The utilization of paper-level classification system on the evaluation of journal impact

Zhesi Shen\textsuperscript{a,}\textsuperscript{*}, Sichao Tong\textsuperscript{a}, Fuyou Chen\textsuperscript{a}, Liying Yang\textsuperscript{a}

\textsuperscript{a}National Science Library, Chinese Academy of Sciences, Beijing 100190, P. R. China

\textbf{Abstract}

CAS Journal Ranking, a ranking system of journals based on the bibliometric indicator of citation impact, has been widely used in meso and macro-scale research evaluation in China since its first release in 2004. The ranking’s coverage is journals which contained in the Clarivate’s Journal Citation Reports (JCR). This paper will mainly introduce the upgraded version of the 2019 CAS journal ranking. Aiming at limitations around the indicator and classification system utilized in earlier editions, also the problem of journals’ interdisciplinarity or multidisciplinarity, we will discuss the improvements in the 2019 upgraded version of CAS journal ranking (1) the CWTS paper-level classification system, a more fine-grained system, has been utilized, (2) a new indicator, Field Normalized Citation Success Index (FNCSI), which ia robust against not only extremely highly cited publications, but also the wrongly assigned document type, has been used, and (3) the calculation of the indicator is from a paper-level. In addition, this paper will present a small part of ranking results and an interpretation of the robustness of the new FNCSI indicator. By exploring more sophisticated methods and indicators, like the CWTS paper-level classification system and the new FNCSI indicator, CAS Journal Ranking will continue its original purpose for responsible research evaluation.

\textit{Keywords:} Journal ranking, Field normalization, Citation Success Index
1. Introduction

1.1. History of CAS Journal Ranking

The CAS journal ranking, an annually released journal ranking by the Center of Scientometrics (CoS), National Science Library of Chinese Academy of Sciences (CAS), is a journal ranking widely used in China. It ranks journals contained in the Clarivate’s Journal Citation Reports (JCR), based on bibliometrics data. We’ll sketch out its history and mainly introduce the upgraded version of the 2019 CAS journal ranking, which we firstly utilize the CWTS paper-level classification system and a new indicator, the Field Normalized Citation Success Index (FNCSI).

The non-Field Normalized impact factor (JIF), which has been widely used as a journal indicator, performs differently in different research domains. Around the year 2000, in practical administrative work, the CoS research group has gradually identified that the impact factor was in a misused situation in most cases at that time in China. Aiming to compare or analyze journals separately in different scientific domains, the CoS research group released the first edition of CAS journal ranking in 2004, becoming popularly used in China. Journals can be grouped by subject area (major areas developed from degree classification by Degree Office of the Ministry of Education of the People’s Republic of China), subject category (the same specific subject categories developed from the JCR journal subject categories in the Web of Science database).

CAS journal ranking has been applied in many cases, varying from supporting related scientific policy-making of institutions to providing journals’ information to researchers. For the institutional level, they can know the performance of scientific output via drawing their distributions in CAS journal ranking, this information can help them when making related policies. Among the cash-per-publication reward policies in China, CAS journal ranking plays a dominant role. Chinese universities usually reward researchers for scientific output, motivating scientific research. Quan et al. (2017) analyze 168 reward policies in China, and they find that there is an increasing trend of adopting CAS journal ranking in Chinese universities from 2005, after the first edition of CAS journal ranking was released. And there are 99 of reward policies taking CAS journal ranking as the reference by 2016. For researchers, CAS journal ranking can help them know journals of targeted fields, from a relatively comprehensive view, when submitting their research output. Additionally, some journals utilize CAS journal ranking as the source of information, about themselves and other journals.
1.2. Limitations of old CAS Journal Ranking

A limitation relates to indicator exists in old CAS journal ranking. For a journal, the citation distributions are skewed, and JIF can be vastly affected by the tail of highly-cited papers. We previously utilize a three-year averaged JIF to alleviate such fluctuation. However, it is still not robust enough against occasionally highly-cited papers.

The second limitation is that the journal classification system used in the old CAS journal ranking is not fine-grained. Regarding citation practices, Garfield (1979) proposes the citation potential which can be defined as the probability of being cited, perform significantly differently in different fields, and we previously use the JCR journal subject categories in the Web of Science database. However, it is still not fine-grained, differences in citation also exist within fields (e.g., citation performs differ between different areas within a medical field in the study by van Eck et al. (2013), based on the subject categories in the WoS database, which we use in old CAS journal ranking).

We plot a science map for journals from all fields (please see Figure 1) with each dot representing a journal and the color representing potential citation. The layout of this map is used in an earlier paper (Shen et al., 2019) based on journal citation network. Here we use journal’s expected JIF as an indicator for potential citation, the detailed formula can be found in the data and method section. The color of each dot is related to the value of the corresponding journal’s expected JIF: the more red/blue the color is, the larger/smaller the value is. Figure 1 indicates a clear distinction between the potential citation between different research fields. We can see the phenomenon of citation performs differ between different areas exist within not only the above studied medical fields but also many other fields, for example, the upper part and the lower part of the Math category obviously perform differently.

We then take journals from JCR category: Statistics & Probability as an example. Looking at Figure 2, each dot represent a journal, we color journals titled with probability in blue, and in general, most blue dots have smaller expected JIF, indicating that distinction of citation potential probably exists between journals from different topic, within Statistics & Probability category, e.g., Probability related journals perform more weakly in citation potential.

A third limitation is typically related to journals’ interdisciplinarity or multidisciplinarity. In addition to multidisciplinary scopes included in more journals from a general view, research topic can span across established disciplines (Leydesdorff, 2007), bringing benefits and challenges, especially in journal impact
Figure 1: Map of scientific journals with expected JIF.

Figure 2: Correlation of JIF and expected JIF for journals in Statistics and Probability category.
studies. Utilizing a more fine-grained classification system and more sophisticated indicators can partly be a solution to this phenomenon. For example, Nature Communications and Science Advance, two famous open access multidisciplinary journals, having similar Journal Impact Factor (JIF), but their amount and distribution of covered topics are quite different. Similar situations will also happen to specialized journals. Also, some research journals will publish a far greater proportion of reviews than others, usually leading to high JIF, this is a forth limitation.

For these limitations above, we make improvements in the upgraded version of the 2019 CAS journal ranking, which has been firstly released in January 2020 on the official website[1]. Refinements in this release include the followings:

• The CWTS paper-level classification system, a more fine-grained system, has been utilized to address the above classification system related problem and journals’ interdisciplinarity related problem.

• Instead of JIF, a new indicator, Field Normalized Citation Success Index (FNCSI), has been used in the upgraded version. On the insensitivity side, compared with other citation impact factors, e.g., the three-year average JIF utilized in earlier editions, it excels no merely in the robustness of the occasional ultra-small number of extremely highly cited publications, but also in the robustness against the wrongly assigned document type.

• In addition, from a paper-level instead of journal-level, we calculate the indicator within article/review type papers.

More detailed information about the above refinements will be discussed later in this paper. Data and Methods section will introduce data coverage, CWTS paper-level classification system and the indicator utilized in the upgraded version of the 2019 CAS journal ranking. Results section includes a small part of the CAS journal ranking result and interpretation regarding the advantage of FNCSI. We finally discuss that attention should be paid on how to use CAS journal ranking appropriately for responsible research evaluation. Ongoing work and future plans will also be discussed.

1www.fenqubiao.com
2. Method and Data

2.1. Journals and citation data

The CAS journal ranking includes the journals which contained in the Clarivate’ Journal Citation Reports (JCR) (JCR2018, 2019). For journals’ citations data, we use Journal Impact Factor contributing items, which released by Clarivate’ Journal Citation Reports. This contains citations in year Y of each article and review, published in years Y-1 and Y-2, which counted towards the journal’s impact factor.

2.2. Paper-level classification data

The data utilized in the CWTS paper-level classification was collected from Clarivate’ Web of Science database, with the document types article and review, which were published between 2000 and 2018, and this classification system only included publications from the SCI and SSCI database. For the details of constructing the CWTS paper-level classification system, we refer to Waltman & van Eck (2012, 2013a) for a more detailed introduction of the classification methods from exploring the relatedness of publications to clustering publications into groups. This classification system consists of three levels - macro, meso, and micro levels - according to different granularity. Here we use the micro-level with about 4,000 clusters. It should be noted that, in the released CWTS paper-level classification data, publications from trade journals and several local journals are excluded, i.e., these journals cannot be evaluated. Here we try to include as many journals as possible, thus for these unclassified publications, we retrieve their related records from WoS and put them into corresponding clusters based on the clusters of the retrieved related records using the majority rule. In total, 99% of publications reported in JCR are included for calculation and 98% of journals having more than 90% of their total publications are included.

2.3. Journal Ranking Indicators

In CAS Journal Ranking 2019, we follow the idea of Citation Success Index (CSI) and extend it to a field normalized version. The original CSI presented to compare the citation capacity between two journals (Stringer et al., 2008, Milojević et al., 2017, Shen et al., 2018), is defined as the probability of a randomly selected paper from one journal having more citations than a randomly selected paper from the other journal. Following the same idea, we propose the Field Normalized Citation Success Index (FNCSI). The FNCSI is defined as the probability that the citation of a paper from journal A is larger than a random paper in the same topics.
and with the same document type from other journals. For the details please refer to the section below. For comparison, we also consider the Field Normalized Impact Factor (FNIF).

2.3.1. Field Normalized Citation Success Index (FNCSI)

For journal A, the probability that the citation of a paper from journal A is larger than a random paper in the same topics and document type from other journals, is defined as below:

$$S_A = P(c_a > c_o | a \in A, o \in O) = \sum_{t,d} P(A^{t,d})P(c_a > c_o | a \in A^{t,d}, o \in O^{t,d})$$ (1)

For a specific research topic t, its FNCSI is defined as below:

$$S'_A = \frac{1}{N_A} \sum_t N_{A^{t,d}} \left[ \frac{\sum_{a \in A^{t,d}, o \in O^{t,d}} 1(c_a > c_o) + \sum_{a \in A^{t,d}, o \in O^{t,d}} 0.5(c_a = c_o)}{N_{A^{t,d}}N_{O^{t,d}}} \right]$$ (2)

Journal A usually involves several research topics from the micro level of the system, then the total FNCSI of Journal A can be summed from its involved topics as below:

$$S_A = \frac{1}{N_A} \sum_t N_{A^{t,d}} S'_A$$ (3)

where $t \in \text{\{topic}_1, \text{topic}_2, \text{topic}_3, ..., \text{\}}$, $d \in \text{\{article, review\}}$, $A^{t,d}$ represents the publications clustered in topic t with document type d in journal A.

2.3.2. Field Normalized Impact Factor (FNIF)

Field Normalized Impact Factor (FNIF) uses the same classification system as FNCSI but uses the commonly used average citation based normalization approach, i.e., each citation is normalized by the average citation of papers in the same topic cluster and with the same document type. For instance, the FNIF of journal A is defined as:

$$F_A = \frac{\sum_{t,d} \sum_{a \in A^{t,d}} c_a / \mu_{t,d}}{N_A}$$ (4)

where $\mu_{t,d}$ is the average citation of papers in topic t with document type d. By comparing the results of FNCSI and FNIF, we can see the advantages of CSI.
2.3.3. Expected JIF

As we mentioned earlier, for each journal, we use expected JIF as an indicator of potential citation:

\[ E_A = \frac{\sum_t \mu_t N_A^t}{N_A} \]  

(5)

where \( \mu_t \) is the average citation of papers in topic \( t \).

3. Results

3.1. Ranking Results

In this section we present the results of CAS Journal Ranking based on FNCSI and the comparisons with other indicators. Table 1 shows the top 20 ranked journals according to FNCSI. Here we only list journals mainly publishing research articles. The top five journals are well-acknowledged in natural and life science. The rest journals belong to different fields and not concentrate on a single field or narrow fields. If we take a look at the publishers of these journals, we can see that this list is dominated by Nature-titled journals, Lancet-titled journals and Cell-titled journals.

The corresponding rankings based on journals’ FNIF values of these top 20 journals are also presented Table 1. Among these journals, the rankings of Cancer Cell, Nature Neuroscience, Cell Metabolism and Nature Immunology are boosted most from the FNCSI indicator, they all climb more than 20 positions. Only Lancet Oncology shows a slight drop in position from the FNCSI indicator. Overall, Journals from medical-related categories mostly have a relatively big gap between these indicators. In Appendix Table 4 we present the top 20 journals both for FNCSI and FNIF.

The correlation among these journal citation indicators are shown in Figure 3, we can see that FNCSI and FNIF are highly correlated (spearman correlation: 0.98, p-value: 0.0). In the lower part of Figure 3 we highlight several journals that having worse rankings in FNCSI compared with FNIF. These journals share a common property that they each have one or several highly cited papers and a majority of poorly cited papers, e.g., Chinese Phys C has one paper cited more than 2000 times but about 70% papers are zero cited (JCR2018, 2019).

Earlier in this article, we discuss the difference of citation potential exists between journals from different topics, within the Statistics & Probability category. Here in Table 2, we give the top 20 ranked journals (which mainly publishing research articles) according to FNCSI in this category. And to some extent, we
Table 1: Top 20 ranked journals according to FNCSI.

| Journal          | Category-WoS                                                                 | FNCSI | FNIF |
|------------------|------------------------------------------------------------------------------|-------|------|
| LANCET           | MEDICINE, GENERAL & INTERNAL                                                 | 1     | 3    |
| NATURE           | MULTIDISCIPLINARY SCIENCES                                                   | 2     | 5    |
| JAMA             | MEDICINE, GENERAL & INTERNAL                                                 | 3     | 4    |
| SCIENCE          | MULTIDISCIPLINARY SCIENCES                                                   | 4     | 9    |
| CELL             | BIOCHEMISTRY & MOLECULAR BIOLOGY/CELL BIOLOGY                               | 5     | 15   |
| WORLD PSYCHIATRY | PSYCHIATRY                                                                   | 6     | 8    |
| LANCET NEUROL    | CLINICAL NEUROLOGY                                                           | 7     | 11   |
| NAT PHOTONICS    | OPTICS/PHYSICS, APPLIED                                                      | 8     | 17   |
| NAT GENET        | GENETICS & HEREDITY                                                          | 9     | 13   |
| NAT MED          | BIOCHEMISTRY & MOLECULAR BIOLOGY/CELL BIOLOGY/MEDICINE, RESEARCH & EXPERIMENTAL | 10    | 21   |
| NAT MATER        | MATERIALS SCIENCE, MULTIDISCIPLINARY/CHEMISTRY, PHYSICAL/PHYSICS, APPLIED/PHYSICS, CONDENSED MATTER | 11    | 12   |
| LANCET ONCOL     | ONCOLOGY                                                                     | 12    | 10   |
| CANCER CELL      | ONCOLOGY/CELL BIOLOGY                                                        | 13    | 38   |
| NAT CHEM         | CHEMISTRY, MULTIDISCIPLINARY                                                 | 14    | 31   |
| NAT NEUROSCI     | NEUROSCIENCES                                                                | 15    | 36   |
| CELL METAB       | CELL BIOLOGY/ENDOCRINOLOGY & METABOLISM                                      | 16    | 51   |
| LANCET RESP MED  | CRITICAL CARE MEDICINE/RESPIRATORY SYSTEM                                    | 17    | 22   |
| NAT IMMUNOL      | IMMUNOLOGY                                                                   | 18    | 58   |
| LANCET DIABETES ENDO | ENDOCRINOLOGY & METABOLISM                                            | 19    | 27   |
| NAT NANOTECHNOL  | NANOSCIENCE & NANOTECHNOLOGY/MATERIALS SCIENCE, MULTIDISCIPLINARY            | 20    | 23   |
can find that journals perform weakly in citation potential have been revealed by FNCSI, such as several well-acknowledged journals like Annals of Statistics, Annals of Probability and Biometrika.

3.2. Robustness

3.2.1. Robust against extremely highly cited publications

The robustness of an indicator represents its sensitivity to changes in the set of publications based on which it is calculated. A robust indicator will not change a lot against the occasional ultra-small number of highly cited publications. To measure the robustness of an indicator we construct several sets of publications for each journal with bootstrapping method and recalculate the indicator and rankings accordingly. For instance, for a journal with \( N \) publications, we randomly selected \( N \) publications with replacement, calculate these indicators, and get a new ranking for each journal. We simulate this procedure for 100 times and obtain 100 rankings for each journal. Figure 4(a) shows the distribution of the obtained rankings of Chinese Physics: C. We can see that the range of ranking from FNCSI varies much less than FNIF. The citation distribution of Chinese Physics: C is highly skewed, with one paper cited about two thousand times and about 70% papers not
Table 2: Top 20 ranked journals in Statistics and Probability category according to FNCSI

| Journal                      | Rank-FNCSI | Rank-Expected JIF | Rank-JIF |
|------------------------------|------------|-------------------|----------|
| ECONOMETRICA                | 1          | 69                | 2        |
| J R STAT SOC B              | 2          | 45                | 3        |
| ANN STAT                    | 3          | 63                | 7        |
| PROBAB THEORY REL           | 4          | 86                | 10       |
| ANN PROBAB                  | 5          | 103               | 15       |
| FINANC STOCH                | 6          | 99                | 22       |
| J AM STAT ASSOC             | 7          | 32                | 4        |
| INT STAT REV                | 8          | 39                | 16       |
| J QUAL TECHNOL              | 9          | 104               | 29       |
| J STAT SOFTW                | 10         | 20                | 1        |
| ANN APPL PROBAB             | 11         | 60                | 28       |
| STOCH ENV RES RISK A        | 12         | 7                 | 8        |
| BRIT J MATH STAT PSY        | 13         | 9                 | 20       |
| TECHNOMETRICS               | 14         | 56                | 21       |
| BIOMETRIKA                  | 15         | 49                | 33       |
| BAYESIAN ANAL               | 16         | 44                | 35       |
| BERNOULLI                   | 17         | 92                | 41       |
| INSUR MATH ECON             | 18         | 65                | 43       |
| EXTREMES                     | 19         | 58                | 25       |
| ECONOMET THEOR              | 20         | 91                | 52       |
To get an overview of the indicators’ robustness, we calculate the relative change of rankings for these indicators. The relative change of ranking is defined as:

$$\Delta = \frac{1}{N} \sum_{j}^{N} \frac{\max\{R_j\} - \min\{R_j\}}{\text{avg}\{R_j\}}$$

where \(\{R_j\}\) is the rankings of journal j obtained from the above simulation. As shown in Fig. 4(b), the relative change of FNCSI is smaller than FNIF implying that FNCSI is more robust than FNIF as FNCSI mainly focus on the central tendency of the citation distribution and is not easily affected by occasional highly cited papers.

3.2.2. Robust against Document Type

Citation patterns are expected to vary a lot across different document types (Price, 1965). When conducting the field normalization, we also consider the document type, thus wrongly assigned document types will affect the journals’ indicators and rankings. To test the sensitivity of indicators against wrongly labeled document types, here we generate a virtual dataset:

- for each journal, we turn its most highly cited paper to the opposite, i.e., Article to Review or Review to Article,
and then we recalculate the journal indicators and obtain the new rankings based on FNCSI and FNIF respectively. The comparison of rankings based on this changed data with the original rankings is shown in Fig. 5. We can see that almost all the orange dots (FNCSI-based) locate closely along the diagonal line while the blue squares (FNIF-based) spread much broader which implying that rankings based on FNCSI are more robust against wrongly labeled document type than rankings based on FNIF.

4. Conclusion and Discussion

In this paper, we briefly describe the CAS Journal Ranking’s history and its practical applications by Chinese universities and institutes in rewarding, promotion and research performance monitoring. We also discuss a number of limitations in earlier editions of the CAS Journal Ranking, and our exploration of solving these problems. To better solve these problems we introduce the new indicator - Field Normalized Citation Success Index - which is used in the CAS Journal Ranking 2019 upgraded version. The FNCSI extends the idea of CSI and uses a fine-grained paper-level classification system to eliminate the citation difference among fields. We also consider the difference citation potential between articles.
and reviews in normalization. A detailed comparison between FNCSI and FNIF indicating that the ranking result obtained from FNCSI is favorable and is robust against extremely highly-cited publications and wrongly assigned document type.

We need to point out, towards to one of the important issue that evaluating citation performance fairly between different research fields, some contributed work has been done from the source-side, which is originated from Zitt & Small (2008), to solve the field normalized issue, including the source normalized impact per paper (SNIP) indicator (Moed 2010), the revised SNIP indicator (Waltman et al. 2013). Comparisons and discussions between the source(citing)-side approach and cited-side approach have been done by Waltman & van Eck (2013b), Waltman & Eck (2013), Ruiz-Castillo (2014), and still have been inconclusive, here we refer to the overviews of these discussions provided by Waltman (2016) and Glänzel et al. (2019). We also plan to do an empirical comparison between these indicators. Besides, as previously mentioned about limitations in the earlier editions of CAS journal ranking, with respect to occasionally highly-cited papers, the revised SNIP indicator has the same problem. Lehmann & Wohlrabe (2017) give an example of the journal Advances in Physics which fluctuates significantly across time based on SNIP and they tried to address this problem by adopting the Elo rating system which takes journals’ historical performance into consideration.

In addition, we have an ongoing exploration of providing journal profiles which will provide more detailed information about journals’ covered topics and facilitate the comparison of journals on a target topic. This journal profile module will be added to the CAS journal ranking in future editions.

Around 1990, China started launching a reward policy to encourage Chinese scholars to join the international research community and publish papers in international journals, mainly the WoS-indexed papers (Peng 2011). Till now, Chinese institutions all have their own reward policy (Quan et al. 2017), and these policy which mostly reference CAS Journal Ranking, has indeed succeeded in promoting China’s international scientific publications in the past period. CAS journal ranking truly promotes understanding more about journals for Chinese policymakers and researchers. However at the same time, we are aware of the inappropriate employ that comes along also has a negative impact as indicators’ function may easily be warped in practical evaluation, even becoming a driving force of research (Campbell 1979, Wouters et al. 2019). We here especially notice its misuse in evaluating individual research, like in those cash reward policies which have been analyzed in earlier study (Quan et al. 2017), most of them take CAS journal ranking, or other bibliometric indicators, as the golden rule instead of as a reference or supporting measures. We here call on any practice of using journal
indicators should meet the criteria proposed by Wouters et al. (2019):

- "Justified. Journal indicators should have only a minor and explicitly defined role in assessing the research done by individuals or institutions (McKenna et al. 2019)."

- Contextualized. In addition to numerical statistics, indicators should report statistical distributions (for example, of article citation counts), as has been done in the Journal Citation Reports since 2018 (Larivire, 2016). Differences across disciplines should be considered.

- Informed. Professional societies and relevant specialists should help to foster literacy and knowledge about indicators. For example, a PhD training course could include a role-playing game to demonstrate the use and abuse of journal indicators in career assessment.

- Responsible. All stakeholders need to be alert to how the use of indicators affects the behaviour of researchers and other stakeholders. Irresponsible uses should be called out.”

Following these criteria, we, the CoS research group, will continue our original purpose for responsible research evaluation, exploring more sophisticated methods and indicators, constantly improving the science of CAS Journal Ranking.

Acknowledgements

The authors thank Dr. Nees J van Eck and CWTS for providing the paper-level classification data, and thank Ms. M. Zhu for valuable discussion.

Author contribution

Conceptualization: SZ, YL
Data Curation: CF, SZ
Formal analysis: SZ, TS
Methodology: SZ, YL, TS
Writing original draft: SZ, TS
Writing review & editing: SZ, YL, SF, CF
Supervision: YL
References

Campbell, D. T. (1979). Assessing the impact of planned social change. *Evaluation and Program Planning, 2*, 67–90.

van Eck, N. J., Waltman, L., van Raan, A. F. J., Klautz, R. J. M., & Peul, W. C. (2013). Citation analysis may severely underestimate the impact of clinical research as compared to basic research. *PLOS ONE, 8*.

Garfield, E. (1979). *Citation Indexing - Its Theory And Application In Science, Technology, And Humanities.*

Gnizel, W., Moed, H. F., Schmoch, U., & Thelwall, M. (2019). Springer handbook of science and technology indicators. . JCR2018 (2019). 2018 journal impact factor, journal citation reports (clarivate analytics, 2019). .

Larivire, V. (2016). A simple proposal for the publication of journal citation distributions. *bioRxiv,* (p. 62109).

Lehmann, R., & Wohlrabe, K. (2017). Who is the 'journal grand master'? a new ranking based on the elo rating system. *Journal of Informetrics, 11*, 800–809. doi:[10.1016/j.joi.2017.05.004](https://doi.org/10.1016/j.joi.2017.05.004)

Leydesdorff, L. (2007). Betweenness centrality as an indicator of the interdisciplinarity of scientific journals. *Journal of the Association for Information Science and Technology, 58*, 1303–1319.

McKiernan, E. C., Schimanski, L. A., Nieves, C. M., Matthias, L., Niles, M. T., & Alperin, J. P. (2019). Use of the journal impact factor in academic review, promotion, and tenure evaluations. *eLife, 8*.

Milojevi, S., Radicchi, F., & Bar-Ilan, J. (2017). Citation success index an intuitive pair-wise journal comparison metric. *Journal of Informetrics, 11*, 223–231.

Moed, H. F. (2010). Measuring contextual citation impact of scientific journals. *Journal of Informetrics, 4*, 265–277. doi:[10.1016/j.joi.2010.01.002](https://doi.org/10.1016/j.joi.2010.01.002)
Peng, C. (2011). Focus on quality, not just quantity. *Nature, 475*, 267–267.

Quan, W., Chen, B., & Shu, F. (2017). Publish or impoverish: An investigation of the monetary reward system of science in China (1999-2016). *Aslib Journal of Information Management, 69*, 486–502.

Ruiz-Castillo, J. (2014). The comparison of classification-system-based normalization procedures with source normalization alternatives in Waltman and van Eck (2013). *Journal of Informetrics, 8*, 25–28.

Shen, Z., Chen, F., Yang, L., & Wu, J. (2019). Node2vec representation for clustering journals and as a possible measure of diversity. *Journal of Data and Information Science, 4*, 79–92.

Shen, Z., Yang, L., & Wu, J. (2018). Lognormal distribution of citation counts is the reason for the relation between Impact Factors and Citation Success Index. *Journal of Informetrics, 12*, 153–157. doi:[10.1016/j.joi.2017.12.007]

Stringer, M. J., Sales-Pardo, M., & Amaral, L. A. N. (2008). Effectiveness of journal ranking schemes as a tool for locating information. *PLoS One, 3*, e1683.

Waltman, L. (2016). A review of the literature on citation impact indicators. *Journal of Informetrics, 10*, 365–391. doi:[10.1016/j.joi.2016.02.007]

Waltman, L., & van Eck, N. J. (2012). A new methodology for constructing a publication-level classification system of science. *Journal of the Association for Information Science and Technology, 63*, 2378–2392.

Waltman, L., & van Eck, N. J. (2013a). A smart local moving algorithm for large-scale modularity-based community detection. *European Physical Journal B, 86*, 471.

Waltman, L., & Eck, N. J. (2013). Source normalized indicators of citation impact: an overview of different approaches and an empirical comparison. *Scientometrics, 96*, 699–716.

Waltman, L., & van Eck, N. J. (2013b). A systematic empirical comparison of different approaches for normalizing citation impact indicators. *Journal of Informetrics, 7*, 833–849. doi:[10.1016/j.joi.2013.08.002]
Waltman, L., van Eck, N. J., van Leeuwen, T. N., & Visser, M. S. (2013). Some modifications to the snip journal impact indicator. *Journal of Informetrics, 7*, 272–285. doi:10.1016/j.joi.2012.11.011.

Wouters, P., Sugimoto, C. R., Larivire, V., McVeigh, M. E., Pulverer, B., de Rijcke, S., & Waltman, L. (2019). Rethinking impact factors: better ways to judge a journal. *Nature, 569*, 621–623.

Zitt, M., & Small, H. (2008). Modifying the journal impact factor by fractional citation weighting: The audience factor. *Journal of the Association for Information Science and Technology, 59*, 1856–1860.

**Appendix**

*Appendix A. Top 20 ranked research journals*

In Table 4, we list the top 20 ranked research journals based on FNCSI and FNIF respectively. Compared with the journals of selected according to FNCSI, the top four journals via FNIF are all medical-related.

*Appendix B. Additional results on robust comparison between FNCSI and FNIF*

In this section, we present some additional results on the robustness of the proposed journal indicators. In Figure 4(b) we have illustrated the relative change of rankings based on FNCSI and FNIF, here we demonstrate some further analysis and results. In Figure 6 we compare the 1st quartile and 3rd quartile rankings obtained from the 100 simulations for each journal. The x-axis is the 1st quartile and the y-axis is the 3rd quartile. We can see for both FNCSI and FNIF, the dots mainly located along the diagonal line implying that the rankings of most journals are stable. When comparing the orange dots (FNCSI) and blue squared (FNIF), we can see the spreading area of orange dots is smaller than the blue squares indicating that rankings based on FNCSI are more stable than rankings based on FNIF when dealing with some special journals.

Journal indicators should also be stable across time as a journal’s reputation and quality will not change dramatically. In Figure 7 we present the evolution of rankings based on JIF, FNIF and FNCSI for the journal *J Math Sociol*. We can see the rankings of JIF and FNIF show a big jump in the year 2018 compared with its rankings in previous years. However, the ranking of FNCSI only increases a little. Here because of data availability, we only calculated the indicators for 2017 and 2018, we will continue to monitor this journal’s performance in 2019 and forthcoming years.
Table 3: Top 20 journals based on FNCSI and FNIF respectively.

| Journal                        | FNCSI | Journal            | FNIF |
|--------------------------------|-------|--------------------|------|
| LANCET                         | 1     | CA-CANCER J CLIN   | 1    |
| NATURE                         | 2     | NEW ENGL J MED     | 2    |
| JAMA-J AM MED ASSOC            | 3     | LANCET             | 3    |
| SCIENCE                        | 4     | JAMA-J AM MED ASSOC| 4    |
| CELL                           | 5     | NATURE             | 5    |
| WORLD PSYCHIATRY               | 6     | PSYCHOL SCI PUBL INT| 6   |
| LANCET NEUROL                  | 7     | Q J ECON           | 7    |
| NAT PHOTONICS                  | 8     | WORLD PSYCHIATRY   | 8    |
| NAT GENET                      | 9     | SCIENCE            | 9    |
| NAT MED                        | 10    | LANCET ONCOL       | 10   |
| NAT MATER                      | 11    | LANCET NEUROL      | 11   |
| LANCET ONCOL                   | 12    | NAT MATER          | 12   |
| CANCER CELL                    | 13    | NAT GENET          | 13   |
| NAT CHEM                       | 14    | PSYCHOL BULL       | 14   |
| NAT NEUROSCI                   | 15    | CELL               | 15   |
| CELL METAB                     | 16    | NAT ENERGY         | 16   |
| LANCET RESP MED                | 17    | NAT PHOTONICS      | 17   |
| NAT IMMUNOL                    | 18    | CIRCULATION        | 18   |
| LANCET DIABETES ENDO           | 19    | FUNGAL DIVERS      | 19   |
| NAT NANOTECHNOL                | 20    | LANCET INFECT DIS  | 20   |
Figure 6: Change of rankings based on FNCSI and FNIF.

Figure 7: Evolution of percentile rank for *J Math Sociol* based on different indicators. The percentile ranking is calculated within the *Mathematics, Interdisciplinary applications* category.