

Predicting Article Citations Using Data of 100 Top-Cited Publications in the Journal Anesthesiology since 2012: A Bibliometric Analysis

Shu-Chen Hsing¹, Tsair-Wei Chien², Willy Chou^{3,4*}

¹Repartment of Respiratory Therapy, Chi-Mei Medical Center, Tainan, Taiwan

²Department of Medical Research, Chi Mei Medical Center, Tainan, Taiwan

³Department of Physical Medicine and Rehabilitation, Chung Shan Medical University, Taichung

⁴Department of Physical Medicine and Rehabilitation, Chiali Chi Mei Hospital, Tainan, Taiwan

***Corresponding Author:** Willy Chou, Department of Physical Medicine and Rehabilitation, Chung Shan Medical University, Taichung, Chiali Chi Mei Hospital, Tainan, Taiwan.

Abstract

Background: Publications regarding the 100 top-cited articles in a given discipline are common, but the association between article topics and their citations are lacking in reports. Whether or not article topics can be used for predicting citations have not been addressed before. This study aims to (1) provide a Sankey diagram for better understanding the association between entities in the study 100 top-cited articles and (2) inspect major medical subject headings (i.e., MeSH terms in PubMed) to help predict article citations.

Methods: We searched PubMed Central and downloaded 100 top-cited abstracts in the journal Anesthesiology since 2012. Article types and MeSH terms were extracted from abstracts. Social network analysis and Sankey diagram analysis were performed to identify article contribution to citations. MeSH terms were then applied to predict the number of article citations. We then examined the prediction power with the correlation coefficients (CCs) between MeSH weights and article citations.

Results: The citation counts for the 100 articles ranged from 37 to 353, with an average of 63.1 citations. The most frequent article types were research support(41%) and comparative study(26%), and the most frequent MeSH terms were methods (33%) and drug effects (29%). The most productive countries were the United States (53%) and Canada (9%). The most cited article (PDID=23364566) with a count of 353 was written by Dr. Apfelbaum from the American Society of Anesthesiologists (US) in 2013. MeSH terms were evident in the prediction power of the number of article citations (CC=0.33, $t=3.44$).

Conclusion: The breakthrough was made by using a Sankey diagram showing the overall views of the 100 top-cited articles. MeSH terms can be used for predicting article citations. Analyzing the 100 top-cited articles could help future academic pursuits and applications in other academic disciplines, not just limited to the journal Anesthesiology.

Key words: bibliometric, citation analysis, social network analysis, medical subject heading, dashboard

1. INTRODUCTION

Anesthesiology is the medical specialty of anesthesia or anesthetics [1]. The practice of anesthesiology is dedicated to the relief of pain and total care of the surgical patient before, during, and after surgery [2]. Anesthesia is thus defined as the medical way to the loss of sensation with or without the loss of consciousness [3].

Many articles regarding anesthesiology have been published in literature. Which articles were

cited most in the past remains unknown. Over 227 articles with the keywords of 100 and cited in the title were searched in Pubmed Central as of October 11, 2020[4]. None were found in the field of anesthesiology. We are thus motivated to conduct the research on the top 100-cited articles in anesthesiology to understand the most dominant countries, frequently proposed topics, and article types in the recent ten years.

Bibliometric research on scholarly journals is frequently focused on research domains (RDs)

and research achievements (RAs) [5-7]. The former can be addressed by using social network analysis (SNA) and medical subject headings (MeSH terms) [8-10]. The latter was frequently evaluated by metrics (e.g., impact factor (IF), h-index [11], or x-index [12]). Whether article types denoted by MeSH terms can be used for prediction article citations is required for verifications.

Most articles regarding the top 100-cited publications addressed such topics as descriptive statistics (DS), RA across countries/institutes over the years, and RD (on article types or for individual authors) [13-15]. Some addressed RAs and RDs for journals through bibliometric analyses [16, 17]. However, all aforementioned articles applied the same simple method and provided information limited to DS, RA, and RD, which is a challenge that should be made through with the citation prediction of article types along with a specific visual representation to the topic of 100 top-cited articles. It is because two aspects in bibliometric analyses are often ignored and criticized: (1) the inability to visualize results through a dashboard highlighting relevant entities [18-20] and (2) the lack of a model for predicting the number of article citations for future studies [21-24].

The objectives of this study were to analyze the 100 top-cited articles from a journal through a systematic search and apply a novel approach involving (1) a picture highlighting the most outstanding entities and (2) a mode capable of predicting the number of citations in the future.

2. METHODS

2.1. Data Source

Two steps were carried out to conduct this study. First, we searched the PubMed database (Pubmed.com, PMC for short) using the keyword (Anesthesiology [Journal]) and downloaded 5427 abstracts since 2012 on October 11, 2020. Second, we selected the 100 top-cited articles. Several figures and tables were produced to illustrate (1) the major topics of Anesthesiology, (2) the main contributors associated with their origin countries/regions, and (3) the prediction power based on article topic denoted by MeSH terms for the number of citations.

Ethical approval was not necessary for this study because all data were obtained from the database publicly available on PMC.

2.2. Data Arrangement for DS and RA

The 100 top-cited articles were categorized into different types (e.g., case reports, clinical trials, comparative study, and journal articles in PMC). Topic categories referring to the MeSH terms [25] in each article were classified by using SNA [26-28].

A contingency table was made to show the main contributors from countries /regions. A choropleth map was produced to highlight the most influential countries/regions of origin for authors based on the number of publications.

2.3. Data Visualization for RD

SNA [26-28] was applied to associate the 100 top-cited articles with topic categories and article types on a visual board to achieve the first goal of this study, highlighting the most outstanding entities.

Pajek software [29] was applied to perform SNA. Relevant entities related to article types provide insights into the overall views when presented on a visual board, which was hard to display using the traditional word cloud technique [30]. A Sankey diagram was particularly used to interpret the association of the most dominant entities in the network.

2.4. Prediction Power on Article Citations

To achieve the second objective of this study, we computed the weights for the number of citations on MeSH terms per article according to the proportions and citations, taking the citations into accounts in SNA [31]. A comparison of MeSH terms in clusters was made using the bootstrapping method [32].

2.5. Statistics

The correlation coefficient (CC) was used to determine the prediction power between the weighted MeSH terms and the original article citations. The CC t-value was calculated using the formula $(=CC * \sqrt{(n-2)/(1-CC*CC)})$. A prediction equation was produced by running a simple regression analysis using MedCalc 9.5.0.0 for Windows (MedCalc Software). The significant level was set at Type I error of 0.05.

3. RESULTS

3.1. All 100 Articles Linked on the Website

The results of the 100 top-cited articles published in anesthesiology since 2012 were included in Reference [33], where readers were

invited to inspect the association among entities in detail.

The citation counts for the 100 articles ranged from 37 to 353 as of October 11, 2020, in PMC, with an average of 63.1 citations. The most

productive countries of origin were the United States (54%) and Canada (9%). The highest number of citations per article was from Canada (81.3) (Table 1). The visual representation using a choropleth map is provided in Figure 1.

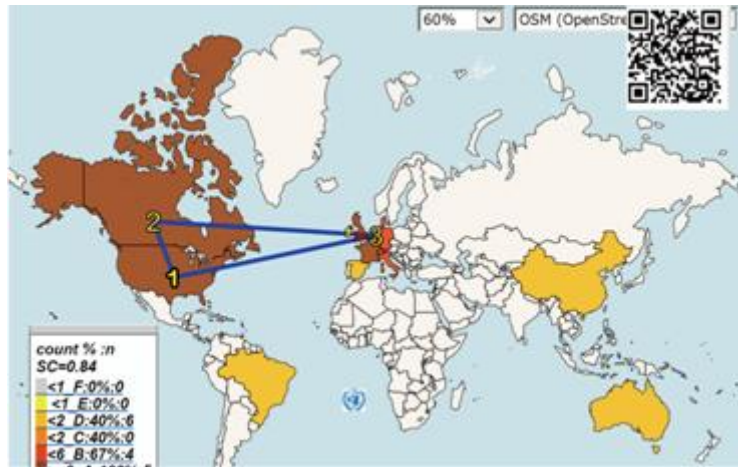


Figure 1: Distribution of these top 100 articles across countries/regions

Table 1: Distribution of publications across countries over the years

Country	2012	2013	2014	2015	2016	n	Ci	IF
U.S	16	16	12	9	1	54	3243	60.06
Canada	4	1	2	1	1	9	732	81.33
Netherlands	4	1	1	1	1	8	577	72.13
France	4	1	1			6	281	46.83
U.K	2	2		1	1	6	332	55.33
Germany	1	3		1		5	306	61.20
Denmark	1	1				2	85	42.50
Italy		1			1	2	168	84.00
Switzerland	2					2	149	74.50
Australia		1				1	48	48.00
Belgium					1	1	51	51.00
Brazil				1		1	46	46.00
China	1					1	42	42.00
Ireland		1				1	40	40.00
Spain			1			1	57	57.00
Total	37	28	17	14	6	100	6311	63.11

Note. $IF=Ci/n$

3.2. Visual Representations Using SNA and the Sankey Diagram

The majority of articles were research support (41%) and comparative study (26%). Most articles were published in the topic categories of methods (33%) and drug effects (29%). The top three MeSH terms of methods, drug effects, and pharmacology were highlighted with triangle lines connected together in Figure 2.

The most cited article (PDID=23364566) with a count of 353 was written by Dr. Apfelbaum from the American Society of Anesthesiologists (US) in 2013 (Figure 3).

For highlighting the association among entities in a picture, the Sankey diagram was drawn in

Figure 3. Only the top dominant entities with the closest relationships were displayed and connected by the curve lines from the left side to the right side. Other weaker cited lines between entities were removed from the diagram. For instance, the year 2013 is merely connected with the highly cited article with PMID=23364566 (=353), which is sequentially linked to the next entity of the US. Bubbles were colored by the clusters and sized by the cited weights. As such, the top three (i.e., the US, research support, and methods) connected by a triangle at the middle-top side are highlighted in the Sankey diagram. Readers are invited to scan the QR-code in Figures to examine the details about the information on entities.

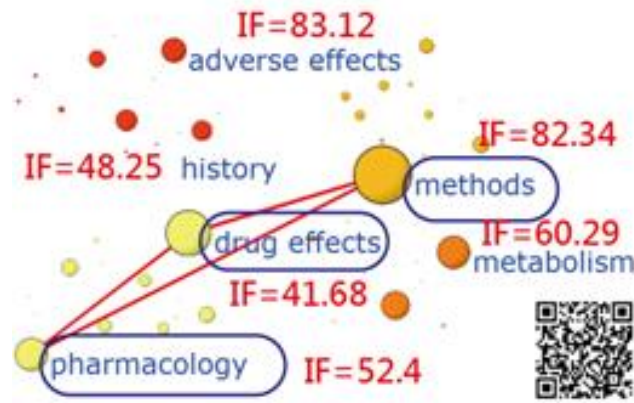


Figure2: The most dominant MeSH terms with high citations and counts in top 100 articles

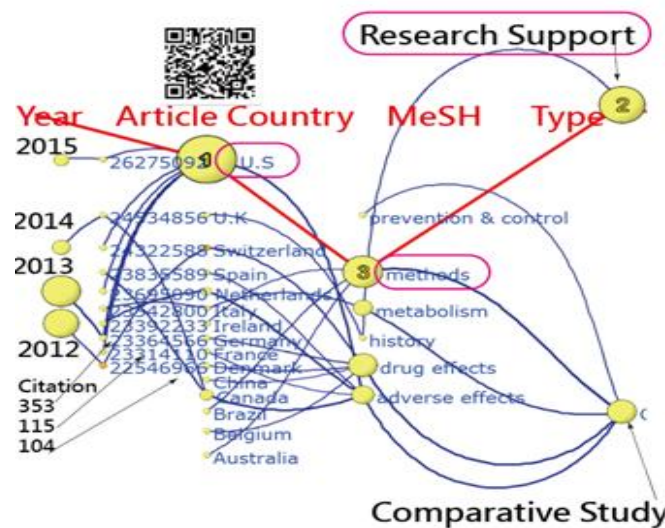


Figure3: Association of entities using the Sankey diagram to display

3.3. Citation Prediction Using Mesh terms

A comparison of MeSH terms in clusters was made using the bootstrapping method [32]. We

can see the highest citations of MeSH terms clusters in Figure 2 are represented by methods and adverse effects

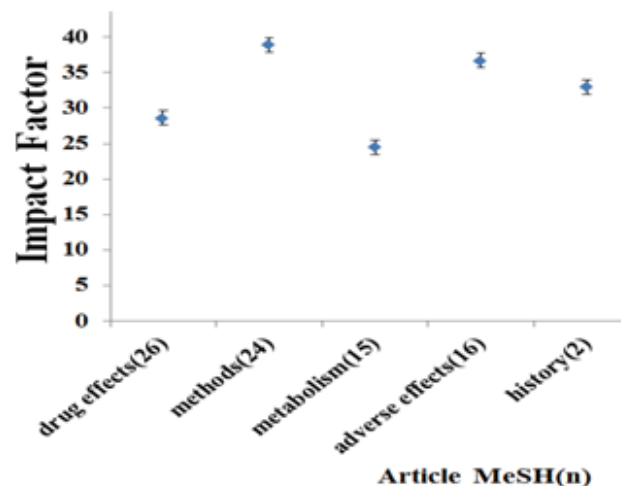


Figure4: Comparison of impact factors among MeSH clusters

MeSH terms were evident in generating prediction power on the number of citations (correlation coefficient = 0.49, $t = 5.62$) in Figure 5. The regression equation is defined as article citation(y) $y = 19.6879 + 0.7514 x$ weight(x) of MeSH term. The slop coefficient appeared statistically significant ($F = 11.83626$, $p = 0.0009$).

The QR-codes are provided with links in Figures. Readers are invited to see the detailed information on the dashboard laid on Google Maps.

4. DISCUSSION

In this study, we downloaded articles from PMC and assessed the most cited papers published in *Anesthesiology* since 2012 to understand the research characteristics of citations in the past. This paper provides an insight into some of the most influential entities, articles (Figure 3), and publication-changing chronicles over the years (Table 1).

The topic categories related to article topics were found using the SNA approach and the Sankey diagram. Nevertheless, the topic category of adverse effects earned the highest mean number of weighted scores ($=83.12$) denoted by MeSH terms (Figure 2). The highest

impact factor is in the cluster of methods (Figure 4). The high mean IF in past studies in the field of methods and adverse effects can be attributed to its important nature in Anesthesiology.

4.1. Features of this Study

We provided a guideline to analyze the 100 top-cited articles in three steps (i.e., DS, RA, and RD). This breakthrough was made using tables and figures (e.g., Table 1 for DS, Figures 1, 2, and 3 for RA, and Figures 4–5 for RD). In particular, visualization dashboards were used to present the results, and a link to the website allowed readers to examine the detailed information on their own. The board equipped with multi-variables (i.e., entities in the Sankey diagram) is superior to the traditional one-variable word cloud technique [30].

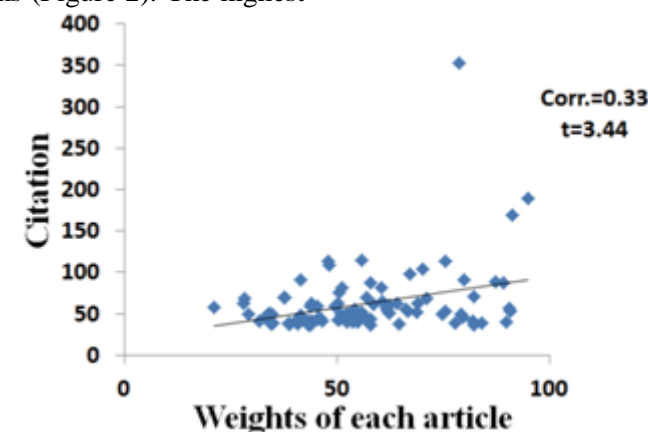


Figure5: the MeSH terms to predict article citations ($P<0.0001$)

Furthermore, both SNA and MeSH terms were sophisticatedly incorporated for interpreting the RA in a succinct representation, as we did in Figures 2 and 3. Visualizing the characteristics of data is important and interesting for readers.

4.2. Limitations and Suggestions

Although findings are promising based on the above analysis, several potential limitations may encourage further research efforts. First, the study data were only downloaded from PMC, different from many other studies that used the combination of science (SCI), Scopus, and Google Scholar. Several journals in PubMed have not been indexed by the database of SCI. In addition, certain SCI-indexed journals have not been included in PubMed. These findings might contribute to a potential bias in the results of publications and citations obtained in this study.

Second, another bias might exist during citation extractions because the number of citations possibly increased as the date elapsed. The article citations might differ if the time available and the citation sources of the data are disparate.

Third, only articles with MeSH terms were included in this study. Others without MeSH terms were excluded from this study despite a potentially large number of citations present.

Fourth, although our cluster analysis using SNA is unique and useful, including all relevant entities on a board without confusing the readers is challenging. The results could be improved further for clarity and understandability.

Fifth, only a single journal was investigated for its RA, RD, and DS in the current study. Although several journals on a specific topic or discipline with their 100-top articles were frequently read in the literature, we performed only the visual representations and citation

prediction approaches for the 100 top-cited documents in a single anesthesiology-related journal. The research methods, particularly with dashboards laid on Google Maps, can be provided to future similar studies.

Finally, article citations have many associated factors, such as the number of references and authors. The finding of MeSH terms associated with article citations is not the only factor influencing article citations. Thus, it is worth studying further about the prediction factors to the highly cited articles in the future.

5. CONCLUSION

The study identified the 100 top-cited articles in the journal of *Anesthesiology* and provided insights for readers. With the novel method of SNA and the Sankey diagram, the bibliometric analysis of 100 top-cited articles can be applied to other academic disciplines in the future, not just limited to the journal *Anesthesiology*.

REFERENCES

- [1] Wang HY, Chien TW, Chow JC, Chou W. Applying Clustering Coefficient to the Pattern of International Author Collaboration in Anesthesiology. *ARC Journal of Anesthesiology* 2017; 2(3), 18-24.
- [2] American Society of Anesthesiologists. What is Anesthesiology? 2017/12/20 available at <https://goo.gl/nAxpXd>
- [3] Chien TW, Kan WC, Wang HY, Chou W. Visualizing Topic Burst and Citation Comparisons among Clusters of Medical Subject Headings on Anesthesiology: Bibliometric Analyses. *ARC Journal of Anesthesiology*. 2018; 3(4):18-24.
- [4] Chien TW. The top 100-cited articles in Pubmed Central. 2020/10/11 available at <https://pubmed.ncbi.nlm.nih.gov/?term=100%5Btitle%5D+and+cited%5Btitle%5D&sort=pubdate>
- [5] Chien TW, Wang HY, Kan WC, Su SB. Whether article types of a scholarly journal are different in cited metrics using cluster analysis of MeSH terms to display: A bibliometric analysis. *Medicine (Baltimore)*. 2019; 98(43):e17631.
- [6] Chien TW, Chang Y, Wang HY. Understanding the productive author who published papers in medicine using the National Health Insurance Database: A systematic review and meta-analysis. *Medicine (Baltimore)*. 2018; 97(8):e9967. doi:10.1097/MD.0000000000009967
- [7] Hsieh WT, Chien TW, Kuo SC, Lin HJ. Whether productive authors using the national health insurance database also achieve higher individual research metrics: A bibliometric study. *Medicine (Baltimore)*. 2020; 99(2):e18631. doi:10.1097/MD.0000000000018631
- [8] DeFosset AR, Mosst JT, Gase LN, Kuo T. The Los Angeles Diabetes Prevention Coalition Experience: Practical Applications of Social Network Analysis to Inform Coalition Building in Chronic Disease Prevention. *J Public Health ManagPract*. 2020; 26(3):270–279. doi:10.1097/PHH.0000000000000958
- [9] Govoeiy B, Agbokounou AM, Camara Y, et al. Social network analysis of practice adoption facing outbreaks of African Swine Fever [published online ahead of print, 2020 Apr 20]. *Prev Vet Med*. 2020; 179:105008. doi:10.1016/j.prevetmed.2020.105008
- [10] Prado AM, Pearson AA, Bertelsen NS, Pagán JA. Connecting healthcare professionals in Central America through management and leadership development: a social network analysis. *Global Health*. 2020; 16(1):34. Published 2020 Apr 15. doi:10.1186/s12992-020-00557-4
- [11] Hirsch JE. An index to quantify an individual's scientific research output. *Proc Natl AcadSci U S A* 2005 Nov 15;102(46):16569-16572
- [12] Fenner T, Harris M, Levene M, Bar-Ilan J. A novel bibliometric index with a simple geometric interpretation. *PLoS ONE* 2018 Jul 10; 13(7):e0200098
- [13] Chien TW, Wang HY, Kan WC, Su SB. Whether article types of a scholarly journal are different in cited metrics using cluster analysis of MeSH terms to display: A bibliometric analysis. *Medicine (Baltimore)*. 2019; 98(43):e17631.
- [14] Chien TW, Chang Y, Wang HY. Understanding the productive author who published papers in medicine using the National Health Insurance Database: A systematic review and meta-analysis. *Medicine (Baltimore)*. 2018; 97(8):e9967. doi:10.1097/MD.0000000000009967
- [15] Hsieh WT, Chien TW, Kuo SC, Lin HJ. Whether productive authors using the national health insurance database also achieve higher individual research metrics: A bibliometric study. *Medicine (Baltimore)*. 2020; 99(2):e18631. doi:10.1097/MD.0000000000018631
- [16] Yin X, Cheng F, Wang X, et al. Top 100 cited articles on rheumatoid arthritis: A bibliometric analysis. *Medicine (Baltimore)*. 2019; 98(8):e14523.
- [17] Jiang Y, Hu R, Zhu G. Top 100 cited articles on infection in orthopaedics: A bibliometric analysis. *Medicine (Baltimore)*. 2019; 98(2):e14067.
- [18] 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 MhealthUhealth. 2018; 6(6):e135. Published 2018 Jun 5. doi:10.2196/mhealth.9581
- [19] Kan WC, Chou W, Chien TW, Yeh YT, Chou PH. The Most-Cited Authors Who Published Papers in JMIR mHealth and uHealth Using the Authorship-Weighted Scheme: Bibliometric Analysis. JMIR MhealthUhealth. 2020; 8(5): e11567.
- [20] Chien TW, Chang Y, Wang HY. Understanding the productive author who published papers in medicine using National Health Insurance Database: A systematic review and meta-analysis. *Medicine (Baltimore)*. 2018; 97(8): e9967.
- [21] Butun E, Kaya M. Predicting Citation Count of Scientists as a Link Prediction Problem [published online ahead of print, 2019 Mar 12]. *IEEE Trans Cybern*. 2019;10.1109/TCYB.2019.2900495.
- [22] Park I, Yoon B. Technological opportunity discovery for technological convergence based on the prediction of technology knowledge flow in a citation network. *J Informetr*. 2018; 12(4):1199-1222.
- [23] Lin HF, Wu XF, Zhang YH. SCI citation analysis and impact factor prediction of JZUS-B in 2008. *J Zhejiang UnivSci B*. 2009; 10(1): 77-78. doi:10.1631/jzus.B0840002
- [24] Ahmad P, Elgamel HAM. Citation Classics in the Journal of Endodontics and a Comparative Bibliometric Analysis with the Most Downloaded Articles in 2017 and 2018. *J Endod*. 2020; 46(8):1042-1051.
- [25] Eisinger D, Tsatsaronis G, Bundschuh M, Wieneke U, Schroeder M. Automated Patent Categorization and Guided Patent Search using IPC as Inspired by MeSH and PubMed. *J Biomed Semantics*. 2013; 4Suppl 1(Suppl1): S3. doi:10.1186/2041-1480-4-S1-S3
- [26] DeFosset AR, Mosst JT, Gase LN, Kuo T. The Los Angeles Diabetes Prevention Coalition Experience: Practical Applications of Social Network Analysis to Inform Coalition Building in Chronic Disease Prevention. *J Public Health ManagPract*. 2020; 26(3):270-279.
- [27] Govoeyi B, Agbokounou AM, Camara Y, et al. Social network analysis of practice adoption facing outbreaks of African Swine Fever [published online ahead of print, 2020 Apr 20]. *Prev Vet Med*. 2020; 179:105008. doi:10.1016/j.prevetmed.2020.105008
- [28] Prado AM, Pearson AA, Bertelsen NS, Pagán JA. Connecting healthcare professionals in Central America through management and leadership development: a social network analysis. *Global Health*. 2020; 16(1):34. Published 2020 Apr 15.
- [29] deNooy W, Mrvar A, Batagelj V. *Exploratory Social Network Analysis With Pajek: Revised and Expanded*, 2nd edn. New York, NY: Cambridge University Press, 2011.
- [30] Atenstaedt R. Word cloud analysis of the BJGP: 5 years on. *Br J Gen Pract*. 2017; 67(658):231-232. doi:10.3399/bjgp17X690833
- [31] Kuo YC, Chien TW, Kuo SC, Yeh YT, Lin JCJ, Fong Y. Predicting article citations using data of 100 top-cited publications in the journal *Medicine* since 2011: A bibliometric analysis. *Medicine (Baltimore)*. 2020; 102(10):e9967.
- [32] Kung SC, Chien TW, Yeh YT, Lin JJ, Chou W. Using the bootstrapping method to verify whether hospital physicians have different h-indexes regarding individual research achievement: A bibliometric analysis. *Medicine (Baltimore)*. 2020 Aug 14;99(33):e21552. doi: 10.1097/MD.00000000000021552. PMID: 32872003; PMCID: PMC7437822.
- [33] Chien TW. The most cited article in journal *Medicine* since 2011. 2020/10/12 available at <http://www.healthup.org.tw/html100/annes2020100.htm>
- [34] Apfelbaum JL, Hagberg CA, Caplan RA, Blitt CD, Connis RT, Nickinovich DG, Hagberg CA, Caplan RA, Benumof JL, Berry FA, Blitt CD, Bode RH, Cheney FW, Connis RT, Guidry OF, Nickinovich DG, Ovassapian A; American Society of Anesthesiologists Task Force on Management of the Difficult Airway. Practice guidelines for management of the difficult airway: an updated report by the American Society of Anesthesiologists Task Force on Management of the Difficult Airway. *Anesthesiology*. 2013 Feb; 118(2):251-70. doi: 10.1097/ALN.0b013e31827773b2. PMID: 23364566.

Citation: Shu-Chen Hsing, Tsair-Wei Chien, Willy Chou, Predicting Article Citations Using Data of 100 Top-Cited Publications in the Journal Anesthesiology since 2012: A Bibliometric Analysis. *ARC Journal of Anesthesiology*. 2020; 5(1): 19-25. DOI: [dx.doi.org/10.20431/2455-9792.0501005](https://doi.org/10.20431/2455-9792.0501005).

Copyright: © 2020 Authors. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.