Systematic Review and Meta-Analysis

Medicine

OPEN

The 100 most cited articles have fewer citations than other bibliometric articles

A pairwise comparison using a temporal bubble graph

Wan-Ting Hsieh, MD[a], Tsair-Wei Chien, MBA[b], Willy Chou, MD[c,d,*]

Abstract

Background: More than 400 articles with the title of 100 top-cited articles (Top100) have been published in PubMed. It is unknown whether their citations are fewer (or more) than those found in other bibliometric studies (Nontop100). After determining article themes using coword analysis, a temporal bubble graph (TBG) was used to verify the hypothesis that the Top100 had fewer citations than the Nontop100.

Methods: Using the Web of Science core collection, the top 50 most cited articles were compiled by Top100 and Nontop100, respectively, based on the research area of biomedicine and bibliometrics only. Coword analysis was used to extract themes. The study results were displayed using 6 different visualizations, including charts with bars, pyramids, forests, clusters, chords, and bubbles. Mean citations were compared between Top100 and Nontop100 using the bootstrapping method.

Results: There were 18 citations in total for the 2 sets of the 50 most cited articles (range 1–134; 5 and 26.5 for Top100 and Nontop100, respectively). A significant difference in mean citations was observed between the 2 groups of Top100 and Nontop100 based on the bootstrapping method (3, 95% confidence interval: [1.18, 4.82]; 26.5, 95% confidence interval: [23.82, 29.18], $P < .001$). The 11 themes were clustered using coword analysis and applied to a TBG, which is composed of 4 dimensions: themes, years, citations and groups of articles. Among the 2 groups, the majority of articles were published in the journal of *Medicine (Baltimore)*, with 9 and 7, respectively.

Conclusion: Eleven themes were identified as a result of this study. In addition, it reveals distinct differences between the 2 groups of Top100 and Nontop100, with the former containing more recently published articles and the latter containing more citations for articles. Clinical and research clinicians and researchers can use bibliometric analysis to appraise published literature and to understand the scientific landmark using TBG in bibliometrics.

Abbreviations: CI = confidence interval, CJA = category, journal impact factor, and authorship, IF = impact factor, Nontop100 = none of Top100, RA = research achievement, SD = standard deviation, SNA = social network analysis, TBG = temporal bubble graph, Top100 = 100 top-cited articles, WoS = Web of Science.

Keywords: 100 top-cited articles, bibliometrics, coword analysis, PubMed, temporal bubble graph, Web of Science

1. Introduction

The number of bibliographic studies published in the field of life sciences and biomedicine has steadily increased over time. There are several reasons behind this rise, including improved accessibility of bibliographic data and software packages that specialize in bibliographic analysis.[1] Over 400 articles with titles of 100 top-cited articles (Top100) were published in PubMed,[2] a free search engine that primarily accesses the MEDLINE database of references and abstracts on life sciences and biomedical topics.

A sharp rise in Top100 publications has also been observed[2] in the past ten years (i.e., [18, 12, 11, 24, 29, 37, 57, 80, 90, 102] by count since 2013). Nonetheless, it is unknown whether their citations are fewer (or more) than their counterparts in other bibliometric studies (Nontop100) based on the research area of biomedicine and bibliometrics. A comparison of article citations between Top100 and Nontop100 is therefore necessary for the study to be conducted.

1.1. Difference in citations between bibliometrics and meta-analysis

The number of meta-analysis studies and systematic reviews has steadily increased, with a total of 5975 and 2119 articles,

The authors have no funding and conflicts of interest to disclose.

The datasets generated during and/or analyzed during the current study are publicly available.

Supplemental Digital Content is available for this article.

[a] Department of Palliative Medicine, Chi-Mei Medical Center, Tainan, Taiwan, [b] Medical Research Department, Chi-Mei Medical Center, Tainan, Taiwan, [c] Department of Physical Medicine and Rehabilitation, Chiali Chi-Mei Hospital, Tainan, Taiwan, [d] Department of Physical Medicine and Rehabilitation, Chung San Medical University Hospital, Taichung, Taiwan.

** Correspondence: Willy Chou, Department of Physical Medicine and Rehabilitation, Chi-Mei Medical Center, 901 Chung Hwa Road, Yung Kung Dist., Tainan 710, Taiwan (e-mail: smilewilly@mail.chimei.org.tw).*

How to cite this article: Hsieh W-T, Chien T-W, Chou W. The 100 most cited articles have fewer citations than other bibliometric articles: A pairwise comparison using a temporal bubble graph. Medicine 2022;101:48(e32101).

Received: 27 October 2022 / Received in final form: 7 November 2022 / Accepted: 8 November 2022

http://dx.doi.org/10.1097/MD.0000000000032101

Key Points

- A coword analysis was carried out to investigate the article themes assigned to each article, which is a novel and modern approach in the literature.
- The TBG enhances the traditional impact beam plot (dot plot) and the temporal bar graph, with 4 dimensions combined in a view.
- Six visualizations were provided in this study, and future relevant studies can easily understand the study results with a quick glance

respectively,[3,4] somewhat lagging behind the number of 7912 for bibliometrics.[5] The availability of online data and software packages for analyses has led to an increase in bibliographic and meta-analysis (or systematic literature review) studies,[2,6] rather than those relying solely on library literature searches. Compared to the article mean citations, a distinct difference was observed for meta-analysis (7.1:5.2).[1] The current study is intended to compare the article mean citations between the Top100 and Nontop100 within bibliographical studies, based on the research area of biomedicine and bibliometrics only.

1.2. Challenges encountered in comparison of citations between Top100 and Nontop100

Comparing article citations between Top100 and Nontop100 presents 3 challenges, including: how to conduct citation comparisons when citations are not normally distributed,[7] classifying article themes in comparison of mean citations (=citations/publications = impact factor [IF]), and displaying citation comparisons on dashboards for a better understanding of detailed features in the 2 groups of Top100 and Nontop100.

It is fortunately possible to obtain the bootstrapping method module from a previous study[7] if the data are not normally distributed when estimating the mean and standard error. By using cluster names in CiteSpace,[8] we are able to mimic the method of determining article themes by coword analysis. Each document can be assigned a theme based on a reference to a previous study,[9] and a timeline clustering map of the articles' cocitation networks can be constructed using CiteSpace (i.e., themes on the row, years on the column, bubbles sized by citations).[10]

Furthermore, temporal bar graphs have been used in bibliometric research.[11–14] It is the disadvantage of temporal bubble graph (TBG)[11] and timeline clustering maps in CiteSpace[8,10] that they cannot be linked to websites as a dashboard does for readers. This study proposes to enhance TBGs by combining their advantages and adding the function of dashboards to TBG.

1.3. Top100 have fewer citations than Nontop100

Reviewing the top 100 most highly cited articles in PubMed,[2] most of these articles focus on a specific topic or discipline and thus limit audiences to those working in that area. On the other hand, Nontop100 articles focus on methods and techniques of analysis, as well as the application of bibliometric metrics (e.g., h-/x-/Y-/ht-index[15–18]). According to previous studies,[19–22] a higher IF has been associated with the publication of reviews and original articles rather than case reports. A rigorous systematic review receives a greater number of citations than a narrative review, whereas case reports with low impact factors are characterized by their focus on a specific topic and are then rarely cited. Therefore, it is proposed and required to verify the hypothesis that the Top100 papers have fewer citations than the Nontop100 papers in the current study.

1.4. Study aims

The aim of this study is to verify the hypothesis that the Top100 articles have fewer citations than the Nontop100 articles, as well as to examine which countries, institutes, departments, authors, and journals dominate the articles in the Top100 and Nontop100.

2. Methods

2.1. Data sources

We searched the keywords of (bibliometric [MeSH Major Topic]) since 2013 in PubMed to ensure articles related to the research area of biomedicine and bibliometrics only. A total of 4574 articles with the PubMed identity number were matched to those in the Web of Science core collection for the extraction of corresponding authors because it is difficult to identify which are corresponding authors in PubMed. The top 50 most cited articles in Web of Science (WoS) were compiled each by Top100 and Nontop100; see Supplemental Digital Content 1, http://links.lww.com/MD/I43.

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

2.2. Four approaches applied to this study

2.2.1. Descriptive statistics in publications. A bar chart was drawn to illustrate the publication trend of articles related to bibliometrics. Pyramid plots in R were used to visualize the top journals with the most publications and mean citations; see Supplemental Digital Content 2, http://links.lww.com/MD/I44.

2.2.2. Research achievements in Top100 and Nontop100. A 4-quadrant plot[23] was applied to present the dominant entities based on the CJAL score[23] determined by the category, journal impact factor, and authorship (CJA) score[24] and the L-index[25] via Equations 1 to 3.

$$\text{CJA} \quad \text{score} = \sum_{i=1}^{n} C_i \times J_i \times A_i, \tag{1}$$

$$\text{CJAL} \quad \text{score} = \sum_{i=1}^{n} C_i \times J_i \times A_i \times \text{L} - \text{index}_i, \tag{2}$$

$$\text{L} - \text{index} = \text{round}\left(\log\left(\frac{\text{Citation}}{A_n \times \text{Age}} + 1\right), 0\right), \geq 1 \tag{3}$$

Three factors are considered in the CJA scores for a published article: the Category (C; e.g., review, original article, case report, etc.), the journal "quality" (J; e.g., impact factor, JIF, or ranking of the journal), and the authorship order denoted by A). The CJAL score is calculated by multiplying each of these 3 aspects as well as the L-index (Equation 3). CJA scores original research articles higher than other types of manuscripts; co-first authors (denoted RP and FP to compute the Y-index RP + FP[26,17]) score higher than other collaborators; for the journal's quality assessment, they use the JIF or SCI/SSCI journal rankings for SCI/SSCI-indexed papers.[24] SCI/SSCI journal rankings are based on JIF in each research domain; therefore, domain-specific journal rankings are usually not significantly different from those based on JIF.[23,24]

The top 10 elements in each entity with CJAL scores in Top100 and Nontop100 are shown on a 4-quadrant radar plot,[23] including countries, institutes, departments, and authors by 2 factors (i.e., RP and FP) on the coordinates. Bubbles were sized by the CJAL score. Accordingly, it is possible to compare the research achievements (Ras) of the top 10 members in each entity with a glance view.

2.2.3. Thematic analysis and comparison of major themes between groups.

2.2.3.1. Themes derived from those 100 articles. By using social network analysis (SNA),[27,28] coword analysis was performed to extract the chief components in clusters as themes (or leaders) in Keywords Plus that were retrieved from the Web of Science core collection. Next, articles were assigned with themes extracted from SNA using equation 4.[9]

$$\text{Theme} = At[\max_{0\le x\le 1} \sum_{i=1}^{L} \sum_{j=1}^{n} (\text{m} < -m+1)_{k\in F}] \quad (4)$$

where L is the number of keywords in article i, n is the number of keywords denoted by keyword k belonging to the theme defined in SNA (i.e., the more coexisting keywords are gathered in an identical cluster). Accordingly, the theme is redirected to the maximal number of keywords (=m) involved in the cluster via Equation 4.

Using a chord diagram,[29] we can understand which themes dominate these 100 articles and their relationship through the color-coded curves to connections; the way to draw the chord diagram is shown in Supplemental Digital Content 2, http://links.lww.com/MD/I44.

2.2.3.2. Comparison of major themes between groups. Comparisons of these themes extracted from the SNA were made for both groups of Top100 and Nontop100 using forest plots.[1,30] Vertical lines represent no effect (e.g., OR = 1). For example, if the confidence intervals (CI) for an individual study (e.g., keyword in this study) overlap with this line, the effect size does not differ from zero (or 1.0) for that study (standard mean difference or odds ratio) at a given level of confidence (e.g., $P < .05$). If the points of the diamond touch the line of no effect, the overall effect cannot be said to differ from no effect at a given level of confidence.[1] With the additional function of zoom-ins and zoom-outs on Google Maps, we incorporated the forest plot on a dashboard to better present the effect on each of these observed studies.

2.2.4. Identification of hypothesis.

To confirm the hypothesis that the Top100 articles have fewer citations than the Nontop100 articles, TBG and forest plots were used.

2.2.4.1. The use of TBG. Similar to CiteSpace,[8] the TBG contains 4 dimensions, namely, themes on the row, years on the column, bubbles sized by article citation, and colors by group. The way to draw the TBG is shown in Supplemental Digital Content 2, http://links.lww.com/MD/I44.

2.2.4.2. The use of bootstrapping method and forest plot. A forest plot was used to compare the mean citations of each theme between the 2 groups using the standard mean difference. The bootstrapping method was used to compute the mean and standard error of citations for each theme in each group. The standard deviation (SD) was then obtained by using Equation 5.

$$\text{SD} = \text{SE} \times \sqrt{n}, \quad (5)$$

where n is the sample size for the theme in either Top100 or Nontop100; the way to draw the forest plot is shown in Supplemental Digital Content 2, http://links.lww.com/MD/I44.

The bootstrapping method[15,31,32] was performed to verify the difference in mean citations for each theme between the groups. A total of 1000 medians were retrieved from random samples of 100 repetitions of median values for each theme in each group. Thus, the mean and 95% CI were obtained to compare differences in mean citations for each theme between groups by inspecting whether the two 95% CI bands were not overlaid.

2.3. Creating dashboards on google maps

All graphs were drawn by author-made modules in Excel (Microsoft Corp). We created pages of HTML used for Google Maps. All relevant CJAL scores for each member can be linked to dashboards on Google Maps. The way to draw visualization involved in this study is deposited in Supplemental Digital Content 2, http://links.lww.com/MD/I44. The bootstrapping method involved with a module is provided in Supplemental Digital Content 3, http://links.lww.com/MD/I45. The study flowchart is shown in Figure 1.

3. Results

3.1. Descriptive statistics

A significant rise in publications regarding bibliometrics in PubMed is evident in Figure 2. We can expect that articles in 2022 would be higher than those in 2021 based on the exponential trend.

In Figure 3, we present a list of the top 13 journals in both the Top100 and Nontop100 groups. There are 9 and 7 articles in the journal of *Medicine*, respectively, which rank top in the 2 Top100 and Nontop100 groups.

When the mean citations are taken into account, the journal *World Psychiatry* ranks at the bottom of Figure 4, with 134 mean citations in Nontop100. As shown at the top of Figure 4, the journal of *Medicine* has mean citations of 9.8 for Top100 articles and 24.7 for Nontop100 articles.

3.2. Research achievements in Top100 and Nontop100

There were 18 citations in total for the 2 sets of the 50 most cited articles (range 1-134; 5 and 26.5 for Top100 and Nontop100, respectively).

The dominant entities in the Top100 are China, Hallym University (South Korea), the medicine department, and the author Vincenzo Montinaro from Italy, with CJAL scores of 69.86, 12.88, 18.08, and 7.80, respectively, in countries, institutes, departments, and authors.

The dominant entities in Nontop100 are China, Huazhong University Science & Technology (China), the medicine department, and the author Dennis F Tompson from the US, with

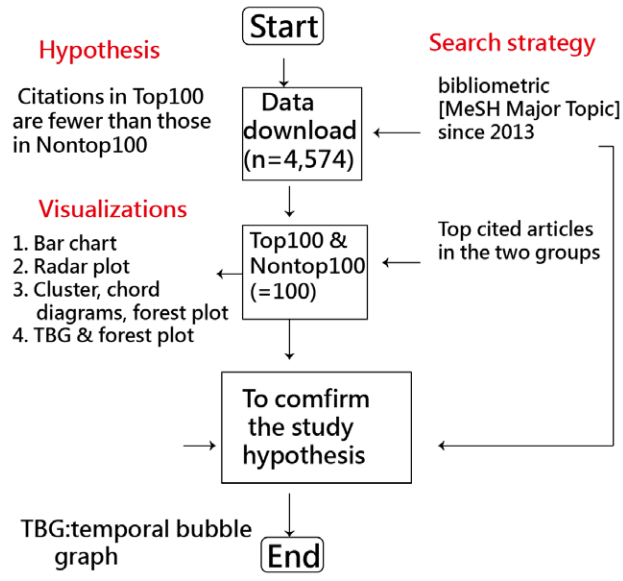

Figure 1. Study flowchart.

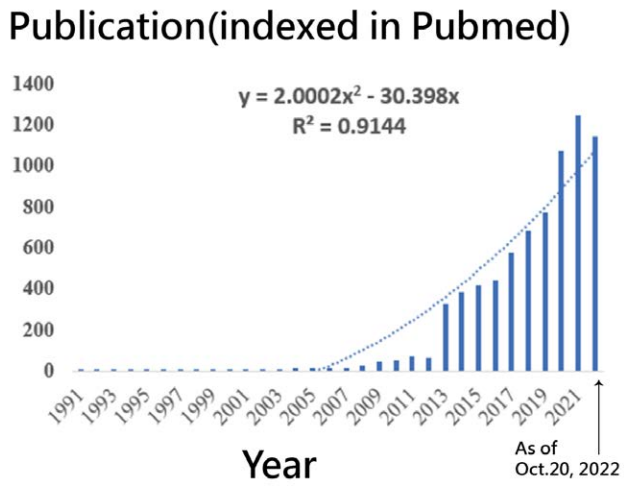

Figure 2. Trend analysis of publications regarding bibliometrics in PubMed (n = 7912).

CJAL scores of 77.74, 18.60, 18.08, and 15.60, respectively, in countries, institutes, departments, and authors.

3.3. Thematic analysis and comparison of major themes between groups

As shown in Figure 5, eleven themes were extracted from the coword analysis of these 100 articles. According to the chord diagram, the majority of keywords are from the theme of citation-analysis, followed by citation-classics and h-index. Colors are used to identify themes. It is important to note that only the top 3 keywords for each cluster are displayed in the chord diagram.

The majority of major keywords in proportional counts observed in themes are in favor of Nontop100, as shown in Figure 6. Only 2 of citation analysis and information science are in favor of Top100.

3.4. Comparison of major themes between groups

A significant difference in mean citations was observed between the 2 groups of Top100 and Nontop100 based on the bootstrapping method (3, 95% CI: [1.18, 4.82]; 26.5, 95% CI: [23.82,29.18], $P < .001$). Readers are invited to scan the QR code and click on the bubble of interest to examine the details about the article information on PubMed.

The 11 themes were clustered using coword analysis and applied to a TBG, which is composed of 4 dimensions: themes, years, citations and groups of articles, as shown in Figure 7.

3.5. Online dashboards shown on google maps

All the QR codes in Figures are linked to the dashboards.[33–39] Readers are suggested to examine the displayed dashboards on Google Maps.

4. Discussion

4.1. Principal findings

We observed that there were 18 citations in total for the 2 sets of 50 most cited articles (range 1–134; 5 and 26.5 for Top100 and Nontop100, respectively). A significant difference in mean citations was observed between the 2 groups of Top100 and Nontop100 based on the bootstrapping method (3, 95% CI: [1.18, 4.82]; 26.5, 95% CI: [23.82, 29.18], $P < .001$). The eleven themes were clustered using coword analysis and applied to a TBG, which is composed of 4 dimensions: themes, years, citations, and groups of articles. Among the 2 groups, the majority of articles were published in the journal of *Medicine (Baltimore)*, with 9 and 7, respectively.

Accordingly, the hypothesis that the Top100 articles have fewer citations than the Nontop100 articles is confirmed.

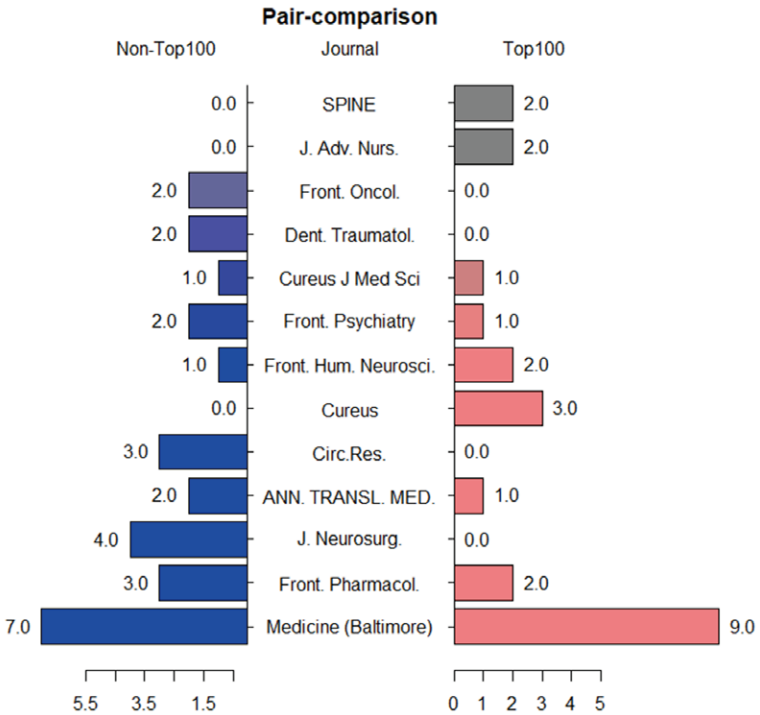

Figure 3. Publications Figure 4. Mean citations in journals for the 2 groups in comparison.

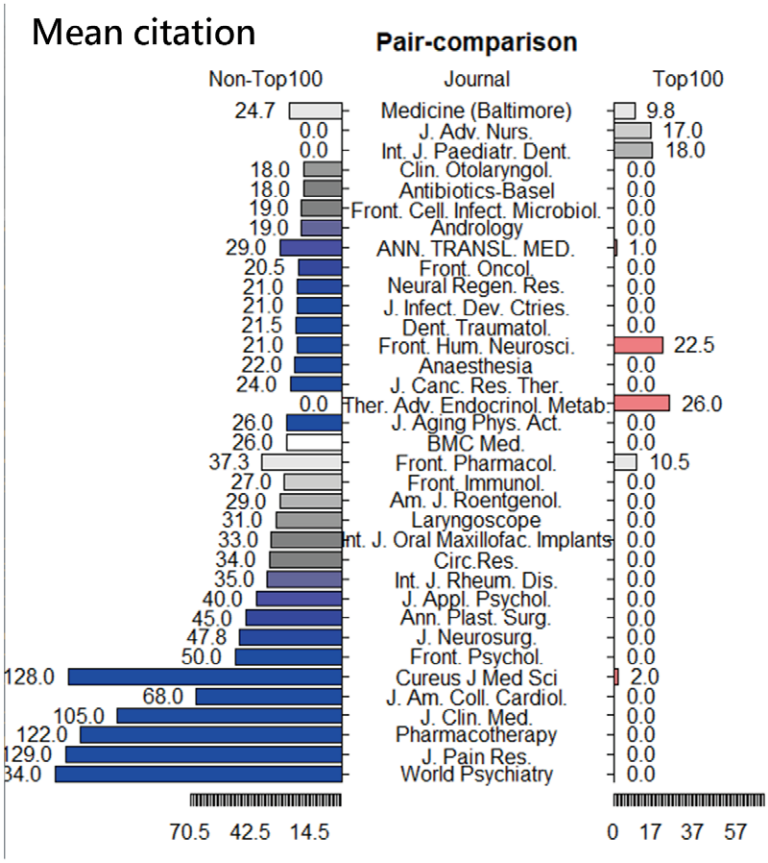

Figure 4. Mean citations in journals for the 2 groups in comparison.

4.2. Additional information

Articles with a higher IF have usually been associated with the publication of reviews and original articles rather than case reports.[19–22] The statement that case reports are rarely cited is questionable, since the number of citations that a case report receives are highly dependent on the content of the report, the type of publication, the journal, and even the topic of the article. For instance, an article classified as a case-report type, submitted on March 18, 2020, and titled "A first case of meningitis and encephalitis associated with SARS-Coronavirus-2" received 1138 and 423 citations in WoS and PubMed, respectively,[40] during the past 2 years and meeting the main strengths noted in accepted manuscripts as the importance "or timeliness" of the problem studies, the quality of the writing, and the soundness of the study design.[41,42]

It is important to note that most Top100 articles using software (e.g., CiteSpace[8] and VOSviewer[43]) with routine reports on article characteristics are case-study types that have low reader interest due to their focus on a particular topic "fishing expeditions in data and conducting systematic reviews that do not provide impactful findings."[44] These articles are thus rarely cited.

The best way to increase citations for Top100 articles is to clearly explain or investigate their novel methodology rather than listing only their entity raking. There is a reason for this: readers are interested in articles that provide concise hypothesis and new insight or significant contribution to the field.[45] For instance, the article would be more interesting if readers were provided with adequate information (e.g., how to conduct this study with visualizations as Supplemental Digital Content 2, http://links.lww.com/MD/I44 in this study) regarding replication of the study.

There have been over 400 publications in PubMed with the titles of 100 top-cited articles.[2] Most of the articles ranked among the top 100 in citations were published in the Journal of *Medicine (Baltimore)*, making it the leading journal in this area.

The characteristics of 100 top-cited articles are commonly visualized with 3 categories of information on descriptive statistics, research domain, and research achievement (RA).[8,44–49] Some studies[8,44–48,49] have applied citation prediction to predict article citations based on the mean citations of article keywords, but the themes have not been definitely classified and clearly visualized through Equation 4.[9]

Additionally, many articles include many Tables and graphs in bibliometrics without applying radar plots[23] and chord diagrams[29] to condense information of interest for readers, as we did in Figures 5, 8, and 9, particularly with the TBG in Figure 7 and forest plots in Figures 6 and 10 to display all 100 articles on a dashboard and save article spaces when compared to those with 100 and 50 articles listed in their studies[50,51] or with 42 Tables and graphs in an article.[52]

4.3. Most cited articles in Top100 and Nontop100

The article cited 32 times was authored by Yeung (Hong Kong) et al[53] and classified it as the theme of VOSviewer in Top100. Based on data provided by WoS in this study, the 100 most-cited articles relevant to neuroscience were identified and characterized. The 100 most-cited articles were mostly research articles published from 1996 to 2000. Stephen M. Smith and Science each had the largest share of these articles. Thirty-seven out of the 100 most-cited articles were interlinked via citations of 1 another, and 41 out of 63 non-interlinked articles could also be categorized under the above 5 topics. It is worth noting that there is no such information regarding VOSviewer that could be directly related to the theme of VOSviewer in the abstract of this study. In contrast, only keywords plus of brain and others were

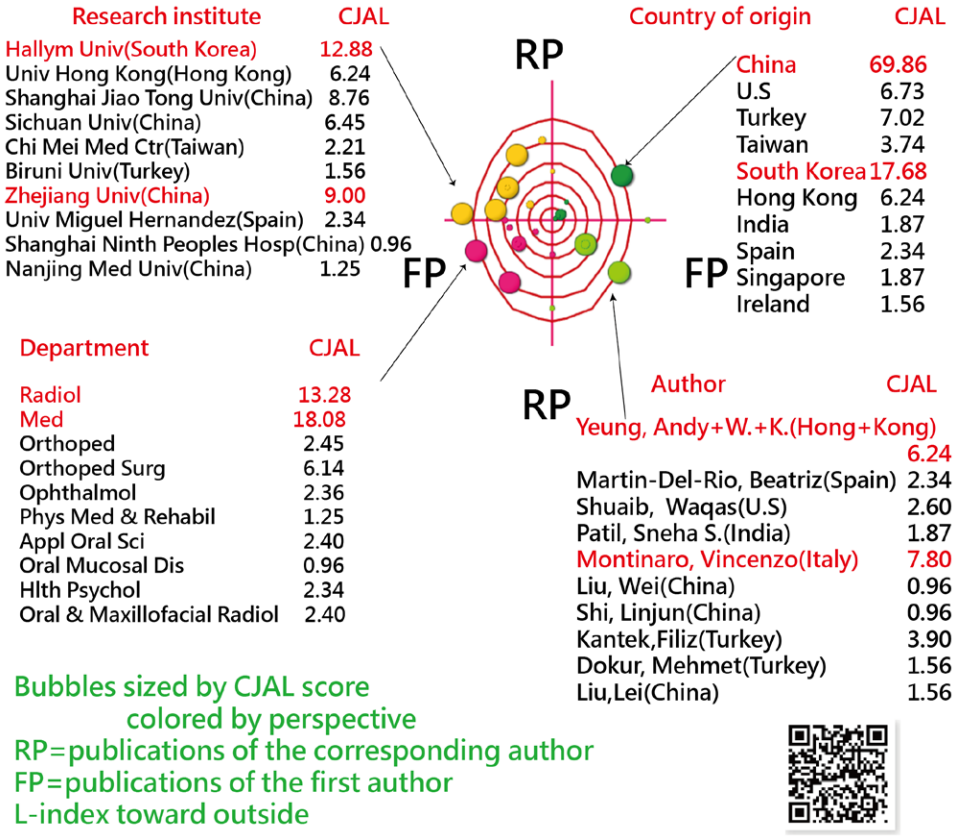

Figure 5. Eleven themes were extracted from coword analysis in these 100 articles.

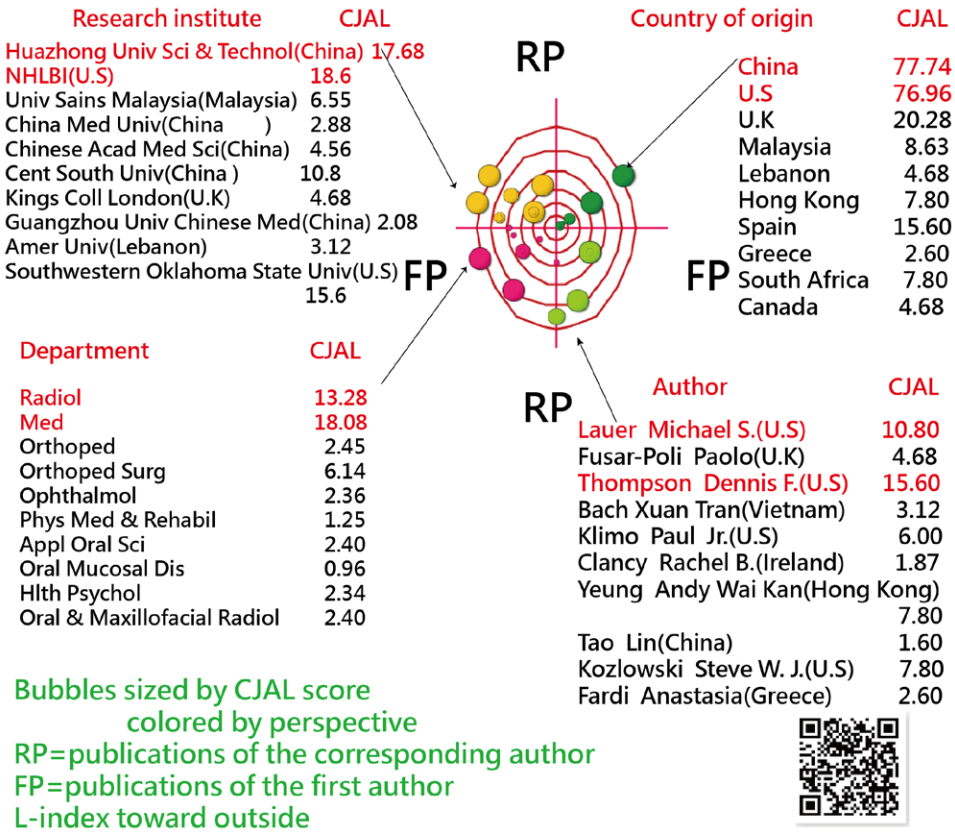

Figure 6. Comparison of proportional counts for major keywords in themes between groups of Top100 and Nontop100. CI = confidence interval.

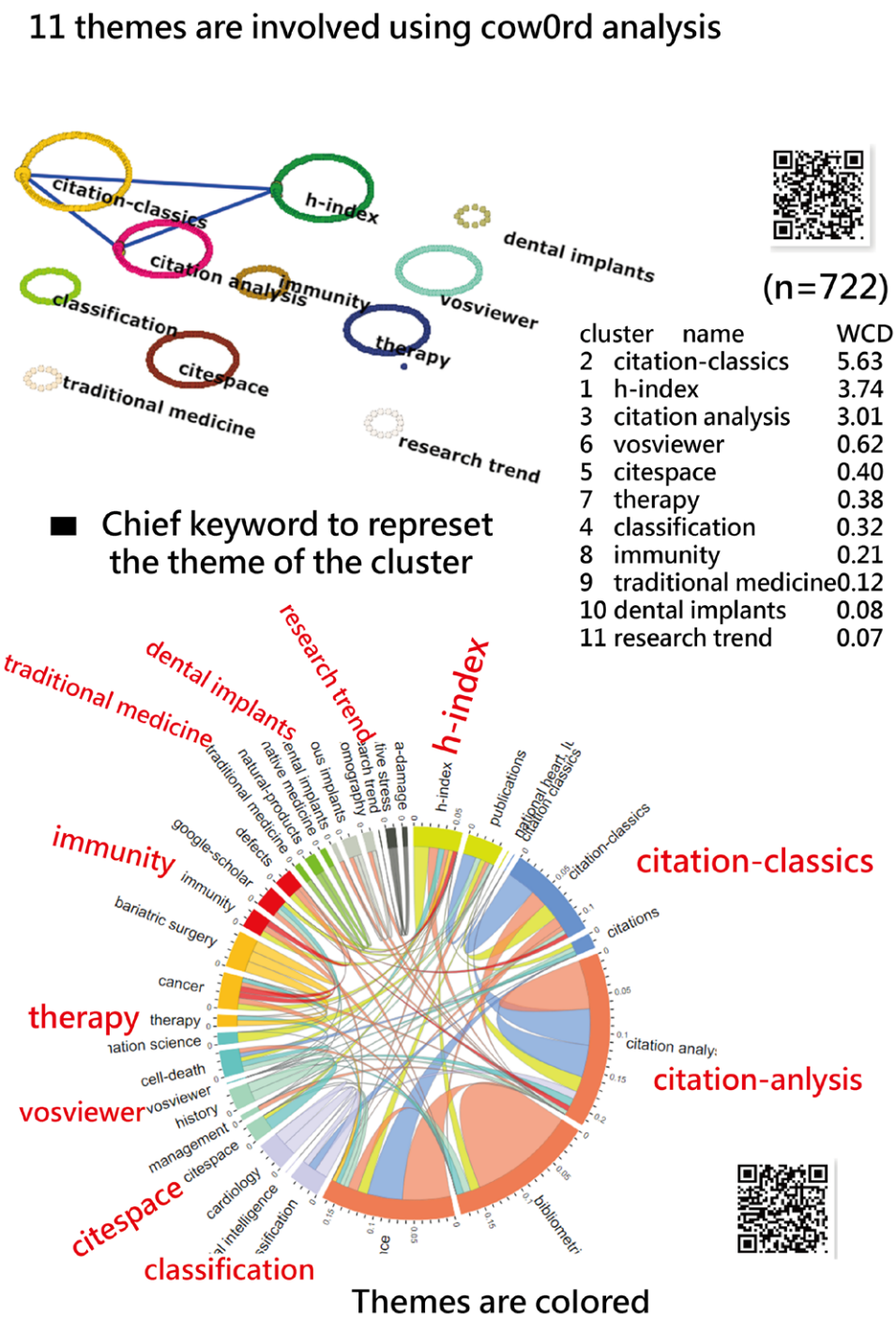

Figure 7. Three dimensions of themes, years, and article citations shown on the TBG. CI = confidence interval, IF = impact factor, SD = standard deviation, TBG = temporal bubble graph.

indexed in this article. The article is classified as VOSviewer that could be found in context instead, which is the feature of the current study using coword analysis to extract article theme via Equation 4.[9]

The article cited 134 times was authored by Fusar-Poli, Paolo (UK)[54] and classified as the theme of classification in Nontop100. Current psychiatric classification is based on ICD/DSM categorical diagnoses, and a promising alternative has been put forward as the "transdiagnostic" approach. A multistep Web of Science literature search was performed to identify all studies that used the word "transdiagnostic" in their title up to May 5, 2018. A total of 111 studies were included, and the quality of the studies was generally low. The conceptual analysis showed that transdiagnostic approaches are grounded more on rediscoveries than on true innovations and are affected by conceptual biases.

A review of 4 productive authors (Lutz Bornmann, Yuh-Shan Ho, Giovanni Abramo, and Ciriaco Andrea D'angelo) with more than 100 bibliometric articles indexed in WoS, with mean citations of 37.6 and median citations of 19. There were no articles entitled with 100 top-cited found in their publications.

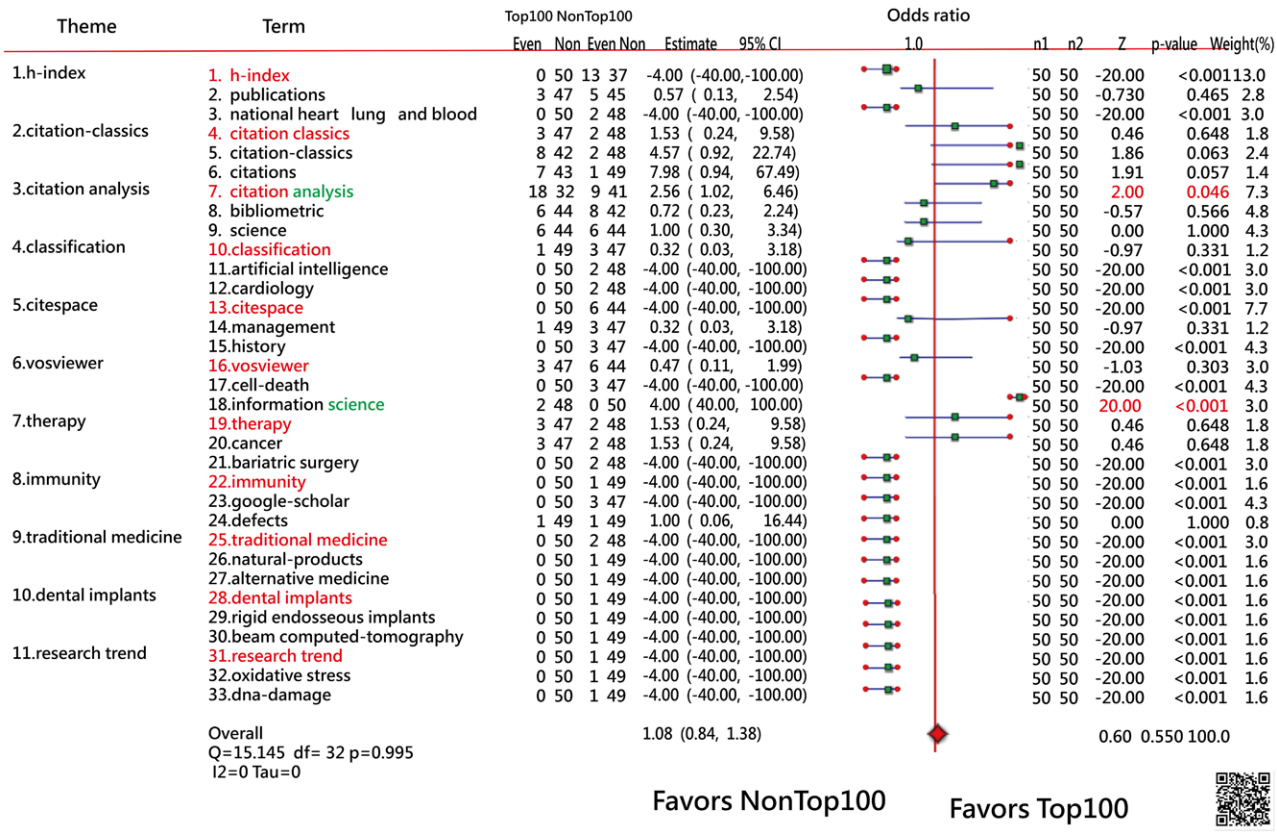

Figure 8. Research achievements in the Top100 for 4 entities using the 4-quadrant radar plot.

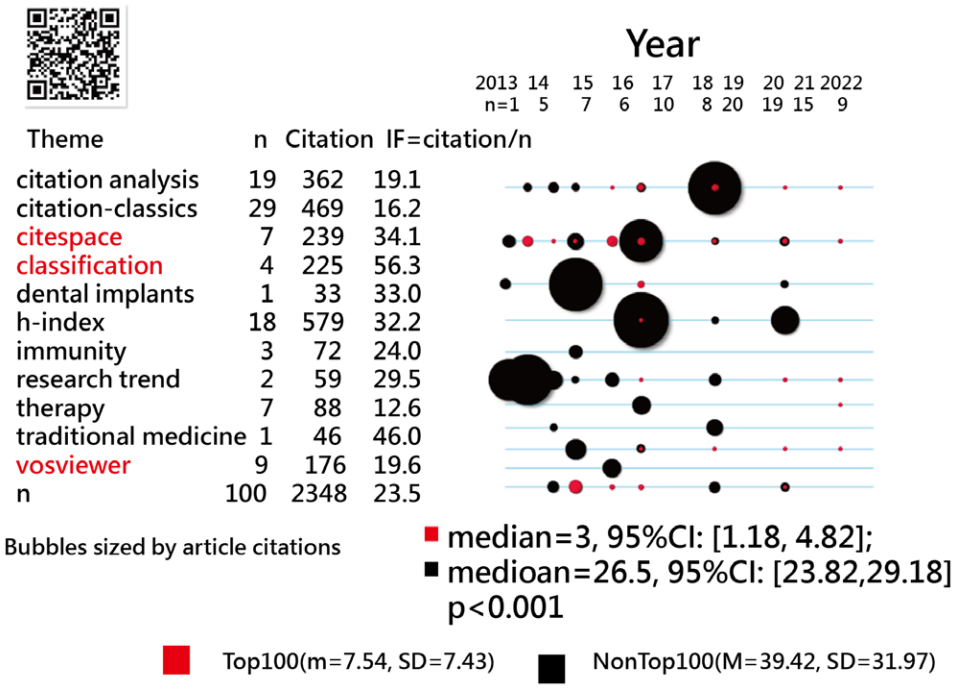

Figure 9. Research achievements in Nontop100 for 4 entities using the 4-quadrant radar plot.

4.4. Implications and changes

Chord diagrams[29] were used to visualize dynamics related to contraceptive use and to bring data into practice. The dashboards (e.g., in Figure 5)[36] provide an easy way to visualize the relationship between elements in themes. As a result of chord diagrams, we gain a clear understanding of the relationship between 2 or more entities (e.g., the themes and clusters shown at the top of Figure 5), something that is rare in bibliometric

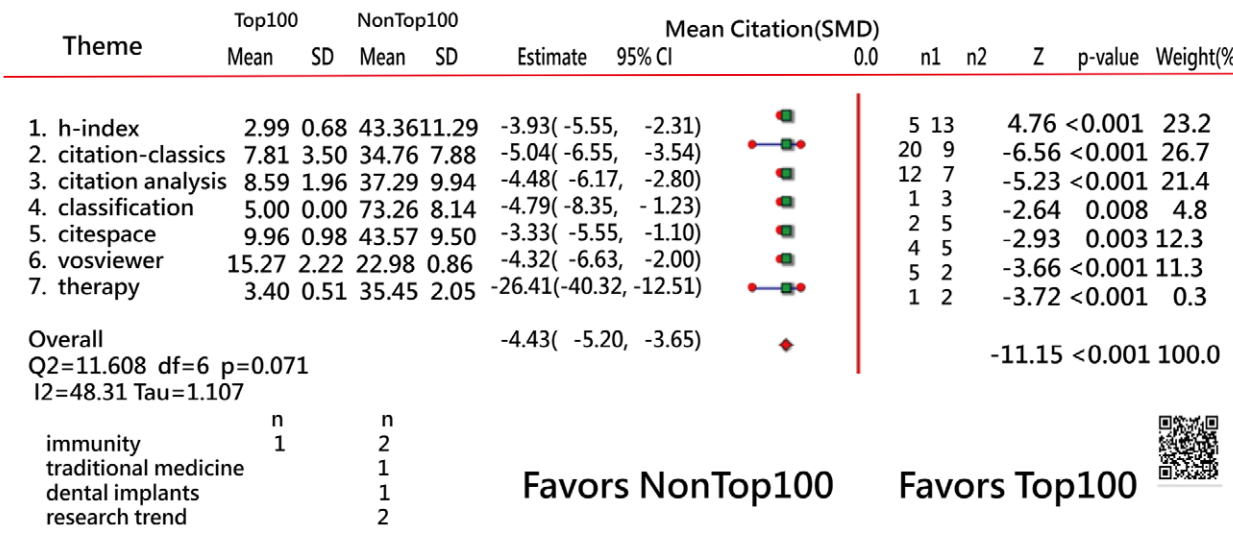

Figure 10. A significant difference in mean citations was observed between the 2 groups of Top100 and Nontop100 across 7 themes in favor of Nontop100. CI = confidence interval, SMD = standard mean difference.

analysis. Supplemental Digital Content 2, http://links.lww.com/MD/I44 provides the R code for reproducing the chord diagram.

Furthermore, chord diagrams could be used by network diagrams to clearly illustrate their network relationships, with more effective representations than the traditional displays, as illustrated at the top of Figure 5.

There are 4 factors that contribute to the CJAL score: subject category, journal impact factor, authorship in positions on the article byline, and article citations. The evaluation of individual RAs has traditionally been based on bibliometric metrics (e.g., h-/x-/Y-/ht-index[15–18]). These metrics have a number of disadvantages, such as assuming that all coauthors contributed equally to the article, regardless of the type of document or journal impact factor. The CJAL score[24] bridges the gap between publications and citations when evaluating the RA beyond those bibliometric metrics.

The current study on Top100 and Nontop100 represents the first attempt to compare the difference in mean citations of articles. The dashboard-type 4-quadrant radar plots depicted in Figures 4 and 8 provide a summary of 4 important entities each in the 2 groups rather than Fables and graphs in traditional bibliometrics. A unique and modern approach of the 4-quadrant radar plots has been applied to bibliometrics before.[23] It is possible to advance bibliometric analysis in this manner in the future.

As seen from the CJAL score, China dominates the bibliometric studies. This study differs from many traditional bibliographical studies in that the publications are computed based on the first and corresponding authors rather than just the first author, as in traditional bibliometric studies. In this study, the dominant entities in the Top100 are China, Hallym University (South Korea), the medicine department, and the author Vincenzo Montinaro from Italy, with CJAL scores of 69.86, 12.88, 18.08, and 7.80, respectively. It is therefore recommended that the CJAL score be used to measure RAs in bibliometric research, particularly when using a radar plot to condense information at a glance.

Traditional bibliographical studies with descriptive statistics, research domain, and RA provided us with a clear understanding of what distinguishes a discipline or field (or topic) from the others and provided insight for physicians and researchers. However, 2 main concerns were frequently overlooked. In such cases, a simplified visualization of all relevant entities is lacking (as in Figures 8 and 9), and a method for displaying all 100 top-cited articles at a glance using the TBG is missing, as shown in Figure 5.

4.5. Limitations and suggestions

A number of issues need to be examined in further research. The first concern is that the Rstudio package used for drawing the chord diagram is not unique and irreplaceable. It is also possible to draw them using several other software packages.

Second, this study uses Google Maps to display dashboards. Since Google Maps requires a paid project key, these installments are not free. It may therefore be difficult for other authors to replicate the usage within a short period of time.

Third, calculating the CJAL score requires considerable computation. The development of this technology will require dedicated software in the future.

Fourth, it has been recommended that the radar plot and CJAL score be combined to simplify article spaces in comparison to other traditional bibliographical studies with many Tables and graphs. However, the RAs are determined by other factors that must also be considered when drawing radar plots in the future.

Fifth, to present the study results, 6 typical visualizations were used, including charts with bars, pyramids, forests, clusters, chords, and bubbles. It is common for bibliometric analysis to be represented visually in a variety of ways. For future studies, it is recommended that more appropriate visual displays be used to facilitate readers' understanding of the study features.

Sixth, article citations should not be solely determined by IF as we compared themes in citations between Top100 and Nontop100 in this study. According to some studies,[55,56] IF should not be used alone as a criterion for evaluating journals. A better assessment of their significance and importance in particular disciplines can be achieved by using Eigenfactor Score, Article Influence Score, and Cited Half-life.

Finally, even though 100 top-cited articles were extracted primarily from WoS and PubMed, the results were different for articles retrieved from other databases (such as Google Scholar and Scopus), while other types of research fields (e.g., engineering and agriculture) were also considered. Future studies should examine whether the Top100 articles have fewer citations than Nontop10 0 articles, as found in this study.

5. Conclusion

A breakthrough was achieved by comparing mean citations in articles of Top100 and Nontop100, which included chord diagrams and the TBG with a demonstration of theme classification.

It is possible to match article themes with author-collaboration clusters (e.g., countries, institutes, and authors) as cluster names labeled in CiteSpace. In future studies, a TBG with 4 dimensions should be applied to 100 top-cited articles in bibliometrics.

Acknowledgments

We thank Enago (www.enago.tw) for the English language review of this manuscript.

Author contributions

WT and TWC initiated the research, collected data, conducted the analysis, and wrote the manuscript. WC contributed to the design of the study and provided critical reviews of the manuscript, and WC and TWC contributed to the interpretation of the results.
Conceptualization: Wan-Ting Hsieh.
Investigation: Willy Chou.
Methodology: Tsair-Wei Chien.

References

[1] Wu JW, Chien TW, Tsai YC, et al. Using the forest plot to compare citation achievements in bibliographic and meta-analysis studies since 2011 using data on PubMed Central: a retrospective study. Medicine (Baltimore). 2022;101:e29213.
[2] PubMed. Over 400 articles with the title of 100 top-cited in PubMed since 1999. Available at: https://pubmed.ncbi.nlm.nih.gov/?term=100%5Btitle%5D%20and%20cited%5Btitle%5D&sort=date&timeline=expanded [access date October 22, 2022].
[3] PubMed. A steady increase in the number of meta-analysis studies. Available at: https://pubmed.ncbi.nlm.nih.gov/?term=(meta%20analysis%5BMeSH%20Major%20Topic%5D&sort=date&timeline=expanded [access date October 22, 2022].
[4] PubMed. A steady increase in the number of meta-analysis studies. Available at: https://pubmed.ncbi.nlm.nih.gov/?term=(Systematic%20Reviews*%5BMeSH%20Major%20Topic%5D)%20or%20(Systematic%20Literature*%5BMeSH%20Major%20Topic%5D)&timeline=expanded [access date October 22, 2022].
[5] PubMed. A steady increase in the number of bibliometrics. Available at: https://pubmed.ncbi.nlm.nih.gov/?term=(bibliometric%5BMeSH%20Major%20Topic%5D&sort=date&timeline=expanded [access date October 22, 2022].
[6] Block JH, Fisch C. Eight tips and questions for your bibliographic study in business and management research. Manag Rev Q. 2020;70:307–12.
[7] Kung SC, Chien TW, Yeh YT, et al. Using the bootstrapping method to verify whether hospital physicians have different h-indexes regarding individual research achievement: a bibliometric analysis. Medicine (Baltimore). 2020;99:e21552.
[8] Chen C, Hu Z, Liu S, et al. Emerging trends in regenerative medicine: a scientometric analysis in Citespace. Expert Opin Biol Ther. 2012;12:593–608.
[9] Chou PH, Lin JCJ, Chien TW. Text mining and forest plots were used to identify similarities and differences between two spine-related journals based on medical subject headings (MeSH terms) and author-specified keywords in 100 top-cited articles. Scientometrics. 2022:10.1007/s11192-022-04549-6.
[10] Yangok C, Li M, Wang Z. A Bibliometric analysis on conservation land trust and implication for China. Int J Environ Res Public Health. 2022;19:12741.
[11] Wu JW, Yan YH, Chien TW, et al. Trend and prediction of citations on the topic of neuromuscular junctions in 100 top-cited articles since 2001 using a temporal bar graph: a bibliometric analysis. Medicine (Baltimore). 2022;101:e30674.
[12] Shen L, Xiong B, Li W, et al. Visualizing collaboration characteristics and topic burst on international mobile health research: bibliometric analysis. JMIR Mhealth Uhealth. 2018;6:e135.
[13] Kleinberg J. Bursty and hierarchical structure in streams. Data Mining Knowl Discov. 2003;7:373–97.
[14] Chow JC, Chien TW, Chou W. Suggestions to the article: demonstrating the ascendancy of COVID-19 research using acronyms. Scientometrics. 2022;127:2897–9.
[15] Hirsch JE. An index to quantify an individual's scientific research output. Proc Natl Acad Sci USA. 2005;98:16569–72.
[16] Fenner T, Harris M, Levene M, et al. A novel bibliometric index with a simple geometric interpretation. PLoS One. 2018;13:e0200098.
[17] Ho YS. Top-cited articles in chemical engineering in science citation index expanded: a bibliometric analysis. Chin J Chem Eng. 2012;20:478–88.
[18] Anderson TR, Hankin RKS, Killworth PD. Beyond the Durfee square: enhancing the h-index to score total publication output. Scientometrics. 2008;76:577–88.
[19] Rodríguez-Lago L, Molina-Leyva A, Pereiro-Ferreirós M, et al. Influence of article type on the impact factor of dermatology journals. Actas Dermosifiliogr. 2018;109:432–8.
[20] Bhandari M, Montori VM, Devereaux PJ, et al.; Hedges Team. Doubling the impact: publication of systematic review articles in orthopedic journals. J Bone Joint Surg Am. 2004;86-A:1012–6.
[21] Nielsen MB, Seitz K. Impact factors and prediction of popular topics in a journal. Ultraschall Med. 2016;37:343–5.
[22] Chien TW, Wang HY, Kan WC, et al. 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:e17631.
[23] Shao Y, Chien TW, Jang FL. The use of radar plots with the Yk-index to identify which authors contributed the most to the journal of Medicine in 2020 and 2021: a bibliometric analysis. Medicine (Baltimore). 2022;101:e31033.
[24] Yeh JT, Shulruf B, Lee HC, et al. Faculty appointment and promotion in Taiwan's medical schools, a systematic analysis. BMC Med Educ. 2022;22:356.
[25] Belikov AV, Belikov VV. A citation-based, author- and age-normalized, logarithmic index for evaluation of individual researchers independently of publication counts [version 1; peer review: 2 approved]. F1000Research. 2015;4:884.
[26] Ho YS. A bibliometric analysis of highly cited articles in materials science. Current Sci. 2014;107:1565–72.
[27] Yie KY, Chien TW, Yeh YT, et al. Using social network analysis to identify spatiotemporal spread patterns of COVID-19 around the world: online dashboard development. Int J Environ Res Public Health. 2021;18:2461.
[28] Kan WC, Chou W, Chien TW, et al. The most-cited authors who published papers in JMIR mHealth and uHealth using the authorship-weighted scheme: bibliometric analysis. JMIR Mhealth Uhealth. 2020;8:e11567.
[29] Finnegan A, Sao SS, Huchko MJ. Using a Chord Diagram to visualize dynamics in contraceptive use: bringing data into practice. Glob Health Sci Pract. 2019;7:598–605.
[30] Yan YH, Chien TW. The use of forest plot to identify article similarity and differences in characteristics between journals using medical subject headings terms: a protocol for bibliometric study. Medicine (Baltimore). 2021;100:e24610.
[31] Efron B. Bootstrap methods: another look at the jackknife. Ann Stat. 1979;7:1–26.
[32] Yie KY, Chien TW, Chen CH, et al. Suitability of h- and x-indices for evaluating authors' individual research achievements in a given short period of years: a bibliometric analysis. Medicine (Baltimore). 2021;100:e25016.
[33] Hien TW. Figure 5 in this study. Available at: http://www.healthup.org.tw/gps/biblioradar.htm [access date October 7, 2022].
[34] Chien TW. Figure 6 in this study. Available at: http://www.healthup.org.tw/gps/biblioradar2.htm [access date October 7, 2022].
[35] Chien TW. Top panel of Figure 7 in this study. Available at: http://www.healthup.org.tw/gps/bibliosna.htm [access date October 7, 2022].
[36] Chien TW. Bottom panel of Figure 7 in this study. Available at: http://www.healthup.org.tw/gps/bibliochord.htm [access date October 7, 2022].
[37] Chien TW. Figure 8 in this study. Available at: http://www.healthup.org.tw/gps/biblioforest.htm [access date October 7, 2022].
[38] Chien TW. Figure 9 in this study. Available at: http://www.healthup.org.tw/gps/biblioibp.htm [access date October 7, 2022].
[39] Chien TW. Figure 10 in this study. Available at: http://www.healthup.org.tw/gps/biblioforest.htm [access date October 7, 2022].
[40] Moriguchi T, Harii N, Goto J, et al. A first case of meningitis/encephalitis associated with SARS-Coronavirus-2. Int J Infect Dis. 2020;94:55–8.
[41] Bordage G. Reasons reviewers reject and accept manuscripts: the strengths and weaknesses in medical education reports. Acad Med. 2001;76:889–96.

[42] Colwell AS, Wong FK, Chung KC. Why do manuscripts get rejected? Plast Reconstr Surg. 2022;150:1169–73.
[43] Van Eck N J, Waltman L. Software survey: VOSviewer, a computer program for bibliometric mapping. Scientometrics. 2009;84:523–38.
[44] Liu P-C, Lu Y, Lin H-H, et al. Classification and citation analysis of the 100 top-cited articles on adult spinal deformity since 2011: a bibliometric analysis. J Chin Med Assoc. 2022;85:401–8.
[45] Yang TY, Chen CH, Chien TW, et al. Predicting the number of article citations on the topic of pemphigus vulgaris with the 100 top-cited articles since 2011: a protocol for systematic review and meta-analysis. Medicine (Baltimore). 2021;100:e26806.
[46] Wang CY, Chien TW, Chou W, et al. Predicting the number of citations of polycystic kidney disease with 100 top-cited articles since 2010: bibliometric analysis. Medicine (Baltimore). 2022;101:e30632.
[47] Chen CH, Chien TW, Yu-Chieh Ho S, et al. Predicting article citations using data from 100 top-cited publications in the field of Psoriasis Vulgaris and biological agents (PVBA) since 1991: a bibliometric analysis. Medicine (Baltimore). 2022;101:e29396.
[48] Lin CH, Chien TW, Yan YH. Predicting the number of article citations in the field of attention-deficit/hyperactivity disorder (ADHD) with the 100 top-cited articles since 2014: a bibliometric analysis. Ann Gen Psychiatry. 2021;20:6.
[49] Wang CY, Li BH, Ma LL, et al. The top-100 highly cited original articles on drug therapy for ventilator-associated pneumonia. Front Pharmacol. 2019;10:108.
[50] Park S, Park BS, Lee YJ, et al. Artificial intelligence with kidney disease: a scoping review with bibliometric analysis, PRISMA-ScR. Medicine (Baltimore). 2021;100:e25422.
[51] Yuan X, Li H, Zhou L, et al. A bibliometric analysis of the 100 most influential papers on peritoneal dialysis. Medicine (Baltimore). 2020;99:e23115.
[52] Radu AF, Bungau SG, Negru PA, et al. In-depth bibliometric analysis and current scientific mapping research in the context of rheumatoid arthritis pharmacotherapy. Biomed Pharmacother. 2022;154:113614.
[53] Yeung AWK, Goto TK, Leung WK. At the leading front of neuroscience: a bibliometric study of the 100 most-cited articles. Front Hum Neurosci. 2017;11:363.
[54] Fusar-Poli P, Solmi M, Brondino N, et al. Transdiagnostic psychiatry: a systematic review. World Psychiatry. 2019;18:192–207.
[55] Roldan-Valadez E, Orbe-Arteaga U, Rios C. Eigenfactor score and alternative bibliometrics surpass the impact factor in a 2-years ahead annual-citation calculation: a linear mixed design model analysis of Radiology, Nuclear Medicine and Medical Imaging journals. Radiol Med. 2018;123:524–34.
[56] Diaz-Ruiz A, Orbe-Arteaga U, Rios C, et al. Alternative bibliometrics from the web of knowledge surpasses the impact factor in a 2-year ahead annual citation calculation: linear mixed-design models' analysis of neuroscience journals. Neurol India. 2018;66:96–104.