Transnational Journal of Pure and Applied Mathematics
Vol. 2, Issue 1, 2019, Pages 21-35
Published Online on January 29, 2019

http://jyotiacademicpress.org

WHETHER MULTIPLY COAUTHORS IN ARTICLES ON PURE AND APPLIED MATHEMATICS CONTRIBUTED MORE CITATIONS

TSAIR-WEI CHIEN[1], **WEI-CHIH KAN**[2,3], **HSIEN-YI WANG**[1] **and WILLY CHOU**[4,5*]

[1]Research Department
Chi-Mei Medical Center
Tainan
Taiwan
e-mail: smile@mail.chimei.org.tw
hsienyiwang@gmail.com

[2]Nephrology Department
Chi-Mei Medical Center
Tainan
Taiwan
e-mail: rockiekan@ntu.edu.tw

[3]Department of Biological Science and Technology
Chung Hwa University of Medical Technology
Tainan
Taiwan

[4]Department of Physical Medicine and Rehabilitation
Chi Mei Medical Center
Tainan
Taiwan
e-mail: ufan0101@ms22.hinet.net

2010 Mathematics Subject Classification: 97R60, 97K80, 97D10.
Keywords and phrases: citation, authorship collaboration, Google maps, social network analysis, Pubmed central.
Communicated by Francisco Bulnes.
Corresponding author*
Received January 13, 2019; Revised January 15, 2019

[5]Department of Recreation and Health-Care Management
& Institute Recreation Industry Management
Chia Nan University of Pharmacy
Tainan
Taiwan

Abstract

Team science in basic science has been accompanied by a trend in the numbers of authors included in scientific publications. Whether numerous coauthors can own more citations from other articles remains unknown. We download 148 abstracts on the topic of pure and applied mathematics from Pubmed Central (PMC) since 2016 and examine whether more number of coauthors in an article byline can earn more citations. Cluster analyses were performed using social network algorithms to classify authors. Bibliometric analyses were conducted to compute individual research achievements (IRA) for selecting the most cited authors. Visual representations were made to show results on Google Maps. We found that (1) 43 articles (29%) have coauthor numbers exceeding 2000; (2) two groups with coauthor numbers greater and less than 2000 present significantly different ($F(1, 38) = 17.89$, $p < 0.001$) in citations based on each with at least one citing article; (3) the cluster with numerous coauthors has lower bibliometrics based on personal IRA; (4) the most cited author is Krzysztof Burnecki from Poland; (5) the dominant nations with higher bibliometrics on pure and applied mathematics are the US, France, and the UK. Social network analysis provides wide and deep insight into the relationships among coauthors. The results can provide readers with knowledge and concept diagram on the topic of pure and applied mathematics in the literature.

1. Introduction

Team science in basic science has been accompanied by a trend in the numbers of authors included in scientific publications [1]. The mean number of individuals listed as authors in articles indexed in PubMed from 1975 to 2016 has increased from 1.9 to 5.67 per article [2]. Authorship trends in research articles published in three leading general medical journals (JAMA, The Lancet, and New England Journal of Medicine) in 2005, 2010, and 2015 have also been verified [3]. The median number of authors per article was increased in all three journals from a range of 8-11 in 2005 to 11-18 in 2015. Whether the trend of

author collaboration, particularly on individual research achievements (IRA), can be generalized to other journals or disciplines, such as the topic of pure and applied mathematics, is still unknown.

There are many metrics used for evaluating author IRA. The h-index The h-index [4] is a simple way to measure both the productivity and citation impact of the publications of a scientist or scholar. The index is defined as the maximum value of h such that the given author has published h papers that have each been cited at least h times in publications [4], see Figure 1. However, many drawbacks were proposed by authors [5-11], such as each author with equal contributions in an article and the h-index without considering the other two parts (i.e., excess and tail citations) in Figure 1.

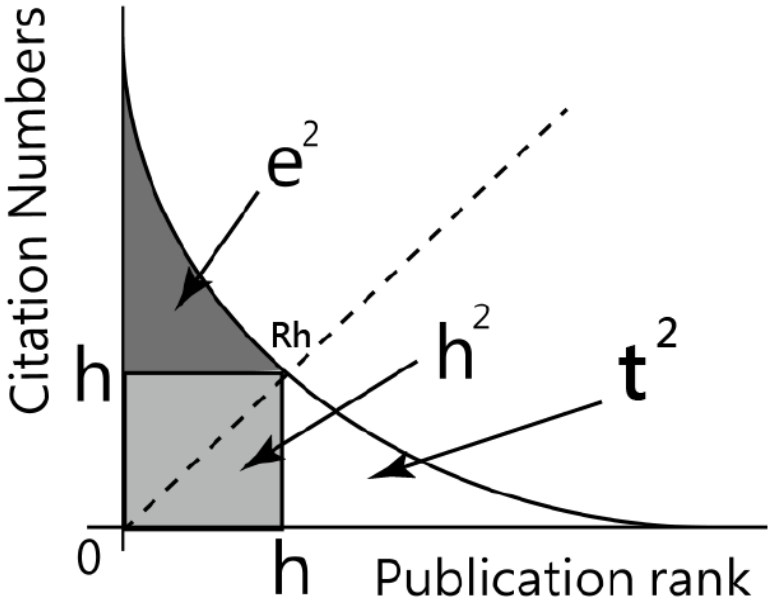

Figure 1. Three parts are divided and related to h-index.

Furthermore, every June, millions of academic scholars pay close attention to the Journal Citation Reports (JCR) ranking the journal impact factor (JIF) for each indexed journal. However, no such author IFs

(AIFs) [12, 13] or bibliometric indices [4, 7-10] have gained scientists' or scholars' attention as much as JIF does annually in the academia. How to apply an appropriate authorship-weighted scheme (AWS) [5, 6, 11] for tracking the dynamics of individual scientific impact and quantifying the coauthor contributions in scientific disciplines is worth studying.

In this study, we aim to present the most cited authors who published articles on pure and applied mathematics and the dominant nations in this field and investigate whether numerous coauthors can own more citations in comparison to other fewer coauthors in an article.

2. Methods

2.1. Data source

We obtained 48 abstracts based on journal article from Pubmed Central (PMC) by searching the keywords of "pure and applied mathematics" [all fields] since 2016. A total number of 133 citing articles matching to the citable papers in PMC were attained. The number of 40 articles were quoted by at least one publication in PMC. All data were downloaded from PMC, which means the study is not necessary for the ethical approval according to the regulation promulgated by the Taiwan Ministry of Health and Welfare.

2.2. Four metrics proposed in this study

The h-index can be divided into three parts [8, 9], see Figure 1. Many modified h-index had been suggested, such as (1) the g-index [10] ($\leq \sum_{i=1}^{g} c_i/g$, where c_i = number of citations to i-th publication); the x-index [7] ($= \sqrt{\max_i(i \times c_i)}$); the L-index [14] ($= \ln(\sqrt{\sum_{i=1}^{n} \frac{c_i}{a_i y_i}}) + 1$), where a_i = number of authors of i-th publication, y_i = age in years of i-th publication, n = number of publications; the h'-index [8] ($= h^* rh$, where $rh = e/t$, and let $t = 1$ if $t < 1$, see Figure 1, perfectionist at $rh > 1$, prolific type at $rh = 1$, and mass-production at $rh < 1$).

Due to the contradiction on h'-index for the results through the formula (= h*rh) which lets h' be greater than $h+1$ (e.g., $h=2$ and $rh=2$ make h' be 4 greater that $h+1=3$). We thus propose the complemental one [i.e., h-plus = $h + rh/(1+rh)$] ranging h-plus between h and $h+1$.

2.3. The AWS for quantifying coauthor contributions

We see the L-index [14] (= $\ln(\sqrt{\sum_{i=1}^{n}\frac{c_i}{a_i y_i}})+1$) applying a_i as the number of authors of i-th publication (i.e., equal size to coauthors). Similarly, other indices mentioned above ignore the author contributions to the article. We assume all coauthor gain equal credits in mathematics discipline using the alphabet ordering of author names [15]. That is, n *Weight* = $\frac{1}{n}$, n = number of coauthor and the sum of authorships = 1 for each paper, which is different from the traditional computation in bibliometrics using weight = 1 for all coauthors.

2.4. Author impact factor (AIF)

Author impact factor (AIF) used for evaluating individual research achievement (IRA) as Equation (1) [13]:

$$AIF = \frac{\sum \text{Cited papers based on } W_j}{\sum \text{Citable papers} \times W_j \text{ in the given yrs}}), \tag{1}$$

The four indices of AIF, $Ag(=\sum_{i=1}^{g} c_i / g)$, x-index(= $\sqrt{\max_i(i \times c_i)}$), and h-plus = $h + rh/(1+rh)$ were used for evaluating IRA in two groups (i.e., low and high citations with a cutting point at 2000 coauthors in an article). One way ANOVA was performed to examine whether numerous coauthors earn more citations on the topic of pure and applied mathematics.

2.5. Social network analysis using Pajek software

In keeping with the Pajek guidelines [16], we applied social network analysis (SNA) to cluster authors. Usually, the relation valued by the weight is defined by the number of connections between two authors [5, 6, 17]. The clusters can be determined by a specific algorithm as named degree centrality.

2.6. Using bootstrapping sampling method to estimate 95% confident intervals

SNA was applied to determine the representative of each cluster. The algorithm of community partition was performed to identify the number of clusters. Each author was, in turn, assigned to the designated cluster represented by the author who owns the highest centrality degree in his/her cluster. As such, each author can be matched to his/her metrics, clusters, and even the affiliated nation by the author-made MS-Excel module.

The bootstrapping method [18] was applied to examine differences in metrics among author clusters. A total of 1000 medians retrieved from the median of the 100 random cased were used to estimate the 95% confidence intervals (CI) for a metric of a given cluster. As such, the difference can be determined by judging the two 95% CI bands separated from each other.

2.7. Creating dashboards on Google Maps

We applied the author-made modules in MS-Excel and the SNA in Pajek to gain the author clusters. The pages of Hyper Text Mark-up Language (HTML) used for Google Maps were created. All relevant bibliometric indices were linked to dashboards on Google Maps.

3. Results

3.1. TASK 1: Presenting the most cited author on nurse bullying

The most cited authors is Krzysztof Burnecki from Poland with three articles cited 7, 3, and 1 each [19-21] until 2018 with relatively high metrics (citable = 0.64, cited = 2.1, AIF = 2.1, Ag = 1.33, h = 1, g = 1, x = 1.15, h-plus = 1.6, r-ratio = 1.51, L = 2.13), see Figure 2. Interested readers are invited to scan the QR-Code in Figure 2 to see the author's publication outputs in PMC by clicking the specific author bobble.

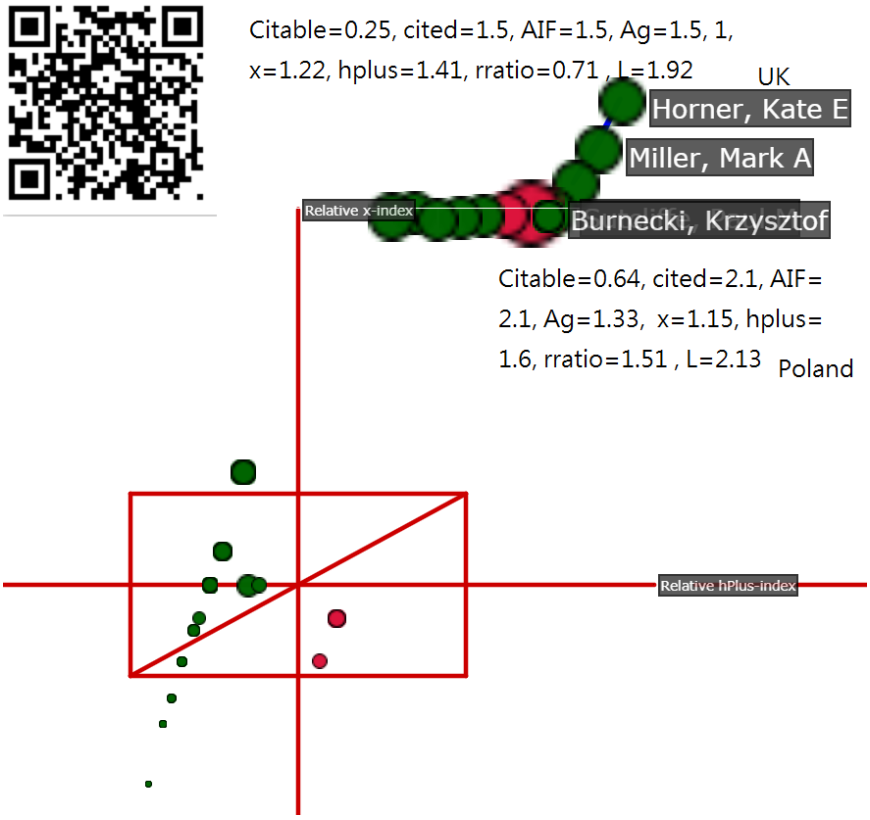

Figure 2. The most cited authors on pure and applied mathematics.

3.2. TASK 2: Selecting the ten top author clusters with high degree centrality

The top 10 author clusters were separated as shown in Figure 3. The representatives with the most degree centrality (DC) are shown for each cluster. The author Aleksander Weron from Poland earns the highest DC, implying more author collations and articles exist since 2016.

The greater number of members in a cluster (i.e., the high citation group) is represented by M. Mikuz from the US. The interested readers are also recommended to scan the QR-code in Figure 3 to see the detailed information in PMC by clicking the word of publication when the specific author bubble is selected.

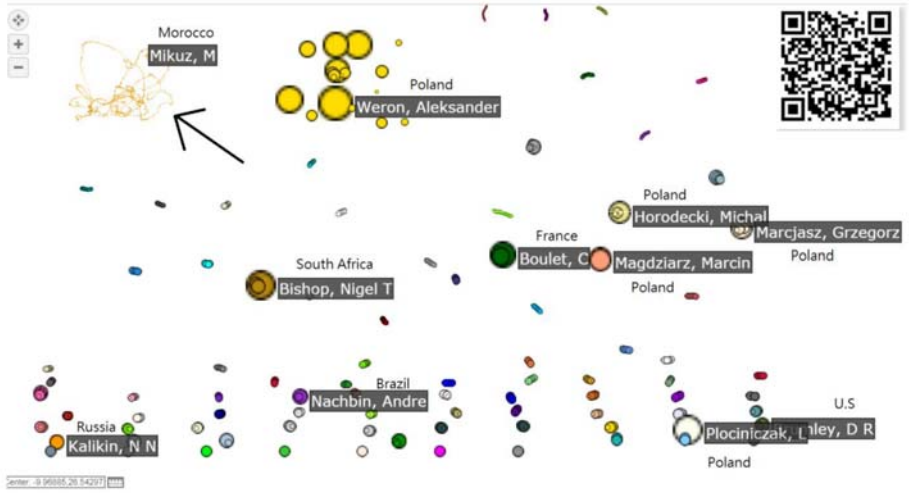

Figure 3. Cluster of author collaborations on the topic of pure and applied mathematics.

3.3. TASK 3: Comparisons of differences in metrics among clusters

The differences in metrics (i.e., x-index, h-plus, Ag, and AIF) were found ($p < 0.05$), see Figure 4, when any two 95% CI bands were separated from each other. We can see the high citation group located at

the first place in both panels has a lower median IRA in comparison to other clusters. However, two groups present significantly different ($F(1, 38) = 17.89$, $p < 0.001$) in citations based on each with at least one citing article. In contrast, two groups present are identical ($F(1, 145) = 3.44$, $p = 0.06$) when data are based on those articles without any citation.

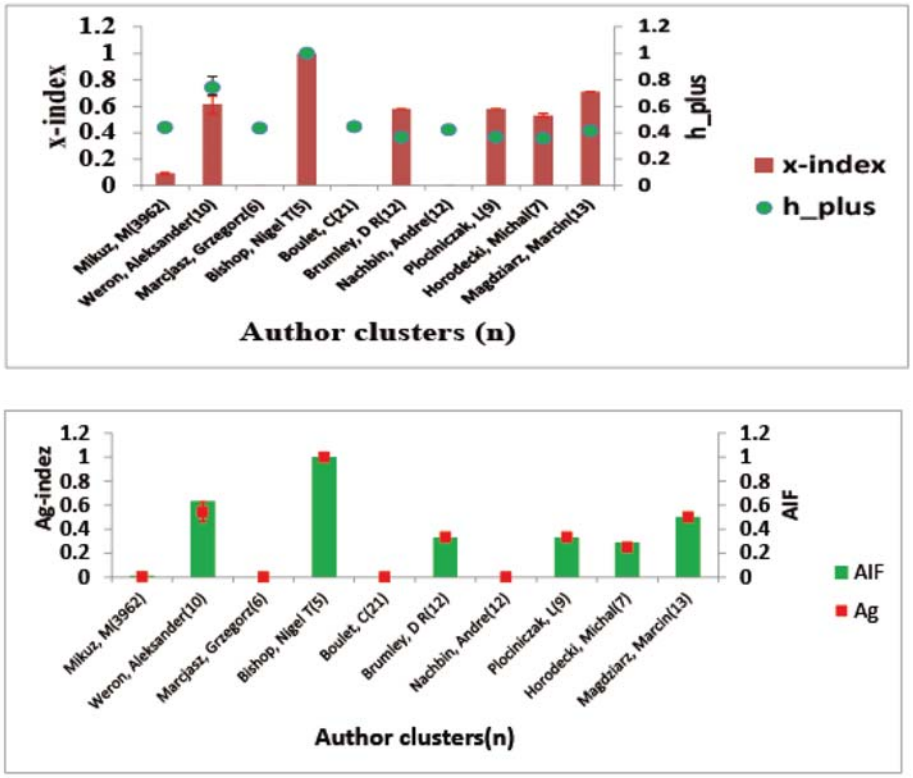

Figure 4. Comparisons of indices among author clusters.

3.4. TASK 4: Overall author IRA based on *x*-index dispersed on a dashboard

The top three counties/areas based on *x*-index [7] are from the US (= 17.96), France (= 10.36), and the UK (= 13.42) shown in Figure 5. The overall *x*-index is 41.24 in the base of individual author *x*-indices in descending order for the nation.

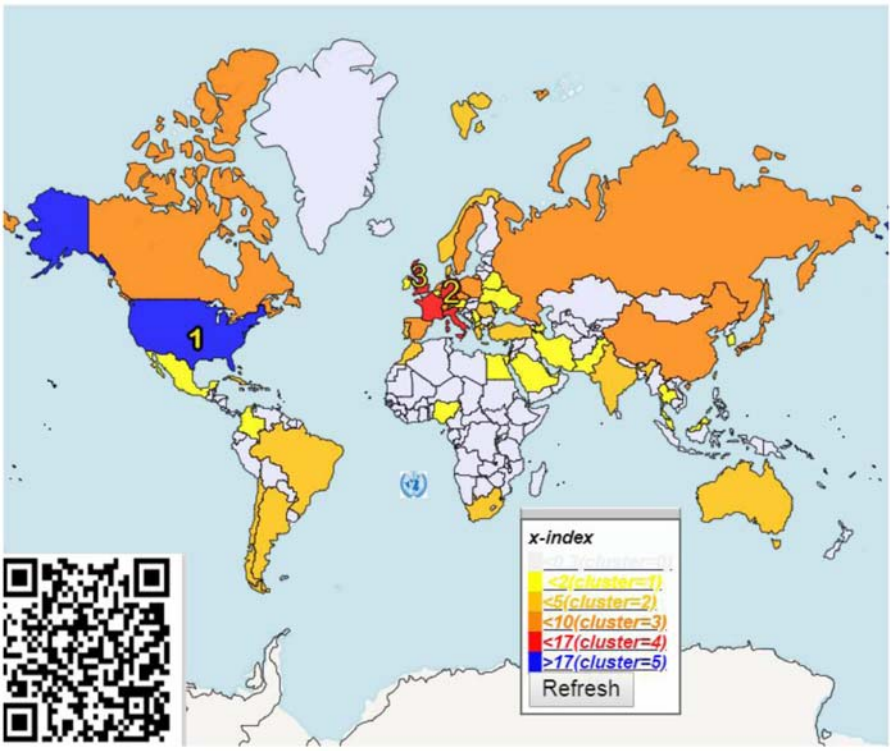

Figure 5. The *x*-indices dispersed around the world.

4. Discussion

4.1. Main findings and implications

It is surprising that 43 articles (29%) have many coauthor exceding 2000 (ranged from 2824 to 2929) included in four journals (i.e., Eur. Phys. J. C. Part Fields; Phys. Rev. Lett.; Eur. Phys. J. C. Part Fields; and Phys. Rev. Lett.) and authored by the two M. Aaboud and G. Aad as primary authors placed at the first order in an article byline.

Two groups with coauthor numbers greater and less than 2000 present significantly different ($F(1, 38) = 17.89$, $p < 0.001$) in citations if only those articles with at least one citing article were included. If the terms were replaced with all articles no matter whether cited by other

articles, no difference was found l ($F(1, 145) = 3.44$, $p = 0.06$). Referring to the small p-value (= 0.06), we confirm that many journal editors would like many coauthors included in an article because a higher probability in citations may occur in the future. That is why author collaborations have been accompanied by a trend in the numbers of authors included in publications [1] and the reason why the mean number of individuals listed as authors in articles has increased from 1.9 to 5.67 per article [2].

However, due to the contributions have bee equally shared by coauthors in the alphabet order in mathematics, the cluster with numerous coauthors has lower bibliometrics in the base of personal IRA because of lower proportional citations and publications leading bibliometrics to relatively lower level.

Furthermore, Google Maps have provided users to capture an overall geospatial visualization in the past [5, 6, 17, 22, 23]. How to apply Google Maps for reporting study results is worth studying in bibliometric analyses, like we did in this study and showed the most cited authors in Figure 2 and the dominant nations with higher bibliometrics on pure and applied mathematics.

4.2. Limitations and future research

Although our findings based on the above analyses have been illustrated, there are several potential limitations that should be overcome in the future. First, all data were linked to PMC which cannot generalize the results to other bibliometric databases and other disciplines.

Secondly, there might be some biases when matching authors' name to calculate the IRA because some different authors with the same name exist. Therefore, the result of author relationship analysis might be influenced by the inaccuracy occurred by the disparate authors with identical names.

Third, many algorithms were used in SNA. The degree centrality used for generating figures might be different if different algorithms were applied.

Fourth, the formula of quantifying coauthor contributions used in this study is assumed all author equal in an article. Any change for the authors we calculated in indices might present distinct results for authors.

Fifth, the data were extracted from *PMC which is* different from other authors using the citation databases-such as the Scientific Citation Index (SCI; Thomson Reuters, New York, NY, USA) and Scopus (Elsevier, Amsterdam, The Netherlands). The results of the most cited authors and nations might be disparate if other databases were applied.

Finally, many other topics besides the one of pure and applied mathematics that should be further investigated on the association between the number of coauthors and citation probabilities in the discernable future.

4.3. Conclusion

The association between the number of coauthors and citations has been verified on the topic of pure and applied mathematics. Other topics or disciplines are recommended to study further using the visualizations, particularly on the type of dashboard on Google Maps. The overall knowledge information provided to readers can be attained, accordingly.

List of Abbreviations

AIF: Author impact factor.

AWS: Authorship-weighted scheme.

DC: Degree centrality.

IF: Impact factors.

IRA: Individual research achievement.

PMC: PubMed Central.

SNA: Social network analysis.

VBA: Visual basic for application.

References

[1] P. Fontanarosa, H. Bauchner and A. Flanagin, Authorship and team science, Journal of the American Medical Association 318(24) (2017), 2433-2437.

DOI: https://doi.org/10.1001/jama.2017.19341

[2] U.S. National Library of Medicine, Number of Authors per MEDLINE®/PubMed® citation, Updated May 16, 2017.

http://www.nlm.nih.gov/bsd/authors1.html

[3] C. C. Muth and R. M. Golub, Trends in Authorship and Team Science in Major Medical Journals, 2005-2015, Presented at: Eighth International Congress on Peer Review and Scientific Publication; September 11, Chicago, IL, 2017.

[4] J. E. Hirsch, An index to quantify an individual's scientific research output, Proceedings of the National Academy of Sciences of the United States of America 102(46) (2005), 16569-16572.

DOI: https://doi.org/10.1073/pnas.0507655102

[5] T. W. Chien, H. Y. Wang, Y. Chang and W. C. Kan, Using Google maps to display the pattern of coauthor collaborations on the topic of schizophrenia: A systematic review between 1937 and 2017, Schizophrenia Research (2018), Sep 24. pii: S0920-9964(18)30573-5.

DOI: https://doi.org/10.1016/j.schres.2018.09.015

[6] T. W. Chien, J. C. Chow, Y. Chang and W. Chou, Applying Gini coefficient to evaluate the author research domains associated with the ordering of author names: A bibliometric study, Medicine 97(39) (2018), e12418.

DOI: https://doi.org/10.1097/MD.0000000000012418

[7] T. Fenner, M. Harris, M. Levene and J. Bar-Ilan, A novel bibliometric index with a simple geometric interpretation, PLoS One 13(7) (2018), e0200098.

DOI: https://doi.org/10.1371/journal.pone.0200098

[8] C. T. Zhang, The h'-index, effectively improving the h-index based on the citation distribution, PLoS ONE 8(4) (2013), e59912.

DOI: https://doi.org/10.1371/journal.pone.0059912

[9] C. T. Zhang, The *e*-index, complementing the *h*-index for excess citations, PLoS One 4(5) (2009), e5429.

DOI: https://doi.org/10.1371/journal.pone.0005429

[10] L. Egghe, Theory and practice of the *g*-index. Scientometrics 69(1) (2006), 131-152

DOI: https://doi.org/10.1007/s11192-006-0144-7

[11] V. Vavryčuk, Fair ranking of researchers and research teams, PLoS One 13(4) (2018), e0195509.

DOI: https://doi.org/10.1371/journal.pone.0195509

[12] G. Lippi and C. Mattiuzzi, Scientist impact factor (SIF): A new metric for improving scientists' evaluation, Annals of Translational Medicine 5(15) (2017), 303.

DOI: https://doi.org/10.21037/atm.2017.06.04

[13] R. K. Pan and S. Fortunato, Author impact factor: Tracking the dynamics of individual scientific impact, Scientific Reports 4 (2014), Article 4880.

DOI: https://doi.org/10.1038/srep04880

[14] A. V. Belikov and V. V. Belikov, A citation-based, author- and age-normalized, logarithmic index for evaluation of individual researchers independently of publication counts, F1000 Research 4 (2015), Article 884.

DOI: https://doi.org/10.12688/f1000research.7070.1

[15] José-Alain Sahel, Quality versus quantity: Assessing individual research performance, Science Translational Medicine 3(84) (2011), 84cm13.

DOI: https://doi.org/10.1126/scitranslmed.3002249

[16] V. Batagelj and A. Mrvar, Pajek - Analysis and Visualization of Large Networks, In M. Jünger and P. Mutzel (Editors) pp.77-103. Graph Drawing Software, Springer, Berlin, 2003.

[17] T. W. Chien, Y. Chang and H. Y. Wang, Understanding the productive author who published papers in medicine using National Health Insurance Database: A systematic review and meta-analysis, Medicine (Baltimore) 97(8) (2018), e9967.

DOI: https://doi.org/10.1097/MD.0000000000009967

[18] B. Efron, Bootstrap methods: Another look at the jackknife, The Annals of Statistics 7(1) (1979), 1-26.

[19] T. Sungkaworn, M. L. Jobin, K. Burnecki, A. Weron, M. J. Lohse and D. Calebiro, Single-molecule imaging reveals receptor-G protein interactions at cell surface hot spots, Nature 550(7677) (2017), 543-547.

DOI: https://doi.org/10.1038/nature24264

[20] A. Weron, K. Burnecki, E. J. Akin, L. Solé, M. Balcerek, M. M. Tamkun and D. Krapf, Ergodicity breaking on the neuronal surface emerges from random switching between diffusive states, Scientific Reports 7 (2017), Article 5404.

DOI: https://doi.org/10.1038/s41598-017-05911-y

[21] G. Sikora, K. Burnecki and A. Wyłomanska, Mean-squared-displacement statistical test for fractional Brownian motion, Physical Review E 95(3) (2017), Article 032110.

DOI: https://doi.org/10.1103/PhysRevE.95.032110

[22] C. H. Sekercioglu, Quantifying coauthor contributions, Science 322(5900) (2008), Article 371.

DOI: https://doi.org/10.1126/science.322.5900.371a

[23] S. Dasgupta, A. S. Vaughan, M. R. Kramer, T. H. Sanchez and P. S. Sullivan, Use of a Google map tool embedded in an internet survey instrument: Is it a valid and reliable alternative to Geocoded address data?, JMIR Research Protocols 3(2) (2014), Article e24.

DOI: https://doi.org/10.2196/resprot.2946

■