Proposal of an individual scientometric index with emphasis on ponderation of the effective contribution of the first author: h-fac índex

Proposta de um índice cientométrico individual, com ênfase na ponderação positiva da participação do primeiro autor: índice h-fac

Francisco Gregori Júnior[1], Moacir Fernandes de Godoy[2], Francisco Ferreira Gregori[3]

Abstract

In the individual assessment of a scientific performance, five scientometric indices have been used most: the h-index, the index g, the h-major index, the contemporary h-index and the normalized h-index. We propose an alternative index ("Index h-fac"), which considers positively the participation of the first author and that, by having a dynamic characteristic, continuously monitors his/her performance and is easily adaptable to particular or individual situations from different research groups. Results from the geometric mean between the original h-index as proposed by Hirsh and a correction factor ("fac", "first author commitment") and, in turn, this value is divided by the mean interval (in years) of all studies. The index emphasizes two scores (X and Y). These scores X and Y were obtained by asking to all 83 cardiovascular surgeons from Southern Brazil (Paraná, Santa Catarina and Rio Grande do Sul) and Specialists, how they realistically estimated, in percentage, their effective contribution in each published paper in which they appeared as first author. Of the total, 80 (96.4%) responded. The average obtained was 78.0% and on this basis, the X score was established as 0.75 and the score Y as 0.25. The new index also considers the total number of citations as first author and as co-author, the average number of coauthors per publication and the total number of papers published. Theoretical examples are presented, discussing the main advantages of application. Serial evaluations in real world situations should be instituted to confirm the diagnostic and prognostic utility of this new index.

Descriptors: Authorship and co-authorship in scientific publications. Bibliometric indicators. Scientific publication indicators. Systems for evaluation of publications.

1. Associate Professor. Head of the Discipline of Cardiac Surgery at Department of Surgical Clinics, Health Sciences Unit, State University of Londrina (UEL), Londrina, PR, Brazil. Creator, Writing and Discussion of the Study.
2. Adjunct Professor of the Cardiology and Cardiovascular Surgery Department at São José do Rio Preto Medical School (FAMERP), Adjunct Learning Professor at Famerp, São José do Rio Preto, São Paulo, SP, Brazil. Writing and Discussion of the study.
3. Mechanical Engineering Student, Technology Institute of Aeronautics (ITA), São José dos Campos, SP, Brazil. Technical Assistance and Discussion.

This study was carried out at Health Sciences Unit at State University of Londrina (UEL), Londrina, PR, Brazil.

Correspondence address:
Moacir Fernandes de Godoy
Rua Garabed Karabashian 570 – São José do Rio Preto, SP, Brasil – CEP 15070-600
E-mail: mf60204@gmail.com

Article received on September 1st, 2012
Article accepted on September 28th, 2012

Gregori Júnior F, et al. - Proposal of an individual scientometric index with emphasis on ponderation of the effective contribution of the first author: h-fac índex

Rev Bras Cir Cardiovasc 2012;27(3):370-6

Abbreviations, acronyms and symbols	
A C	Total citations generated by a scientific production in which the researcher is the first author
CoC	Total citations generated by a scientific production in which the researcher is coauthor
h	Original index, as proposed by Hirsh
CoN	Average number of coauthors per study published
PS	Total number of studies published
AI	Average interval of time between the year of each publication and the current year
X	Weighting score for participation as first author
Y	Weighting score for participation as coauthor

Resumo

Na avaliação individual do desempenho científico, cinco índices cientométricos têm sido mais utilizados: o índice h, o índice g, o índice h-major, o índice h contemporâneo e o índice h normalizado. Propomos um índice alternativo ("Índice h-fac"), que pondera positivamente a participação do primeiro autor e que, por ter característica dinâmica, monitora continuamente seu desempenho, sendo facilmente adaptável a situações particulares individuais ou de diferentes grupos de pesquisa. Resulta da média geométrica entre o índice h original, conforme proposto por Hirsh, e um fator de correção ("fac"; "*first author commitment*"), sendo essa média, por sua vez, dividida pelo intervalo médio (em anos), de todos os estudos. O índice dá ênfase a dois escores (X e Y). Esses escores X e Y foram obtidos perguntando-se a todos os 83 cirurgiões cardiovasculares da Região Sul do Brasil (Paraná, Santa Catarina e Rio Grande do Sul) com Título de Especialista, em quanto estimavam, percentualmente, de modo realista, sua efetiva contribuição em cada trabalho publicado no qual apareciam como primeiro autor. Do total, 80 (96,4%) responderam. A média obtida foi 78,0% e, com base nisso, estabeleceu-se para o escore X o valor de 0,75 e, consequentemente, o valor de 0,25 para o escore Y. São considerados também o número total de citações como primeiro autor e como coautor, a quantidade média de coautores por publicação e o número total de trabalhos publicados. Apresentam-se exemplos teóricos discutindo-se as principais vantagens da aplicação. Avaliações seriadas e em situação de mundo real deverão ser instituídas visando confirmar a utilidade diagnóstica e prognóstica desse novo índice.

Descritores: Autoria e co-autoria na publicação científica. Indicadores bibliométricos. Indicadores de produção científica. Sistemas de avaliação das publicações.

INTRODUCTION

The term Bibliometrics deals with the application of mathematical and statistical methods, books and other media, relating mainly to the management of libraries and documentation centers. [1] On the other hand, the term Scientometrics refers to the analysis of the quantitative aspects of generation, dissemination and use of scientific information. The scientific data of any primary scientometric research are represented by all authors, their studies, their bibliographies and citations they receive. [2]

The individual scientific output has been, in recent years, assessed with the aid of various indices, all aiming at quantifying the academic merit of a particular researcher [3,4]. Among these indices, one of the most used is the "h-index" proposed by Jorge E. Hirsh, professor of physics at the University of California in 2005. According to Hirsh, a researcher has determined h index if h studies of total studies appearing in publications as author or co-author have been cited at least h times. For example: the h-index is 5 if the researcher has at least 5 publications with 5 citations each [5].

H-index is on a single number, incorporating both quantity (number of publications) and quality (citations or visibility), and therefore has an advantage over these indices separately and on other measures, such as "number of significant studies", "number of citations of those most cited, etc. [6].

This is an index easily applicable and practical, but on the other hand, some distortions may reveal, for which reason has appeared in the literature to suggest other indices in order to minimize the problem.

H-index does not take into account the number of citations to a given article has above h index. It is a consequence of the definition of the h-index of which the top publications have at least h^2 citations, but the current number may be much higher. Thus, if a particular author

Gregori Júnior F, et al. - Proposal of an individual scientometric index with emphasis on ponderation of the effective contribution of the first author: h-fac índex

Rev Bras Cir Cardiovasc 2012;27(3):370-6

has an h-index of 20 is irrelevant if some of his most cited articles have 50, 100 or 500 citations.

Egghe [7] disagree on this point and believes that a measure to indicate the overall quality of a researcher should include the citation counts of the most cited articles proposing, therefore, the "g-index". He defines the g-index as the highest number of publications that together received a total of citations equal to or greater than g^2. According that definition he concluded that whenever the g index is equal to or greater than the h index because the h index does not take into account how many citation above h^2 while in the g index it is considered. The author believes that this small modification would retain all the advantages of the h-index without increasing the difficulty of calculating and eliminating what he considered a disadvantage.

Apart from the actual amount of citations, other authors turn to the interaction between researchers and their impact on h-index. Usually, in the authorship of a scientific study can be found three situations: those who are the first author, those who are called corresponding authors and, finally, collaborators without a special role. It has been noted a progressive increase in situations where multiple authors are considered "first-author" given their relative contribution or the existence of more than one corresponding author, which can create internal conflicts between groups.

Hu et al. [6] concerned with the complexity found in the patterns of collaboration between researchers introduce an h-index based on the function performed by the investigator at each study, called h-major ("h-maj"), which takes into account only the articles in which the researcher played a very important or central role. They define the h-major index of a researcher as having a "m" value if m is the number of publications that the researcher has with relevant contributions with m citations. For example, if a researcher has a original h-index of 20, but only in 8 of these 20 had involvement considered of great importance, then his h-major index will be 8. It should be noted that the "g" index mentioned above is always equal to or greater than the original h-index, while the h-major index will always be equal or less than that index. They propose that the h-maj index be used as a supplemental index to the original h-index, especially in areas where it is common occurrence of multiple key authors or multiple corresponding authors. Hu et al. [6] could not to objectively quantify what they call the relevant or central participation in performing scientific study and it is also not clear in related publications. The h-index never regresses original value, since the number of citations never decreases, but this fact may constitute a drawback of the method, since there is no impact on the index in case of a break in productivity.

To solve this fact, it was proposed the contemporary h-index, which adds a score inversely related to the age of each article cited. Sidiropoulos et al. [8] proposed a $S^c(i)$ score for a specific item (i) based on the citation count according to the formula:

$$S^c_i = \gamma.(AC - AP + 1)^{-\delta}.(C_i)$$

Since γ and δ are coefficients with fixed values set forth by proponents of this contemporary h-index as 4 and 1, respectively, CY is the current year, PY is the publication year of the article and Ci is the number of citations that the article received.

Thus, citations received by an article in the current year are considered with the factor 4, while citations of studies published for 4 years already have their weight factor reduced to just one, dropping rapidly to only 0.67 when the article cited enters its sixth year of publication [2,8]:

$$S^c_i = 4.(2012 - 2012 + 1)^{-1}.(C_i) = 4.(C_i)$$
$$S^c_i = 4.(2012 - 2009 + 1)^{-1}.(C_i) = 1.(C_i)$$
$$S^c_i = 4.(2012 - 2007 + 1)^{-1}.(C_i) = 0,67.(C_i)$$

By this formulation, a particular researcher will have a h^c contemporary h-index when h^c of his N_p studies published obtain a $S^c_i \geq h^c$ score and the other $(N_p - h^c)$ articles have a $S^c_i < h^c$ score, numerically exemplifying, a particular researcher will have a h contemporary index 5 when 5 of their articles published Sci obtain a $S^c_i \geq 5$ score and the other $(N_p - 5)$ articles have a $S^c_i < 5.</h$ score.

The normalized h-index uses the feature to normalize the number of citations of each article dividing that number by the number of citations of authors of the article and, therefore, seeks to provide a better approximation of the single impact of each author [2].

The "e" index proposed by Zhang [9], differentiates researchers who have the same h index, counting also the surplus of citations, or that is, the impact of other publications that are not in those that comprised the h index.

The h individual indices try to reduce the effects of co-authorship (hi index) dividing the h index by the number of authors, trying to better assess the impact by author [10,11].

Finally, the rates of citation-weighted according the age of the manuscript (AWCR, AWCRpA and aW-index) relate the number of citations weighted with the age of the article, where the number of citations generated is divided by the age of article [12].

Thus, we see that no modification of the h-index weights, more concretely, the effective participation of the first author in the production of scientific study.

The aim of this study is to present a new index ("h-fac Index"), in which a correction factor is introduced in order to consider positively the participation of the first author usually responsible for the conception of the idea,

Gregori Júnior F, et al. - Proposal of an individual scientometric index with emphasis on ponderation of the effective contribution of the first author: h-fac índex

Rev Bras Cir Cardiovasc 2012;27(3):370-6

preparation methods and effective performa of the trial until the final conclusions.

METHODS

The h-fac index would result from the geometric mean between the original h-index, as proposed by Hirsh, and a correction factor ("fac" from English "first author commitment") and this mean, in turn, divided by the interval average time (in years) of all publications since the year of each publication to the current year.

The "fac" component would be the sum of the number of citations generated by a number of publications in which the researcher is the first author multiplied by a "X" score, which represents the appreciation by participation as an author on the study, plus the sum of number of citations in which the researcher is coauthor, multiplied by a "Y" score that represents the appreciation by participation as co-author, divided by the total number of coauthors that comprise the scientific researcher, all divided by the total number of articles published (Equations 1 and 2):

$$fac = \frac{CA.(X) + (CoC.(Y))/CoN}{TA} \quad \text{(Eq.1)}$$

$$h\text{-}fac \; index = \frac{\sqrt{h.(fac)}}{AI} \quad \text{(Eq.2)}$$

Where:

h = Original h-index, as proposed by Hirsh

CA = Total citations generated by a scientific production in which the researcher is the first author

CoC = Total citations generated by a scientific production in which the researcher is coauthor

CoN = Average number of coauthors per study published

PS = Total number of studies published

X = Weighting score for participation as first author

Y = Weighting score for participation as coauthor

AI = Average interval of time between the year of each publication and the current year

Square root of h. (Fac) = Geometric mean between the h-index and the fac-factor

To get the "X" and "Y" scores, we performed a survey among cardiac surgeons in Southern Brazil (states of Paraná, Santa Catarina and Rio Grande do Sul) with Specialist by the Brazilian Society of Cardiovascular Surgery (hence who have published at least one scientific study). The questions were posed by electronic means for surgeons to quantify their effective participation in a study that is the first author, based on the following: A) 0-20% B) 30-40% C) 50-60 %, D) 70-80% E) 90-100%.

RESULTS

Of the total of 83 surgeons included in the study, 80 (96.4%) answered. The values distributed by the ranges of choice are shown in Table 1.

Table 1. Breakdown of responses indicating the quantification of effective participation in scientific study when placed in the position of first author, under the self-assessment of 83 cardiovascular surgeons in southern Brazil.

Option	Range	Answers	%
A	0-20	0	0
B	30-40	0	0
C	50-60	10	12.1
D	70-80	48	57.8
E	90-100	22	26.5
X	Did not answer	3	3.6

The mean indicated value of 78.1% with a range of variation from 55% to 95%, with most of the options located in the range of 70 to 80% of participation, being then adopted as weighting score "X" for first author the value of 0.75 and thus the weighting score "Y" of 0.25 to participate as co-author. The final formulation for the "fac" was defined as follows (Equation 3):

$$fac = \frac{CA.(0.75) + (CoC.(0.25)/CoN}{TA} \quad \text{(Eq.3)}$$

The following is an example of how would be the h-fac index for two researchers with the same number of articles published by generating the same number of citations, with the same original Hirsch "h"index, with the same average number of coauthors by publication and with the same average time interval between the year of each publication and the current year (Table 2 and Figure 1).

Table 2. Indicators of researchers 1 and 2

	Researchers 1	Researchers 2
N Total of Articles [TA]	20	20
N of articles as author	15	5
N of articles as coauthor	5	15
N mean of coauthors [NCo] per article	5	5
N citations as author [CA] (assuming 6 citations per article)	90	30
N of citations as coauthor [CoC] (assuming 6 citations per article)	30	90
h-index	6	6
AI	9	9

Gregori Júnior F, et al. - Proposal of an individual scientometric index with emphasis on ponderation of the effective contribution of the first author: h-fac índex

Rev Bras Cir Cardiovasc 2012;27(3):370-6

$$fac = \frac{CA.(0,75) + [CoC.(0,25)]/ CoN}{TA}$$

$$\text{h-fac index } fac = \frac{\sqrt{h.(fac)}}{AI}$$

Researcher 1

$$fac = \frac{90.(0,75)+[30.(0,25)]/5}{20} = (67,5+1,5)/20 = 3,45$$

$$\text{h-fac index } fac = \frac{\sqrt{6 . 3,45}}{9} = 4,55/9 = 0,51$$

Researcher 2

$$fac = \frac{30.(0,75) + [90.(0,25)]/5}{20} = (22,5 + 4,5)/20 = 1,35$$

$$\text{h-fac index } fac = \frac{\sqrt{6 . 1,35}}{9} = 2,85/9 = 0,32$$

Fig. 1 - Development of the calculations of h-index for researchers 1 and 2

Another example of the application of h-fac index would be to compare the productivity of academic researchers with two different h-indices taking into account the weighting factor described above (Table 3 and Figure 2).

Table 3. Indicators of researchers 3 and 4.

	Researcher 3	Researcher 4
N Total of Articles [TA]	30	20
N of articles as author	15	5
N of articles as coauthor	15	15
N mean of coauthors [NCo] per article	5	3
N citations as author [CA]	60 (assuming 4 citations per article)	30 (assuming 6 citations per article)
N of citations as coauthor [CoC]	60	90
h-index	4	6
AI	9	9

$$fac = \frac{CA.(0,75) + [CoC.(0,25)]/ CoN}{TA}$$

$$\text{h-fac index } fac = \frac{\sqrt{h.(fac)}}{TA}$$

Researcher 3

$$fac = \frac{60.(0,75)+[60.(0,25)]/5}{30} = (45,0+3,0)/30 = 1,6$$

$$\text{h-fac index } fac = \frac{\sqrt{4 . (1,6)}}{9} = 2,53/9 = 0,28$$

Researcher 4

$$fac = \frac{30.(0,75) + [90.(0,25)]/5}{20} = (22,5 + 7,5)/20 = 1,50$$

$$\text{h-fac index } fac = \frac{\sqrt{6 \times 1,5}}{9} = 3/9 = 0,33$$

Fig 2 - Development of the calculations of h-index for researchers 3 and 4

The mean time intervals (Ti) were established as being equal to the two researchers, in order to observe that the influence of the score of the first author ("X") on the final index only.

DISCUSSION

In the analysis of a medical curriculum vitae, the production of full scientific studies and their publication is what draws the most attention. Initially, we highlight the articles published internationally, a showcase that generates greater disclosure for two main reasons: high possibility of penetration by the importance of the scientific journal and language, usually English, in which the study is published [13,14].

The real contribution of a researcher for the advancement of scientific knowledge can only be valued by the same number of citations he generates, because that is the most accurate indicator that the study was actually read and forwarded some information.

On the other hand, it seems unfair that, while generally having a predominance of activity of one of the authors of a particular study, all receive equal benefit with respect to citations. Hence, there is the proposal to use the most valuable and appropriate scores.

The "X" score: Evaluation of effective participation of the first author

In our proposal, a prominent factor for determining the "X" score, which because it was found based on the evaluation of the real world and with practically all cardiovascular surgeons from the south of the country, it should translate the effective equity share of participation of the study's first author. A noteworthy fact is that this score almost certainly represents a profile not only regional but national, extensive, to all Brazilian cardiovascular surgeons and perhaps worldwidely, precisely by the characteristic of performance and personality of surgeons. Moreover, in activities related to Clinic or the Basic Chairs, this score quite possibly will be different, since the participation distribution is more equitable between the authors and coauthors.

It is quite possible that, for clinical studies, the "X" score occupies an intermediate position, and for basic research, is almost uniform the participation by all, with minimal predominance of the first author. Similar surveys performed in the present study with surgeons, extended to those two other areas, and will be important to confirm our prediction.

Anyway, once established this score for each area, the "fac"component, indicative of the degree of commitment of each individual researcher with published work, will be easily calculated and therefore the h-fac index, allowing both peer assess as well as a self-assessment of academic performance over time.

Referring to the theoretical examples above mentioned ("Researcher 1" and "Researcher 2"), it appears that those two very similar situations and with original h-index equal to six for both researchers, the h-fac indices are different, favoring widely the first author.

Gregori Júnior F, et al. - Proposal of an individual scientometric index with emphasis on ponderation of the effective contribution of the first author: h-fac índex

Rev Bras Cir Cardiovasc 2012;27(3):370-6

The extra weight of the X score of first author appreciated his participation in the preparation of studies, since, in most cases, the first author is the founder and many others, also the main contributor in the collection, assessment, writing and final conclusions. The way to perform these values and how to weigh it properly, without a doubt, is the most difficult, but the model presented in this study appears to circumvent the problem.

We would add that the geometric mean of the h-index and the original fac distributes mathematically correct form the contributions of each component of the formula, strengthening the information contained in the h-fac index.

When comparing the researchers with different original h-index, our proposition also appears to assist in the matter. As exemplified in the "Researcher 3" and "Researcher 4", it is noted by those examples given that despite the "Researcher 3" having an original h-index lower than the "Researcher 4" at a ratio of 4 to 6 (0.67), his productive capacity in terms of effective contribution to the study, based on the weighting factor has improved (ratio 0.28/0.33 = 0.85).

In articles with a single author, the author's score, of course, will be 1 and those with two authors, sometimes the first author is cited by alphabetical order, as they participate in a similar way in the design and performance of the study. In such cases, the factors may not be valid, however, it is not what happens in most cases. Additionally, it can always assign a differential score, as 0.5 for each, exceptionally for the publication in question.

As already mentioned above, the the "X" score may eventually receive a value not so different from the "Y" score when the team of researchers concerned the distribution of tasks in the production of the study is more equitable, as often happens in basic research area. However, whatever the score attributed to the first author, his weighting will be positive and always greater than that attributed to participation as co-author.

The h-fac index proposed has the following advantages:

1st. Pondering positively the commitment and participation of the first author ("X" score);

2nd. Pondering negatively the number of co-authors, inhibiting the introduction of "ghosts" co-authors (CoN);

3rd. Pondering positively the number of citations per volume of publications (AC + CoC);

4th. Pondering negatively the number of publications without relevance (PS);

5th. Pondering the authors positively with the lowest average time interval between the publication date and the current date, indicating that they stays productive;

6th. Including all the benefits of predecessors global indices (using all citations, making valuable the first author and meeting the contemporary) without appreciably increasing the difficulty of calculation, which can still be easily implemented in conventional spreadsheets; 7th. Covering also the characteristics of adaptability to particular conditions and might serve as an informative performance of most individuals or research groups. 8th. Finally, being related to the average time interval (in years), of all publications, from the year of publication until the current year, gives a dynamic nature to the proposed index and making valuable the ongoing maintenance of the productive state, because with no entries of new publications the denominator will only increase tending to reduce the h-fac.

Therefore, the introduction of the h-fac index complementing the Hirsch"h" index may significantly collaborate, in our view, for a fairer assessment of the scientific productivity of a given author.

In medium-term, the formation of a database with real-world situations can provide important analysis of the predictive value of the combination of the correction factor proposed to the original h-index.

CONCLUSION

The h-index fac, in view of its value attributes to the first author, using all the citations, dynamism and contemporaneity, may constitute a useful tool to assess scientific production of cardiovascular surgeons in southern Brazil with Specialist Title, and, most likely, could be adapted and extended for application to quantify the productivity of individual researchers from other areas of medical knowledge.

REFERENCES

1. Ruiz MA, Greco OT, Braile DM. Fator de impacto: importância e influência no meio editorial, acadêmico e científico. Rev Bras Cir Cardiovasc. 2009;24(3):273-8.

2. Repanovici A. Measuring the visibility of the university's scientific production through scientometric methods: an exploratory study at the Transilvania University of Brasov, Romania. Perform Meas Metr. 2011;12(2):106-17.

3. Mudanças dos critérios Qualis!. Rev Bras Cir Cardiovasc. 2010;25(2):III-V. DOI: http://dx.doi.org/10.1590/S0102-76382010000200002

4. Castro F. Impacto nacional. Rev Bras Cir Cardiovasc. 2011;26(4):676-7.

Gregori Júnior F, et al. - Proposal of an individual scientometric index with emphasis on ponderation of the effective contribution of the first author: h-fac índex

Rev Bras Cir Cardiovasc 2012;27(3):370-6

5. Hirsch JE. An index to quantify an individual's scientific research output. Proc Natl Acad Sci U S A. 2005;102(46):16569-72.

6. 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. J Inform Sci. 2010;36(1):73-85.

7. Egghe L. Theory and practice of the g-index. Scientometrics. 2006;69(1):131-52.

8. Sidiropoulos A, Katsaros D, Manolopoulos Y. Generalized h-index for disclosing latent facts. ArXiv DI/06070,2006,1.

9. Zhang CT. The e-index complementing the h-index for excess citations. PLoS One. 2009,4(5):e5429.

10. Batista PD, Campiteli MG, Kinouchi O, Martinez AS. It is possible to compare researches with different scientific interests? Scientometrics. 2006,68:179-89.

11. Schreiber M. A modification of the h-index: the hm-index accounts for multi-authored manuscripts. J Informetrics. 2008,2:211-6.

12. Jin BH. The AR-index: complementing the h-index. ISSI newsletter. 2007,3:6.

13. Meneghini R, Packer A. Dossier of the Brazilian Journal "Revista Brasileira de Cirurgia Cardiovascular" (Brazilian Journal of Cardiovascular Surgery) submitted to the Journal of Citation Report/ISI, aimed at Indexation on JCR. Revista Brasileira de Cirurgia Cardiovascular Rev Bras Cir Cardiovasc. 2005;20(2):IV-VI. DOI: http://dx.doi.org/10.1590/S0102-76382005000200003

14. Souza EPS, Cabrera SEM, Braile DM. Artigo do futuro. Rev. Bras Cir Cardiovasc . 2010;25(2):141-8. DOI: http://dx.doi.org/10.1590/S0102-76382010000200003