Application of a feature extraction and normalization method to improve research evaluation across clinical disciplines

Rui Liu[1#], Qian Liu[2#], Jianwei Shi[3], Wenya Yu[3], Xin Gong[4], Ning Chen[5], Yan Yang[2], Jiaoling Huang[3], Zhaoxin Wang[3,6]

[1]Shanghai Tenth People's Hospital of Tongji University, Shanghai, China; [2]School of Economics and Management, Tongji University, Shanghai, China; [3]School of Public Health, Shanghai Jiao Tong University School of Medicine, Shanghai, China; [4]Shanghai East Hospital Affiliated to Tongji University School of Medicine, Shanghai, China; [5]School of Medicine, Tongji University, Shanghai, China; [6]General Practice Center, Nanhai Hospital, Southern Medical University, Foshan, China

Contributions: (I) Conception and design: J Huang, Z Wang; (II) Administrative support: J Shi, W Yu, X Gong; (III) Provision of study materials or patients: R Liu; (IV) Collection and assembly of data: R Liu; (V) Data analysis and interpretation: N Chen, Y Yang; (VI) Manuscript writing: All authors; (VII) Final approval of manuscript: All authors.

[#]These authors contributed equally to this work.

Correspondence to: Jiaoling Huang; Zhaoxin Wang. School of Public Health, Shanghai Jiao Tong University School of Medicine, South Chongqing Road 227, Shanghai 200025, China. Email: jiaoling_huang@sina.com; supercell002@sina.com.

Background: To deal with the large disparity across disciplines using impact factor, which is widely used in hospitals and has recently come under attack for distorting good scientific practices, we propose a set of systematic methods to improve the equality of research evaluations of various clinical disciplines.

Methods: We used bibliometric information on 18 clinical disciplines from 2016 to 2018. We first sought to clarify disciplinary characteristics with the aim of identifying the characteristic fields for each clinical discipline, and we constructed a keyword database. To minimize the disparity across various clinical disciplines, we used normalized evaluation, referring to the calculation of the normalized coefficient of a specific discipline, to enable a relatively clear evaluation across different disciplines.

Results: Feature extraction was performed, and over 700,000 journals were retrieved each year. Using this information, the journal correlation coefficient was calculated. From 2016 to 2018, oncology had the largest normalized coefficient (0.133, 0.136, 0.146 respectively), which reflects the highest correlation between the characteristic journals of the discipline. The findings showed a clear distinction in journal coverage and journal correlations for different disciplines.

Conclusions: The new evaluation indicator and normalized process measure different features of disciplines, providing a basis for the further balancing of evaluations, and considering differences across disciplines.

Keywords: Research evaluation; clinical discipline; feature extraction; normalization

Introduction

Official administration organizations tend to prefer to allocate more funding to those deemed to be "excellent performers" as research groups that perform better in the evaluation process are thought to be more likely to achieve better scientific results in the near future. Thus, research evaluation analyses conducted at the individual or institutional level provide useful information that can be used not only for the evaluation of the scientific performance of individuals (e.g., professors) and research institutions (e.g., hospitals, colleges, and even official scientific research institutions) but also for detailed future planning of funding

allocation. For more than 30 years, this trend has resulted in debate regarding how to conduct these evaluations more scientifically (1-5). Impact factors (IF) are widely used in the evaluation process, as it fits well with the opinion we have in each field of the best journals in our specialty (6). However, even the creator of IFs, Eugene Garfield, stated that this metric is not appropriate for assessing the importance or significance of individual works (7).

In the field of medicine, scientific research performance has become a crucial index for measuring the comprehensive strength of hospitals: how should the comprehensive scientific research strength of hospitals be measured, especially for hospitals that are affiliated with universities? China has provided its own answer. Nanjing University took the lead in using the Science Citation Index (SCI) as the core index for scientific research evaluation at the end of the 20th century. Since then, the number of SCI papers and their IFs has gradually become the only or most heavily weighted index for evaluating the scientific research level of hospitals and even individuals, which promoting doctors to write more effective and efficient SCI papers to benefit their own promotion and personal acknowledgment. In Chinese hospitals, all departments, including internal medicine, obstetrics and gynecology, surgery, and pediatrics, share the same criteria: the number of SCI papers and their corresponding IFs.

Although IFs are not an optimal measure of the quality of an individual's scientific articles, they are used as a tool for the comparative evaluation of different departments in more than 90% of Chinese hospitals. There are 3 main problems with using only the IFs to compare scientific research performance among different disciplines and departments. First, this approach fails to take into account that some disciplines are in a weaker position and that the scientific researchers in these disciplines are at a disadvantage in the evaluation practice. For example, the Journal Citations Report 2018 states that there were 129,352 SCI articles related to general and internal medicine, but only 13,474 articles related to critical care medicine (https://jcr.clarivate.com/JCRHomePageAction.action?, accessed 08/20/2019). There were nearly 10 times more articles in the former discipline than in the latter. Second, most of the collections in SCI prioritize journals in English, but a substantial number of secondary disciplines, such as family planning, have distinct Chinese characteristics and include little research at the international level. Consequently, this kind of evaluation obviously falls short. Third, for less popular subjects, such as health engineering, local

epidemiology, tropical medicine, and other disciplines, related research is very limited, so there are fewer matching SCI journals, compared with other disciplines, leading to limited choices for researchers seeking to publish their scientific results. In the long term, this puts these researchers at a disadvantage in scientific evaluations and discourages research on less popular subjects.

However, relatively few studies have focused on evaluation across disciplines in hospitals, with a few exceptions, such as a study conducted by Bordons *et al.* that analyzed the structure and research performance of teams in 2 biomedical subfields: pharmacology and pharmacy and the cardiovascular system (8). Recently, using references to measure "interdisciplinarity" has been the focus of attention. Mishra and Torvik suggested that, by using a set of Medical Subject Headings (known as MeSH), it was easier to capture complex trends in how the phenomenon of publishing novel, interdisciplinary articles changes across an author's career (9). Wang also argued that using references made it possible to evaluate the effect of cooperation intensity on knowledge creation (10). Leydesdorff *et al.* proposed a new indicator, DIV, which assesses the variety, balance, and disparity of disciplines independently (11). For this indicator, the Gini coefficient is used to achieve a balance. Results indicate that, compared with earlier indicators, DIV is an improved method for distinguishing between interdisciplinary knowledge integration versus knowledge diffusion. However, to date, researchers have rarely measured the specific differences between disciplines, and limited empirical evidence has been examined in hospitals. Leydesdorff *et al.* also noted that the boundaries between disciplines are fluid and difficult to define (11). Defining a discipline concretely seems to be particularly complicated in a changeable environment that advocates global collaboration and interdisciplinary partnership.

We believe that both individuals and scientific organizations are in urgent need of such definitions. This is particularly the case for hospitals, which increasingly focus on scientific performance and tend to allocate different levels of funding to departments using these indexes. Therefore, to address this complex phenomenon, in the present study we used normalization method, referring to the calculation of the normalized coefficient of a specific discipline, to enable a relatively clear evaluation across different disciplines based on computer science technology. We sought to answer the following questions: (I) how should the characteristics of medical disciplines corresponding to different departments be defined? (II)

Table 1 Articles retrieved in 18 clinical medicine subdisciplines from 2016 to 2018

Disciplines	2016	2017	2018
General internal medicine	51,133	49,358	49,287
Pediatrics	54,650	55,757	51,382
Geriatrics gerontology	74,004	73,501	64,682
Neurosciences neurology	76,983	77,897	76,288
Psychiatry	42,814	44,096	42,491
Dermatology	25,075	25,567	24,502
Radiology, nuclear medicine, medical imaging	45,189	45,722	43,475
Clinical laboratory	54,629	53,991	49,640
Nursing	11,453	11,360	11,598
Surgery	97,889	100,432	95,708
Obstetrics gynecology	34,504	35,434	33,992
Ophthalmology	21,343	20,790	20,784
Otorhinolaryngology	9,339	9,004	8,835
Oncology	115,148	120,803	116,365
Rehabilitation medicine	10,033	11,034	10,990
Sport sciences	8,853	9,141	9,123
Anesthesiology	12,834	12,908	12,360
Emergency medicine	8,256	8,612	8,034

What are the characteristics of the distribution of SCI journals in various disciplines, and what are the implications of this for scientific researchers' contributions? (III) How can innovations be made in theoretical methods to balance the disciplinary differences among departments?

Methods

Discipline selection and definition of characteristics

In the retrieval analysis, we comprehensively considered the 18 subdisciplines of clinical medicine based on the Web of Science classification of research areas and the National Standard Disciplines of China, which include general internal medicine; pediatrics; geriatrics and gerontology; neurosciences and neurology; psychiatry; dermatology; radiology, nuclear medicine, and medical imaging; clinical laboratory; nursing; surgery; obstetrics and gynecology; ophthalmology; otorhinolaryngology; oncology; rehabilitation medicine; sport sciences; anesthesiology; and emergency medicine. Included articles had to satisfy the following two criteria: (I) the journal where the article was published was included in the Journal Citation Reports and the Web of Science Core Collection; and (II) because we considered a period of 3 years (from 2016 to 2018), each journal had to be included in the Journal Citation Reports for these 3 consecutive years.

Information on the final articles retrieved after excluding journals with missing values is presented in *Table 1*. The IFs were taken from the Journal Citation Reports database (2016, 2017, and 2018 editions). To clarify disciplinary characteristics, we first defined each of the subdisciplines of clinical medicine on the basis of the research content, research methods, and applications. We then analyzed the key elements of each discipline through a meta-analysis of the literature. Finally, we selected the characteristic fields for each clinical discipline, creating the keyword database. The process included the following steps: (I) reviewing the classic international textbooks as a foundation of the keyword database with the aim of determining the basic research content of each discipline; (II) conducting definition extension by defining core keywords through the

definition given by experts of this discipline as a supplement; and (III) identifying international core journals to screen for "hot" keywords in LetPub and the Web of Science database in order to keep up with the research hotspot.

The keyword database was then reviewed by experts with more than 10 years of scientific experience in their fields. After 3 rounds of review, a revised database was established to provide a basis for using feature extraction to match each discipline's journal catalog afterwards.

Feature extraction process

The aim of the present study was to improve the traditional evaluation system used in Chinese hospitals and provide references for researchers to get further paper submission smoother, with a narrow focus on the clinical medicine subdisciplines that are included among the departments in all Chinese hospitals. To respond to the question of how to define the characteristics of medical disciplines, a variety of methods for analyzing interdisciplinary differences have been reported in the literature, including in a number of studies examining observations of disparities across a wide variety of specific domains (economics, medical subject categories, library and information science, and science and technology studies) (12-14). The authors of these previously published articles called for the use of more detailed classifications at the subject level. However, none of these existing studies has tried to cover all the journals in a given discipline in defining the discipline's characteristics, and most research has remained at the stage of qualitative or semiquantitative analysis. Creating an adequate delineation of disciplines to provide a framework within which not only articles but also personal performance can be thoroughly compared is a less discussed, but equally important, issue.

In this study, we present the novel method of feature extraction. This method is based on the sorting algorithm of an internet search engine. It refers to evaluating and sorting the correlations between relevant sentences in the retrieved text and keywords according to their degree of match, determined by, for example, the location and frequency of keywords in the keyword database appearing in the text. After crawling each page, the search engine extracts the features of the page and builds an index. In the ranking of search results, results with a more relevant feature have a higher ranking. The present study also set a weight index for the occurrence position of the keyword field to provide another reference for sorting. Articles in which the keyword field was more important have a higher ranking. We used this method to establish the correlation ranking of each journal and the above disciplines, and we used the IF as the weight to be included in the evaluation criteria of the discipline-related journal ranking.

The feature extraction process can be divided into the following 3 procedures: the gripping layer, the preprocessing layer, and the computing layer. These 3 layers are explained in more detail later and visualized schematically in *Figure 1*. The first step is to create a gripping system, which is built in Python language, and all data are stored in MongoDB. The purpose of establishing the gripping layer is to input the keywords of the characteristics of each discipline from the keyword database, unify the input format, and de-duplicate the keywords. The gripping layer adopts the Spider class and the Item class of the Scrapy framework to complete the construction, and the Web of Science is used as the crawling source. Spider is the primary class responsible for crawling and can be customized to point to different sites. Item is used to structure the captured items. The fields required for this research included article titles, keywords, abstracts, the position of keywords in the text, the journal where the article was published, and the year the article was published. In the preprocessing layer, we used SQL statements, the Scrapy framework, and the NLTK (Natural Language Toolkit) for natural language processing to complete the preprocessing and calculation of the fetching data. For all articles in the database, we used the term frequency-inverse document frequency algorithm to calculate the weight of keywords—that is, to obtain the weight dictionary of the corresponding keywords for each article. Finally, in the computing layer, the journal correlation coefficient was calculated. The algorithm was as follows: the number of articles captured in a certain journal in a given year in the preprocessing layer was used as the denominator, x_{ia}, and the number of all articles captured in that journal in that year, x_{ib}, was taken as the numerator and then divided by the weighted average of all the articles captured in the year, β_i. The equation for the journal correlation coefficient, Y_i, was as follows:

$$Y_i = \frac{\sum x_{ia} \times \beta_i}{x_{ia}} \times \frac{x_{ia}}{x_{ib}} = \frac{\sum x_{ia} \times \beta_i}{x_{ib}} \quad [1]$$

We defined the integrated impact coefficient, γ_i, of journal i as:

$$\gamma_i = Y_i \times I_i \quad [2]$$

Here, I_i is the IF of journal i in a given year.
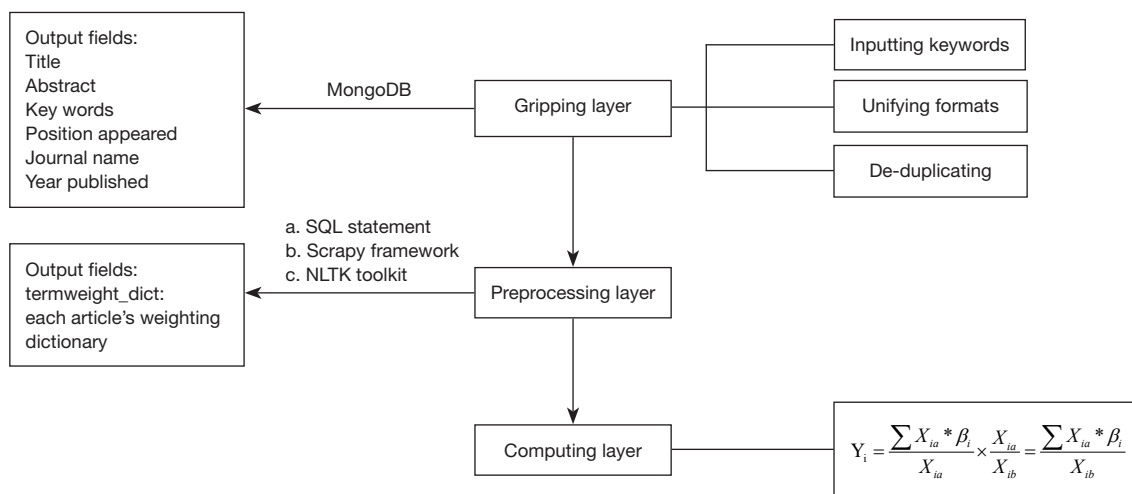
Figure 1 Key procedures and details of feature extraction. NLTK, Natural Language Toolkit.

Normalized evaluation

The whole range of IFs differed not only over different time periods but also across different research fields (15). For example, according to the CD-ROM version of the science and social science editions of the Journal Citation Reports for 2018, in the field of management, the highest IF was 12.289 for the *Academy of Management Annals*, whereas the highest IF in nanoscience and nanotechnology was 74.499 for *Nature Reviews Materials*. There was also a large gap between the average IFs in these fields. The average IF in nanoscience and nanotechnology was 6.795, nearly 3 times higher than the average IF in management (2.983). To minimize the gap in citation numbers across different fields, Rehn *et al.* proposed the item-oriented field-normalized citation score, which was later referred to as the normalized citation score (NCS) (16). This well-received field-normalized indicator is calculated by dividing the number of citations to a specific paper by the average citation rate of the papers published in the same discipline category and year. The equation for the NCS of a given paper is as follows:

$$\mu_i = \frac{1}{n} \sum_{i=1}^{n} C_i \qquad [3]$$

$$NCS_i = \frac{C_i}{\mu_i} \qquad [4]$$

Here, C_i denotes the overall number of citations to paper i, and n refers to the number of papers in the same discipline category as paper i. Therefore, μ_i is the total average of citation numbers in a specific discipline category. The NCS is defined by the specific values of C_i and μ_i.

Using the feature extraction process described earlier, we included all journals in all fields in an efficient, systematic, and low-cost method, ultimately obtaining the journal correlation coefficient. Because of the wide use of the NCS, we then explored how to convert the standardized scores of papers in various disciplines. Through the process of transforming a dimensional expression into a dimensionless scalar, we obtained a normalized weighted system for each discipline, enabling the construction of an interdisciplinary and comprehensive evaluation system of papers and the realization of uniform and fair evaluation across different clinical disciplines. The calculation procedures were as follows.

The mean weight of a specific discipline was obtained by calculating the sum of the influence factors and the journal correlation coefficient of the selected journals. Y_i was the journal correlation coefficient of journal i, and I_i was the IF of journal i (journal i was included when it belonged to the selected discipline). We defined the weight of a given discipline i as:

$$p_i = \frac{1}{n} \sum_{i=1}^{n} (Y_i \cdot I_i) = \frac{1}{n} \sum_{i=1}^{n} \gamma_i \qquad [5]$$

The sum of all the weights of every discipline was defined as:

$$T_0 = \sum_{i=1}^{n} \rho_i \qquad [6]$$

The calculation of the normalized coefficient of a specific discipline (discipline-normalized coefficient, DNC_i) was

defined as:

$$DNC_i = \rho_i / T_0 \qquad [7]$$

Statistical analysis

Statistical analyses were performed using Python 3.7 (Python Software Foundation, Wilmington, Delaware, USA). SQL statements, Scrapy framework, and NLTK (Natural Language Toolkit) for natural language processing were used to tackle the feature extraction process.

Results

Overview of retrieved journals in 3 consecutive years

Table 2 provides an overview of all journals retrieved using the method discussed earlier over 3 consecutive years (2016, 2017, and 2018). Using feature extraction, we can see that in 2016, at least 1,144 journals were retrieved among 18 disciplines, with a maximum number of 5,870 (more than 5 times higher than the minimum). Using our keyword database, we found the papers that were highly related to a certain discipline, with 5,520 papers related to oncology and only 390 classified into the field of nursing. Referring to the results for 2017 and 2018, the total numbers of journals retrieved using our method changed slightly. Neurosciences and neurology remained the top discipline, with the largest total number of journals in both years, and there was a sharp difference between this discipline and the discipline with the lowest total number of journals in 2016 and 2017, sport sciences, which ranked last in 2018, with only 1,070 journals found. The same trend was observed in the total number of papers retrieved in 2017 and 2018. However, the number of papers in the field of nursing decreased slightly in 2017 (n=348), and then slightly increased in 2018 (n=386). The number of papers in the discipline of oncology increased stably, from 5,592 in 2017 to 5,691 in 2018.

To generate an approximation of a discipline that is sufficient for analysis, we aimed to concentrate on the top 50 journals, including the journals ranked in the top 50 by IFs and journal correlation coefficients. We adopted this simplification because of the aim of the method, which was not only to exhaustively display all related publications but to generate an overview of a particular research policy context and guide further evaluations and contributions for researchers in different fields. In 2016, oncology had the

highest average IF (mean: 8.717, range, 4.041–33.900), and the field of general internal medicine had the second highest average IF (mean: 8.359, range, 1.847–72.406). In contrast, nursing was identified as the discipline with the lowest average IF (mean: 1.797), including journals with IFs from 1.261 to 3.755. In 2018, general internal medicine took the place of oncology as the discipline with the highest average IF, with the mean IF for general internal medicine increasing from 9.081 in 2017 to 9.461 in 2018. Nursing remained in last place, but saw a slight increase in the average IF, from 1.841 in 2017 to 1.974 in 2018. Examining the fluctuations in the journal correlation coefficients over these 3 consecutive years, it can be seen that 9 disciplines showed a slight increase during this timeframe, and no discipline steadily decreased throughout the analyzed period. The average journal correlation coefficient fluctuated from 0.025 to 1. From 2016 to 2018, surgery had the highest average journal correlation coefficients (range, 0.750–0.754), and the lowest average journal correlation coefficients belonged to clinical laboratory (range, 0.215–0.217).

Analysis of differences across disciplines

Table 3 compares the top 3 journals ranked by IFs and journal correlation coefficients. Considering that changes over 2 years might not be obvious, we constructed a cross-year comparison; that is, 2016 and 2018 data were selected for analysis. We also clearly displayed the corresponding journal correlation coefficient for each journal in a given discipline. For example, our process of feature extraction identified the *New England Journal of Medicine* as having the highest IF, but its related journal correlation coefficient was much lower compared with the journals ranked in the top 3. In obstetrics and gynecology, the *New England Journal of Medicine* ranked first based on IFs, but its journal correlation coefficient, 0.043, was relatively low. In contrast, the highest journal correlation coefficient in this discipline was 0.826 and belonged to *Gynecologic Oncology*; this was nearly 20 times higher than the correlation coefficient of the *New England Journal of Medicine*. A similar phenomenon was observed in many disciplines: Journals with high IFs did not have proportionately high journal correlation coefficients.

Comparing differences across the 18 disciplines, each discipline had its own featured journals, and their ranges of correlation coefficients varied substantially. Ophthalmology had the highest journal correlation coefficient in both 2016 and 2018, with the *Annual Review of Vision Science* ranking first; this journal's correlation coefficient reached 1.

Table 2 Information about retrieved journals by discipline in 2016, 2017, and 2018

Disciplines	Total journals	Retrieved papers (minimum/ maximum)	Top 50 impact factors		Top 50 correlation coefficients	
			Average	(Minimum/ maximum)	Average	(Minimum/ maximum)
2016						
General internal medicine	3,997	(1/4,026)	8.359	(1.847/72.406)	0.487	(0.386/0.707)
Pediatrics	4,825	(1/1,375)	4.527	(2.254/18.392)	0.647	(0.528/0.920)
Geriatrics/gerontology	4,978	(1/1,802)	4.8	(2.960/19.896)	0.489	(0.247/0.800)
Neurosciences/neurology	5,870	(1/2,162)	6.741	(3.552/26.284)	0.605	(0.429/0.839)
Psychiatry	3,825	(1/1,032)	5.894	(3.295/15.307)	0.714	(0.571/0.860)
Dermatology	3,713	(1/1,204)	4.668	(2.351/13.081)	0.542	(0.279/0.767)
Radiology/nuclear medicine/medical imaging	5,102	(1/2,220)	6.178	(2.617/19.896)	0.499	(0.345/0.919)
Clinical laboratory	5,793	(1/1,893)	7.679	(3.786/19.896)	0.217	(0.116/0.561)
Nursing	1,332	(1/390)	1.797	(1.261/3.755)	0.659	(0.540/0.891)
Surgery	5,258	(1/1,711)	5.843	(2.953/19.896)	0.750	(0.629/0.996)
Obstetrics/gynecology	3,727	(1/1,081)	5.857	(2.443/72.406)	0.641	(0.436/0.826)
Ophthalmology	2,852	(1/4,811)	6.25	(2.466/72.406)	0.699	(0.538/1)
Otorhinolaryngology	1,963	(1/414)	3.798	(1.568/13.081)	0.439	(0.025/0.755)
Oncology	5,029	(1/5,520)	8.717	(4.041/33.900)	0.699	(0.610/0.889)
Rehabilitation medicine	1,612	(1/522)	3.245	(1.769/19.651)	0.459	(0.269/0.929)
Sport sciences	1,144	(1/1,357)	3.415	(1.872/19.651)	0.376	(0.275/0.541)
Anesthesiology	1,752	(1/1,223)	4.268	(1.803/44.405)	0.574	(0.178/0.878)
Emergency medicine	1,732	(1/485)	6.289	(2.036/72.406)	0.341	(0.047/0.789)
2017						
General internal medicine	3,903	(21/3,892)	9.081	(2.029/79.258)	0.492	(0.368/0.739)
Pediatrics	4,877	(1/2,886)	3.885	(2.028/20.773)	0.647	(0.503/0.900)
Geriatrics/gerontology	5,027	(1/1,730)	4.771	(2.917/16.834)	0.498	(0.250/0.802)
Neurosciences/neurology	6,123	(1/2,192)	6.848	(3.653/27.138)	0.611	(0.441/0.846)
Psychiatry	3,919	(1/1,774)	5.848	(3.476/16.642)	0.704	(0.584/0.829)
Dermatology	3,810	(1/1,384)	4.858	(2.564/13.258)	0.54	(0.297/0.922)
Radiology/nuclear medicine/medical imaging	5,387	(1/2,010)	6.524	(2.758/23.425)	0.503	(0.347/0.907)
Clinical laboratory	5,716	(1/1,658)	7.788	(3.950/23.425)	0.215	(0.110/0.571)
Nursing	1,333	(1/348)	1.841	(1.242/3.656)	0.655	(0.526/0.905)
Surgery	5,251	(1/1,529)	5.947	(2.792/23.425)	0.75	(0.629/0.925)
Obstetrics/gynecology	3,730	(1/1,777)	6.016	(2.434/79.258)	0.628	(0.444/0.841)
Ophthalmology	2,868	(1/4,260)	6.446	(2.464/79.258)	0.708	(0.526/1)
Otorhinolaryngology	1,893	(1/355)	3.88	(1.664/13.258)	0.433	(0.024/0.800)

Table 2 (*continued*)

Table 2 (*continued*)

Disciplines	Total journals	Retrieved papers (minimum/ maximum)	Top 50 impact factors		Top 50 correlation coefficients	
			Average	(Minimum/ maximum)	Average	(Minimum/ maximum)
Oncology	5,149	(1/5,592)	9.273	(4.204/36.418)	0.693	(0.612/0.910)
Rehabilitation medicine	1,729	(1/459)	3.512	(1.863/23.425)	0.453	(0.284/0.908)
Sport sciences	1,140	(1/1,319)	3.599	(2.042/23.425)	0.379	(0.276/0.647)
Anesthesiology	1,732	(1/964)	4.496	(1.924/47.661)	0.585	(0.205/0.867)
Emergency medicine	1,772	(1/571)	6.745	(2.141/79.258)	0.35	(0.048/0.844)
2018						
General internal medicine	3,620	(1/4,022)	9.461	(1.994/70.67)	0.494	(0.385/0.703)
Pediatrics	4,525	(1/1,944)	4.695	(2.256/19.233)	0.66	(0.535/0.908)
Geriatrics/gerontology	4,695	(1/1,738)	4.811	(3.08/18.639)	0.492	(0.257/0.806)
Neurosciences/neurology	5,917	(1/2,217)	7.003	(3.749/28.755)	0.616	(0.437/0.881)
Psychiatry	3,595	(1/1,097)	6.04	(3.488/18.329)	0.71	(0.556/0.843)
Dermatology	3,575	(1/1,356)	4.92	(2.190/14.110)	0.544	(0.320/0.918)
Radiology/nuclear medicine/medical imaging	5,027	(1/2,080)	6.7	(2.895/23.239)	0.513	(0.358/0.909)
Clinical laboratory	5,497	(1/1,268)	8.075	(3.812/23.239)	0.215	(0.109/0.609)
Nursing	1,196	(1/386)	1.974	(1.327/3.570)	0.677	(0.550/0.960)
Surgery	5,103	(1/1,879)	6.147	(2.903/23.239)	0.754	(0.640/0.926)
Obstetrics/gynecology	3,606	(1/1,244)	5.834	(2.413/70.670)	0.644	(0.425/0.834)
Ophthalmology	2,539	(1/4,421)	6.508	(2.509/70.670)	0.712	(0.545/1)
Otorhinolaryngology	1,822	(1/283)	3.926	(1.819/14.110)	0.433	(0.026/0.814)
Oncology	4,872	(1/5,691)	9.292	(4.117/35.386)	0.7	(0.624/0.890)
Rehabilitation medicine	1,590	(1/547)	3.575	(1.907/23.239)	0.475	(0.302/0.944)
Sport sciences	1,070	(1/1,165)	3.779	(2.054/23.239)	0.38	(0.279/0.591)
Anesthesiology	1,576	(1/853)	4.731	(1.958/51.273)	0.585	(0.178/0.879)
Emergency medicine	1,559	(1/589)	6.992	(2.247/70.670)	0.377	(0.051/0.802)

However, in the disciplines of clinical laboratory and sport sciences, the journal correlation coefficients were lower (0.561 and 0.541, respectively). When considering 2 years, the list of the top 3 journals ranked by IFs remained almost unchanged. However, slight differences were seen in the list ranked by journal correlation coefficients from 2016 to 2018. In 2018, taking the discipline of psychiatry as an example, the *Journal of Affective Disorders*, *Depression and Anxiety*, and *Bipolar Disorders* appeared first in the list, with correlation coefficients of 0.843, 0.824, and 0.821, respectively.

Wagner stated that "interdisciplinary" areas of scientific practice have not been well studied, although these areas are growing rapidly (17). The underlying concept of "disciplines" are social baselines used to allocate the privileges and responsibilities of expertise, as well as resources. Our comparison of the characteristics of different disciplines in 3 recent years suggests a clear distinction between the coverage of different disciplines, which provides evidence for questioning the fairness of

Table 3 Top 3 journals and corresponding journal correlation coefficients of 18 clinical medicine disciplines in 2016 and 2018

Disciplines	2016				2018			
	Top 3 journals (by IF)	Correlation coefficient	Top 3 journals (by correlation coefficient)	Correlation coefficient	Top 3 journals (by IF)	Correlation coefficient	Top 3 journals (by correlation coefficient)	Correlation coefficient
General internal medicine	N Engl J Med	0.392	TER ARKHIV	0.707	N ENGL J MED	0.36	TER ARKHIV	0.703
	LANCET	0.254	KOREAN J INTERN MED	0.619	LANCET	0.23	SHOCK	0.662
	JAMA-J AM MED ASSOC	0.317	IRAN RED CRESCENT ME	0.409	JAMA-J AM MED ASSOC	0.32	INTERNAL MED	0.615
Pediatrics	GASTROENTEROLOGY	0.028	PEDIATRIC DIABETES	0.920	GASTROENTEROLOGY	0.0234	PEDIATRIC DIABETES	0.908
	BLOOD	0.052	PEDIATR TRANSPLANT	0.828	BLOOD	0.0508	PEDIATR TRANSPLANT	0.868
	J ALLERGY CLIN IMMUN	0..178	CHILD CARE HLTH DEV	0.743	J ALLERGY CLIN IMMUN	0.161	PEDIATRIC OBESITY	0.848
Geriatrics/ gerontology	J AM COLL CARDIOL	0.034	NEUROBIOL AGING	0.800	J AM COLL CARDIOL	0.021	NEUROBIOL AGING	0.806
	BLOOD	0.022	INT J GERIATR PSYCH	0.788	BLOOD	0.024	INT J GERIATR PSYCH	0.803
	DIABETES OBES METAB	0.239	FRONT AGING NEUROSCI	0.758	AGING CELL	0.733	FRONT AGING NEUROSCI	0.798
Neurosciences/ neurology	LANCET NEUROL	0.642	NEUROBIOL AGING	0.800	LANCET NEUROL	0.635	J HEADACHE PAIN	0.881
	J AM COLL CARDIOL	0.027	INT J GERIATR PSYCH	0.788	EUR HEART J	0.033	CEPHALALGIA	0.861
	EUR HEART J	0.031	FRONT AGING NEUROSCI	0.758	J AM COLL CARDIOL	0.027	EPILEPSY BEHAV	0.854
Psychiatry	JAMA PSYCHIAT	0.753	EPILEPSY BEHAV	0.839	LANCET PSYCHIAT	0.556	J AFFECT DISORDERS	0.843
	AM J PSYCHIAT	0.571	EPILEPSY RES	0.832	JAMA PSYCHIAT	0.707	DEPRESS ANXIETY	0.824
	MOL PSYCHIATR	0.685	CEPHALALGIA	0.818	AM J PSYCHIAT	0.643	BIPOLAR DISORD	0.821
Dermatology	J ALLERGY CLIN IMMUN	0.092	MELANOMA RES	0.767	ANN RHEUM DIS	0.067	MELANOMA RES	0.918
	ANN RHEUM DIS	0.063	J COSMET LASER THER	0.702	J ALLERGY CLIN IMMUN	0.081	AM J CLIN DERMATOL	0.720
	NAT COMMUN	0.009	ACTA DERMATOVENER CR	0.691	NAT COMMUN	0.013	J COSMET LASER THER	0.681

Table 3 (*continued*)

Table 3 (*continued*)

Disciplines	2016				2018			
	Top 3 journals (by IF)	Correlation coefficient	Top 3 journals (by correlation coefficient)	Correlation coefficient	Top 3 journals (by IF)	Correlation coefficient	Top 3 journals (by correlation coefficient)	Correlation coefficient
Radiology/nuclear medicine/ medical imaging	J AM COLL CARDIOL	0.084	J CONTEMP BRACHYTHER	0.919	EUR HEART J	0.054	J CONTEMP BRACHYTHER	0.909
	EUR HEART J	0.085	BRACHYTHERAPY	0.840	GASTROENTEROLOGY	0.024	BRACHYTHERAPY	0.856
	GASTROENTEROLOGY	0.021	INVEST RADIOL	0.808	J AM COLL CARDIOL	0.090	J COMPUT ASSIST TOMO	0.827
Clinical Laboratory	J AM COLL CARDIOL	0.040	CLIN EPIGENETICS	0.561	EUR HEART J	0.042	CLIN EPIGENETICS	0.609
	EUR HEART J	0.058	JOURNAL OF CLIN LAB ANALYSIS	0.461	CIRCULATION	0.061	HLA	0.464
	CIRCULATION	0.071	CLIN LAB	0.452	GASTROENTEROLOGY	0.047	CLIN CHIM ACTA	0.449
Nursing	INT J NURS STUD	0.777	NURS EDUC TODAY	0.891	INT J NURS STUD	0.832	J NURS MANAGE	0.906
	PLOS ONE	0.002	NURSE EDUC PRACT	0.890	BMC GERIATRICS	0.101	NURS EDUC TODAY	0.901
	EUR J CARDIOVASC NUR	0.396	J NURS MANAGE	0.857	PLOS ONE	0.004	INT NURS REV	0.893
Surgery	J AM COLL CARDIOL	0.141	J CRANIO MAXILL SURG	0.996	EUR HEART J	0.073	J LAPAROENDOSC ADV S	0.926
	EUR HEART J	0.087	EJSO	0.929	GASTROENTEROLOGY	0.113	J ARTHROPLASTY	0.889
	GASTROENTEROLOGY	0.114	J LAPAROENDOSC ADV S	0.903	J AM COLL CARDIOL	0.144	EJSO	0.879
Obstetrics/ gynecology	N ENGL J MED	0.043	GYNECOL ONCOL	0.826	N ENGL J MED	0.046	GYNECOL ONCOL	0.834
	J AM COLL CARDIOL	0.007	BMC WOMENS HEALTH	0.824	J AM COLL CARDIOL	0.010	PREGNANCY HYPERTENS	0.832
	CLIN CANCER RES	0.130	GYNECOL OBSTET INVES	0.809	CLIN CANCER RES	0.055	JOURNAL OF GYNECOL ONCOL	0.830
Ophthalmology	N ENGL J MED	0.009	ANNU REV VIS SCI	1.000	N ENGL J MED	0.013	ANNU REV VIS SCI	1.000
	ANN RHEUM DIS	0.013	J GLAUCOMA	0.870	ANN RHEUM DIS	0.010	J GLAUCOMA	0.864
	NAT COMMUN	0.005	J CATARACT REFR SURG	0.831	NAT COMMUN	0.006	OPHTHALMOLOGY	0.816

Table 3 (*continued*)

Table 3 (*continued*)

Disciplines	2016				2018			
	Top 3 journals (by IF)	Correlation coefficient	Top 3 journals (by correlation coefficient)	Correlation coefficient	Top 3 journals (by IF)	Correlation coefficient	Top 3 journals (by correlation coefficient)	Correlation coefficient
Otorhinolaryngology	J ALLERGY CLIN IMMUN	0.032	TRENDS HEAR	0.755	J ALLERGY CLIN IMMUN	0.032	AUDIOL NEURO-OTOL	0.814
	EUR RESPIR J	0.001	AUDIOL NEURO-OTOL	0.709	EUR RESPIR J	0.001	EAR HEARING	0.675
	CLIN CANCER RES	0.005	INT FORUM ALLERGY RH	0.688	CLIN CANCER RES	0.005	INT FORUM ALLERGY RH	0.673
Oncology	LANCET ONCOL	33.900	BREAST CANCER RES TR	0.889	LANCET ONCOL	0.527	BREAST CANCER RES TR	0.890
	J CLIN ONCOL	24.008	CLIN BREAST CANCER	0.878	J CLIN ONCOL	0.695	CLIN BREAST CANCER	0.838
	JAMA ONCOLOGY	20.871	CLIN LUNG CANCER	0.869	CANCER DISCOV	0.481	BONE MARROW TRANSPLV	0.828
Rehabilitation medicine	EUR HEART J	0.003	TOP STROKE REHABIL	0.929	EUR HEART J	0.013	TOP STROKE REHABIL	0.944
	SPORTS MED	0.547	J HEAD TRAUMA REHAB	0.847	BRIT J SPORT MED	0.447	INT J REHABIL RES	0.758
	BRIT J SPORT MED	0.422	REHABILITATION	0.807	SPORTS MED	0.519	EUR J PHYS REHAB MED	0.750
Sport sciences	EUR HEART J	0.003	J ORTHOP TRAUMA	0.541	EUR HEART J	0.005	J ORTHOP TRAUMA	0.591
	SPORTS MED	0.403	STRENGTH COND J	0.525	BRIT J SPORT MED	0.429	J AGING PHYS ACTIV	0.563
	BRIT J SPORT MED	0.373	J SPORT HEALTH SCI	0.524	SPORTS MED	0.439	ARTHROSCOPY	0.480
Anesthesiology	JAMA-J AM MED ASSOC	0.009	CLIN J PAIN	0.878	JAMA-J AM MED ASSOC	0.009	PAIN PRACTICE	0.879
	INTENS CARE MED	0.437	PEDIATR ANESTH	0.858	INTENS CARE MED	0.461	CLIN J PAIN	0.865
	CRIT CARE MED	0.109	SCHMERZ	0.840	COCHRANE DB SYST REV	0.030	PEDIATR ANESTH	0.856
Emergency medicine	N ENGL J MED	0.009	ACAD EMERG MED	0.789	N ENGL J MED	0.009	SCAND J TRAUMA RESUS	0.802
	JAMA-J AM MED ASSOC	0.019	EMERG MED CLIN N AM	0.780	JAMA-J AM MED ASSOC	0.019	EUR J EMERG MED	0.776
	CIRCULATION	0.006	SCAND J TRAUMA RESUS	0.743	CIRCULATION	0.006	ACAD EMERG MED	0.771

IF, impact factor.

　　　Ann Transl Med 2021;9(20):1580 | https://dx.doi.org/10.21037/atm-21-5046

judging productivity using IF alone. Furthermore, it can be concluded that, in the analysis of the variation tendencies of IFs and journal correlation coefficients, some disciplines exhibit a steady publication feature, which may inform further contributions for researchers to find a more suitable journal to publish their original articles.

Results of the normalized evaluation

For the fair evaluation of papers across different disciplines in clinical medicine, the normalized coefficient of each secondary discipline was calculated using the normalized evaluation method, represented by DNC_i. Here, the 18 disciplines are numbered from 1 to 18, and the specific discipline numbers are as follows: (I) general internal medicine; (II) pediatrics; (III) geriatrics and gerontology; (IV) neurosciences and neurology; (V) psychiatry; (VI) dermatology; (VII) radiology, nuclear medicine, and medical imaging; (VIII) clinical laboratory; (IX) nursing; (X) surgery; (XI) obstetrics and gynecology; (XII) ophthalmology; (XIII) otorhinolaryngology; (XIV) oncology; (XV) rehabilitation medicine; (XVI) sport sciences; (XVII) anesthesiology; and (XVIII) emergency medicine. As mentioned earlier, to achieve an approximate outline of a discipline that is sufficient for analysis, we focused on only the top 50 journals, including journals ranked in the top 50 using the multiple of each journal's IF and correlation coefficient. We used the product of the IF and the journal correlation coefficient for this sorting for 2 main reasons. First, some journals had extremely high IFs, but through extraction, we observed that these journals were relatively less correlated with the specific disciplines. Consequently, it would be difficult for scientific researchers to publish in such journals, which only published a few articles related to their disciplines. To minimize the gap between the IF and the journal's correlation with a given discipline, we used the normalized evaluation. The foundation of this method is using the multiples of 2 indexes for assessing the journals' overall characteristics.

Table 4 shows 2 indexes from 2016 to 2018: the weight of a given discipline, ρ_i, and the normalized coefficient of a specific discipline, DNC_i. The normalized coefficients of the disciplines differ from each other in the horizontal comparison, and these coefficients vary with time in the vertical comparison. In these 3 years, oncology had the highest normalized coefficient, which reflects the highest correlation between the characteristic journals

of the discipline. Additionally, it is easier for oncology scholars to publish high-quality SCI articles. In contrast, the normalized coefficients of otorhinolaryngology and emergency medicine are very low. It is difficult for scholars in these disciplines to publish articles with high IFs. Evidence of this can be found in the number of journals retrieved that are related to these disciplines, as mentioned earlier. During these 3 years, the normalized coefficient of each discipline generally showed a steady increase, whereas there was a U-shaped change for a low number of disciplines, reflecting the gradual increase of the specificity of these disciplines, where journals gradually began to focus on the characteristic research of the discipline.

Discussion

The results of our empirical study show that the disciplines of clinical medicine are not represented equally, and that the characteristics of the journals of each discipline vary substantially. As the findings of the present study only relate directly to the disciplines of clinical medicine, we cannot assume that they apply to other disciplines. However, although our analysis was restricted to these disciplines, the results obtained explore the variation of discipline features and indicate broader implications.

Advantages of the method's focus on clinical medicine

Compared with traditional IF evaluation, the application illustrated in the present study illustrates our method's results, and the focus on clinical medicine can facilitate the evaluation of scientific researchers in Chinese hospitals. The application of the combined methods proposed here has provided numerous types of information on each discipline, with a total of 702,996 journals retrieved in 2016 (although there is some overlap in journals across different disciplines). Firstly, the advantage of the innovative feature extraction method used in this study is that the most relevant journals in each discipline are obtained through computer dynamic recognition and retrieval. Among the 18 examined disciplines of clinical medicine, there are many that intersect with public health and basic medicine, and the examined disciplines also intersect with each other. The feature matching method can quickly select the journals that are directly related to a particular discipline from the vast number of SCI journals. Secondly, through further analysis of the selected journals, this method can identify

Table 4 The weight of a given discipline (ρ_i) and the normalized coefficient of a specific discipline (DNC_i) for 18 clinical medicine disciplines in 2016, 2017, and 2018

i	2016		2017		2018	
	ρ_i	DNC_i	ρ_i	DNC_i	ρ_i	DNC_i
1	3.125	0.092	3.357	0.096	3.446	0.104
2	1.590	0.047	1.629	0.046	1.670	0.051
3	1.702	0.050	1.731	0.049	1.672	0.051
4	2.949	0.087	2.969	0.085	3.078	0.093
5	3.424	0.101	3.440	0.098	3.515	0.107
6	1.330	0.039	1.427	0.041	1.528	0.046
7	1.533	0.045	1.516	0.043	1.810	0.055
8	0.946	0.028	0.988	0.028	0.949	0.029
9	1.024	0.030	1.055	0.030	1.195	0.036
10	2.868	0.085	2.887	0.082	3.026	0.092
11	1.793	0.053	1.724	0.049	1.762	0.053
12	1.760	0.052	1.873	0.053	1.945	0.059
13	0.787	0.023	0.816	0.023	0.857	0.026
14	4.493	0.133	4.775	0.136	4.829	0.146
15	1.180	0.035	1.264	0.036	1.312	0.040
16	0.950	0.028	1.030	0.029	1.130	0.034
17	1.652	0.049	1.796	0.051	1.840	0.056
18	0.715	0.021	0.774	0.022	0.881	0.027

the discipline's characteristics "hot" research areas and new research progress. Thirdly, also most importantly, through discipline selection and characteristics retrieval, we found the citation and research popularity differences among disciplines, then we performed the normalization method to further balance those disparities in disciplines during evaluations.

Potential application in the near future

The proposed method can be characterized as a widely applicable approach that, beyond being used to assess the scientific performance of researchers in hospitals, could be expanded for use in other disciplines, institutions, and databases. For example, this method is also applicable to universities, which also face multidisciplinary academic evaluations. The current university evaluation is dominated by IF, which is obviously unfair for less popular

subjects as well. From the perspective of researchers in a particular specialty, the adjusted evaluation system may at first seem a bit small because the adjustments it makes to the original IFs may be seen as a minor change. However, in the reality of the evaluation of individual scientists, substantial gaps do exist between different disciplines, and narrowing these gaps is of great value. The rapid growth of "big data" provides new opportunities and challenges for the field of bibliometrics (18). More logical and applicable approaches are needed to improve the evaluation of researchers, institutions, and publishing agencies, particularly in China.

Potential improvements for designing a more precise method

In the present study, we presented a theoretical model of feature extraction and normalization for the research

performance of 18 disciplines for a theoretical examination, see Waltman *et al.* (19); for an empirical study, see Waltman *et al.* (20). The method design can be further refined at each phase to better fit the demands of future investigations for wider disciplines, such as management or economics. Furthermore, more advanced techniques could be used to increase the accuracy of the delineation. For example, how can journals with minimal relevance be optimally filtered out to conserve resources? Such questions are worthy of further investigation.

Refinements that could be made to the method design to better fit the demands of wider acceptance and further application in other fields include the following.

Source selection

In the present study, we selected only journals from the Web of Science, because almost all clinical medical research is included in this database, which is becoming the core data source for assessment in the evaluation of researchers in clinical medicine. However, to construct an overview of other disciplines, the database selection needs to be changeable to include 1 or more appropriate databases to exhaust the disciplinary characteristics. As Frandsen and Nicolaisen noted, the results of bibliometric research based on data retrieved from databases are affected by differences in the database coverage (21). In the field of management, for example, it would be reasonable to change the focus to more related databases, such as the Social Sciences Citation Index (founded by the American Institute for Scientific Information) and the Chinese Social Sciences Citation Index (founded by the Chinese Social Science Research Evaluation Institute, covering the fields of law, management, economics, history, and political science).

Keyword selection

Selected articles could be included in a more refined way by excluding discipline-specific stop words in extracting information from abstracts or full text (22). Additionally, as proposed by Milojevic *et al.*, key phrases could be extracted from titles and transferred into a common list of key title words (23). Furthermore, although the full text of articles or abstracts is generally available for analyzing the characteristics of a discipline, more detailed criteria could be used instead of, or in addition to, the keyword list based on title words. For example, Waltman and van Eck proposed using natural language processing techniques to derive sets of terms characterizing fine-grained clusters of

publications (24).

Field weight setting

For the selection and subsequent processing of the field database, in the present study we chose a simple approach where all data fields for a certain discipline were treated equally. However, future work could explore whether fields with more dynamic differentiation might produce better results. In summary, different ways of setting field weights are some of the factors to consider in developing a more reasonable evaluation system.

Footnote

Conflicts of Interest: All authors have completed the ICMJE uniform disclosure form (available at https://dx.doi.org/10.21037/atm-21-5046). The authors have no conflicts of interest to declare.

Ethical Statement: The authors are accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved.

Open Access Statement: This is an Open Access article distributed in accordance with the Creative Commons Attribution-NonCommercial-NoDerivs 4.0 International License (CC BY-NC-ND 4.0), which permits the non-commercial replication and distribution of the article with

the strict proviso that no changes or edits are made and the original work is properly cited (including links to both the formal publication through the relevant DOI and the license). See: https://creativecommons.org/licenses/by-nc-nd/4.0/.

References

1. Costas R, van Leeuwen TN, van Raan AF. The "Mendel syndrome" in science: durability of scientific literature and its effects on bibliometric analysis of individual scientists. Scientometrics 2011;89:177-205.
2. Leydesdorff L. Alternatives to the journal impact factor: I3 and the top-10% (or top-25%?) of the most-highly cited papers. Scientometrics 2012;92:355-65.
3. Fraley RC, Vazire S. The N-pact factor: evaluating the quality of empirical journals with respect to sample size and statistical power. PLoS One 2014;9:e109019.
4. Comins JA, Hussey TW. Detecting seminal research contributions to the development and use of the global positioning system by reference publication year spectroscopy. Scientometrics 2015;104:575-80.
5. Budd JM. Measuring Research: What Everyone Needs to Know. College & Research Libraries 2018;79:853-4.
6. Hoeffel C. Journal impact factors. Allergy 1998;53:1225.
7. Garfield E. The history and meaning of the journal impact factor. JAMA 2006;295:90-3.
8. Bordons M, Zulueta MA. Comparison of research team activity in two biomedical fields. Scientometrics 1997;40:423-36.
9. Mishra S, Torvik VI. Quantifying Conceptual Novelty in the Biomedical Literature. Dlib Mag 2016. doi: 10.1045/september2016-mishra.
10. Wang J. Knowledge creation in collaboration networks: Effects of tie configuration. Research Policy 2016;45:68-80.
11. Leydesdorff L, Wagner CS, Bornmann L. Interdisciplinarity as diversity in citation patterns among journals: Rao-Stirling diversity, relative variety, and the Gini coefficient. Journal of Informetrics 2019;13:255-69.
12. van Leeuwen TN, Medina CC. Redefining the field of economics: Improving field normalization for the application of bibliometric techniques in the field of economics. Research Evaluation 2012;21:61-70.
13. Schmidt M, Sirtes D. Transforming the Heterogeneity of Subject Categories into a Stability Interval of the MNCS. Proceedings of Issi 2015 Istanbul: 15th International Society of Scientometrics and Informetrics Conference; 2015:365-71.
14. Leydesdorff L, Bornmann L. The operationalization of "fields" as WoS subject categories (WCs) in evaluative bibliometrics: The cases of "library and information science" and "science & technology studies". J Assoc Inf Sci Technol 2016;67:707-14.
15. Althouse BM, West JD, Bergstrom CT, et al. Differences in Impact Factor Across Fields and Over Time. J Am Soc Inf Sci Technol 2009;60:27-34.
16. Rehn C, Kronman U, Wadskog D. Bibliometric indicators – Definitions and usage at Karolinska Institutet. Stickholm, Sweden: Karolinska Institutet University Library; 2014.
17. Wagner H, Finkenzeller T, Würth S, et al. Individual and team performance in team-handball: a review. J Sports Sci Med 2014;13:808-16.
18. Ding Y, Zhang G, Chambers T, et al. Content-based citation analysis: The next generation of citation analysis. J Assoc Inf Sci Technol 2014;65:1820-33.
19. Waltman L, van Eck NJ, van Leeuwen TN, et al. Towards a new crown indicator: Some theoretical considerations. J Informetr 2011;5:37-47.
20. Waltman L, van Eck NJ, van Leeuwen TN, et al. Towards a new crown indicator: an empirical analysis. Scientometrics 2011;87:467-81.
21. Frandsen TF, Nicolaisen J. Intradisciplinary differences in database coverage and the consequences for bibliometric research. J Am Soc Inf Sci Technol 2008;59:1570-81.
22. Makrehchi M, Kamel MS. Automatic extraction of domain-specific stopwords from labeled documents. Advances in Information Retrieval 2008;4956:222-33.
23. Milojevic S, Sugimoto CR, Yan EJ, et al. The Cognitive Structure of Library and Information Science: Analysis of Article Title Words. J Am Soc Inf Sci Technol 2011;62:1933-53.
24. Waltman L, van Eck NJ. A new methodology for constructing a publication-level classification system of science. J Am Soc Inf Sci Technol 2012;63:2378-92.

(English Language Editor: R. Scott)