Volume 1 • Issue 1 • 2019

Research Article

Journal of Gynaecology and Paediatric Care

ISSN: 2652-9718

Using Choropleth Maps to Show the Most Dominant Nations and Highly Cited Authors on the Topic of Gynaecology and Paediatric Care from 2013 to 2017 in Pubmed Central

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*

[2]*Nephrology Department, Chi-Mei Medical Center, Tainan, Taiwan*

[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*

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

***Corresponding Author:** Willy Chou, Chi-Mei Medical Center, 901 Chung Hwa Road, Yung Kung Dist., Tainan, Taiwan.

Received: February 02, 2019; **Published:** February 11, 2019

Abstract

Objective: To explore the most dominant nations and highly cited authors on the topic of gynaecology and paediatric care (GPC) from 2013 to 2017 in Pubmed Central (PMC) using choropleth maps.

Method: Authors and their affiliated countries/areas were extracted from the PMC based on the keywords "Gynaecology and paediatric care" in all fields between 2013 and 2017. Citations were based on articles indexed in PubMed Central (PMC) in 2018 and preceding five years. Differences in citations among author clusters were examined using the bootstrapping method. Social network analysis was performed to separate author clusters. A visual dashboard for the most-cited authors was shown on Google Maps.

Results: We observed that (1) the dominant countries with higher x-index were the United States (24.62), the United Kingdom (10.87), and Canada (10.76), (2) the most frequently cited paper (PMID=26799652) was that with 976 citations since 2016 and authored by Daniel J Klionsky with x-inex=24.84 from the US, (3) Differences were observed in bibliometric indices ($p<0.05$) among author clusters.

Conclusions: The dominant nations were determined by the citation indices instead of traditional publications only. The author-weighted scheme (AWS) applied in this study is unique for improving evaluation of individual research achievements (IRA) in a fair and reasonable manner.

Keywords: *Citation; Authorship collaboration; Google Maps; Social network analysis; PubMed Central; x-index*

Abbreviations: AIF: author impact factor; AWS: authorship-weighted scheme; CI: confidence intervals; DC: degree centrality; IF: impact factors; IRA: individual research achievement; PMC: PubMed Central; SNA: Social network analysis; VBA: visual basic for application

Citation: Tsair-Wei Chien, Wei-Chih Kan, Hsien-Yi Wang and Willy Chou. (2019). Using Choropleth Maps to Show the Most Dominant Nations and Highly Cited Authors on the Topic of Gynaecology and Paediatric Care from 2013 to 2017 in Pubmed Central. *Journal of Gynaecology and Paediatric Care* 1(1).

Introduction

Team work in science has been accompanied by a trend in the numbers of authors included in article bylines. Many authors [1-7] applied social network analyses (SNA) for exploring author collaborations in publications. The dominant countries/areas in science articles were found mainly from the United States and Europe [8,9] based on publications also. As of January 30 in 2019, more than 860 articles were searched by the keywords "cited article or paper [title]" in Pubmed Central (PMC). None, but the two [3,4], investigated individual research achievements (IRA) for authors using an appropriate author-weighted scheme (AWS) for quantifying their contributions in article bylines.

Although many types of AWS [9-13] have proposed in the past, we have not aware of any which can be observed and applied for fairly allocating author credits and reporting author IRAs in scientific disciplines. The second essential tool for addressing IRA is the bibliometric index. Despite h-index [14] is a simple and effective index that has been widely used to evaluate academic output of scientists, the h-index suffers from drawbacks, such as assuming all coauthors contributing equally in an article [3,4] and difficult to differentiate the IRAs among authors due to many with identical the integer h-indexes [15].

Numerous metrics have been proposed for use to individuals in the literature, such as x-index [16], g and Ag-index [17], Rh-index [18], e-index [19], and h'-index [20]. However, the mostly preliminary challenge is to be incorporated with the AWS in use. Otherwise, the IRA would be unfair and unreasonable when compared to each other. The dominant nations in scientific fields are also challenged by including publications only instead of involving both citations and publications.

Pediatric gynaecology or pediatric gynecology in the British English spelling is the medical practice dealing with the health of the vagina, vulva, uterus, and ovaries of infants, children, and adolescents. Over 303 articles were retrieved from PMC using the keywords "Pediatric gynaecology [MeSH Major Topic]' on January 30, 2019. Whether the US and Europe are still dominant on the topic of gynaecology and paediatric care is worth exploring this issue in this study.

In June of every year, millions of academic scholars pay close attention to the Journal Citation Reports ranking the journal impact factor (JIF) for the each-indexed journal. No such author IFs (AIFs) [11,12] or bibliometric indices [14, 16-20] can attract the interest of authors as much as JIF does annually in the academia. How to apply an appropriate AWS [3,4] to track the dynamics of individual scientific impact and quantify the co-author contributions in scientific disciplines is worth studying.

In this study, we aim to present (1) the dominant countries based on x-index (see the next section), (2) the most frequently cited papers and authors, and (4) differences in metrics among author clusters using SNA and the bootstrapping method as examination approaches.

Materials and Methods

Data sources

By searching the PubMed database in PMC, we downloaded 1400 abstracts and the author countries/regions from papers published with the topic of gynaecology and paediatric care (GPC) between 2013 and 2017. We applied an author-made Microsoft Excel VBA module to analyze the data-driven contents and present study results. All papers published in PMC based on the type of article were included regarding GPC. All other materials, such as letters to editors, were excluded from this study. Due to all data downloaded from PMC, the study required no ethical approval according to the regulation of the Taiwan Ministry of Health and Welfare.

Four metrics proposed in this study

The h-index can be divided into three parts [19,20] (i..e, h-excess, h-tail, and the h-square area. Many modified h-indexes have been raised by author. The four (e.g., x.Ag, AIF, and h-plus) were used in this study. The h-plus is derived from the h'-index(=h * h-excess/h-tail).Due to the contradiction in rational logic for the h' possibly beyond h and h+1(e.g., h'=5, when h=2 and ratio-h= h-excess/h-tail=3), the h-plus (=h+ratio-h/(1+ratio-h and let t =1 if t<1) was proposed in this study. Accordingly, the h-plus value is always between h and h+1(e.g., h-plus=2.75, when h=2 and ratio-h= h-excess/h-tail=3) and can be complemental to h-index when authors with identical h (called iso-hindex) [19, 20].

AWS for quantifying co-author contributions

We obtain the unique formula $(W_i = \frac{\exp(\gamma_i)}{\sum_{i=0}^{m}\exp(\gamma_i)} = \frac{2.72\hat{}\gamma_i}{\sum_{i=0}^{m}2.72\hat{}\gamma_i}$, (1))

for quantifying coauthor credits in each article. The AIF [11,12] is defined as below:

$$AIF = \frac{\sum Cited.papers.based.on W_j.}{\sum Citable.papers. \times W_j .in.the.given.yrs}), \quad (2)$$

In Eq. (1) giving the first (=exp (m), primary) and last (=exp (m-1) corresponding or supervisory authors with more credits, where m+1 = number of authors of an article. The summation for all co-author weights is equal to 1.0. If γ =0 was assigned to all co-authors, all authors should share equal sizes of contributions to the article, similar to the attached to all co-authors. In contrast, the denominator and numerator in Eq. (1) were replaced with 1/i, where i denotes the ordering of author names in an article. The AWS is called harmonic allocation of authorship credit [17].

Social network analysis using Pajek software

In complying with the Pajek software requirements [22], this study applied SNA [1-7] to cluster authors. Usually, the relation valued by the weight is defined by the number of connections between the two authors. The clusters can be determined by a specific algorithm called degree centrality in Pajek.

Using bootstrapping sampling method to estimate 95% confidence intervals

For comparing differences in metrics among author clusters, we illustrated authors with the highest degree centrality (DC) in their clusters as the representatives. The bootstrapping method [23] was applied to examine differences in metrics among author clusters. A total of 1,000 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. Thus, the difference can be determined by judging the two 95% CI bands separated from each other. .

Creating dashboards on Google Maps

We applied the author-made modules in MS Excel and the SNA in Pajek to separate the author clusters. HTML pages were created for Google Maps. All relevant bibliometric indices were linked to dashboards on Google Maps.

Results and Discussion

Task 1: the dominant countries based on x-index

The dominant countries with higher x-index were the United States (24.62), the United Kingdom (10.87), and Canada (10.76). A simple legend at the right bottom side in Figure 1 shows the proportion of counts in countries/areas around the world, which was rarely reported in traditional choropleth maps. Two cumulative lines of count frequency and total x-index in strata display distinctly different, indicating most countries/areas with fewer research achievements in GPC. It is worth noting that the calculation of x-index for each nation is based on the author x-index in descending order.

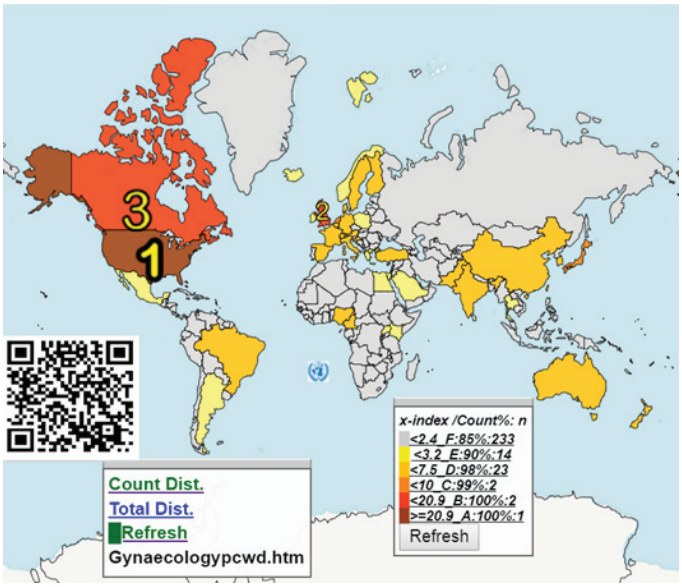


Figure 1: *Using x-index to present the choropleth maps around countries/areas.*

The Gini coefficient [3,26] ranging from 0 to 1.0 was applied to interpret the disparity(=0.90, the higher, the worse) of the counts among strata in Figure 2. Interested readers are suggested to click the QR-code in the respective figure. Animated dashboards on Google Maps were particularly designed for readers who can examine the x-index for each nation/area in Figure 1 when the nation/area is clicked.

Task 2: author collaborations clustered using SNA in comparison with metrics

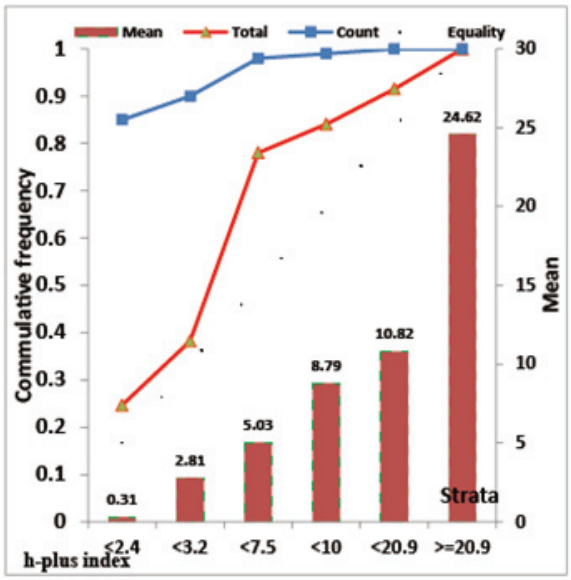
The top ten author clusters are shown in Figure 3, in which we can see the most number of authors are gathered at the left top side with 7167 authors represented by the author van der Ralf J P Valk from Netherland, indicating the closer in relations within a cluster and the less in collaborations between clusters.

Task 3: comparisons of differences in metrics among author clusters

The differences in metrics (i.e., x index, h-plus, Ag, and AIF) were found (p <.05), as shown in Figure 4 when any two 95% CI bands

were separated from each other. The cluster represented by the author Bostjan Leskovar has the lower impact factor among the clusters, as shown in the bottom panel of Figure 4. However, the highest h-plus and x index is the author cluster of "Alan Barrett" (n=14), which indicates relatively higher metrics in the median.

One article with 2467 coauthors has been cited 976 times in PMC.



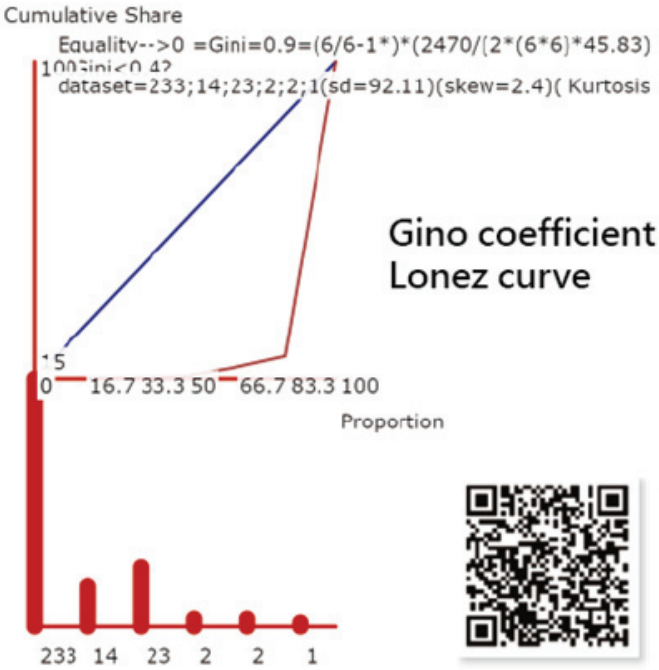


Figure 2: *Legends to complement the choropleth map.*

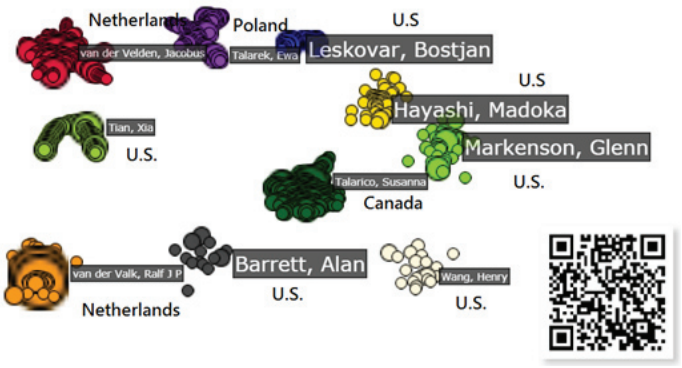


Figure 3: *The top ten author clusters dispersed on a dashboard.*

Task 4: the most frequently cited papers and authors

The most frequently cited paper (PMID=26799652) [27] was that with 976 citations since 2016 and authored by Daniel J Klionsky with x-inex=24.84 from the US, at the right top side in Figure 5. If the bubble was clicked, a series of metric appear in a box showing that the weighted citable=0.63 with a single article as the first author, the weighted citation=616.95= 976* 0.63, AIF=617, Ag=616.95, h-index=1, g-index=1, x-index=24.84=√1*617, h-plus=1.96, ratio-h=24.82. Other authors can be examined by clicking the bubble of interest.

Discussions

According to Hirsch [14], the impact factor (i.e., citations/publications, in Eq. (2)) usually rewards low productivity and penalizes high productivity, as shown in the button size Figure 5. The IRA cannot be measured by either publications or citations alone. The combined effects using bibliometrics mentioned above have certain disadvantages and limitations, such as assuming all coauthor contributions equal and allocating more weights on either outputs or citations. If h-index was applied, we suggest using the h-plus first, followed by the x index, h-plus index, or impact factor, as shown in Figures 5, in accordance to the preference of a research institute.

We have illustrated the use of AWS onto quantifying coauthor contributions in an article. The suggested AWS in Eq. (1) [3,4] that implies giving more importance to the first (=exp (m), primary) and the last (=exp(m-1) corresponding or supervisory) authors. The others (middle authors) have thus made smaller contributions to the articles. As such, Vavryčuk's combined weighted scheme [10] (or harmonic credits [1]) is a special case of AWS we mentioned in Eq. (1). For example, replacing the denominator and the numerator with 1/i that can be a harmonic type of AWS, where i represents the ordering of author names.

We are surprised at findings on the GPC topic in PMC. Three articles [27-29] were named with over 2000 coauthors, 2467, 4107, and 4107, respectively. Except for the most highly cited paper (PMID=26799652) [27] mentioned above, the other two with PMID= 25673413 and 24097068 [30,31] were cited 520, and 611 times, respectively. Interested readers are recommended to read those articles with more citations in references.

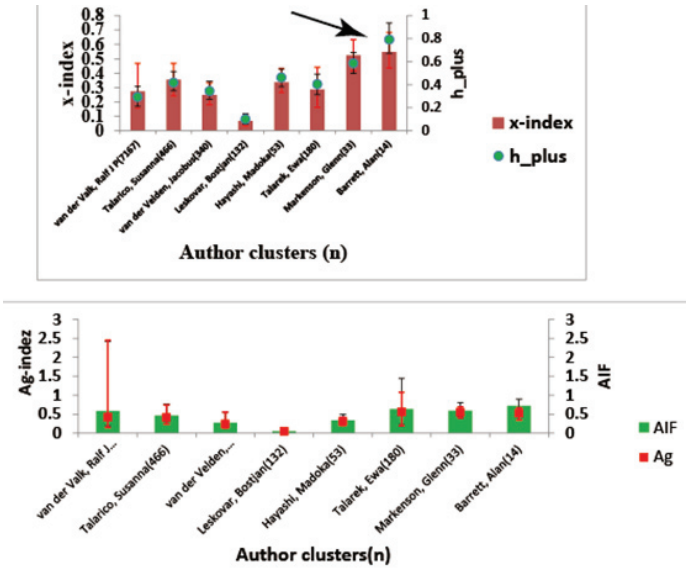


Figure 4: *Comparison of metrics among author clusters.*

Limitations

Although our findings have been illustrated above, several potential limitations should be overcome in the future. First, all data were downloaded from the PMC, which cannot generalize the results to other bibliometric databases and disciplines.

Second, biases might occur when matching authors' names to calculate the IRA because, in some cases, different authors have the same name but with disparate author identity. Therefore, the result of author relationship analysis using SNA might be influenced by inaccuracy as a result of false author classification in this study.

Third, many SNA algorithms were applied by users. The degree centrality used in generating the partitioned clusters might vary in different algorithms were applied. Fourth, the formula of quantifying co-author contributions used in this study (e.g.,Eq. (1) assumed that all authors made different contributions to an article and the first and the corresponding with the most parts. Any change for the rule in author contributions will affect the results of the metrics we computed in this study.

Fifth, the data were extracted from PMC, which is different from other studies applying those common citation databases, such as the Scientific Citation Index (Thomson Reuters, US) and Scopus (Elsevier, Netherlands) or even the Google Scholar. The results for the most frequently cited authors and countries might vary if other databases were used.

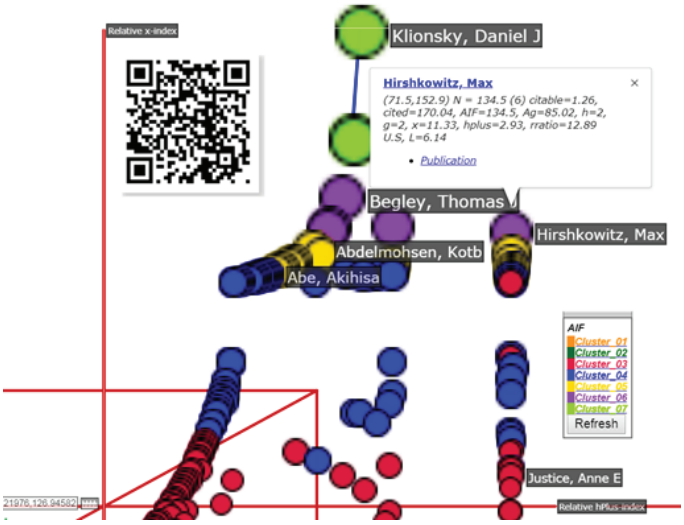


Figure 5: *The most cited authors on a dashboard to display.*

Conclusion

The dominant nations were determined by the bibliometric indices instead of those traditional publications applied only. The author-weighted scheme (AWS) applied in this study is unique for improving evaluation of individual research achievements (IRA) in a fair and reasonable manner. Many other topics besides the given GPC should be further investigated applying the AWS to characterize the features and patterns onto other disciplines in the future.

Acknowledgments

This study was supported by the grant of Chi Mei Medical Center, Taiwan.

Conflict of interest

No any financial interest or any author conflict of interest exists.

References

1. Lu Y, Figler B, Huang H, Tu YC, Wang J, Cheng F. (2017). Characterization of the mechanism of drug-drug interactions from PubMed using MeSH terms. PLoS One. Apr 19; 12 (4):e0173548.

2. Chien TW, Chang Y, Wang HY. (2018). Understanding the productive author who published papers in medicine using National Health Insurance Database: A systematic review and meta-analysis. Medicine (Baltimore). 97(8): e9967.
3. 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.
4. Chien TW, Wang HY, Chang Y, Kan WC. (2018). Using Google Maps to display the pattern of coauthor collaborations on the topic of schizophrenia: A systematic review between 1937 and 2017.Schizophr Res. pii: S0920-9964(18)30573-5.
5. Sadoughi F, Valinejadi A, Shirazi MS, Khademi R. (2016). Social Network Analysis of Iranian Researchers on Medical Parasitology: A 41 Year Co-Authorship Survey. Iran J Parasitol. 11 (2):204-212.
6. Osareh F, Khademi R, Rostami MK, Shirazi MS. (2014). Co-authorship Network Structure Analysis of Iranian Researchers' scientific outputs from 1991 to 2013 based on the Social Science Citation Index (SSCI). Collnet J Scientometr Info Manag. 8 (2): 263–71.
7. Liu X, Bollen J, Nelson ML, Van de Sompel H. (2005). Co-authorship networks in the digital library research community. Info Process Manag. 41 (6): 1462–80.Leydesdorff L, Wagner C, Park HW, Adams J. International collaboration in science: the global map and the network. CoRR abs/1301.0801 (2013).
8. Glänzel W, Schlemmer B. (2007). National research profiles in a changing Europe (1983–2003). An exploratory study of sectoral characteristics in the Triple Helix. Scientometrics 70(2), 267-275.
9. Vavryčuk V. (2018). Fair ranking of researchers and research teams. PLoS One. 5; 13 (4):e0195509.
10. Sekercioglu CH. (2008). Quantifying coauthor contributions. Science 322 (5900): 371.
11. Lippi G G, Mattiuzzi C. (2017). Scientist impact factor (SIF): a new metric for improving scientists' evaluation? Ann Transl Med. 5(15): 303.
12. Pan RK, Fortunato S. (2014). Author Impact Factor: tracking the dynamics of individual scientific impact. Sci Rep. 4: 4880.
13. Hagen NT. (2008). Harmonic allocation of authorship credit: source-level correction of bibliometric bias assures accurate publication and citation analysis. PLoS One. 3 (12):e4021.
14. Hirsch JE. (2005). An index to quantify an individual's scientific research output. Proc Natl Acad Sci USA 102:16569-72.
15. Huang MH, Chi PS. (2010). 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. 8(2):1-10.
16. Fenner T, Harris M, Levene M, Bar-Ilan J. (2018). A novel bibliometric index with a simple geometric interpretation. PLoS One. 13 (7): e0200098.
17. Egghe L. (2006). Theory and practice of the g-index. Scientometrics 69:131-15.
18. Jin BH, Liang LM, Rousseau R, Egghe L. (2007). The R- and AR-indices: Complementing the h-index. Chinese Science Bulletin 52: 855–863.
19. Zhang CT. (2009). The e-index, complementing the h-index for excess citations. The e-index, complementing the h-index for excess citations. PLoS One 4 (5):e5429.
20. Zhang CT. (2013). The h'-Index, Effectively Improving the h-Index Based on the Citation Distribution. PLoS ONE 8(4): e59912.
21. American and British English spelling differences. Gynecology is the American spelling, but it is also common in international contexts, e.g., International Federation of Gynecology and Obstetrics and International Society of Ultrasound in Obstetrics and Gynecology.
22. BatageljV, MrvarA. (2003). Pajek - Analysis, and Visualization of Large Networks. In Jünger, M., Mutzel, P., (Eds.) pp.77-103. Graph Drawing Software, Springer, Berlin.
23. Efron B. (1979). Bootstrap methods: Another look at the jackknife. The Annals of Statistics 7(1): 1-26.
24. Cromley RG, Cromley EK. (2009). Choropleth map legend design for visualizing community health disparities. Int J Health Geogr. Sep 24; 8:52. doi: 10.1186/1476-072X-8-52.
25. Cromley RG, Ye Y. (2006). Ogive-based legends for choropleth mapping. Cartogr Geogr Inform Sci. 33:257–268. doi: 10.1559/152304006779500650.
26. Gini C. Concentration and dependency ratios (in Italian). English translation in Rivista di Politica Economica 1909; 87 (1997): 769–789.
27. Klionsky DJ, Abdelmohsen K, Abe A, Abedin MJ, Abeliovich H, Acevedo Arozena A, etc. (2016). Guidelines for the use and interpretation of assays for monitoring autophagy (3rd edition). Autophagy. 12 (1):1-222. doi: 10.1080/15548627.2015.1100356.

28. Veitenhansl M, Stegner K, Hierl FX, Dieterle C, Feldmeier H, Gutt B,etc. (2004). 40th EASD Annual Meeting of the European Association for the Study of Diabetes: Munich, Germany, 5-9 September 2004.Diabetologia. 47 (Suppl 1):A1-A464.
29. Melander A, Olsson J, Lindberg G, Salzman A, Howard T, Stang P,etc. (1999). 35th Annual Meeting of the European Association for the Study of Diabetes: Brussels, Belgium, 28 September-2 October 1999. Diabetologia. 42 (Suppl 1):A1-A330.
30. Locke AE, Kahali B, Berndt SI, Justice AE, Pers TH, Day FR, etc. (2015). Genetic studies of body mass index yield new insights for obesity biology. Nature. Feb 12; 518 (7538):197-206. doi: 10.1038/nature14177.
31. Willer CJ, Schmidt EM, Sengupta S, Peloso G, Gustafsson S, Kanoni S, etc. (2013). Discovery and refinement of loci associated with lipid levels. Nat Genet. 45 (11):1274-1283. doi: 10.1038/ng.2797. Epub 2013 Oct 6.