Author's response to reviews

Title: Geographic trends of scientific output and citation practices in psychiatry

Authors:

Artemis Igoumenou (a.igoumenou@qmul.ac.uk)
Klaus Ebmeier (klaus.ebmeier@psych.ox.ac.uk)
Nia Roberts (nia.roberts@bodleian.ox.ac.uk)
Seena Fazel (seena.fazel@psych.ox.ac.uk)

Version: 2
Date: 20 October 2014

Author's response to reviews: see over
Dear Editor in Chief,

We are grateful for the opportunity to respond to the comments of the three reviewers. Their comments were very helpful and we think our revised paper has improved considerably as a consequence.

We will address the recommended revisions point by point.

Reviewer 1; Ms Samantha Horswill

Major Compulsory Revisions

1. “It is unclear why the authors chose to analyse the entire European continent together, but did not include Canada in with the American analysis.”

We wanted to test a hypothesis that seems to relate specifically to US researchers, which is cited in the introduction. Nevertheless, the reviewer may be correct that North America could be considered as a whole, and we have done this in a sensitivity analysis. North America published 20.061 records during the study period, receiving 226.822 citations (TGCS), hence having 11.3 citations per paper. We repeated our analysis of 1000 selected articles including North American and European articles. The results are shown on Appendix 4, as follows:

<table>
<thead>
<tr>
<th></th>
<th>North America</th>
<th>Europe</th>
<th>Rest of the world</th>
<th>Anonymous</th>
<th>Total citations (n= 3393)</th>
</tr>
</thead>
<tbody>
<tr>
<td>North America</td>
<td>1294 (0.54)</td>
<td>675 (0.28)</td>
<td>324 (0.14)</td>
<td>84 (0.04)</td>
<td>2377 (0.70)</td>
</tr>
<tr>
<td>Europe</td>
<td>248 (0.24)</td>
<td>596 (0.59)</td>
<td>144 (0.14)</td>
<td>28 (0.03)</td>
<td>1016 (0.30)</td>
</tr>
</tbody>
</table>
As shown on Appendix 4, the selected 1,000 records (500 from North America and 500 from Europe) were cited on 3393 occasions. Seventy per cent (n=2,377) of these 3393 records cited North American articles and 30% (n=1,016) European. We found that North Americans cite North Americans more often (54%, n=1,294) and Europeans cite more European papers (59%, n= 596) ($x^2=314$, d.f.=1, p<0.001). Both Americans and Europeans receive citations from the rest of the world less often than from their own continent (n= 324, 14% and n= 144, 14% respectively; $x^2=314$, d.f.=2, p<0.001).

We also explored numbers of self-citations. From the total of 3,393 citations that our 1,000 records received, 691 (20%) were self-citations. We found no difference between the self-citation practices of the Americans (20% of their citations were self-citations) and the Europeans (21% were self-citations) ($x^2=0.01$, d.f.=1, p=0.919).

In addition, we investigated how many records of the 3,393 publications were international research collaborations and found that 22% (n=746) were collaborations between two or more countries.

The above results are similar to those between US and Europe (Table 5). It seems that considering North America (US and Canada) rather than US alone does not materially alter our conclusions.

2. “Further to that point, it is unclear why the additional analysis only included American and European articles when Canada performed better than the UK in the citations/paper metric. The authors are careful throughout the article not to state a preference for any one type of metric, but given that only American and European
analysis are conducted the implication is that total global citation score and number of records is better”

As in point 1, the reviewer may be correct that North America could be considered as a whole, and we have done this in a sensitivity analysis. We have addressed this point together with point 1.

3. “The authors never discuss whether potential biases may have influenced the 1000-paper sampling technique. Again, the inclusion of other countries (Canada, Australia) from is curious given the higher population count than Europe excluding Germany/UK.”

This is a helpful comment and we have addressed it in our revised Discussion (p. 16, line 338-343)

“In our sub-study we selectively looked at 1000 articles, 500 from the US and 500 from Europe. This was an arbitrary decision based on where the main concentrations of research-active universities were. Thus, this sub-study was limited in not being to assess important research countries such as Canada or Australia. Future work should include other countries such as Canada, Australia, and China, as their impact is increasingly recognised.”

4. “The countries represented in this paper are predominantly Western. With the exception of one comment on language (p.11, line 264) the authors do not discuss this in any great detail. I found it particularly surprising that China was not represented in any way if for no other reason than sheer population size. In order to be less Western-centric, discussion of this omission and any potential others is needed”. 
The reviewer is quite right – it is surprising that China is not represented despite its population size and recent research productivity. In our methods, we looked at all countries but decided to limit our results to the top 10 in research productivity. China came 21st with regards to productivity (468 records, TGCS 4254), following Spain, France, Sweden, Austria, Denmark, Brazil, Israel, Turkey, Finland, and Norway. This is likely to change over time as China’s research infrastructure develops, and we have noted this in the limitations section (p16, lines 330-334).

5. “A number of other metrics exist for evaluating research merit-including the h-index (as in Carleton, Parkenson, & Horswill, 2012), which deals with publication quality and accounts for such issues as self-citation. I recognize the authors considered institutions in general rather than particular individuals at these institutions; however, I saw no discussion of the flaws in the metrics they did use or consideration of additional metrics that could be used”.

This is a very helpful comment. As the reviewer noted, we did not use metrics such as the h-index, as they are focused on individuals rather than our focus which was countries and institutions. However, we agree that some discussion of the limitations of the metrics we used is relevant and have revised the manuscript accordingly (p6, lines 136-152):

“All from the GCS, we also calculated the average number of citations per paper. Many measures have been used to quantify an individual researcher’s citation impact, some of
which provide also information about the quality of the publications, such as the h-index and the g-index [29-31]. We chose the GCS over other metrics as we focused on the scientific productivity and impact at country and university level rather than at researcher level. A major advantage of the GCS scores is that they are readily available from the Web of Science, as is the h-index. Citations are, of course, only one measure of scientific impact and do not distinguish between the quality of the publication, the impact of the journal citing and whether the citation is positive or negative. Citations have some of the limitations of the other measures of academic output. As with other measures, citations can be magnified or deflated according to researchers productivity concerning articles included in WoS database, (31), their scientific collaborations, and can be influenced by individuals in the same institution or group of institutions citing each other. Citations can potentially be exposed to errors due to articles being cited in different ways, variations in author affiliations or multiple author affiliations, and variations in institutional naming or indeed omission of country/institute of origin. (32, 33)”

Minor Essential Revisions

1. “The manuscript should be thoroughly reviewed for grammar and punctuation. For example, the commas need to be revised on p. 2, lines 50-53. There are several instances of these types of errors throughout.”

The manuscript was thoroughly reviewed for grammar and punctuation.

2. “Rationale should be provided for why the dates were restricted to 2004-2009 (p. 5, line 118).”
The manuscript was amended accordingly (p5, lines 98-100):

“A time period of 5 years was considered to be long enough to provide a large number of articles. We collected our data in 2011; hence we allowed publications at least 2 years to receive citations.”

3. “The authors state on p. 5, lines 125-126 that Web of Science was chosen due to its capability for “citation analysis,” but then state on lines 130-131 that citation analysis was not yet available.”

We amended our statement accordingly (p. 5, lines 106-122):

“We chose to use Web of Science (within the Web of Knowledge), because of its potential to systematically provide data for informetric analysis, including publication and citation analysis, as well as the fact that it comprehensively includes databases, such as the Science Citation Index, the Social Sciences Citation Index, the Journal Citation Reports, and the Essential Sciences Indicators [4].

At the time of our collection of data and analysis, the Web of Science did not provide the facility for analysis of citation practices in large datasets; hence we used HistCite, a complementary tool. We extracted all relevant articles with their references, citations and authors’ affiliations from Web of Science and imported the data into a software tool for infometric analysis of citation linkage, HistCite. HistCite provides citation linkage between scientific papers with tables and graphs that assist visualising the flow of publications and citations within a scientific field [4, 28]. Using HistCite, we identified numbers of scientific publications for countries and institutions, characteristics of papers that attract more or fewer citations, and the flow of citations between different scientific publications.”
4. “Page 6, lines 160-161: one of the stated goals of the manuscript was to generate the top ten institutes in psychiatry in terms of publication productivity. However, in practice, the authors generated the top ten American and top ten European institutions (excluding Canadian, Australian, etc. Institutions) and did not finalize this list. Either the stated goal or the displayed results need modifying.”

This was an oversight on our part. The table of the top 10 institutes in psychiatry with regards to productivity is now included in the manuscript that includes Canadian, Australian and other countries.

“Table 2: Rankings of top 10 research institutions ordered by total number of records produced (ranking in brackets).

<table>
<thead>
<tr>
<th>#</th>
<th>Institution</th>
<th>TGCS</th>
<th>Records</th>
<th>Citations/paper</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Harvard Univ</td>
<td>38910 (1)</td>
<td>1948 (1)</td>
<td>20.0 (1)</td>
</tr>
<tr>
<td>2</td>
<td>Inst Psychiatry</td>
<td>16730 (2)</td>
<td>1346 (2)</td>
<td>12.4 (8)</td>
</tr>
<tr>
<td>3</td>
<td>Univ Pittsburg</td>
<td>15420 (3)</td>
<td>970 (3)</td>
<td>15.9 (6)</td>
</tr>
<tr>
<td>4</td>
<td>Yale Univ</td>
<td>14633 (4)</td>
<td>894 (4)</td>
<td>16.4 (4)</td>
</tr>
<tr>
<td>5</td>
<td>Columbia Univ</td>
<td>14381 (6)</td>
<td>882 (5)</td>
<td>16.3 (5)</td>
</tr>
<tr>
<td>6</td>
<td>NIMH</td>
<td>14451 (5)</td>
<td>810 (6)</td>
<td>17.8 (2)</td>
</tr>
<tr>
<td>7</td>
<td>Toronto Univ</td>
<td>8408 (9)</td>
<td>743 (7)</td>
<td>11.3 (9)</td>
</tr>
<tr>
<td>8</td>
<td>Univ Calif Los Angeles</td>
<td>12634 (7)</td>
<td>741 (8)</td>
<td>17.1 (3)</td>
</tr>
<tr>
<td>9</td>
<td>Univ Calif San Diego</td>
<td>9942 (8)</td>
<td>716 (9)</td>
<td>13.9 (7)</td>
</tr>
<tr>
<td>10</td>
<td>Univ Munich</td>
<td>6010 (10)</td>
<td>651 (10)</td>
<td>9.2 (10)</td>
</tr>
</tbody>
</table>
5. “Page 10, lines 225-227: did the authors intend to generate the number of collaborations in the 1,000 sampled papers, rather than the number of collaborations cited by these papers? The former would be useful information as well. Do those who collaborate internationally tend to cite other international collaborations?”

We intended to investigate the proportion of the citing papers that were collaborations between two or more countries. We did not collect data so that we could answer the last question from the reviewer: Do those who collaborate internationally tend to cite other international collaborations? Nevertheless, this is a very interesting research question. Future research could focus on exploring whether researchers that collaborate internationally cite other international collaborations. We now state this in our revision (p15, lines 309-311).

6. “The first paragraph in the Discussion (lines 240-247) introduce several new citations that should have been present in the introduction.”

Thank you for this useful comment, and we agree. We have added these citations to the intro (p 4, line 75-80)

“Within psychiatry, there has been research on local, national and international scientific productivity focusing on specific subspecialties and the most popular research areas within them [3, 4, 5-11], on individual institutions [12-14], journals (15, 16), the contribution of different countries [8, 17], practices [18] and early recognition of high quality researchers [19, 20]. However, there has been little research examining trends in citation practice [21].”
7. “Page 13, line 297: the APA 6th edition manual indicates that gender neutral terminology is required in scientific publications. As such, revision of “his scientific network” is required.”

This was amended accordingly (lines 329):

“We did not explore the citation practices between authors and their scientific network, [ ]”
Reviewer 2; Chris Piotrowski

“Although the authors acknowledge the limitation of their select group of 50 journals, I think that this actually begs another important question: What are the top cited ‘Second-order’ or ‘B’ level psychiatry journals?....and what countries tend to publish their ‘psychiatric’ findings in these less-regarded publications? This is important as many of these journals are collaborative publications with adjunctive fields like nursing, mental health, psychology, social work, and psychotherapy. This secondary literature plays an important function in a vast, ever emerging field like Psychiatry.”

“The methodology is quite sound and appropriate procedures were used. However, I would have preferred that ‘minor’ research like ‘letters’ or ‘comments’ be excluded from the analysis. Also, as noted in the Table, 4855 articles could not be categorized by country. It is speculative how this might affect the results.”

The reviewer is quite right, and we amended our limitations accordingly (p 16, lines 334-338). A large number (n=4855) of articles could not be categorised by country. This is an important limitation of using a web-based data base to collect our data. As described in point 5 of the response to the previous reviewer’s comments, “Citations can potentially be exposed to errors due to articles being cited in different ways, variations in author affiliations or multiple author affiliations, and variations in institutional naming or indeed omission of country/institute of origin [32,33].” It is not certain whether the articles that we could not categorise would be disproportionately from one country, and therefore the effect on the results is not known.
We also added to our limitations following the reviewer’s helpful comment (p 16, lines 322, 325): “We limited our study to the top 50 journals in psychiatry – which would miss the leading journals in important related fields such as nursing, social work and clinical psychology. Alternative approaches to selecting journals can be considered in future research.”

“Oh, the findings are based on the data analysis and seem quite appropriate. Although I would like more emphasis (Table 1) on the great discrepancy between the Total# of citations for the U.S. (202,781) versus other countries........ e.g. U.K. (62,486). “

Thank you for your helpful comment. As shown on table 1, US produce 3 times more publications than the UK and receive 3 times more citations. The average citation per article number is not very different from the UK (11.5 versus 9.7).

“Furthermore, Google Scholar is today a much acceptable ‘metric’ in citation analysis. Could the authors speculate how their methodology, if applied in Google Scholar (GS), might compare to their findings or influence their conclusions? Comparisons between Web of Science/ Google Scholar are frequent areas of study in bibliometric journals like Scientometrics and Informetrics.”

(Page 15, Lines 318- 321)

Thank you for this very helpful comment. We carefully avoided to compare the Web of Science with other databases used for citation analysis, such as the Scopus or the Google Scholar, as they are all using different sources to generate their data, hence are not directly
comparable. Similarly, research on which database is preferable remains controversial. [32, 34, 35]
Reviewer 3; Michel Thibodeau

Moderate comment (Minor Essential Revisions)

1. “Some of the methods used by the authors are described in the results section, and may be better suited for the methods section. I found myself at times thinking that the methods section was missing some information, and I later found this information in the results section.”

This was a very useful comment. The manuscript was amended accordingly:

- (p6, line 127-130)

“As some of the institutions consist of a number of subdivisions, hence different authors’ affiliations, we explored whether institutions subsumed into greater local units were more productive and if their research was cited more often than others.”

-(p6, line 136-172)

“From the GCS, we also calculated the average number of citations per paper. Many measures have been used to quantify an individual researcher’s citation impact, some of which provide also information about the quality of the publications, such as the h-index and the g-index [29-31]. We chose the GCS over other metrics as we focused on the scientific productivity and impact at country and university level rather than at researcher level. A major advantage of the GCS scores is that they are readily available from the Web of Science, as is the h-index. Citations are of course only one measure of scientific impact and do not distinguish between the quality of the publication, the impact of the journal citing and whether the citation is positive or negative. Citations present with some of the
limitations of the other measures of academic output. As with other measures, citations can be magnified or deflated according to researchers productivity concerning articles included in WoS database, (31), their scientific collaborations, and can be influenced by individuals in the same institution or group of institutions citing each other. Citations can potentially be exposed to errors due to articles being cited in different ways, variations in author affiliations or multiple author affiliations, and variations in institutional naming or indeed omission of country/institute of origin. (32, 33)"

As HistCite does not automatically provide information about the details of citation practices, such as the distribution of citations between countries and institutions, and the number of self-citations, we extracted this information manually in a stratified sample of 1000 records. In order to gain representation of our initial sample of 51,072 articles, we collected the first 500 articles from the USA (as presented on HistCite in chronological order), and equally the first 500 articles from Europe (200 records from UK, 200 from Germany and 100 from other European countries). We sampled the earliest articles included in our initial search, as they had a better chance to generate citations since being published. For the purposes of this sub-study, we used the Local Citation Score (LCS), which is the number of citations that a paper received within the field of psychiatry. Since research collaboration between countries is an increasing phenomenon, we also investigated the proportion of the publications in our sample that resulted from such collaboration. In a sensitivity analysis we combined Canada with the US (500 articles) to make North America and we investigated citation practices in North America and Europe.

We furthermore attempted to identify subject trends by examining the most common key words in the titles of publications. HistCite automatically generates results on the most common key words as they were identified by the article authors.”
Minor comments (Discretionary Revisions):

1. “There is an extra space in the last line of the first paragraph in methods section.”

This error was amended.

2. “The authors report doing a representative sampling of records. What did this sampling include? The authors may wish to demonstrate how their sampling is indeed representative. The authors could, for example, gather a second sample and demonstrate the similarity between both samples, although this is not mandatory. (I noticed upon further reading that the nature of the representativeness is described in the results. Can this be moved to the methods section? It would seem more appropriate).”

Thank you for this helpful comment. The manuscript was amended as follows (p7, line 154-168):

“As HistCite does not automatically provide information about the details of citation practices, such as the distribution of citations between countries and institutions, and the number of self-citations, we extracted this information manually in a stratified sample of 1000 records. In order to gain representation of our initial sample of 51,072 articles, we collected the first 500 articles from the USA (as presented on HistCite in chronological order), and equally the first 500 articles from Europe (200 records from UK, 200 from Germany and 100 from other European countries). We sampled the earliest articles included in our initial search, as they had a better chance to generate citations since being published. For the purposes of this sub-study, we used the Local Citation Score (LCS), which is the number of citations that a paper received within the field of psychiatry. Since research collaboration between countries is an increasing
phenomenon, we also investigated the proportion of the publications in our sample that resulted from such collaboration. In a sensitivity analysis we combined Canada with the US (500 articles) to make North America and we investigated citation practices in North America and Europe.

3. “The authors write “who cites whom in psychiatry?” The article may benefit from this question being broken down into more specific questions. For example, are the authors referring to self-citation? Or perhaps which authors from what countries cite authors from which other countries?”

We agree with the reviewer and we clarified this point (p8, lines 177-178):

“Who cites whom in psychiatry (self citations, within and between countries citation trends)?”

4. “In table 4, the meaning of the total citations column is not immediately apparent. What are these numbers (and percentages) referring to?”

This was amended in the table (p23, line 507). “The 1000 articles considered received 3504 citations.”

5. “The authors may wish to further elaborate the “local” units described in the results section. Were these based solely on geography? The authors could elaborate on this in the methods section rather than in results section. The authors also write that “the rankings were broadly confirmed”, but it is unclear what rankings they are referring to. Presumably this refers to the ranking using the “non-localized” units, but this is not obvious.”
Thank you for this comment. We addressed it as follows (p.6, line 127-130):

“As some of the institutions consist of a number of subdivisions, hence different authors’ affiliations, we explored whether institutions subsumed into greater local units were more productive and if their research was cited more often than others.”

And (p. 10, line 208-209)

“The rankings of the institutions were broadly confirmed when the various subdivisions were combined.”

6. “The authors apparently used the 500 oldest papers as the sample for citation flow. It seems possible that publication trends have changed in the last 10 years. The authors could mention this as a limitation, if appropriate.”

This was a very helpful comment. We amended the manuscript accordingly (see page 16, line 343-345):

“We also used the oldest articles for the sub-study. It is possible that the trends are changing or will change in the future, especially with regards to the recent appearance of eastern countries in the citation map.

7. “The authors mentioned using Pearson correlations in the methods section, but I did not notice any correlations in the results section. Please correct me if I am mistaken, or remove mention of the Pearson correlations.”

This was amended; Pearson correlations mention was removed.

8. “The authors mention “When counting separately for UK publications (n=693), we found that British publications more often cited other Europeans (30%), but cited American (22%)”. It is unclear whether these differences are statistically significant,”
or only informal observations. If possible, perhaps the authors could conduct chi square tests (or similar) as they did for other proportions.”

Thank you for the clarification. The statistics mentioned are only descriptive statistics.

9. “research collaborations and found that 22% (n=755) of were” seems grammatically incorrect.

This typo was corrected (P11, lines 235-237):

“In addition, we investigated how many records of the 3,504 publications were international research collaborations and found that 22% (n=755) were collaborations between two or more countries.”

10. “The authors mention examining subject trends in the results, but this topic did not get any attention in the methods section. The authors present a few topics (schizophrenia, treatment, etc.) but not others. How were these topics chosen? It seems unclear why these topics were chosen, how they were chosen, and why others were not included.”

Thank you for this helpful comment. We added the explanation on the methods section (p 8, Line 170-172)

“We furthermore attempted to identify subject trends by examining the most common key words in the titles of publications. HistCite automatically generates results on the most common key words as they were identified by the article authors.”

11. “The authors write in discussion: “One notable difference was that between Germany and the Netherlands, where citation rates differ two-fold”. Can the authors
Citation rates were added in parenthesis (p. 13, line 272):

“One notable difference was that between Germany and the Netherlands, where citation rates differ two-fold (citations per publication 5.0 and 11.4 respectively).”

12. “Line 297 (second to last paragraph) – “his” scientific work should likely be “his or her”.”

This line was amended accordingly (line 329): “We did not explore the citation practices between authors and their scientific network, [ ]”

13. “Some interesting information is presented in the appendices. The authors may wish to include these as regular tables if the journal allows it.”

We have followed the journal’s recommendations with regards to the appendices.

Thank you for these very helpful comments, and we look forward to hearing from you.