Biostatistics Research
https://ojs.wiserpub.com/index.php/BSR/

UNIVERSAL WISER PUBLISHER

Research Article

Author Impact Factor: A Framework for Evaluating Authorship and Scientific Contribution

Brian C. Drolet[1*], **Alan T. Makhoul**[2]

[1]Department of Plastic Surgery, Department of Biomedical Informatics and Center for Biomedical Ethics and Society, Vanderbilt University Medical Center, Nashville, Tennessee, United States
[2]Vanderbilt University School of Medicine, Nashville, Tennessee, United States
E-mail: brian.c.drolet@gmail.com

Received: 16 August 2021; **Revised:** 27 October 2021; **Accepted:** 8 November 2021

Abstract: In the decade since Hirsch defined *h*, there has been widespread acceptance of the *h*-index as a bibliometric indicator. Although the *h*-index has been validated in numerous applications and settings, the bibliometric has some important limitations. Most importantly, the *h*-index does not account for authors' individual contributions to manuscripts within their *h*-defining body of work. Since each author makes a variable contribution to a piece of scientific work, an author-adjusted index would more fairly reflect scholarly productivity. We propose the author impact factor (AIF), which accounts for authorship position and number of co-authors, to adjust the *h*-index and more fairly account for contributions to the body of scientific work. The AIF is calculated from the *h*-index and an author's proportional contribution (α) to each *h*-defining manuscript. The α is based on authorship position and the number of co-authors. Using the golden ratio (φ), the calculation of α for each *h*-index defining manuscript is simple and axiomatic. To demonstrate the utility of this index, we calculated the AIF for a sample of high-impact scientists. The results show that the AIF maintains all the benefits of the *h*-index while adjusting the bibliometric for author-specific factors. Therefore, AIF more accurately reflects total "research output" and can be used to better compare authors' scientific contributions.

Keywords**:** citation, bibliometric, index, impact

1. Introduction

Bibliometrics are analytical methods used to study the interconnectivity and structure of scholarly work. With the rapid growth of scientific knowledge, greater attention is being given to "evaluative bibliometrics" that measure scientific productivity and research impact [1-5]. Bibliometric indices like journal impact factor and author-specific metrics, such as citation frequency and the "*h*-index", have been described and are commonly used to measure the importance of a journal or individual scientist's research output [6-8]. To be effective, these indices must provide an accurate and valid measurement of the variable (i.e., journal or author impact) for which they were created. Despite widespread usage, the validity of common bibliometrics remains widely debated [9-14].

One of the most widely used measurements of individual research impact is the *h*-index. Described by Jorge Hirsch in 2005, it was the first bibliometric to account for two factors related to a researcher's scholarly activity: productivity (i.e., number of papers) and impact (i.e., frequency of citation). The metric was defined as *h*, which is the number of papers (N_p) that have been cited at least *h* times [6]. In other words, if an author's *h*-index is 10, that author has at least

10 papers that have each been cited 10 times.

Hirsch's method allows for a single metric to combine productivity and impact in equal weight. In doing so, the *h*-index controls for "inflation" that might result from a small number of highly cited manuscripts (i.e., citation-based system) or from many low impact manuscripts (i.e., publication-based system). The result is a power function, in which an author with an *h*-index of X has at least X publications and X^2 total citations.

With the rapid advancement of online cataloging, computation of bibliometrics (including *h*-index) has become rapid and simple. For example, GoogleScholar ® reports that Dr. Hirsch has published nearly 300 papers with more than 22,000 total citations; and he has an *h*-index of 63. Neither total citation count nor the volume of publications increases the *h*-index without a concomitant increase in the N_p that has been cited at least *h* times [6]. Mathematically, the *h*-index appears to have many benefits over one-dimensional bibliometrics [15]. The *h*-index also appears to have real-world validity, with numerous studies demonstrating a correlation between *h* and application success, promotion, academic rank and even future publication in numerous specialties [16-25]. Nevertheless, the *h*-index has been criticized for various weaknesses and numerous revisions have been proposed, even by Hirsch himself [26-28].

By design, the *h*-index measures individual scholarly productivity (i.e., N_p and citation count) irrespective of co-authorship. Consequently, the *h*-index promotes the listing of additional co-authors regardless of contribution. Because researchers often cite their own works, papers with many co-authors may also receive more citations. To track collaboration, other bibliometrics, such as the collaborative coefficient, have been created that measure the mean number of authors per paper and the proportion of multi-author works. However, opportunities for improvement remain.

In particular, the *h*-index does not account for proportional contributions among any number of co-authors on a single manuscript. The *h*-index assigns the same work-product for a first author as all other authors on the manuscript. However, this is rarely an accurate reflection of each author's contribution, which is ultimately the purpose of the bibliometric.

In this study, we sought to define and validate a revised, author-specific bibliometric, called the author impact factor (AIF), which can be used to measure individual research productivity with greater sensitivity to author contribution.

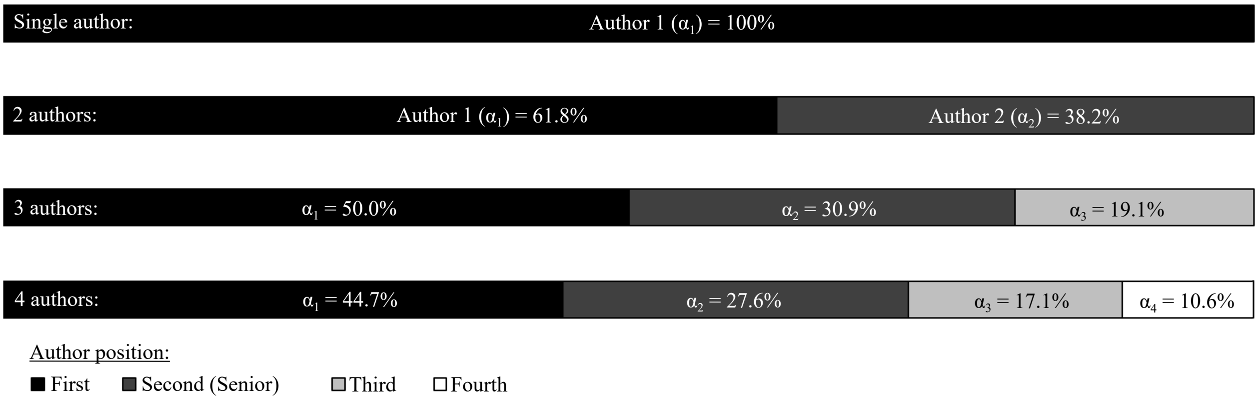
2. Framework

In the current form, the *h*-index can be graphed as a two-dimensional Cartesian function, which grows only as an integer when both the N_p and the concomitant citation frequency of specific papers increases. We propose the addition of an author's proportional contribution (α) to adjust for authorship position on each manuscript comprising a scientist's total *h* (as defined by Hirsch). To define α, we propose that there is a fixed, proportional difference in an author's contribution to a publication based on the sequentially listed position in the publication. In other words, each author contributes a factor (φ) greater than the subsequently listed author (or φ times less than the preceding author). A simple distribution curve is shown in Table 1 when the golden ratio (φ = 1.61803…) is applied to this function.

Table 1. Distribution of author's proportional contribution (α)

Number of authors	Author position							
	First	Second (Senior)	Third	Fourth	Fifth	Sixth	Seventh	Eighth
1	100.0%							
2	61.8%	38.2%						
3	50.0%	30.9%	19.1%					
4	44.7%	27.6%	17.1%	10.6%				
5	42.0%	25.9%	16.0%	9.9%	6.1%			
6	40.5%	25.0%	15.5%	9.5%	5.9%	3.6%		
7	39.6%	24.4%	15.1%	9.3%	5.8%	3.6%	2.2%	
8	39.0%	24.1%	14.9%	9.2%	5.7%	3.5%	2.2%	1.3%

Also called the Divine Function, φ is found frequently in nature, science and mathematics, making it a reasonable adjusting factor [29, 30]. The golden ratio can naturally be applied to the authorship sequence, in which the proportional contribution (ratio) of any two sequential authors is equal to the ratio of the prior and subsequent authors (Figure 1).



$$\alpha_n = \varphi^{N-n} \div \sum_{n}^{N-1} \varphi^{N}$$

$$\text{N = total number of authors, n = author position, } \varphi = (1+\sqrt{5}) \div 2 = 1.61803\ldots$$

$$\text{Sum of authors' proportional contributions } (\alpha): \sum_{o}^{N} \alpha = \alpha_1 + \alpha_2 + \ldots + \alpha_N = 100\%$$

$$\text{Fixed proportional ratio } (\varphi): \alpha_1 \div \alpha_2 = \alpha_2 \div \alpha_3 = \ldots = \alpha_{N-1} \div \alpha_N = 1.61803$$

Figure 1. Author's proportional contribution (α)

In a single author paper, the first author warrants 100% proportional credit (α = 1.0). For multiple author papers, each subsequent author is given fractional credit (α_n) towards the total research "product" as determined by the proportional contribution formula (Figure 1).

This framework supposes that authors are listed sequentially (i.e., credited) in descending order of their contribution to the total work. Although this is the most common method of indicating credit, it is not standard for all fields [31-35]. An important exception to the rule of sequential author listing applies to the "senior" or "last" author. Although a senior author may make more limited contributions to an individual manuscript, their contribution to the research endeavor at large cannot be underestimated. Most research endeavors require pre-requisite investments from a senior author, such as grant funding, laboratory space, a patient population, prior research groundwork, mentorship, or any number of other critical (and sometimes intangible) factors. Therefore, wherever the senior author is listed in the "last" position, we propose that this should be valued as the second most important contribution following the first author.

The AIF is then calculated by averaging the α from each paper comprising the author's *h*-index defining manuscripts (as defined by Hirsch) and then multiplying this average α to the *h*-index. Since α is always positive but less than or equal to 1, the α-adjustment will decrease the *h*-index to represent the author's proportional contribution to their index.

3. Framework application

To test the real-world application of this method, a sample of high impact authors (*h*-index greater than 150) from Google Scholar was cross-referenced with SCOPUS. All reference data for each author was imported into Microsoft Excel (Redmond, WA) and α was calculated for each *h*-index defining manuscript. The average α was calculated and multiplied by the *h*-index to create the AIF for each author. Additional descriptive statistics were calculated to demonstrate the correlation of AIF to author position and number of co-authors.

Table 2 demonstrates that AIF is significantly different from the *h*-index for ranking the overall impact of individual authors. The highest AIF author had 40 fewer *h*-index defining manuscripts than others on the list but significantly more first and senior author manuscripts. Higher AIF authors generally had fewer co-authors.

Table 2. High impact author AIF and *h*-index comparison

Author initials	Author impact factor	*H*-index	*H*-index rank	Avg. author proportional contribution (α)	Avg. # co-authors	No. first	No. senior	Total citations (10^3)
PB	61.43	160	16	38.39%	4.54	51	69	173.0
SS	55.97	187	7	29.93%	3.42	16	141	210.0
MM	54.09	150	26	36.06%	4.33	50	74	123.7
RK	51.69	179	8	28.88%	5.20	66	49	258.6
ZW	49.99	151	25	33.10%	3.75	28	86	152.0
KF	49.07	163	14	30.10%	3.90	49	48	168.8
MG	48.92	194	5	25.22%	5.02	11	128	175.9
PR	45.75	155	22	29.51%	28.12	59	44	192.1
XY	44.76	158	18	28.33%	3.54	18	116	143.2
TR	43.58	167	11	26.09%	3.87	19	90	128.0
GK	42.57	160	15	26.61%	16.93	16	119	158.4
RL	38.64	178	9	21.71%	5.04	11	90	211.4
FH	37.20	158	17	23.55%	22.08	32	63	155.1
AH	37.01	151	24	24.51%	3.83	8	91	167.1
CF	34.52	156	21	22.13%	3.94	15	52	148.9
CT	32.64	151	23	21.62%	15.50	6	92	135.5
RF	31.91	194	4	16.45%	6.63	1	93	167.9
BV	31.49	202	3	15.59%	9.29	12	82	306.2
MM	31.40	164	13	19.15%	5.59	9	79	171.3
GC	30.00	207	2	14.49%	8.43	21	33	252.6
MS	26.92	210	1	12.82%	6.30	10	33	170.8
DL	24.83	156	20	15.91%	15.58	11	63	156.7
EL	23.53	192	6	12.26%	37.93	11	54	287.3
RF	22.51	164	12	13.73%	5.54	2	66	133.3
CC	21.46	156	19	13.76%	10.32	4	64	177.9
KK	17.43	171	10	10.19%	10.51	9	33	198.1

These data indicate that the AIF can adjust for author position and number of co-authors, two critical factors in a scientist's contribution to the literature. Importantly, the AIF considers not only all the variables of the *h*-index but also these important author specific variables.

4. Discussion

As indicated by the International Committee of Medical Journal Editors (ICMJE), contributions from authors include conceptualization, research design and execution, data acquisition and analysis, manuscript authorship and oversight [36, 37]. Upon completion, each author listed on a manuscript has contributed variably to the total work represented by the publication. The *h*-index cannot account for the differing contributions by each author. A first author is considered equal to any middle author when using this evaluative bibliometric. A modified, author-adjusted impact factor would more accurately reflect scientific contribution and productivity. Using our framework (AIF), a single-author manuscript represents the most significant proportional contribution ($\alpha = 1$) for an author. As the number of authors grows, the proportional contribution of each subsequent author (including the first author) decreases.

There are several benefits to adjusting the *h*-index for author contribution. First, the AIF more fairly and accurately represents the contributions of all authors. The *h*-index is fundamentally flawed in awarding the same 'credit' to a first-author as an N^{th}-author. If AIF became the standard measure, scientists would be discouraged from adding marginally contributing co-authors to a manuscript because of the deflating effect on α for all authors, including the higher-order authors. This effect is most dramatically seen in the middle and later authors. In fact, when using AIF to judge productivity, there is little value (i.e., contribution to total AIF) to being listed below the fifth author. Therefore, authors are encouraged to limit authorship to only significant contributors, thereby limiting ghost authorship and academic inflation. Unlike collaborative bibliometrics, such as the collaborative coefficient and the collaborative index, which incentivize co-authorship, the AIF promotes rigorous authorship criteria. This discussion is particularly relevant as evaluative bibliometrics become more commonly used in academic promotion, and may even play a causative role in academic success. Moreover, honorary-authorship and other borderlines, or outright unethical, authorship practices would likely disappear.

5. Limitations

The α-adjustment in AIF has several important limitations. First, α assumes that there is a standard "author contribution" to an individual manuscript for each author, which is dependent on the author position and the number of authors. For any individual manuscript, this description may not be accurate. However, in aggregate this assumption appears valid.

While choosing φ may seem arbitrary, the golden ratio is found throughout math, science, and nature. More importantly, the outcome for real-world application appears to be a reasonable model (Table 1). Although other ratios may also be valid, the diminishing value of subsequent authorship is the key characteristic of the AIF.

For larger studies, which involve many authors, α may overly devalue the later authors' contributions relative to the earlier position authors. This may argue for a logarithmic function for high author manuscripts. However, we believe this would diminish the application of the function as a simple modification to the *h*-index, designed to create a useable bibliometric index.

Lastly, the AIF, by design, disincentivizes the addition of co-authors. While this may be seen as punitive, the AIF may ultimately promote more meaningful contributions among collaborators.

6. Conclusion

Evaluative bibliometrics, like the *h*-index, can be used to measure contributions to the body of science and compare the impact or productivity of individual authors. However, the *h*-index is limited to measuring the number of publications and frequency of citation; it does not account for authorship position and number of co-authors. Using an author's proportional contribution factor (α) computed from the golden ratio, the *h*-index can be "author-adjusted" to measure productivity more fairly and accurately by accounting for individual authors' contribution to the work. The AIF maintains all the benefits of the *h*-index but adds a factor accounting for authors' contributions.

References

[1] Narin F. *Evaluative Bibliometrics: The Use of Publication and Citation Analysis in the Evaluation of Scientific Activity*. Washington D.C.: Computer Horizons Publishing; 1976.

[2] Bornmann L, Williams R. How to calculate the practical significance of citation impact differences? An empirical example from evaluative institutional bibliometrics using adjusted predictions and marginal effects. *Journal of Informetrics*. 2013; (7): 562-574. Available from: doi:10.1016/j.joi.2013.02.005.

[3] Sengupta I. The growth of biophysical literature. *Scientometrics*. 1985; (8): 365-375. Available from: doi:10.1007/BF02018059.

[4] Bornmann L, Mutz R. Growth rates of modern science: A bibliometric analysis based on the number of publications and cited references. *Journal of the Association for Information Science and Technology*. 2015; 66(11): 2215-2222. Available from: doi:10.1002/asi.23329.

[5] Larsen PO, Von Ins M. The rate of growth in scientific publication and the decline in coverage provided by Science Citation Index. *Scientometrics.* 2010; 84: 575-603. Available from: doi:10.1007/s11192-010-0202-z.

[6] 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; 102: 16569-16572. Available from: doi:10.1073/pnas.0507655102.

[7] Eugene G. Citation indexes for science: A new dimension in documentation through association of ideas. *Science*. 1955; 122: 108-111. Available from: doi:10.1126/science.122.3159.108.

[8] Garfield E. The history and meaning of the journal impact factor. *JAMA*. 2006; 295: 90-93. Available from: doi:10.1001/jama.295.1.90.

[9] Seglen PO. Why the impact factor of journals should not be used for evaluating research. *BMJ*. 1997; 314: 497. Available from: doi:10.1136/bmj.314.7079.497.

[10] Zhang CT. The *e*-index, complementing the *h*-index for excess citations. *PLoS One.* 2009; 4: e5429. Available from: doi:10.1371/journal.pone.0005429.

[11] Bornmann L, Mutz R, Daniel HD. Are there better indices for evaluation purposes than the *h* index? A comparison of nine different variants of the *h* index using data from biomedicine. *Journal of the American Society for Information Science and Technology*. 2008; 59: 830-837. Available from: doi:10.1002/asi.20806.

[12] Bornmann L, Daniel HD. What do we know about the *h* index? *Journal of the American Society for Information Science and Technology*. 2007; 58: 1381-1385. Available from: doi:10.1002/asi.20609.

[13] Bornmann L, Daniel HD. The state of *h* index research. *EMBO Reports*. 2009; 10: 2-6. Available from: doi:10.1038/embor.2008.233.

[14] Engqvist L, Frommen J. The *h*-index and self-citations. *Proceedings of the National Academy of Sciences*. 2008; 99: 11270-11274. Available from: doi:10.1016/j.tree.2008.01.009.

[15] Hirsch JE. Does the *h* index have predictive power? *Proceedings of the National Academy of Sciences*. 2007; 104: 19193-19198. Available from: doi:10.1073/pnas.0707962104.

[16] Cronin B, Meho L. Using the *h*-index to rank influential information scientists. *Journal of the American Society for Information Science and Technology*. 2006; 57: 1275-1278. Available from: doi:10.1002/asi.20354.

[17] Rad AE, Brinjikji W, Cloft HJ, Kallmes DF. The *h*-index in academic radiology. *Academic Radiology*. 2010; 17: 817-821. Available from: doi:10.1016/j.acra.2010.03.011.

[18] Svider PF, Choudhry ZA, Choudhry OJ, Baredes S, Liu JK, Eloy JA. The use of the *h*-index in academic otolaryngology. *The Laryngoscope*. 2013; 123: 103-106. Available from: doi:10.1002/lary.23569.

[19] Lee J, Kraus KL, Couldwell WT. Use of the *h* index in neurosurgery. *Journal of Neurosurgery*. 2009; 111: 387-392. Available from: doi:10.3171/2008.10.jns08978.

[20] Bornmann L, Daniel H-D. Does the *h*-index for ranking of scientists really work? *Scientometrics*. 2005; 65: 391-392. Available from: doi:10.1007/s11192-005-0281-4.

[21] Spearman CM, Quigley MJ, Quigley MR, Wilberger JE. Survey of the *h* index for all of academic neurosurgery: Another power-law phenomenon? *Journal of Neurosurgery*. 2010; 113: 929-933. Available from: doi:10.3171/2010.4.JNS091842.

[22] Pagel P, Hudetz J. *H*-index is a sensitive indicator of academic activity in highly productive anaesthesiologists: Results of a bibliometric analysis. *Acta Anaesthesiologica Scandinavica*. 2011; 55: 1085-1089. Available from: doi:10.1111/j.1399-6576.2011.02508.x.

[23] Svider PF, Pashkova AA, Choudhry Z, Agarwal N, Kovalerchik O, Baredes S, et al. Comparison of scholarly impact among surgical specialties: An examination of 2429 academic surgeons. *The Laryngoscope.* 2013; 123: 884-889. Available from: doi:10.1002/lary.23951.

[24] Rezek I, McDonald RJ, Kallmes DF. Is the *h*-index predictive of greater NIH funding success among academic radiologists? *Academic Radiology.* 2011; 18: 1337-1340. Available from: doi:10.1016/j.acra.2011.06.017.
[25] Acuna DE, Allesina S, Kording KP. Future impact: Predicting scientific success. *Nature.* 2012; 489: 201-202. Available from: doi:10.1038/489201a.
[26] Hirsch J. An index to quantify an individual's scientific research output that takes into account the effect of multiple coauthorship. *Scientometrics*. 2010; 85: 741-754. Available from: doi:10.1007/s11192-010-0193-9.
[27] Schreiber M. A modification of the *h*-index: The h_m-index accounts for multi-authored manuscripts. *Journal of Informetrics*. 2008; 2: 211-216. Available from: doi:10.1016/j.joi.2008.05.001.
[28] Hu X, Rousseau R, Chen J. In those fields where multiple authorship is the rule, the *h*-index should be supplemented by role-based *h*-indices. *Journal of Information Science*. 2010; 36: 73-85. Available from: doi:10.1177%2F0165551509348133.
[29] Livio M. *The Golden Ratio.* New York: Broadway Books; 2003.
[30] Markowsky G. Misconceptions about the golden ratio. *The College Mathematics Journal.* 1992; 23: 2-19. Available from: doi:10.2307/2686193.
[31] Engers M, Gans JS, Grant S, King SP. First-author conditions. *Journal of Political Economy*. 1999; 107: 859-883. Available from: doi:10.1086/250082.
[32] Baerlocher MO, Newton M, Gautam T, Tomlinson G, Detsky AS. The meaning of author order in medical research. *Journal of Investigative Medicine*. 2007; 55: 174-180. Available from: doi:10.2310/6650.2007.06044.
[33] Lake DA. Who's on First? Listing authors by relative contribution trumps the alphabet. *PS: Political Science & Politics*. 2010; 43: 43-47. Available from: doi:10.1017/S104909651099080X.
[34] Frandsen TF, Nicolaisen J. What is in a name? Credit assignment practices in different disciplines. *Journal of Informetrics.* 2010; 4: 608-617. Available from: doi:10.1016/j.joi.2010.06.010.
[35] Waltman L. An empirical analysis of the use of alphabetical authorship in scientific publishing. *Journal of Informetrics.* 2012; 6: 700-711. Available from: doi:10.1016/j.joi.2012.07.008.
[36] International Committee of Medical Journal Editors. *Recommendations for the conduct, reporting, editing, and publication of scholarly work in medical journals*. Available from: http://www.icmje.org/recommendations/ [Accessed 11th November 2021].
[37] Shapiro DW, Wenger NS, Shapiro MF. The contributions of authors to multiauthored biomedical research papers. *JAMA*. 1994; 271: 438-442. Available from: doi:10.1001/JAMA.1994.03510300044036.