The most cited MeSH terms and authors who published papers in Pubmed Central on the topic of medicine and health using bibliometric analyses

Running title: Patterns of author collaboration on medicine and health

ABSTRACT

Aims: We visualized the current state of research on publication outputs and citations in the field of medicine and health to uncover topic burst and citations among medical subject headings(MeSH) clusters.
Study design: a bibliometric analysis.
Place and Duration of Study: Using Pubmed indexed articles to inspect the characteristics of topics on medicine and health since 1969.
Methodology: Selecting 156 abstracts, author names, countries, and MeSH terms on January 10, 2019, from Pubmed Central(PMC) based on the terms of medicine and health in the title since 1969, we applied the x-index and impact factor to evaluate author individual research achievements and compute MeSH bibliometric performances. The bootstrapping method was used to estimate the median and its 95% confidence intervals and make differences in metrics among MeSH clusters. The dominant nations were selected using the x-index to display on a dashboard. We programmed Microsoft Excel VBA routines to extract data. Google Maps and Pajek software were used for displaying graphical representations.
Results: We found that (1)the dominant countries/areas are the Unlited States, Taiwan, and Australia; (2)the author Grajales, Francisco Jose 3rd form Canada has the most cited metrics such as author IF=39.46 and x-index=6.28; (3)the MeSH terms of organization & administration, standards, and prevention & control gain the top three degree centralities among MeSH clusters; (4) No any differences in metrics were found among MeSH clusters; (5) the article(PMID= 24518354) with three MeSH term of delivery of health care, social media, and software and published in 2014 was cited most at least 62 times.
Conclusion: Social network analysis provides wide and deep insight into the relationships among MeSH terms. The MeSH weighted scheme and x-index were recommended to academics for computing MeSH citations in the future.

Keywords: x-index, medical subject headings, Google Maps, social network analysis, medicine and health, bibliometric analysis.

1. INTRODUCTION

Medicine and health is an important issue in healthcare settings. The term of medicine is defined as the science and practice of the diagnosis, treatment, and prevention of disease. Medicine encompasses a variety of healthcare practices evolved to maintain and restore health by the prevention and treatment of illness.

Health, on the other hand, is defined as a state of complete physical, mental and social well-being and not merely the absence of disease or infirmity[1][2]. Health can be further defined as the ability to adapt and manage physical, mental and social challenges throughout life[3].

As of January 10, 2019, more than 33.609 papers were published on Pubmed.com by searching the keyword “medicine and health” and 321 in the paper title including “medicine and health,” which present the importance of author collaborations in academics in the past. On the journal perspective, the keywords of (medicine[journal] or health[journal]) were searched and found 22,075 articles. One article regarding the most-cited publications on the topic of medicine and health has been published[4]. The most cited topic burst and authors have not been addressed in the field of medicine and health.

Given the importance of medicine and health, many scientific researchers[5-7] have focused on reviewing related literature to identify the characteristics and status of health in recent years. However, much of these efforts were only considering specific subfields of medicine and health using traditional descriptive analyses or systematic reviews. Even some[8-10] have conducted citation analyses on articles related to medicine and health; none reports topic burst incorporated with citations on medicine and health until now.

Papers on the bibliometric perspective of medicine and health can provide readers with methods of analyzing data of scientific literature quantitatively and then gain knowledge of the meta-analysis related to the research questions [11,12]. The combined use of methodologies that give information on different aspects of scientific output is generally recommended [13]. Also, discussions relating to the collaborative status and overall topic burst on medicine and health remain relatively scarce.

Due to the citations representing the importance and eminence on an academic topic, we are thus interested in following four issues, including (1) which nations was dominant in the field of medicine and health; (2) which medical subject headings(MeSH) were cited most by papers in recent years; (3)is any difference in scientometrics among MeSH clusters;(4)which article was cited most in the pas

We aim to apply x-index (= $\sqrt{\max_i(i \times c_i)}$, where i is the i-th publication in descending order of citations, and ci is the citation at the i-th publication)[14] that can effectively improve the h-index(≤citation at the i-th publication)[15] in bibliometric analyses[16,17] and investigate the four questions mentioned above. Google Maps[18.19] will be applied to the study for illustrating results on dashboards.

2. MATERIAL AND METHODS

2.1. Data Sources

We programmed Microsoft Excel VBA (visual basic for applications) modules[20,21] to extract abstracts and their corresponding coauthor names as well as the countries/areas of the first authors for each article on January 20, 2019, from Pubmed Central(PMC) based on the topic of medicine and health. Only those abstracts published by the keyword medicine and health [title] and labeled with journal article were included. Others like those labeled with Published Erratum, Editorial or without author nation name were excluded from this study. A total of 156 eligible abstracts(i.e., journal articles and entitled including medicine and health) were downloaded from PMC.

As this study did not involve the examination or treatment of patients or a review of patient records, it was exempt from review and approval by our research ethics committees.

2.2 The MeSH weighted scheme and the x-index used for quantifying citations

The way for quantifying MeSH terms would be defined first in this study. As the author-weighted scheme[20,21], we define the MeSH weighted scheme as Wi=1/n and ci=Wi*pi for citations and publications, respectively, where n=number of articles, ci, and pi as citations and publications on the i-th article.

The x-index[14] is used for improving h-index by emphasizing the citations in h-core articles. For instance, one article with 100 citations has h=1 and x=10, respectively. In contrast, 100 articles with one citation have h=1 and x=10 for each article. The MIF, like the journal impact factor, is defined as Eq.(1) with the inclusion of weights mentioned above:

$$W_j\ MIF = \frac{\sum Cited.papers.based.on.\times W_j.in.a.given.year.and.the.proceeding.5.yrs}{\sum Citable.papers.\times W_j.in.the.given.5.yrs}. \quad (1)$$

The weighted publications are set at 1.0 if the values<1.0 for avoiding the MIF inflated too much.

2.3 Social network analysis using Pajek software

In keeping with the Pajek guidelines [22] using social network analysis(SNA), we defined a MeSH term as a node(or an actor) that is connected to another counterpart at another node through the edge of a line. Usually, another weight is defined by the number of connections between two nodes. For instance, Notes(A and B) having 3 times of connects are graphically represented by the statement and symbols with {A B 3}. A series of those co-occurrence statements can thus construct a social network.

2.4. Graphical representations to Report

A visual display with the publication outputs labeled by the1st author nations was made for presenting the distribution of nations on medicine and health. The quantity is colored by the size of publications. The most cited MeSH terms sized by MIF and colored by MIF were shown based on both axes(i.e., x-index on the x-axis and MIF on the y-axis),

2.5 MeSH clusters using SNA to separate

SNA was applied to determine the representative for each cluster. The algorithm of community partition was performed to separate clusters. Each MeSH was assigned to the corresponding cluster which was similar to the author analysis[20,21] and sized with the highest degree centrality(DC) within the cluster. As such, each MeSH contributes equally proportional parts(i.e., 1/n) in an article and then shows the total DC in the cluster network with the respective metrics.

The bootstrapping method [23] was applied to examine differences in metrics among MeSH 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 is determined by judging the two 95% CI bands separated from each other.

2.6 Creating dashboards on Google Maps

We applied the author-made modules in MS-Excel and the SNA in Pajek to gain the MeSH 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 AND DISCUSSION

3.1 TASK1: The dominant nations on anesthesiology around the world

The dominant countries/areas on medicine and health are the US, Taiwan, and Australia, see Figure 1. The author Grajales, Francisco Jose 3rd form Canada has the most cited indices including author IF=39.46 and x-index=6.28, see Figure 2

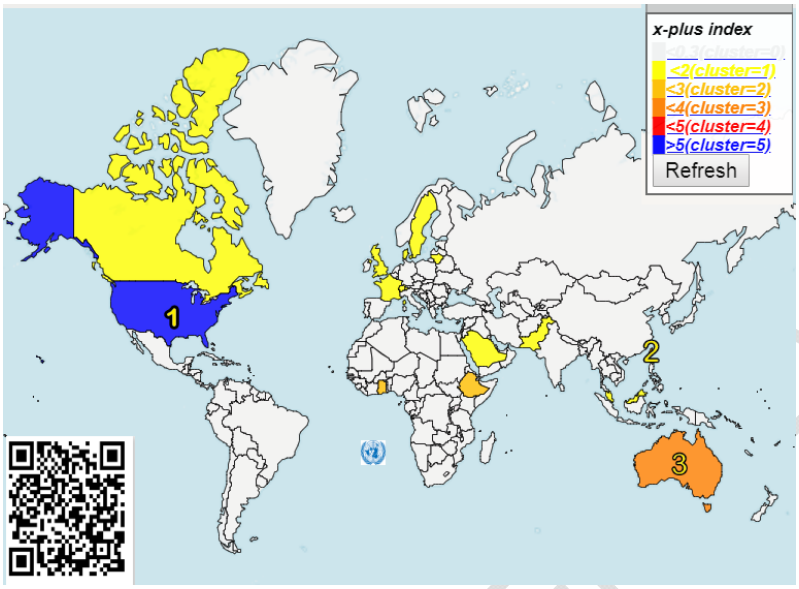


Fig. 1. x-index shown for nation distribution around the world

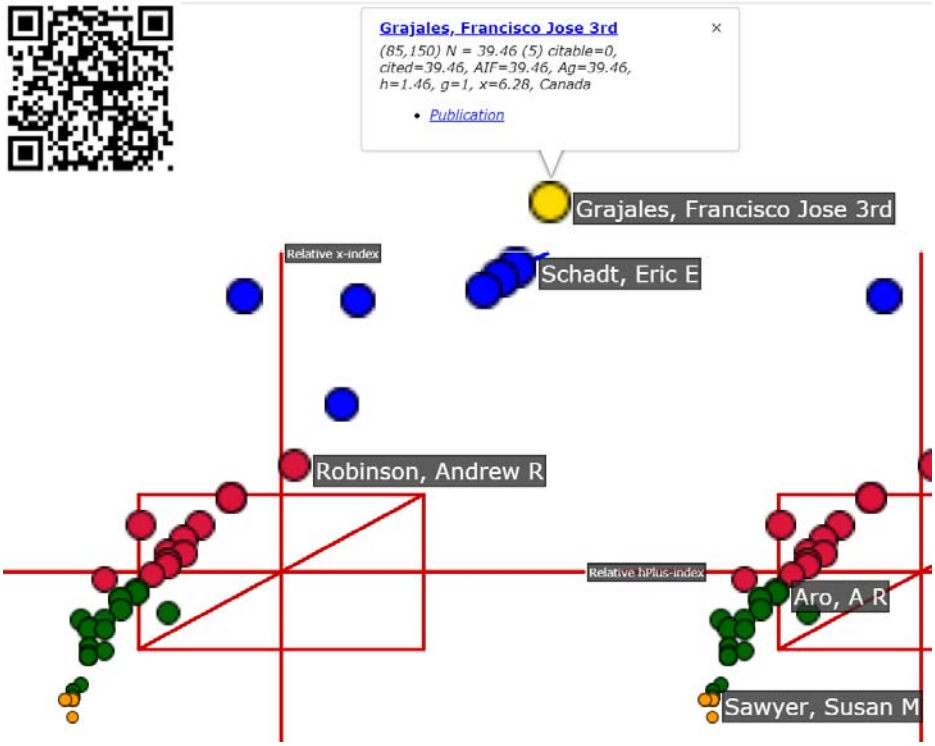


Fig. 2. The most cited authors shown on a dashboard

3.2 TASK2: selecting the ten top MeSH clusters with high degree centrality

The MeSH terms of organization and administration, standards, prevention and control gain the top three degree centralities among clusters. Interested readers are invited to scan the QR-Code in Figure 3 to see the MeSH terms regarding relevant outputs in PMC by clicking the specific MeSH bobble.

The most cited MeSH term is social media with relatively higher metrics such as total weighted citations =20.67, MIF=20.67, x=4.55, see Figure 4.

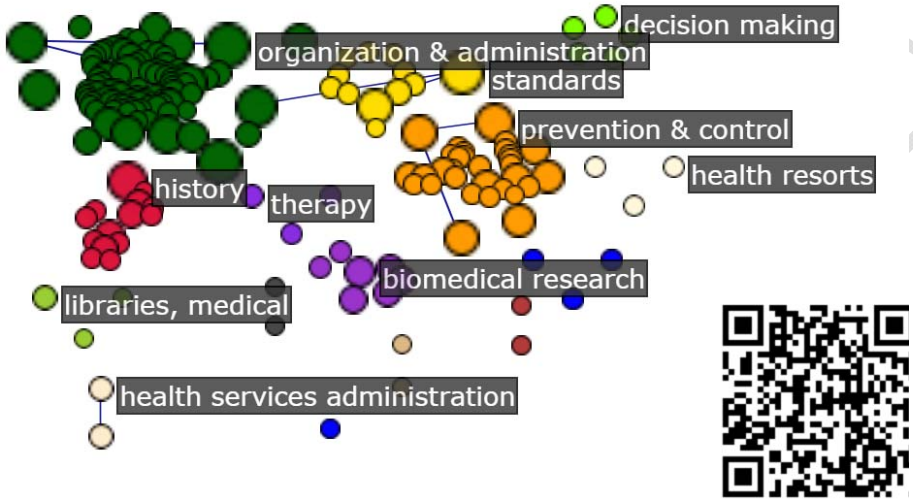


Fig. 3. Mesh clusters dispersed on a dashboard

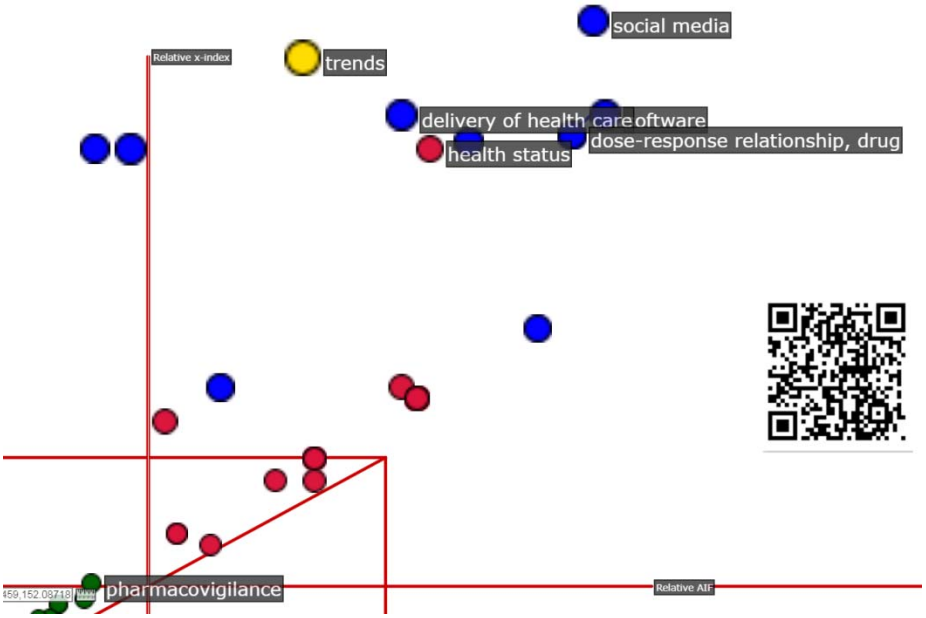
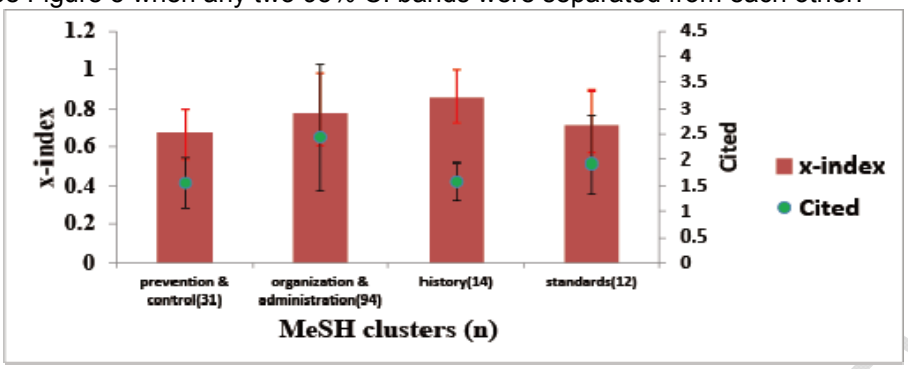


Fig. 4. The most cited MeSH terms shown on a dashboard

3.3 TASK3: Comparisons of differences in metrics among MeSH clusters

No difference in metrics (i.e., x-index, citations, MIF, and Ag-index[24]) were found ($p > .05$), see Figure 5 when any two 95% CI bands were separated from each other.



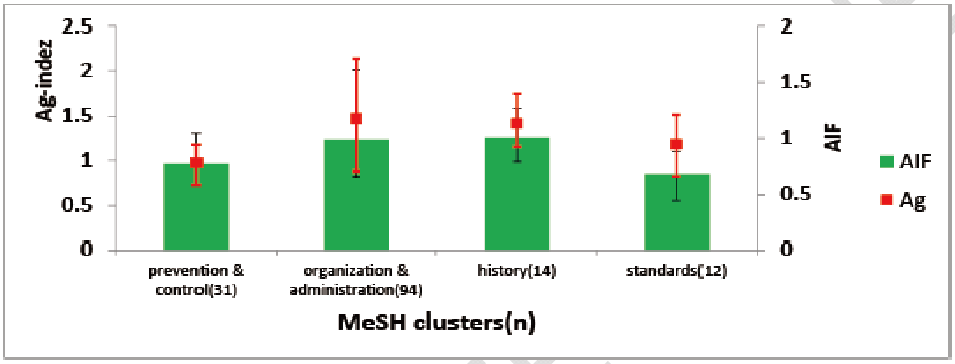


Fig. 5. Comparison of MeSH clusters using the bootstrapping method

3.5 TASK4: The most cited article on medicine and health

The article(PMID= 24518354) with three MeSH term of delivery of health care, social media, and software [25] published in 2014 has been cited at least 62 times and.

3.6 Discussion

We observed that (1)the dominant countries/areas were the US, Taiwan, and Australia, which can be an outline of the dominant nations in the field of medicine and health, but different from those studies[25,26] addressing the dominant nations in the US and Europe for the science discipline.

The most cited author on the topic of medicine and health was Grajales, Francisco Jose 3rd whose articles are worth reading for readers interested in the field of medicine and health.

The three MeSH terms of organization & administration, standards, and prevention & control gain the top three degree centralities based on the cluster category, which findings are worth understanding the main topics in the fields of medicine and health in the past literature.

The finding of no any difference in metrics among MeSH clusters tells us which types of MeSH term which are equal in bibliometric indices. The way of comparisons in metric difference can be applied to author clusters or institutes in the future.

The article(PMID= 24518354) published in 2014 was cited most at least 62 times[27] in the past. The readers who are interested in the field of medicine and health are suggested to read the one with a a vast number of citations in the literature.

Although the h-index[15] is a simple and popular author-level metric used for combining both productivity and citations for evaluating individual research achievements, one of its shortcomings is less discriminative power[28] due to many with identical value in an integer unit. Although many concepts of bibliometrics have already been proposed in the past [14,16, 17,24], we have not seen any that can be effectively applied to the scientific disciplines. We suggest using the one of x-index that can effectively improve h-index with less discriminative power for evaluating authors and MeSH terms in the past.

We also demonstrated the utility of the x-index using the MeSH weighted scheme for quantifying contributions among MeSH terms in an article, which was never seen in literature before. In Figure 3, we see the MeSH-based x-indexes and MeSH-based impact factor shown on a dashboard using Google Maps to display, which is worth recommending to authors for use in the future.

The bibliometric indices are dependent on both quantity (i.e., the number of publications) and quality(i.e., the number of articles being cited), which are suitable for use on the topics of medicine and health as we did in this study. In comparison to the author-based bibliometrics[20, 21], the MeSH-based metrics gain higher values than the author indices because of a larger number of occurrences in MeSH terms than authors. On Figures 2 and 3, it is easy to see the topic burst regarding medicine and health exist in the past.

As above, the x-index[14] has many advantages compared to the h index. For instance, ten publications with ten citations each have an identical h-index and x-index at 10(or $\sqrt{}$ 10*10 for x-index). In contrast, One publication with 100 citations leads to a difference in h-index(=1) and x-index(=10=$\sqrt{}$ 1*100). On the other hand, 100 publications with only one citation each have different results in h-index(=1) and x-index(=10=$\sqrt{}$100*1)[14].

The most worth-noting feature is the general weighted scheme that can be totally congruent with the real-world scenario in practice. That is, the contributions were determined by the weights(=1/n) instead of all with an identical value(=1) no matter the ordering of MeSH terms.

The second feature is the intrinsic dynamic character along with the periods changed by years, like the journal citation report(JCR) locating JIF each year in June, to examine the change of IF.

The reason we applied x-index in this study is the strength of the index in practice. According to the illustration in the study of Fenner and his colleagues[14], the x-index can truly extend the feature of an author or MeSH term with quality and quantity achievements in academics as mentioned above.

The importance of these metrics to the scientific community can be referred to other disciples. For instance, members in their departments have gain many types of awards, no matter what types of golden or silver medals in a period, The items as the publicaionts and the weights(e.g., credits for golden and silver medals) as the citations can be used for evaluating the achievements for individuals or groups(e.g., departs). The dashboards shown on Google Maps that can be an assessment of indices produc relevant information for individuals or groups in decision-making.

Although findings are based on the above analysis, there are still several potential limitations that may encourage further research efforts. First, all data were extracted from the PubMed database. There might be some biases of understanding the matched MeSH terms because of some different terms with the asterisk represented by major MeSH in the article, which will affect the result of MeSH relationship analysis by the accuracy of the indexing MeSH terms.

Second, many algorithms have been applied to SNA. We merely applied the algorithm of degree centrality in Figures. Any changes in the algorithm used in this study might present a different pattern and judgment to the results. [29]

Third, the data extracted from PMC cannot be generalized to other major citation databases—such as the Scientific Citation Index (SCI; Thomson Reuters, New York, NY, USA) and Scopus (Elsevier, Amsterdam, The Netherlands). Such as the most cited authors are determined by the paper selections on Pubmed.

4. CONCLUSION

Social network analysis provides a wide and deep insight into the relationships among MeSH terms. The MeSH weighted scheme and x-index can be applied to academics for computing MeSH citations in the future. The metrics such as h-index or x-index that can be applied other fields or administrations for evaluating individual (or department) achievements using the award items(e..g, golden medal) and the weighted scores(e.g., 50 points for one golden medal) .

ETHICAL APPROVAL, CONSENT

As this study did not involve the examination or treatment of patients or a review of patient records, it was exempt from review and approval by our research ethics committees.

REFERENCES

1. Grad FP. The Preamble of the Constitution of the World Health Organization". Bulletin of the World Health Organization 2002; 80 (12): 982.

2. World Health Organization. Constitution of the World Health Organization – Basic Documents, Forty-fifth edition, Supplement, October 2006.

3. Huber M, Knottnerus, JA, Green L, Horst H van der, Jadad AR, Kromhout D, Leonard BLK, Loureiro MI. How should we define health?. BMJ 2011; 343: d4163. doi:10.1136/bmj.d4163

4. Tian J, Li M, Lian F, Tong X.The hundred most-cited publications in microbiota of diabetes research: A bibliometric analysis.Medicine (Baltimore). 2017 Sep;96(37):e7338. doi: 10.1097/MD.0000000000007338.

5. Orhan C, Van Looveren E, Cagnie B, Mukhtar NB, Lenoir D, Meeus M.Are Pain Beliefs, Cognitions, and Behaviors Influenced by Race, Ethnicity, and Culture in Patients with Chronic Musculoskeletal Pain: A Systematic Review.Pain Physician. 2018 Nov;21(6):541-558.

6. Hussen S, Wachamo D, Yohannes Z, Tadesse E.Prevalence of chlamydia trachomatis infection among reproductive age women in sub Saharan Africa: a systematic review and meta-analysis.BMC Infect Dis. 2018 Nov 26;18(1):596. doi: 10.1186/s12879-018-3477-y.

7. Téblick S, Ruymaekers M, Van de Casteele E, Nadjmi N.Effect of Cleft Palate Closure Technique on Speech and Middle Ear Outcome: A Systematic Review.J Oral Maxillofac Surg. 2018 Sep 29. pii: S0278-2391(18)31099-1. doi: 10.1016/j.joms.2018.09.027.

8. Zyoud SH.Estimates of global research productivity in using nicotine replacement therapy for tobacco cessation: a bibliometric study.Global Health. 2018 Jan 30;14(1):14. doi: 10.1186/s12992-018-0335-z.

9. Al-Jabi SW.Global Trends in Aspirin Resistance-Related Research from 1990 to 2015: A Bibliometric Analysis.2015: Basic Clin Pharmacol Toxicol. 2017 Dec;121(6):512-519. doi: 10.1111/bcpt.12840. Epub 2017 Jul 25.

10. Rowe N.Tracing the 'grey literature' of poster presentations: a mapping review.Health Info Libr J. 2017 Apr 6:106-124. doi: 10.1111/hir.12177.

11. Shen L, Xiong B, Li W, Lan F, Evans R, Zhang W. Visualizing Collaboration Characteristics and Topic Burst on International Mobile Health Research: Bibliometric Analysis.JMIR Mhealth Uhealth. 2018 Jun 5;6(6):e135.

12. Pritchard A. Statistical bibliography or bibliometrics? J Doc. 1969 Jan;25(4):348–349.

13. Van Leeuwen T, Visser M, Moed H, Nederhof T, Van Raan A. The Holy Grail of science policy: exploring and combining bibliometric tools in search of scientific excellence. Scientometrics. 2003 Jun;57(2):257–280. doi: 10.1023/A:1024141819302.

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

15. Hirsch JE. An index to quantify an individual's scientific research output. Proc Natl Acad Sci USA 2005;102:16569-72.

16. Zhang CT. The h'-Index, Effectively Improving the h-Index Based on the Citation Distribution. PLoS ONE 2013; 8(4): e59912.

17. Zhang CT. The e-index, complementing the h-index for excess citations.The e-index, complementing the h-index for excess citations.PLoS One 2009;4(5):e5429.

18. Dasgupta S, Vaughan AS, Kramer MR, Sanchez TH, Sullivan PS. Use of a Google Map Tool Embedded in an Internet Survey Instrument: Is it a Valid and Reliable Alternative to Geocoded Address Data?JMIR Res Protoc. 2014 Apr 10;3(2):e24.

19. Kobayashi S, Fujioka T, Tanaka Y, Inoue M, Niho Y, Miyoshi A. A geographical information system using the Google Map API for guidance to referral hospitals. J Med Syst. 2010 Dec;34(6):1157-60.

20. Chien TW, Wang HY, Chang Y, Kan WC. Using Google Maps to display the pattern of coauthor collaborations on the topic of schizophrenia: A systematic review between 1937 and 2017.Schizophr Res. 2018 Sep 24. pii: S0920-9964(18)30573-5.

21. Chien TW, Chow JC, Chang Y, Chou W. Applying Gini coefficient to evaluate the author research domains associated with the ordering of author names: A bibliometric study.Medicine: September 2018 - Volume 97 - Issue 39 - p e12418

22. BatageljV, MrvarA. Pajek - Analysis, and Visualization of Large Networks. inJünger, M., Mutzel, P., (Eds.) pp.77-103. Graph Drawing Software, Springer, Berlin, 2003.

23. Efron B. Bootstrap methods: Another look at the jackknife. The Annals of Statistics 1979; 7(1): 1-26.

24. Egghe L. Theory and practice of the g-index. Scientometrics 2006; 69:131-15Lindsey D. Further evidence for adjusting for multiple authorship. Scientometrics. 1982; 4(5): 389–395.

25. Leydesdorff L, Wagner C, Park HW, Adams J.International collaboration in science: the global map and the network.CoRR abs/1301.0801 (2013)

26.Glänzel W, Schlemmer B. National research profiles in a changing Europe (1983–2003) An exploratory study of sectoral characteristics in the Triple Helix. Scientometrics 2007; 70(2), 267-275.

27. Grajales F 3rd1, Sheps S, Ho K, Novak-Lauscher H, Eysenbach G.Social media: a review and tutorial of applications in medicine and health care.J Med Internet Res. 2014 Feb 11;16(2):e13. doi: 10.2196/jmir.2912.

28. Huang MH, Chi PS. A comparative analysis of the application of h-index, g-index, and a-index in institutional-level research evaluation. Journal of Library and Information Studies 2010; 8(2):1-10.

29. Chien, T. W., ChihKan, W., Wang, H. Y., & Chou, W. (2019). Comparisons of citations among clusters of medical subject headings using visualizing topic burst on neuropsychiatry: a bibliometric analysis. *Ann Neuropsychiatry*, *1*, 101.

DEFINITIONS, ACRONYMS, 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