

Report

The Collection and Use of Research Metrics for Assessment and Evaluation

Joint Committee on Metrics
January 2018



UNIVERSITY
OF MANITOBA

Report

The Collection and Use of Research Metrics for Assessment and Evaluation

Prepared by the Joint Committee on Metrics,
University of Manitoba and University of Manitoba Faculty Association
January 2018

Joint Committee on Metrics Members:

Charlotte Enns (Co-Chair), Sherri Vokey (Co-Chair),
Fletcher Baragar, James Blatz, Colin Garroway, Peter Nickerson

Table of Contents

Background: Joint Committee on Metrics.....	1
Summary of Committee Activities.....	2
Principles of Agreement.....	4
Recommendations	8
Selected Bibliography on Bibliometrics.....	9
Appendix. Appendix H.....	19

Background: Joint Committee on Metrics

The Joint Committee on Metrics was struck at the request of the University of Manitoba and the University of Manitoba Faculty Association (UMFA) during collective bargaining in 2016. The purpose of the committee was to decide whether the language proposed in the *Letter of Understanding* (known as Appendix H¹ and found at the end of this report) should be included in the Collective Agreement.

The University and Faculty Association were to each appoint three members to the Committee. The committee was comprised of the following members:

University-Appointed Members	UMFA-Appointed Members
James Blatz, PhD, PEng Faculty of Engineering Professor, Civil Engineering	Fletcher Baragar, PhD Faculty of Arts Associate Head & Associate Professor, Economics
Charlotte Enns (Co-Chair), PhD (EAF&P) Faculty of Education Associate Dean (Graduate & Research)	Colin Garroway, PhD Faculty of Science Assistant Professor, Biological Science
Peter Nickerson, MD, FRCPC, FCAHS Rady Faculty of Health Sciences Vice Dean - Research	Sherri Vokey (Co-Chair), MA, MLIS Health Sciences Libraries Head & Associate Librarian, Neil John Maclean Health Sciences Library

¹ University of Manitoba & University of Manitoba Faculty Association. Appendix H: Letter of Understanding Re: Joint Committee on Metrics. Available from: http://www.umfa.ca/images/pdfs/member-resources/Appendix_H.pdf

Summary of Committee Activities

Committee Meetings

The committee met six times between April and December 2017, on the following dates:

- April 3
- April 27
- May 26
- July 17
- October 26
- December 6

Committee Activities

Summer Student

A graduate student from the Archival Studies M.A. program, Christopher Kshyk, was hired from May to August 2018 to complete a comprehensive literature review on the use of metrics to measure the quality of research and scholarly work.

SSHRC Webinar

Several committee members participated in the SSHRC webinar on October 26, 2018 titled “Assessing Impacts in the Humanities and Social Sciences”, led by Peter Severinson and David Phillips.

Invited Speaker

A public presentation by Dr. Vincent Larivière, CRC on the Transformations of Scholarly Communication, was organized for all University of Manitoba faculty members on November 23, 2017.

Bibliography

The bibliography was initiated by the summer student and completed by co-chair, Sherri Vokey, to provide a detailed database of publications related to the evaluation of research metrics.

Committee Vote & Summary

The final committee meeting on December 6, 2017 included a vote on the following motion:

“To move to recommend the immediate addition of all language included within section 7 of Appendix H (19.D.1.8.5. through 35.5.7) of the 2016-2017 University of Manitoba – University of Manitoba Faculty Association Collective Agreement.”

The motion failed with three members voting for, and three members voting against the inclusion of the proposed language in Appendix H.²

NO	YES
James Blatz	Fletcher Baragar
Charlotte Enns	Colin Garroway
Peter Nickerson	Sherri Vokey

Note: The committee felt it was important to state that the vote result did not reflect the unanimous agreement amongst all committee members regarding the issues of bibliometrics, and that the disagreement centred around the specific language being proposed for the Collective Agreement. The attached Principles, Recommendations and Bibliography can provide guidelines for future work in this area.

² University of Manitoba & University of Manitoba Faculty Association. Appendix H: Letter of Understanding Re: Joint Committee on Metrics. Available from: http://www.umfa.ca/images/pdfs/member-resources/Appendix_H.pdf

Principles of Agreement

After careful consideration of the scholarly literature and consultation with experts in metrics and research assessment, the Joint Committee on Metrics is unanimous in its agreement with the following ten principles.

1. Regular assessment and evaluation is important for both guidance and career progression.

The committee agreed that proper assessment and evaluation of individuals for tenure, promotion, and performance evaluations is critical for the success of faculty, librarians, instructors, and the university. It was agreed that peer review is the foundation of assessment. The committee further agreed that empirically tested metrics published in the peer-reviewed bibliometric literature that were demonstrably objective and unbiased measures of well-defined aspects of research quality could serve as an additional important component of evaluation. The committee discussed hypothetical scenarios where individual biases of those performing peer review may have negative effects on tenure and promotion evaluations. There was some disagreement about whether the peer review system was inherently biased or whether it was individuals within the system that biased the process of peer review. No clear pathway through which metrics could be used to correct individual biases was identified.

2. Bibliometrics should be objective.

The committee agreed very early after formation that journal level metrics (i.e. JIF: Journal Impact Factor) cannot indicate the quality of individual articles in a journal. The committee agreed that there must be an empirically demonstrable and clear relationship between a metric and a precisely defined aspect of the quality of an individual's research for it to be usable for tenure, promotion, and performance evaluations. This relationship between metric and quality should be supported by up-to-date research in the peer-reviewed bibliometric literature.

For example: metric X is being used because it accurately measures research quality Y as demonstrated in the peer-reviewed bibliometric research literature Z. Custom-designed metrics are never to be used until thoroughly tested and assessed. The San Francisco Declaration on Research Assessment's (DORA)³ recommendation on this issue is as follows: "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."⁴

3. Bibliometrics should be discipline-appropriate.

The Committee agreed that there is no logical connection between the volume of publications produced and their quality or importance. It was agreed that after scholarly works are read and assessed via the peer-review process, that it may be desirable or sensible to assess the volume of output relative to those publications (or any other unit of output) if they are deemed to be

³ American Society of Cell Biology (ASCB). San Francisco Declaration on Research Assessment (DORA). Available from: <http://www.ascb.org/dora/>

⁴ American Society of Cell Biology (ASCB). San Francisco Declaration on Research Assessment (DORA): General Recommendation #1: <http://www.ascb.org/wp-content/uploads/2017/07/sfdora.pdf>

of approximately equal quality. This is a customary part of the peer-review process that would remain unaffected by the addition of the proposed language found in Appendix H (see article 19.D.1.8.4 on page 63 of the 2016-2017 Collective Agreement)⁵. This does not contravene the language included in 19.D.1.8.1 on page 63 of the 2016-2017 Collective Agreement: "...quantitative analysis and statistics used to assess the quality of a body of work". The additional language in Appendix H that is being voted on for addition into the Collective Agreement is thus not related to this common task.

4. *Bibliometrics should not be biased.*

The number of citations a paper receives is nearly entirely predicted by the size of the sub-field within which it is published⁶. Therefore, used in this way, citations are not an indicator of quality of an individual piece of research's quality.

A researcher's cumulative citations are best predicted by the length of time they have been doing research and the number of papers they have published. Time spent doing research is not an indicator of the quality of an individual's research.

Citation rates and the citation life (time to integration and recognition) of papers vary widely by discipline and sub-discipline. Citation peaks vary across sub-disciplines, and most often many years after publication. Thus, the timing of tenure and promotion applications will happen too early for papers in the majority of disciplines to be recognized and cited. This metric thus negatively affects researchers in fields where citations accrue more slowly, such as in the Arts and Humanities.

The Committee agreed that the citation explicitly named in the language of the Collective Agreement (i.e. 19.D.1.8.1) is a flawed metric that should not be used. Further, the Committee noted that it was unfortunate that the *h*-index was explicitly named *as an example of a metric that could be used* in the language of the Collective Agreement and that this was not eligible for discussion, revision, and vote.

5. *The h-index and its relatives do not indicate the quality of an individual researcher.*

The committee reached early agreement about the issues and defects associated with the *h*-index. Gingras described the *h*-index as being like a bad thermometer that can't go down and can only go up very slowly.⁷ The *h*-index is highly correlated with years in a position and limited by the number of papers published. It also suffers from all biases associated with citation and paper counting as these are the components of the metric. The Committee noted that it was unfortunate that the *h*-index was explicitly named in the language of the Collective Agreement (i.e. 19.D.1.8.1) and that this was not eligible for discussion, revision, and vote.

⁵ University of Manitoba & University of Manitoba Faculty Association. Collective Agreement – UM/UMFA 2016-2017. Available from: <http://www.umfa.ca/member-resources/collective-agreement>

⁶ Gingras Y. *Bibliometrics and Research Evaluation: Uses and Abuses*. Cambridge: The MIT Press; 2014. pp 15-18; 20-21.

⁷ Gingras Y. *Bibliometrics and Research Evaluation: Uses and Abuses*. Cambridge: The MIT Press; 2014. pp 42-44.

6. *Bibliometrics alter the focus of research in undesirable ways due to incentivising research that maximises an individual's performance based on a particular metric or set of metrics.*

This phenomenon exemplifies *Goodhart's Law*: that is, statistically speaking, once a metric becomes a goal, it ceases to be a valid measure of what it was intended to measure. Once quantitative metrics are instituted then there is incentive to maximize one's ranking. Such incentives change research behaviour in an undesirable way, moving it in directions that maximize the metric performance and not necessarily the needs of society or the curiosity-driven interests of a researcher.

Additionally, when metrics aim to indicate 'impact', (e.g., publication in international journals, citations) then the incentive becomes tied to doing research on high profile topics that will be published in high visibility journals to indicate impact, as per the metric. Judging research quality in this way ignores the fact that important and high-quality research is needed throughout society, not just on flashy and international issues or regions. The quality of this sort of research cannot be judged by 'visibility' metrics. This is an undesirable outcome because research is largely locally funded (provincially and nationally) and setting goals that incentivizes non-local research creates a system where we as a society are often paying salaries and funding research on the issues that most directly benefits others.⁸

This phenomenon was discussed in Yves Gingras' book when he presented a scenario in which Canadian academic economists may tend to disproportionately study American and global economies at the expense of research on Canadian issues. This was due to striving for high visibility research and to the detriment of the society that funded the research”

“An Economist who wants to “maximize” its number of citations would thus tend to study the economy of the United States rather than that of France or Canada, which are of little interest to American journals of economics that happen to be the most cited.”⁹

7. *Numerical components of bibliometrics are gender-biased.*

All bibliometrics use some combination of the number of papers, the number of citations, or authorship order in their calculations. Larivière et al. summarized all research included in ISI's Web of Science database (>5 million records 2008 - 2013) across all disciplines.¹⁰ The analysis demonstrated that in all disciplines and all localities, men published more papers than women, women held fewer senior authorship positions on papers than men, and women were cited less often. This work was a summary of the raw data on publication practices and so this bias is inherent to all bibliometrics that rely on these values. The committee agreed that systematic discrimination of any sort in evaluation is unacceptable.

⁸ Gingras Y. *Bibliometrics and Research Evaluation: Uses and Abuses*. Cambridge: The MIT Press; 2014. pp 54-57.

⁹ Gingras Y. *Bibliometrics and Research Evaluation: Uses and Abuses*. Cambridge: The MIT Press; 2014. p 54.

¹⁰ Larivière V, Ni C, Gingras Y, Cronin B, Sugimoto CR. *Bibliometrics: Global Gender Disparities in Science*. *Nature News*. 2013;504(7479):211.

8. Databases used to calculate bibliometrics are not well-correlated at the individual level.

This point was elucidated in a seminar presented by Vincent Larivière at the University of Manitoba on November 23, 2017.¹¹ Bibliometrics are reliably calculable when the same database is used. However, database information is not well correlated when correlations are assessed at the individual level. This lack of correlation makes the use of metrics for the assessment of individuals lack robustness: they vary considerably from one database to the next and there is no objective basis for deciding which database is best or provides greater accuracy.

9. There are currently no metrics that suit the criteria for usability in the evaluation of individuals.

Despite an effort to uncover metrics that met the committee's criteria, it was agreed, after extensive research and consultation with the scholarly literature, that there are no current bibliometrics suitable for the evaluation of individual research quality. There was clear utility for the use of metrics to assess 'populations' of individuals. However, the use of these metrics for assessing individuals suffered from the textbook statistical and logical fallacy termed the 'ecological fallacy.' This well-characterized statistical logical fallacy occurs during the interpretation of statistical results and states that one cannot draw inferences about the nature of individuals from inferences about the group to which an individual is a member. While the utility of bibliometrics at the group level was in some cases clear, the Committee was not able to identify a viable or valid way for group level indicators to be translated to individuals level indicators of research quality.

¹¹ Larivière V. Impact Factors and Bibliometrics: What Everyone Should Know: An Open Seminar on Bibliometrics and Research Evaluation [Seminar]. Winnipeg: University of Manitoba, Joint Committee on Metrics; Nov. 23, 2017.

Recommendations

In light of the numerous flaws and issues presently associated with the application of research metrics in the evaluation of scholarly work, the Joint Committee agreed on a set of principles and guidelines that are necessary in order for a metric to be considered valid.

Much of the Committee's work in this area was informed by Gingras' *Essential Characteristics of a Good Indicator*.¹²

Principles of Evaluation for Bibliometrics

1. Peer review is foundational.
2. Bibliometrics should strive for objectivity.
3. Bibliometrics should be customized to the discipline.
4. Bibliometrics should minimize bias.
5. Bibliometrics should maximize objectivity.

Guidelines in the Use of Bibliometrics

1. Bibliometrics that are flawed should be avoided.
2. Bibliometrics should be empirically validated measures related to well-defined features of research or researcher quality.
3. Bibliometrics should be field-weighted.

¹² Gingras Y. *Bibliometrics and Research Evaluation: Uses and Abuses*. Cambridge: The MIT Press; 2014. pp 71-79.

Selected Bibliography on Bibliometrics

1. Abramo G, Cicero T, D'Angelo CA. Should the research performance of scientists be distinguished by gender? *Journal of Informetrics*. 2015;9(1):25–38.
2. Abramo G, Cicero T, D'Angelo CA. Are the authors of highly cited articles also the most productive ones? *Journal of Informetrics*. 2014;8(1):89–97.
3. Abramo G, D'Angelo CA. Ranking research institutions by the number of highly-cited articles per scientist. *Journal of Informetrics*. 2015;9(4):915–23.
4. Abramo G, D'Angelo CA. Does your surname affect the citability of your publications? *Journal of Informetrics*. 2017;11(1):121–7.
5. Abramo G, D'Angelo CA, Costa F Di. National research assessment exercises: A comparison of peer review and bibliometrics rankings. *Scientometrics*. 2011;89(3):929–41.
6. Abramo G, D'Angelo CA, Rosati F. The importance of accounting for the number of co-authors and their order when assessing research performance at the individual level in the life sciences. *Journal of Informetrics*. 2013;7(1):198–208.
7. Abramo G, D'Angelo CA, Rosati F. Career advancement and scientific performance in universities. *Scientometrics*. 2014;98(2):891–907.
8. Abramo G, D'Angelo CA, Rosati F. Measuring institutional research productivity for the life sciences: The importance of accounting for the order of authors in the byline. *Scientometrics*. 2013;97(3):779–95.
9. Abramo G, D'Angelo CA, Viel F. The suitability of h and g indexes for measuring the research performance of institutions. *Scientometrics*. 2013;97(3):555–70.
10. Adler NJ, Harzing A-W. When Knowledge Wins: Transcending the Sense and Nonsense of Academic Rankings. *Academy of Management Learning & Education*. *Academy of Management*; 2009;8(1):72–95.
11. Agrawal AA. Corruption of journal Impact Factors. *Trends in Ecology & Evolution*. 2005;20(4):157.
12. Anninos LN. Research Performance Evaluation: Some Critical Thoughts on Standard Bibliometric Indicators. *Studies in Higher Education*. 2014;39(9):1542–61.
13. Ardanuy J. Sixty years of citation analysis studies in the humanities (1951-2010). *Journal of the American Society for Information Science and Technology*. 2013;64(8):1751–5.
14. Arlinghaus R. Are Current Research Evaluation Metrics Causing a Tragedy of the Scientific Commons and the Extinction of University-Based Fisheries Programs? *Fisheries*. 2014;39(5):212–5.

15. Barnes CS. The construct validity of the h-index. *Journal of Documentation*. 2016;72(5):878–95.
16. Barnes C. The emperor’s new clothes: the h-index as a guide to resource allocation in higher education. *Journal of Higher Education Policy and Management*. 2014;36(5):456–70.
17. Belter CW. Bibliometric indicators: opportunities and limits. *Journal of the Medical Library Association : JMLA*. 2015;103(4):219–21.
18. Bolli T, Schläpfer J. Job mobility, peer effects, and research productivity in economics. *Scientometrics*. Springer Netherlands; 2015;104(3):629–50.
19. Bornmann L. Do altmetrics point to the broader impact of research? An overview of benefits and disadvantages of altmetrics. *Journal of Informetrics*. 2014;8(4):895–903.
20. Bornmann L. The problem of citation impact assessments for recent publication years in institutional evaluations. *Journal of Informetrics*. 2013;7(3):722–9.
21. Bornmann L, Daniel H-D. The state of h index research. Is the h index the ideal way to measure research performance? *EMBO reports*. 2009;10(1):2–6.
22. Bornmann L, Marx W. How should the societal impact of research be generated and measured? A proposal for a simple and practicable approach to allow interdisciplinary comparisons. *Scientometrics*. 2014;98(1):211–9.
23. Bornmann L, Mutz R, Neuhaus C, Daniel HD. Citation counts for research evaluation: Standards of good practice for analyzing bibliometric data and presenting and interpreting results. *Ethics in Science and Environmental Politics*. 2008;8(1):93–102.
24. Buchanan AL, Hérubel J-PVM. Disciplinary Culture, Bibliometrics, and Historical Studies. *Behavioral & Social Sciences Librarian*. 1997;15(2):37–53.
25. Butler L. Using a balanced approach to bibliometrics: Quantitative performance measures in the Australian Research Quality Framework. *Ethics in Science and Environmental Politics*. 2008;8(1):83–92.
26. Butler L, McAllister I. Metrics or peer review? Evaluating the 2001 UK research assessment exercise in political science. *Political Studies Review*. 2009;7(1):3–17.
27. Campbell P. Escape from the impact factor. *Ethics in Science and Environmental Politics*. 2008;8(1):5–7.
28. Carpenter CR, Cone DC, Sarli CC. Using publication metrics to highlight academic productivity and research impact. *Academic Emergency Medicine*. 2014;21(10):1160–72.

29. Carpenter CR, Cone DC, Sarli CC. Using publication metrics to highlight academic productivity and research impact. Gaddis GM, editor. *Academic Emergency Medicine*. 2014;21(10):1160–72.
30. Castleden H, Sylvestre P, Martin D, McNally M. “I Don’t Think that Any Peer Review Committee . . . Would Ever ‘Get’ What I Currently Do”: How Institutional Metrics for Success and Merit Risk Perpetuating the (Re)production of Colonial Relationships in Community-Based Participatory Research Involving Indigenous People in Canada. *International Indigenous Policy Journal*. 2015;6(4).
31. Chang YW. A comparison of citation contexts between natural sciences and social sciences and humanities. *Scientometrics*. 2013;96(2):535–53.
32. Cheung WWL. The economics of post-doc publishing. *Ethics in Science and Environmental Politics*. 2008;8(1):41–4.
33. Choudhri AF, Siddiqui A, Khan NR, Cohen HL. Understanding bibliometric parameters and analysis. *Radiographics*. 2015;35(3):736–46.
34. Ciriminna R, Pagliaro M. On the use of the h-index in evaluating chemical research. *Chemistry Central Journal*. 2013;7(1):132.
35. Costas R, Bordons M. The h-index: Advantages, limitations and its relation with other bibliometric indicators at the micro level. *Journal of Informetrics*. 2007;1(3):193–203.
36. Council of Canadian Academies. *Informing Research Choices: Indicators and Judgement*. Ottawa; 2012. Available from: [http://www.scienceadvice.ca/uploads/eng/assessments and publications and news releases/science performance/scienceperformance_fullreport_en_web.pdf](http://www.scienceadvice.ca/uploads/eng/assessments_and_publications_and_news_releases/science_performance/scienceperformance_fullreport_en_web.pdf)
37. De Rijcke S, Rushforth A. To Intervene or Not to Intervene; Is That the Question? On the Role of Scientometrics in Research Evaluation. *Journal of the Association for Information Science and Technology*. 2015;66(9):1954–8.
38. Delgado-López-Cózar E, Cabezas-Clavijo Á. Ranking journals: Could Google Scholar Metrics be an alternative to journal citation reports and Scimago journal rank? *Learned Publishing*. 2013;26(2):101–14.
39. Derrick GE, Haynes A, Chapman S, Hall WD. The Association between Four Citation Metrics and Peer Rankings of Research Influence of Australian Researchers in Six Fields of Public Health. *PLOS ONE*. Public Library of Science; 2011;6(4):e18521.
40. Devos P. Research and bibliometrics: A long history... *Clinics and Research in Hepatology and Gastroenterology*. 2011;35(5):336–7.
41. Diem A, Wolter SC. The Use of Bibliometrics to Measure Research Performance in Education Sciences. *Research in Higher Education*. 2013;54(1):86–114.

42. Dupps WJ. Impact of citation practices: Beyond journal impact factors. *Journal of Cataract and Refractive Surgery*. 2008;34(9):1419–21.
43. Economic & Social Research Council. Taking Stock: A Summary of ESRC's Work to Evaluate the Impact of Research on Policy and Practice. 2009. Available from: <http://www.esrc.ac.uk/files/research/research-and-impact-evaluation/taking-stock-a-summary-of-esrc-s-work-to-evaluate-the-impact-of-research-on-policy-and-practice/>
44. Ellegaard O, Wallin JA. The bibliometric analysis of scholarly production: How great is the impact? *Scientometrics*. 2015;105(3):1809–31.
45. Federation for the Humanities and Social Sciences. Approaches to Assessing Impacts in the Humanities and Social Sciences. Ottawa; 2017. Available from: <http://www.ideas-ideas.ca/issues/research-impact>
46. Federation for the Humanities and Social Sciences. The Impacts of Humanities and Social Science Research. Ottawa; 2014. Available from: <http://www.ideas-ideas.ca/sites/default/files/2014-10-03-impact-project-draft-report-english-version-final2.pdf>
47. Ferber MA, Brün M. The gender gap in citations: Does it persist? *Feminist Economics*. 2011;17(1):151–8.
48. Franco G. Research evaluation and competition for academic positions in occupational medicine. *Archives of Environmental and Occupational Health*. 2013;68(2):123–7.
49. Gagolewski M. Scientific impact assessment cannot be fair. *Journal of Informetrics*. 2013;7(4):792–802.
50. Gingras Y. *Bibliometrics and Research Evaluation: Uses and Abuses*. Cambridge: The MIT Press; 2016. 119 p.
51. González-Valiente CL, Pacheco-Mendoza J, Arencibia-Jorge R. A review of altmetrics as an emerging discipline for research evaluation. *Learned Publishing*. 2016;29(4):229–38.
52. Gorraiz J, Gumpenberger C. A flexible bibliometric approach for the assessment of professorial appointments. *Scientometrics*. 2015;105(3):1699–719.
53. Gruber T. Academic sell-out: how an obsession with metrics and rankings is damaging academia. *Journal of Marketing for Higher Education*. 2014;24(2):165–77.
54. Guthrie S, Wamae W, Diepeveen S, Wooding S, Grant J. *Measuring research: A guide to research evaluation frameworks and tools*. Santa Monica: RAND Corporation; 2013. Available from: <https://www-rand-org.uml.idm.oclc.org/pubs/monographs/MG1217.html>
55. Habibzadeh F, Yadollahie M. Journal weighted impact factor: A proposal. *Journal of Informetrics*. 2008;2(2):164–72.

56. Hack TF, Crooks D, Plohman J, Kepron E. Research citation analysis of nursing academics in Canada: Identifying success indicators. *Journal of Advanced Nursing*. 2010;66(11):2542–9.
57. Haeffner-Cavaillon N, Graillot-Gak C. The use of bibliometric indicators to help peer-review assessment. *Archivum Immunologiae et Therapiae Experimentalis*. 2009;57(1):33–8.
58. Harnad S. Validating research performance metrics against peer rankings. *Ethics in Science and Environmental Politics*. 2008;8(1):103–7.
59. Harzing A-W, van der Wal R. Google Scholar as a new source for citation analysis. *Ethics in Science and Environmental Politics*. 2008;8(1):61–73.
60. Haustein S, Larivière V. The Use of Bibliometrics for Assessing Research: Possibilities, Limitations and Adverse Effects. In: Welpel IM, Wollersheim J, Ringelhan S, Osterloh M, editors. *Incentives and Performance*. Cham: Springer International Publishing; 2015. p. 121–39
61. Henneken EA, Accomazzi A. Linking to Data - Effect on Citation Rates in Astronomy. *CoRR*. 2011;abs/1111.3. Available from: <http://arxiv.org/abs/1111.3618>
62. Higginson AD, Munafò MR. Current Incentives for Scientists Lead to Underpowered Studies with Erroneous Conclusions. *PLoS Biology*. 2016;14(11).
63. Hirsch JE. An index to quantify an individual's scientific research output. *Proceedings of the National Academy of Sciences of the United States of America*. 2005 Nov 15;102(46):16569–72.
64. Hjørland B. Citation analysis: A social and dynamic approach to knowledge organization. *Information Processing & Management*. 2013;49(6):1313–25.
65. Holden G, Rosenberg G, Barker K. Bibliometrics. *Social Work in Health Care*. 2005;41(3–4):67–92.
66. Hood WW, Wilson CS. The literature of bibliometrics, scientometrics, and informetrics. *Scientometrics*. 2001;52(2):291–314.
67. Htoo THH, Na J-C. Disciplinary differences in altmetrics for social sciences. *Online Information Review*. 2017;41(2):235–51.
68. Johnson MH, Cohen J, Grudzinskas G. The uses and abuses of bibliometrics. *Reproductive BioMedicine Online*. 2012;24(5):485–6.
69. Jonkers K, Derrick GE. The Bibliometric Bandwagon: Characteristics of Bibliometric Articles Outside the Field Literature. *Journal of the American Society for Information Science and Technology*. 2012;63(4):829–36.

70. Kähler O. Combining peer review and metrics to assess journals for inclusion in Scopus. *Learned Publishing*. 2010;23(4):336–46.
71. Kelly CD, Jennions MD. The h index and career assessment by numbers. *Trends in Ecology and Evolution*. 2006;21(4):167–70.
72. Khodiyar VK, Rowlett KA, Lawrence RN. Altmetrics as a means of assessing scholarly output. *Learned Publishing*. 2014;27(5):S25–32.
73. Kokko H, Sutherland WJ. What do impact factors tell us? *Trends in Ecology & Evolution*. 1999;14(10):382–4.
74. Kostoff RN. Use and misuse of metrics in research evaluation. *Science and Engineering Ethics*. 1997;3(2):109–20.
75. Krauss J. Journal self-citation rates in ecological sciences. *Scientometrics*. 2007;73(1):79–89.
76. Lando T, Bertoli-Barsotti L. A new bibliometric index based on the shape of the citation distribution. *PLoS ONE*. 2014;9(12).
77. Larivière V, Kiermer V, Maccallum CJ, Mcnutt M, Patterson M, Pulverer B, et al. A simple proposal for the publication of journal citation distributions. *bioRxiv*. 2016; Available from: <http://biorxiv.org/content/early/2016/09/11/062109>
78. Larivière V, Ni C, Gingras Y, Cronin B, Sugimoto CR. Bibliometrics: Global gender disparities in science. *Nature*. 2013;504(7479).
79. Lawrence PA. The politics of publication. *Nature*. 2003;422(6929):259–61.
80. Lawrence PA. The mismeasurement of science. *Current Biology*. 2007;17(15):R583–5.
81. Lawrence PA. Lost in publication: How measurement harms science. *Ethics in Science and Environmental Politics*. 2008;8(1):9–11.
82. López-Cózar ED, Robinson-García N, Torres-Salinas D. The google scholar experiment: How to index false papers and manipulate bibliometric indicators. *Journal of the Association for Information Science and Technology*. 2014;65(3):446–54.
83. Lovegrove BG, Johnson SD. Assessment of Research Performance in Biology: How Well Do Peer Review and Bibliometry Correlate. *BioScience*. 2008;58(2):160–4.
84. Meho LI, Sonnenwald DH. Citation ranking versus peer evaluation of senior faculty research performance: A case study of Kurdish scholarship. *Journal of the American Society for Information Science*. 2000;51(2):123–38.
85. Metcalfe NB. Serious bias in journal impact factors. *Trends in ecology & evolution*. 1995;10(11):461.

86. Michels C, Schmoch U. Impact of bibliometric studies on the publication behaviour of authors. *Scientometrics*. 2014;98(1):369–85.
87. Mongeon P, Paul-Hus A. The journal coverage of Web of Science and Scopus: a comparative analysis. *Scientometrics*. 2016;106(1):213–28.
88. Morton S. Progressing research impact assessment: A “contributions” approach. *Research Evaluation*. 2015;24(4):405–19.
89. Neufeld J, von Ins M. Informed peer review and uninformed bibliometrics? *Research Evaluation*. 2011;20(1):31–46.
90. Olden JD. How do ecological journals stack-up? Ranking of scientific quality according to the h index. *Écoscience*. 2007 Jan 1;14(3):370–6.
91. Ortega JL. Are peer-review activities related to reviewer bibliometric performance? A scientometric analysis of Publons. *Scientometrics*. 2017;112(2):947–62.
92. Pendlebury DA. The use and misuse of journal metrics and other citation indicators. *Archivum Immunologiae et Therapiae Experimentalis*. 2009;57(1):1–11.
93. Pendlebury DA, Adams J. Comments on a critique of the Thomson Reuters journal impact factor. *Scientometrics*. 2012;92(2):395–401.
94. Polit DF, Northam S. Impact factors in nursing journals. *Nursing Outlook*. 2011;59(1):18–28.
95. Reece RL, Hardy MC. Moving Beyond Metrics: A Primer for Hiring and Promoting a Diverse Workforce in Entomology and Other Natural Sciences. *Annals of the Entomological Society of America*. 2017;110(5):484–91.
96. Régibeau P, Rockett KE. Research assessment and recognized excellence: simple bibliometrics for more efficient academic research evaluations. *Economic Policy*. 2016;31(88):611–52.
97. Rey-Rocha J, Martín-Sempere MJ, Martínez-Frías J, López-Vera F. Some Misuses of Journal Impact Factor in Research Evaluation. *Cortex*. 2001;37(4):595–7.
98. Rothfus M, Sketris IS, Traynor R, Helwig M, Stewart SA. Measuring Knowledge Translation Uptake Using Citation Metrics: A Case Study of a Pan-Canadian Network of Pharmacoepidemiology Researchers. *Science & Technology Libraries*. 2016;35(3):228–40.
99. Rousseau R, Hooydonk G Van. Journal production and journal impact factors. *Journal of the American Society for Information Science*. 1996;47(10):775–80.
100. Schoonbaert D, Roelants G. Citation analysis for measuring the value of scientific publications: quality assessment tool or comedy of errors? *Tropical Medicine & International Health*. 1996;1(6):739–52.

101. Scully C, Lodge H. Impact factors and their significance; overrated or misused? *British Dental Journal*. 2005;198(7):391–3.
102. Seglen PO. Citations and journal impact factors: questionable indicators of research quality. *Allergy*. 1997;52(11):1050–6.
103. Shapiro FR. Origins of bibliometrics, citation indexing, and citation analysis: The neglected legal literature. *Journal of the American Society for Information Science*. 1992;43(5):337–9.
104. Smaldino PE, McElreath R. The natural selection of bad science. *Royal Society Open Science*. 2016;3(160384). Available from: <http://dx.doi.org/10.1098/rsos.160384>
105. Smith DR, Hazelton M. Bibliometrics, citation indexing, and the journals of nursing. *Nursing and Health Sciences*. 2008;10(4):260–5.
106. Smith KM, Crookes E, Crookes PA. Measuring research “impact” for academic promotion: issues from the literature. *Journal of Higher Education Policy and Management*. 2013;35(4):410–20.
107. Statzner B, Resh VH, Kobzina NG. Low impact factors of ecology journals: don’t worry. *Trends in Ecology & Evolution*. 1995;10(5):220.
108. Steele C, Butler L, Kingsley D. The publishing imperative: the pervasive influence of publication metrics. *Learned Publishing*. 2006;19(4):277–90.
109. Symonds MRE, Gemmell NJ, Braisher TL, Gorringer KL, Elgar MA. Gender differences in publication output: Towards an unbiased metric of research performance. *PLoS ONE: Public Library of Science*; 2006;1(1):e127.
110. Tahamtan I, Safipour Afshar A, Ahamdzadeh K. Factors affecting number of citations: a comprehensive review of the literature. *Scientometrics*. 2016;107(3):1195–225.
111. Taylor J. The assessment of research quality in UK universities: Peer review or metrics? *British Journal of Management*. 2011;22(2):202–17.
112. Taylor M, Perakakis P, Trachana V. The siege of science. *Ethics in Science and Environmental Politics*. 2008;8(1):17–40.
113. The Expert Panel on Academic Recognition of Team Science in Canada. *Academic Recognition of Team Science: How to Optimize the Canadian Academic System*. Ottawa; 2017. Available from: <http://www.caahs-acss.ca/academic-recognition-of-team-science-how-to-optimize-the-canadian-academic-system/>
114. Thompson DF, Walker CK. A descriptive and historical review of bibliometrics with applications to medical sciences. *Pharmacotherapy*. 2015;35(6):551–9.
115. Todd PA, Ladle RJ. Hidden dangers of a “citation culture.” *Ethics in Science and Environmental Politics*. 2008;8(1):13–6.

116. Togia A, Tsigilis N. Impact factor and education journals: a critical examination and analysis. *International Journal of Educational Research*. 2006;45(6):362–79.
117. Tol RSJ. Identifying excellent researchers: A new approach. *Journal of Informetrics*. 2013;7(4):803–10.
118. Traynor M, Rafferty AM. Bibliometrics and a culture of measurement. *Journal of Advanced Nursing*. 2001;36(2):167–8.
119. Tsikliras AC. Chasing after the high impact. *Ethics in Science and Environmental Politics*. 2008;8(1):45–7.
120. University of Waterloo Working Group on Bibliometrics. White Paper on Bibliometrics, Measuring Research Outputs through Bibliometrics. Waterloo; 2016. Available from: [https://uwspace.uwaterloo.ca/bitstream/handle/10012/10323/Bibliometrics White Paper 2016 Final_March2016.pdf?sequence=4](https://uwspace.uwaterloo.ca/bitstream/handle/10012/10323/Bibliometrics%20White%20Paper%202016%20Final_March2016.pdf?sequence=4)
121. Van Dalen HP, Henkens K. Intended and unintended consequences of a publish-or-perish culture: A worldwide survey. *Journal of the American Society for Information Science and Technology*. 2012;63(7):1282–93.
122. Van Raan AFJ. Fatal attraction: Conceptual and methodological problems in the ranking of universities by bibliometric methods. *Scientometrics*. 2005;62(1):133–43.
123. Vanclay JK. Impact factor: Outdated artefact or stepping-stone to journal certification? *Scientometrics*. 2012;92(2):211–38.
124. Vanclay JK. Factors affecting citation rates in environmental science. *Journal of Informetrics*. 2013;7(2):265–71.
125. Von Bohlen und Halbach O. How to judge a book by its cover? How useful are bibliometric indices for the evaluation of “scientific quality” or “scientific productivity”? *Annals of Anatomy*. 2011;193(3):191–6.
126. Wainer J, Billa C, Goldenstein S. Invisible work in standard bibliometric evaluation of computer science. *Communications of the ACM*. 2011;54(5):141.
127. Wainer J, Vieira P. Correlations between bibliometrics and peer evaluation for all disciplines: The evaluation of Brazilian scientists. *Scientometrics*. 2013;96(2):395–410.
128. Wallin JA. Bibliometric Methods: Pitfalls and Possibilities. *Basic & Clinical Pharmacology & Toxicology*. 2005;97(5):261–75.
129. Wang J, Veugelers R, Stephan P. Bias against novelty in science: A cautionary tale for users of bibliometric indicators. *Research Policy*. 2017;46(8):1416–36.
130. Webster NR. Bibliometrics and assessing performance and worth. *British Journal of Anaesthesia*. 2011;107(3):306–7.

131. Wendl MC. H-index: however ranked, citations need context. *Nature*. 2007;449(7161):403.
132. Wilsdon J, Allen L, Belfiore E, Campbell P, Curry S, Hill S, et al. *Metric Tide: Report of the Independent Review of the Role of Metrics in Research Assessment and Management*. 2015. Available from: <http://www.hefce.ac.uk/pubs/rereports/year/2015/metrictide/>
133. Wixted B, Beaudry C. "Capturing the impacts" of Research: Discussion paper. 2012. Available from: http://www.sshrc-crsh.gc.ca/about-au_sujet/publications/Compendium_e.pdf
134. Yu G, Wang L. The self-cited rate of scientific journals and the manipulation of their impact factors. *Scientometrics*. 2007;73(3):321–30.
135. Zitt M, Bassecoulard E. Challenges for scientometric indicators: Data demining, knowledge-flow measurements and diversity issues. *Ethics in Science and Environmental Politics*. 2008;8(1):49–60.

Appendix. Appendix H

Appendix H

LETTER OF UNDERSTANDING

Re: Joint Committee on Metrics

BETWEEN:

THE UNIVERSITY OF MANITOBA (Hereinafter referred to as “the University”)

-and-

THE UNIVERSITY OF MANITOBA FACULTY ASSOCIATION (Hereinafter referred to as “UMFA”)

1. The University of Manitoba believes that, in light of new provisions in the collective agreement enhancing collegial participation in the creation of tenure and promotion criteria, Members should have the ability to decide on the appropriate use (if any) of research metrics (as defined in s. 19.D.1.8.1) in tenure, promotion, and performance evaluation processes.
2. The University of Manitoba Faculty Association believes that any such requirements pose risks of unfair, inaccurate, and/or discriminatory processes of assessment in tenure, promotion, and performance review. It further believes that such requirements pose significant risks to the quality and integrity of research, scholarship, and education. It believes that therefore, no such requirement should be imposed on Members, nor should Members be penalized for refusing to submit such metrics as part of tenure, promotion, or performance review.
3. The University of Manitoba and the University of Manitoba Faculty Association agree that a Joint Committee will be struck to examine issues related to the collection and use of research metrics in evaluative processes (such as tenure, promotion, and performance review).
4. The University and the Association will each nominate three members (each with academic rank below the level of a Dean, and who are not on the UMFA executive) to the committee, including one to act as co-chair.
5. In the course of conducting its work, the Joint Committee will:
 - a) examine the academic and scholarly literature relevant to the appropriateness, risks and benefits of the collection and use of research metrics as a means of assessment;

- b) examine the treatment of research metrics at other Canadian Universities;
 - c) consult with members of the academic community representing a diversity of disciplines;
 - d) consult with members of the academic community with experience and/or relevant academic expertise relating to the collection and use of research metrics.
6. The Joint Committee will submit a report to the President of the University and the President of the Association containing its findings and recommendations no later than December 31, 2017.
 7. Should the committee recommend it, with no more than two (2) members dissenting, the University and the Association agree to the immediate addition of the following to the Collective Agreement:
 - 19.D.1.8.5 No tenure criteria shall require the submission of a research metric, nor shall they include any standard or recommended expectation based on a quantitative measure.
 - 19.D.1.8.6 Research metrics will only be used as part of evaluation and/or assessment when such metrics are personally and voluntarily submitted by the Member.
 - 19.D.1.8.7 There shall be no penalty or adverse inference to any Member for the non-submission of research metrics as part of the tenure application.
 - 20.A.2.5.5 No promotion criteria shall require the submission of a research metric, nor shall they include any standard or recommended expectation based on a quantitative measure.
 - 20.A.2.5.6 Research metrics will only be used as part of evaluation and/or assessment when such metrics are personally and voluntarily submitted by the Member.
 - 20.A.2.5.7 There shall be no penalty or adverse inference to any Member for the non-submission of research metrics as part of the promotion application.
 - 20.B.1.6.5 No promotion criteria shall require the submission of a research metric, nor shall they include any standard or recommended expectation based on a quantitative measure.
 - 20.B.1.6.6 Research metrics will only be used as part of evaluation and/or assessment when such metrics are personally and voluntarily submitted by the Member.
 - 20.B.1.6.7 There shall be no penalty or adverse inference to any Member for the non-submission of research metrics as part of the promotion application.

- 35.5.5 No evaluation criteria shall require the submission of a research metric, nor shall evaluation criteria include any standard or recommended expectation based on a quantitative measure.
- 35.5.6 Research metrics will only be used as part of evaluation and/or assessment when such metrics are personally and voluntarily submitted by the Member.
- 35.5.7 There shall be no penalty or adverse inference to any Member for the non-submission of research metrics as part of an evaluation.

8. Notwithstanding s. 7, the Joint Committee shall be free to come to any other recommendations or conclusions it desires, including alternative wording for the principles described in s. 7. Any recommendations under this section will be advisory only.

9. This letter of understanding shall be attached to and form part of the Collective Agreement.

10. This letter of understanding will expire on January 12, 2018. This letter of understanding shall be attached to and form part of the Collective Agreement and the period in which s. 10(4) of the Labour Relations Act applies.

11. The terms of this letter of understanding are a term and condition of employment. This letter of Understanding is intended to survive the expiry of the Collective Agreement.

DATED at the City of Winnipeg in the
Province of Manitoba

University of Manitoba

this ____ day of _____ 2017

University of Manitoba Faculty Association