



Check for updates

BRIEF REPORT

Journal- or article-based citation measure? A study of academic promotion at a Swiss university [version 1; peer review: 2 approved, 1 approved with reservations]

Nicole Steck¹, Lukas Stalder^{id}2, Matthias Egger^{1,3}

¹Institute of Social and Preventive Medicine, University of Bern, Bern, Switzerland

²Dean's office, Medical Faculty, University of Bern, Bern, Switzerland

³Department of Population Health Sciences, Bristol Medical School, University of Bristol, Bristol, UK

V1 First published: 01 Oct 2020, 9:1188
<https://doi.org/10.12688/f1000research.26579.1>
Latest published: 01 Oct 2020, 9:1188
<https://doi.org/10.12688/f1000research.26579.1>

Abstract

In academia, decisions on promotions are influenced by the citation impact of the works published by the candidates. The Medical Faculty of the University of Bern used a measure based on the journal impact factor (JIF) for this purpose: the JIF of the papers submitted for promotion should rank in the upper third of journals in the relevant discipline (JIF rank >0.66). The San Francisco Declaration on Research Assessment (DORA) aims to eliminate the use of journal-based metrics in academic promotion. We examined whether the JIF rank could be replaced with the relative citation ratio (RCR), an article-level measure of citation impact developed by the National Institutes of Health (NIH). An RCR percentile >0.66 corresponds to the upper third of citation impact of articles from NIH-sponsored research. We examined 1525 publications submitted by 64 candidates for academic promotion at University of Bern. There was only a moderate correlation between the JIF rank and RCR percentile (Pearson correlation coefficient 0.34, 95% CI 0.29-0.38). Among the 1,199 articles (78.6%) published in journals ranking >0.66 for the JIF, less than half (509, 42.5%) were in the upper third of the RCR percentile. Conversely, among the 326 articles published in journals ranking <0.66 regarding the JIF, 72 (22.1%) ranked in the upper third of the RCR percentile. Our study demonstrates that the rank of the JIF is a bad proxy measure for the actual citation impact of individual articles. The Medical Faculty of University of Bern has signed DORA and replaced the JIF rank with the RCR percentile to assess the citation impact of papers submitted for academic promotion.

Keywords

Relative Citation Ratio, Journal Impact Factor, DORA Declaration

Open Peer Review

Reviewer Status

Invited Reviewers			
	1	2	3
version 1			
01 Oct 2020	report	report	report

1. **Sandra L. Schmid** ^{id}, Chan Zuckerberg Biohub, San Francisco, USA
University of Texas Southwestern Medical Center, Dallas, USA
2. **Bernd Pulverer** ^{id}, European Molecular Biology Organization (EMBO), Heidelberg, Germany
3. **Leonhard Held**, University of Zurich, Zurich, Switzerland
Eva Furrer ^{id}, University of Zurich, Zurich, Switzerland

Any reports and responses or comments on the article can be found at the end of the article.



This article is included in the **Science Policy Research gateway**.

Corresponding author: Lukas Stalder (lukas.stalder@meddek.unibe.ch)

Author roles: **Steck N:** Conceptualization, Data Curation, Formal Analysis, Investigation, Methodology, Visualization, Writing – Original Draft Preparation; **Stalder L:** Conceptualization, Data Curation, Investigation, Methodology, Writing – Review & Editing; **Egger M:** Conceptualization, Funding Acquisition, Investigation, Methodology, Project Administration, Supervision, Writing – Original Draft Preparation, Writing – Review & Editing

Competing interests: No competing interests were disclosed.

Grant information: This study was funded by intramural funds of the Dean's office of the Medical Faculty. NS and ME were supported by the Swiss National Science Foundation (grant 17481). We are grateful to the 64 candidates who submitted their applications for promotion.

The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Copyright: © 2020 Steck N *et al.* This is an open access article distributed under the terms of the **Creative Commons Attribution License**, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

How to cite this article: Steck N, Stalder L and Egger M. **Journal- or article-based citation measure? A study of academic promotion at a Swiss university [version 1; peer review: 2 approved, 1 approved with reservations]** F1000Research 2020, 9:1188 <https://doi.org/10.12688/f1000research.26579.1>

First published: 01 Oct 2020, 9:1188 <https://doi.org/10.12688/f1000research.26579.1>

Introduction

In academia, decisions on promotion to senior positions are influenced by the work published by the candidate. The assessment of publication lists should be systematic, using standardized criteria, and straightforward¹. Bibliometric measures such as the journal impact factor (JIF) or its rank within a given field meet this requirement. The JIF or its rank reflect citations to all articles published in the journal, rather than citations to the article submitted by the candidate. Of note, it was initially created as a tool to help librarians identify the journals they should subscribe to, and not as a measure of the scientific quality or impact of research². In 2013, the San Francisco Declaration on Research Assessment (DORA) was published with the aim to improve the way research output is evaluated². Research should be assessed on its own merits rather than based on the journal in which it is published. The first general recommendation of DORA says “Do not use journal-based metrics, such as Journal Impact Factors, as a surrogate measure of the quality of individual research articles, to assess an individual scientist’s contributions, or in hiring, promotion, or funding decisions.”². As of May 2020, more than 1,900 organizations and over 15,000 individuals have signed the DORA declaration.

There is thus growing consensus that the JIF is not a good measure to assess individual research papers. Efforts have been underway for several years to find a practical measure by which the citation impact of papers can be individually evaluated^{3–7}. In 2016 Hutchins *et al.*⁸, presented the Relative Citation Ratio (RCR), an article-level measure of citation impact which compares the citations to the article of interest with the articles in the network of co-cited articles⁸. The RCR is benchmarked to National Institutes of Health (NIH)-funded research⁹: an article with an RCR equal to 1.0 is at the median for NIH-funded articles in this year¹⁰. The NIH provides access to the RCR and percentile of papers indexed in the PubMed bibliometric database on a website¹¹. Several studies used the RCR to assess the citation impact of researchers, for example, vascular surgeons within the NIH¹², biomedical scientists in one country¹³ or papers from scientific publications produced by the United States Food and Drug Administration (FDA)¹⁴.

The University of Bern signed the DORA declaration in January 2016. Therefore, the Medical Faculty decided to review its practice for internal promotion, where the assessment of citation impact was based on the rank of the JIF¹⁵. In 2018, a working group of the Medical Faculty examined whether the RCR could replace the ranking of the JIF as a decision-making aid for hiring, tenure and promotion decisions. The present study aimed to investigate the effects of switching from journal-based JIF-ranking to the RCR in the assessment of the papers submitted by candidates.

Methods

Academic promotion

In medical faculties in Switzerland, the habilitation degree and promotion to associate professor are essential steps in an academic career. The habilitation degree was introduced in the first half of the 19th century to ensure the quality of academic teaching and research at German universities. Today the

habilitation is a post-doctoral qualification, which is required for independent teaching and supervision of doctoral students and to obtain an associate or full professorship in many European countries, including Switzerland¹⁶. At the University of Bern, the degree is conferred based on an application which includes a list of papers and a summary of the work highlighting the applicant’s contributions in research and teaching. An academic committee reviews the application, and the candidate presents and discusses his/her research at a faculty meeting. A similar process is followed for promotion to associate professor.

Until 2019, the faculty used the rank of the JIF¹⁵ for its assessment of applications for promotion to habilitation or associate professorships, and also The Journal Citation Reports (JCR)¹⁵ rank journals based on the JIF within subject categories, for example, oncology, surgery or nursing. Per university regulations¹⁷, candidates for habilitations needed at least ten original articles with a JIF rank in the upper third of the relevant discipline, and among the ten papers four as first or last author. The successful habilitation is a prerequisite for promotion to associate professor. The guidelines for promotion to associate professor required at least six additional original papers published in journals of the upper third of the JIF-based ranking, with at least three as first or last author¹⁸.

Study sample and data sources

The Dean’s office of the Medical Faculty of the University of Bern compiled the publication lists submitted by a randomly selected 34 candidates for habilitation and 30 candidates for associate professor in 2017 and 2018. For each paper, we recorded the JIF of that year and its ranking in the corresponding field. The data were obtained from the JCR of Clarivate Analytics¹⁵. The Relative Citation Ratio (RCR) and the RCR percentile were obtained from the iCite tool¹¹. Since the results for the papers submitted by candidates for the habilitation degree and for associate professorship were similar, we combined the data in the analysis.

Analysis

We assessed the papers submitted by the candidates and the number of first- or last-author papers. We calculated the number with a JIF ranking in the upper third (the cutoff defined in the regulations) and the number with an RCR percentile >66%. To visualize the distribution of RCR percentile and JIF ranking by candidate, we used beam plots¹⁹ and kernel density estimation (Epanechnikov distribution, bandwidth=5.0). We examined the relation between RCR percentiles and JIF rank in scatterplots and calculated the Pearson correlation coefficient and its confidence interval.

The calculation of the RCR and its percentile requires the article of interest to be cited so that these citations can be compared with those received by the articles in the co-citation network⁸. Papers in the second year after publication or more recent papers with five or more citations receive a provisional RCR¹¹. We included both articles with definitive and provisional RCRs in the main analysis. For each candidate with articles from both categories, we calculated the difference between the articles with definitive RCR and all articles, including provisional

RCRs, and combined the differences using random-effects meta-analysis. All statistical analyses were done in Stata version 15 (StataCorp, College Station, TX, USA).

Results

The 64 candidates submitted 1,903 original articles, including 801 (42.1%) first- or last-author papers. A total of 134 papers (7.0%) had no JIF, and 328 (17.2%) had no RCR; 378 papers had to be excluded. A total of 1,525 articles were included in the analyses, including 625 (41.0%) first- or last-author papers and

223 (14.6%) articles with a provisional RCR. At the time of the download of the bibliometric data (12 September 2018) the total number of citations to the 1,525 articles was 45,119.

The relation between journal rank and article RCR

The beam plot in [Figure 1](#) shows, row-wise for each candidate, the JIF ranking and RCR percentile of all 64 candidates. The kernel density estimation at the bottom of [Figure 1](#) shows the differences in distribution of JIF ranking and RCR percentiles. As expected from the university regulations, the majority of



Figure 1. Beam plots and Kernel density estimation of JIF rank and RCR percentile of all original publications submitted by candidates for academic promotion at the University of Bern. JIF rank (left panel) and RCR percentile (right panel) are shown for each article submitted by candidates for habilitation (1-34) and associate professorship (35-64). Each candidate corresponds to one row. Kernel density estimation (epanechnikov, bandwidth=5.0) for JIF rank (left panel) and RCR percentile (right panel) are shown below. The broken lines show rank 0.66 (left panel) and RCR percentile 66 (right panel).

papers were in the upper third of the JIF rank for all candidates. In contrast, they were more evenly distributed across percentiles of the RCR. Of note, some candidates had few or no articles above the 66th percentile of the RCR. Overall, 1,199 (78.6%) of 1,525 papers had a JIF rank in the upper third, and 581 (38.1%) had an RCR percentile above 66. Among the 625 first- and last-author papers, 489 (78.2%) had a JIF rank in the upper third, and 233 (37.3%) had an RCR percentile above 66. The beam plot and the kernel density estimation for first- and last-author papers was similar to the plot for all papers (see *Extended data: Figure S1²⁰*).

Figure 2 shows a scatterplot of the RCR percentile against the JIF ranks of all 1,525 papers submitted by applicants for habilitation and candidates for the associate professorship with data on both indicators. The correlation coefficient was 0.34 (95% CI 0.29 to 0.38). The scatter plot is divided into four quadrants by cutoffs 0.66 (for JIF rank) and 66 (for RCR percentile). Among the articles published in journals with a JIF ranking in the upper third, 57.5% (690 of 1,199) did not have an RCR percentile above 66 (blue quadrant in **Figure 2**). Conversely, 22.8% (72 of 326) of articles published in journals with a JIF rank in the lower two thirds (<0.66) had an RCR percentile above 66 (pink quadrant in **Figure 2**). The results for first- and last-author papers were similar: the correlation coefficient was 0.31 (95% CI 0.24 to 0.38), and the percentages of papers in the

blue and pink quadrants were 59.1% (289 of 489) and 24.3% (33 of 136), respectively (*Extended data: Figure S2²⁰*).

Definitive versus provisional RCRs

In total 57 candidates had both definitive and provisional RCRs (**Figure 3**). The meta-analysis of the differences between definitive and all RCRs across candidates gave an overall weighted mean difference of -0.04 (95% CI -0.13 to 0.04). There was thus no evidence of a systematic bias due to provisional RCRs, and no heterogeneity between candidates (I squared 0.0%).

Discussion

This analysis of papers submitted to promotion committees at a Swiss university illustrates that the rank of the journal's impact factor within its discipline is a bad proxy measure of the citation impact of individual articles. Many articles published in higher impact journals were cited less than their companion papers in the co-citation network. Whereas the majority of the papers submitted by candidates for the habilitation or an associate professorship were, by university regulation, published in journals with a JIF that ranked in the upper third of its field, only about 40% of these papers had an RCR percentile in the upper third. Furthermore, 20–25% of the papers that did not meet the requirement for the JIF rank (below 0.66) were, in fact, more impactful than their peers in

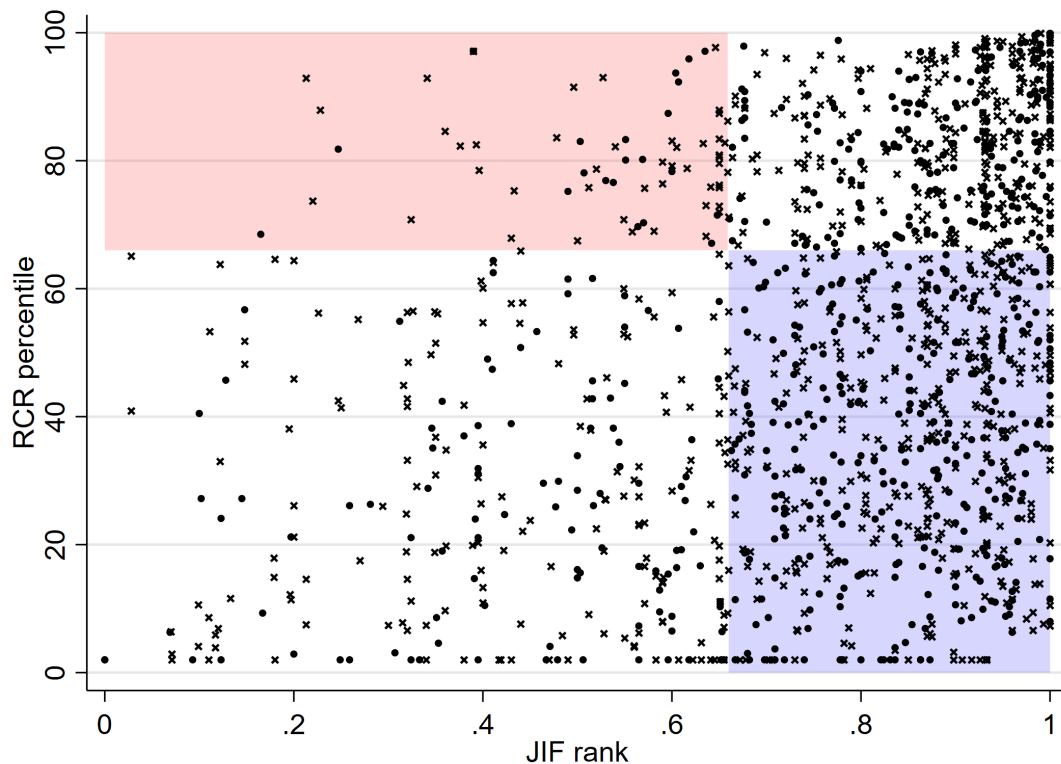


Figure 2. Scatter plot of RCR percentile against JIF rank of publications submitted by candidates for promotion at the University of Bern. Publications are shown as point or cross for candidates for habilitation and associate professorship, respectively. Cutoffs of 0.66 for the JIF rank (as per university regulations) and 66 for the RCR percentile define four quadrants. The pink top-left quadrant shows the publications that have an RCR percentile >66 but were published in a journal with a JIF ranking <0.66. The blue quadrant shows the papers published in a journal with a JIF rank >0.66 but had an RCR percentile <66.

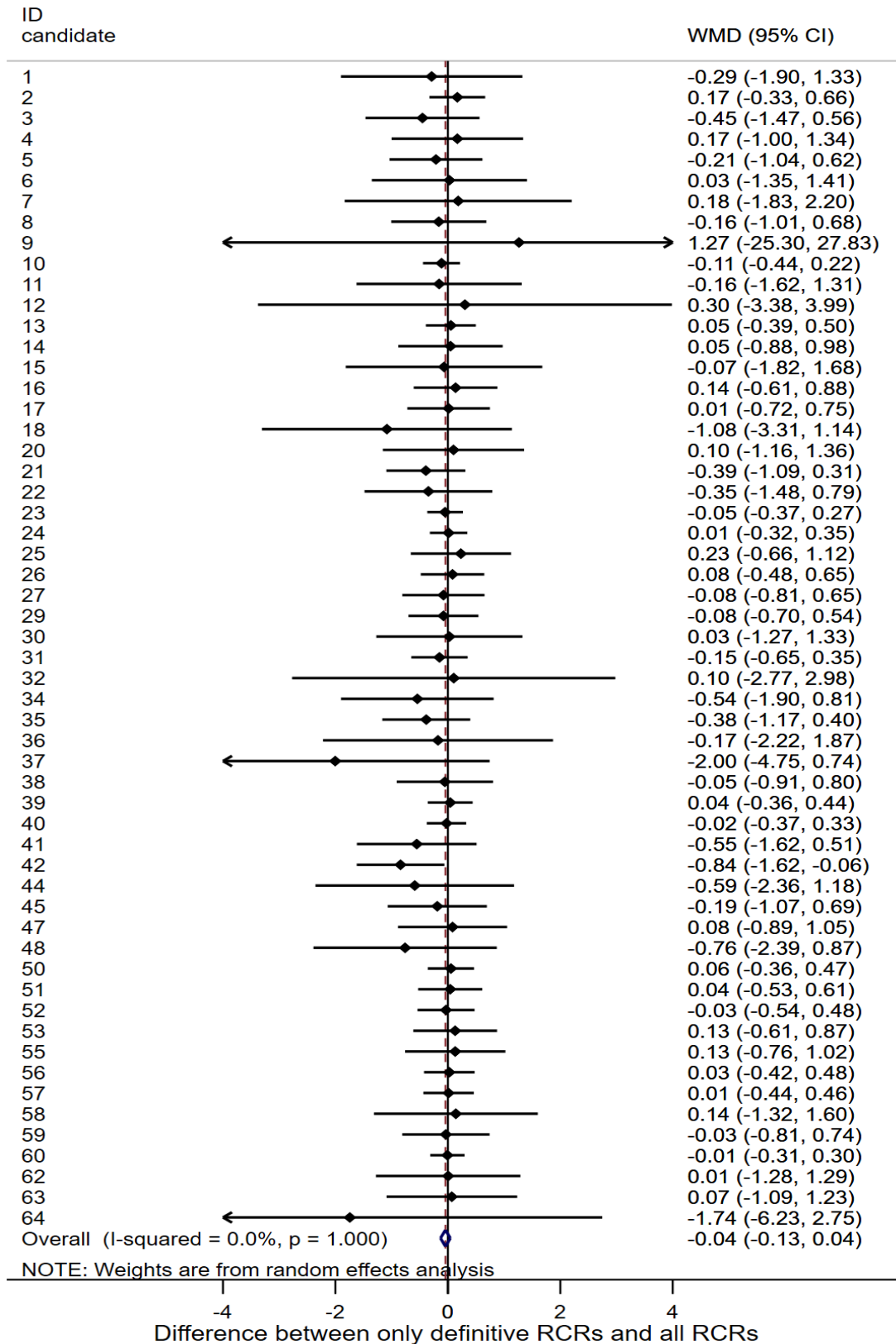


Figure 3. Meta-analysis of differences between means of definitive RCRs and all RCRs (definitive and provisional). For each candidate submitting articles with definitive and provisional RCRs the weighted mean difference (WMD, filled diamonds) and its 95% confidence interval (horizontal line) were calculated. The estimates were combined in a random-effects meta-analysis model. The empty diamond at the bottom shows the combined estimates from the meta-analysis of all candidates.

the co-citation network. Our study thus confirms the findings of the RCR developer's analysis of 80,000 papers: "Though journals with the highest impact factors have the highest median RCR, influential papers can be found in virtually all journals"⁸. Unsurprisingly, the correlation between the journal-based measure, the JIF, and the article-based measure, the RCR, was weak.

Several previous studies used the RCR to assess the citation impact of different groups of researchers^{12–14}. To the best of our knowledge, this is the first study investigating differences between a journal-based metric and an article-based metric of citation impact in the context of academic promotion. We used a large "real-world" dataset of over 1,500 papers submitted by candidates for academic promotion at a large Swiss medical faculty. Our results provide further empirical evidence supporting the San Francisco Declaration on Research Assessment (DORA)³, and challenges the practices at a Swiss university that relied inappropriately on the JIF. The results indicate that moving from a journal-based measure to an article-based measure is feasible. Indeed, the regulations of the Medical Faculty at the University of Bern have since been revised. The new regulations state that the assessment of candidates must follow the DORA principles. The committee analysing the papers should examine the novelty of the research question, the suitability of the methods, the interpretation of results and their relevance to the field. The evaluation should be based on the scientific content of the work. Article-based impact measures or qualitative indicators for the impact of the research (e.g. influence on policy and practice) may complement the assessment. Specifically, the regulations state that two or more of the papers with first or last authorship should have an RCR of 1 or higher. Also, the regulations explicitly state that "the journal and its impact factor will not be considered"²¹. The regulations for associate professor and titular professor were revised in the same spirit. They also refer to DORA³.

The RCR is based on citations and shares all the limitations of using citations as a proxy for impact. For example, the number of citations is influenced by factors unrelated to the quality of the research. The impact outside academia, including for political decision-making, is not well captured by citations^{22–24}. Furthermore, unlike the JIF, the RCR requires time to allow citations to the article of interest to appear in the literature, which may limit its use in the context of academic promotion. In our study, only 328 of 1,903 (17.2%) had to be excluded because no RCR was available. Furthermore, within candidates, the inclusion of provisional RCRs of recent papers did not influence their mean RCR, indicating that provisional RCRs can be included in assessments. Of note, the developers of the RCR have shown that the RCR is usually very stable after one year¹¹. The RCR is based on Medline and therefore not suitable for assessing non-biomedical literature¹⁰, for example, on medical ethics or teaching.

Moreover, when using the RCR, it should not be forgotten that the reference is the papers financed by the NIH⁸. While an

RCR of 1.0 corresponds to the median of the citations of NIH-funded articles, the median of all papers has an RCR of around 0.37¹¹, subject to annual fluctuations. Another important point of criticism regarding the RCR is that "papers may be penalized rather than rewarded for receiving interdisciplinary citations"²⁵. If a paper from a low-citation field is published in a journal from a high-citation field, this could reduce its RCR²⁵. However, a comparison with other bibliometric indicators did not support this criticism²⁶. Also, the developers of the RCR found a good agreement between metric and expert reviewer scores⁸.

In conclusion, we hope that our study will serve as a model to other researchers who, in the spirit of the DORA³ intend to challenge research assessment practices at medical and other faculties that rely inappropriately on Journal Impact Factors and contribute to promoting best practice that focuses on the value and influence of specific research outputs.

Data availability

Underlying data

The original raw data is composed of personal data of individuals applying for academic promotion and can therefore not be shared. The data without the identifying variables is available:

Open Science Framework: Journal- or article-based citation measure? A study of academic promotion at a Swiss university, <https://doi.org/10.17605/OSF.IO/H7SKN>²⁰.

Extended data

Open Science Framework: Journal- or article-based citation measure? A study of academic promotion at a Swiss university, <https://doi.org/10.17605/OSF.IO/H7SKN>²⁰.

This project includes the following extended data:

- **Figure S1. Beam plots and Kernel density estimation of JIF rank and RCR percentile of first- and last-author publications submitted by candidates for academic promotion at the University of Bern.** JIF (left panel) and RCR percentile (right panel) are shown for each article submitted by candidates for habilitation (1-34) and associate professorship (35-64) as first or last author. Each candidate corresponds to one row. Kernel density estimation (epanechnikov, bandwidth=5.0) for JIF rank (left panel) and RCR percentile (right panel) are shown below. The broken lines show rank 0.66 (left panel) and RCR percentile 66 (right panel).
- **Figure S2. Scatter plot of RCR percentile against JIF rank of publications submitted by candidates for promotion at the University of Bern as first or last authors.** Publications are shown as point or cross for candidates for habilitation and associate professorship, respectively. Cutoffs of 0.66 for the JIF rank (as per university regulations) and 66 for the RCR percentile define four quadrants. The pink top-left quadrant shows the

publications that have an RCR percentile >66 but were published in a journal with a JIF ranking <0.66. The blue quadrant shows the papers published in a journal with a JIF rank >0.66 but had an RCR percentile <66.

Data are available under the terms of the [Creative Commons Zero “No rights reserved” data waiver](#) (CC0 1.0 Public domain dedication).

Acknowledgements

We thank the Dean's office of the Medical Faculty of the University of Bern for providing the data for the study. We are grateful to the working group of the Medical Faculty at University of Bern (Martin Bachmann, Thomas Geiser, Andrea Huwiler) for helpful discussions of this study. We thank Stephen Curry and Michael Hill for their constructive comments on a previous version of the manuscript.

References

- Adler R, Ewing J, Taylor P: **Citation Statistics: A Report from the International Mathematical Union (IMU) in Cooperation with the International Council of Industrial and Applied Mathematics (ICIAM) and the Institute of Mathematical Statistics (IMS)**. *Statistical Science*. 2009; **24**(1): 1–14. [Reference Source](#)
- DORA – San Francisco Declaration on Research Assessment (DORA). [Accessed 3 July 2020]. [Reference Source](#)
- Moed HF, Burger WJM, Frankfort JG, et al.: **The use of bibliometric data for the measurement of university research performance**. *Res Policy*. 1985; **14**(3): 131–149. [Publisher Full Text](#)
- Zitt M, Small H: **Modifying the journal impact factor by fractional citation weighting: The audience factor**. *J Am Soc Inf Sci Technol*. 2008; **59**(11): 1856–1860. [Publisher Full Text](#)
- Bornmann L, Leydesdorff L: **The validation of (advanced) bibliometric indicators through peer assessments: A comparative study using data from InCites and F1000**. *Journal of Informetrics*. 2013; **7**(2): 286–291. [Publisher Full Text](#)
- Waltman L, Yan E, van Eck NJ: **A recursive field-normalized bibliometric performance indicator: An application to the field of library and information science**. *Scientometrics*. 2011; **89**(1): 301–314. [PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Waltman L, van Eck NJ, van Leeuwen TN, et al.: **Towards a new crown indicator: Some theoretical considerations**. *Journal of Informetrics*. 2011; **5**(1): 37–47. [Publisher Full Text](#)
- Hutchins BI, Yuan X, Anderson JM, et al.: **Relative Citation Ratio (RCR): A New Metric That Uses Citation Rates to Measure Influence at the Article Level**. Vaux DL editor. *PLoS Biol*. 2016; **14**(9): e1002541. [PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Ian Hutchins B, Baker KL, Davis MT, et al.: **The NIH open citation collection: A public access, broad coverage resource**. *PLoS Biol*. 2019; **17**(10): e3000385. [PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Surkis A, Spore S: **The relative citation ratio: What is it and why should medical librarians care?** *J Med Libr Assoc*. 2018; **106**(4): 508–513. [PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- NIH Office of Portfolio Analysis. NIH-Website iCite. [Reference Source](#)
- Davis FM, Obi AT, Gallagher KA, et al.: **Assessing the academic influence of vascular surgeons within the National Institutes of Health iCite database**. *J Vasc Surg. Mosby Inc.*; 2020; **71**(5): 1741–1748.e2. [PubMed Abstract](#) | [Publisher Full Text](#)
- Spiroski M: **Relative citation ratio of top twenty Macedonian biomedical scientists in pubmed: A new metric that uses citation rates to measure influence at the article level**. *Open Access Maced J Med Sci*. 2016; **4**(2): 187–193. [PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Schneider JA, Miklos AC, Onken J, et al.: **An Analysis of Recent FDA Oncology Scientific Publications**. *Oncologist*. 2020; **25**(3): 266–270. [PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Journal Impact Factor - Journal Citation Reports. In: *Web of Science Group*. [cited 6 Jun 2020]. [Reference Source](#)
- Weineck SB, Koelblinger D, Kiesslich T: **[Medical habilitation in German-speaking countries : Quantitative assessment of content and elaboration of habilitation guidelines]**. *Chirurg*. 2015; **86**(4): 355–365. [PubMed Abstract](#) | [Publisher Full Text](#)
- Medizinische Fakultät der Universität Bern: **Habilitationsreglement der Medizinischen Fakultät der Universität Bern**. 2016. [Reference Source](#)
- Medizinische Fakultät der Universität Bern. Assoziierte Professur 1 Richtlinien, Anforderungen, Verfahren, 1997.
- Bornmann L, Marx W: **Distributions instead of single numbers: Percentiles and beam plots for the assessment of single researchers**. *J Assoc Inf Sci Technol*. 2014; **65**(1): 206–208. [Publisher Full Text](#)
- Steck N, Stalder L, Egger M: **Journal- or article-based citation measure? A study of academic promotion at a Swiss university**. 2020. <http://www.doi.org/10.17605/OSF.IO/H7SKN>
- Medizinische Fakultät der Universität Bern: **Habilitationsreglement der Medizinischen Fakultät der Universität Bern**. 2019; 8–8. [Reference Source](#)
- Ravenscroft J, Liakata M, Clare A, et al.: **Measuring scientific impact beyond academia: An assessment of existing impact metrics and proposed improvements**. *PLoS One*. 2017; **12**(3): e0173152. [PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Hogan AM, Winter DC: **Changing the Rules of the Game: How Do We Measure Success in Social Media?** *Clin Colon Rectal Surg*. 2017; **30**(4): 259–263. [PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Liverani M, Hawkins B, Parkhurst JO: **Political and Institutional Influences on the Use of Evidence in Public Health Policy. A Systematic Review**. *PLoS One*. 2013; **8**(10): e77404. [PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Waltman L: **NIH's new citation metric: A step forward in quantifying scientific impact?** [cited 28 Apr 2020]. [Reference Source](#)
- Bornmann L, Haunschild R: **Relative Citation Ratio (RCR): An empirical attempt to study a new field-normalized bibliometric indicator**. *J Assoc Inf Sci Technol*. 2017; **68**(4): 1064–1067. [Publisher Full Text](#)

Open Peer Review

Current Peer Review Status:   

Version 1

Reviewer Report 29 October 2020

<https://doi.org/10.5256/f1000research.29342.r72324>

© 2020 Held L et al. This is an open access peer review report distributed under the terms of the [Creative Commons Attribution License](#), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.



Leonhard Held

Epidemiology, Biostatistics and Prevention Institute (EBPI), Center for Reproducible Science (CRS), University of Zurich, Zurich, Switzerland

Eva Furrer 

Epidemiology, Biostatistics and Prevention Institute (EBPI), Center for Reproducible Science (CRS), University of Zurich, Zurich, Switzerland

General comments

This is an interesting paper that compares journal- with article-based citation measures for the promotion of Swiss academics. The results from this analysis seem to have been used as a basis to change the current system for academic proportion at the University of Berne.

In the following we outline some aspects of the present study that could be improved to further justify its use for decision making. Some additional discussion of other promotional criteria related to teaching, academic service (such as peer review) and more general contributions for the benefit of society (e.g. software, policy briefs, open data etc) could also be added.

A consistent use of either "paper" or "article" would add more clarity to the paper.

Specific comments

Introduction, 2nd paragraph:

- "The RCR is benchmarked to National Institutes of Health (NIH)-funded research: an article with an RCR equal to 1.0 is at the median for NIH-funded articles in this year."
Change "this year" to "the same year"

Methods, Academic promotion, 2nd paragraph

- "Until 2019, the faculty used the rank of the JIF for its assessment of applications for promotion to habilitation or associate professorships, [...]"
Surely the applications have been assessed in other ways too and the JIF has only been used to quantify the publication output of the candidates?
- "Per university regulations, candidates for habilitations needed at least ten original articles"
Change "habilitations" to "habilitation"

- *"The successful habilitation is a prerequisite for promotion to associate professor."*
This has been mentioned in the paragraph above and can be deleted.

Study sample and data sources

- *"The Dean's office of the Medical Faculty of the University of Bern compiled the publication lists submitted by a randomly selected 34 candidates for habilitation and 30 candidates for associate professor in 2017 and 2018."*
Change "submitted by a randomly selected" to "submitted by randomly selected".
This paragraph suggests that the study sample was not just a simple convenience sample. More details are needed on how this random sample has been obtained: how large was the underlying source population, why the sample size of 64 and how exactly was the sampling done?
- *"For each paper, we recorded the JIF of that year and its ranking in the corresponding field."*
Change "that year" to "the corresponding year"
- *"Since the results for the papers submitted by candidates for the habilitation degree and for associate professorship were similar, we combined the data in the analysis."*
What does similar mean here exactly?

Analysis

The following issues need attention:

- Does a study protocol (with pre data collection analysis plan) exist?
- The data provided on OSF is not easy to read in - please separate explanation of variables from the actual data file, preferably not in Excel.
- For "Promotion=AssocProf" the variables ("First_author", "Last_author") sometimes have the values (1, .) resp. (., 1) where "." presumably means "missing". What is the difference to (1, 0) resp. (0, 1) for "Promotion=Habil"? If a candidate is a first author then she/he cannot be a last author and vice versa.
- The code of the analysis should be made available to make the analysis reproducible.
- Would it be possible and useful to follow the STROBE reporting guidelines on observational studies to some extent?
- Figure 1: The use of kernel smoothing along with the beam plots has the disadvantage that it extends beyond the domain boundaries (0-1). Also the y-scale of the left plot (kernel density estimation of JIF rank) is wrong, as the density does not integrate to unity. A simple histogram would provide the same information and is methodologically more appropriate.
- One of the key results of the paper is the weak positive correlation between JIF rank and RIC percentile. However, in view of the scatter plot in Figure 2 a Spearman or Kendall rank correlation seems more appropriate than a Pearson correlation.
- For the meta analysis it is written "For each candidate with articles from both categories, we

calculated the difference between the articles with definitive RCR and all articles, including provisional RCRs, and combined the differences using random effects meta analysis." This seems to indicate that for each candidate an average RCR of all articles with definitive RCR has been subtracted from an average RCR of all definitive and provisional articles. In Figure 3 it is mentioned that a weighted mean difference has been used. This needs to be clarified in more detail. Moreover, it needs to be made clearer to what purpose the meta analysis is used. All methods for which results are provided in the Results section should be described in the methods section, specifically the I-squared and the corresponding test. Figure 3 would profit from an indication which side corresponds to higher RCR for articles including only definitive RCR and which side for higher RCR for articles including definitive and provisional RCR.

- We are also wondering if the Hartung-Knapp or the traditional DerSimonian-Laird method is used for random effects meta-analysis, Hartung-Knapp is known to perform better, see <http://www.biomedcentral.com/1471-2288/14/25>.
- In the Results Section we learn that missing values are present. A discussion in the Methods Section on the treatment of missing values is needed. Specifically, an explanation why no imputation methods have been considered and some discussion of possible bias this may have caused.

Results, 1st paragraph

- 378 papers had to be excluded: how many authors does this affect and on average how many papers per author?

The relation between journal rank and article RCR

- A more consistent section header would be "The relation between JIF rank and article RCR percentile".

1st paragraph

- *"The beam plot in Figure 1 shows, row-wise for each candidate, the JIF ranking and RCR percentile of all 64 candidates."*
The plot shows the JIF ranking and RCR of all papers of a candidate not of the candidates themselves.

2nd paragraph

- 72 divided by 326 is 22.1% not 22.8%.

Definitive versus provisional RCRs

- *"In total 57 candidates had both definitive and provisional RCRs (Figure 3)."*
Again, it is not the candidates but their papers that have RCRs.

Discussion, 2nd paragraph

- *"Our results provide further empirical evidence supporting the San Francisco Declaration on Research Assessment (DORA), and challenges the practices at a Swiss university that relied inappropriately on the JIF."*
Change "challenges" to "challenge", "practices" to "former practices"
We think this paragraph could be made clearer how exactly the results of this study

influenced the faculty board in their decision to change practices. Was the study done before the decision and was the decision conditional on its results?

3rd paragraph

- *"For example, the number of citations is influenced by factors unrelated to the quality of the research."*

Can you give a reference for this statement?

- *"The RCR is based on Medline and therefore not suitable for assessing non-biomedical literature, for example, on medical ethics or teaching."*

It would be better to stick with the term "Pubmed" and not use Medline, this will be confusing to readers not familiar with biomedical research.

4th paragraph

- *"Moreover, when using the RCR, it should not be forgotten that the reference is the papers financed by the NIH."*

Change "the reference is the papers" to "the reference are the papers".

Last paragraph

- You imply that the change in Berne was due to the initiative of researchers. Is this the case or was it a management decision? Hence the public you want to address with this article are not individual researchers but researchers who are now in managerial positions, right?

Is the work clearly and accurately presented and does it cite the current literature?

Yes

Is the study design appropriate and is the work technically sound?

Yes

Are sufficient details of methods and analysis provided to allow replication by others?

Partly

If applicable, is the statistical analysis and its interpretation appropriate?

Partly

Are all the source data underlying the results available to ensure full reproducibility?

Yes

Are the conclusions drawn adequately supported by the results?

Yes

Competing Interests: LH currently receives several grants from the Swiss National Science Foundation (SNSF) - Matthias Egger (the last author) is currently president of the National Research Council of the SNSF.

Reviewer Expertise: LH: Biostatistics, Reproducibility and Replicability, Bayesian Biostatistics,

Infectious Disease Epidemiology EF: Biostatistics, Reproducibility, Scientific Rigor and Science Policy

We confirm that we have read this submission and believe that we have an appropriate level of expertise to confirm that it is of an acceptable scientific standard, however we have significant reservations, as outlined above.

Reviewer Report 19 October 2020

<https://doi.org/10.5256/f1000research.29342.r72318>

© 2020 Pulverer B. This is an open access peer review report distributed under the terms of the [Creative Commons Attribution License](#), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.



Bernd Pulverer 

European Molecular Biology Organization (EMBO), Heidelberg, Germany

This is a very valuable analysis underlying an important research assessment policy change at the University of Bern: a switch from assessing the top third research papers published by candidates for academic promotion as classified by the Clarivate 'Journal Impact factor' to the 'Relative Citation Ratio' (RCR), which uses a subject/citation network weighed algorithm. The analysis is based on 1,525 paper by 64 individuals and represents a sufficiently large dataset to yield meaningful results. The article concludes that RCR and JIF based rankings of papers do not show a significant correlation. This is important, but the authors should be careful not to conclude from the presented data that RCR is necessarily a good metric for academic performance.

I suggest the authors consider enhancing the study by addressing the following points:

Major:

1. The papers analyzed are from 2017/18 and the analysis from 2018. I suggest to re-run the citation counts for 2020. This will ensure that citation rates to most medical papers have peaked and will allow the authors to extend the comparison between JIF and RCR to a third important metric: actual citations to the papers analyzed (I suggest using a fixed time window of e.g. 18 months).
2. Given the dramatic differences in the assessment by RCR vs. JIF and the policy change by Bern, it would be helpful to others considering adopting this policy change to specify how many candidates would have been reclassified in the academic assessment as a result of this change.
3. The RCR is presented without any critical discussion as the clear article-level indicator of choice. Please discuss limitations at least briefly e.g. the citation network is heavily based on reference lists, which can in principle be 'gamed' by authors to increase RCR rankings.
4. On p. 5 the authors conclude that candidates were more evenly distributed across percentiles using the RCR. This may not affect habilitation and academic performance

evaluation, but it may affect hiring, where rank lists are used, since candidate ranking appear less pronounced.

Minor:

1. The Abstract refers to a 'moderate correlation', while p.7 refers to a 'weak' correlation. fig shows a very weak correlation at best.
2. It would be helpful to specify in more detail what disciplines were covered and if any differences were noted between subject areas in the analyzed dataset.
3. JIF and JIF rank within a Clarivate attributed field are treated as similarly problematic: I suggest to note that field ranking is even more problematic as the journals included are v. patchy e.g. general interest journals are not included.
4. It would be useful to briefly describe how the faculty plans to implement this significant policy change, especially during the transition period. Equally, to describe in more detail if other attributes are considered such as teaching quality (since 'habilitation' is described as a degree to 'ensure the quality of academic teaching and research' (p. 3).
5. V. minor: add a comma for 1,525, in the abstract; add 'in principle' to line 6, p. 3; add 'the decl. states that' to line 14, p.3; change 'had no JIF' to 'were publ. in jnl. without JIF' & add 'Thus, after RCR: p. 4, l. 6.

Additional suggestions for further-reaching analyses:

1. Candidates may chose different journals to publish if their assessment is based on RCR; a retrospective analysis if the journal profile of the faculty changes in the future would be fascinating.

Is the work clearly and accurately presented and does it cite the current literature?

Yes

Is the study design appropriate and is the work technically sound?

Yes

Are sufficient details of methods and analysis provided to allow replication by others?

Yes

If applicable, is the statistical analysis and its interpretation appropriate?

Yes

Are all the source data underlying the results available to ensure full reproducibility?

Yes

Are the conclusions drawn adequately supported by the results?

Partly

Competing Interests: steering committee, DORA

Reviewer Expertise: Molecular Biology; Scientific Publishing in biomedical sciences. Co-author of DORA declaration

I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard.

Reviewer Report 14 October 2020

<https://doi.org/10.5256/f1000research.29342.r72322>

© 2020 Schmid S. This is an open access peer review report distributed under the terms of the [Creative Commons Attribution License](#), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.



Sandra L. Schmid 

¹ Chan Zuckerberg Biohub, San Francisco, CA, USA

² Department of Cell Biology, University of Texas Southwestern Medical Center, Dallas, TX, USA

This is a very straight-forward comparison of the article-specific bibliometric RCR with the journal specific metric JIF. The results clearly support the conclusion that the JIFG is NOT a reliable metric of the impact of individual papers published in the journal and therefore should NOT be used as a tool to assess promotion and tenure. This is a very important, yet not unexpected, finding. It should be widely distributed and acted upon, just as the University of Bern has altered its policies.

Is the work clearly and accurately presented and does it cite the current literature?

Yes

Is the study design appropriate and is the work technically sound?

Yes

Are sufficient details of methods and analysis provided to allow replication by others?

Yes

If applicable, is the statistical analysis and its interpretation appropriate?

Yes

Are all the source data underlying the results available to ensure full reproducibility?

Yes

Are the conclusions drawn adequately supported by the results?

Yes

Competing Interests: No competing interests were disclosed.

Reviewer Expertise: I am a cell biologist and served as Chair of the Departments of Cell Biology at

The Scripps Research Institute and UT Southwestern for a total of ~20 years, hiring and promoting faculty. I was also a co-author and original signator of the DORA Proclamation.

I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard.

The benefits of publishing with F1000Research:

- Your article is published within days, with no editorial bias
- You can publish traditional articles, null/negative results, case reports, data notes and more
- The peer review process is transparent and collaborative
- Your article is indexed in PubMed after passing peer review
- Dedicated customer support at every stage

For pre-submission enquiries, contact research@f1000.com

F1000Research