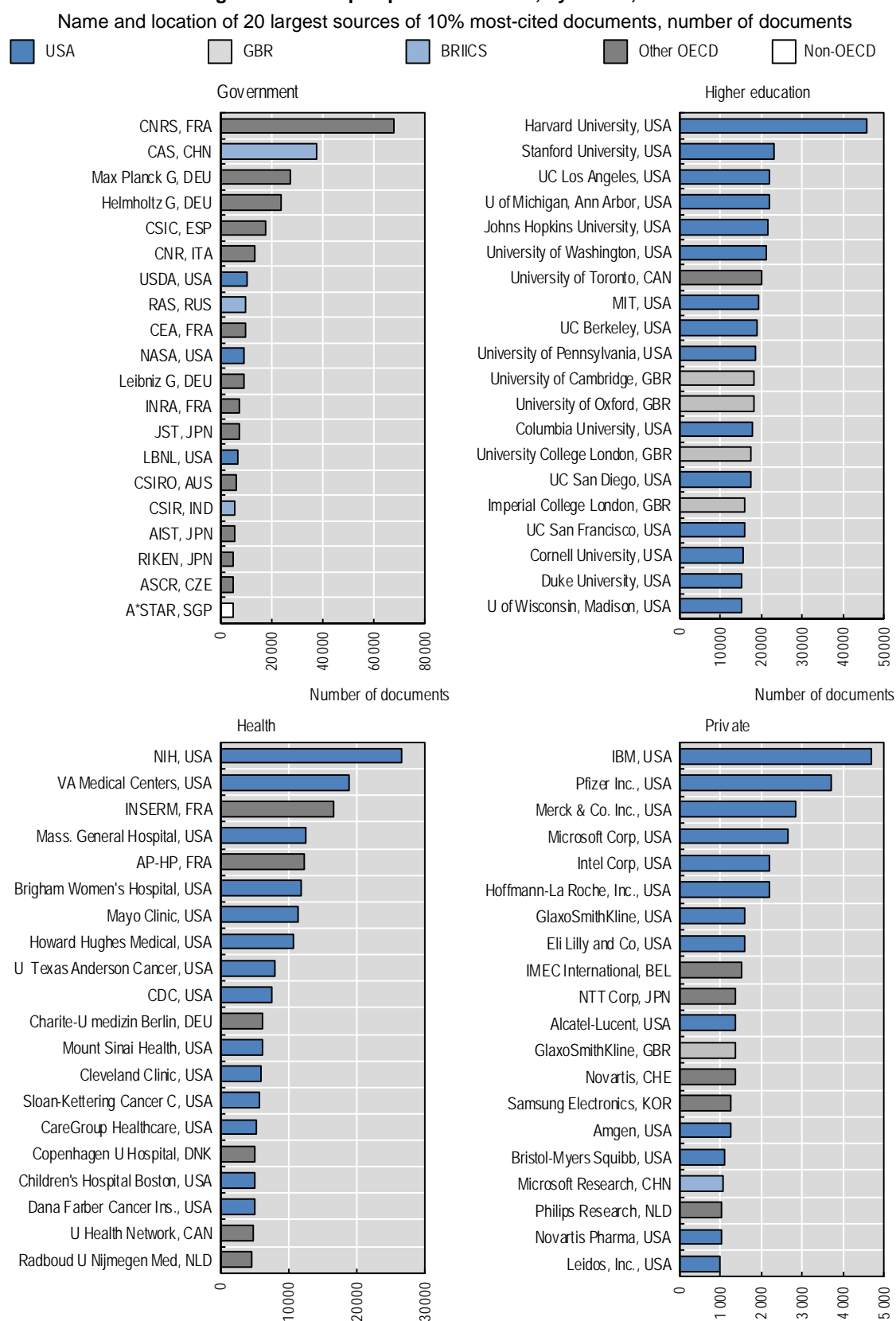
The scale of scientific publishing in “private” institutions is far more limited than for other sectors. Multinationals active in the pharmaceutical and ICT sectors take the top positions, while companies from engineering-based R&D intensive sectors are absent from the list.

Definitions and measurability

The definitions and approaches used for these institution-based indicators are equivalent to those presented earlier in this publication. Top 10% most cited publications are defined by reference to a subject-based norm, to take into account different citation patterns by field. Having accounted for this, this indicator shows which institutions account for the largest number of publications that qualify in the normalised 10% “high impact” group. It does not represent a measure average performance for these institutions.

These results are not intended as league tables, but to help provide a picture of the extent to which high impact publications are concentrated in a number of major institutions, as well as how this pattern compares or differs across sectors.

The presentation of these figures is also aimed at illustrating the classification decisions that underpin the results presented within this chapter. It is possible to note, for example, that “Health” institutions include government institutions like the National Institutes of Health in the United States or INSERM in France, as well as number of medical service providers that could be potentially classified into different sectors, including business. In the case of the “Private” sector, the presence of multinationals raises questions about the degree of consolidation of data (some companies feature in different locations) or classification decisions at the margin with other sectors, as in the case of IMEC that is headquartered in Belgium.

Figure 3.12. Top impact institutions, by sector, 2003-12
[Download figure/data and additional data](#)

Chapter notes and references

Country scope

In this report, the geographical coverage has been limited to the OECD countries and the BRIICS (Brazil, the Russian Federation, India, Indonesia, China and South Africa). For presentational reasons, other economies have been excluded.

References

EC, IMF, OECD, UN & World Bank (2009), *System of National Accounts 2008*, New York, <http://unstats.un.org/unsd/nationalaccount/docs/SNA2008.pdf>.

Moya-Anegón, F., C. López-Illescas and H. Moed (2014), “How to interpret the position of private sector institutions in bibliometric rankings of research institutions”, *Scientometrics*, Vol. 98, pp 283–298.

OECD (2015a), *Frascati Manual 2015: Guidelines for Collecting and Reporting Data on Research and Experimental Development*, The Measurement of Scientific, Technological and Innovation Activities, OECD Publishing, Paris, <http://dx.doi.org/10.1787/9789264239012-en>.