

EDITORIAL

**An updated h-index measures both the primary
and total scientific output of a researcher**

Octavian Bucur^{1,}, Alex Almasan^{2,*}, Roman Zubarev³, Mark Friedman⁴, Garth L. Nicolson⁵, Pavel Sumazin⁶, Mircea Leabu⁷, Barbara S. Nikolajczyk⁸, Dorina Avram⁹, Tanja Kunej¹⁰, George A. Calin¹¹, Andrew K. Godwin¹², Hans-Olov Adami¹³, Peter G. Zaphiropoulos¹⁴, Des R. Richardson¹⁵, Gerold Schmitt-Ulms¹⁶, Håkan Westerblad¹⁷, Megan Keniry¹⁸, Georges E. R. Grau¹⁹, Salvatore Carbonetto²⁰, Radu V. Stan²¹, Aurel Popa-Wagner²², Kasumov Takhar²³, Beverly W. Baron²⁴, Paul J. Galardy²⁵, Feng Yang²⁶, Dipak Data²⁷, Oluwole Fadare²⁸, KT Jerry Yeo²⁹, Georgiana R. Gabreanu³⁰, Stefan Andrei³¹, Georgiana R. Soare³², Mark A. Nelson³³, Elisa A. Liehn^{34,*}*

¹Department of Pathology, Harvard Medical School and Beth Israel Deaconess Medical Center, Boston, MA, USA; ²Department of Cancer Biology, Lerner Research Institute & Department of Radiation Oncology, Taussig Cancer Institute, Cleveland Clinic, Cleveland, Ohio, USA; ³Department of Medical Biochemistry & Biophysics, Karolinska Institutet, Stockholm, Sweden; ⁴Department of Pathology, St. Luke's-Roosevelt Hospital Center, Beth Israel Medical Center and Icahn School of Medicine at Mount Sinai, New York, NY, USA; ⁵Department of Molecular Pathology, The Institute for Molecular Medicine, Huntington Beach, CA, USA; ⁶Department of Pediatrics, Baylor College of Medicine, Houston, TX, USA; ⁷Department of Cellular and Molecular Medicine, University of Medicine and Pharmacy "Carol Davila" and Victor Babes National Institute of Pathology, Bucharest, Romania; ⁸Department of Microbiology, Boston University School of Medicine Boston, MA, USA; ⁹Department of Medicine, Division of Pulmonary, Critical Care and Sleep Medicine, University of Florida, Gainesville, FL, USA; ¹⁰Department of Genetics, Animal Biotechnology and Immunology, Domzale, Slovenia; ¹¹Department of Experimental Therapeutics and Leukemia & Center for RNA Interference and Non-Coding RNAs, University of Texas MD Anderson Cancer Center, Houston, TX, USA; ¹²Department of Pathology and Laboratory Medicine, University of Kansas Cancer Center Kansas City, KS, USA; ¹³Department of Medical Epidemiology and Biostatistics, Karolinska Institute, Stockholm, Sweden; ¹⁴Department of Biosciences and Nutrition, Karolinska Institute, Sweden; ¹⁵Department of Pathology and Bosch Institute, University of Sydney, Sydney, Australia; ¹⁶Department of Laboratory Medicine & Pathobiology, Tanz Centre for Research in Neurodegenerative Diseases, University of Toronto, Toronto, Canada; ¹⁷Department of Physiology and Pharmacology, Karolinska Institute, Stockholm, Sweden; ¹⁸Biology Department, University of TEXAS Pan American, Edinburg, TX, USA; ¹⁹Department of Pathology, University of Sydney, Sydney, Australia; ²⁰Centre for Research in Neuroscience, McGill University Health Sciences Centre, Montreal, Canada; ²¹The Geisel School of Medicine at Dartmouth, Dartmouth College, Lebanon, NH, USA; ²²Department of Psychiatry, Rostock University Medical School, Rostock, Germany; ²³Department of Gastroenterology and Hepatology, Cleveland Clinic, Cleveland, OH, USA; ²⁴Department of Pathology, University of Chicago, Chicago, IL, USA; ²⁵Pediatric Hematology/Oncology, Mayo Clinic Transplant Center, Mayo Clinic, Rochester, MI, USA; ²⁶Department of Molecular and Cellular Biology, Baylor College of Medicine, Houston, TX, USA; ²⁷Biochemistry Division, CSIR-Central Drug Research Institute, Lucknow, India; ²⁸Department of Pathology, University of California, San Diego, San Diego, CA, USA; ²⁹Department of Pathology, University of Chicago, Chicago, IL, USA; ³⁰"Victor Babes" National Institute of Pathology and Biomedical Sciences, Bucharest, Romania; ³¹Service de Médecine Interne, CNHO des Quinze-Vingts, Paris, France; ³²Department of Obstetrics and Gynaecology, Bucharest University Emergency Hospital, Bucharest, Romania; ³³Department of Pathology, The University of Arizona, Tucson, AZ, USA; ³⁴Institute for Molecular Cardiovascular Research (IMCAR), RWTH Aachen University, Aachen, Germany;

*Corresponding authors:

Octavian Bucur, MD, PhD, Department of Pathology, Harvard Medical School and Beth Israel Deaconess Medical Center, 330 Brookline Ave, Boston, MA, 02215, USA. E-mail: obucur@bidmc.harvard.edu

Alexandru Almasan, PhD, Department of Cancer Biology, Lerner Research Institute, Cleveland Clinic, 9500 Euclid Avenue, Cleveland, OH 44195, USA. Tel: +1 216 444 9970; Fax: +1 216 445 6269; E-mail: almasaa@ccf.org

Elisa A. Liehn, MD, PhD, Institute for Molecular Cardiovascular Research (IMCAR), University Hospital Aachen, Rheinisch-Westfälische Technische Hochschule Aachen, Pauwelsstrasse 30, 52074 Aachen, Germany. Tel.: 0049-241-80 35983; Fax: 0049-241-80 82716; E-mail: eliehn@ukaachen.de

Abstract

The growing interest in scientometry stems from ethical concerns related to the proper evaluation of scientific contributions of an author working in a hard science. In the absence of a consensus, institutions may use arbitrary methods for evaluating scientists for employment and promotion. There are several indices in use that attempt to establish the most appropriate and suggestive position of any scientist in the field he/she works in. A scientist's Hirsch-index (h-index) quantifies their total effective published output, but h-index summarizes the total value of their published work without regard to their contribution to each publication. Consequently, articles where the author was a primary contributor carry the same weight as articles where the author played a minor role. Thus, we propose an updated h-index named Hirsch(p,t)-index that informs about both total scientific output and output where the author played a primary role. Our measure, $h(p,t) = h(p)h(t)$, is composed of the h-index $h(t)$ and the h-index calculated for articles where the author was a key contributor; i.e. first/shared first or senior or corresponding author. Thus, a $h(p,t) = 5,10$ would mean that the author has 5 articles as first, shared first, senior or corresponding author with at least 5 citations each, and 10 total articles with at least 10 citations each. This index can be applied in biomedical disciplines and in all areas where the first and last position on an article are the most important. Although other indexes, such as r- and w-indexes, were proposed for measuring the authors' output based on the position of researchers within the published articles, our simpler strategy uses the already established algorithms for h-index calculation and may be more practical to implement.

Introduction

The current scientific community needs strategies to rank the individual contribution of authors in order to most accurately appreciate the value of their efforts. Scientometry has developed as a useful field for objectively evaluating published scientific work, mainly in hard sciences, including biomedical research. However, as it is expected in science, a number of criticisms were proposed for many of the scientometric indices. Although none of these indices is perfect by itself, a complex assessment can be helpful to most accurately appreciate any

author's published work. Moreover, a careful evaluation is critical for the career of scientists, as well as for employers in academia and science alike.

Why do we need an updated h-index?

The h-index, introduced by professor J.E. Hirsch in 2005, is now widely accepted and employed. The h-index of a scientist has index x if x of his/her published articles have at least x citations each and his/her remaining articles have $\leq x$ citations each¹. The h-index is a measure of the total effective research output of a scientist¹. However, the h-index does not take into consideration the order of the authors within the published articles, which in some cases can lead to over- or under-evaluation of the importance of an author's scientific output^{2,3}. An author's rank is important in most research fields, including biomedical research, with the first and last (senior or corresponding) authors being the most important contributors to an article. Thus, an updated h-index that includes both a measure of the primary output of an author (as the first, shared first, senior or corresponding author) and the total output of the author (all articles regardless of an author's position – the current h-index) is needed and would be a useful tool in the evaluation of biomedical publication output.

Currently proposed indexes for evaluation of research output based on author rank.

Other indexes have been proposed for evaluation of research output based on an author's rank, including the r-index and w-index^{2,3,4}. The *revised h-index (r-index)* for biomedical research was proposed in 2012 by professor A.A. Romanovski, where the first and last author positions are evaluated as being 4 times more important than middle author positions, within a published article². Another index, the *weighted h-index (w-index)* was proposed in 2009 by professor C.T. Zhang³ and is calculated taking into account the position of each author and giving a weighted coefficient of 1 for the first and last (corresponding) authors and linear decreasing coefficient numbers for the authors in positions 2, 3, ..., $n-1$, where n is the number of the authors. Both methods are useful in many cases, but may prove inaccurate in some specific situations^{2,3}, including ones where shared first authors or multiple corresponding/senior authors are concerned.

$$\mathbf{Hirsch(p,t) = h(p,t) = h(p), h(t)}$$

h(p) is the h-index calculated based on the articles as first/shared first and last/shared last author (primary output)

h(t) is the h-index calculated based on all articles published (known as h-index, or total research output)

Example: h(p,t)=5,10 means that the author has an h-index of 5 for his/her articles published as first and last author, while having an h-index of 10 for all articles

Our proposed h(p,t) index measures both the primary output [h(p) index] and total output [h(t) index].

We propose here an updated Hirsch factor, termed Hirsch(p,t), or h(p,t), that calculates both the primary research output (h(p) index) and the total research output (h(t) index, currently known as h-index) of an author. h(p,t) = h(p),h(t) is composed of the h-index h(t) calculated for all articles published by the author, and the h-index h(p) calculated for articles where the author is a key contributor: i.e. first/shared first or last/shared last (senior or corresponding) author.

For example, an author with 10 published articles, of which 2 are as first author and on 1 he/she is the last (corresponding) author, each of the articles being cited more than 10 times, would have an h(p,t)-index of h(p,t)=2+1,10=3,10. This means that the author has 3 articles as first and last author with at least 3 citations each (primary research output), and 10 total articles with at least 10 citations each (total research output).

The h(p,t) index will be very helpful in distinguishing the authors with high primary research output from those with low primary research output, both when their total h indexes have similar or very distinct values.

How will this updated h-index help the research community and evaluation of a researcher's scientific output?

Despite its flaws, the currently used h-index is a useful way of measuring a researcher's total publication output, whereas the primary output as first, shared first, last and shared last author is not measured^{2,3,5,6}. Since the employed algorithm used for the h-index calculation is already recognized as useful by the scientific community and is widely used, we believe that addition of the primary

research output information to this would be also useful. This can be done by using the same h-index calculation for the total output, adding primary output, and showing them together in the same term.

Why do we think that the h(p,t) index should be employed by the research community?

Using an already accepted algorithm plus additional information on primary research output could allow the scientific community to accept this small, but significant improvement faster than would likely be the case with other proposed and less tested indexes. Moreover, other proposed indexes may miss the already proven usefulness of the h-index.

In our opinion, assigning weight coefficients for the position of authors within an article (r-index or w-index)^{2,3}, although useful in many cases, is definitely prone to errors. For example, it is hard to correctly estimate and assign quantitative weight coefficients to compare the work of authors in the same manuscript, even when the contribution of each author is described (which is only true in a small number of journals). Assigning the same weight coefficient for the authors on the same position from different papers could be helpful, but it is not completely accurate. However, such quantitative indexes may complement h- or h(p,t)-indexes in the future.

It is worth pointing out that our strategy takes into consideration not only the first and last authors, but also the shared first or shared last (senior or corresponding) author positions. Although the automatic detection of the first and last author is not a problem, automatic detection of the shared first author or shared senior/corresponding authors by established databases, such as Scientific Citation Index, Google Scholar or Scopus, may be a challenge. A potential

solution would be the enabling of limited editing by authors to manually identify shared authorship contributions to improve the accuracy of the h(p,t) index.

In conclusion, we describe a useful and simpler strategy of evaluating researchers' scientific output based on the calculation of the h-index for both the primary and total research output. This strategy can be applied in all fields of research where the first and last authors are the most important authors within published articles, including biomedical research.

Acknowledgements:

Dr. *Octavian Bucur* was supported by the Lady Tata Memorial Trust Award, London, UK. Dr. *Alexandru Almasan* is funded by the NCI grant RO1 CA184137-01. Dr. *Tanja Kunej* is supported by the Slovenian Research Agency (ARRS) through the Research programme (P4-0220). Dr. *George Calin* is The Alan M. Gewirtz Leukemia & Lymphoma Society Scholar. Work in Dr. George Calin's laboratory is supported in part by the NIH/NCI grants 1UH2TR00943-01 and 1 R01 CA182905-01, the UT MD Anderson Cancer Center SPORE in Melanoma grant from NCI (P50 CA093459), Aim at Melanoma Foundation and the Miriam and Jim Mulva research funds, the Brain SPORE (2P50CA127001), the Center for Radiation Oncology Research Project, the Center for Cancer Epigenetics Pilot project, a 2014 Knowledge GAP MDACC grant, a CLL Moonshot pilot project, the UT MD Anderson Cancer Center Duncan Family Institute for Cancer Prevention and Risk Assessment, a SINF grant in colon cancer, the Laura and John Arnold Foundation, the RGK Foundation and the Estate of C. G. Johnson, Jr. Dr. *Andrew K. Godwin* is the Chancellor's Distinguished Chair in Biomedical Sciences Endowed Professor and a Kansas Bioscience Authority Eminent Scholar. Work in Dr. Godwin's program is supported in part by the NCI (R21CA186846, R01CA106588, R01CA140323, R01CA140323-S1, U01CA113916, & P30CA168524), the NIH (UL1TR000001), the Braden's Hope Foundation, the Vicki Welsh Fund, and the OVERRUN Ovarian Cancer Foundation. Dr. *Des Richardson* is funded by the National Health and Medical Research Council (NHMRC) Senior Principal Research Fellowship and Project Grants. *Peter G. Zaphiropoulos* is supported by the Swedish Childhood Cancer Foundation (PROJ12/032) and the AFA Insurance (Dnr: 130014). Dr. *Håkan Westerblad* is funded by the Swedish Research Council and Swedish National Center for Sports Research. Dr. *Radu V. Stan* is funded by the NIH CA172983 (PI: Stan), CA023108 (PI: Israel) grants. *Takhar Kasumov* is supported by the NIH/NIGM (1R01 GM112044-01A1), NIH/NHBLI (1R21 HL114407-

01A1) grants. *Georgiana R. Gabreanu* is funded by the Research Project: PN 09.33 - 01.01. *Dorina Avram* is funded by the NIH (R01 AI067846, R21 AI112492). *Gerold Schmitt-Ulms* is funded by the Canadian Institutes for Health Research (CIHR). *Paul J. Galarly* is supported by the National Cancer Institute (CA151351), the Multiple Myeloma Research Foundation, the Gabrielle's Angel Foundation for Cancer Research, and the Hyundai Hope on Wheels Foundation. Dr. Galarly is a former ASH Basic Research Scholar, a past recipient of the Howard Hughes Medical Institute Physician Scientist Early Career Award, and a former Harriet H. Samuelsson Foundation Pediatric Cancer Research Scientist. *Dipak Datta* is funded by the CSIR-Central Drug Research Institute Fund. Dr. *Elisa A. Liehn* is funded by the Interdisciplinary Centre for Clinical Research IZKF Aachen within the Faculty of Medicine at RWTH Aachen University.

Conflict of Interest:

The authors declare no conflict of interest.

Keywords:

Updated h-index, Hirsch index, h(p,t), h(p), h(t), primary scientific output

References:

1. Hirsch JE. An index to quantify an individual's scientific research output. *Proc Natl Acad Sci U S A*. 2005 Nov 15;102(46):16569-72.
2. Romanovsky AA. Revised h index for biomedical research. *Cell Cycle*. 2012 Nov 15;11(22):4118-21.
3. Zhang CT. A proposal for calculating weighted citations based on author rank. *EMBO Rep*. 2009 May;10(5):416-7
4. Sekercioglu CH. Quantifying coauthor contributions. *Science*. 2008 Oct 17;322(5900):371
5. Saleem T. The Hirsch index - a play on numbers or a true appraisal of academic output? *Int Arch Med*. 2011 Jul 7;4:25
6. Carbon CC. The Carbon_h-factor: predicting individuals' research impact at early stages of their career. *PLoS One*. 2011;6(12):e28770.

Citation: Bucur O et al. An updated h-index measures both the primary and total scientific output of a researcher. *Discoveries* 2015, Jul-Sep; 3(3): e50. DOI: 10.15190/d.2015.42

DISCOVERIES is a peer-reviewed, open access, online, multidisciplinary and integrative journal, publishing high impact and innovative manuscripts from all areas related to MEDICINE, BIOLOGY and CHEMISTRY; © 2015, Applied Systems