



Using citation-based indicators to compare bilateral research collaborations

Downloaded from: <https://research.chalmers.se>, 2024-11-19 09:39 UTC

Citation for the original published paper (version of record):

Pohl, H. (2024). Using citation-based indicators to compare bilateral research collaborations. *Scientometrics*, 129(8): 4751-4770. <http://dx.doi.org/10.1007/s11192-024-05087-z>

N.B. When citing this work, cite the original published paper.



Using citation-based indicators to compare bilateral research collaborations

Hans Pohl^{2,1}

Received: 11 September 2023 / Accepted: 13 June 2024 / Published online: 22 July 2024
© The Author(s) 2024

Abstract

A standard approach to compare research collaborations between pairs of countries is to look at the citations accrued by all publications with authors from both countries. This approach is often misleading, as aspects only marginally related to the collaboration between the country pairs may bias the result considerably. Among them, the main aspect is the number of co-authors. Publications with many co-authors have on average higher citation impact. If the mix of co-publications between two countries has a high share of such publications, the citation impact will likely be high. Moreover, publications with many co-authors tend to include many countries and are thus only to a limited extent characterising the actual collaboration between the selected pair of countries. The purpose of this study is to develop methods for comparisons of country pairs useful for policy makers, who use SciVal or similar tools. Five methods to compare international collaboration are developed and tested. It is noted that the standard approach for comparisons deviates the most. Fractional methods to calculate the citation impact are recommended, as they allow for the use of citations to all co-publications with a higher weight on the citations to publications in which the country pair dominates. As fractionalisation is laborious to carry out based on SciVal data, a more convenient option is also suggested, which is to use co-publications with maximum 10 co-authors. Elsevier should introduce better methods for comparisons of international collaborations and, until this has been made, help its users understand the limitations of the standard approach featured in SciVal. A by-product of the study is that international co-publications deliver a higher citation impact also when publications with the same number of co-authors are compared.

Keywords International co-publications · Scientometrics · Citation impact · Fractionalisation

✉ Hans Pohl
hans.pohl@lindholmen.se

¹ Chalmers University of Technology, Gothenburg, Sweden

² Lindholmen Science Park AB, Lindholmospiren 3-5, Box 8077, 402 78 Gothenburg, Sweden

Introduction

Which countries should we prioritize in our portfolio of international research collaborations? One straight-forward approach to answer this typical policy question would be to compare the citation impact of co-publications involving each pair of countries. This is easily done in tools such as Elsevier's SciVal. Unfortunately, such an approach tends to give misleading results. The main distorting factor is the positive correlation between the number of co-authors and the citation impact (Glänzel & Thijs, 2004; Puuska et al., 2014; Waltman & van Eck, 2015). If the mix of co-publications involving the two countries is dominated by publications with many co-authors, the citation impact will probably be high. Publications involving many co-authors and countries not only have a large influence on the overall citation impact, they are also the least relevant to consider when studying a specific pair of countries as they also represent research from several other countries.

This study was triggered by three related aspects. First, there is a clear need for fact-based selections of preferred partner countries in research. Second, not only the author but probably many others have used SciVal or similar tools rather uncritically to analyse the citation impact of co-publications in attempts to compare different international collaborations (Kahn, 2018). Third, there appears to be limited guidance in the existing literature for how to make this type of comparisons, especially if the ambition is to rely on tools available to policy makers. The aim of the study is therefore to provide such guidance obtained through an investigation of different approaches to compare the citation impact of collaborations between country pairs, within the limits of broadly available tools such as SciVal. The resulting recommendations take both the quality of the comparison and the work associated with the analysis into account. A by-product of the study is a comparison of national and international co-publications, are the latter delivering a higher citation impact also when the influence of the number of co-authors is removed?

Studies have investigated various dimensions associated with international co-publications. They are more cited (Confrreira et al. 2016; Kamalsky and Plume 2013; van Raan, 1998; Wagner et al., 2017) and they have higher numbers of co-authors (Kahn, 2018). Shortcomings of the field-weighted citation impact or the mean normalized citation score (MNCS) in combination with full count approaches are explained in Waltman and van Eck (2015). Partly to reduce the influence of varying co-author numbers, fractional instead of full count methods are recommended for the analysis of institutions or countries (Aksnes et al., 2012). Potter et al (2020) suggest a method to study the citation impact of a country through a separate analysis of different types of publications, essentially divided into groups based on their numbers of countries involved. Among the few studies specifically addressing international collaboration between different country pairs, Adams and Gurney (2018) differentiate between co-publications based on the number of countries involved. Their analysis illustrates that the partner country to the United Kingdom generating the highest citation impact differed if exclusively bilateral, or publications with additional countries, were considered. Previous literature highlights that the citation impact of international co-publications should be analysed with the co-author numbers in consideration, either through a fractional count or by dividing the publications into different groups based on the number of co-authors or countries involved.

It could be argued that policy makers should not select partner countries and rather leave the decision to the researchers themselves. Moreover, it should be noted that a selection of partner countries should be guided by many different types of data, where the citation impact of existing co-publications is just one possibility. However, if we care about

how the research performs, an analysis of research collaborations between countries is still relevant. Even though all research collaborations do not lead to co-publications, it is widely acknowledged that such data has relevance for the study of international collaboration (Glänzel et al., 1999). If the collaborative research has generated highly cited publications, it is probable that also future collaboration will deliver such publications.

Following a review of previous studies of citation impact indicators and their use on international co-publications, data and methods used in this study are described. Thereafter, an analysis comprising nine countries is made of how the number of co-authors and countries involved in the publications influence the citation impact. Five approaches are tested and the results are discussed. Finally, conclusions follow with recommendations to policy makers as well as providers of tools for publication analysis.

Previous research: citation impact indicators and international co-publications

There is a wide range of indicators based on the citations a publication receives. A direct count of all citations of an author's publications might be used on individual level but for larger entities, other slightly more elaborated indicators are often preferred. One group of such indicators is the share of publications in the top X citation percentile. To account for differences in the publication and citation traditions, this indicator can also be field-normalised. An article-level indicator accounting for the field of study as well as the publication year and the type of publication is the field-weighted citation impact (FWCI), which will be used as the main indicator in this study. A similar concept is used for the mean normalized citation score (MNCS).

Purkayastha et al (2019) describe the FWCI in some detail. It uses the All Science Journal Classification category (334 level) for the field normalisation, which means that all publications in each field are normalised by the average number of citations in the field. It also takes the publication year and the type of publication into account. A baseline of 1.0 is thus obtained.

International co-publications and their associated citation impact have been studied a lot and given the steadily increasing shares of international co-publications (Wagner et al., 2015), it is not surprising. Glänzel and Schubert (2001) studied publications in chemistry and noted that international co-publications on average were more cited than national ones. However, in the selection of 35 countries, a few had higher (Canada, Morocco, Switzerland, Taiwan) or very similar (UK, USA) citation rates for domestic compared to international publications. In another paper by Glänzel (2001), covering a broader data set, the on average positive citation impact of international co-publications is confirmed. However, the positive influence of international collaboration on the national citation indicators is not systematic, there are cases where both partners do not exhibit advantages. Many other studies find that international co-publications are rewarded with higher impact (Confrreira et al. 2016; Kahn, 2018; Kamalsky and Plume 2013; van Raan, 1998; Wagner et al., 2017).

Using data from the extensive 2021 Research Excellence Framework in the United Kingdom, Thelwall et al (2024) managed to show that international co-publications typically have a higher quality than national (co-)publications. This was achieved using peer review data, thereby, at least not explicitly, avoiding the use of citation indicators.

The number of co-authors increases in all parts of science (Thelwall & Mafflahi, 2022). One small but often not negligible share of international co-publications is hyper-authored.

Such publications have up to several thousands of co-authors and involve sometimes more than 100 countries. These publications influence several indicators, not least the citation impact. According to Thelwall (2020), the citation impact of large consortia publications is twice the world average. Kahn (2018) argues that science policy for collaboration between the BRICS countries has been distorted by hyper-authored co-publications and thus the policy is not targeting the involved countries research specialisations. He argues for a separate analysis of publications with very many co-authors.

In a study of whole count and fractionalised methods by Aksnes et al (2012), countries with high shares of international co-publications have a larger difference between whole and fractional citation counts than countries with low shares of international co-publications. They argue for fractionalised methods when countries are to be compared.

Puuska et al (2014) analyse how the number of co-authors relates to the citation impact for six scientific fields using data for Finland. All fields exhibit a positive relationship between the number of co-authors and the citation impact.

A pedagogic and detailed explanation of how full count and various fractionalisation methods relate to the citation impact is given in Waltman and van Eck (2015). A “full count bonus” is calculated and it is shown that field-normalisation in combination with full count approaches might produce misleading results. The relationships between the citation impact and the number of co-authors, organisations or countries are plotted, indicating that an increase in any of the three factors leads to a higher citation impact. As long as this relationship is true, the full count bonus increases with higher numbers of authors, organisations and/or countries involved. It is also demonstrated that the full count bonus has increased steadily over the period 1980–2010 on the country level. Also on the country level, a fractional approach leads to a slightly lower citation impact score compared to the full count approach for all 25 countries included, however with some variation. This variation is to a large extent similar to the reduction in publication volumes, when comparing full count and fractionalised publication volumes per country. The authors consider it absolutely essential to use fractional counting instead of full counting at the country and organisational levels.

Addressing another dimension of the differences between full count and fractionalized data, Perianes-Rodriquez et al. (2016) forward VOSviewer co-authorship network maps based on each type of data. They differ considerably. Interestingly, a full count network map based on publications with up to 20 institutions looks similar to the fractionalised one.

Closer to the purpose of this study, one approach to handle varying co-author numbers caused by international collaboration is forwarded by Adams et al (2019) and Potter et al (2020). Acknowledging the influence of co-author numbers, they suggest a split of the publications of an entity into five groups, single author, national collaboration, and international collaboration with one, two or many partner countries. For each group a category normalised citation impact is calculated and normalised within the group.

Directly relating to the purpose of this study, Adams and Gurney (2018) study international co-publications including the United Kingdom. They look at bilateral as well as all publications including a selection of other countries. The normalised citation impact is lower for the bilateral publications and whereas co-publications with the United States give the highest citation impact if strictly bilateral co-publications are considered, co-publications with France give higher citation impact if all co-publications including France and the United Kingdom are considered. They recommend that publications with more than 20 countries involved should be analysed separately.

Previous literature demonstrates that the citation impact for international co-publications is influenced by the co-author number or the closely related number of countries

involved. It also argues that the citation impact is higher for international co-publications, at least on a full count basis. One conclusion from the study of previous literature is that the citation impact of international co-publications should be analysed with the co-author numbers in consideration, either through a fractional count or by dividing the publications into different groups based on the number of co-authors or countries involved.

Methodology and data

Outline of five approaches to be tested

Given the ambition to help policy makers in their work with publication analyses, the study is limited to approaches possible to carry out in SciVal, or similar tools. This excludes a development of completely new advanced indicators. However, given the emphasis on fractionalisation in previous literature, one such approach is tested, even though it requires some efforts to calculate.

In SciVal, the co-author number filters use the limits 10, 50, 100 and in some cases also 1,000 co-authors. These filters are available on the institutional level but unfortunately, there is no such filter on the country level. However, to study co-publications with a specified maximum co-author number, publication sets can be made. One limitation of publication sets is that they can include up to 100,000 publications.

The following approaches will be tested:

1. All co-publications including country A and B (the standard approach)
2. Co-publications including country A and B excluding publications with more than 100 co-authors
3. Co-publications including country A and B excluding publications with more than 10 co-authors
4. Co-publications including country A and B and up to two more countries
5. All co-publications including A and B and a country-level fractionalisation of the citation impact.

The standard approach is what you can get directly in SciVal's Collaboration Module and thus the approach that is probably used by most policy makers. One factor triggering this study was that the FWCI when using this approach is very high, especially for collaboration with countries not considered established academic super-powers. In Fig. 1, an illustration of how the collaboration between two countries is illustrated, in this case between Sweden and Bulgaria. The FWCI for Sweden, Bulgaria and the publications including both countries are central indicators. The Collaboration Module is not visible in the new SciVal navigation introduced in June 2024 but the data presented under Collaboration - Current collaborators is similar to what Fig. 1 illustrates.

The second approach is used to reflect the recommendations to remove or treat hyper-authored publications separately (Adams & Gurney, 2018; Kahn, 2018). As the analysis in next section will show, publications with more than 100 co-authors also tend to involve many countries.

The third approach uses the co-author filter with the lowest available number. How does it change the citation impact when all co-publications with more than 10 co-authors are excluded?

Collaboration with Bulgaria

Year range: 2019 to 2021

Shortcuts

Research Performance	Areas of collaboration	Current co-authors	Potential co-authors
<div> Metric guidance Add to Reporting Export </div>			
	Sweden	Co-authored	Bulgaria
	<p>1,057 ▲</p> <p>co-authors with Bulgaria</p> <p>1.61</p> <p>Field-Weighted Citation Impact</p>	<p>653 ▲</p> <p>publications</p> <p>7.25</p> <p>Field-Weighted Citation Impact</p>	<p>531 ▲</p> <p>co-authors with Sweden</p> <p>0.98</p> <p>Field-Weighted Citation Impact</p>
Authors	90,846 ▲	–	14,494 ▲
Scholarly Output	143,586 ▲	–	21,448 ▲
Views count (from Scopus)	3,789,610	99,898	659,259
Field-Weighted Views Impact	1.33	9.23	1.83
Citation Count	1,805,091	43,892	137,981
Number of Institutions	177		42

Fig. 1 Collaboration between two countries as presented in SciVal

The fourth and fifth approaches require considerably more efforts. Both use the number of countries involved in each publication, which when the data exports for this study were made were not directly available in SciVal. Using the export function in SciVal, which allows for an export of up to 100,000 publications at a time, a count of the country numbers was made in Excel. The selection of maximum four countries per publication in the fourth approach is motivated by a wish to maintain a focus on the two partner countries studied, without ignoring the fact that some countries appear to be important enablers of trilateral or quadrilateral collaborations. With maximum four countries, the two partner countries represent 50% of all countries involved in the collaboration, whereas with a higher number, they represent a minority. A strict focus on bilateral collaborations would in many cases reduce the co-publication volumes considerably and bias the analysis towards fields typically addressed in small collaborations. The fourth approach is to some extent similar to the grouping proposed by Adams et al (2019).

The fifth approach deserves a more detailed discussion. Even though not explicitly focusing on international co-publications, several studies suggest fractionalisation of the citations. Fractional counting on author level means that the citations are divided between the authors following some scheme, where an equal split probably is the most common, but many different approaches exist (Gaufriau, 2021). To facilitate the

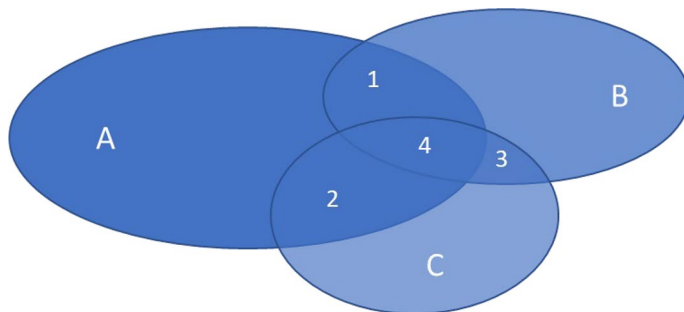


Fig. 2 A world with three countries

discussion of fractional approaches, a very simplified world with only three countries A, B and C is depicted in Fig. 2. The overlapping areas indicate international co-publications. Country A's international co-publications are the areas 1 + 2 + 4. Strictly bilateral publications between country A and B are in the area 1.

Further, let us assume two publications, one national with two authors based in country A, and one international with one author in country A and one in country B. If we use fractional author-level counting, the nationally co-authored publication will contribute twice as much to country A's total citation count as the international one. This is true also when the fractional counting is implemented on country level. With three co-authors, whereof one in country A, one third of the citations are counted for country A. From this very simple example follows that a comparison of national and international co-publications for a specific country will yield clearly different results when author or country level fractionalisation is used.

The essence of this study is to answer the following question: For country A, do the co-publications with country B or country C deliver the highest citation impact? Obviously, the trilateral publications in area 4 contribute to both collaborations and as publications with more countries involved on average receive a higher citation impact, the share of such publications in the collaboration of interest plays a role. If the share of trilateral publications among all international publications is larger for the collaboration with country B than country C, the collaboration between A and B benefits more from the presumably higher impact of the trilateral publications.

In the real world, the main reason why Sweden had a very high FWCI for co-publications with Bulgaria is that the collaboration had a large share of publications with many co-authors. Out of the 653 co-publications illustrated in Fig. 1, 242 had more than 100 co-authors.

When the collaborations A–B and A–C are compared, it appears reasonable to count the citations of the trilateral co-publications to a lower extent. If citations to bilateral publications are counted fully (they include the targeted pair of countries to 100%), the citations relating to trilateral publications could be included to 67%, assuming an equal contribution from each country. One author in each country means that two are involved in the selected country pair, which is $2/3$ of the total. For publications with four countries involved, the country pair represents $2/4$ of the total, or 50%. This type of fractionalisation on country level is analogous to the fractionalisation approaches proposed in previous literature but adapted for international co-publications only.

The fractionalisation approach is far from perfect. One simplification is that all countries involved in a publication are assumed to participate to the same extent. A better way would be to weigh the citations based on the number of authors from each country. Further complicating factors include that some authors have multiple affiliations in different countries and that the position of the author in the publication might matter for how the citations should be distributed. However, given the purpose of this study to develop methods being possible to use by policymakers, such fine-tuning of the approach had to be excluded.

It is possible to weigh the citations accrued by different types of international co-publications differently. The full count approach equals a full weight for all types of co-publications. A weight directly related to the number of countries involved as per above (weight = $2/\text{number of countries involved}$) reduces the impact of the citations to publications with many countries involved considerably, with 10 countries involved the weight is 20%. Other approaches, such as excluding publications with many authors or countries involved, correspond to a full weight of citations to all publications up to a certain limit

and zero weight for those above. In the next section, the relationships between the number of authors, countries and the associated citation impact will be investigated, which among others will tell how the weights relate to the increase in citation impact as the numbers of countries increase.

Data

Nine countries were used to test the selected approaches. The countries were selected based on the following. The two largest countries in terms of publication volume were selected (United States and China), as they represent a large share of the total publication volume in the world. Thereafter, five countries with a very high field-weighted citation impact in 2016 were selected as they might be the most likely ones not enjoying higher citation impact from international collaboration (Denmark, Netherlands, Singapore, Sweden, and Switzerland). Finally, two countries with rapidly growing publication volumes representing the potential future research nations (Indonesia and Saudi Arabia) were selected to see if they are different in the aspects studied (Pohl, 2020). When looking at the collaboration networks of these nine countries, all 48 countries having more than 10,000 Scopus publications in 2016 were included in the comparison. A threshold is motivated as the co-publication volumes between some countries would otherwise be very small and the citation impact rather random. For a discussion of small countries, Potter et al (2020).

To form a basis to better understand the different approaches to rank partner countries, the selected countries were initially studied with a focus on their mix of publications with different co-author and country numbers and the associated citation impact. The number of co-authors per type of publication was also studied, which in this case means that the number of co-authors for national, bilateral, trilateral and so on were calculated. Even though not the main purpose of the study, it was also investigated if international co-publications give a higher citation impact also when the influence of the number of co-authors was removed.

Data comes from the Scopus database, which is among the largest curated abstract and citation databases (Baas et al., 2020) or, to use a more neutral source, the publication database with the broadest coverage (Burnham, 2006). In the analysis of Scopus data, SciVal and Excel were used. SciVal is a tool provided by Elsevier to facilitate the use of publication data (Elsevier, 2022). Publication data in SciVal are available from 1996. In SciVal, international co-publications include at least two authors and affiliations in at least two countries. The year 2016 was used and data was extracted in the beginning of 2022, which means that there has been considerable time for citations to accumulate. All types of publications were included.

Results

Detailed study of the citation impact of international co-publications

In Table 1 below, some basic publication data for the country selection are presented.

In 2016, the United States was still the largest country in terms of publications, being involved in more than 22% of all publications world-wide. China was number two. The other countries have much lower volumes. The FWCI including self-citations is very high for several countries, which is not surprising as it was a selection criterion when the sample

Table 1 Basic publication data

All Scopus publications 2016

Country	Volume	Share of world (%)	Citation impact incl. selfcitations (FWCI)	Citation impact excl. selfcitations (FWCI)	International co-publications (field-weighted)
China	501,834	16.4	0.97	0.36	0.54
Denmark	28,961	0.9	1.92	1.59	1.51
Indonesia	12,740	0.4	0.97	0.56	0.79
Netherlands	61,871	2.0	1.86	1.50	1.51
Saudi Arabia	21,403	0.7	1.51	1.23	1.95
Singapore	22,259	0.7	1.83	1.57	1.60
Sweden	43,009	1.4	1.74	1.42	1.55
Switzerland	47,721	1.6	1.87	1.54	1.66
United States	690,443	22.6	1.44	0.75	0.85

of countries was designed. However, when excluding self-citations on country level, it is predominantly the large countries with low shares of international co-publications that get drastically lower values. Self-citations on country level means that citations from the same country as at least one of the co-author affiliations are not counted. As this punishment of size appears irrelevant, self-citations are not considered in the following analysis. In the last column, the field-weighted internationalisation score is high for countries with high FWCI. This value is to be interpreted the same way as the FWCI, which means that the world average is 1.0 (for details, see Pohl et al., 2014).

The distribution of national and international publications per co-author number is presented with one figure per country in Table 2. For each number of authors from 1 (only national publications) to 10, the share is calculated as a percentage of all publications including the country in 2016. The intervals 11–20 and more than 20 are also presented. On the right axis, the FWCI is indicated and the line shows the value for each co-author number.

On the one hand, the patterns relating to co-author numbers differ between the countries. National publications show a decreasing share for all countries except China and Indonesia, which have the highest share of national publications with 3 or 4 co-authors. The international co-publications peak at 3 or 4 co-authors in all 9 countries. All countries in Europe have relatively high shares of publications with more than 10 co-authors, whereas only China and the United States have 2% or more national publications with more than 10 co-authors. It should also be noted that the international co-publications dominate when the co-author numbers increase.

The share of international co-publications with more than 10 co-authors varies between less than 3% (China) to approximately 15% (Switzerland). This has clear implications on how the approaches 3 and 4, as described in sub-Sect. “[Outline of five approaches to be tested](#)”, influence the citation impact.

On the other hand, the curve for the FWCI has a similar shape for all nine countries. The scales differ but it is obvious that the FWCI increases with the number of co-authors per publication and that publications with more than 20 co-authors differ substantially from other publications.

Figure 1 displays the share of all publications and field-weighted citation impact (FWCI) for different numbers of co-authors across eight countries: China, Denmark, Indonesia, Netherlands, Saudi Arabia, Singapore, Sweden, Switzerland, and United States. The charts are organized in a 4x2 grid. Each chart plots three series: Int'l (solid blue line with dots), Nt'l (dashed blue line with circles), and FWCI (solid red line). The x-axis represents the number of co-authors (1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11-20, >20). The left y-axis represents the share of all publications (0% to 16% for China, 0% to 12% for Denmark, 0% to 30% for Indonesia, 0% to 10% for Netherlands, 0% to 16% for Saudi Arabia, 0% to 12% for Singapore, 0% to 12% for Sweden, 0% to 10% for Switzerland, and 0% to 18% for United States). The right y-axis represents the field-weighted citation impact, FWCI (0.00 to 7.00 for China, 0.00 to 9.00 for Denmark, 0.00 to 18.00 for Indonesia, 0.00 to 8.00 for Netherlands, 0.00 to 10.00 for Saudi Arabia, 0.00 to 12.00 for Singapore, 0.00 to 8.00 for Sweden, 0.00 to 7.00 for Switzerland, and 0.00 to 6.00 for United States). In all countries, the share of publications generally decreases as the number of co-authors increases, while the FWCI increases significantly for higher co-author counts, particularly for >20 co-authors.

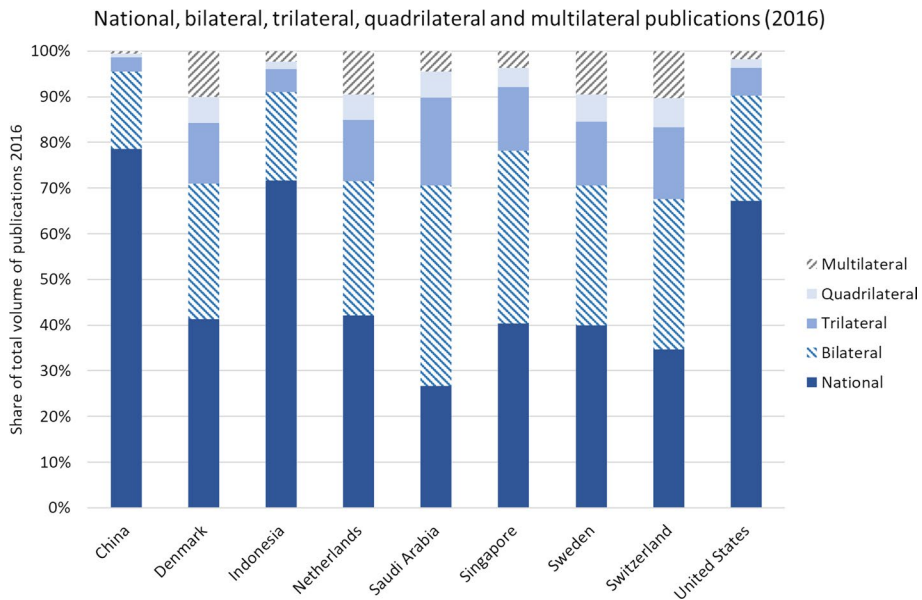


Fig. 3 Proportions of publications with one, two, three, four and more than four countries involved

One rule of thumb is that a large country in terms of scientific publications has a low share of international co-publications (Aksnes et al., 2012). In Fig. 3, this is to some extent confirmed, as the two largest countries, China and the United States, both have a large share of national publications. Saudi Arabia has the highest shares of bilateral and trilateral co-publications. Denmark, Netherlands, Sweden and Switzerland all have approximately 10% “multilateral” co-publications, which in this figure and in the following is defined as co-publications with more than four countries involved.

When looking specifically at the international co-publications, China has almost 80% bilateral such publications, which is much higher than the other countries in the sample. The FWCI for each number of countries involved in the publications is indicated in Fig. 4. All countries exhibit an increasing citation impact when the number of countries involved increases. Publications with more than 4 countries involved differ substantially with their very high citation impact.

When normalising the citation impact in line with approach 5 described in 3.1, the differences in citation impact for the three categories bilateral, trilateral and quadrilateral publications become smaller. For most countries, the opposite order is resulting with the highest normalised citation impact for bilateral co-publications (counted 100%) and the lowest for quadrilateral (counted 50%).

The number of co-authors per country number is presented in Fig. 5. Here, international means all international co-publications as per definition, and not only those with more than four countries involved. Despite this, the averages for international co-publications are quite high for countries like Denmark and Switzerland. China has the highest and Saudi Arabia the lowest number of co-authors for national publications. What could be noted is that the numbers of co-authors per bilateral, trilateral and quadrilateral publications are rather similar in all countries. International publications exhibit larger

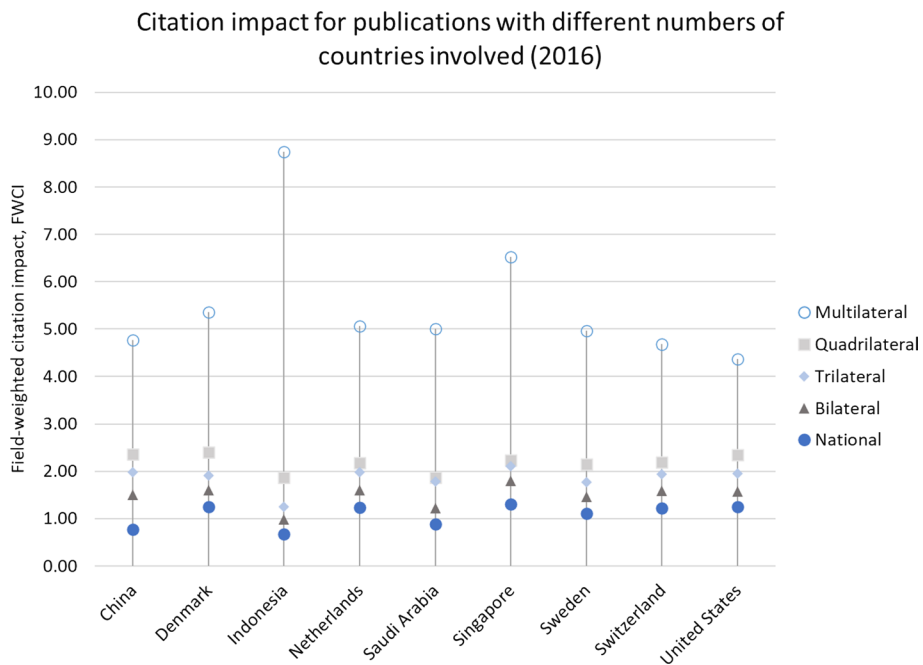


Fig. 4 Citation impact for publications with 1, 2, 3, 4 and more countries involved

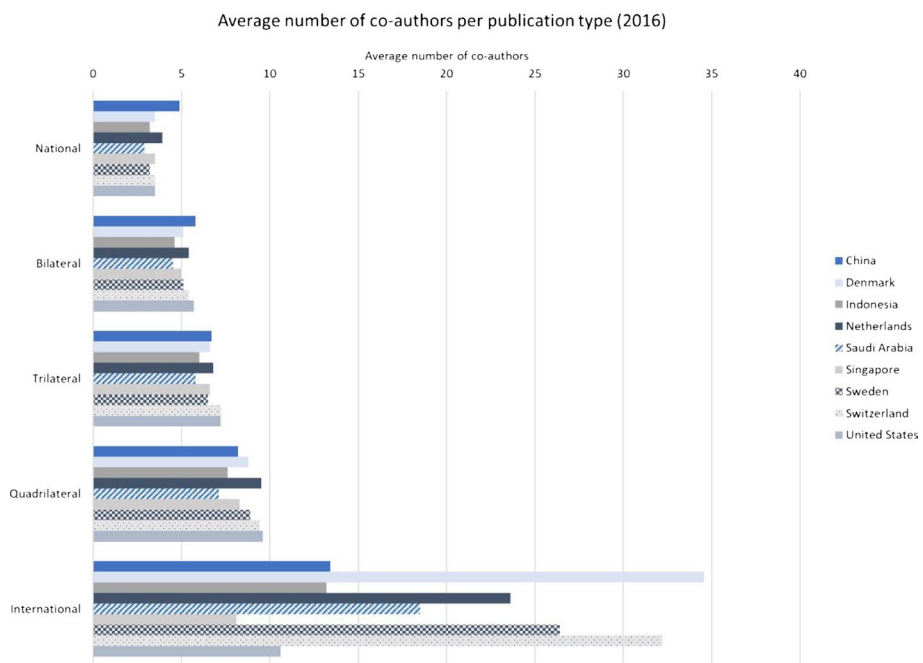


Fig. 5 Average number of co-authors per type of publication and country

differences. A slightly more detailed analysis of co-author numbers per publication type is available in Appendix.

Are international co-publications more cited?

Given the substantial attention in previous literature devoted to whether international co-publications are more cited than national ones and the lack of approaches comparing publications with the same co-author numbers, this was tested using data as per above. A linear regression model was used for publications with 2 – 10 co-authors to investigate if there is a significant difference between national and international publications in terms of the FWCI. As FWCI values are highly skewed, a logarithmic transformation was made, using $\ln(\text{FWCI} + 1)$. This is in line with other studies of similar type, among them Puuska et al (2014).

The equation $\ln(\text{FWCI} + 1) = a + b \cdot \text{Co-authors} + c \cdot \text{International}$ was used, where Co-authors represents the number of co-authors per publication and International is 1 for international co-publications, otherwise 0. In Table 3, some results are summarized.

For all countries, both b and c are positive and p very small. It means that both co-author numbers and whether the co-publication is international or not contribute significantly (***) and positively to the FWCI. Interestingly, China and the United States have the strongest positive correlation between the citation impact and the number of co-authors ($b = 0.035$). China is also the country which most clearly benefits from international collaboration ($c = 0.269$).

The omission of publications with higher co-author numbers has as illustrated Table 2 a negative impact on the citation impact for international co-publications. It means that this test is more challenging for international co-publications than if all publications were included.

Empirical test of five approaches to compare international partnerships

In this sub-section, the approaches described in 3.1 are calculated using publications for 2016. As a complete listing of all values for five approaches for all 9 countries and

Table 3 Regression results for the comparison of national and international co-publications

Country	Publications	a		b		c	
				Value	p	Value	p
China	466,265	0.26	0.035	0		0.269	0
Denmark	22,082	0.54	0.025	2.47E-43		0.103	1.88E-37
Indonesia	11,065	0.31	0.023	6.06E-15		0.157	1.6E-49
Netherlands	47,621	0.52	0.027	1.4E-110		0.106	1.45E-84
Saudi Arabia	18,194	0.38	0.034	3.84E-59		0.121	1.18E-29
Singapore	18,811	0.53	0.029	3.36E-38		0.107	2.42E-29
Sweden	33,404	0.50	0.028	7.82E-94		0.080	4.99E-35
Switzerland	35,967	0.52	0.028	1.88E-87		0.084	2.42E-34
United States	526,131	0.48	0.035	0		0.056	3.2E-242

the selection of 47 partner countries would be space consuming, only the collaborations between the 9 countries analysed above and two more are included in the tables. The additional two are those with the highest rank in the standard approach (number 1), which means the FWCI when all co-publications with both countries are counted fully. They are placed in the last positions in the tables. Moreover, only results for three countries are presented here. In Appendix, tables for all 9 countries are available.

In the standard approach for United States and its partner countries, the collaborations with the highest citation impact are with Indonesia, Ireland and Pakistan, see Table 4. China gets the lowest rank.

When co-publications with more than 100 co-authors are removed, the rank of Indonesia and Pakistan changes substantially, whereas Ireland still ranks number 2. For other countries, the ranking improves and the Netherlands top the ranking.

A comparison of the FWCI for co-publications with maximum 100 and 10 co-authors respectively, shows that the values are as expected clearly lower for the latter. Singapore and China climb the rankings, whereas Indonesia, Ireland and Pakistan lose several positions.

Quadrilateral collaboration typically yields a slightly higher FWCI than co-publications with maximum 10 co-authors. Exceptions are collaborations with Ireland and Pakistan. Data illustrated in Fig. 5 tells that quadrilateral co-publications including the United States on average have 9.61 co-authors. Given that the limit to maximum four countries allows for a larger spread in the co-author numbers, this might be the reason why this category tends to give a higher FWCI than publications with maximum 10 co-authors. However, the differences in FWCI and ranks are relatively small.

Finally in Table 4, the FWCI normalised with the country number per publication is listed. The highest ranked country in this approach is Canada with $FWCI_{\text{frac}}$ 1.61 (not in the table) closely followed by Singapore. China has a clearly better rank in the fractionalised column. China was the largest partner to the United States. 74% of the publications were bilateral, which is a very high proportion, and they had a FWCI equalling

Table 4 Co-publications with the United States (2016)

		United States			
Partner country		Field-weighted citation impact, FWCI (rank)			
Approach no	1	2	3	4	5
Publications included	All	Max 100 coauthors	Max 10 coauthors	Max 4 countries	All (fractionalised)
Indonesia	5.66 (1)	2.33 (24)	1.46 (28)	1.49 (29)	1.07 (29)
Denmark	3.67 (7)	3.02 (3)	1.96 (6)	2.16 (5)	1.45 (10)
Netherlands	3.61 (10)	3.20 (1)	2.16 (1)	2.28 (2)	1.58 (3)
Sweden	3.37 (17)	2.85 (8)	1.94 (8)	2.03 (8)	1.38 (13)
Switzerland	3.27 (20)	2.91 (5)	2.14 (2)	2.22 (3)	1.49 (8)
Singapore	3.03 (27)	2.50 (19)	2.05 (4)	2.14 (6)	1.61 (2)
Saudi Arabia	2.67 (37)	2.11 (31)	1.82 (13)	1.84 (18)	1.34 (17)
China	1.91 (47)	1.82 (38)	1.65 (23)	1.72 (22)	1.51 (5)
Ireland	4.28 (2)	3.10 (2)	1.88 (9)	1.87 (17)	1.37 (15)
Pakistan	4.13 (3)	2.24 (26)	1.38 (36)	1.35 (38)	1.00 (36)

1.59. In the fractionalised approach, bilateral publications have a large impact so even if the FWCI for these publications was a bit lower than it was for some other partner countries to the United States, it contributed to a better rank.

In Table 5, China's collaboration with the United States shows very similar results with a gradual improvement in rank from the left to the right, in both cases starting at the bottom (rank 47) and reaching rank 5. China and Saudi Arabia appear to have a collaboration generating high citation impact. Singapore is another strong partner, as it is also to the United States. All the top-ranked countries according to the standard approach land close to the bottom of the list when a limit of maximum four countries or fractionalisation is used. The ranks and positions in the columns “Max 10 co-authors” and “Max 4 countries” are similar in all cases except for Colombia. The limited co-publication volume (440 publications) appears to have some highly cited publications with maximum 10 co-authors from more than four countries.

Indonesia got a very prominent rank in the rankings for the United States and China (and several other countries, see Appendix) according to the standard approach. From Indonesia's perspective, many other countries give even higher FWCI, see Table 6. This lack of symmetry is not surprising, as the high ranking of Indonesia in the other countries' tables predominantly reflects that a large share of Indonesia's international co-publications involves many countries and authors. The extremely high FWCI values in the standard approach should be noted, they are as the column Max 100 co-authors shows to a large extent caused by publications with more than 100 co-authors. A low number of co-publications makes the results volatile. There are several ways to increase the publication volumes included, one is to use several years, another is to exclude countries with few publications. Here, this has not been done.

The tables in Appendix for the other six countries are in line with the results presented above. An exclusive look at publications with up to ten co-authors or with four countries give similar ranks and the standard approach including all publications with country A and B, deviates the most from the other approaches.

Table 5 Co-publications with China (2016)

China					
Partner country	Field-weighted citation impact, FWCI (rank)				
Approach no	1	2	3	4	5
Publications included	All	Max 100 coauthors	Max 10 coauthors	Max 4 countries	All (fractionalised)
Indonesia	10.78 (1)	2.24 (34)	1.12 (44)	1.04 (45)	0.77 (40)
Switzerland	4.55 (24)	3.14 (7)	2.24 (3)	2.09 (4)	1.17 (14)
Netherlands	4.09 (30)	2.66 (20)	1.60 (28)	1.55 (34)	1.16 (16)
Denmark	4.02 (31)	2.26 (33)	1.69 (18)	1.77 (12)	1.19 (11)
Saudi Arabia	4.01 (32)	3.00 (11)	2.81 (1)	2.76 (1)	1.83 (1)
Sweden	3.69 (34)	2.59 (24)	1.68 (20)	1.74 (15)	1.26 (8)
Singapore	2.73 (39)	2.29 (32)	2.09 (4)	2.18 (3)	1.82 (2)
United States	1.91 (47)	1.82 (44)	1.65 (22)	1.72 (17)	1.51 (5)
Argentina	9.42 (2)	5.53 (2)	1.01 (46)	0.99 (46)	0.71 (41)
Colombia	9.15 (3)	7.46 (1)	1.93 (6)	1.57 (30)	0.59 (46)

Table 6 Co-publications with Indonesia (2016)

Partner country	Indonesia				
	Field-weighted citation impact, FWCI (rank)				
Approach no	1	2	3	4	5
Publications included	All	Max 100 coauthors	Max 10 coauthors	Max 4 countries	All (fractionalised)
Denmark	27.11 (21)	1.99 (33)	1.41 (14)	0.86 (36)	0.95 (35)
Sweden	21.94 (24)	3.21 (14)	2.15 (5)	2.05 (7)	1.28 (12)
Switzerland	20.62 (27)	3.87 (8)	2.09 (6)	2.19 (6)	1.11 (21)
Saudi Arabia	17.06 (31)	1.67 (41)	1.02 (29)	1.05 (25)	1.02 (31)
Singapore	12.25 (35)	2.65 (22)	1.43 (13)	1.29 (17)	1.10 (22)
China	10.78 (36)	2.24 (30)	1.12 (25)	1.04 (26)	0.77 (44)
Netherlands	7.16 (42)	1.88 (38)	1.23 (23)	1.23 (19)	0.98 (34)
United States	5.66 (43)	2.33 (28)	1.46 (12)	1.49 (13)	1.07 (27)
Israel	82.12 (1)	4.04 (7)	1.38 (15)	0.22 (45)	2.12 (2)
Ireland	79.16 (2)	1.90 (37)	0.56 (45)	0.40 (43)	1.95 (3)

Discussion

The main aim of this study was to provide guidance how to compare the citation impact of collaborations between country pairs, within the limits of broadly available tools such as SciVal. Implicitly, it is assumed that the citation impact is related to the quality of the publications and thus to at least one aspect of the international collaboration. Given the substantial interest in comparisons of the citation impact for national versus international (co-)publications, an effort was also made to contribute to this discussion.

Starting with the latter question, almost all comparisons of the citation impact for national versus international publications result in a higher impact for the international ones (Confraira et al. 2016; Kahn, 2018; Kamalsky and Plume 2013; van Raan, 1998). Peer review also indicates that international publications typically are of higher quality (Thellwall et al., 2024). As international co-publications have a higher number of co-authors, as illustrated in Fig. 5, and higher co-author numbers on average generate a higher citation impact, see Table 2 and Fig. 4, it was considered relevant to investigate if the higher citation impact only depends on the higher co-author numbers or if there is a difference also when publications with the same co-author numbers are compared. This was investigated for all nine countries in this study for all co-author numbers from 2 – 10. The result was that yes, international co-publications were more cited also when the influence from the number of co-authors was removed.

Previous studies highlight the large differences in how countries publish and collaborate (Adams & Gurney, 2018). To support interpretation of the results, a brief analysis of nine countries' publication profiles was made with a focus on the co-author and country numbers involved in the publications. The analysis confirmed large variations. One factor having an impact on how the share and mix of international co-publications is constituted is the size of the country in terms of the annual number of publications. However, even for countries with approximately the same publication volumes, the differences are far from negligible. For example, the share of China's international co-publications

involving more than four countries was only 3%, whereas the corresponding share for the United States was 6%. The share of such publications for Denmark, the Netherlands, Sweden, and Switzerland was 16–17%. Despite these differences, it was noted that the relationship between the number of co-authors and the number of countries involved was relatively stable, at least up to and including quadrilateral co-publications.

With inspiration from literature, five approaches to calculate the citation impact for country pairs were developed and tested. Which recommendations emanate from the comparison of the five approaches?

From a theoretical perspective, the fractionalisation approach is most appealing. Fractionalisation is recommended by several researchers (Aksnes et al., 2012; Waltman & van Eck, 2015). In this study, an adapted version of fractionalisation specifically for international collaborations was developed. It normalises the citations per publication based on the number of countries involved with 100% of the citations included for purely bilateral collaborations, 67% for trilateral and so on. Compared to the other approaches tested, it has the advantage that all publications are considered with a gradual decrease of the influence on the citation impact as the number of countries involved increase. Obviously, the weights for the citation scores for different numbers of countries could and should be discussed, there might be arguments to have a steeper or less steep reduction of the weights. But as a starting point, the proposed model appears sound.

From a practical perspective, fractionalisation required some efforts as the number of countries per publication was not directly available in SciVal. To get the number of countries per publication, a count of the countries affiliated in each publication was made in Excel. For the larger countries, this was associated with some work, as the publications had to be exported in several portions to be merged in Excel. In this regard, the close dialogue with Elsevier appears to have been fruitful, since the data-intensive part of this study was carried out in 2022, the number of countries involved in the publications have been added to the list of possible items to export from SciVal.

The approach only considering publications with up to four countries involved, requires similar efforts as the fractionalisation approach and is thus not considered interesting for practical use. However, the test was useful as it showed that the existing filter “maximum 10 co-authors” gives similar results in terms of the citation impact.

If fractionalisation is considered too demanding, the recommendation is to focus on publications with up to 10 co-authors when comparing the citation impact between country pairs. An exclusion of all publications with more than 10 co-authors is relatively easy to carry out in SciVal.

Adams and Gurney (2018) studied international collaborations explicitly and their results are in line with this study. They recommend an analysis of publications with more than 20 collaborating countries separately, which is somewhat similar to approach 2 in this study (exclude publications with more than 100 co-authors). However, as the five approaches compared in this study show, the rankings may change considerably when focusing more directly on the actual pair of countries.

A final recommendation is to avoid a full count of all publications involving country A and B and their citations, which is the standard approach in SciVal. It gave the highest citation impact and differed the most from the other approaches. For collaborations between small and medium sized countries, the mix of the co-publications has a large influence on the citation impact. For example, if both countries have just one author each involved in a big research consortium such as Atlas with many publications annually and more than 1,000 co-authors per publication, it might improve citation impact for the co-publications

substantially. Even though it in some cases might be relevant to use the full count approach, it must be used with caution and probably not for comparisons of different country pairs.

As always, this study has several limitations, among them that it is based on only a few countries and it is only using Scopus data for one year. The main limitation is closely linked to the purpose of the study, which is to use tools and methods available to policy-makers to carry out the analysis.

On a higher level, one policy implication relating to this study is the role of publication analysis experts versus the availability of tools allowing policymakers to carry out the analyses themselves. Given the rapid development of AI-supported tools to facilitate advanced analyses, this question will probably become even more important in the future. My experience from the user-friendly SciVal is predominantly positive. Even though there are pitfalls such as the one highlighted in this study, a large majority of the insights that SciVal delivers to policymakers and others are not at all problematic. Another benefit of such user-friendly tools is that the use of publication data comes closer to the users of the results, which typically increases the relevance of the analyses.

To conclude the discussion, there are many ways to compare international partnerships, even when only scientific co-publications are used. This study illustrates some important aspects to have in mind. The choice of approach depends on the purpose of the study and the resources available. Based on the tests carried out in this study, it is recommended to use country fractionalisation and as a second-best approach, to limit the co-publication set to include up to 10 co-authors. Elsevier should develop SciVal to make it easier to perform relevant comparisons of country pairs and make sure that SciVal customers do not use the full count approach for this purpose without knowing its potentially misleading results.

Conclusions

International research collaboration is often promoted and there are many good reasons for this. This study addressed one of the questions that a policymaker may ask; with which country should we intensify our collaboration, and compared different approaches to use citation data for the comparison of country pairs. Previous literature forwards many indicators trying to capture the value of a publication. It shows that citations are positively associated with the number of co-authors as well as the number of countries involved. Fractionalised approaches are recommended, which means that the citations are split between the researchers or countries involved. Broadly available tools to analyse publications do not offer fractionalisation. The contribution of this study is that it develops and tests practical approaches to answer the policymaker's question, acknowledging the limitations of available tools such as SciVal from Elsevier. The analysis of empirical data for nine selected countries shows clear differences in how their publications are distributed and confirms that international co-publications, even when correcting for the number of co-authors, exhibit a higher citation impact. Five approaches were tested, and it was highlighted that the standard approach, which is the default one in SciVal, gives the potentially most misleading results. Country-level fractionalisation is recommended for the comparison of country pairs but as this approach is associated with considerable work in SciVal and Excel, a second-best approach using all co-publications with maximum 10 co-authors is also suggested. Elsevier and other providers of tools to study publications should help their users understand the potentially misleading results of the standard approach to study international collaboration and introduce better alternatives.

Conflict of interest

The author declares no external funding and no conflicts of interests.

Acknowledgements The author acknowledges generous support from The Swedish Foundation for International Cooperation in Research and Higher Education, STINT, which as an employer provided resources and inspiration for this study.

Declarations

Funding Open access funding provided by Chalmers University of Technology.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Adams, J., & Gurney, K. A. (2018). Bilateral and multilateral coauthorship and citation impact: Patterns in UK and US international collaboration. *Frontiers in Research Metrics and Analytics*, 3, 12.
- Adams, J., Pendlebury, D., Potter, R. and Szomszor, M. (2019), 'Multi-authorship and research analytics', Global Research Report, Institute for Scientific Information, Clarivate Analytics.
- Aksnes, D. W., Schneider, J. W., & Gunnarsson, M. (2012). Ranking national research systems by citation indicators. A comparative analysis using whole and fractionalised counting methods. *Journal of Informetrics*, 6(1), 36–43.
- Baas, J., Schotten, M., Plume, M., Côté, G., & Karimi, R. (2020). Scopus as a curated, high-quality bibliometric data source for academic research in quantitative science studies. *Quantitative Science Studies*, 1(1), 377–386. https://doi.org/10.1162/qss_a_00019
- Burnham, J. F. (2006). Scopus database: a review. *Biomedical Digital Libraries*. <https://doi.org/10.1186/2F1742-5581-3-1>
- Confraria, H., Mira Godinho, M. and Wang, L. (2016), 'Determinants of citation impact: A comparative analysis of the Global South versus the Global North', UNU-MERIT Working Paper Series, #2016–029, Maastricht, The Netherlands.
- Elsevier (2022), *SciVal*. <https://www.elsevier.com/solutions/scival>
- Gauffriau, M. (2021). Counting methods introduced into the bibliometric research literature 1970–2018: A review. *Quantitative Science Studies*, 2(3), 932–975.
- Glänzel, W. (2001). National characteristics in international scientific co-authorship relations. *Scientometrics*, 51(1), 69–115.
- Glänzel, W., & Schubert, A. (2001). Double effort = Double impact? A critical view at international co-authorship in chemistry. *Scientometrics*, 5(7), 199–214.
- Glänzel, W., Schubert, A., & Czerwon, H. J. (1999). A bibliometric analysis of international scientific cooperation of the European Union (1985–1995). *Scientometrics*, 45(2), 185–202. <https://doi.org/10.1007/2Fb02458432>
- Glänzel, W., & Thijs, B. (2004). Does co-authorship inflate the share of self-citations? *Scientometrics*, 61, 395–404.
- Kahn, M. (2018). Co-authorship as a proxy for collaboration: a cautionary tale. *Science and Public Policy*, 45(1), 117–123. <https://doi.org/10.1093/2Fscipol/2Fscx052>
- Kamalski, J., & Plume, A. (2013). *Comparative benchmarking of European and US research collaboration and researcher mobility: A report prepared in collaboration between Science Europe and Elsevier SciVal Analytics*. Elsevier.
- Perianes-Rodriguez, A., Waltman, L., & van Eck, N. J. (2016). Constructing bibliometric networks: A comparison between full and fractional counting. *Journal of Informetrics*, 10, 1178–1195.

- Pohl, H. (2020). Collaboration with countries with rapidly growing research: Supporting proactive development of international research collaboration. *Scientometrics*, 122(1), 287–307. <https://doi.org/10.1007/2Fs11192-019-03287-6>
- Pohl, H., Warnan, G., & Baas, J. (2014). Level the playing field in scientific collaboration with the use of a new indicator: Field-weighted internationalization score. *Research Trends*, 39, 3–8.
- Potter, R. W. K., Szomszor, M., & Adams, J. (2020). Interpreting CNCIs on a country-scale: The effect of domestic and international collaboration type. *Journal of Informetrics*, 14(4), 101075. <https://doi.org/10.1016/j.joi.2020.101075>
- Purkayastha, A., Palmaro, E., Falk-Krzesinski, H., & Baas, J. (2019). Comparison of two article-level, field-independent citation metrics: Field-weighted citation impact (FWCI) and relative citation ration (RCR). *Journal of Informetrics*, 13, 625–642.
- Puuska, H.-M., Muhonen, R., & Leino, Y. (2014). International and domestic co-publishing and their citation impact in different disciplines. *Scientometrics*, 98, 823–839.
- Thelwall, M. (2020). Large publishing consortia produce higher citation impact research but co-author contributions are hard to evaluate. *Quantitative Science Studies*, 1(1), 290–302. https://doi.org/10.1162/qss_a_00003
- Thelwall, M., & Maflahi, N. (2022). Research coauthorship 1900–2020: Continuous, universal, and ongoing expansion. *Quantitative Science Studies*, 3(2), 331–344.
- Thelwall, M., Kousha, K., Abdoli, M., Stuart, E., Makita, M., Wilson, P., & Levitt, J. (2024). Which international co-authorships produce higher quality journal articles? *J Assoc Inf Sci Technol*. <https://doi.org/10.1002/asi.24881>
- Van Raan, A. F. J. (1998). The influence of international collaboration on the impact of research results. *Scientometrics*, 42(3), 423–428.
- Wagner, C. S., Park, H. W., & Leydesdorff, L. (2015). The continuing growth of global cooperation networks in research: A conundrum for national governments. *PLoS ONE*, 10(7), e0131816. <https://doi.org/10.1371/2Fjournal.pone.0131816>
- Wagner, C. S., Whetsell, T. A., & Leydesdorff, L. (2017). Growth of international collaboration in science: revisiting six specialties. *Scientometrics*, 110(3), 1633–1652. <https://doi.org/10.1007/2Fs11192-016-2230-9>
- Waltman, L., & van Eck, N. J. (2015). Field-normalized citation impact indicators and the choice of an appropriate counting method. *Journal of Informetrics*, 9(4), 872–894.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.