

Not all International Collaboration is Beneficial: The Mendeley Readership and Citation Impact of Biochemical Research Collaboration¹

Pardeep Sud, Mike Thelwall

Statistical Cybermetrics Research Group, School of Mathematics and Computer Science, University of Wolverhampton, Wulfruna Street, Wolverhampton WV1 1SB, UK.

Biochemistry is a highly funded research area that is typified by large research teams and is important for many areas of the life sciences. This article investigates the citation impact and Mendeley readership impact of biochemistry research from 2011 in the Web of Science according to the type of collaboration involved. Negative binomial regression models are used that incorporate, for the first time, the inclusion of specific countries within a team. The results show that, holding other factors constant, larger teams robustly associate with higher impact research, but including additional departments has no effect and adding extra institutions tends to reduce the impact of research. Although international collaboration is apparently not advantageous in general, collaboration with the USA, and perhaps also with some other countries, seems to increase impact. In contrast, collaborations with some other nations associate with lower impact, although both findings could be due to factors such as differing national proportions of excellent researchers. As a methodological implication, simpler statistical models would have found international collaboration to be generally beneficial and so it is important to take into account specific countries when examining collaboration.

Introduction

According to the American Chemical Society, biochemistry is “the study of the structure, composition, and chemical reactions of substances in living systems” (ACS, 2013) and has applications in food science, pharmacology and toxicology. As with all research areas, and particularly those that are highly funded, it is important to identify factors that can help scientists to conduct the highest possible quality research. This may help individual scientists to design better research and policy makers to fund or incentivise more effective types of research. It has long been recognised that collaboration is increasingly common (Wuchty, Jones, & Uzzi, 2007) and essential to many areas of science (Price, 1963). Collaboration also leads to higher citation impact in most areas of science (Persson, Glänzel, & Danell, 2004; Wuchty et al., 2007). Intuitively, collaboration may be beneficial because a combination of different skills, perhaps from different research areas, may be necessary to tackle difficult issues, especially applied problems (Gibbons et al., 1994).

Whilst collaboration in general is known to be beneficial in many research areas, there are different types and they are not all equally advantageous. For example, international teamwork seems to lead to higher impact research in many areas, including biochemistry, but inter-institutional collaboration seems to not affect research impact (Didegah & Thelwall, 2013). It is important to recognise that not all nations are equal in research and that some tend to produce articles with above average impact (Bornmann &

¹ This is a preprint of an article to be published in the Journal of the Association for Information Science and Technology © copyright 2014 John Wiley & Sons, Inc

Leydesdorff, 2013). Hence, findings about the importance of international collaboration may obscure or overly generalise trends at the international level that are in fact a combination of different national-level trends. In addition, citation impact is not the only indicator of the value of research and so other metrics, such as patents (Oppenheim, 2000), downloads (Pinkowitz, 2002) or readership (Li, Thelwall, & Giustini, 2012) for articles may reveal different information about the value of research collaboration in fostering research impact.

In order to identify beneficial and non-beneficial types of collaboration for biochemistry research in a way that avoids the limitations of reliance upon a single impact metric and avoids treating all nations as equal, this article uses two impact measures and regression models that include factors for the presence of specific nations within an authorship team. These models are used to identify general types of collaboration that are beneficial as well as whether international collaboration in general is a good thing, or whether it associates with more citations for some nations but not for others.

The goal of the research is therefore to give new answers to the old question of which types of collaboration associate with higher impact research based upon (i) identifying specific countries involved in international collaboration, rather than just focusing on the extent of collaboration, (ii) using a statistical model that is able to differentiate between collaboration in general and different types of collaboration, and (iii) using multiple impact metrics. Whilst many papers have addressed this question for various fields, and a few have also tackled either (i), (ii) or (iii), none have previously used both (i) and (ii) together, with or without (iii). Biochemistry was chosen for this analysis because it is an area in which collaboration and funding seem to be important. For example, a vast majority (91%) of WoS Biochemistry articles in 2011 declared a funding source and 99% had at least two authors. In addition, biochemistry also has substantial coverage in Mendeley (128,679 articles in the Biochemistry subcategory, as of 29 November, 2013).

- **Research Question:** Does international collaboration in general associate with higher impact research in biochemistry, after accounting for the number of departments and institutions as well as the number of collaborators?

Background

Research collaboration is the contribution of multiple people within a research project, although each may perform different types of work (Katz & Martin, 1997). Contributors to a research project may have separate tasks, such as supervision, data processing, data analysis or literature reviewing. Individual researchers may also tend to play the same specific roles within their collaborations, such as mentoring junior researchers (Bozeman & Corley, 2004) or providing advice (Sonnenwald, 2007). Not all forms of collaboration lead to co-authorship (Cronin, 2012; Cronin, Shaw, & La Barre, 2003), however, and a decision about whether someone should be listed as an author may depend on field norms, specific agreements or the whim of the main author, and authorship may even be assigned on an honorary basis (Flanagin et al., 1998).

Research team size

Throughout most areas of science, articles with larger numbers of co-authors tend to attract more citations (Frenken, Hölzl, & de Vor, 2005; Glänzel, 2002; Leimu & Koricheva, 2005; Persson et al., 2004; Wuchty et al., 2007) although there is no universally acknowledged reason why this is the case and the underlying reasons may vary between fields. One general possibility is that collaborative research is more cited because it is of higher quality.

Alternatively, it may attract more citations as a by-product of the larger number of co-authors available to publicise or self-cite it, although this does not seem to be a large factor (Aksnes, 2003b). It is also possible that researchers that collaborate more attract more citations from former collaborators (Wallace, Larivière, & Gingras, 2012). One large scale multidisciplinary study of Italian research found that articles with more authors tended to attract more citations and better peer review judgements, with the latter suggesting that collaboration genuinely benefits research (Franceschet & Costantini, 2010), at least in terms of the quality of the scholarly outputs produced.

Collaborations could theoretically be more powerful because of the greater range of skills within the team (Gibbons et al., 1994) or because better researchers tend to collaborate more and hence are more likely to be in a team. As perhaps an example of the latter, large research grants tend to be awarded to more collaborative researchers (Bozeman & Corley, 2004). Conversely, independence is valued for researchers (Fox & Faver, 1984), particularly in the humanities, and it seems that collaborations will diminish this. Moreover, a study of highly cited library and information science researchers found that their collaborative papers were not more highly cited than their solo articles, suggesting that the ability of good researchers to participate in teams may be more important than the participation itself (Levitt & Thelwall, 2009). In other words, better researchers may participate more in teams and the teams may produce better work as a result, even though those researchers may have produced equally good research had they chosen not to collaborate.

From a statistical perspective, larger teams may tend to produce more highly cited articles because, other factors being equal, they are more likely to contain an eminent or senior author, *the* presence of which is known to associate with more citations in some areas (Haslam et al., 2008). In other words, the size of a larger team may be less important than its increased likelihood of containing an important researcher.

The internationalism of research collaboration

The importance of international collaboration has been recognised for a long time. For example, one early study found that internationally collaborative European research had a citation impact that matched that of the USA (Narin, Stevens, & Whitlow, 1991) and Chinese-authored molecular biology is substantially more highly cited if internationally collaborative (Ma & Guan, 2005). Similarly, highly cited research tends to be more international and more collaborative than average (Aksnes, 2003a), and the more countries contributing to co-authoring an article, the more citations it is likely to attract (Bote, Vicente, Olmeda-Gómez, & Moya-Anegón, 2013). Although international research seems to receive more self-citations than national research, the difference is too small to account for its higher citation impact, at least for astronomy (Van Raan, 1998). Nevertheless, international collaboration might help to give an article a wider audience, potentially attracting disproportionately many citations from all countries represented in the authorship team (Lancho-Barrantes, Bote, Vicente, Rodríguez, & de Moya Anegón, 2012).

International collaboration might associate with more citations because it is more likely to be published in higher impact journals, for example if they are international rather than national in scope. This does not seem to be the case, however, since a regression model found that international collaboration associated with more citations (for Biology & Biochemistry, Chemistry and Social Sciences) even after taking into account the impact factors of the journals publishing the articles (Didegah & Thelwall, 2013). This model also

took into account a range of other document properties (e.g., abstract readability) and so provides robust evidence of the importance of internationalism in research.

At the level of individual institutional research teams, international collaboration associates with increased success, at least for European life sciences research, and this extends to hidden international collaboration in the sense of employing foreign researchers within a research group (Barjak & Robinson, 2007).

International collaboration is not equally beneficial to all countries, with some countries benefiting much more from it than others (Glänzel, 2001; Lancho-Barrantes, Guerrero-Bote, & de Moya-Anegón, 2013). Moreover, the extent to which international collaboration is beneficial can vary between fields within an individual country (Glänzel, 2001). In support of this, a later study of economics found that international collaboration and inter-state collaboration within the USA tended to be beneficial for most countries and states, but the extent of the benefit varied greatly between them (Levitt & Thelwall, 2010).

Some studies have differentiated between the types of countries forming international collaborations. An investigation of Columbia found, counterintuitively, that collaborations with poorer nations tended to produce more outputs than did collaborations with richer nations (Ordonez-Matamoros, Cozzens, & Garcia-Luque, 2011).

Impact metrics

Probably the ultimate gold-standard for the value of an academic article is peer judgements, although these have limitations, such as the potential for nepotism or otherwise biased or uninformed conclusions (Lee, Sugimoto, Zhang, & Cronin, 2013). Peer judgements are rarely available on a large scale, however, which is a practical limitation for many studies. This may change to some extent, however, if national research assessment exercises start to publish formal evaluations of research (Franceschet & Costantini, 2010).

Most empirical studies of the value of research collaboration have used citation impact as their sole indicator of value. Although citations are an established and reasonable indicator of the impact of scientific work, they have a number of technical (Meho & Yang, 2007) and conceptual (Seglen, 1998) limitations, some of which are relevant to assessing the value of collaboration. In addition to the issue of self-citations discussed above, citation counts may undervalue applied research that is found useful by people that are not publishing scientists, but this problem can be ameliorated for inter-field comparisons by normalising citations by field (Moed, 2010). Interdisciplinary collaborative research has been claimed to be particularly useful for solving applied problems (Gibbons et al., 1994) and so may be particularly disadvantaged by citation-based indicators.

Another way to assess the impact of an article is to count how many times it has been read. This can be achieved to some extent with download statistics (Moed, 2005), although such values ignore the readers of print versions of articles and assume that all downloaders are humans and that they read the downloaded papers. A practical limitation is that download statistics are not freely available for all academic articles and are also not always fully comparable between all academic publishers (COUNTER, 2011), although they can be useful for collections of articles in specific repositories (Shuai, Pepe, & Bollen, 2012).

Readership indicators can be gathered from free websites in which people register articles that they have an interest in, including online reference managers. Although users of such sites are presumably a minority of article readers and the uptake of such resources probably varies between fields and between nations, it seems that social web article readership statistics should provide useful alternative metrics for article impact within

individual fields. Although many other altmetrics (Priem, Piwowar, & Hemminger, 2012) are available, social reference site readership seems to be the most widespread (other than Twitter) (Thelwall, Haustein, Larivière, & Sugimoto, 2013) and registering an article in Mendeley seems to be a better indicator of use than tweeting, since Twitter contains spam (e.g., Thomas, Grier, Song, & Paxson, 2011) and typical tweets of academic articles merely echo article titles or a very brief summary (Thelwall, Tsou, Weingart, Holmberg, & Haustein, 2013). In particular, several studies have shown Mendeley readership counts to correlate with article citations (Li & Thelwall, 2012; X. Li et al., 2012; Mohammadi & Thelwall, in press).

Data and Methods

Data was gathered 25-27 November 2013 from Mendeley and WoS for pure biochemistry articles published in 2011. The year 2011 was chosen as a compromise between citations and readers. An earlier year would have given more time for the articles to accumulate citations and hence would give more statistically powerful citation data. In contrast, earlier years would probably have given lower Mendeley readership counts because of the lower Mendeley user base, assuming that members tend to bookmark current articles much more frequently than older articles. Mendeley usage appeared to peak and stabilise in late 2011, according to Google Trends (for the query Mendeley) and so it seems likely that articles from 2012 would have more readers but articles from 2011 should not have substantially fewer Mendeley readers, and articles from 2012 could be expected to have substantially fewer citations at the time of data collection (November 2013).

Records for all biochemistry publications of document type *article* from the Web of Science were downloaded, after eliminating articles that were also classified as something else (e.g., editorial material, review). Articles classified within additional subject areas were also removed before downloading, so the basic data set consists of documents within WoS that were published in 2011 and have the single classification Biochemistry and the single document type classification as article (n=13578). Multiply-classified articles were removed to reduce the risk that the results would be due to differing patterns of collaboration for different biochemistry subfields. In other words, in any multi-field collection of documents, any statistical differences found between countries (e.g., in citation or readership rates) may be due to countries specialising in fields with different properties (e.g., country 1 specialising in a highly read field, and country 2 specialising in a little read field, leading to a [correct but unhelpful] conclusion that country 2's research is read less). The removal of multiply classified articles reduces but does not eliminate this risk. In general, it is desirable to have as homogenous a sample of documents as possible to reduce the chance that statistical differences in the sample are spurious.

For each article extracted, the number of different authors, departments, institutions and countries represented was counted by processing the address field (C1). Institutions were taken to be the first entity named in the comma-separated address list, departments were taken to be the second and countries to be the last (state affiliations were excluded for the USA). Manual inspection of the results suggested that this was correct, although in a few cases of non-university institutions, the institution was a company and the department was a geographic location, presumably of a laboratory. The departments had various different subunit names, including department (36%), institute (13%), school (12%), centre (8%), lab (6%) and college (5%). If authors had more than one affiliation then only the first affiliation was recorded on the assumption that the first listed

affiliation was likely to be the most important. The address fields did not always contain addresses for all authors and so the number of authors in the author field was matched with the number of authors with addresses in the address field and any mismatches were flagged and the associated records removed (1369 cases out of 13578, leaving 12209 as the basic sample size for the citation models). Also flagged were cases of authors with multiple addresses. Records were then marked when any of the addresses were from any of the ten most numerous countries in the data set (USA, China, UK, Germany, Japan, France, Canada, Italy, India, Spain), counting only the first addresses for authors with multiple addresses. Papers where the first author's first address was in the USA were also noted. Full text matching was used to identify the number of different countries, institutions and departments for each paper on the assumption that these would tend to be standardised within individual papers even though department and institution names are known to not be standard between papers.

Mendeley queries were constructed to identify records, if any, for each WoS biochemistry article. Each query consisted of the full article title in quotes, together with the first author last name and the publication year (i.e., 2011). In a few cases (406 or 3%) multiple records matched and only the first one was used. Manual checks did not identify any false matches between the queries and the documents returned by Mendeley. For example, the paper with the most identical results "Glycome Diagnosis of Human Induced Pluripotent Stem Cells Using Lectin Microarray" had 8 correct matches and no false matches. The total number of Mendeley readers was recorded for each document identified. In some cases, no document was found in Mendeley. This could indicate that there were no Mendeley readers or that the query failed, for example due to a typo in WoS or Mendeley, or due to a complex title that caused the queries to fail. Hence the zeros in this dataset were ignored on the assumption that an unknown proportion of them were false. After this step, an additional 2601 records were removed and the final sample size for readership analysis was 9608 articles with Mendeley readers and consistent author affiliations (2905 out of the original 13578 articles had no Mendeley readers found).

Because the goals of the analysis include differentiating between different correlating factors, regression analysis is needed rather than simple correlation tests. Citation data is known to be highly skewed and the dependant variable (either citations or readers) is ordinal and so count data regression is more appropriate than linear regression. The default model for count data regression uses the Poisson distribution but this was not appropriate because the dispersion of the data was greater than the mean (almost double for both citations [14.3 vs. 7.5] and for readers [22.6 vs. 12.4], although the latter ignores the zero truncation) and so a negative binomial model was used instead. Citation data sometimes have excess zeros and so both zero inflated and standard models were tested. The difference between them was statistically insignificant and visual inspection of the data suggested that there did not seem to be too many zeros and so a standard negative binomial model was used (for other disciplines the choice of model might be clearer). For the Mendeley readership data, all the zeros were removed from the data since some could be false and a zero-truncated negative binomial model was applied instead.

The coefficient estimates in regression models can be unreliable when sets of variables are close to being linear combinations of each other (i.e., collinearity). A common metric to test for this is the Variance Inflation Factor (VIF). This was calculated for all the variables in the data set and the highest values were 5.6 (institution count), 5.5 (department count) and 2.2 (country count) with all other values being below 1.7. A common rule of

thumb is that VIF values above 10 indicate serious collinearity (O'Brien, 2007) and so the data is acceptable but these values indicate that the department and institution count coefficients in the models below should be treated with an element of caution because the model might "confuse" them between each other a little (they have a Pearson correlation with each other of 0.891).

For both impact indicators (citations and readership) complete models were fitted that included all of the hypothesised factors. In addition, models were also fitted with subsets of the data to give additional tests of the importance of individual factors and to assess whether partially misleading results could be obtained from omitting some factors. The original models are available on figshare.org² and the coefficients are transformed into more easily interpretable values in the results section.

Results

As reported in Table 1, the general collaboration variables all correlate positively with the number of readers and the number of citations, although the correlations are not high. Using the standard models that exclude individual countries (tables 2 and 3; *All valid data* models), it can be seen that more citations and readers significantly associate with larger numbers of authors, smaller numbers of institutions and larger numbers of countries. In comparison to the correlations (Table 1), the model is able to show that increasing the number of institutions within a project without increasing the number of authors, countries or departments will lower the expected citation and readership impact of the research. It is also able to show that, other factors being equal, the number of departments represented in a project is irrelevant to readership and citation impact.

Table 1. Spearman correlations for all articles with all author addresses extracted and Mendeley records found (n=9608)*.

	Citations	Readers	Authors	Depts.	Institutions	Countries
Citations	1	0.534	0.121	0.076	0.058	0.083
Readers		1	0.044	0.031	0.020	0.068
Authors			1	0.594	0.529	0.311
Depts.				1	0.864	0.489
Institutions					1	0.572
Countries						1

*All correlations are significant at p=0.001.

When restricting the data to collaborations with participation from the USA or with a first author from the USA, the results are different (tables 2 and 3; *Includes USA* and *USA first author* models). In these cases the number of countries is irrelevant, rather than being a substantial and highly significant factor. The results are similar for collaborations including the UK (tables 2 and 3; *Includes UK* models), but additional departments are a disadvantage whereas additional institutions and countries might be an advantage – perhaps reflecting a different research culture in the UK to the USA. For collaborations including China, the only

2

important factor appears to be getting other countries involved (tables 2 and 3; *Includes China* models).

Table 2. Coefficients for negative binomial regressions for the number of citations. Each column represents a single model and each percentage is the change in the expected mean citations for a unit increase in the variable. All models include the same variables but use different sets of articles based on their author affiliations, as described in the column headings.

	Model 1: All valid data	Model 2: Includes USA	Model 3: USA first author	Model 4: Includes China	Model 5: Includes UK
Authors	7%***	8%***	8%***	5%***	6%***
Depts.	0%	-2%	-1%	2%	-17%**
Institutions	-9%***	-6%**	-7%**	-8%	11%
Countries	17%***	4%	-6%	60%***	5%
Sample size	12209	5003	4191	1480	1095

***sig. at p=0.001, **sig. at p=0.01, *sig. at p=0.05

Table 3. Coefficients for zero truncated negative binomial regressions for the number of readers. Each column represents a single model and each percentage is the change in the expected mean readers for a unit increase in the variable. All models include the same variables but use different sets of articles based on their author affiliations, as described in the column headings.

	Model 1: All valid data	Model 2: Includes USA	Model 3: USA first author	Model 4: Includes China	Model 5: Includes UK
Authors	5%***	6%***	6%***	5%**	5%***
Depts.	1%	3%	2%	10%	-16%*
Institutions	-9%***	-8%*	-10%**	-15%	14%*
Countries	19%***	6%	-1%	67%***	14%*
Sample size	9608	4147	3499	1067	808

***sig. at p=0.001, **sig. at p=0.01, *sig. at p=0.05, all estimated from the t-distribution.

The full model that includes variables for the presence of the ten countries most represented in biochemical research (i.e., a binary variable for each country indicating the presence of at least one author from the country) confirms that the number of countries in a project is not important but that specific countries are important (tables 4 and 5; *All valid data* models). Whilst the total number of countries in a project is not significant, the presence of several countries is, with some associating with more citations and readers and some associating with fewer citations and readers. Hence the *countries* factor in the basic model for all valid data obscures the underlying pattern that more international collaborations are likely to produce higher impact research because they are more likely to include a high impact country (e.g., the USA) rather than because additional countries are always beneficial.

A model was built including only papers where all authors had a single affiliation (tables 4 and 5; *Single affiliation* models) to check whether the results might be affected by

ignoring additional affiliations for authors. The model coefficients are broadly similar, suggesting that the overall conclusions would not be affected by the decision to exclude the additional affiliations.

Considering again only collaborations involving the USA (tables 4 and 5; *Includes USA* models), it seems that collaborations benefit only from the inclusion of the UK, and possibly also France and Germany, but including India and China might be a disadvantage (from a statistical perspective, but see the conclusions for a fuller discussion of the issue). If the first author is from the USA, however (tables 4 and 5; *USA first author* models), no country gives a statistically significant disadvantage but collaboration with the UK, Germany, Canada and Spain may generate higher impact work. The results for the UK and China are less reliable due to the lower number of data points but it is clear that both countries benefit from collaboration with the USA.

Table 4. Coefficients for negative binomial regressions for the number of citations. Each column represents a single model and each percentage is the change in the expected mean citations for a unit increase in the variable. All models include the same variables (except where they are constant) but use different sets of articles based on their author affiliations, as described in the column headings.

	Model 1: All valid data	Model 2: Single affiliation	Model 3: Includes USA	Model 4: USA first author	Model 5: Includes China	Model 6: Includes UK
Authors	7%***	7%***	8%***	8%***	5%***	6%***
Depts.	-1%	-5%*	-1%	-1%	1%	-13%*
Institutions	-8%***	-5%.	-7%**	-7%**	-7%	7%
Countries	-1%	-7%*	-6%.	-7%	13%	-12%*
USA incl.	54%***	64%***	-	-	71%***	59%***
China Incl.	-21%***	-22%***	4%	0%	-	-17%
UK incl.	61%***	87%***	60%***	17%	30%	-
Germany Incl.	49%***	57%***	14%.	-1%	112%*	42%**
Japan Incl.	-1%	8%	-1%	1%	7%	-22%
France Incl.	13%**	19%*	21%*	4%	15%	30%.
Canada Incl.	17%***	14%*	6%	0%	35%	77%***
Italy Incl.	0%	13%	-12%	-19%	-81%.	21%
India Incl.	-12%*	-3%	-7%	9%	-90%	-38%
Spain Incl.	-1%	-6%	-4%	-6%	-19%	-43%**
Sample size	12209	6752	5003	4191	1480	1095

***sig. at p=0.001, **sig. at p=0.01, *sig. at p=0.05 (all estimated from the t-distribution) – constant in the model.

Table 5. Coefficients for zero truncated negative binomial regressions for the number of readers. Each column represents a single model and each percentage is the change in the expected mean readers for a unit increase in the variable. All models include the same variables (except where they are constant) but use different sets of articles based on their author affiliations, as described in the column headings.

	Model 1: All valid data	Model 2: Single affiliation	Model 3: Includes USA	Model 4: USA first author	Model 5: Includes China	Model 6: Includes UK
Authors	6%***	6%***	6%***	6%***	5%**	5%***
Depts.	0%	-8%**	2%	1%	10%	-13%*
Institutions	-8%***	-1%	-8%*	-10%**	-14%	13%*
Countries	2%	-6%	-2%	-11%*	20%	4%
USA incl.	46%***	55%***	-	-	66%*	38%**
China Incl.	-40%***	-44%***	-24%*	-5%	-	-23%*
UK incl.	47%***	68%***	45%***	32%*	63%	-
Germany Incl.	60%***	72%***	21%*	48%**	38%	20%*
Japan Incl.	-2%	5%	4%	6%	-37%	-20%*
France Incl.	5%	31%**	9%	5%	39%	1%
Canada Incl.	13%*	8%	-7%	19%*	77%	48%*
Italy Incl.	-21%***	-18%*	-8%	-4%	-100%	-24%*
India Incl.	-37%***	-29%***	-61%**	-14%	-77%	-69%*
Spain Incl.	16%*	-1%	-7%	25%*	-82%	-39%*
Sample size	9608	5182	4147	3499	1067	808

***sig. at $p=0.001$, **sig. at $p=0.01$, *sig. at $p=0.05$ (all estimated from the t-distribution) – constant in the model..

Limitations

This study covers only one year (2011) and only WoS publications in biochemistry and so the results may be different for other years and may change over time. For example, the apparent disadvantage of collaborating with China may disappear as its science system matures. Another limitation is that the two impact indicators used may be biased towards the USA because WoS is known to have national and language biases that favour the USA (Van Leeuwen, Moed, Tijssen et al., 2001) and the USA often seems to be an early user of new web technologies so it may be the largest user of Mendeley. Moreover, both indicators may not fully reflect applied contributions to research and so it is possible, for example, that Chinese research that is not internationally collaborative tends to focus on the needs of domestic industries. The results also apply only to relatively pure biochemistry research that has been published in a journal that is only in the Biochemistry WoS category. Finally, the results only concern one area of science and, although it seems likely that similar findings would be found widely within the physical and medical sciences in WoS, this hypothesis has not been tested.

Conclusions

Confirming previous similar studies, the results show that larger collaborations are more successful at attracting citation impact in biochemistry but that, other factors being equal, additional institutions tend to associate with fewer citations.

Overall and other factors being equal, the results suggest that biochemists in the USA should seek, and funders should encourage, larger collaborations but not more highly multi-institutional collaborations. For such researchers, inter-institutional collaboration may be a disadvantage but there is no reason to encourage or discourage inter-departmental collaboration and international collaboration. In support of the latter point, a previous study found Harvard-authored publications not to benefit from international collaboration (Gazni & Didegah, 2011). In contrast, whilst biochemists outside the USA should seek to develop larger teams for projects, the number of institutions may not be relevant but the model suggests they should seek to collaborate with the USA or another nation that conducts high impact biochemistry research. Although the model has not tested for this, it seems likely that this result should not be taken as blanket encouragement to collaborate with Americans, but reflects the fact that more successful researchers are more likely to be found in the USA, either because they are better or because they are more highly funded, or both. In practice, the underlying lesson may be the importance of collaborating with successful researchers, which tend to be in the USA, at least for biochemistry. Similarly, countries that appeared not to be advantageous to collaborate with may simply have a lower proportion of excellent researchers at the moment and so any blanket encouragement to avoid collaborating with them would be inappropriate, at least without substantial additional evidence. These conclusions are of a statistical nature and should be interpreted as advice to contribute to the decision making process for building research teams and should not override individual decisions that stem from the nature of a particular project.

An important implication of the methods used to derive the findings reported above is that statistical methods can suggest that international collaboration itself is an advantage whereas a more detailed model *with the same data* finds that international collaboration is only an advantage with specific nations. This casts doubt on the many previous findings about the generic value of international collaboration based upon models that don't differentiate between nations. This finding appears to contradict that of a previous much larger-scale study with a more internationally balanced data source (Scopus) which found that international collaboration was beneficial for all countries (Bote et al., 2013). The later study did not take into account the size of collaboration teams as a separate factor, however, and so the universal advantage that it found for international collaboration could be, at least in part, a generic citation impact advantage for more collaborative research. Combining the results of this article (Bote et al., 2013) and the current article, international collaboration may statistically always be advantageous compared to not collaborating, but it is not always advantageous compared to domestic collaboration.

For future research it would be useful to test a range of different research fields to see whether similar lessons can be found in each case. In particular, it would be useful to identify whether international collaboration is ever beneficial irrespective of the capabilities of the countries involved, which seems unlikely. It would also be helpful to assess whether the country-specific advice in the current article would break down, as seems likely, at the level of individual scientists. In other words, as discussed above, is the underlying factor the capabilities of, and geographic distribution of, individual scientists, rather than generic national characteristics?

References

- ACS (2013). Biochemistry. Retrieved, 2013, from <http://www.acs.org/content/acs/en/careers/whatchemistsdo/careers/biochemistry.html>
- Aksnes, D. W. (2003a). Characteristics of highly cited papers. *Research Evaluation*, 12(3), 159-170.
- Aksnes, D. W. (2003b). A macro study of self-citation. *Scientometrics*, 56(2), 235-246.
- Barjak, F., & Robinson, S. (2007). International collaboration, mobility and team diversity in the life sciences: Impact on research performance. *Social Geography Discussions*, 3(1), 121-157.
- Bornmann, L., & Leydesdorff, L. (2013). Macro-indicators of citation impacts of six prolific countries: InCites data and the statistical significance of trends. *PloS ONE*, 8(2), e56768. doi:10.1371/journal.pone.0056768
- Bote, G., Vicente, P., Olmeda-Gómez, C., & Moya-Anegón, F. (2013). Quantifying the benefits of international scientific collaboration. *Journal of the American Society for Information Science and Technology*, 64(2), 392-404.
- Bozeman, B., & Corley, E. (2004). Scientists' collaboration strategies: Implications for scientific and technical human capital. *Research Policy*, 33(4), 599-616.
- COUNTER. (2011). Counting online usage of networked electronic resources. Retrieved December 14, 2011 from: <http://www.projectcounter.org/>.
- Cronin, B. (2012). Collaboration in art and science: Approaches to attribution, authorship and acknowledgment. *Information & Culture*, 47(1), 18-37.
- Cronin, B., Shaw, D., & La Barre, K. (2003). A cast of thousands: Coauthorship and subauthorship collaboration in the 20th century as manifested in the scholarly journal literature of psychology and philosophy. *Journal of the American Society for Information Science and Technology*, 54(9), 855-871.
- Didegah, F., & Thelwall, M. (2013). Which factors help authors produce the highest impact research? Collaboration, journal and document properties. *Journal of Informetrics*, 7(4), 861-873.
- Flanagin, A., Carey, L. A., Fontanarosa, P. B., Phillips, S. G., Pace, B. P., Lundberg, G. D., & Rennie, D. (1998). Prevalence of articles with honorary authors and ghost authors in peer-reviewed medical journals. *JAMA: The Journal of the American Medical Association*, 280(3), 222-224.
- Fox, M. F., & Faver, C. A. (1984). Independence and cooperation in research: The motivations and costs of collaboration. *The Journal of Higher Education*, 55(3), 347-359.
- Franceschet, M., & Costantini, A. (2010). The effect of scholar collaboration on impact and quality of academic papers. *Journal of Informetrics*, 4(4), 540-553.
- Frenken, K., Hözl, W., & de Vor, F. (2005). The citation impact of research collaborations: The case of European biotechnology and applied microbiology (1988–2002). *Journal of Engineering and Technology Management*, 22(1), 9-30.
- Gazni, A., & Didegah, F. (2011). Investigating different types of research collaboration and citation impact: A case study of Harvard University's publications. *Scientometrics*, 87(2), 251-265.
- Gibbons, M., Limoges, C., Nowotny, H., Schwartzman, S., Scott, P., & Trow, M. (1994). *The new production of knowledge*. London, UK: Sage.
- Glänzel, W. (2001). National characteristics in international scientific co-authorship relations. *Scientometrics*, 51(1), 69-115.

- Glänzel, W. (2002). Coauthorship patterns and trends in the sciences (1980-1998): A bibliometric study with implications for database indexing and search strategies. *Library Trends*, 50(3), 461-473.
- Haslam, N., Ban, L., Kaufmann, L., Loughnan, S., Peters, L., Whelan, S., & Wilson, A. (2008). What makes an article influential? Predicting impact in social and personality psychology. *Scientometrics*, 76(1), 169-185.
- Katz, J. S., & Martin, B. R. (1997). What is research collaboration? *Research Policy*, 26(1), 1-18.
- Lancho-Barrantes, B. S., Bote, G., Vicente, P., Rodríguez, Z. C., & de Moya Anegón, F. (2012). Citation flows in the zones of influence of scientific collaborations. *Journal of the American Society for Information Science and Technology*, 63(3), 481-489.
- Lancho-Barrantes, B. S., Guerrero-Bote, V. P., & de Moya-Anegón, F. (2013). Citation increments between collaborating countries. *Scientometrics*, 94(3), 817-831.
- Lee, C. J., Sugimoto, C. R., Zhang, G., & Cronin, B. (2013). Bias in peer review. *Journal of the American Society for Information Science and Technology*, 64(1), 2-17.
- Leimu, R., & Koricheva, J. (2005). Does scientific collaboration increase the impact of ecological articles? *Bioscience*, 55(5), 438-443.
- Levitt, J., & Thelwall, M. (2009). Citation levels and collaboration within library and information science. *Journal of the American Society for Information Science and Technology*, 60(3), 434-442.
- Levitt, J., & Thelwall, M. (2010). Does the higher citation of collaborative research differ from region to region? A case study of economics. *Scientometrics*, 85(1), 171-183.
- Li, X., & Thelwall, M. (2012). F1000, Mendeley and traditional bibliometric indicators. In É. Archambault, Y. Gingras & V. Larivière (Eds.), *17th international conference on science and technology indicators (STI 2012)* (pp. 541-551). Montréal: Science-Metrix and OST.
- Li, X., Thelwall, M., & Giustini, D. (2012). Validating online reference managers for scholarly impact measurement. *Scientometrics*, 91(2), 461-471.
- Ma, N., & Guan, J. C. (2005). An exploratory study on collaboration profiles of Chinese publications in molecular biology. *Scientometrics*, 65(3), 343-355.
- Meho, L. I., & Yang, K. (2007). Impact of data sources on citation counts and rankings of LIS faculty: Web of Science versus Scopus and Google Scholar. *Journal of the American Society for Information Science & Technology*, 58(13), 2105-2125.
- Moed, H. F. (2005). Statistical relationships between downloads and citations at the level of individual documents within a single journal. *Journal of the American Society for Information Science & Technology*, 56(10), 1088-1097.
- Moed, H. F. (2010). Measuring contextual citation impact of scientific journals. *Journal of Informetrics*, 4(3), 265-277.
- Mohammadi, E., & Thelwall, M. (in press). Mendeley readership altmetrics for the social sciences and humanities: Research evaluation and knowledge flows. *Journal of the American Society for Information Science and Technology*,
- Narin, F., Stevens, K., & Whitlow, E. S. (1991). Scientific co-operation in Europe and the citation of multinationally authored papers. *Scientometrics*, 21(3), 313-323.
- O'Brien, R. (2007). A caution regarding rules of thumb for Variance Inflation Factors. *Quality & Quantity*, 41(5), 673-690.
- Oppenheim, C. (2000). Do patent citations count? In B. Cronin, & H. B. Atkins (Eds.), *The web of knowledge: A festschrift in honor of Eugene Garfield* (pp. 405-432). Metford, NJ: Information Today Inc. ASIS Monograph Series.

- Ordonez-Matamoros, G., Cozzens, S. E., & Garcia-Luque, M. (2011). North-South and South-South research collaboration: What differences does it make for developing countries?-the case of Colombia. In *2011 Atlanta Conference on Science and Innovation Policy*, (pp. 1-10). Los Alamitos: IEEE Press.
- Persson, O., Glänzel, W., & Danell, R. (2004). Inflationary bibliometric values: The role of scientific collaboration and the need for relative indicators in evaluative studies. *Scientometrics*, *60*(3), 421-432.
- Pinkowitz, L. (2002). Research dissemination and impact: Evidence from web site downloads. *Journal of Finance*, *57*(1), 485-499.
- Price, D. J. d. S. (1963). *Little science, big science*. New York: Columbia University Press.
- Priem, J., Piwowar, H. A., & Hemminger, B. M. (2012). Altmetrics in the wild: Using social media to explore scholarly impact. *ArXiv.Org*, <http://arxiv.org/abs/1203.4745v1>.
- Seglen, P. O. (1998). Citation rates and journal impact factors are not suitable for evaluation of research. *ACTA Orthopaedica Scandinavica*, *69*(3), 224-229.
- Shuai, X., Pepe, A., & Bollen, J. (2012). How the scientific community reacts to newly submitted preprints: Article downloads, Twitter mentions, and citations. *PloS ONE*, *7*(11), e47523. doi:10.1371/journal.pone.0047523
- Sonnenwald, D. H. (2007). Scientific collaboration. In B. Cronin (Ed.), *Annual Review of Information Science and Technology* (41st ed., pp. 643-681). Medford, NJ: Information Today.
- Thelwall, M., Haustein, S., Larivière, V., & Sugimoto, C. R. (2013). Do altmetrics work? Twitter and ten other social web services. *PloS ONE*, *8*(5), e64841. doi:10.1371/journal.pone.0064841
- Thelwall, M., Tsou, A., Weingart, S., Holmberg, K., & Haustein, S. (2013). Tweeting links to academic articles. *Cybermetrics*, *17*(1), <http://cybermetrics.cindoc.csic.es/articles/v17i1p1.html>.
- Thomas, K., Grier, C., Song, D., & Paxson, V. (2011). Suspended accounts in retrospect: an analysis of twitter spam. In *Proceedings of the 2011 ACM SIGCOMM conference on Internet measurement* (pp. 243-258). New York: ACM Press.
- Van Leeuwen, T. N., Moed, H. F., Tijssen, R. J., Visser, M. S., & Van Raan, A. F. (2001). Language biases in the coverage of the Science Citation Index and its consequences for international comparisons of national research performance. *Scientometrics*, *51*(1), 335-346.
- Van Raan, A. F. J. (1998). The influence of international collaboration on the impact of research results. *Scientometrics*, *42*(3), 423-428.
- Wallace, M. L., Larivière, V., & Gingras, Y. (2012). A small world of citations? The influence of collaboration networks on citation practices. *PloS ONE*, *7*(3), e33339.
- Wuchty, S., Jones, B., & Uzzi, B. (2007). The increasing dominance of teams in production of knowledge. *Science*, *316*, 1036-1038.