

The Geographical Gap in Leading Medical Journals - a Computational Audit

Oscar Brück¹

1 Hematoscope Lab, Helsinki University Hospital, Helsinki, Finland

ADDRESS FOR CORRESPONDENCE

Dr. O. Brück, Hematoscope Lab, Helsinki University Hospital, Comprehensive Cancer Center, Haartmaninkatu 8, P.O. Box 700, Helsinki FIN-00290, Finland.

LENGTH 2671 words, 3 figures and 1 table

ABSTRACT

Background:

Auditing geographical representation in medical publishing could help to mitigate possible national and regional disparities.

Methods:

Using the Web of Science indexing database, we collected bibliometric data of original research articles published between 2010-2019 in *The New England Journal of Medicine*, *Nature Medicine*, *Journal of the American Medical Association*, *The BMJ*, and *The Lancet*. We studied the corresponding authors' geolocation in regard to publication and citation count, their temporal evolution, and the journals' and citing organizations' nationality.

Results:

We identified 10,558 articles. Based on the nationality of the corresponding authors' institutes, only 32 countries published more than 10 publications in 10 years equaling to 98.9% of all publications. English-speaking countries USA (48.2%), UK (15.9%), Canada (5.3%), and Australia (3.2%) were most represented, but with a declining trend in recent years. Normalized to their accumulated citations, 9/32 countries were associated with $\geq 10\%$ publication excess, of which USA (n=1,174 publications) and UK (n=410) accounted for 85.7%. Similar findings were replicated at the municipal level where all top 10 most productive cities were located in USA (n=7), UK (n=2), or Canada (n=1), and 21 out of 25 most productive cities published more articles than predicted based on their accumulated citations. Finally, we discovered that both journals published, and researchers cited more commonly research conducted in the same country.

Discussion:

NOTE: This preprint reports new research that has not been certified by peer review and should not be used to guide clinical practice.

52 The audit revealed Anglocentric dominance, domestic preference occurring in both
53 journals and citation selection, and increased geographical representation in recent
54 years in medical publishing.

55 **BACKGROUND**

56 Defined as the yearly-averaged citation frequency of research articles, journal impact
57 factors (JIF) were initially designed to rank journals by their influence¹. JIFs have also
58 been used for non-intended purpose. At the individual level, a successful publication
59 track record has been interpreted as an ability to conduct high-quality research
60 permitting both career advancements and increased chances of research funding. At
61 the institutional level, harboring scientists with publications in journals with high JIF
62 creates a positive feedback loop, which further attracts human and financial resources
63 creating polarization among research centers.

64
65 Monitoring diversity and equity in academic publishing are fundamental to mitigate
66 possible disparities. By signing the “Joint commitment for action on inclusion and
67 diversity in publishing” more than 50 publishers representing 15,000 scientific journals
68 will participate by collecting data on gender and ethnical representation and identify
69 bias in the editorial and review processes². However, decades of publication metadata
70 stored in journal indexing databases such as the Web of Science (WoS), Scopus,
71 PubMed, and Google Scholar could reveal the extent and temporal development of
72 inequity in medical publishing.

73
74 In this retrospective cohort study, we examined possible bias related to the
75 geographical location of the conducted research to the number of publications and
76 citations. We also interrogated publishing and citing patterns based on the
77 geographical origin of the conducted research.

78
79

80 **METHODS**

81 *Data collection*

82 We selected the five journals publishing mainly original articles and ranked highest in
83 the Journal Citation Reports 2022 JIF in the field of medicine. These included *New*
84 *England Journal of Medicine* (NEJM), *Nature Medicine* (NatMed), *Journal of the*
85 *American Medical Association* (JAMA), *The BMJ*, and *Lancet*. We included all original
86 articles (n=10,558) published in 2010-2019. More recent publications were not
87 included due to unequal opportunities of being cited and to avoid bias related to
88 COVID-19 pandemic. Using the WoS database provided by Clarivate Plc and the query
89 “(((SO=(NATURE MEDICINE OR LANCET OR NEW ENGLAND JOURNAL OF
90 MEDICINE OR JAMA JOURNAL OF THE AMERICAN MEDICAL ASSOCIATION OR
91 BMJ BRITISH MEDICAL JOURNAL)) AND DT=(Article)) AND PY=(2010-2019))”, we
92 downloaded publication metadata including author names, page length of the article,
93 address of the corresponding author, and total number of citations. We could also
94 query international citing patterns based on the digital object identifiers of the 10,558
95 original articles. National population data were downloaded from the World Bank
96 database <https://data.worldbank.org/indicator/SP.POP.TOTL>, and municipal data from
97 the world.cities dataset of the utils R package.

98
99

99 *Preprocessing variables of interest*

100 We defined the impact of an article by its average citation count per year. We text-
101 mined the number of authors, by calculating the frequency of the semicolon “;” delimiter
102 between author names and added 1. To geolocate the primary institutes where the
103 research has been performed, we identified the latitude and longitude coordinates of
104 the address(es) of the corresponding author(s) with the ggmap library employing

105 Google's Geocoding API. For non-successful matches, we geolocated only the city
106 and country of the address.

107

108 *Indices*

109 We defined the international research impact (IRI) as the relationship between the total
110 number of publications (10-fold logarithm transformed) and their cumulative citations
111 (10-fold logarithm transformed) grouped by the research institute of the corresponding
112 authors. In this study, IRI was used separately when comparing countries and cities.

113

114 We defined the Domestic Self-Citation Index (DSCI) as:

$$115 \frac{\frac{\# \text{ Citations of country } \alpha \text{ from country } \alpha}{\# \text{ Citations in total from country } \alpha}}{\frac{\# \text{ Citations of country } \alpha \text{ from all countries except country } \alpha}{\# \text{ Citations in total of all other except country } \alpha \text{ from country } \alpha}}$$

116 where "# Citations" represents the number of citations and α any country of interest.

117

118 *Statistical analysis*

119 We performed comparisons of two continuous variables with the Wilcoxon rank-sum
120 test (unpaired, two-tailed) and three or more continuous variables with the Kruskal-
121 Wallis test. For categorical variables, we used the χ^2 test. We adjusted p values with
122 the Benjamini–Hochberg correction. We performed hierarchical clustering with
123 Euclidean distance metrics and the Ward.d2 method. To compare two linear
124 regression slopes, we tested the T-test significance of the interaction term. We
125 conducted statistical analyses and visualizations with R 3.5.1. using base, tidyverse,
126 fastDummies, maps, reshape2, ggmap, data.table, countrycode, ggpubr, ggrepel,
127 rstatix, ggdendro and dendextend libraries.

128

129

130 **RESULTS**

131 *National publication productivity does not imply high citation frequency*

132 Of 10,558 original articles published between 2010-2019 in *NEJM* (n=2,966; 28.1%),
133 *JAMA* (n=1,833; 17.4%), *NatMed* (n=1,577; 14.9%), *Lancet* (n=2,462; 23.3%), and
134 *BMJ* (n=1,720; 16.3%), corresponding authors were affiliated to institutes from 77
135 countries. Each article included 1.2 mean affiliations [range 1-15] equaling to 10,732
136 total unique entries, of which 100.0% (n=10,730) could be geolocated. When focusing
137 on countries with ≥ 10 original articles in 10 years in any of these five medical journals,
138 only 32 countries were included representing 98.9% (n=10,613) affiliations highlighting
139 the geographical exclusivity of medical research.

140

141 Of these, English-speaking countries were overrepresented with almost 3/4 of
142 publications from the United States of America (USA, 48.2%), the United Kingdom (UK,
143 15.9%), Canada (5.3%), and Australia (3.2%, Table 1). National population-normalized
144 publication frequencies correlated with total publication number (corr 0.69 p<0.001,
145 Spearman test). Following population-normalization, the most represented countries
146 included Denmark (31.1 articles per million people), UK (25.1), and Switzerland (24.8,
147 Table 1).

148

149 When examining the international research impact (IRI) by countries, publication count
150 explained 96.5% of total citations corresponding to an excellent statistical
151 correspondence (Fig. 1A). However, we also observed unexpected discrepancies. We
152 identified 19/32 countries with over 10% publication excess or deficit when normalized

153 by their citation frequencies. Based on the relative number of articles, Spain published
154 78.7% (n=76) and Brazil 62.1% (n=26) fewer and Bangladesh 54.3% (n=6) too many
155 articles than expected. Based on the absolute number of articles, 1,174 (23.0%) and
156 410 (24.3%) excess articles were published by corresponding authors affiliated to an
157 organization located in the USA and UK, respectively. Together, USA and UK
158 accounted for 85.7% of all articles published in excess (n=1,847).

159
160 Publications with a corresponding author from Saudi Arabia were associated with
161 highest median yearly citations (49.4) per publication given their emphasis on the
162 documentation of the Middle East Respiratory Syndrome (MERS, 7 out of 11 total
163 publications from Saudi Arabia). However, even without MERS-related articles, the
164 median yearly citation of publications from Saudi Arabia was the highest 41.8. Other
165 countries with most elevated median yearly citation per publication included Belgium
166 (40.2), and China (39.6). Lowest median citations were associated with publications
167 geolocated in South Asia (Pakistan 11.6, Bangladesh 14.6, India 21.3) and
168 Scandinavia (Norway 19.7, Denmark 21.2, Sweden 21.3).

169
170 When extrapolating to continental level, publication exclusivity was even more
171 pronounced, with 89.5% of all articles originating either from the Americas (53.4%) or
172 Europe (36.1%, Fig. 1B). However, the 5.2% publications originating from Asian
173 countries were associated with 15.6% higher citation frequencies compared to median
174 (27.5) and other continents (Fig. 1B-C).

175
176 We then studied how the number of publications per country has evolved during the
177 study time period. By measuring the slope of a fitted linear regression, yearly
178 publishing productivity in the studied journals decreased steadily in the USA (8.0 fewer
179 articles per year), UK (5.7), Canada (1.7), and Italy (1.2, Fig. 1D). However, the number
180 of publications originating from China (2.5) and Israel (0.89) rose, but with a lower
181 inclination. The IRI patterns varied also temporally (Supplementary Fig. 1). Countries
182 with numerous publications (USA, UK, Germany, Australia) gradually stepped up their
183 yearly citation frequencies in line with rising medical journal impact factors. The
184 average citations decreased for some countries for limited periods (France,
185 Netherlands, China) before resuming their original increasing inclination.
186 Heterogenous citation patterns were found in countries with 1-10 yearly publications in
187 average.

188
189

190 *Publications from elite research institutes tend to accumulate fewer citations per article*
191 The international publication selection bias was even more pronounced when studying
192 the regional location of the corresponding authors' institutes (Fig. 2A). The most
193 distinct publication hot spots were situated in Northeastern USA, Central Europe, and
194 the UK with additional dense centers in Southwestern USA and East Asia. On the
195 contrary, the number of citations per article did not replicate similar geographical
196 patterns (Fig. 2A). The 10 most represented cities in terms of publication number
197 covered 36.5% of all publications and were all located in the Anglosphere: USA (n=6),
198 UK (n=3) and Canada (n=1). The three cities with over 300 publications in 2010-2019
199 were Boston, USA (n=1,209, 12.2% of all publications), London, UK (n=648, 6.5%)
200 and New York, USA (n=378, 3.8%).

201
202 Next, we examined municipal IRIs by including only cities with ≥ 10 publications in 10
203 years retaining 94.0% of the number of publications (n=9,327). Publication number

204 explained 83.6% of the variability in total accumulated citations (Fig. 2B). Multiple
205 metropolitan cities known for their established research institutes were
206 overrepresented in terms of publication frequencies. Unexpectedly, 21 out of the 25
207 most productive cities accumulated less citations than predicted with linear regression
208 suggesting selection bias. Cities with the highest absolute publication excess included
209 Boston, USA (n=728 excess articles), London, UK (n=285) and Oxford, UK (n=76). On
210 the contrary, cities with the highest publication deficit included Seattle, USA (n=76),
211 Houston, USA (n=27) and Los Angeles, USA (n=6).

212
213 While the publication count decreased in the USA and UK during the 10-year follow-
214 up, similar patterns were not as evident when examining the 25 most productive
215 institutes (Supplementary Fig. 3A). The publishing dynamics varied with some cities
216 publishing gradually more (New York, Stanford, Philadelphia), some less (London,
217 Baltimore, Ann Arbor), but for most the direction was inconclusive. As in the national
218 level, the average yearly citations-per-article increased also at the institute level and
219 the slope tendency varied more in cities with fewer publications (Supplementary Fig.
220 3B). Houston stood out with higher citations/publication across follow-up, whereas
221 publications originating from Boston, Toronto, and Copenhagen accumulated fewer
222 citations compared to other institutes.

223
224 The absolute number of publications correlated moderately with population-normalized
225 publication frequencies (corr 0.41, $p < 0.001$). However, the residuals of the linear
226 regression model between publication number and citations increased (corr 0.37,
227 $p < 0.001$) by publication number signifying that articles originating from established
228 research organizations tended to accumulate fewer citations than statistically would be
229 expected. Finally, institutes from capital cities did not accumulate more citations per
230 article compared to non-capital cities indicating that the governmental status or
231 population of the city are poor predictors of IRI (Fig. 2C and Supplementary Fig. 2).

232 233 234 *Institute nationality affects both journalistic and citation patterns*

235 Given overlap in the scientific scope of the studied five medical journals, we reasoned
236 that the geographical coverage of publications should not differ. However, we observed
237 substantial preponderance of publications from English-speaking countries (UK, New
238 Zealand and Australia) in both UK-based *BMJ* and *Lancet* journals (Fig. 3A). Moreover,
239 publications from Scandinavian countries (Norway, Sweden and Denmark),
240 Netherlands and Canada were overrepresented in *BMJ*. Publications from Germany,
241 Israel, and Italy were most common in *NatMed*, which is part of the British-German
242 Springer publishing group. Publications from institutes located in the USA were most
243 represented in *NatMed* and US-based journals *JAMA* and *NEJM*.

244
245 To conclude, we audited national citation patterns. First, countries with most
246 publications were the most active to cite as expected (Fig. 3B, horizontal lines of
247 recurring balloons). USA was the most common citing country in 30/32 of the most
248 frequently publishing countries. Articles from China and Singapore were cited more
249 commonly in publications with a corresponding author based in China.

250
251 Secondly, countries tended to show domestic preference when citing (Fig. 3B, diagonal
252 balloons). We observed highest absolute domestic self-citation (#domestic
253 citations/#all citations) among researchers affiliated to an institute in USA, UK, or South
254 Africa (Supplementary Fig. 4). To highlight the proportion of self-citations compared to

255 the proportion of citations originating from abroad, we introduce the domestic self-
256 citation index (DSCI, see Methods). The highest DSCI associated with China, USA,
257 and Japan, and the lowest with Austria, New Zealand, Kenya, and Greece (Fig. 3C).
258 National DSCI correlated with their total accumulated citation count (corr 0.70,
259 $p < 0.001$) and publication count (corr 0.46, $p = 0.015$), but not with the absolute count
260 (corr 0.032, $p = \text{ns}$) nor proportion of excess articles (corr -0.060, $p = \text{ns}$). Domestic self-
261 citation accounted for 74.5% of total national citation count (corr 0.86, $p < 0.001$)
262 indicating the serious bias related to citation count for evaluating manuscript impact.

263
264 Lastly, we observed that national citation patterns formed distinct clusters (Fig. 3B,
265 horizontal colored groups). The largest cluster was composed of the Anglosphere,
266 European countries, South Korea, and Saudi Arabia. The second largest cluster
267 included developing countries whereas China, Japan, and Singapore formed the third
268 cluster. In summary, similar geographical location, development index and health
269 challenges account for citation trends at the national level.

271 272 **DISCUSSION**

273 Available bibliometric data can reveal important information on the equity and
274 representation in medical research. Longitudinal data permits the study of temporal
275 trends and can be repurposed for monitoring of diversity. Here, we presented
276 geographical publication disparity in five leading medical journals based on JIFs
277 between 2010-2019.

278
279 We limited the scope of this study to available geographical data recorded in scientific
280 indexing databases. Therefore, for example ethnical, career-stage, and gender
281 disparity was not examined. Similarly, the number of submissions and geographical
282 coverage of reviewers are unavailable, while these could help to further interpret the
283 results.

284
285 First, we described the overrepresentation of distinct countries and research
286 organizations. Almost 3/4 of all publications with a corresponding author originated
287 from an institute in USA, UK, Canada, or Australia indicating medical publishing is
288 Anglocentric. Similar findings have been reported both in leading medical journals
289 between 1971-2005 and in the field of pathology between 2000-2006 suggesting both
290 an interdisciplinary and long-lasting phenomenon^{3,4}. We showed that the number of
291 articles from institutes in USA, UK and Canada was decreasing, while publications
292 from China and Israel were more frequent reflecting a gradual transition in international
293 representation.

294
295 Second, by comparing IRIs our analysis highlighted an undocumented discrepancy
296 favoring notably highly productive institutes and countries. For instance, 86% of all
297 articles published in excess based on their IRI originated from USA or UK. At the
298 municipal level, 21/25 of the leading centers were found to publish more articles than
299 expected. In particular, the number of publications with a corresponding author located
300 in Boston exceeded 60% its predicted quantity.

301
302 Multiple factors could be involved. For instance, previous studies have indicated a
303 pronounced correlation between national gross domestic product and publishing in top
304 medical journals^{5,6}. The publication fees require considerable research funding, which
305 are primarily available in high-income countries making scientific publishing an

306 unintended luxury. Moreover, our data indicated higher proportions of articles from UK,
307 Australia, Canada and New Zealand in UK-based journals (*BMJ* and *Lancet*), while
308 corresponding authors from USA were overrepresented in American journals (*JAMA*
309 and *NEJM*). UK-affiliated research published in UK-based journals (*BMJ*, 28.1%;
310 *Lancet*, 34.8%; mean 30.8%) was almost 5-fold compared to the average proportion
311 in *NEJM*, *NatMed* and *JAMA* (6.3%). Importantly, journals such as *Lancet* advice
312 authors to consider diversity when inviting coauthors⁷. Yet, further comparative studies
313 on submission rates are required to better understand the reasons and to exclude
314 domestic preference at review or editorial level, which could undermine public trust in
315 scientific publishing and possibly found new journals from underrepresented countries.
316 Importantly, our data indicated that less-publishing countries and institutes tended to
317 accumulate higher citations, which could reflect both interest and ultimately elevate
318 further JIFs. Diversity and inclusion of developing countries would benefit the scientific
319 community as demonstrated during the SARS, MERS, and COVID-19 epidemics,
320 facilitate the adoption of health policies globally, and diversify medical research by
321 increasing international collaborations⁸.

322
323 When examining international citation patterns, we identified considerable domestic
324 preference. The nationality of the research institute accounted for 74.5% of the
325 international citation variance. In addition, the DSCI varied 8-fold between the studied
326 countries emphasizing that further research in citation patterns is needed. The results
327 are in line with a study reporting that 85% of the variability of the scientific journalism
328 coverage of 22 newspapers focused on domestic research⁹. However, we also
329 observed international citing clusters reflecting likely similar research interests and
330 collaborations.

331
332 In summary, this computational audit of author affiliations in top medical journals
333 revealed the Anglocentric dominance embodied as publication excess in relation to
334 their accumulated citation and domestic preference by journals. Based on longitudinal
335 data, the international representation gap in top medical journals is gradually
336 decreasing.

337 **ADDITIONAL INFORMATION**

338

339 **ACKNOWLEDGMENTS**

340 The author wishes to thank Susanna Lallukka-Brück for her patience, understanding,
341 and insightful comments. The author is grateful for Olli Dufva and to the members of
342 the Hematoscope Lab for discussion and comments. Certain data included herein are
343 derived from Clarivate Web of Science. © Copyright Clarivate 2022. All rights reserved.

344

345 **AUTHORS' CONTRIBUTIONS**

346 Conception and design; Collection and assembly of data; Data analysis; Manuscript
347 writing; Manuscript editing; Data interpretation; Final approval of manuscript: O.B.

348

349 **ETHICS APPROVAL AND CONSENT TO PARTICIPATE**

350 The study complied with the Declaration of Helsinki.

351

352 **CODE AND DATA AVAILABILITY**

353 Codes and a 100-row data example are available at
354 <https://github.com/obruck/International-Research-Impact>. Raw data can be
355 downloaded from Clarivate Web of Science, with instructions provided in the Github
356 repository.

357

358 **COMPETING INTERESTS**

359 O.B. declares no Competing Non-Financial Interests but the following Competing
360 Financial Interests: consultancy fees from Novartis, Sanofi, and Amgen, outside the
361 submitted work.

362

363 **FUNDING INFORMATION**

364 This study was supported by research grants from the Helsinki University Hospital.

365

366 **REFERENCES**

- 367 1 Garfield E. The History and Meaning of the Journal Impact Factor. *JAMA* 2006;
368 **295**: 90–3.
- 369 2 Else H, Perkel JM. The giant plan to track diversity in research journals. *Nature*
370 2022; **602**: 566–70.
- 371 3 Fritzsche FR, Oelrich B, Dietel M, Jung K, Kristiansen G. European and US
372 publications in the 50 highest ranking pathology journals from 2000 to 2006. *J*
373 *Clin Pathol* 2008; **61**: 474 LP – 481.
- 374 4 Falagas ME, Alexiou VG. An analysis of trends in globalisation of origin of
375 research published in major general medical journals. *Int J Clin Pract* 2008; **62**:
376 71–5.
- 377 5 Halpenny D, Burke J, McNeill G, Snow A, Torreggiani WC. Geographic Origin
378 of Publications in Radiological Journals as a Function of GDP and Percentage
379 of GDP Spent on Research. *Acad Radiol* 2010; **17**: 768–71.
- 380 6 Man JP, Weinkauff JG, Tsang M, Sin JHDD. Why do Some Countries Publish
381 More Than Others? An International Comparison of Research Funding, English
382 Proficiency and Publication Output in Highly Ranked General Medical Journals.
383 *Eur J Epidemiol* 2004; **19**: 811–7.
- 384 7 The Lancet Information for Authors.
385 <https://www.thelancet.com/pb/assets/raw/Lancet/authors/tl-info-for-authors.pdf>.
386 2022.
- 387 8 Ghani M, Hurrell R, Verceles AC, McCurdy MT, Papali A. Geographic, Subject,
388 and Authorship Trends among LMIC-based Scientific Publications in High-
389 impact Global Health and General Medicine Journals: A 30-Month Bibliometric
390 Analysis. *J Epidemiol Glob Health* 2020; **11**: 92–7.
- 391 9 Casino G, Rius R, Cobo E. National citation patterns of NEJM, The Lancet,
392 JAMA and The BMJ in the lay press: a quantitative content analysis. *BMJ Open*
393 2017; **7**: e018705.
- 394
395

396 **TABLES**

397 Table 1. Publication metrics by country.

Country	# Citations	# Citations/ Year	# Articles	# Population	# Citations/ Article/ Population/ Year	# Articles/ Population	# Excess Articles	% Excess Articles
Australia	20443.2	28.9	339	25.7	1.1	13.2	-22.1	-6.5
Austria	1594.3	29.8	36	9.0	3.3	4.0	-1.9	-5.2
Bangladesh	162.4	14.6	11	166.3	0.1	0.1	6.0	54.3
Belgium	5111.6	40.2	82	11.6	3.5	7.1	-24.1	-29.4
Brazil	3094.9	37.1	42	214.0	0.2	0.2	-26.1	-62.1
Canada	30632.0	28.4	560	38.2	0.7	14.6	43.8	7.8
China	15919.5	39.6	225	1412.4	0.0	0.2	-64.5	-28.7
Denmark	6633.5	21.2	182	5.9	3.6	31.1	48.5	26.6
Finland	1679.8	26.1	40	5.5	4.7	7.2	0.3	0.8
France	24133.5	33.7	386	67.5	0.5	5.7	-32.1	-8.3
Germany	21318.0	36.0	376	83.1	0.4	4.5	1.3	0.3
Greece	733.3	31.9	14	10.7	3.0	1.3	-5.1	-36.2
India	1278.1	21.3	34	1393.4	0.0	0.0	2.8	8.4
Ireland	722.1	29.5	17	5.0	5.9	3.4	-1.8	-10.7
Israel	2047.4	22.0	47	9.4	2.3	5.0	-0.3	-0.5
Italy	6699.9	35.3	133	59.1	0.6	2.3	-1.7	-1.3
Japan	5982.8	26.9	112	125.7	0.2	0.9	-9.9	-8.8
Kenya	453.4	23.7	17	55.0	0.4	0.3	4.5	26.7
Netherlands	17730.4	25.9	374	17.5	1.5	21.3	55.6	14.9
New Zealand	2240.5	27.8	53	5.1	5.4	10.3	1.8	3.4
Norway	1935.5	19.6	62	5.4	3.6	11.5	17.0	27.5
Pakistan	377.4	11.6	13	225.2	0.1	0.1	2.4	18.5
Saudi Arabia	638.4	49.4	12	35.3	1.4	0.3	-4.9	-40.6
Singapore	1173.0	23.2	23	5.5	4.3	4.2	-5.9	-25.6
South Africa	2368.6	25.5	58	60.0	0.4	1.0	4.3	7.3
South Korea	1656.9	39.1	34	51.7	0.8	0.7	-5.2	-15.3
Spain	8909.0	39.6	97	47.3	0.8	2.0	-76.3	-78.7
Sweden	7303.3	21.2	199	10.4	2.0	19.1	53.6	26.9
Switzerland	10218.5	29.9	216	8.7	3.4	24.8	20.4	9.4
Thailand	946.8	26.7	17	70.0	0.4	0.2	-6.9	-40.6
UK	85627.5	24.4	1691	67.3	0.4	25.1	410.5	24.3
USA	305130.0	28.0	5111	331.9	0.1	15.4	1174.5	23.0

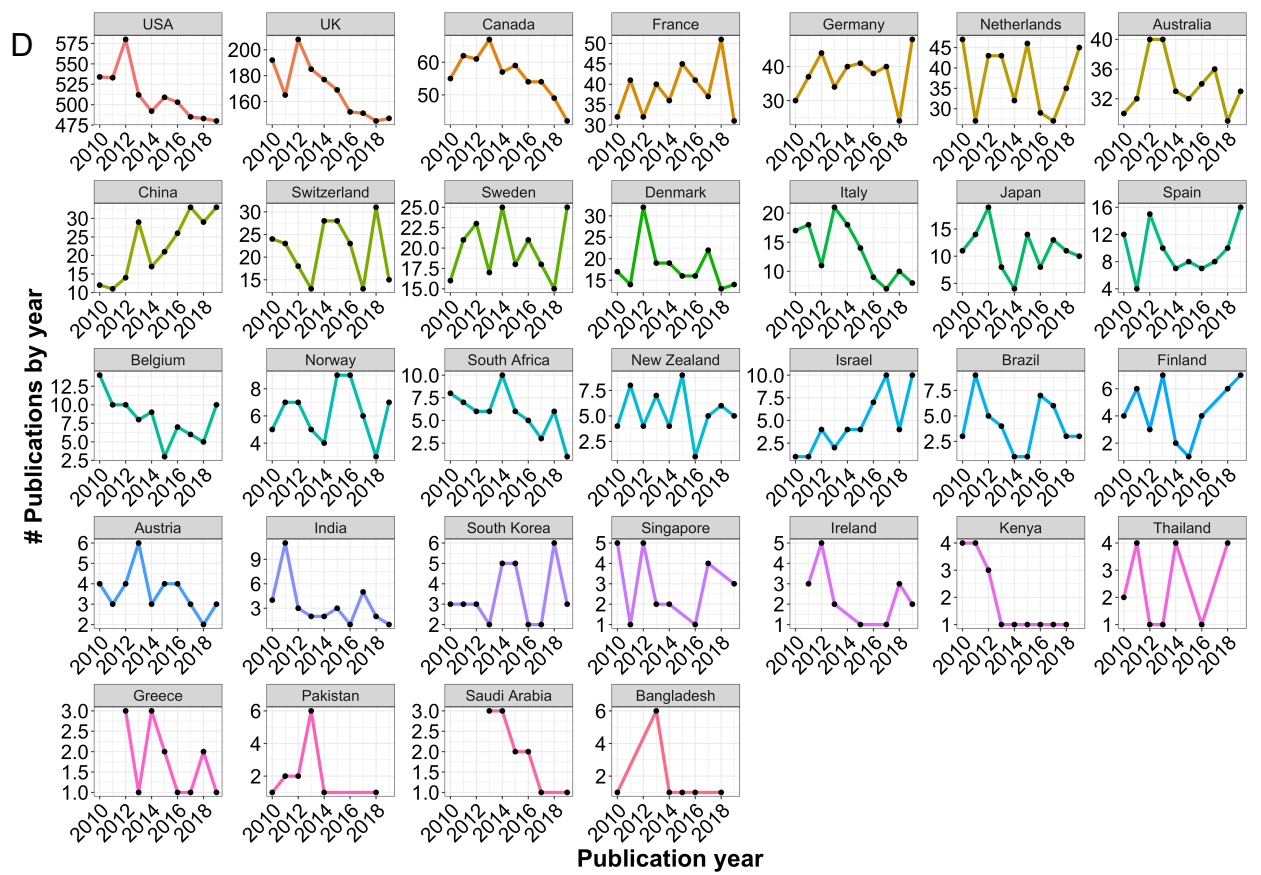
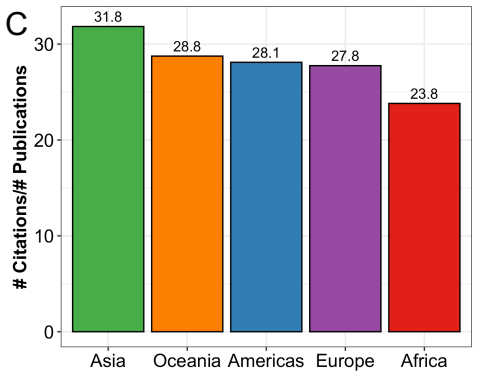
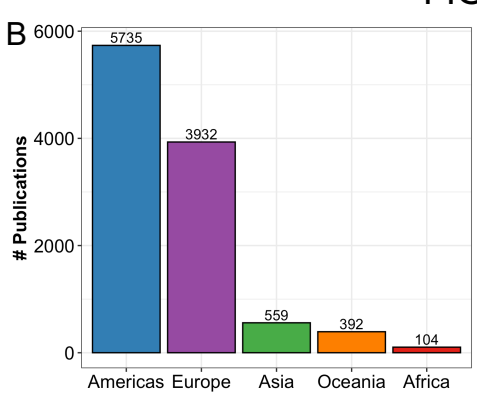
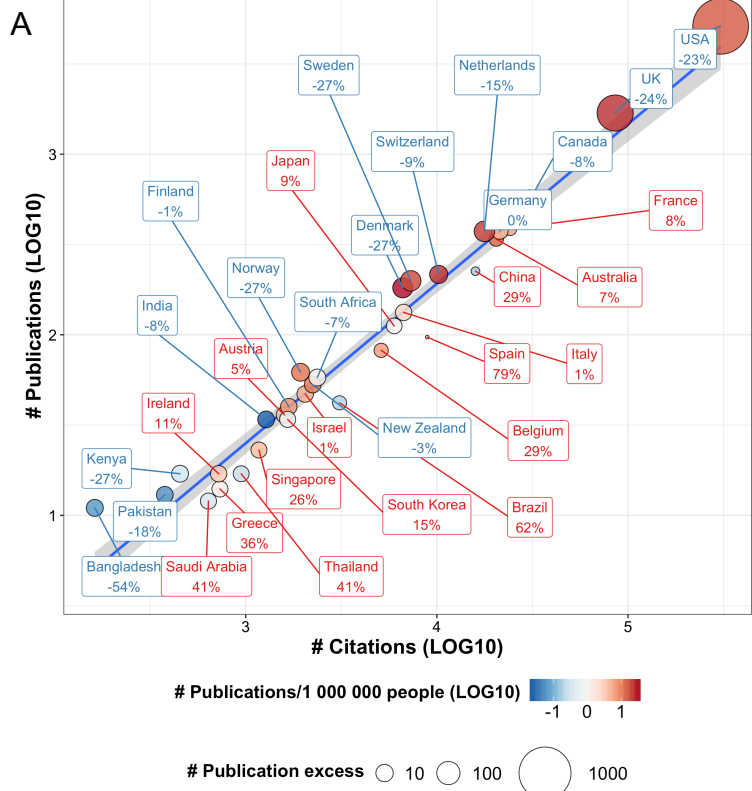
398

399 **FIGURE LEGENDS**

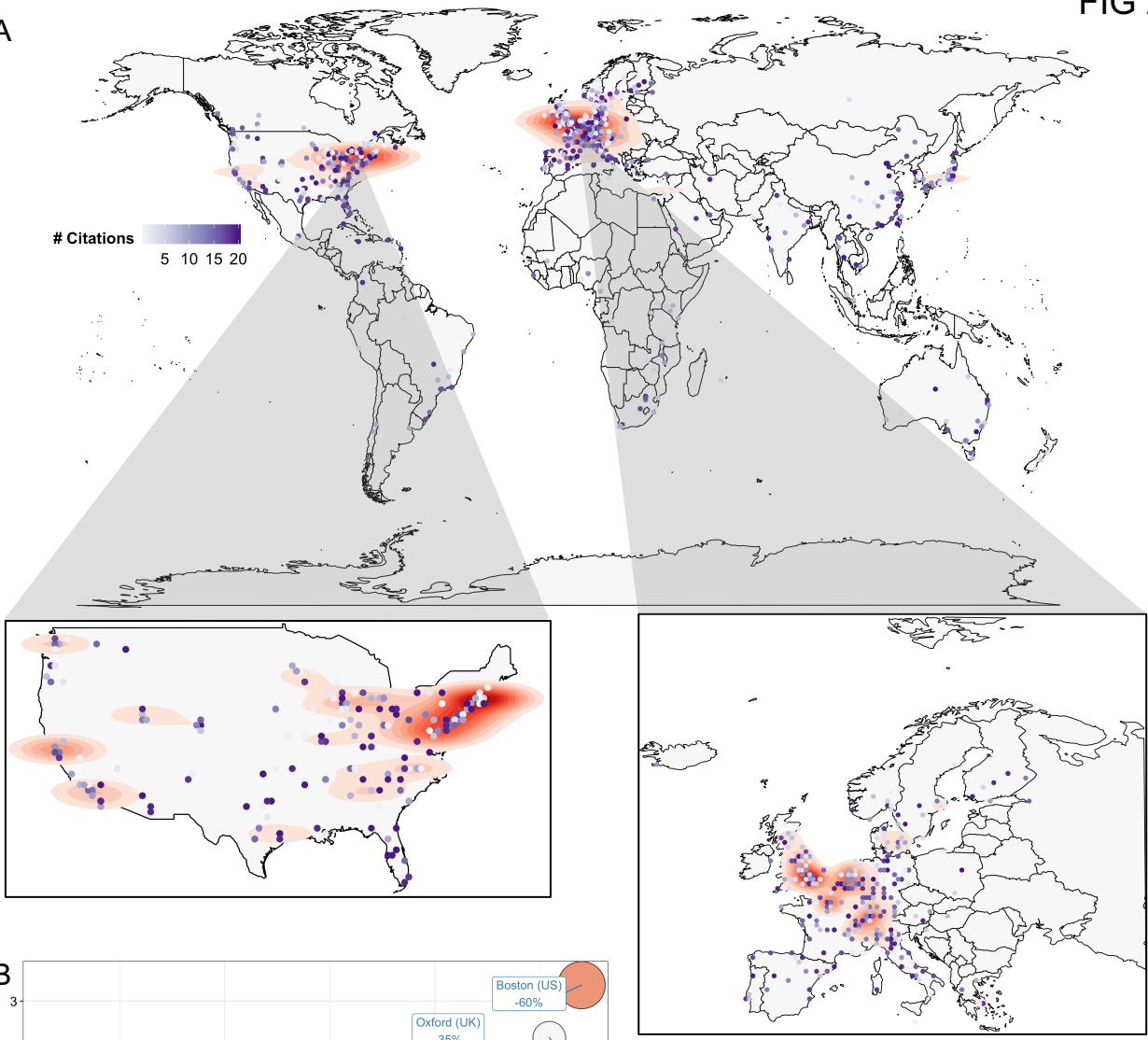
400 **Figure 1. Medical publishing at the national level.** (A) Linear regression of the
401 number of citations (LOG10-transformed) and publications (LOG10-transformed) of
402 the top 32 most productive countries. The percentage and font color indicate the
403 proportion of excess (positive number and red font) or deficient (negative number and
404 blue font) publications compared to the predicted number based on the
405 citation/publication count. (B) Bar plots illustrating the number of publications and (C)
406 citations per publication by continent. (D) Line plot illustrating the number of
407 publications by their publishing year from the top 32 most productive countries.

408
409 **Figure 2. Medical publishing at the municipal level.** (A) A world map with density
410 contours indicating hot spots based on the number of publications. Blue-colored points
411 reflect the average number of citations in that area. (B) Linear regression of the number
412 of citations (LOG10-transformed) and publications (LOG10-transformed) by cities. The
413 percentage and font color indicate the proportion of excess (positive number and
414 country name in red) or deficient number of publications compared to the predicted
415 number based on the citation/publication count. Country name abbreviations are
416 added for city names, which are found in more than one country. (C) Box plot of the
417 number of citations per publication in capital vs. non-capital cities (Paired Wilcoxon
418 test).

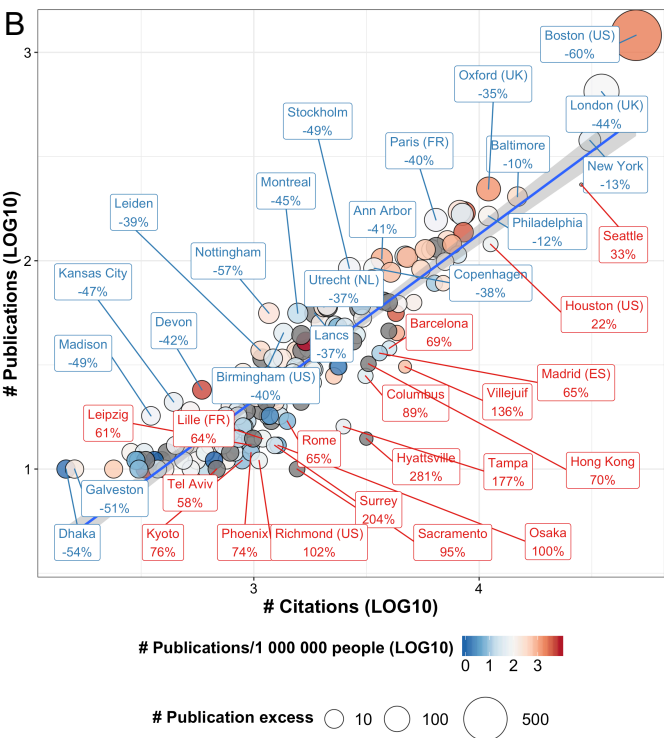
419
420 **Figure 3. National journal and citation fingerprints.** (A) Bar plot illustrating the
421 proportion of publications in medical journals originating from different countries. Only
422 significant results of the Kruskal-Wallis analysis are plotted. (B) Balloon plot illustrating
423 the international citation patterns of articles published in leading medical journals. The
424 balloon size reflects the proportion which countries have cited the most any given
425 country (column-wise). The balloon color reflects the absolute number of citations any
426 given country has received. Citing countries have been grouped by hierarchical
427 clustering (Euclidean distance). (C) Bar plot illustrating the same data as in the
428 previous plot but proportioned to how much other countries have cited any given
429 country in relation to how much they have cited all others countries. See Methods for
430 the definition of the Domestic Self-Citation Index.



A



B



C

