Informatics in Medicine Unlocked 15 (2019) 100166

Contents lists available at ScienceDirect

Informatics in Medicine Unlocked

journal homepage: www.elsevier.com/locate/imu

Comparison of several author indices for gauging academic productivity

Edward J. Ciaccio[a,*], Govind Bhagat[a,b], Benjamin Lebwohl[a], Suzanne K. Lewis[a], Carolina Ciacci[c], Peter H. Green[a]

[a] Department of Medicine – Celiac Disease Center, Columbia University College of Physicians and Surgeons, New York, NY, USA
[b] Department of Pathology and Cell Biology, Columbia University College of Physicians and Surgeons, New York, NY, USA
[c] Department of Medicine, Surgery, and Dentistry, Scuola Medica Salernitana, Università degli Studi di Salerno, Italy

ARTICLE INFO

Keywords:
Author index
Author rank
Celiac disease
c-index
h-index

ABSTRACT

Background: Many author indices exist to gauge academic productivity. Several of these indices are calculated based upon an author's scholarly publication record, but the measurement methodology to calculate each index varies considerably, and the precise function being used, as well as the end result, is often complex and difficult to assess.

Method: Two straightforward methods to weigh author productivity from the publication and citation record were evaluated as possible means for providing a clearer assessment of scholarly activity. The author characteristic index (termed c-index) assigns author rank for each publication based upon author position. The characteristic prime (c') -index normalizes author rank from author position, so that the total weight per publication is unity. The top 10 scholars with keyword 'celiac disease' in the Google Scholar database were then assessed using these metrics. Rankings according to total number of publications, h-index, and c- and c'-indices were compared, then tabulated along with total papers included for assessment, and mean values per paper for author position, number of authors, citations, and year of publication.

Results: The order of the top ten authors with keyword 'celiac disease' varied substantially depending upon whether the h-index, c-index, or c'-index was used as a gauge. The characteristic indices assign credit to authors according to their position in an author list. The affiliated metrics provided a more complete picture of scholarly activity.

Conclusions: Academic achievement by scholars, based upon quantitative publication characteristics, has recently become of interest for evaluating job candidates, for determining work performance, and for bestowing awards and honors. The characteristic indices as described herein are readily calculated and interpreted, and may improve the assessment of scholarly activity.

1. Introduction

Quantitative metrics are increasingly being utilized as measures of personal achievement [1]. These including credit score, daily exercise parameters, and fitness parameters, including body mass index (BMI) [2]. Quantitative metrics pertaining to academic achievement are now often implemented to assess job applicants and job performance, and in consideration for bestowing awards and other honors [3]. Yet, it is often the case that the calculation, and the value calculated, are not readily understandable and interpretable, and/or their impact on scholarly performance is not readily elucidated.

The h-index has become a pervasive measure of academic achievement [4]. It is based upon author publications, and the number of citations for each of the publications. The h-index can be defined as follows. A particular scientist's peer-reviewed publications are first listed in order from highest to lowest number of citations. The data used for calculation of the h-index are simply the number of citations per publication, displayed from highest to lowest citations, and the row number of each publication. Thus for example, if the top three author publications x, y, and z have the following values:

x (100 citations) 1
y (50 citations) 2
z (25 citations) 3

The important parameters for calculating the h index are: 100 citations: row 1, 50 citations: row 2, 25 citations: row 3. When the complete list is displayed, not just the top three as above, then the last

* Corresponding author. Columbia University, P&S 7-445, 630 West 168th Street, New York, NY, 10032, USA.
E-mail address: ciaccio@columbia.edu (E.J. Ciaccio).

https://doi.org/10.1016/j.imu.2019.100166
Received 6 December 2018; Received in revised form 1 March 2019; Accepted 3 March 2019
Available online 08 March 2019

row in the list in which the number of citations is still greater than or equal to the row number is the h-index. Thus, the h-index is an integer value.

However, the h-index depends in part on the source by which publications and citations are included and counted for each person. Two commonly used sources are the Science Citation Index (SCI), and Google Scholar. The SCI is rigorous in the sense that journals included in this source must pass a number of stringent criteria. The SCI is also exclusive; one must have access to a subscription to view the entries. It is not open access. Whereas, Google Scholar is more inclusive of new and upcoming academic journals as sources, and anyone with access to the Internet can view it. Furthermore, Google Scholar may include publications other than just those that are peer-reviewed, including conference papers and US patents. A particular subject's h-index will often be higher in Google Scholar as compared with SCI, because many more journals are included in Google Scholar, and citations from these journals are counted. Regardless of source, the h-index can be time-varying. Whenever a new citation appears for a particular author, it is used in the calculation of the h-index and contributes to a change in its value. Uncommonly, the h-index can also decrease, if there is a retraction of a publication, or when the indexing agency makes an error and updates. Many different variations on the h-index have been developed [5].

The h-index is limited in several ways. It is simple, and does not consider any measurement beyond counting *h* papers with *h* citations. Thus the index does not consider the field of publication. A particular researcher may publish in a single field or in two or many fields. Furthermore, the profile of the citation curve is not considered. For example, if a particular scientist has published only four papers, but each of these papers has 1000 citations, the h-index = 4. Likewise, if a scientist has published only four papers, but each of these papers has 4 citations, the h-index = 4. There is a difference in the academic achievement of these individuals, yet the h-index would fail to show it. Moreover, the h-index does not account for author position in the author list, nor the relative contribution of a particular author for any particular published paper [6]. In some journals, author contribution is stated in the paper [7]. However, it is rare that these contributions, which may include data acquisition, data handling, data analysis, inspiration for the study, reference gathering, paper writing, and reading and correcting the paper, are assigned percentage values. Even if they were, such percentage values would likely be subjective and biased.

A number of variations with respect to the h-index have been introduced in the literature to account for other variables [8]. It is generally agreed upon that the first author and perhaps the senior (last) author, contribute the most work to any particular paper [9]. The second author and the second-to-last author may be considered to have contributed the next largest amount toward the study, and so on, so that the middle authors may be considered to have lesser importance in developing the article [10]. Herein, a variation of the h-index is implemented which is readily calculated, easy to understand, and considers author position in the list as a measure of contribution to each study. Although it does not remedy the case above where an author has four papers with either 1 or 1000 citations each, it can be considered as a suggested first step in improving author index metrics.

2. Method

Since Google Scholar is readily available to the public, as it does not have a paywall, and because it is easy to use and to refer to, it was selected for our study to assess author indexing [11]. A major background of most of the authors of this study is celiac disease. The keyword 'celiac disease' was thus selected as the topic for which to assess author indices. Google Scholar ranks authors based on total number of publications in the journals it indexes [12]. The top 10 authors in Google Scholar who listed 'celiac disease' as one of five keywords to describe their work were assessed. The publications of these authors included those pertaining to celiac disease, but also those related to any other topic that a particular author has investigated during the course of academic activity. Thus some of these authors, although listing celiac disease as an area of interest and expertise, may have many more publications in fields only distantly related to celiac disease, or completely unrelated to celiac disease. Authors are identified by a letter, based on total number of publications in Google Scholar.

Table 1
Traditional Author Indices, used in Google Scholar.

Author	total cites	h-index
A	24914	78
B	17859	63
C	16404	69
D	10034	45
E	7281	43
F	5410	35
G	5349	22
H	3710	38
I	3175	30
J	2872	28

In Table 1, the top 10 authors describing their work by including 'celiac disease' as a keyword are noted. The order of the authors is determined from their total number of publications, according to Google Scholar, at the time of the analysis. Publication number increases as more publications are attributed to each author. They can also sometimes decrease - when Google Scholar errs in its tally, followed by correction, or if a paper were to be retracted. The h-index is also shown. The ranking would differ when using h-index to assess scholarly activity, as compared to total number of publications. The work of some of these authors mostly pertains to celiac disease. However, the work of other of the authors may be divided between two main subjects, such as celiac disease and cardiology, or it may be divided among many subjects. The listings of Table 1 do not consider author position in their publications.

To consider author position, a characteristic index of scholarly activity that assigns weight was utilized. This technique has also been commonly used in prior work. The position *p* for calculating the characteristic index is:

1st or last author: $p = 1$

2nd or 2nd to last author: $p = 2$

3rd or 3rd to last author: $p = 3$

…

n or nth to last author: $p = n$ (1)

The weighted contribution of the author to a particular paper $wc = 1/p$. Thus for first or last authorship *wc* is 1/1 = 1, for second or second-to-last authorship wc is 1/2 = 0.5, for third or third-to-last authorship *wc* is 1/3 = 0.33, and so on. In Table 2, the top cited papers of author J, Table 1, are given as an example, since the h-index is relatively low, so that the number of rows needed for illustration will be relatively short. Shown for each paper, columns from left to right, are the author's location in the author list, the total number of authors on the paper, the number of citations attributed to the paper, its year of publication, the row number (which is also the sum of contributions to the h-index), the author's position in the author list according to the weighting paradigm described above (for example if listed at location 6 out of 7 authors, position is 2, if listed at location 7 of 9 authors, position is 3), the weight according to the characteristic index *wc*, and the sum of weighted contributions according to the characteristic index. The elements of the citations column decreases monotonically, while those of row number and the sum of weighted contributions increase monotonically. As described in the Introduction, the h-index is defined

Table 2
Calculation of characteristic index (c-index).

Location	Total	# Cites	Year	Row # = Σ wh	Position	wc = 1/p	Σ wc
6	7	449	1999	1	2	0.50	0.50
1	2	167	2001	2	1	1.00	1.50
3	10	114	2009	3	3	0.33	1.83
1	3	106	1994	4	1	1.00	2.83
4	13	99	2012	5	4	0.25	3.08
7	9	95	2006	6	3	0.33	3.42
4	5	90	2009	7	2	0.50	3.92
1	3	75	2007	8	1	1.00	4.92
3	7	71	2009	9	3	0.33	5.25
7	10	68	2010	10	4	0.25	5.50
1	10	67	2007	11	1	1.00	6.50
1	6	60	2010	12	1	1.00	7.50
4	6	60	1997	13	3	0.33	7.83
4	5	59	2012	14	2	0.50	8.33
1	3	55	2005	15	1	1.00	9.33
5	9	47	2008	16	5	0.20	9.53
4	8	38	2011	17	4	0.25	9.78
1	1	38	2000	18	1	1.00	10.78
1	6	36	2009	19	1	1.00	11.78
1	5	34	2010	20	1	1.00	12.78
1	5	33	2004	21	1	1.00	13.78
1	7	32	2011	22	1	1.00	14.78
1	6	31	2010	23	1	1.00	15.78
5	10	31	2008	24	5	0.20	15.98
1	4	30	2001	25	1	1.00	16.98
1	3	30	2001	26	1	1.00	17.98
1	6	28	2011	27	1	1.00	18.98
7	10	28	2011	28*	4	0.25	19.23
3	3	28	1992	29	1	1.00	20.23
1	5	27	2012	30	1	1.00	21.23
1	2	27	2008	31	1	1.00	22.23
1	5	26	2012	32	1	1.00	23.23
1	5	25	2010	33	1	1.00	24.23*
1	5	24	2013	34	1	1.00	25.23

Location = location of the author in the author list, Total = number of authors, # Cites = number of citations (time dependent), Year = year of publication, Row # = Σ wh = sum of contributions per paper to form the h-index, Position = author position, wc = 1/p = weighting for the characteristic index calculation, Σ wc = sum of contributions per paper to form the c-index.

as the row after which the value in the citation column is less than the row number. For the h-index calculation, author weight *wh* is unity for all papers regardless of author position, and thus the sum of h weighted contributions (Σ wh, Table 2) increases by 1 in every row. The h-index is therefore the row after which the value in the citation column becomes less than Σ wh. The value is 28, an integer, noted by an asterisk. Similarly, the c-index is defined as the row after which the value in the citation column becomes less than the sum of c-index weighted contributions (column Σ wc, Table 2). Note that the sum of c contributions increases fractionally, except when the author is listed as first or last author, in which case the weighted contribution equals 1, the same as for the h-index. The crossover of Σ wc with citations, the c-index, is also marked by an asterisk, and it is a real numbered value (24.23).

In Table 2, when the contribution of the author to each paper is small (i.e., author's position is toward the middle of the author list), the sum of weighted contributions (Σ wc) will increase more slowly than the h index, and thus the crossing with the citations column will occur further down in the list (c-index is smaller than h-index). However, if the author location for all published papers were to be either first or last in the list of authors, then *wc* would equal 1 in each row, and the sum of contributions column would be the same as for the h-index (Σ wc = Σ wh). Hence:

$$c\text{-}index \leq h\text{-}index \tag{2}$$

Furthermore, the h-index is always an integer, while the c-index is always a real number.

As a variant of the c-index, although also weighted according to author position, the characteristic prime index (c'-index) is fractional and based upon a harmonic series. To normalize so that the total contribution for all authors per paper equals unity, author weighting is calculated as follows:

1 author: $1/x = 1$

2 authors: $1/x + 1/x = 1$

3 authors: $1/x + 1/2x + 1/x = 1$

4 authors: $1/x + 1/2x + 1/2x + 1/x = 1$

5 authors: $1/x + 1/2x + 1/3x + 1/2x + 1/x = 1$

6 authors: $1/x + 1/2x + 1/3x + 1/3x + 1/2x + 1/x = 1$

7 authors: $1/x + 1/2x + 1/3x + 1/4x + 1/3x + 1/2x + 1/x = 1$ (3)

To calculate the c'-index, first solve for x. Then:

$$y = author\ position \cdot x \tag{4}$$

The weighed c' contribution is $wc' = 1/y$. An example is shown in Table 3 (author J, as in Table 2). For the paper with the most cites, in row 1, the author is 6th of 7 total authors. Based upon the equations shown above, for 7 authors, $x = 3.92$. The author position $p = 2$, and $y = p \cdot x = 2 \cdot 3.92 = 7.84$, which is noted in Table 3. The weighted c' contribution wc' = 0.13. The sum of contributions for the c'-index (Σ wc') is shown. The point after which the citations value crosses Σ wc' is the c'-index (not shown for brevity). In the Results section, graphs were constructed of citations, sum of c contributions (c-index), and sum of c' contributions (c'-index).

Since the c'-index weighs author contribution as a smaller fraction per paper as compared with the c-index, its calculated value will generally be smaller than the c-index. In general:

$$c'\text{-index} \leq c\text{-index} \leq h\text{-index} \tag{5}$$

However, if all of the publications of a particular author were single author papers, then:

$$c'\text{-index} = c\text{-index} = h\text{-index} \tag{6}$$

3. Results

The c-index and affiliated statistics for celiac disease authors are provided in Table 4. The c-index values (Table 4) are less than h-index values (Table 1) for all authors. Other information is shown including the average position of the author in the list of authors for all publications (Av Pos), average number of authors per publication (Av Au/

Table 3
Example of characteristic prime index (c'-index).

Location	Total	# Cites	Year	Row # = Σ wh	y	wc' = 1/y	Σ wc'
6	7	449	1999	1	7.84	0.13	0.13
1	2	167	2001	2	2.00	0.50	0.63
3	10	114	2009	3	13.71	0.07	0.70
1	3	106	1994	4	2.50	0.40	1.10
4	13	99	2012	5	20.16	0.05	1.15
7	9	95	2006	6	13.11	0.08	1.23
4	5	90	2009	7	6.66	0.15	1.38
1	3	75	2007	8	2.50	0.40	1.78
3	7	71	2009	9	11.76	0.09	1.86
7	10	68	2010	10	18.28	0.05	1.92

Location = location of the author in the author list, Total = number of authors, # Cites = number of citations (time dependent), Year = year of publication, Row # = Σ wh = sum of contributions per paper to form the h-index, y = an intermediate parameter used in calculating the c'-index, wc' = weighting for the characteristic prime index calculation, Σ wc' = sum of contributions per paper to form the c'-index.

Table 4
The c-index and Affiliated Statistics.

Author	c-index	Av Pos	Av Au/paper	Cites/paper	Year	# Papers
A	66.1	2.0	6.6	223.1	2003.0	88
B	35.2	4.2	13.5	187.8	2009.6	84
C	47.8	2.9	12.1	134.9	2005.2	94
D	36.0	2.1	8.5	154.2	2002.7	54
E	35.8	2.2	6.6	76.4	2001.2	54
F	32.6	3.1	11.1	109.5	2011.6	38
G	17.3	16.5	97.9	196.9	2003.7	26
H	20.6	3.4	10.8	72.1	2007.6	47
I	25.0	1.5	5.0	70.4	2011.8	32
J	24.2	1.9	6.0	66.8	2006.5	33

Table 5
The c'-index and Affiliated Statistics.

Author	c'-index	Av Pos	Av Au/paper	Cites/paper	Year	# Papers
A	35.5	1.9	6.2	165.8	2004.0	132
B	14.9	4.5	13.2	121.2	2010.4	141
C	22.7	2.9	11.9	96.3	2006.0	152
D	17.0	2.0	8.2	101.6	2003.9	91
E	20.0	2.1	6.1	54.6	2002.3	97
F	19.4	2.5	9.4	80.6	2011.7	58
G	7.4	11.5	67.8	135.3	2004.1	39
H	8.0	3.1	10.1	55.2	2008.8	66
I	13.7	1.7	5.4	46.8	2012.3	59
J	13.2	1.8	5.7	53.3	2006.5	46

paper), average number of citations per paper (Cites/paper), average year of publication (Year), and total number of publications needed for the c-index calculation (# Papers). The c-index value for Author J is 33, the same as the number of rows needed to reach the c-index, Table 2. The average position, authors per paper, and cites per paper were calculated for Author J using 33 papers, and so on for the other authors in the table. The c'-index and affiliated statistics for celiac disease authors are shown in Table 5. The c'-index values (Table 5) are lower than c-index values (Table 4) for all authors. A longer list of publications was used to calculate the c'-index as compared with the c-index (compare # Papers, Tables 4 and 5). Because of the different number of papers needed to calculate the c-versus c'-indices, the values of the average parameters in Tables 4 and 5 differ, as well as the indices themselves. Since more publications are used to calculate the c'- as compared with the c-index, and since the list of publications is ranked according to citations, the average cites per paper is smaller for the c'-index as compared with the c-index. Of note, the author labeled G has large average values for position and authors per paper. This was due to the fact that one paper in particular skewed the data, in which the author was listed as author number 383 of 2366 total authors listed. This paper, published in 2016, had been cited 3916 times at the time of this analysis. This skewed the average authorship position for G to greater than 10, Tables 4 and 5, as compared with the average authorship position of other authors, approximately 2–4. It also skewed the average number of authors per paper for G to greater than 50, Tables 4 and 5, yet the average authors per paper of other authors is approximately 5–10. This skew has very little effect on the c- and c'-indices - it is just used for weighting one of the values in the calculation. However, as is evident in Tables 4 and 5, it does affect the mean values for author position and average authors per paper, as it should. The h-index, in comparison, is not calculated by the number of authors per paper, and thus is unaffected by this statistic.

In Table 6, the rank of the 10 authors with celiac disease as a keyword is shown for the various indices. The authors A-J are listed according to total citations from Google Scholar. The h-, c-, and c'- indices only partially correspond to the ranking by total number of citations per author. Author A is ranked in first place for all indices. The

Table 6
Rank of authors by various indices.

Author	total cites	h-index	c-index	c'-index
A	1	1	1	1
B	2	3	5	6
C	3	2	2	2
D	4	4	3	5
E	5	5	4	3
F	6	7	6	4
G	7	10	10	10
H	8	6	9	9
I	9	8	7	7
J	10	9	8	8

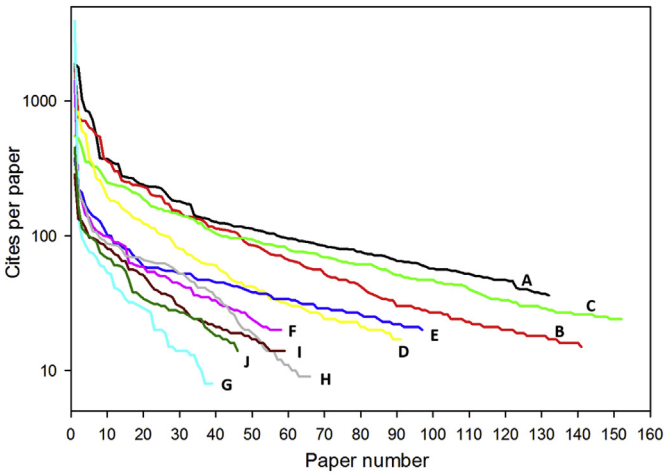

Fig. 1. Graph of cites per paper versus paper number, for 10 authors with keyword 'celiac disease' in Google Scholar. Papers are ordered from most to lesser citations. Thus, papers having the most citations are to the left on the graph. Toward the right on the graph, many author papers have similar, albeit lesser, number of citations. The ordinate axis has a log 10 scale.

c- and c'- indices change the ordering more than does the h-index, likely due to the inclusion of author position as a part of the weighting.

Table 6 displays rankings, which are integer values. In contrast, continuous values for some of the parameters used for calculation can be displayed as profiles. In Fig. 1, citations per paper are shown versus paper number, where papers are numbered according to their citation count. The author identifier is denoted next to each trace. The paper number at the end of each trace is equal to the number of papers used for the c'-index calculation, for reference. Author A has the highest number of citations per paper across the graph. Similarly, author B has a high citation count for all papers. Authors D and I have intermediate citations, while author J has a relatively low number of citations per paper, although higher than the author identified as G. Author G's first paper, as mentioned prior, had 3916 citations as of the time of the analysis; thus paper number 1 spikes to a high level for this author (the ordinate axis, Cites per paper, is based on log 10).

In Fig. 2, the running tabulation of the sum of paper contributions for c-index calculation (Σ wc) is shown. Traces are toward the top of the graph for authors A, D, I, and J as well as other authors, indicating that the position of these authors on each paper is relatively high. For the c'-index, Fig. 3 traces of Σ wc' are toward the top of the graph for authors A, I, J, as well as some other authors, indicating that both the position of these authors on each paper is relatively high, and that the total number of authors per paper is relatively low.

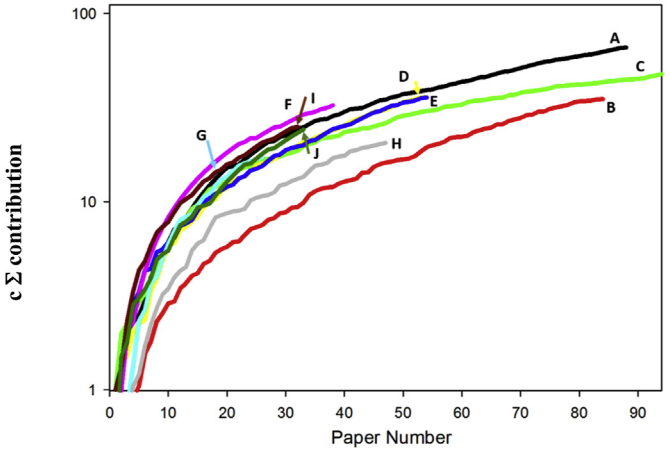

Fig. 2. Graphs of c-index sum of contributions versus paper number for 10 authors (same authors as in Fig. 1). The contributions for c-index are based on author order and are fractional (see Methods). The ordinate axis has a log 10 scale.

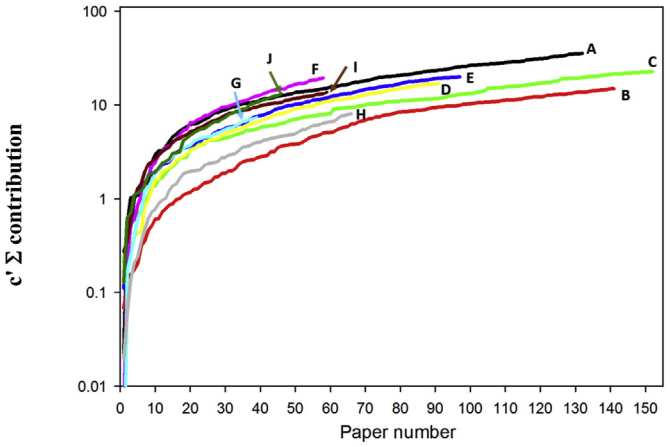

Fig. 3. Graphs of c'-index sum of contributions versus paper number for 10 authors (same authors as in Fig. 1). The contributions for c'-index are based on author order, they are normalized to a total weighting of unity for each paper, and they are fractional (see Methods). The ordinate axis has a log 10 scale.

4. Discussion

4.1. Summary

In this study, several author indices were computed and compared for the top 10 authors who listed 'celiac disease' as a keyword on Google Scholar. Google Scholar ranks authors based on total number of publications in a particular field or topic. Author ranking using h-, c-, and c'- indices varied as compared to total number of publications (Table 6). The h-index ranks authors by determining *h* publications with *h* or more citations. However, the characteristic indices c- and c'- modify the calculation by altering the weighting of author contribution for each paper (Tables 2 and 3). The c-index calculation considers author position in weighting author contribution. The c'-index calculation considers not only author position in weighting author contribution, but also total authors per paper. The h-index is an integer value, while the c- and c'- indices are real numbers. Besides calculation of these indices, mean values of auxiliary parameters were also tabulated, i.e., mean values of author position, total authors per paper, citations per paper, and publication year. The papers from which the mean values were computed were those that were utilized for index calculation (Tables 4 and 5). Together, all of these parameters provide a broad picture of publication achievement. Additional detailed information was graphed, based on monotonically increasing or decreasing parameters. Specifically, the citations per paper, sum of contributions in forming the c-index, and sum of contributions in forming the c'-index, were graphed versus paper number (Figs. 1–3 respectively).

4.2. Other weighted versions of the h-index

The h-index can be identical for authors with different order ranking in the byline of publications. This deficiency has resulted in various investigators proposing modifications to the h-index. Herein, two straightforward methods to rank author productivity based upon byline order were implemented. Other investigators have similarly considered fractional weighing-based author order, including [13,14]. An equal-weighting method has been suggested in which a paper is counted fractionally according to the inverse of the total number of authors [15]. It has also been proposed that the first author should receive lone significance for weighting in a modified h-index [16].

The h-index can be identical for authors with very different total citation characteristics. For papers cited fewer times than h, or citations of an individual paper above h, no contribution is made to the h-index [17]. Inclusion of the citation profile in weighting can assist in remedying this oversight [18]. Characterization of the excess citations beyond the h^2 citations of h-core papers (the *h* papers with *h* citations) was found useful to compare the output of a group of scientists having otherwise identical h-index [17,19]. The h-index can also be weighted based upon citation impact [20].

4.3. Quantitative relationship of indices and parameters to author

The author ranking depicted in Table 6 is likely reflective of job duties. Based on research into the backgrounds of those authors with highest rankings, and also those with substantial variation in ranking based on index type, author A is a senior investigator and lab director, which is evident by their typical location as last author. Likewise, authors B and D are senior investigators and lab directors, although not as senior as A. This is indicated by their high ranking in the indices, and they are also often listed last in the author order for publications, in accord with their senior status. Author I is a relatively early investigator and lab director, as is evident from the average year of publications for papers used to calculate the characteristic indices (Tables 4 and 5). Author I has many first and many last author status in the author rankings, befitting transition to senior investigator. Author J is a senior investigator but acts in an auxiliary capacity. Moreover, much of the work of author J is published in biomedical engineering journals as methodology, which is not often as well cited as papers published in clinical journals. Author J was first author in most of the publications, which is reflective of the description of new techniques by this author in these papers. In comparison between the h-index versus c- and c'- indices, author I's rank jumped from 9th place to 7th place, and author J from 10th place to 8th place, for the 10 authors (Table 6). This reflects the fact that many of the publications of authors I and J have few co-authors, and their position on each paper is often 1 or 2 (1st or last, or 2nd or 2nd-to-last location in the ordering). Hence, the mean author position of authors I and J rank the highest among the 10 authors (Tables 4 and 5). Author A also ranked highly in terms of mean position on publications, owing to often being listed last on papers, which, along with the large number of cites per paper, resulted in c- and c'- indices of greater magnitude as compared to the other authors. (Tables 4 and 5). Although author B tended to publish papers with more authors, these papers were often of high impact, with a large number of citations per paper, and a high profile in terms of the graphs of cites per paper (Fig. 1), contribution per paper to c-index (Fig. 2), and contribution per paper to c'-index (Fig. 3).

The information contained in Tables 4 and 5, and in Figs. 1–3, include a much more complete picture of author productivity as compared to the h-index alone. The average year of publication (see Tables

4 and 5) is partly indicative of the degree of seniority of the author. The average position on each paper is shown in the tables and is representative of the number of original studies contributed to by the author (i.e. as first or senior author). The number of authors per paper describes whether the studies were done by a relatively small group, or resulted from an effort by a large number of investigators. The citations per paper is expressive of the impact of the author studies. The traces of cites per paper, and sum of contributions per paper (Figs. 1–3) show whether the single value parameters are consistent among the collective body of work of each author. For example, although the number of citations per paper tend to show a slow decline for most authors, there is a sharper drop-off for authors G and H (Fig. 1) which is reflective of many citations being attributed to relatively few publications in these authors' portfolios.

4.4. Limitations

The data presented in this study was based on a limited number of 10 authors. Although these authors were listed as the top 10 authors in Google Scholar with keyword 'celiac disease' based upon total number of publications, their respective contributions to celiac disease research may not be as large as seemingly indicated. This is because the number of research papers each author contributes related to celiac disease can vary markedly. Some authors may devote their research solely to celiac disease investigation, while other authors may have one or more additional research topics to which they contribute to and publish frequently. Different fields have differing citation rates [21]. To investigate celiac disease research contributions more fully, the number of publications could be counted for all authors listed in Google Scholar, and those authors with a top number of celiac disease publications would then be included for computing indices. Though, this would be more difficult to implement, as the information is not readily available in a public database. Although two ways to weight author contribution for each paper were used for calculations herein (c- and c'-indices), based on author position and total number of authors per paper, there are numerous other possible methods to weight author contribution per paper. Furthermore, other factors were excluded in the calculations, including journal rank [22]. Although the citation profile per paper cannot be included in single-value indices, it was shown graphically (Fig. 1). Furthermore, the way in which author contribution per paper affected each author profile was shown in Figs. 2 and 3. Another factor that should be considered in future studies, concerns whether the author publications are original research or reviews, editorials, or case reports. This might require weighting the article type according to its importance. Although the c- and c'-indices depend upon author position, there is a lack of objective measure, and perhaps a lack of transparency, in assigning effort attribution in terms of authorship location on the byline. If however the error from the optimal assessment, which is unknown, is randomly distributed, it should not substantially affect the index values.

5. Conclusions

Measures of academic productivity are highly variable in terms of the measurement parameters used, and the resulting ranking of authors [23]. Some metrics are complex and difficult to calculate [8]. Many of these indices use citation number as a foundation for calculation. The basic unit of measurement is perhaps total citations [24], which is used by Google Scholar. However, this metric could be due to a large number of citations in one paper (if the number of author publications = 1) or from many papers. By calculating h papers which reach h citations, the h-index also considers the profile of citations over a number of papers. Herein, author position on each paper was also considered (c-index) and additionally, number of authors per paper was considered (c'-index). All of these indices could be improved, as presently they only provide a single number to state the productivity of each author. By including additional parameters (Tables 4 and 5) a more complete picture of author productivity is achieved. Furthermore, by graphing traces of monotonically changing variables used in the calculations (cites per paper, and weighted contribution of the author per paper; Figs. 1–3) the entire profile of authors' most cited publications is rendered. There are additional factors that may be included to enhance author indexing, such as journal rank in which papers are published. Still, this latter parameter may not be so important, because a paper cited many times is likely to be excellent and impact the field of study, regardless of the journal in which it is published.

Compliance with ethical standards

* Disclosure of potential conflicts of interest - All authors declare that they have no conflict of interest.

Ethical approval: This article does not contain any studies with human participants or animals performed by any of the authors.

Ethical Statement

Data was obtained per university guidelines. No clinical human or animal data were used.

Acknowledgement

None.

References

[1] Abbott A, Cyranoski D, Jones N, Maher B, Schiermeier Q, Van Noorden R. Metrics: do metrics matter? Nature 2010;465:860–2.
[2] Dalzill C, Nigam A1, Juneau M, Guilbeault V, Latour E, Mauriège P, Gayda M. Intensive lifestyle intervention improves cardiometabolic and exercise parameters in metabolically healthy obese and metabolically unhealthy obese individuals. Can J Cardiol 2014;30:434–40.
[3] Carpenter CR, Cone DC, Sarli CC. Using publication metrics to highlight academic productivity and research impact. Acad Emerg Med 2014;21:1160–72.
[4] Hirsch JE. An index to quantify an individual's scientific research output. Proc Natl Acad Sci USA 2005;102:16569–72.
[5] 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. J Assoc Inf Sci Technol 2008;59:830–7.
[6] Wan JK, Hua PH, Rousseau R. The pure h-index: calculating an author's h-index by taking co-authors into account. COLLNET J Sci Inf Manag 2007;1:1–5.
[7] Bates T, Anić A, Marušić M, Marušić A. Authorship criteria and disclosure of contributions: comparison of 3 general medical journals with different author contribution forms. JAMA 2004;292:86–8.
[8] Alonso S, Cabrerizo FJ, Herrera-Viedma E, Herrera F. h-Index: A review focused in its variants, computation and standardization for different scientific fields. J Inf 2009;3:273–89.
[9] Drenth JP. Multiple authorship: the contribution of senior authors. JAMA 1998;280:219–21.
[10] Shapiro DW, Wenger MD, Shapiro MD. The contributions of authors to multiauthored biomedical. JAMA 1994;271:438–42.
[11] Vision TJ. Open data and the social contract of scientific publishing. Bioscience 2010;60:330.
[12] Harzing AW, Van der Wal R. Google Scholar as a new source for citation analysis. Ethics Sci Environ Pol 2008;8:61–73.
[13] Egghe L, Rousseau R, Van Hooydonk G. Methods for accrediting publications to authors or countries: consequences for evaluation studies. J Assoc Inf Sci Technol 2000;51:145–57.
[14] Zhang C-T. A proposal for calculating weighted citations based on author rank. EMBO Rep 2009;10:416–7.
[15] Schreiber M. To share the fame in a fair way, hm modifies h for multi-authored manuscripts. New J Phys 2008;10:040201.
[16] Butson MJ, Yu PK. The first author h-index (h(fa)-index): levelling the field for small and large institute medical and science scholars. Australas Phys Eng Sci Med 2010;33:299–300.
[17] Brown OR. The h b-index, a modified h-index designed to more fairly assess author achievement. Redox Rep 2012;17:176–8.
[18] Zhang C-T. The h'-index, effectively improving the h-index based on the citation distribution. PLoS One 2013;8:e59912.
[19] Zhang C-T. The e-index, complementing the h-index for excess citations. PLoS One 2009;4:e5429.
[20] Egghe L, Rousseau R. An h-index weighted by citation impact. Inf Process Manag 2008;vol 44:770–80.
[21] Anderson T, Hankin R, Killworth P. Beyond the Durfee square: enhancing the h-

index to score total publication output. Scientometrics 2008;76:577–88.
[22] Braun T, Glänzel W, Schubert A. A Hirsch-type index for journals. Scientometrics 2006;69:169–73.
[23] Abramo G, D'Angelo CA. How do you define and measure research productivity? Scientometrics 2014;101:1129–44.
[24] Serenko A, Bontis N. Meta-review of knowledge management and intellectual capital literature: citation impact and research productivity rankings. Knowl Process Manag 2004;11:185–98.